Plant Systematics and Evolution (2022) 308:12
https://doi.org/10.1007/s00606-021-01802-y

ORIGINAL ARTICLE

Genetic structure and demographic history of *Allium mongolicum* based on SSR markers

Xiaoke Hu[1] · Jing Hu[1] · Yinghua Zhang[1] · Shengxiu Jiang[1] · Qiushi Yu[1]

Received: 28 March 2021 / Accepted: 29 December 2021 / Published online: 28 February 2022

Abstract

The population genetic structure was proved to be influenced by both historical events and gene flow, which can be examined by analyzing intraspecific differentiation and spatiotemporal population dynamics of a species at large spatial scale. In this study, we surveyed the population structure and evolutionary dynamics of 38 natural populations of a desert plant, *Allium mongolicum* Regel (Liliaceae), in northwestern China, using nine microsatellite loci. High genetic diversity was observed within populations, with mean allele diversity and the expected heterozygosity values of 0.689 and 0.655, respectively. The value of expected heterozygosity ($He = 0.655$) was higher than that of observed heterozygosity ($Ho = 0.317$), which indicated an excess of homozygosity within populations, which could be due to inbreeding. The gene flow among populations was high ($N_m = 1.245$) while the genetic differentiation among populations was low ($F_{st} = 0.169$). However, a distinct regional-scale differentiation was discerned among three geographical regions. Our results further detected considerably restricted but asymmetric gene flow among the three regions. The demographic dynamic analysis also detected an ancient population contraction and subsequent expansion during 0.11–0.33 Ma. These results suggested that both gene flow and population contraction/expansion caused by climate oscillations in ancient time played an important role in forming the population structure and accelerating the regional-scale differentiation of the desert plant. Our results further enforce the idea that the aridification and subsequent desert expansion/contraction since Pleistocene have greatly promoted the habitat fragmentation, and subsequent plant differentiation in northwestern China.

Keywords Desert plant · Gene flow · Genetic differentiation · Liliaceae · Northwestern China

Introduction

The geographical patterns of genetic variation within and among populations are commonly correlated with species demographic history such as range contraction/expansion, habitat fragmentation due to vicariance, genetic drift and gene flow (Schaal et al. 1998). Therefore, population structure of species has always been one of the key study components in the field of evolutionary ecology. Up to now, numerous studies have been carried out to reveal genetic structure and diversity of endangered plant species which mainly concentrated on specific environmental conditions, such as the Hengduan Mountains, the Qinghai–Tibet Plateau (QTP) and its adjacent highlands in China (e.g., Zhang et al. 2005; Wan et al. 2016; Cun & Wang 2010; Liu et al. 2013). Desert plants have long been growing under the special habitat in deserts, which have undergone a long-term process of adaptive evolution under environmental pressures such as extreme high temperature, drought and flowing sand, and subsequently, formed a unique genetic pattern involving regional-scale differentiation and range expansion during glaciation onto the vast temperate desert and steppe due to adaption to the extreme environment (Yu et al. 2013; Meng et al. 2015; Qian et al. 2016; Salameen et al. 2018). Owing to climate changes, especially the more recent aridification and desert expansion, the range expansion–contraction cycles formed the genetic structure and diversity of many species in desert regions of northwestern China (Meng et al. 2015). Meanwhile, a few geographical barriers in desert area caused by mountains, deserts and continuous oasis probably

Handling editor: Ying-xun Qiu.

✉ Qiushi Yu
yqs528@126.com

[1] State Key Laboratory Breeding Base of Desertification and Aeolian Sand Disaster Combating, Gansu Desert Control Research Institute, Lanzhou, China

had dramatic effects on the population structure and genetic diversity of desert species (e.g., Rebering et al. 2010; Guo et al. 2010; Li et al. 2012; Yu et al. 2013). For example, a large-scale drifting desert with many megadunes, which is a potential geographical barrier, could restrict population spread of some desert plants preferring the habitat of fixed sand land, and consequently, exerts influence on distribution and differentiation of these species (Yu et al. 2013; Wang et al. 2013).

Allium mongolicum Regel (Liliaceae) is a high-quality perennial forage herb mainly occurring in a few desert regions in northwestern China (He et al. 2007). It has a wide distribution, ranging from western Liaoning province to northeastern Xinjiang autonomous region, and occupies a large variety of ecological conditions in desert region, including flowing sandy land, arid and semiarid grassland, droughty gravel hillsides (Zhang et al. 2014; Wang et al. 1980). This species has strong resistance to shifting sand, drought and low temperature and plays an extremely important role in maintaining ecosystem stability in some desert regions (Ma 1994). Particularly, its leaves contain several kinds of amino acids and essential mineral elements (Sechenbater and Liu 2002), and therefore, it is a favorable wild green vegetable in desert regions (Ma et al. 2006; Liu et al. 2007; Yang et al. 2007). *Allium mongolicum* has a distinct ability to develop and reproduce themselves clonally under the extreme arid desert environments. Furthermore, this plant gets into temporary dormancy when an extreme drought occurs and thrives again rapidly when the suitable conditions are available (Zhao 2010). These biological characteristics of the species facilitate the colonization and growth of its populations in heterogeneous habitats with stressful conditions (Chidumayo 2006; Roiloa and Retuerto, 2007; Wang et al. 2008; You et al. 2013). According to the results of study by Zhang (2010), the mating system of *A. mongolicum* is mainly outcrossing with a few self-compatible individuals. It has high pollen viabilities and fertilization rates as well as high stigma receptivity. The pollen life is about 8 h and the best pollination time is 3–4 h after flowering. Pollination duration could last for about 5 days, best pollination time is about 2 days after flowering. The seeds are mainly dispersed through flowing sand, which probably caused limited dispersal distance compared to pollen (Bacles et al. 2006). In our previous study (Zhang et al. 2017), we used chloroplast DNA variations to survey the population structure and phylogeographical history of *A. mongolicum*. The results revealed three distinct groups, i.e., the east-central group (ECG), the west-central group (WCG) and the western group (WG), and each of the groups occupied a distinct geographical region with a specific dominant chlorotype. The desert species has a distinct regional-scale differentiation and it had probably experienced a sudden regional-scale range expansion/recolonization in the Quaternary glaciers (Zhang et al. 2017). It is known that chloroplast DNA is maternally inherited in angiosperms, and it is only dispersed through seeds (Wolfe et al. 1987), while nuclear DNA (e.g., nuclear microsatellite loci) is biparentally inherited and dispersed via both seeds and pollen (Palmé et al. 2003). Therefore, the level of gene flow only mediated by seeds is probably lower than that of the gene flow of both seeds and pollen, which possibly caused different population structure and differentiation pattern of the same species (Petit et al. 1993; Zhou et al. 2017a). Therefore, we further examined the population structure and genetic diversity of *A. mongolicum* based on biparentally inherited microsatellites in this study.

Microsatellites, also known as simple sequence repeats (SSRs), were extensively used for revealing genetic diversity and population structure of species (e.g., Walter and Epperson 2004; Ghebru et al. 2002; Falahati-Anbaran et al. 2007), due to high hypervariability, codominance, ubiquity and reproducibility in eukaryotic genomes (Wang et al. 1994; Yang et al. 2015). In the present study, we used nine polymorphic microsatellite loci isolated by de novo sequencing to investigate the genetic diversity and population structure of 38 *A. mongolicum* populations across its distributional range. The main aims are to: (1) compare the genetic diversity within and among populations, especially between different geographical regions of *A. mongolicum*, (2) reveal the population structure in the whole geographical distributions based on SSR markers and (3) evaluate the population dynamics of the desert plant in response to ancient climate oscillations.

Materials and methods

Population sampling

A total of 38 natural populations of *A. mongolicum* were sampled for this study (Online resource 1, 10 individuals per population). These populations represented almost the entire distributional range and different habitats of this species except for the southwestern Mongolia (Fig. 1: a). According to Wang et al. (1980), this species is widely distributed in the deserts of northern and northwestern China and its natural range extends to southwestern Mongolia which is located to the northwest of the sampled populations. For some difficulties of going abroad, we did not sample the populations in Mongolia. However, this does not affect the assessment of the interspecific differentiation and genetic diversity presented in this study, since these populations only occur at the edge of the distributional range in Inner Mongolia and probably represent only a small part of the total genetic pool of this species. The distance among populations sampled in this study is usually over 50 km, and all individuals in each population were usually separated from

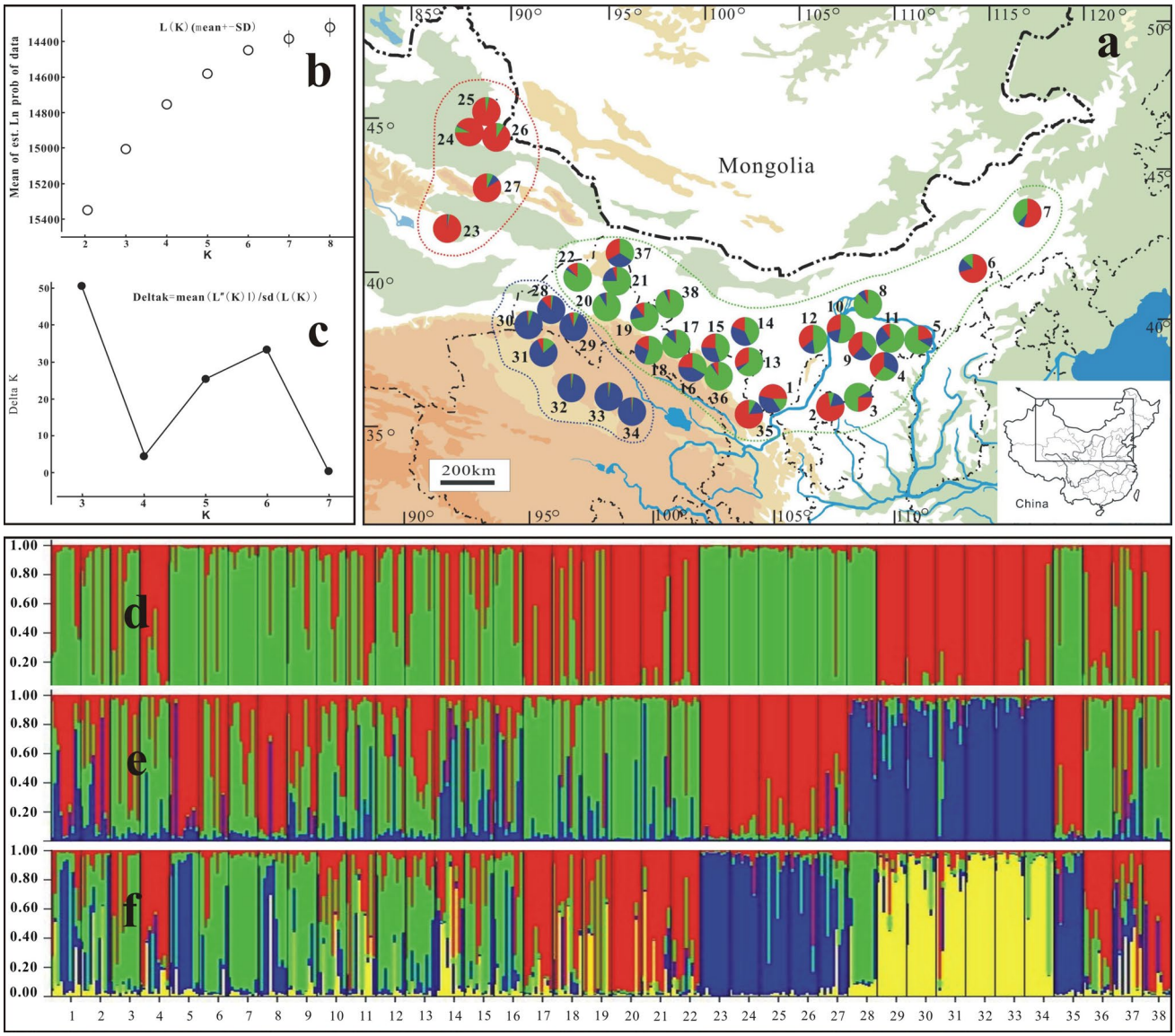

Fig. 1 Map of the sampling sites, geographical distribution of genotypes, and the summary plot of estimates based on SSR polymorphisms. **a** Sampling sites and distribution of polymorphic frequencies in surveyed populations ($K=3$). **b** Plot of mean likelihood L(K) and variance per K value from STRUCTURE on a dataset containing 380 individuals genotyped for 9 polymorphic microsatellite loci. **c** Evanno et al. (2005) plots for detecting the number of K groups that best fit the data. **d, e, f** Bar plots of individual Bayesian assignment probabilities of microsatellites for *Allium mongolicum* sampling sites using STRUCTURE for two, three and four clusters, respectively ($K=2$, 3, 4, respectively)

each other by more than 50 m, which aimed to ensure that samples closely related to the population and minimized the chance of sampling clones. Fresh healthy leaves for DNA extraction were sampled in the field and rapidly dried with silica gel. The latitude, longitude and altitude at each collection center were recorded using an Etrex GIS monitor (Garmin, Taiwan).

DNA extraction, amplification and fluorescence detection

Total genomic DNA was extracted from 20 mg of leaf material dried with silica gel using DNAsecure Plant Kit (DP305-03, Tiangen Biotech, Beijing, China). A total of fifteen polymorphic microsatellite loci of *A. mongolicum* were successfully isolated based on de novo sequencing using Illumina MiSeq platform in our previous study (Hu et al. 2017). However, being amplified in the population, six of the fifteen SSRs showed significant segregation distortion, and these distorted markers were excluded from the following population scanning. Consequently, nine microsatellite loci (Online resource 2) were used for genotyping of the whole pupulations sampled in the present study. SSR PCRs were performed on an ABI-2720 analyzer, in a 20 μL reaction volume containing the following components: 0.5 μL DMSO, 2 μL 10×PCR Buffer (2 mmol/L), 0.5 μl dNTP (2 mmol/L), 0.5 μL Primer F and Primer R (10 μmol/L), 0.2 μL TransStart Taq Polymerase, 1μL DNA, 14.8 μL Sterilized ddH_2O. PCR amplification was programmed at 95 °C for 5 min; 30 cycles of 95 °C for 30 s, 52/54/58 °C for 30 s and 72 °C for 30 s; and a final extension at 72 °C for 5 min. PCR products were checked on 1.5% (w/v) agarose gels containing ethidium bromide and visualized by AlphaImager (Version 4.0.1) for subsequent analysis.

The successfully amplified products were further used for subsequent PCR fluorescence reaction. Each microliter of PCR product that was labeled with HEX or 6-FAM (diluted 5–10 times with sterilized distilled water depending on the amount of PCR product) was mixed with 9.5 μl of deionized formamide (Hi-Di Formamide, Applied Biosystems) and 1.0 μl of fourfold-diluted GeneScan 400HD ROX size standard (Applied Biosystem), followed by denaturation at 95 °C for 5 min and immediate chilling on ice. Fluorescent signals of PCR products were detected on an ABI PRISM 3730 Genetic Analyzer (Applied Biosystems) in GeneScan (Version 3.0) mode. The results were viewed one by one using the software Penk Scanner and further analyzed using the program GeneMapper V4.0.

Genetic diversity and population structure analyses

We calculated the number of different alleles (N_a), number of effective alleles (N_e), observed heterozygosity value (H_o), expected heterozygosity value (H_e), Shannon's information index (I), private alleles and the percentage of polymorphic loci (PPL) for each population using GenALEx 6.503 (Peakall and Smouse 2006). We also estimated the gene flow (N_m) over all populations for each locus. The polymorphism information content (PIC) for each locus was calculated using the program Cervus 3.0.7 (Marshall et al. 1998) with the lowest expected frequency 5. A model-based clustering analysis was performed using the software STRUCTURE V2.3.4 to infer cryptic population structure and the number of clusters (Pritchard et al. 2000). This program employs a Bayesian algorithm to infer the true number of clusters (K) in a sample of individuals. K values were explored from 2 to 8 with each set having 15 repeats, using an admixture model. Each simulation started from 100 random initial conditions and each run consisted of a burn-in period of 100,000 iterations, followed by 1,000,000 Monte Carlo Markov Chain (MCMC) replicates. The output results of structure analysis were further imported into the web-based program STRUCTURE HARVESTE (available at http://taylor0.biology.ucla.edu/structureHarvester/) for the choice of the appropriate K value (Earl and vonHoldt 2012; Evanno et al. 2005). The program provides a fast way to assess and visualize likelihood values across multiple values of K and hundreds of iterations for easier detection of the number of genetic groups that best fit the data. The best K value was estimated using the delta K method (Evanno et al. 2005) and by choosing the smallest K value after the log probability of the data values [ln Pr (X|K)] reached a plateau (Pritchard et al. 2000). The Clumpp v1.1.2 program (Jakobsson and Rosenberg 2007) was used to combine the results from the 15 repetitions of the best K. The Distruct v1.1 program (Rosenberg 2004) was used to graphically display the results produced by Clumpp. We further performed analyses of molecular variance (AMOVA) using the program Arlequin version 3.0 to assess genetic differentiation among and within populations, and among groups, as well as average allele diversity over loci in each population, with significant tests by a nonparametric procedure involving 1000 permutations (Excoffier et al. 1992; 2005).

To examine the genetic isolation by geographic distance, Mantel tests with 10,000 permutations were used to reveal correlation between matrices of geographic and Nei genetic distances (Nei 1987) among all populations, as implemented in GenAlEx 6.5 software (Peakall and Smouse 2006). The geographical distance was calculated based on the coordinates of each location. We also performed Mantel tests in each geographical region clustered by STRUCTURE.

Demographic history analyses

Populations that have experienced a recent reduction of their effective population size exhibit a correlative reduction of the allele number and heterozygosity at polymorphic loci. But the allele number is reduced faster than the heterozygosity value (H_E). Thus, the H_E becomes larger than the heterozygosity (H_{EQ}) expected at mutation-drift equilibrium because H_{EQ} is calculated on the base of the allele number (Cornuet and Luikart 1996). We used the program BOTTLENECK version 1.2.02 (Piry et al. 1999) to analyze heterozygosity excess or heterozygosity deficit at each locus

under the infinite allele model (IAM), two-phased mutation model (TPM) and step-wise mutation model (SMM), respectively. Both the sign test and the Wilcoxon's signed rank test were used to determine if a population exhibits a significant number of loci with heterozygosity excess (Cornuet and Luikart 1996; Luikart and Cornuet 1998; Luikart 1997).

Furthermore, in order to infer the early demographic histories of the species, we used approximate Bayesian computation (ABC) in DIYABC v1.0 (Cornuet et al. 2010) to examine four designed plausible scenarios of demographic changes based on data set obtained for the nine nuclear SSR loci. These four scenarios are based on the results of the related demographic analyses based on chloroplast DNA variation in our previous study (Zhang et al. 2017), for instance, the Tajima's *D* and Fu's *F*s statistical tests as well as mismatch distribution analysis indicated that the desert species had experienced a sudden regional-scale range expansion/recolonization in the Quaternary glaciers. The four scenarios designed in the present study were simulated under the framework of a single population by assuming the same initial population size (NA). Scenario 1 assumed an ancient bottleneck event (t2) with a subsequent smaller stable population size (N2) and a recent expansion (t1, N1). Scenario 2 assumed an ancient population growth (t2) and a subsequent larger but stable population size (N3), followed by a recent bottleneck (t1, N1). Scenario 3 assumed an ancient population growth (t2) followed by a larger but stable population size (N4), and finally a recent expansion (t1, N1). In our previous study (Zhang et al. 2017), a rapid regional-scale range expansion was tested only in the main distributional areas of *Allium mongolicum* (the ECG and WCG populations). In order to infer population dynamics in unique geographical distributions with contrasting habitats (e.g. the QTP), we further designed the fourth scenario which assumed an ancient bottleneck event (t2) with a subsequent smaller stable population size (N5) and a recent more severe bottleneck (t1, N1) (Fig. 2). The prior distributions of the demographic parameters are listed in Online resource 3.

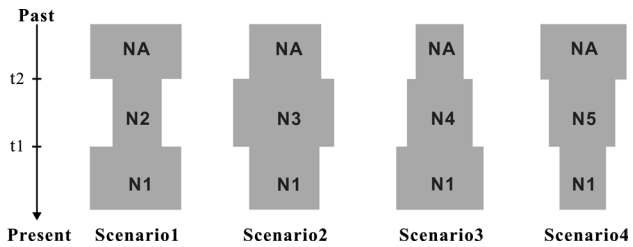

Fig. 2 Tested historical demographic models proposed for the three populations groups of *Allium mongolicum*. Scenario 1 assumed an ancient bottleneck event (t2) with a subsequent smaller stable population size (N2) and a recent expansion (t1, N1). Scenario 2 assumed an ancient population growth (t2) and a subsequent larger but stable population size (N3), followed by a recent bottleneck (t1, N1). Scenario 3 assumed an ancient population growth (t2) followed by a larger but stable population size (N4), and finally a recent expansion (t1, N1). Scenario 4 assumed an ancient bottleneck event (t2) with a subsequent smaller stable population size (N5) and a recent more severe bottleneck (t1, N1)

Finally, in order to determine gene flow among geographical groups of *A. mongolicum*, we used the software package MIGRATE version 3.7.3 (Beerli and Felsenstein 1999; Beerli and Palczewski 2010; http://popgen.sc.fsu.edu) to estimate the scaled effective population size, θ_w, of population groups clustered by STRUCTURE ($4N_e\mu$, where N_e is the effective population size and μ is the mutation rate) and the effective number of migrants ($2\ N_e m$, where m is the effective migration number) per generation. Because no estimates of the mutation rates of nuclear microsatellites are available for *Allium*, we used mean nuclear microsatellite mutation rate (2.5×10^{-4}) to estimate the effective population size of *A. mongolicum*. The mutation rate of $\mu = 2.5 \times 10^{-4}$ per site per generation was calculated by (Pereyra et al. 2009) based on the upper (5.2×10^{-4}) and lower (1.1×10^{-4}) 95% confidence limits of mutation rate for nuclear microsatellite loci of plants (Vigouroux et al. 2002). The maximum-likelihood analysis was performed based on 20 short chains (10,000 trees) and three long chains (1,000,000) with 10,000 discard trees per chain (burn-in).

Results

SSR polymorphism

Nine SSR loci used in this study were highly polymorphic (Table 1). The average number of different alleles (*Na*) at each locus ranged from 3.316 at AM51 to 8.477 at AM35, with the mean number of alleles 5.649. The number of effective allele (*Ne*) at each locus was in a range of 2.005–6.359 with a mean value of 3.986. The Shannon's information index (*I*) ranged from 0.799 to 1.935, with the lowest value at AM51 and the highest value at AM35. The observed and expected heterozygosity values ranged from 0.003–0.463 and 0.440–0.826, respectively. The highest values of *Ho* and *He* were observed at the locus AM35 (*Ho* = 0.463, *He* = 0.826), while the lowest values of *Ho* and *He* were observed at the locus AM03 and AM51, respectively (*Ho* = 0.003, *He* = 0.440). The polymorphic information content (*PIC*) is usually used for estimating the degree of allelic variabilities at a gene locus (Avval 2017). The *PIC* value calculated for each locus in this study was in a range of 0.773–0.976 with an average of 0.890. The *PIC* value of each SSR locus was more than 0.5, indicating that these screened loci had high properties to discriminate polymorphism of the species (Botstein et al. 1980). The *F*-statistics

Table 1 Genetic diversity indices, *F*-statistics and estimates of gene flow for each locus in *Allium mongolicum*

Locus	*N*	*Na*	*Ne*	*I*	*Ho*	*He*	*PIC*	F_{is}	F_{it}	F_{st}	N_m
AM03	380	4.868	3.716	1.252	0.003	0.603	0.923	0.996	0.997	0.357	0.451
AM14	380	4.289	2.594	1.035	0.258	0.526	0.773	0.510	0.601	0.187	1.089
AM33	380	4.553	2.782	1.118	0.366	0.567	0.876	0.354	0.489	0.209	0.948
AM35	380	8.447	6.359	1.935	0.463	0.826	0.965	0.439	0.516	0.137	1.570
AM39	380	7.868	5.907	1.841	0.287	0.803	0.976	0.643	0.702	0.165	1.268
AM51	380	3.316	2.005	0.799	0.421	0.440	0.777	0.042	0.200	0.164	1.270
AM55	380	7.500	5.467	1.782	0.311	0.784	0.958	0.604	0.667	0.159	1.324
S32	380	3.947	2.678	1.079	0.358	0.588	0.826	0.391	0.477	0.141	1.522
S56	380	6.053	4.370	1.587	0.384	0.756	0.935	0.492	0.555	0.124	1.761
mean	380	5.649	3.986	1.381	0.317	0.655	0.890	0.497	0.578	0.183	1.245

N number of individuals; *Na* number of different alleles; *Ne* number of effective alleles; *I* Shannon's information index; *Ho* observed heterozygosity; *He* expected heterozygosity; *PIC* polymorphic information content; F_{is} inbreeding coefficient at individual population level; F_{it} inbreeding coefficient at total population level; F_{st} population differentiation; N_m gene flow estimated from $N_m = 0.25(1-F_{st})/F_{st}$

showed that most loci had obvious inbreeding coefficient both at individual population level (F_{is}) and at total population level (F_{it}), with the mean values of 0.497 and 0.578 for F_{is} and F_{it}, respectively. Gene flow (N_m) ranged from 0.451 to 1.761, with the high mean value of 1.245, implying significant inbreeding coefficient level of *A. mongolicum*.

Genetic diversity and population structure

Genetic diversities among 38 populations of *A. mongolicum* are listed in Table 2. The average diversity (N_d) over loci per population ranged from 0.5421 to 0.8023 with an average of 0.6892. The mean number of polymorphic loci (*Na*) and the mean number of effective alleles (*Ne*) ranged from 4.0 to 7.4 and 2.6 to 5.6, respectively. The mean number of private alleles (N_p) was in a range of 0–0.4 with an average of 0.1 for the whole populations sampled in this study. At the species level, the range of Shannon's information index was 0.991–1.701 with an average of 1.381. The population 3 contained the highest genetic polymorphism (*Na* = 7.444, *Ne* = 5.653, *I* = 1.701), while the population 33 had the lowest genetic diversity (*Na* = 4.0, *Ne* = 2.568, *I* = 0.991). The observed and expected heterozygosity values were in a range of 0.1–0.444 and 0.515–0.762, respectively. Furthermore, the value of observed heterozygosity of each population was significantly lower than that of expected heterozygosity, and the overall mean value of expected heterozygosity (*He* = 0.655) was significantly higher than that of observed heterozygosity (*Ho* = 0.317).

In the clustering analysis based on Bayesian algorithm, the delta K was used for estimating the number of clusters, *K* (Evanno et al. 2005), due to the continuously increasing value of the log likelihood L(*K*) (Fig. 1: b). The delta K showed the highest peak at *K* = 3 by Evanno's method (Fig. 1: c), indicating that the appropriate *K* value should be 3 (Online resource 4, 8–9). Based on the detailed study of clusters, all populations sampled in this study were clustered into three groups (Table 2; Fig. 1: a): the first group (group 1) included populations 1–22 and 35–38, which occupied almost the whole geographical distribution of *A. mongolicum* except ranges in Xinjiang and Qinghai. The second group (group 2) consisted of populations 23–27. All the populations in this group occurred in Xinjiang, which are geographically far from the first group populations. The third group (group 3) was composed of populations 28–34, and most of which distributed on the QTP and its northern edge. For the population groups, diversity parameters such as N_d, N_p, *I* and *He* were clearly different from each other (Table 2). The results of AMOVA revealed that approximately 83.1% of the total SSR variation occurred within populations, whereas only about 5.3% occurred among groups (Table 3). For the whole distribution, variation within populations was obviously higher than that of among populations (11.6%). Furthermore, variance among populations ($F_{ST} = 0.169$) and within groups ($F_{SC} = 0.108$) was low. For each group clustered based on Bayesian algorithm, most variation also occurred within populations (79.93–89.58%), and variance among populations was low (10.42–20.07%).

In addition, using Mantel tests, the results did not reveal a significant correlation between geographic and genetic distance ($r = 0.190$, $p = 0.190$, Fig. 3: a) at the overall scale. Similarly, no significantly positive correlation was revealed in the two geographical regions (Fig. 3: c, d) apart from group 1 ($r = 0.199$, $p = 0.031$, Fig. 3: b). Although there was an obvious linear relationship between the geographic and genetic distances in the group 2 ($r = 0.661$), the *p* value of Mantel test was high ($p = 0.217$), suggesting that genetic isolation by geographic distance was not obvious.

Table 2 Genetic diversity in 38 natural populations of *Allium mongolicum*. Population codes are the same as in Online resource 1

Groups	Pop	N	N_d	N_p	*N*a	*N*e	*I*	*H*o	*H*e	PPL(%)	*N*pFreq
1	1	10	0.7608	0.222	6.778	4.907	1.613	0.300	0.723	100	0.050
	2	10	0.7830	0.222	7.000	5.216	1.652	0.300	0.744	100	0.100
	3	10	0.7684	0.111	7.444	5.653	1.701	0.200	0.730	100	0.050
	4	10	0.6813	0.000	5.111	3.261	1.297	0.367	0.647	100	0.000
	5	10	0.6883	0.000	4.778	3.390	1.272	0.322	0.654	100	0.000
	6	10	0.7789	0.222	6.556	4.694	1.613	0.278	0.740	100	0.100
	7	10	0.7368	0.222	6.333	4.233	1.505	0.278	0.700	100	0.050
	8	10	0.7187	0.111	5.667	4.019	1.439	0.244	0.683	100	0.100
	9	10	0.8023	0.222	6.778	4.818	1.668	0.244	0.762	100	0.075
	10	10	0.7474	0.000	6.667	5.198	1.594	0.267	0.710	100	0.000
	11	10	0.7345	0.000	6.556	4.543	1.544	0.356	0.698	100	0.000
	12	10	0.7509	0.000	6.556	5.071	1.578	0.300	0.713	100	0.000
	13	10	0.6789	0.111	5.556	3.658	1.341	0.344	0.645	100	0.100
	14	10	0.7462	0.111	6.778	5.129	1.593	0.389	0.709	100	0.150
	15	10	0.7181	0.000	6.556	4.888	1.521	0.300	0.682	100	0.000
	16	10	0.7362	0.111	6.556	4.869	1.530	0.333	0.699	100	0.050
	17	10	0.7006	0.333	5.111	3.632	1.346	0.356	0.666	100	0.250
	18	10	0.7620	0.000	6.444	4.412	1.558	0.356	0.724	100	0.000
	19	10	0.7298	0.000	6.111	4.532	1.478	0.378	0.693	100	0.000
	20	10	0.6585	0.000	5.000	3.360	1.257	0.411	0.626	100	0.000
	21	10	0.6813	0.000	5.778	3.958	1.377	0.400	0.647	100	0.000
	22	10	0.6819	0.111	5.889	3.789	1.401	0.400	0.648	100	0.050
	35	10	0.6766	0.000	4.556	3.405	1.252	0.444	0.643	100	0.000
	36	10	0.5661	0.444	4.556	3.642	1.134	0.100	0.538	77.8	0.150
	37	10	0.7415	0.000	6.111	4.353	1.507	0.278	0.704	100	0.000
	38	10	0.6988	0.000	5.333	3.803	1.345	0.411	0.664	100	0.000
Mean (Group1)		10	0.7203	0.098	6.022	4.324	1.466	0.321	0.684	99.15	0.049
2	23	10	0.7070	0.111	5.222	3.597	1.361	0.433	0.672	100	0.200
	24	10	0.6579	0.111	4.889	3.372	1.264	0.311	0.625	100	0.050
	25	10	0.6883	0.111	5.222	3.436	1.320	0.422	0.654	100	0.500
	26	10	0.5871	0.444	5.222	3.582	1.184	0.278	0.558	88.9	0.088
	27	10	0.5947	0.000	5.111	3.694	1.197	0.322	0.565	100	0.000
Mean (Group2)		10	0.6650	0.155	5.133	3.536	1.265	0.353	0.615	97.8	0.168
3	28	10	0.5953	0.111	4.556	2.703	1.115	0.278	0.566	100	0.100
	29	10	0.5795	0.000	5.000	2.988	1.153	0.311	0.551	88.9	0.000
	30	10	0.6123	0.000	4.889	3.205	1.191	0.211	0.582	100	0.000
	31	10	0.6918	0.111	5.444	3.978	1.359	0.378	0.657	100	0.100
	32	10	0.5900	0.000	4.000	2.615	1.063	0.222	0.561	100	0.000
	33	10	0.5421	0.111	4.000	2.568	0.991	0.167	0.515	100	0.050
	34	10	0.6175	0.000	4.556	3.311	1.168	0.344	0.587	100	0.000
Mean (Group3)		10	0.6041	0.0476	4.635	3.052	1.148	0.273	0.574	98.4	0.036
Mean (all Pop.)		10	0.6892	0.096	5.649	3.986	1.381	0.317	0.655	98.83	0.062

N sample size; N_d average allele diversity over loci; N_p mean number of private alleles; *N*a mean number of different alleles; *N*e mean number of effective alleles; *I* Shannon's information index; *H*o observed heterozygosity; *H*e expected heterozygosity; *PPL* percentage of polymorphic loci; *N*pFreq private allele frequency

Demographic history

The BOTTLENECK analysis indicated that the values of measured heterozygosity (H_E) at most loci of *A. mongolicum* were higher than that of expected average heterozygosity (H_{EQ}) under both IAM and TPM models. Under SMM models, H_E values at loci AM03, AM14, AM33 and AM55 were lower than H_{EQ} values while the

Table 3 Analyses of molecular variance (AMOVA) for populations and population groups of *Allium mongolicum* based on SSR data

Grouping of regions	Source of variation	d.f	SS	VC	Percent variation	Fixation index
Whole distribution	Among groups	2	95.875	0.19826Va	5.31	$F_{ST}=0.169$
	Among populations within groups	35	410.537	0.43140Vb	11.56	$F_{SC}=0.122$
	Within populations	722	2239.400	3.10166Vc	83.13	$F_{CT}=0.053$
Group1	Among populations	25	271.567	0.37985 Va	10.42	$F_{ST}=0.104$
	Within populations	494	1613.300	3.26579 Vb	89.58	
Group2	Among populations	4	67.070	0.69914 Va	20.07	$F_{ST}=0.201$
	Within populations	95	264.550	2.78474 Vb	79.93	
Group3	Among populations	6	71.900	0.46325 Va	14.56	$F_{ST}=0.146$
	Within populations	133	361.550	2.71842 Vb	85.44	

d.f. degrees of freedom; SS sum of squares; VC variance components; F_{ST} variance among populations; F_{SC} variance among populations within groups; F_{CT}, variance among groups relative to total variance

other loci had higher H_E values than H_{EQ}. However, we did not find significant differences between H_E and H_{EQ} ($P>0.05$) under three models (Table 4). The sign test and the Wilcoxon's signed rank test indicated that only two populations (2 and 38) exhibited a significant or extremely significant deviation from mutation-drift equilibrium under IAM model, exhibiting significant heterozygosity excess. Under TPM and SMM models, however, the two statistical tests above did not detect the significant deviation from mutation-drift equilibrium in the two populations (2 and 38). Under IAM model, the Wilcoxon's signed rank test showed that there were almost half of the populations in this study with a significant deviation from mutation-drift equilibrium, exhibiting prominent heterozygosity excess (Online resource 5).

In the DIYABC analysis, a distinct population contraction/expansion was discerned for different geographical regions (population groups) (Online resource 6). The most highly favored scenario was scenario 1, which showed a good fit for two groups (group 1 and 2) clustered by STRU CTURE. In this case, the population contractions occurred about 0.33 Ma and 0.28 Ma, respectively, and the population expansions happened about 0.14 Ma and 0.11 Ma, respectively. However, the designed scenario 4 was clearly favored by group 3 (QTP group), which had probably experienced an ancient bottleneck event at about 0.2 Ma and a recent more severe population contraction at about 0.12 Ma. The estimates on population expansions are approximately consistent with our previous results based on chloroplast DNA variation (Zhang et al. 2017).

In addition, the directionality of gene flow among three groups of *A. mongolicum* was examined and an estimate obtained with MIGRATE using the pooled data for each group (Table 5). The group 1 had the largest estimated effective population size ($N_e=665$), whereas the group 2 had the lowest one ($N_e=224$). The effective population size of the group 3 ($N_e=276$) was slightly higher than that of group 2. The estimated gene flow among the three groups by MIGRATE ranged from 0.0194 to 0.1048. The effective number of migrants ($2N_em$) from group 1 to group 2 was 0.0415 while the opposite direction gene flow between the two groups was 0.1006. There were also different asymmetric gene flows between group 1 and group 3 ($2N_em_{13}=0.1048$, $2N_em_{31}=0.0183$), and between group 2 and group 3 ($2N_em_{23}=0.0396$, $2N_em_{32}=0.0194$), respectively. Given all shared polymorphisms arised from gene flow, rather than from ancestral polymorphisms, there were obviously asymmetric gene flows among these three geographical population groups.

Discussion

Genetic diversity varied in different regions

Our results revealed a high level of genetic variation in *A. mongolicum* (Table 2, $He=0.655$), compared with other common plants ($He=0.150$) (Cole 2003), and most of the genetic variation is found within populations (Table 3). In particular, our results revealed remarkable genetic differences in different regions of the species (Table 2; Fig. 1: a). The genetic diversity of group 1 is higher than that of the other groups. This obvious genetic difference among different regions could have resulted from the heterogeneous habitats caused by environmental factors (e.g., altitude and desert formation) in the whole distributional area of *A. mongolicum*. The three population groups in this study lie in different climatic or geographical zones: the group 1 continuously lies in arid and semiarid deserts in northern China, while the group 2 mostly occurs in extremely arid gravel sand in Xinjiang where is geographically far from the former. The group 3 mainly occurs on the fringe of QTP where there are a series of mountains and valleys which probably formed barriers to gene flow among populations.

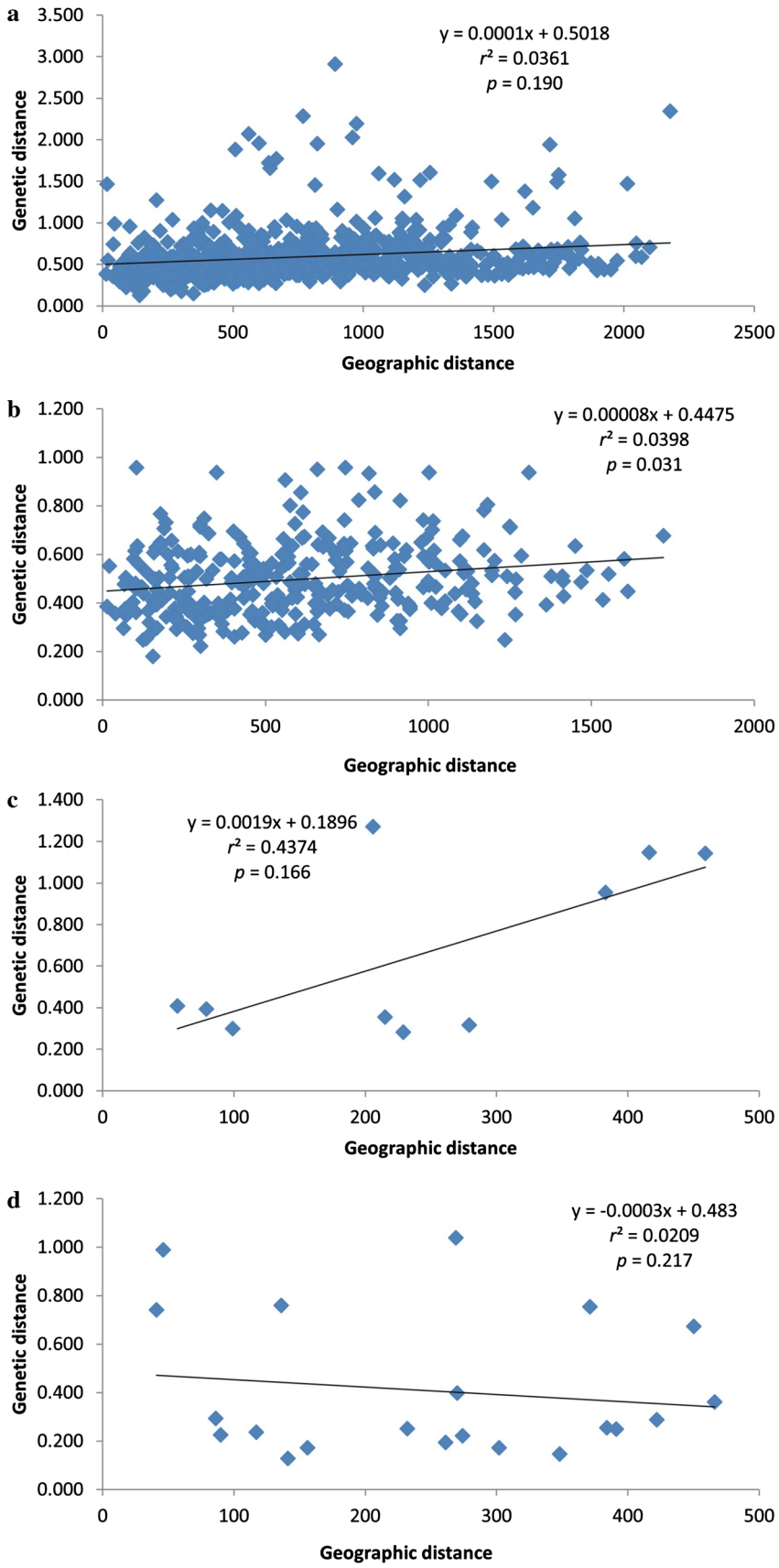

Fig. 3 Scatterplots of Mantel tests based on the geographical and Nei genetic distances. **a** across the entire populations; **b** in the first geographical region (group 1); **c** in the second geographical region (group 2); **d** in the third geographical region (group 3)

Table 4 Bottleneck analysis of *Allium mongolicum* based on microsatellites

Locus	H_E	IAM model			TPM model			SMM model		
		H_{EQ}	DH/sd	*P*	H_{EQ}	DH/sd	*P*	H_{EQ}	DH/sd	*P*
AM03	0.942	0.874	0.918	0.301	0.914	0.623	0.319	0.949	0.265	0.351
AM14	0.760	0.782	− 0.189	0.389	0.831	− 0.740	0.346	0.871	− 1.392	0.313
AM33	0.878	0.865	0.125	0.464	0.914	− 0.403	0.401	0.955	− 1.053	0.308
AM35	1.216	1.157	1.006	0.270	1.183	0.640	0.297	1.204	0.188	0.336
AM39	1.192	1.127	1.083	0.292	1.155	0.713	0.347	1.179	0.294	0.371
AM51	0.670	0.672	− 0.036	0.384	0.721	− 0.492	0.347	0.765	− 1.018	0.322
AM55	1.168	1.121	0.938	0.293	1.151	0.561	0.326	1.177	0.088	0.326
S32	0.863	0.765	0.733	0.315	0.818	0.296	0.344	0.862	− 0.197	0.325
S56	1.116	1.032	1.042	0.303	1.073	0.655	0.352	1.106	0.168	0.383

H_E measured heterozygosity based on allele frequencies; H_{EQ} expected average heterozygosity based on allele number; sd standard deviation of he mutation-drift equilibrium distribution of the heterozygosity; DH = H_E-H_{EQ}

Table 5 Effective population size (N_e) in three *Allium mongolicum* groups, and effective migration rates (N_em) among three groups, estimated by coalescent theory and a maximum-likelihood-based approach

Groups	θ_w	M_{12}	M_{21}	M_{13}	M_{31}	M_{23}	M_{32}	N_e	$2N_em_{12}$	$2N_em_{21}$	$2N_em_{13}$	$2N_em_{31}$	$2N_em_{23}$	$2N_em_{32}$
1	0.6650	–	0.3025	–	0.0550	–	–	665	–	0.1006	–	0.0183	–	–
2	0.2242	0.3702	–	–	–	–	0.1731	224	0.0415	–	–	–	–	0.0194
3	0.2761	–	–	0.7595	–	0.2869	–	276	–	–	0.1048	–	0.0396	–

The range estimates of 95% confidence limits are not shown

N_e effective population size, μ mutation rate ($\mu = 2.5 \times 10^{-4}$ /site/generation), $2N_em$ the effective number of migrants per generation from one group to another, e.g., $2N_em_{12}$ is the effective number of migrants from group 1 to group 2, while $2N_em_{21}$ is the effective number of migrants from group 2 to group 1

These contrasting habitats (e.g. drought, flowing sand and mountains) among different geographical groups probably played a critical role in forming different levels of genetic diversity in different regions (Schluter 2000; Rundle and Nosil 2005). The result of pairwise population *Nei* genetic distance analysis also indicated that the most genetic distances among groups are higher than within groups, with the highest genetic distance (2.908) between population 33 and 36 and the lowest genetic distance (0.179) between population 19 and 20 (Online resource 7).

The breeding system is expected to strongly influence the population structure and genetic diversity of plant populations. In this study, we found a high level of population inbreeding coefficient (F_{is}) at each locus for all sampled populations (Table 1), and we also detected distinctly higher value of expected heterozygosity than that of observed heterozygosity in each population (Table 2), indicating an excess of homozygosity and inbreeding of the desert plant. However, the excess of homozygosity in *A. mongolicum* is probably attributed to inbreeding rather than clonal reproduction because of the following reasons. Firstly, we observed extensive inbreeding within each natural population and inbreeding had been proved to increase population homozygosity (Charlesworth and Willis 2009; Liu et al. 2017). Secondly, *A. mongolicum* is a typical clonal plant which occurs in heterogeneous habitats with stressful conditions including extremely arid gobi, fixed or semi-fixed sandy land and semi-arid grassland. This herb's clonal ability and its temporary dormancy under extremely arid conditions probably lead the species to survive and recover mainly through clonal ramets after severe environmental changes and allow for occupation of new space (Yu et al. 2008; Wang et al. 2008; Moola and Vasseur 2009; You et al. 2013). However, the reproduction by cloning increases population heterozygosity (Zhou et al. 2017b; Gaut et al. 2018). Clonal propagation can lead to the accumulation of recessive deleterious mutations but without decreasing fitness, and a reduction of genetic load which originates partly from the decreased *Ne* during a bottleneck. Consequently, it reduces the efficacy of genome-wide selection (Charlesworth and Willis 2009), which may in turn increase population heterozygosity (Lohmueller 2014; Henn et al. 2015). Thirdly, the plant distribution is continuous in its main distributional ranges (Fig. 1: a. populations in central Inner Mongolia, Hexi corridor of Gansu, and Ningxia). There are few geographical barriers for gene flow in its distributional regions. Actually, we did detect higher levels of gene flow (N_m) among populations (Table 1). Considering the clonal reproduction, pollen and seed characteristics,

geographical proximity to one another and habitat continuity, we believe that the main cause of higher homozygosity is inbreeding rather than clonal reproduction. The clonal reproducing individuals could have balanced some levels of heterozygosity deficit in this species.

Regional-scale differentiation based on SSR markers

The aridification and subsequent desert expansion/contraction during the Pleistocene glaciations has probably resulted in range fragmentation of many desert species in Northwest China (Meng et al. 2015), furthermore, promoted allopatric divergence among isolated populations, and possibly speciation (Sun and Li 2003; Willis and Niklas 2004; Guo et al. 2010; Al-Shehbaz et al. 2006; Su et al. 2011; Ma and Zhang 2012; Lu 2015). In this study, the STRUCTURE analysis based on SSR markers separated all surveyed populations into three groups (Fig. 1; Online resource 4), which were supported by the results of AMOVA analysis as well as genetic diversity indices (Tables 2, 3). A clustering analysis (Fig. 1) and AMOVA result indicated that there was a clear genetic differentiation among the three geographical regions ($F_{CT}=0.053$) although most of the genetic variation (83%) existed within populations. These results all together implied that there was a distinct regional-scale differentiation among the three population groups of *A. mongolicum*. The low among-groups differentiation (F_{CT}) was probably caused by the late differentiation rather than gene flow, considering the restricted gene flow among the three geographical regions estimated by the MIGRATE (Table 5).

This regional-scale differentiation has probably resulted from desert expansion/contraction, or distributional fragmentation caused by aridification (Meng et al. 2015). Actually, the population group 2 is not only geographically far from the group 1, but also has a unique habitat characteristic which is significantly different from the group 1. The habitat of group 1 characterizes flowing sand or fixed sand with many other accompanying herbs, while the group 2 occupies the extremely arid gobi habitat with few other plants. However, most populations of group 3 mainly occur on the QTP where there are a series of high mountains and deep valleys which probably had important influences on the population expansion and gene flow among regional groups of *A. mongolicum* (Fig. 1: a). The Mantel test did not reveal a significantly positive correlation between geographical and genetic distance ($r=0.190$, $p=0.190$, Fig. 3), indicating that the genetic isolation caused by geographical distance played a minor role in accelerating the regional-scale differentiation of the species. Therefore, it is highly likely that the three groups diverged from each other due to environmental changes and habitat fragmentation caused by aridification and subsequent desert expansion/contraction, and that they may have survived in different refugia and experienced ecological divergence (Meng et al. 2015).

The clusters of all populations based on SSR markers in the present study were partly congruent with the population groups based on chloroplast DNA (cpDNA) variations in our previous study (Zhang et al. 2017). In this study, SSR variations did not discern the two population groups (the ECG and the WCG) separated by cpDNA variations (Zhang et al. 2017), namely most of the populations in Inner mongolia, Gansu and Ningxia (populations 1–22, 35–38) were grouped into the same one (Group 1), and the populations on Qinghai–Tibet Plateau were grouped separate one (Group 3). This incongruence of regional-scale differentiation between cpDNA and SSR variations probably resulted from the two following reasons. First, chloroplast DNA is maternally inherited in angiosperm (Wolfe et al. 1987) and its gene flow mainly mediated by seeds (Schaal et al. 1998; Palmè et al. 2003), while SSR variation is biparentally inherited and its gene flow mediated via both seeds and pollen (Petit et al. 1993). Therefore, the rates of gene flow between cpDNA and SSR markers are very different due to their different inherited modes (Palmè et al. 2003; Zhou et al. 2017a), which probably caused different clustering of *A. mongolicum* populations. In general, biparentally inherited SSR variations should have experienced higher levels of gene flow than that of maternally inherited cpDNA. Consequently, the higher gene flow of SSR variations caused the genetic homogeneity (Fig. 1: a; Table 1) between the two geographically contiguous groups (ECG and WCG) clustered by cpDNA variations. The second reason is probably hybridization. There are few distinct geographical barriers between the two geographically contiguous groups (ECG and WCG). Therefore, a high frequency of past hybridization mediated by pollen between the two cpDNA lineages was expected to induce discordance of genetic pattern between nuclear microsatellite markers and cpDNA haplotypes (Wegener et al. 2019). The pollen-mediated hybridization among populations rapidly homogenized nuclear genotypes and eroded boundaries between the two parapatric population groups (Glotzbecker et al. 2016).

Demographic dynamics and gene flow among geographical regions

The historical dynamics undoubtedly left genetic imprints in their extant populations (Hewitt 1996; Avise 2000; Abbott et al. 2000). Recently colonized species may experience a population bottleneck, resulting in a more reduction in the number of alleles than that of expected heterozygosity at polymorphic loci (Luikart and Cornuet 1998). In this study, the BOTTLENECK test did not detect the distinct recent bottleneck in most of *A. mongolicum* populations surveyed, indicating that the plant populations have not experienced recent demographic reductions, although the IAM model

indicated significant heterozygote excess in the two populations (2 and 38) (Online resource 5). However, the DIYABC analysis indicated that *A. mongolicum* populations have undergone demographic changes during the Quaternary glaciers or interglaciers (Online resource 6). The populations in group 1 and 2 experienced an ancient contraction event at about 0.33–0.28 Ma, and a subsequent expansion at about 0.14–0.11 Ma, while the populations in group 3 have experienced two contractions at about 0.20–0.12 Ma. The estimate of the demographic expansion time based on SSR variations is approximately consistent with the previous result inferred from the chlorotypes of the plant (Zhang et al. 2017), and especially, this estimate of expansion time is also quite consistent with the Last Interglacial period (ca. 0.14–0.12 Ma before present), indicating the population expansion of this species was probably correlated with the climate oscillations of this stage. Some studies suggested that the aridification process since the Quaternary caused the desert expansion in northwestern China, and subsequently, created some novel beneficial habitats for population expansions of some desert species which are more suitable for surviving in sandy land (Bush et al. 2004; Sun and Liu 2006; Qian et al. 2016). *Allium mongolicum* is a typical psammophyte, with a higher drought tolerance and a preference to desert habitats (Zhang et al. 2014). Our results further enhanced the hypothesis that the enlargement of deserts in northwestern China since the Pleistocene has probably accelerated population expansion of the desert plant. In addition, the different scenario of demographic dynamic of group 3 could have been linked to the special surviving environment of the Qinghai–Tibet Plateau, e.g., extreme low temperature, or high altitude. However, more researches are needed to explore the causes which led to the different population dynamics between the two types of environments with contrasting ecological conditions.

To further investigate the demographic history of the species, we estimated the effective population sizes of the three geographical population groups and effective migration rates among them. Our results detected considerably restricted gene flow (Table 5) among the three geographical regions. Considering comprehensive factors including environmental difference, geographical barriers, pollen and seed dispersal mechanism and distance, as well as pollen longevity, we suppose that the extremely limited gene flow among the three population groups might have been caused by environmental difference or geographical barriers (mountains and/or valleys), which may have contributed to genetic differentiation among them (Dutech et al. 2004; Meng et al. 2015). Furthermore, we detected obviously asymmetric gene flow among these three geographical population groups, e.g., the estimation of migration rates indicated that the gene flow from group 2 to 1 was higher than that in the opposite direction (Online resource 6), which was probably correlated with monsoon climate in northwestern China (Yin et al. 2015). The monsoon climate oscillations have frequently affected the vegetation transition of desert ecosystems and the temporal-spatial evolution of deserts in the Asia interior (Qian et al. 2015). A previous study showed that the East Asia monsoon system could have significantly affected intraspecific divergence, gene flow and regional population dynamics of a typical desert shrub *Reaumuria soongarica* (Yin et al. 2015). It is highly likely that unbalanced alternate actions between summer and winter monsoons affected the dynamic ranges of the deserts in northwestern China, and consequently, caused the asymmetric gene flow among the three geographical population groups of *A. mongolicum*.

Information on Electronic Supplementary Material

Online Resource 1. Sampling sites and sample size for 38 populations of *Allium mongolicum*.
Online Resource 2. The size range, repeat motifs, and primer sequences of 9 SSR loci in *Allium mongolicum* developed for this study.
Online Resource 3. Parameters used as prior settings in DIYABC analysis.
Online Resource 4. Output of the Evanno method results. Yellow highlight is performed dynamically on the website and shows the largest value in the Delta *K* column.
Online Resource 5. Analysis on mutation-drift equilibrium of *Allium mongolicum* based on microsatellite.
Online Resource 6. Estimated posterior distributions of the parameters according to Approximate Bayesian Computation (ABC) for the best scenario of the demographic history for the four groups of *Allium mongolicum* divided by STRUCTURE.
Online Resource 7. Genetic distance (below diagonal) and genetic identity (above diagonal) among *Allium mongolicum* populations.
Online Resource 8. Raw STRUCTURE output of the Evanno method used to estimate the best *K* values.
Online Resource 9. Rate of change of the likelihood distribution in the Evanno method used to estimate the best *K* values with STRUCTURE.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00606-021-01802-y.

Acknowledgements The authors thank Meng Xu (Northwest University) for his help in data analysis, Fanglan He, Jinhu Zhang and Boyou Zhao (Gansu Desert Control Research Institute) for their helps in field sampling, and two anonymous reviewers for their constructive suggestions on revision. We also thank professor Yongfeng Zhou (Agricultural Genomics Institute at Shenzhen, CAAS) for his English improvement and some precious suggestions on discussion in the manuscript. This work was supported by the National Natural Science Foundation of China (Grant Number 31360089, 32060235).

Funding This study was funded by the National Natural Science Foundation of China (Grant Number 31360089, 32060235).

Declarations

Conflict of interest The authors declare that they have no conflict interest.

References

Abbott RJ, Smith LC, Milne RI, Crawford RM, Wolff K, Balfour J (2000) Molecular analysis of plant migration and refugia in the Arctic. Science 289:1343–1346. https://doi.org/10.1126/science.289.5483.1343

Al-Shehbaz IA, Beilstein MA, Kellogg EA (2006) Systematics and phylogeny of the Brassicaceae (Cruciferae): an overview. Pl Syst Evol 259:89–120. https://doi.org/10.1007/s00606-006-0415-z

Avise JC (2000) Phylogeography: the history and formation of species. Harvard University Press, London

Avval SE (2017) Assessing polymorphism information content (PIC) using SSR molecular markers on local species of Citrullus colocynthis. Case study: Iran, Sistan-Balouchestan Province. J Molec Biol Res 7:42–46

Bacles CFE, Lowe AJ, Ennos RA (2006) Effective seed dispersal across a fragmented landscape. Science 311:628. https://doi.org/10.1126/science.1121543

Beerli P, Felsenstein J (1999) Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. Genetics 152:763–773. https://doi.org/10.1073/pnas.081068098

Beerli P, Palczewski M (2010) Unified framework to evaluate panmixia and migration direction among multiple sampling locations. Genetics 185:313–326. https://doi.org/10.1534/genetics.109.112532

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Amer J Human Genet 32:31–34

Bush ABG, Little EC, Rokosh D, White D, Rutter NW (2004) Investigation of the spatio-temporal variability in Eurasian Late Quaternary loess-paleosol sequences using a coupled atmosphere-ocean general circulation model. Quatern Sci Rev 23:481–498. https://doi.org/10.1016/j.quascirev.2003.08.009

Charlesworth D, Willis JH (2009) The genetics of inbreeding depression. Nat Rev Genet 10:783–796. https://doi.org/10.1038/nrg2664

Chidumayo EN (2006) Fitness implications of clonal integration and leaf dynamics in a stoloniferous herb, *Nelsonia canescens* (Lam.) Spreng (Nelsoniaceae). Evol Ecol 20:59–73. https://doi.org/10.1007/s10682-005-3873-9

Cole CT (2003) Genetic variation in rare and common plants. Annual Rev Ecol Evol Syst 34:213–237. https://doi.org/10.1146/annurev.ecolsys.34.030102.151717

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144:2001–2014. https://doi.org/10.3892/ijo_00000551

Cornuet JM, Ravigne A, Estoup A (2010) Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v. 1.00.). BMC Bioinf 11:401. https://doi.org/10.1186/1471-2105-11-401

Cun YZ, Wang XQ (2010) Plant recolonization in the Himalaya from the southeastern Qinghai-Tibetan Plateau: geographical isolation contributed to high population differentiation. Molec Phylogen Evol 56:972–982. https://doi.org/10.1016/j.ympev.2010.05.007

Dutech C, Joly HI, Jarne P (2004) Gene flow, historical population dynamics and genetic diversity within French Guianan populations of a rainforest tree species, *Vouacapoua americana*. Heredity 92:69–77. https://doi.org/10.1038/sj.hdy.6800384

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genet Resources 4:359–361. https://doi.org/10.1007/s12686-011-9548-7

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. Molec Ecol 14:2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol Bioinf 1:47–50. https://doi.org/10.1177/117693430500100003

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491. https://doi.org/10.0000/genetics.org/content/131/2/479

Falahati-Anbaran M, Habashi AA, Esfahany M, Mohammadi SA, Ghareyazie B (2007) Populaltion genetic structure based on SSR markers in alfalfa (*Medicago sativa* L.) from various regions contiguous to the centres of origin of the species. J Genet 86:59–63. https://doi.org/10.1007/s12041-007-0008-9

Gaut BS, Seymour DK, Liu Q, Zhou Y (2018) Demography and its effects on genomic variation in crop domestication. Nat Pl 4:512–520. https://doi.org/10.1038/s41477-018-0210-1

Ghebru B, Schmidt RJ, Bennetzen JL (2002) Genetic diversity of Eritrean sorghum landraces assessed with simple sequence repeat (SSR) markers. Theor Appl Genet 105:229–236. https://doi.org/10.1007/s00122-002-0929-x

Glotzbecker GJ, Walters DM, Blum MJ (2016) Rapid movement and instability of an invasive hybrid swarm. Evol Appl 9:741–755. https://doi.org/10.1111/eva.12371

Guo YP, Zhang R, Chen CY, Zhou DW, Liu JQ (2010) Allopatric divergence and regional range expansion of *Juniperus sabina* in China. J Syst Evol 48:153–160. https://doi.org/10.1111/j.1759-6831.2010.00073.x

He F, Liu S, Yan Z, Ma Q (2007) The resources distribution, protection and utilization of wild *Allium mongolicum* Regel. Wild Pl Res 26:14–17

Henn BM, Botigué LR, Bustamante CD, Clark AG, Gravel S (2015) Estimating the mutation load in human genomes. Nat Rev Genet 16:333–343. https://doi.org/10.1038/nrg3931

Hewitt GM (1996) Some genetic consequence of ice ages, and their role in divergence and speciation. Biol J Linn Soc 58:247–276. https://doi.org/10.11111/j.1095-8312.1996.tb01434

Hu J, Hu XK, Zhang Q, Zhang JH, Fan BL, Yu QS (2017) Development of SSR molecular markers for *Allium mongolicum*. Genes Genom 39:1387–1394. https://doi.org/10.1007/s13258-017-0601-0

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23:1801–1806. https://doi.org/10.1093/bioinformatics/btm233

Li ZH, Chen J, Zhao GF, Guo YP, Kou YX, Ma YZ, Wang G, Ma XF (2012) Response of a desert shrub to past geological and climatic change: a phylogeographic study of *Reaumuria soongorica* (Tamaricaceae) in western China. J Syst Evol 50:351–361. https://doi.org/10.1111/j.1759-6831.2012.00201.x

Liu J, Möller M, Provan J, Gao LM, Poudel RC, Li DZ (2013) Geological and ecological factors drive cryptic speciation of yews in a biodiversity hotspot. New Phytol 199:1093–1108. https://doi.org/10.1111/nph.12336

Liu Q, Zhou Y, Morrell P, Gaut B (2017) Deleterious variants in Asian rice and the potential cost of domestication. Molec Biol Evol 34:908–924. https://doi.org/10.1093/molbev/msw296

Liu SW, Zhao T, Yang ML (2007) Analysis of the volatile oil from stem of *Allium mongolicum* Regel by GC–MS. West China J Pharm Sci 22:313–314

Lohmueller KE (2014) The distribution of deleterious genetic variation in human populations. Curr Opin Genet Developm 29:139–146. https://doi.org/10.1016/j.gde.2014.09.005

Lu W (2015) Phylogeography of *Nitraria roborowskii* based on cpDNA sequences (Master dissertation). Shihezi University, Shihezi

Luikart G (1997) Usefulness of molecular markers for detecting population bottlenecks and monitoring genetic change. PhD Thesis, University of Montana, Missoula

Luikart G, Cornuet JM (1998) Empirical evaluation of a test for identifying recently bottlenecked population from allele frequency data. Conservation Biol 12:228–237. https://doi.org/10.1111/j.1523-1739.1998.96388.x

Ma QL, Liu SZ, Yan ZZ, He FY (2006) Photosynthetic physiological characteristics of planted *Allium mongolicum*. Acta Bot Bor-Occid Sin 26:127–132

Ma SM, Zhang ML (2012) Phylogeography and conservation genetics of the relic *Gymnocarpos przewalskii* (Caryophyllaceae) restricted to northwestern China. Conservation Genet 13:1531–1541. https://doi.org/10.1071/BT11055

Ma YQ (1994) Inner Mongolia Flora, 2nd edn. Inner Mongolia People's Press, Huhhot

Marriage TN, Hudman S, Mort ME, Orive ME, Shaw RG, Kelly JK (2009) Direct estimation of the mutation rate at dinucleotide microsatellite loci in *Arabidopsis thaliana* (Brassicaceae). Heredity 103:310–317. https://doi.org/10.1038/hdy.2009.67

Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. Molec Ecol 7:639–655. https://doi.org/10.1046/j.1365-294x.1998.00374.x

Meng HH, Gao XY, Huang JF, Zhang ML (2015) Plant phylogeography in arid Northwest China: retrospectives and perspectives. J Syst Evol 53:33–46. https://doi.org/10.1111/jse.12088

Moola FM, Vasseur L (2009) The importance of clonal growth to the recovery of *Gaultheria procumbens* L. (Ericaceae) after forest disturbance. Pl Ecol 201:319–337. https://doi.org/10.1007/s11258-008-9496-9

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Palmé AE, Semerikov V, Lascoux M (2003) Absence of geographical structure of chloroplast DNA variation in sallow, *Salix caprea* L. Heredity 91:465–474. https://doi.org/10.1038/sj.hdy.6800307

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molec Ecol Notes 6:288–295. https://doi.org/10.1111/j.1471-8286.2005.01155.x

Pereyra RT, Bergström L, Kautsky L, Johannesson K (2009) Rapid speciation in a newly opened postglacial marine environment, the Baltic Sea. BMC Evol Biol 9:70

Petit RJ, Kremer A, Wagner DB (1993) Finite island model for organelle and nuclear genes in plants. Heredity 71:630–641. https://doi.org/10.1038/hdy.1993.188

Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. J Hered 90:502–503. https://doi.org/10.1093/jhered/90.4.502

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Qian CJ, Yin HX, Shi Y, Zhao JC, Yin CL, Luo WY, Dong ZB, Chen GX, Yan X, Wang XR, Ma XF (2016) Population dynamics of *Agriophyllum squarrosum*, a pioneer annual plant endemic to mobile sand dunes, in response to global climate change. Sci Rep 6:1–12. https://doi.org/10.1038/srep26613

Rebering CA, Schneeweiss GM, Bardy KE, Schonswetter P, Villasenor JL, Overnayer R, Stuessy TF, Weiss-Schneeweiss H (2010) Multiple Pleistocene refugia and Holocene range expansion of an abundant southwestern American desert plant species (*Melampodium leucanthum*, Asteraceae). Molec Ecol 19:3421–3443. https://doi.org/10.1111/j.1365-294X.2010.04754.x

Roiloa SR, Retuerto R (2007) Responses of the clonal *Fragaria vesca* to microtopographic heterogeneity under different water and light conditions. Environm Exp Bot 61:1–9. https://doi.org/10.1016/j.envexpbot.2007.02.006

Rosenberg NA (2004) Distruct: a program for the graphical display of population structure. Molec Ecol Notes 4:137–138. https://doi.org/10.1046/j.1471-8286.2003.00566.x

Rundle HD, Nosil P (2005) Ecological speciation. Ecol Lett 8:336–352. https://doi.org/10.1111/j.1461-0248.2004.00715.x

Salameen FA, Habibi N, Kumar V, Amad SA, Dashti J, Talebi L, Doaij BA (2018) Genetic diversity and population structure of *Haloxylon salicornicum* moq. in Kuwait by ISSR Markers. PLoS ONE 13:e0207369. https://doi.org/10.1371/journal.pone

Schaal BA, Hayworth DA, Olsen KM, Rauscher JT, Smith WA (1998) Phylogeographic studies in plants: problems and prospects. Molec Ecol 7:465–474. https://doi.org/10.1046/j.1365-294x.1998.00318.x

Schluter D (2000) The ecology of adaptive radiation. Oxford University Press, Oxford

Sechenbater LXM (2002) Nutritinal compositions and ethnobotany of *Allium mongolicum* Regel. Grassl China 24:52–54. https://doi.org/10.1088/1009-1963/11/5/313

Su ZH, Zhang ML, Sanderson SC (2011) Chloroplast phylogeography of *Helianthemum songaricum* (Cistaceae) from northwestern China: implications for preservation of genetic diversity. Conservation Genet 12:1525–1537. https://doi.org/10.1007/s10592-011-0250-9

Sun H, Li ZM (2003) Qinghai-Tibet Plateau uplift and its impact on tethys flora. Advances Earth Sci 18:852–862

Sun JM, Liu TS (2006) The age of the Taklimakan desert. Science 312:1621. https://doi.org/10.1126/science.1124616

Vigouroux Y, Jaqueth JS, Matsuoka Y, Smith OS, Beavis WD, Smith JSC, Doebley J (2002) Rate and pattern of mutation at microsatellite loci in Maize. Molec Biol Evol 19:1251–1260. https://doi.org/10.1093/oxfordjournals.molbev.a004186

Walter R, Epperson B (2004) Microsatellite analysis of spatial structure among seedlings in populations of *Pinus strobes* (Pinaceae). Amer J Bot 91:549–557. https://doi.org/10.2307/4123683

Wan DS, Feng JJ, Jiang DC, Mao KS, Duan YW, Miehe G, Opgenoorth L (2016) The quaternary evolutionary history, potential distribution dynamics, and conservation implications for a Qinghai-Tibet Plateau endemic herbaceous perennial, *Anisodus tanguticus* (Solanaceae). Ecol Evol 6:1977–1995. https://doi.org/10.1002/ece3.2019

Wang FT, Tang T, Chen SC, Xu JM, Liang SY, Tsi ZH, Lang KY, Mao ZM, Shue LZ (1980) Liliaceae. Flora of China (14). Science Press, Beijing, pp, 224–226

Wang N, Yu FH, Li PX, He WH, Liu FH, Liu FH, Liu JM (2008) Clonal integration affects growth, photosynthetic efficiency and biomass allocation, but not the competitive ability, of the alien invasive *Alternanthera philoxeroides* under severe stress. Ann Bot (Oxford) 101:671–678. https://doi.org/10.1093/aob/mcn005

Wang Q, Abbott RJ, Yu QS, Lin K, Liu JQ (2013) Pleistocene climate change and the origin of two desert plant species, *Pugionium cornutum* and *Pugionium dolabratum* (Brassicaceae), in northwest China. New Phytol 199:277–287. https://doi.org/10.1111/nph.12241

Wang Z, Weber JL, Zhang G, Tanksley SD (1994) Survey of plant short tandem DNA repeats. Theor Appl Genet 88:1–6. https://doi.org/10.1007/bf00222386

Wegener JE, Pita-Aquino JN, Atutubo J, Moreno A, Kolbe JJ (2019) Hybridization and rapid differentiation after secondary contact between the native green anole (*Anolis carolinensis*) and the introduced green anole (*Anolis porcatus*). Ecol Evol 9:4138–4148. https://doi.org/10.1002/ece3.5042

Willis KJ, Niklas KJ (2004) The role of quaternary environmental change in plant macroevolution: the exception or the rule? Philos Trans Roy Soc B Biol Sci 359:159–172. https://doi.org/10.1098/rstb.2003.1387

Wolfe KH, Li WH, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondria, chloroplast, and nuclear DNAs. Proc Natl Acad Sci USA 84:9054–9058. https://doi.org/10.1073/pnas.84.24.9054

Yang T, Fang L, Zhang X, Hu J, Bao S, Hao J, Li L, He Y, Jiang J, Tian S, Zong X (2015) High-throughput development of SSR markers from Pea (*Pisum sativum* L.) based on next generation sequencing of a purified Chinese commercial variety. PLoS ONE 10:2. https://doi.org/10.1371/journal.pone.0139775

Yang Z, Hao L, Zhang F (2007) Seed germination and changes in storage substance contents of *Allium mongolicum* Regel. Pl Physiol Commun 43:173–175

Yin HX, Yan X, Shi Y, Qian CJ, Li ZH, Zhang W, Wang LR, Li Y, Li XZ, Chen GX, Li XR, Nevo E, Ma XF (2015) The role of East Asian monsoon system in shaping population divergence and dynamics of a constructive desert shrub *Reaumuria soongarica*. Sci Rep 5:15823. https://doi.org/10.1038/srep15823

You WH, Yu D, Liu CH, Xie D, Xiong W (2013) Clonal integration facilitates invasiveness of the alien aquatic plant *Myriophyllum aquaticum* L. under heterogeneous water availability. Hydrobiologia 718:27–39. https://doi.org/10.1007/s10750-013-1596-4

Yu FH, Wang N, He WM, Chu Y, Dong M (2008) Adaptation of rhizome connections in drylands: increasing tolerance of clones to wind erosion. Ann Bot (Oxford) 102:571–577. https://doi.org/10.1093/aob/mcn119

Yu QS, Wang Q, Wu GL, Ma YZ, He XY, Wang X, Xie PH, Hu LH, Liu JQ (2013) Genetic differentiation and delimitation of *Pugionium dolabratum* and *Pugionium cornutum* (Brassicaceae). Pl Syst Evol 299:1355–1365. https://doi.org/10.1007/s00606-013-0800-3

Zhang Q, Chiang TY, Liu JQ, Abbott RJ (2005) Phylogeography of the Qinghai-Tibetan Plateau endemic *Juniperus przewalskii* (Cupressaceae) inferred from chloroplast DNA sequence variation. Molec Ecol 14:3513–3524. https://doi.org/10.1111/j.1365-294X.2005.02677.x

Zhang YH, Liu SZ, Yu QS, He FL, Zhang JH (2014) Geographical distribution and seed characteristics of *Allium mongolicum* in China. J Desert Res 34:391–395

Zhang YH, Yu QS, Zhang Q, Hu XK, Hu J, Fan BL (2017) Regional-scale differentiation and phylogeography of a desert plant *Allium mongolicum* (Liliaceae) inferred from chloroplast DNA sequence variation. Pl Syst Evol 303:451–466. https://doi.org/10.1007/s00606-016-1383-6

Zhao JH (2010) A study on ecological adaptability and reproduction feature of three wild *Allium* plants. PhD Thesis, Inner Mongolia Agricultural University, Huhhot

Zhou Y, Duvaux L, Ren G, Zhang L, Savolainen O, Liu J (2017a) Importance of incomplete lineage sorting and introgression in the origin of shared genetic variation between two closely related pines with overlapping distributions. Heredity 118:211–220. https://doi.org/10.1038/hdy.2016.72

Zhou Y, Massonnet M, Sanjak J, Cantu D, Gaut B (2017b) Evolutionary genomics of grape domestication. Proc Natl Acad Sci USA 114:11715–11720. https://doi.org/10.1073/pnas.1709257114