

Genetic diversity and fine-scale spatial genetic structure of the near-threatened *Pinus gerardiana* in Gardiz, Afghanistan

Sayed Jalal Moosavi¹, Katharina Birgit Budde¹, Markus Mueller¹, Oliver Gailing¹

¹ Forest Genetics and Forest Tree Breeding, Georg-August-University Goettingen, Goettingen, Germany

Corresponding author: Oliver Gailing (ogailin@gwdg.de)

Academic editor: Myriam Heuertz ♦ Received 2 March 2022 ♦ Accepted 28 June 2022 ♦ Published 5 October 2022

Abstract

Background and aims – Chilgoza pine (*Pinus gerardiana*) is a near-threatened tree species from the north-western Himalayas. This species is the economically most important pine in Afghanistan because of its edible nuts; however, its distribution range is disjunct and restricted to a few isolated regions. The IUCN lists Chilgoza as a near threatened species because of overexploitation of its nuts and a declining population trend. This research is the first in-depth analysis of the genetic variability and structure of Chilgoza in Afghanistan using microsatellite markers.

Material and methods – We tested cross-amplification of 44 SSR markers developed for pine species. Eight polymorphic EST-SSRs were genotyped in a natural Chilgoza population in Gardiz, Afghanistan. To evaluate the genetic diversity, fine-scale spatial genetic structure (SGS), signatures of bottleneck events, and the effective population size, 191 trees were sampled and genotyped. Based on the diameter at breast height, individuals were classified as young or old trees.

Key results – Genetic variation in the whole population was moderate. For individual markers, H_e ranged from 0.130 to 0.515 (mean = 0.338) and H_o from 0.118 to 0.542 (mean = 0.328). The expected heterozygosity in young trees was slightly lower than in old trees. The SGS was stronger for young trees ($Sp = 0.0100$) than for old trees ($Sp = 0.0029$). Heterozygosity excess analysis detected no recent population size reduction, but the M ratio revealed an ancient and prolonged bottleneck in the Chilgoza population.

Conclusion – Identification of suitable EST-SSRs for future studies of natural Chilgoza populations provides important tools for the conservation of the species. Despite the moderate genetic variation in Gardiz, scarcity of natural regeneration is likely to reduce the genetic variation and adaptability in future generations. Our results indicated a slight decrease in genetic diversity and stronger SGS in young trees calling for conservation measures fostering natural regeneration.

Keywords

Chilgoza, effective population size, fine-scale spatial genetic structure, genetic diversity, microsatellites, *Pinus gerardiana*

INTRODUCTION

Tree species are the dominant species in forest ecosystems worldwide, providing essential ecosystem services, such as soil and water conservation, and habitat and resources for associated species (e.g. Petit and Hampe 2006; Bonan 2008). Understanding the impacts of human land-use intensification and climate change on tree populations has been the focus of numerous studies during the last decades (e.g. reviewed in Milad et al. 2011). However, most studies focused on economically important and common tree

species, while rare and scattered tree species have received much less attention (Song et al. 2019). Furthermore, most studies focused on European and North American ecosystems, while ecosystems in other parts of the world are still underrepresented (Martin et al. 2012).

Pinus gerardiana Wall. ex D. Don (English: Chilgoza; Persian: Jalghoza, meaning “40 nuts”) is economically and ecologically important in Afghanistan. Nevertheless, it is threatened by overexploitation. Mitigating this threat by conservation and maintaining genetic diversity requires genetic information, which is not yet available.

Especially in Afghanistan, the economy and livelihood of the local population living close to these forests depend on the pine nuts trade. Nuts of this tree are edible (Fig. 1E) with high nutritive value and are used as a food ingredient in desserts, sauces, and salads (Cai et al. 2017; Singh et al. 2021). It obtains the highest price among all nuts in Afghanistan (~30 \$/kg, personal survey in November 2021 on the market in Paktia province). In recent years, the Chilgoza pine nut price increased due to increasing export (Comprehensive Agriculture and Rural Development Facility [CARD-F] 2017). Additionally, Chilgoza pine forests provide fuelwood, medicine, and pasture for the local population.

According to the International Union for Conservation of Nature (IUCN) Red List version 3.1, Chilgoza is considered near threatened due to habitat fragmentation and a decreasing population trend (Farjon 2013). Especially in Afghanistan, Chilgoza pine forests are under severe threat due to poor natural regeneration. The main factor leading to poor natural regeneration is the overexploitation of pine cones (Kumar et al. 2016). Harvesting by heavy pruning (removal of branches), grazing livestock, and logging further contribute to the population decline (Ahmed et al. 1991; Siddiqui et al. 2009; Groninger 2012; Kumar et al. 2013). Since Chilgoza produces 30 to 118 seeds per cone (which is directly related to the environmental conditions), a high number of seeds is generated on each tree. However, a sufficient number of cones must be left on each crown each year to promote natural regeneration (Shalizi et al. 2016) and ensure the populations' survival in the next generations. Chilgoza seed viability and germination rates are low, and seedlings may be damaged by many factors, including a shortage of soil moisture, intense heat, desiccating winds and overgrazing (Kumar et al. 2013). Natural regeneration in the region ranges from no regeneration to poor/fair regeneration (Groninger and Ruffner 2010; Shalizi et al. 2018). Afghanistan's political instability and local warfare (started in 1978) are other important factors directly related to deforestation. An increase in illegal logging (for smuggling and firewood), poor forest management, lack of community involvement and awareness, conversion to agricultural and urban use, and limited investment in reforestation are the consequences of war and instability.

Chilgoza pine occurs in dry valleys of eastern Afghanistan, contiguous northern and north-western Pakistan, north-western India, and Tibet and Xizang province of China (Critchfield and Little 1966; Bonner and Karrfalt 2008) (Fig. 2A). It mainly grows in an altitudinal range of 2000 to 3350 m a.s.l. (Thomas 2019; Singh et al. 2021). In Afghanistan, the species is limited to the eastern forest complex (Fig. 2A) (Singh et al. 1973; Bonner & Karrfalt 2008; Groninger 2012). It forms a closed belt below *Cedrus deodara* (Roxb. ex D. Don) G. Don populations, and in most places, it grows in mixed stands with *Quercus dilatata* Royle and *Quercus baloot* Griff. In the eastern surroundings of Gardiz city, the species forms monospecific stands (Masumy 2006; Wingard et al. 2010).

The species is wind-pollinated, and the large, wingless seeds (Fig. 1E) are most likely dispersed by the large spotted nutcracker *Nucifraga multipunctata*, endemic to the distribution area of *P. gerardiana* (Lanner 1990; Madge et al. 2020). Typically, the seeds of bird-dispersed pines are stored in closed cones that the birds crack open to collect and store in caches. In other better-studied stone pine species, e.g. in *Pinus cembra* L., the behaviour of the seed-dispersing bird was shown to determine seedling recruitment at the upper altitudinal range limits of the tree species (Neuschulz et al. 2018). Seeds can be dispersed over quite long distances by the birds, as shown by Bekku et al. (2019), for *Pinus parviflora* Siebold & Zucc. in Japan, which can affect the spatial genetic structure of the tree species. The abundance of the essential seed dispersal agent, the large spotted nutcracker, is unknown in the study region.

The basic prerequisite for developing suitable strategies to protect genetic resources is the study of genetic variability. Therefore, a primary objective of conservation genetics is to estimate the level and distribution of genetic variation in endangered species to optimize sampling strategies for conserving and managing genetic resources (Pautasso 2009). To identify the level of genetic diversity, genetic markers are indispensable. Simple sequence repeat (SSR, also called microsatellite) markers are commonly used to estimate genetic diversity, population genetic structure, and differentiation in numerous plant species. Microsatellites have increasingly become the markers of choice for endangered and threatened species due to their codominant, highly polymorphic nature, and cost efficiency when processing high numbers of samples. Also, EST-SSRs developed from expressed sequence tags have a clear potential in basic evolutionary applications such as population genetic analysis, and are more likely to be transferable across taxonomic boundaries than intergenic SSRs (Ellis and Burke 2007). They reside in or near coding DNA and, consequently, they should be more conserved than non-coding genomic sequences (Liewlaksaneeyanawin et al. 2004). Therefore, a higher transferability of EST-SSR markers than of random nuclear SSRs is expected between related species. Transferability and polymorphism decrease in more distantly related species (Chagné et al. 2004).

Studies of natural genetic variation in *P. gerardiana* are scarce, and only a few genomic resources are available, e.g. the chloroplast genome published by Cronn et al. (2008). So far, there are no species-specific nuclear microsatellites available for *P. gerardiana*. However, genomic resources and gene-based microsatellites (EST-SSRs) with high potential for transferability across pine species have been developed for the related pine species *Pinus bungeana* Zucc. ex Endl. (Duan et al. 2017). Investigations on Chilgoza pine ecology (Kumar et al. 2013), seed germination (Saeed and Thanos 2006), genetic variability of phenotypic traits (Kant et al. 2006a), seed protein properties (Cai et al. 2013), storage and drying treatments (Malik et al. 2013; Thakur et al. 2014), and

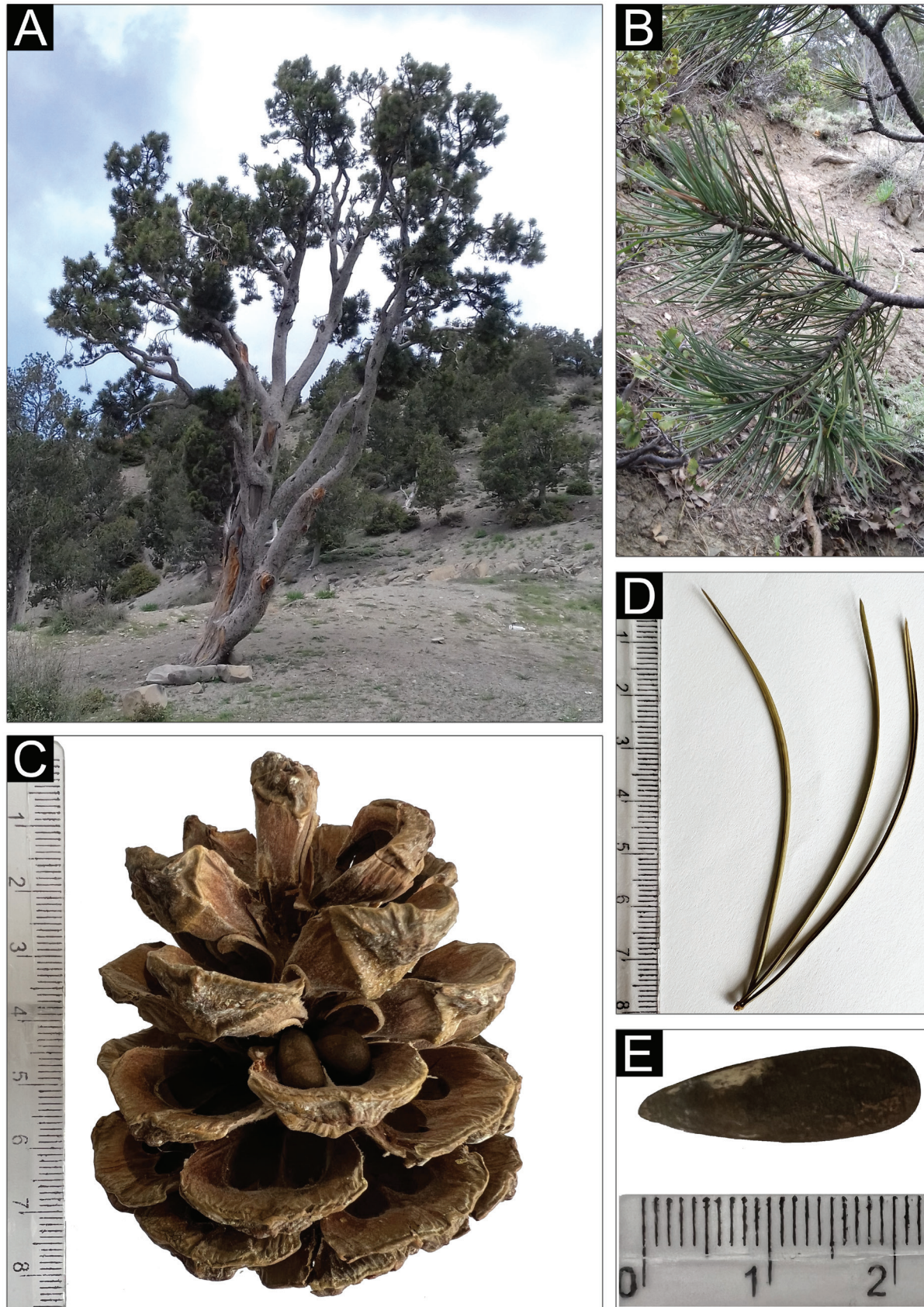


Figure 1. *Pinus gerardiana*. A. *Pinus gerardiana* tree in the sampling area, Gardiz, Afghanistan. B. Branches. C. A cone with two nuts inside. D. Needles in a fascicle of three. E. Nut. Photos by Sayed Jalal Moosavi.

morphometric characters (Ranot and Shrama 2015) have been carried out in other countries, especially in India and Pakistan, and also in Afghanistan (Khurram and Shalizi 2016; Safari et al. 2017; Alami and Mousavi 2020). In natural populations of *P. gerardiana*, high levels of genetic variation at RAPD (Randomly Amplified Polymorphism DNA) markers were found within populations in India (Kant et al. 2006b). However, as far as we know, Chilgoza pine populations in Afghanistan have never been thoroughly studied at the molecular level.

Here, we set out to identify suitable EST-SSR markers for *P. gerardiana* and to characterize the genetic variation in a natural stand in Gardiz, Afghanistan. We specifically compared different age classes of this vulnerable tree species. Due to overexploitation and a lack of natural regeneration, we hypothesized that younger trees might harbour lower levels of genetic diversity than older trees. Given the importance of Chilgoza pines for the local population in Afghanistan and the declining population trend, a better understanding of the genetic structure is crucial.

MATERIAL AND METHODS

Study population and plant material

We collected Chilgoza pine needle samples (Fig. 1D) from a total of 191 trees in four subpopulations in the vicinity of Gardiz, the capital of Paktia province, Afghanistan (Fig. 2B). The distribution map of Chilgoza, sampling area, and also the location of each individual are shown in Fig. 2. The maps were created using ArcMap v.10.8 (Esri, USA). The distance from Gardiz to the sampling location was ca 25 km. The maximum distance between sampled trees in this natural forest of the eastern forest complex was 726 m. The two southern subpopulations are adjacent and located in the lower elevation (lowest sample at 2617 m; Table 1), while the two northern subpopulations are spatially separated (highest sample at 2837 m; Table 1). We collected samples from all trees (DBH > 6 cm) in four subpopulations which were accessible from the road (see Fig. 2C) which assured the collection of trees in close vicinity, as well as more distant trees. Security in most forest regions is not stable; therefore, the sampling area was selected based on security and accessibility.

Diameter at breast height (DBH at 1.40 m) and GPS (GPSmap 60CSx, Garmin, USA) coordinates of each tree were recorded. In the near absence of natural regeneration, we only sampled trees with a DBH > 6 cm. The largest and smallest DBH were 65.6 cm and 6.36 cm, respectively. In the DBH histogram, the class intervals were adjusted to 5 cm to visualize the DBH distribution in more detail (Supplementary file 1: Fig. S1). We followed the definition by Ahmed et al. (1991), who considered Chilgoza pines with > 6 cm diameter as adult trees. The species is characterized by extremely slow growth (Ahmed and Sarangezai 1991), although growth and DBH

distribution depend on site conditions and growth stage. According to Ahmed et al. (1991), based on tree ring analysis, the growth rate of Chilgoza is very low, 43 years/cm or ca 0.2 mm/year in seedlings, and ca 0.8 mm/year was the average growth rate of all investigated trees in the mentioned study. Based on the range of DBH, according to Shalizi et al. (2018), the age of the Chilgoza trees in Paktia province is about 241 years at 51.6 cm, therefore the oldest sampled trees in this study could be about 250 years old. All sampled trees represent adult trees, however, to avoid overlapping generations in further analyses we grouped the trees into two DBH size classes, representing most likely younger (DBH = 6.36–28.65 cm) and older trees (29.28–65.57 cm).

Needles were stored in individual paper envelopes, and packages of 2 g silica gel (Tamad Kala, Iran) were placed inside the envelopes to dry the needles. After drying the samples in Afghanistan and transferring them to the University of Goettingen, Germany, the DNA of needles was extracted.

DNA extraction, marker testing, and genotyping

Genomic DNA was extracted according to the producer's protocol for silica gel dried needles with the DNeasy 96 Plant Kit (Qiagen, Germany). DNA quality and quantity were tested in 1.0% agarose gels stained with Roti®Gelstain (Carl ROTH, Germany) and then visualized under UV light and compared to a Lambda DNA size marker (Roche, Germany). Isolated DNA was used directly for PCR amplification without dilution.

Polymerase chain reactions (PCR) were performed with M13 tails (5'-CACGACGTTGTAAAACGAC-3') and dye labelled adaptors complementary to forward primers (Schuelke 2000; Kubisiak et al. 2013) and a PIG-tail (Brownstein et al. 1996) added to the reverse primers, using a touchdown program with the following protocol: first denaturation at 95°C for 15 min, followed by ten cycles including a denaturation step of 1 min at 94°C, an annealing step at 60°C for 1 min (-1°C per cycle), an extension step at 72°C for 1 min, then 25 cycles with the same denaturation and extension time and temperature, but 50°C annealing for 1 min, and a final extension at 72°C for 20 min.

The PCR mix was composed of 1.0 µL genomic DNA (ca 10 ng/µL), 1.5 µL 10x reaction buffer B (Solis BioDyne, Estonia), 1.5 µL MgCl₂ (25 mM), 1.0 µL dNTPs (2.5 mM each dNTP), 0.2 µL (5 U/ µL) HOT FIREPol® Taq DNA polymerase (Solis BioDyne, Estonia), 0.2 µL tailed (Schuelke 2000; Kubisiak et al. 2013) forward primer (5 pM/µL), 0.5 µL PIG-tailed (Brownstein et al. 1996) reverse primer (5 pM/µL), 1 µL (5 pM/µL) dye labelled (6-FAM or HEX) M13 primer, and HPLC grade H₂O (filled up to a volume of 13 µL).

Forty-four SSR primers, including 19 EST-SSRs from *P. bungeana* (Duan et al. 2017), eight chloroplast SSRs (cpSSRs) from *Pinus thunbergii* Parl. (Vendramin et al. 1996), six EST-SSRs and six genomic SSRs from *Pinus taeda*

L., and five EST-SSRs from *Pinus halepensis* Mill. (Elsik et al. 2000; Auckland et al. 2002; Liewlaksaneeyanawin et al. 2004; Leonarduzzi et al. 2016) were tested in eight samples from all subpopulations for amplification and detection of polymorphism. Amplification of almost all markers was successful as visualized on agarose gels (Supplementary file 1: Table S1). However, when separating amplification products on an ABI 3130xl Genetic Analyzer (Applied

Biosystems, USA), polymorphism was observed in only eight markers from *P. bungeana* (Table 2), the closest relative of *P. gerardiana* (Yang et al. 2016). Ultimately, the eight polymorphic markers were selected and genotyped in all samples. The markers 33255, 34533, and 66538 were amplified in a multiplex, while the remaining markers were amplified in separate PCRs (Supplementary file 1: Table S1).

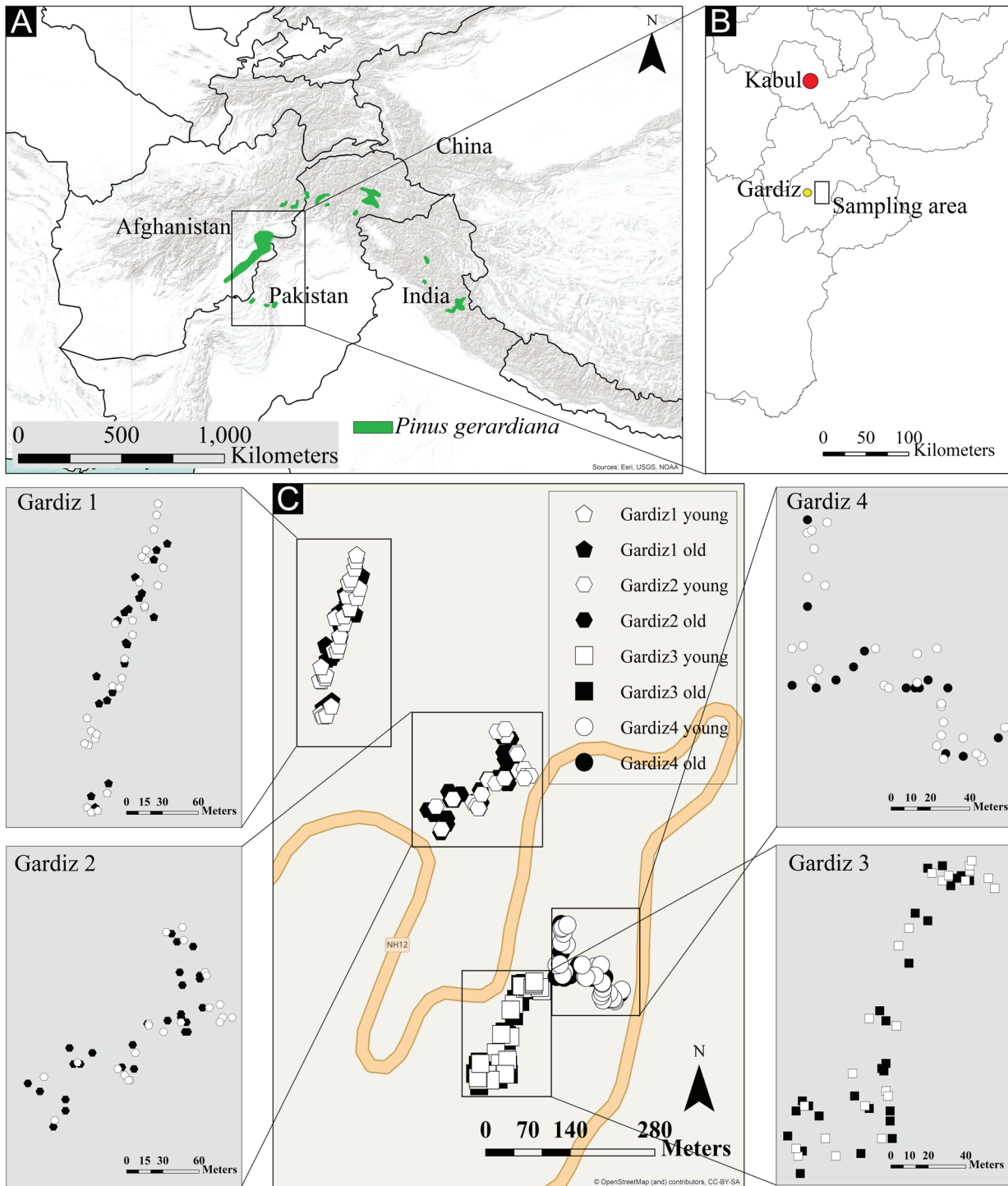


Figure 2. *Pinus gerardiana* maps. A. Distribution map of *Pinus gerardiana*, redrawn from Critchfield and Little (1966). B. Location of the sampling area in Afghanistan. C. Location of each *Pinus gerardiana* individual sampled in the eastern forest complex, in Gardiz, Afghanistan in the vicinity of national highway 12 (NH12).

Table 1. Characteristics of the four sampled subpopulations in the *Pinus gerardiana* population close to Gardiz, Afghanistan.

Sub-population	Number of samples	Mean altitude (m)	Geographical coordinate	Mean DBH (cm)	Area (ha)
Gardiz1	50	2820.08	33°28'49.6"N, 69°23'05.6"E	29.79	1.26
Gardiz2	49	2714.86	33°28'40.7"N, 69°23'12.6"E	33.36	1.00
Gardiz3	50	2676.08	33°28'30.2"N, 69°23'14.0"E	31.22	1.08
Gardiz4	42	2636.48	33°28'33.9"N, 69°23'18.0"E	29.94	1.22

Table 2. Characteristics of the polymorphic EST-SSRs used in this study. ^M, multiplex PCRs; ^S, separate PCRs; ^F, 6-FAM fluorescent dye; ^H, HEX fluorescent dye.

Marker name	Repeat motif	Observed size range (bp)	Sequence (5'-3')
33255 ^{M,F}	(AAGGC) ₅ (GAG) ₅	200–220	F: TCAGCAACCAAACCATACCA R: TGCACTCGCTCCCTATCTTT
34533 ^{M,F}	(CTCACC) ₆	265–277	F: ATCTCGGCCAATTTGTGCATC R: TTGGTCCACCTTTCATCCTC
66538 ^{M,F}	(GGGCGA) ₄	295–301	F: ATATTGATCAGGCGAGGCAG R: GGATTGTTGCAGGTTTTTCGT
24177 ^{S,H}	(GGCTGC) ₄	262–292	F: CTGGGGAGTATGCACACCTT R: CAGTATCAACAGCAAGCCCA
7028 ^{S,H}	(TTC) ₈	230–251	F: AGCCATTTCTTCTGCTTCCA R: TTTTCACCCATTCTCCTTCG
10962 ^{S,F}	(TA) ₁₁	273–277	F: CGGCCTTTCACTTCTGGTAG R: TGCTGACAAACAAACCGAGA
3534 ^{S,F}	(AT) ₁₂	283–289	F: AAGCATCTGCACCTATGGG R: GTGGAATTGAGATCGGCTGT
72763 ^{S,H}	(AAAACC) ₄	274–282	F: GGCAATTCTGCAGTAGCCTC R: ATGGTCTGTCCATTTCCGGTG

Table 3. Genetic variation over all samples for each locus. N, number of successfully genotyped samples; N_a , number of alleles; N_{ae} , effective number of alleles; H_e , expected heterozygosity; H_o , observed heterozygosity; F_{is} , inbreeding coefficient; p values of F_{is} .

Locus	#N	N_a	N_{ae}	H_e	H_o	F_{is}	p value (F_{is})
33255	177	7	1.920	0.480	0.542	-0.130	0.004
34533	189	3	1.240	0.190	0.175	0.085	0.230
66538	191	2	1.300	0.228	0.230	-0.010	0.942
24177	188	8	2.070	0.515	0.468	0.093	0.038
7028	191	8	1.890	0.472	0.466	0.013	0.780
10962	184	3	1.640	0.389	0.364	0.066	0.294
3534	180	4	1.440	0.304	0.261	0.142	0.023
72763	187	3	1.150	0.130	0.118	0.096	0.232

Table 4. Genetic variation over seven loci without null alleles for young (DBH < 29 cm) and old (DBH > 29 cm) trees in *Pinus gerardiana* from Gardiz. N, number of samples; N_a , number of alleles; N_{ae} , effective number of alleles; A_R (k = 28), allelic richness of a standardized sample of 28 gene copies; H_e , expected heterozygosity; H_o , observed heterozygosity; F_{is} , inbreeding coefficient; p values of F_{is} .

Cohort	#N	N_a	N_{ae}	A_R (k = 28)	H_e	H_o	F_{is}	p value (F_{is})
Gardiz (young)	101	4.430	1.590	3.200	0.334	0.319	0.046	0.130
Gardiz (old)	90	4.430	1.620	3.170	0.354	0.358	-0.011	0.719
Mean		4.430	1.605	3.180	0.344	0.338		

In order to pool the PCR products from different loci with similar sizes, the PCR reactions were carried out with M13 primers labelled with fluorescent dyes 6-FAM (blue) or HEX (green, see Table 2). Fragment analysis was performed on an ABI 3130xl Genetic Analyzer with the internal size standard GS 500 ROX (Applied Biosystems, USA). GeneMapper v.4.0 (Applied Biosystems, USA) was used for visualization and fragment size calling of the PCR products.

Genetic variation and differentiation

To check for the presence of null alleles, MICRO-CHECKER v.2.2.3 was used (Van Oosterhout et al. 2004). Genetic diversity estimates, number of alleles per locus (N_a), effective number of alleles (N_{ae}), allelic richness on a standardized sample of 28 gene copies (A_R), and observed and expected heterozygosity (H_o , H_e), as well as the inbreeding coefficient (F_{is}) were calculated using SPAGeDi 1.5d (Hardy and Vekemans 2002) for two diameter classes (6.36–28.65 cm: young trees, 29.28–65.57 cm: old trees) over all loci and also for each SSR marker separately. The significance of differences of genetic diversity parameters, observed and expected heterozygosity, among the four subpopulations and the two diameter classes, were tested using a Kruskal-Wallis test with multiple comparisons implemented in the R package *pgirmess* v.1.6.7 (Giraudeau 2021). Also, associations between individual heterozygosity (number of heterozygote loci/sample) and DBH, and between altitude and DBH were evaluated by Pearson's correlation coefficient. To visualize the result of the correlation between heterozygosity and DBH, a scatter smooth plot was generated in the R package *stats* v.4.0.5 (R Core Team 2021).

Genetic population structure was assessed and quantified in two steps. First, a Principal Coordinates Analysis (PCoA) was performed in GenALEX v.6.5 (Peakall and Smouse 2012) based on Nei's unbiased genetic distances (Nei 1978) between individual samples. In addition, the population genetic structure based on EST-SSRs was inferred using STRUCTURE v.2.3.4 (Pritchard et al. 2000). Specifically, we used the admixture model and correlated allele frequencies. The Markov Chain Monte Carlo (MCMC) iterations included a burn-in period of 10,000 iterations followed by 100,000 Markov Chain iterations with ten replicates for $K = 1$ to $K = 4$. In order to infer the most likely number of clusters (K), the Evanno method (Evanno et al. 2005) implemented in STRUCTURE HARVESTER v.0.6.94 was used (Earl and vonHoldt 2012). The STRUCTURE results were uploaded to the CLUMPAK pipeline (Kopelman et al. 2015), and cluster assignment bar plots were created by averaging over all replicates per K .

To assess the fine-scale SGS using SPAGeDi 1.5d (Hardy and Vekemans 2002), the Loiselle kinship coefficient F (Loiselle et al. 1995) was estimated for all pairs of samples and regressed on the logarithm of spatial distances. Significance was assessed by comparing the

regression slope b_F with its distribution obtained from 10,000 permutations of individuals among locations. The strength of the fine-scale SGS was computed as $Sp = -b_F / (F_1 - 1)$, where b_F is the regression slope, and F_1 is the mean kinship coefficient of pairs of individuals belonging to the first distance class (Vekemans and Hardy 2004). This analysis was performed on each diameter class separately to avoid parent-offspring pairs. In SPAGeDi, we applied 10 m intervals for each distance class to ensure that at least thirty individual pairs were represented in each distance class. To further evaluate the spatial genetic structure, we performed a Spatial Principal Component Analysis (sPCA, Jombart et al. 2008) implemented in the R package *adegenet* v.2.1.5 (Jombart 2008). We tested for global structure (G-test) reflecting autocorrelation due to family patches or clines and local structure (L-test) indicating elevated differentiation between neighboring individuals in young and old trees, using 1,000 Monte Carlo simulations.

In addition, to test for signals of past demographic changes, the T_2 statistic implemented in INEst v.2.2 (Chybicki and Burczyk 2009), originally described by Cornuet and Luikart (1996), was used. Since microsatellite markers and only eight loci were used in this study, an appropriate test and model were the Wilcoxon signed-rank test (recommended if the number of markers is less than 20) and the Two-Phase Model (TPM, Piry et al. 1999). However, some microsatellites are known to rather follow an Infinite Allele Model (IAM). As we had no prior knowledge about the most adequate mutation model for our EST-SSRs, we report the test statistics of both models, TPM and IAM. INEst v.2.2 provides improved p value estimation for the Wilcoxon signed-rank test based on 10^6 permutations (Chybicki 2017). We used the default settings for the TPM (proportion of multi-step mutations = 0.22 and average multi-step mutation size = 3.1) to assess bottleneck effects in INEst v.2.2. In addition, we report the M ratio (Garza and Williamson 2001), the ratio of the number of observed alleles over the number of expected alleles in the allele size range, which was also assessed using INEst. Both tests were run on the seven loci that did not show any signs of null alleles.

The effective population size (N_e) was estimated using the linkage disequilibrium method (Hill 1981; Waples 2006; Waples and Do 2010), as implemented in NeEstimator v.2.1 (Do et al. 2014). We estimated the effective population size (N_e) separately for the old and young cohorts and reported parametric confidence intervals based on a chi-square approximation (Waples 2006). We used a threshold of 0.02 as the lowest allele frequency to estimate the N_e values.

RESULTS

DBH sizes were 6.36 cm to 65.57 cm (mean 31.19 cm). Few individuals showed very small or very large

DBH. Intermediate DBH values were most common (Supplementary file 1: Fig. S1).

Amplification of almost all EST-SSRs (18 of 19 markers) originally developed for *P. bungeana* was successful (95%). Eight of them were polymorphic (Supplementary file 1: Table S1), showing a high transferability rate of EST-SSR markers from *P. bungeana* to *P. gerardiana*.

While high transferability of the markers from different species was observed in Chilgoza pine (Supplementary file 1: Tables S1, S2), SSR markers from more distantly related *Pinus* species were not polymorphic. All eight cpSSRs from *P. thunbergii* amplified but were not polymorphic. Four out of five (80% of transferability) tested EST-SSR markers from *P. halepensis* amplified monomorphic bands. Of the six genomic SSR markers transferred from *P. taeda* none showed amplification. However, five out of the six EST-SSRs derived from *P. taeda* amplified but were monomorphic.

Single markers showed a broad range of F_{is} values (-0.130–0.142). Two markers had negative F_{is} values (33255 and 66538, Table 3). MICRO-CHECKER detected possible null alleles at low frequency at only one locus (3534).

The levels of genetic diversity were very similar between old and young trees (Table 4) and also between the subpopulations (Supplementary file 1: Table S3). A positive and low, but non-significant F_{is} value ($F_{is\text{-young}} = 0.046$, $p = 0.130$) indicated a slightly lower number of heterozygotes than expected under Hardy Weinberg equilibrium observed in the younger cohort. A negative but also non-significant F_{is} value ($F_{is\text{-old}} = -0.011$, $p = 0.719$) was observed in the old cohort indicating a higher number of observed heterozygotes than expected. The difference between observed and expected heterozygosity, reflected in F_{is} was not significant for any of the subpopulations or age classes (Supplementary file 1: Tables S3, S4, S5).

A total of 38 alleles were observed across 191 individuals. Based on the Kruskal-Wallis test, observed and expected heterozygosity differences among the four subpopulations and two age cohorts were not significant. Correlation analyses did not reveal a significant association between individual heterozygosity and DBH in the four subpopulations (Supplementary file 1: Fig. S2) and also not for elevation and DBH (data not shown).

PCoA (Principal Coordinates Analysis) based on Nei's unbiased genetic distance did not show a clustering by subpopulation or DBH class, indicating no genetic structure (Supplementary file 1: Fig. S3). Likewise, the STRUCTURE analysis did not identify any genetic clusters (Supplementary file 1: Fig. S4). The highest probability of our data was observed for $K = 1$ (Supplementary file 1: Figs S5, S6).

The fine-scale SGS indicated a significant family structure for young trees ($Sp = 0.0100$, $p = 0.0003$) and no significant family structure for the old trees (Table 5, Fig. 3). The results of the sPCA showed similar patterns. We detected significant global structures for both old and young trees, however, the pattern was more pronounced

for young trees compared to the old trees as indicated by a higher eigenvalue of the first sPCA axis (Table 5, Supplementary file 1: Figs S7, S8).

A significant heterozygosity excess in populations is interpreted as a recent bottleneck. However, under TPM and also under IAM, negative $T2$ values indicated no signs of a recent reduction in population size in young and old cohorts. Wilcoxon signed-rank test under TPM and IAM for heterozygosity excess rather pointed to a deficiency in heterozygosity (Table 6), indicating a population expansion in the past. The Garza-Williamson's M ratio, most suitable for detecting more ancient bottleneck signals, indicated a past bottleneck event in Gardiz based on the samples from both cohorts (Table 6).

The effective population size (when allowing for 0.02 as the lowest allele frequency) was estimated to be slightly lower for the old cohort ($N_{e,LD,0.02} = 101.3$, 95% CI = 47.2–553.8) than for the younger cohort ($N_{e,LD,0.02} = 401.1$, 95% CI = 101.8–infinite) but the confidence intervals largely overlapped.

DISCUSSION

Transferability of markers

The de novo development of SSRs is a time and cost-consuming procedure, especially for non-model species for which genomic resources are scarce. Most SSR primers are species-specific; hence, they cannot be transferred to other species. However, EST-SSRs are the best choice for obtaining high-quality gene-based nuclear microsatellite markers for non-model species because of their high transferability across closely related species (Ellis and Burke 2007). Due to the location of EST-SSRs in or close to expressed genes, they often show a lower genetic variation than genomic (non-genic) microsatellites. Our results are in accordance with these expectations, as we observed a high transferability of the markers between pine species, although most were monomorphic. Out of 44 markers, including 30 EST-SSRs, eight cpSSRs, and six genomic SSRs from different pine species, eight EST-SSRs from the sister species *P. bungeana* amplified well and showed polymorphism; hence, they were selected for further analysis.

Transferability of markers reflected the relatedness among the taxa. EST-SSRs from *P. bungeana* were transferrable (95%) and polymorphic (47%), reflecting the close systematic relationship between Chilgoza pine and *P. bungeana*, belonging to section *Quinquefoliae* and subsection *Gerardianae*. *Pinus thunbergii* and *P. halepensis* belong to the section *Pinus* but different subsections (*Pinus* and *Pinaster*, respectively). Nevertheless, almost all the markers derived from them amplified but were monomorphic in *P. gerardiana*. Since the chloroplast genome is more conserved than the nuclear genome (Ni et al. 2018), cpSSRs have even higher transferability rates in related species and even between distantly related species

Table 5. Characterization of the fine-scale spatial genetic structure using eight microsatellite markers in *Pinus gerardiana* in Gardiz, Afghanistan for young and old trees. F_{ij} , multilocus kinship coefficient between individuals of the first distance class (Loiselle et al. 1995); b_f , regression slope of F_{ij} on natural log distance; S_p , quantification of the fine-scale SGS; eig.sPCA, eigenvalue of the first sPCA axis.

Cohort	F_{ij}	b_f	S_p	p value (b_f)	eig.sPCA	p value (G_{-test})
Gardiz (young)	0.0847	-0.0092	0.0100	0.0003	0.0905	0.0010
Gardiz (old)	0.0619	-0.0027	0.0029	0.2011	0.0632	0.0300

Table 6. Genetic bottleneck tests in *Pinus gerardiana* cohorts in Gardiz under the Two-Phase (TPM) and the Infinite-Allele Model (IAM) using the Wilcoxon signed-rank test, as well as the M ratio approach. The T_2 statistic (combined Z-score in INEst) and M ratio (observed MR and MR_{eq} in an equilibrium population averaged over loci) are reported, and the p values are based on 10^6 permutations.

Cohort	T_2 under TPM		T_2 under IAM		M -ratio		
	T_2	p value	T_2	p value	MR	MR_{eq}	p value
Gardiz (young)	-3.5375	0.9844	-0.9428	0.9448	0.6017	0.8875	0.0079
Gardiz (old)	-3.5998	0.8906	-0.5240	0.7115	0.7065	0.8942	0.0253

(Ginwal et al. 2011). We observed successful amplification of monomorphic bands from cpSSRs derived from *P. thunbergii*. Our findings are consistent with other transferability rates reported in *Pinus* spp. where transfer rates increased with decreasing evolutionary distance and vice versa (Liewlaksaneeyanawin et al. 2004). Chagné et al. (2004) observed high transfer rates between two American pines (94.2% between *Pinus taeda* and *Pinus radiata* D.Don) and a lower rate for more distantly related species (64.6% between *Pinus taeda* and *Pinus canariensis* C.Sm. ex DC.). We successfully transferred eight polymorphic EST-SSR markers from *P. bungeana* to *P. gerardiana*, which will be valuable and essential tools for population and conservation genetic studies of this near-threatened pine species.

Genetic variation and structure

Higher heterozygosity in the *P. gerardiana* population in Gardiz compared to *P. bungeana* using partly the same

EST-SSRs was detected. Duan et al. (2017) revealed $H_e = 0.163$ and $H_o = 0.179$ estimates in *P. bungeana* compared to $H_e = 0.344$ and $H_o = 0.338$ in our study. Genetic diversity in other pine species was moderate and similar to our results based on EST-SSRs, e.g. for *P. halepensis* ($H_e = 0.380$) (Leonarduzzi et al. 2016) and *Pinus koraiensis* Siebold & Zucc. ($H_e = 0.311$) (Li et al. 2020). However, heterozygosity was slightly higher for *Pinus massoniana* Lamb. and *Pinus hwangshanensis* W.Y.Hsia ($H_e = 0.571$ and 0.488 , respectively) (Zhang et al. 2014), *Pinus sylvestris* L. ($H_e = 0.402$) (Fang et al. 2014), and *Pinus dabeshanensis* W.C.Cheng & Y.W.Law ($H_e = 0.458$) (Xiang et al. 2015). These findings indicate that the genetic diversity in Chilgoza pine in Gardiz can be categorized as moderate among pine species. Compared to *P. bungeana* (Duan et al. 2017), based on 64 samples from six populations, the mean number of alleles of selected markers was higher in Chilgoza pine in a single population (the mean number of alleles N_A for six common markers in *P. bungeana* is 1.630 in contrast to 4.430 in Chilgoza pine in Gardiz).

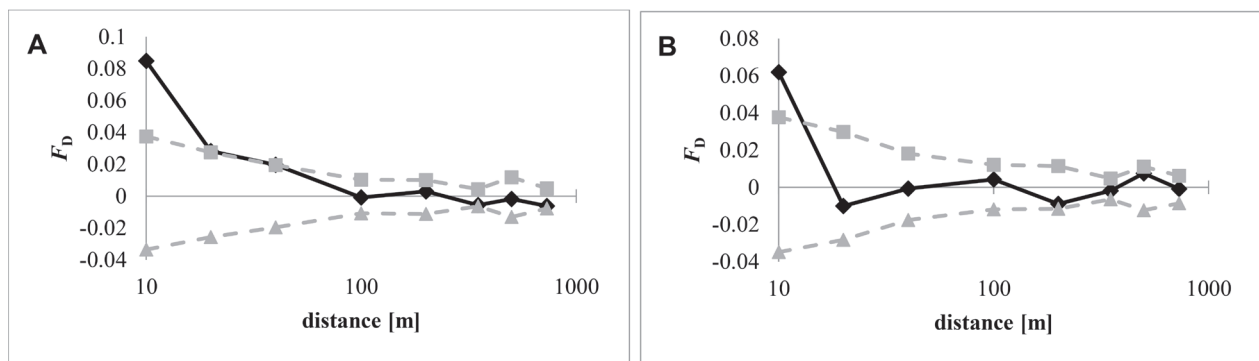


Figure 3. Spatial genetic structure in young (A) and old (B) *Pinus gerardiana* cohorts using EST-SSRs. The kinship coefficient per distance class, F_{ij} , was plotted against the logarithm of geographical distances between individuals.

The effective number of alleles N_{ae} in Chilgoza for young and old trees (1.595 and 1.620, respectively) was slightly higher than in *P. bungeana* (1.271). However, expected heterozygosity was slightly lower in young trees than in old trees in Gardiz, but no significant inbreeding was observed in the young and old trees. However, the lack of natural regeneration is likely to cause a loss of genetic diversity in the next generations.

PCoA and STRUCTURE analysis across the individuals revealed no significant structure or clustering among subpopulations for young or old tree cohorts. This is in line with our expectations as no population structure at this small spatial scale was expected for a wind-pollinated and bird-dispersed tree species. The large spotted nutcracker most likely disperses the seeds of Chilgoza pine, and large dispersal distances are expected. We found a moderate fine-scale SGS in the young trees ($Sp = 0.0100$) but no significant family structure in the older trees. In agreement with this, the sPCA indicated a significant global structure in both cohorts but with a slightly stronger family structure in young trees than in old trees. Usually, a weak, or non-significant fine-scale SGS is common in *Pinus* species (González-Martínez et al. 2002) and the fine-scale SGS found in the young tree cohort is similar to patterns found in other pine species. As a comparison, the Sp -statistic varies widely among pine species and populations: it was higher in two fragmented populations ($Sp = 0.0196$ and 0.0264) than in two continuous populations ($Sp = 0.0104$ and 0.0065) of maritime pine (*Pinus pinaster* Aiton) on the Iberian Peninsula (De-Lucas et al. 2009). Şofletea et al. (2020) reported Sp values between 0.0011 and 0.0207 in Scots pine (*P. sylvestris*) populations of the Romanian Carpathian Mountains, while González-Díaz et al. (2017) found Sp values ranging from -0.0006 to 0.0119, with higher Sp values in juvenile trees than in old trees in two of three Scots pine populations in Scotland and Siberia. Budde et al. (2017) also reported a wide range of Sp values in *P. halepensis* from the Iberian Peninsula, ranging from -0.0070 to 0.0169 under two distinct fire regimes characterized by differences in fire recurrence. In *P. cembra*, a bird dispersed pine species, very low Sp values indicated no significant SGS in four populations in the Alpes, and only one population showed a significant SGS with an Sp value of 0.0088 (Mosca et al. 2018). In a large-scale review, Gelmi-Candusso et al. (2017) compared the fine-scale spatial genetic structure of 68 zoochorous plant species with otherwise different life-history traits and found Sp values ranging from 0.0030 to 0.0240 for wind-pollinated and bird-dispersed (synzoochorous) tree species which is well in line with our results.

When comparing different age cohorts, previous studies have found no congruent difference in the fine-scale SGS. Kitamura et al. (2018) found stronger SGS in younger age classes of wind dispersed *Picea jezoensis* Carrière than in mature trees. However, Berens et al. (2014) observed the opposite trend in bird dispersed *Prunus africana* (Hook.f.) Kalkman, and e.g. Zhou and

Chen (2010) did not find any difference between age cohorts in bird dispersed fig species. Here, we observed stronger SGS in the younger cohort than in the older cohort, which might reflect less efficient gene flow during the last decades. Overexploitation might have directly impacted the seed dispersal (Vander Wall 2003; Kumar et al. 2013). Pruning during cone harvesting can lead to fewer strobili and a severe reduction in seeds left to disperse. If fewer individuals contributed to the next generation due to overexploitation, this could have increased SGS as seed shadows are expected to show less overlap than in stands with a higher density of reproducing trees. SGS is usually stronger in populations with lower density (Vekemans and Hardy 2004). While we do not have direct population density estimates, young (101 individuals) and old trees (90 individuals) were sampled randomly in the same area and distance classes in the SGS analyses were represented by very similar numbers of pairwise comparisons in both cohorts.

Overexploitation may also have significantly reduced the nuts available to the large spotted nutcracker. This reduction in food may have reduced the population of the birds or forced them to migrate or even go extinct in the region. Over the past decades, a decline of this main seed dispersal agent could also have led to an increase in SGS. Data about the abundance, ecology, and behaviour of the large spotted nutcracker in this region would be very valuable. Additionally, stronger SGS could increase homozygosity (Rico and Wagner 2016) due to mating of related individuals, which might be reflected in a slight decrease in heterozygosity from the old to the young cohort.

Bottleneck signals and effective population size

Climate oscillations have changed the distribution ranges and population sizes of species in the past, including forest trees. Many temperate tree species shifted towards the lower latitudes, resulting in a sometimes severe range reduction (Wells 1983; Łabiszak et al. 2021). Pines experienced an ancient bottleneck with the beginning of the Last Glacial Maximum (LGM) (Oline et al. 2000; Al-Rabab'ah and Williams 2004; Naydenov et al. 2011; Łabiszak et al. 2021). In Gardiz, we found, as expected, that both cohorts showed very similar signs of past population size change. The Wilcoxon test did not indicate a recent decline in population size. However, the low M ratio value indicated a more ancient bottleneck, likely followed by a population expansion. The T_2 statistic is most suitable for detecting recent and weak bottlenecks (Piry et al. 1999), while the M ratio is more appropriate for detecting ancient and more severe bottleneck events that lasted several generations (Williamson-Natesan 2005). The results of the Wilcoxon signed-rank test for heterozygosity excess showed a heterozygosity deficiency for five out of seven markers which could be interpreted as a recent population expansion only a few generations

ago or a recent introduction of rare alleles by gene flow (Luikart and Cornuet 1998; Piry et al. 1999).

While we did not find signs of a recent bottleneck, we observed fewer trees in smaller size classes, indicating that natural regeneration is not as abundant as some decades ago. Typically, the smaller size classes should be the most abundant ones (Ahmed et al. 1991; Khan et al. 2015). However, during the last decade, a decreasing population trend has been reported for range-wide populations of *P. gerardiana* (Farjon 2013).

Estimating effective population size, N_e , and monitoring its changes over time is crucial in assessing the risk of extinction for endangered species (Ellstrand and Elam 1993). In Europe, the use of contemporary N_e estimates has been proposed to develop genetic monitoring programs of forest tree species (Aravanopoulos et al. 2015). When N_e is low or declining in a population, the causes of the decline must be identified, and measures should be taken to reverse the trend (Wang et al. 2016). In Chilgoza pine in Gardiz, we found contemporary N_e values of ca 101 and 401 and the confidence intervals of both cohorts were very wide and overlapped. Therefore, regarding the wide intervals, the Chilgoza population maintained similar levels of effective population size during the last decades. However, contemporary N_e mostly reflects the effective population size in the parental generations (Waples and Do 2010) but does not reflect the more recent decline in natural regeneration. To compare N_e between studies and species, the generation time, mutation rates, markers and methods should be the same. Rajora and Zinck (2021), using microsatellites, found a range of 136 to 655 (mean = 340, 95% CI range = 99–infinite) for $N_{e_{LD_{0.01}}}$ and from 69 to 222 (mean = 145, 95% CI range = 51–1234) for $N_{e_{LD_{0.03}}}$ over eight *Pinus strobus* L. populations, which is in the range of our result. Also, in *P. cembra*, based on linkage disequilibrium estimates, using nSSRs, the effective population size in the Tatra Mountains ranged from 26.4 to 4354.6 (mean = 509.47, 95% CI range = 16.30–infinite), in the Carpathians it ranged from 12.50 to 11325.80 (mean = 1618.16, 95% CI range = 6.50–infinite) and in the Alps it ranged from 24.80 to 991.10 (mean = 396.06, 95% CI range = 13.30–infinite) (Dzialuk et al. 2014).

The total forest cover of Afghanistan was strongly reduced compared to its original state as a result of longtime instability during the last few decades (Groninger 2006; Shalizi et al. 2018). This trend continued, even after 2001, with the persistence of instability affecting the population of Chilgoza pines in Paktia. While forests covered a large part of the country until ca 2000 years ago, they are now mostly limited to a small part of eastern Afghanistan. With the very fragile ecosystems of Afghanistan, when the forest is lost, it may never recover (Saba 2001). Therefore, conservation measures of the government and local people are of vital importance. During the past few years, in provinces with Chilgoza forests (Paktia, Paktika, and Khost), some conservation practices, including direct seeding and home nursery projects, were established by the Ministry of Agriculture, Irrigation and Livestock

(MAIL). With the implementation of these projects, the MAIL expects ca 20 Kha of Chilgoza forests to be restored in the next ten years (MAIL 2019).

CONCLUSIONS

In this research, we collected 191 samples from Chilgoza pines in Paktia province, one of the most insecure locations in Afghanistan (Groninger 2012). The collection of samples from forests located far away from the cities, the safe places, was impossible; therefore, sampling was a risky, dangerous and challenging part of this research. As we could not go to central parts of the forest, we collected the samples from locations near the main road.

Younger trees are mostly missing which could possibly result in a strong population size decline in the long term. However, despite signatures of an ancient bottleneck, the Chilgoza stands currently contain moderate genetic variation and an effective population size similar to that of other pine species. Furthermore, the environmental conditions, such as well-drained and sandy soils in eastern Afghanistan, are favourable for Chilgoza pines, making them superior and dominant compared to other tree species (Shalizi et al. 2016). However, none of the accepted forest management methods, e.g. rotational cone harvesting, individual tree protection or partial cone harvesting, has been implemented in Afghanistan (Shalizi et al. 2016). Local tribes control the Chilgoza natural forests and treat them as private nut-producing orchards for commercial purposes and not as natural ecosystems (Groninger and Ruffner 2010; Shalizi et al. 2016). Therefore, forest conservation projects in coordination with the local people need to be implemented and should also focus on the protection and more sustainable use of natural stands.

The sustainable management of forests requires a better understanding of the specific features of forest trees and their genetic diversity. Therefore, provenance trials should be established in Paktia, Paktika, and Khost provinces which produce 86% of Afghanistan's Chilgoza pine nuts (CARD-F 2017). To improve and maintain the genetic variability of Chilgoza, as the only species with edible pine nuts in Afghanistan, additional reforestation and population conservation programs should be implemented. Our results show that the sampling area is suitable for conservation-targeted projects, appears genetically diverse, but comparative studies with other populations throughout the species distribution range are unfortunately missing. More samples from other areas are needed to better understand the distribution of genetic diversity in *P. gerardiana* and could be analysed with the same eight EST-SSRs. In conclusion, ex situ conservation measures and the protection and sustainable management of Chilgoza forests in situ in the area are necessary to maintain genetic diversity by fostering natural regeneration.

DATA AVAILABILITY STATEMENT

EST-SSR data, GPS coordinates, and diameter at breast height measures were deposited in the Göttingen Research Online repository (<https://doi.org/10.25625/MGKPZA>).

ACKNOWLEDGEMENTS

We thank the German Academic Exchange Service (DAAD) for its financial support of Sayed Jalal Moosavi. We would like to thank Christine Radler and Alexandra Dolynska for their technical support in the laboratory, Dr Sayed Isamel Moosavi, Zahra Moosavi, and Somayah Moosavi for their support and help with sampling in Gardiz.

REFERENCES

- Ahmed M, Ashfaq M, Amjad M, Saeed M (1991) Vegetation structure and dynamics of *Pinus gerardiana* forests in Balouchistan, Pakistan. *Journal of Vegetation Science* 2(1): 119–124. <https://doi.org/10.2307/3235904>
- Ahmed M, Sarangezai AM (1991) Dendrochronological approach to estimate age and growth rate of various species from Himalayan region of Pakistan. *Pakistan Journal of Botany* 23(1): 78–89.
- Al-Rabab'ah MA, Williams CG (2004) An ancient bottleneck in the Lost Pines of central Texas. *Molecular Ecology* 13(5): 1075–1084. <https://doi.org/10.1111/j.1365-294X.2004.02142.x>
- Alami K, Mousavi SY (2020) Afghan Chehelghoza (*Pinus gerardiana* L.) pine nut diet enhances the learning and memory in male rats. *Nutrition and Dietary Supplements* 12: 277–288. <https://doi.org/10.2147/NDS.S278350>
- Aravanopoulos FA, Tollefsrud MM, Graudal L, Koskela J, Kätzel R, Soto A, Nagy L, Pilipovic A, Zhelev P, Božic G (2015) Development of genetic monitoring methods for genetic conservation units of forest trees in Europe. European forest genetic resources programme (EUFORGEN) and Bioversity International, Rome.
- Auckland L, Bui T, Zhou Y, Shepherd M, Williams C (2002) Conifer microsatellite handbook. Texas A&M University, College Station, Texas, 1–57.
- Bekku YS, Kurokouchi H, Matsuki Y, Nishi N, Lian C (2019) Genetic structure of *Pinus parviflora* on Mt. Fuji in relation to the hoarding behavior of the Japanese nutcracker. *Ecosphere* 10(4): e02694. <https://doi.org/10.1002/ecs2.2694>
- Berens DG, Braun C, González-Martínez SC, Griebeler EM, Nathan R, Böhning-Gaese K (2014) Fine-scale spatial genetic dynamics over the life cycle of the tropical tree *Prunus africana*. *Heredity* 113(5): 401–407. <https://doi.org/10.1038/hdy.2014.40>
- Bonan GB (2008) Forests and climate change: forcings, feedbacks, and the climate benefits of forests. *Science* 320(5882): 1444–1449. <https://doi.org/10.1126/science.1155121>
- Bonner FT, Karrfalt RP (2008) The woody plant seed manual. US Department of Agriculture, Forest Service, Washington, DC, 1–1223.
- Brownstein MJ, Carpten JD, Smith JR (1996) Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. *Biotechniques* 20(6): 1004–1010. <https://doi.org/10.2144/96206st01>
- Budde KB, González-Martínez SC, Navascués M, Burgarella C, Mosca E, Lorenzo Z, Zabal-Aguirre M, Vendramin GG, Verdú M, Pausas JG (2017) Increased fire frequency promotes stronger spatial genetic structure and natural selection at regional and local scales in *Pinus halepensis* Mill. *Annals of Botany* 119(6): 1061–1072. <https://doi.org/10.1093/aob/mcw286>
- Cai L, Cao A, Luo Z, Mao L, Ying T (2017) Ultrastructure characteristics and quality changes of low-moisture Chilgoza pine nut (*Pinus gerardiana*) during the near-freezing-temperature storage. *CyTA-Journal of Food* 15(3): 466–473. <https://doi.org/10.1080/19476337.2017.1297962>
- Cai L, Xiao L, Liu C, Ying T (2013) Functional properties and bioactivities of pine nut (*Pinus gerardiana*) protein isolates and its enzymatic hydrolysates. *Food and Bioprocess Technology* 6(8): 2109–2117. <https://doi.org/10.1007/s11947-012-0885-7>
- CARD-F (2017) Pine nut's value chain (an assessment report) focus on Loya Paktia region (Paktia, Paktika and Khost provinces), Kabul.
- Chagné D, Chaumeil P, Ramboer A, Collada C, Guevara A, Cervera MT, Vendramin GG, Garcia V, Frigerio J-M, Echt C (2004) Cross-species transferability and mapping of genomic and cDNA SSRs in pines. *Theoretical and Applied Genetics* 109(6): 1204–1214. <https://doi.org/10.1007/s00122-004-1683-z>
- Chybicki IJ (2017) INEST 2.2. http://www.ukw.edu.pl/pracownicy/strona/igor_chybicki/software_ukw/ [accessed 20.10.2021]
- Chybicki IJ, Burczyk J (2009) Simultaneous estimation of null alleles and inbreeding coefficients. *Journal of Heredity* 100(1): 106–113. <https://doi.org/10.1093/jhered/esn088>
- Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* 144(4): 2001–2014. <https://doi.org/10.1093/genetics/144.4.2001>
- Critchfield WB, Little EL (1966) Geographic distribution of the pines of the world. US Department of Agriculture, Forest Service, Washington, DC, 1–97. <https://doi.org/10.5962/bhl.title.66393>
- Cronn R, Liston A, Parks M, Gernandt DS, Shen R, Mockler T (2008) Multiplex sequencing of plant chloroplast genomes using Solexa sequencing-by-synthesis technology. *Nucleic Acids Research* 36(19): e122. <https://doi.org/10.1093/nar/gkn502>
- De-Lucas AI, González-Martínez SC, Vendramin GG, Hidalgo E, Heuertz M (2009) Spatial genetic structure in continuous and fragmented populations of *Pinus pinaster* Aiton. *Molecular Ecology* 18(22): 4564–4576. <https://doi.org/10.1111/j.1365-294X.2009.04372.x>

- Do C, Waples RS, Peel D, Macbeth GM, Tillett BJ, Ovenden JR (2014) NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Molecular Ecology Resources* 14(1): 209–214. <https://doi.org/10.1111/1755-0998.12157>
- Duan D, Jia Y, Yang J, Li Z-H (2017) Comparative transcriptome analysis of male and female conelets and development of microsatellite markers in *Pinus bungeana*, an endemic conifer in China. *Genes* 8(12): 393. <https://doi.org/10.3390/genes8120393>
- Dzialuk A, Chybicki I, Gout R, Mączka T, Fleischer P, Konrad H, Curtu AL, Sofletea N, Valadon A (2014) No reduction in genetic diversity of Swiss stone pine (*Pinus cembra* L.) in Tatra Mountains despite high fragmentation and small population size. *Conservation Genetics* 15(6): 1433–1445. <https://doi.org/10.1007/s10592-014-0628-6>
- Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4: 359–361. <https://doi.org/10.1007/s12686-011-9548-7>
- Ellis J, Burke J (2007) EST-SSRs as a resource for population genetic analyses. *Heredity* 99(2): 125–132. <https://doi.org/10.1038/sj.hdy.6801001>
- Ellstrand NC, Elam DR (1993) Population genetic consequences of small population size: implications for plant conservation. *Annual Review of Ecology and Systematics* 24(1): 217–242. <https://doi.org/10.1146/annurev.es.24.110193.001245>
- Elsik CG, Minihan VT, Hall SE, Scarpa AM, Williams CG (2000) Low-copy microsatellite markers for *Pinus taeda* L. *Genome* 43(3): 550–555. <https://doi.org/10.1139/g00-002>
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14(8): 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Fang P, Niu S, Yuan H, Li Z, Zhang Y, Yuan L, Li W (2014) Development and characterization of 25 EST-SSR markers in *Pinus sylvestris* var. *mongolica* (Pinaceae). *Applications in Plant Sciences* 2(1): 1300057. <https://doi.org/10.3732/apps.1300057>
- Farjon A (2013) *Pinus gerardiana*. The IUCN Red List of Threatened Species e.T34189A2850009. <https://doi.org/10.2305/IUCN.UK.2013-1.RLTS.T34189A2850009.en>
- Garza JC, Williamson EG (2001) Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology* 10(2): 305–318. <https://doi.org/10.1046/j.1365-294x.2001.01190.x>
- Gelmi-Candusso TA, Heymann EW, Heer K (2017) Effects of zoochory on the spatial genetic structure of plant populations. *Molecular Ecology* 26(21): 5896–5910. <https://doi.org/10.1111/mec.14351>
- GINWAL HS, CHAUHAN P, BARTHWAAL S, SHARMA A, SHARMA R (2011) Cross-species amplification and characterization of *Pinus* chloroplast microsatellite markers in *Cedrus deodara* Roxb. *Silvae Genetica* 60(1-6): 65–69. <https://doi.org/10.1515/sg-2011-0009>
- Giraudeau P (2021) pgirmess: spatial analysis and data mining for field ecologists. R package version 1.7.0. <https://CRAN.R-project.org/package=pgirmess> [accessed 25.02.2022]
- González-Díaz P, Jump AS, Perry A, Wachowiak W, Lapshina E, Cavers S (2017) Ecology and management history drive spatial genetic structure in Scots pine. *Forest Ecology and Management* 400: 68–76. <https://doi.org/10.1016/j.foreco.2017.05.035>
- González-Martínez S, Gerber S, Cervera M, Martínez-Zapater J, Gil L, Alía R (2002) Seed gene flow and fine-scale structure in a Mediterranean pine (*Pinus pinaster* Ait.) using nuclear microsatellite markers. *Theoretical and Applied Genetics* 104(8): 1290–1297. <https://doi.org/10.1007/s00122-002-0894-4>
- Groninger JW (2006) Forestry and forestry education in Afghanistan. *Journal of Forestry* 104(8): 426–430. <https://doi.org/10.1093/jof/104.8.426>
- Groninger JW (2012) Reforestation strategies amid social instability: lessons from Afghanistan. *Environmental Management* 49(4): 833–845. <https://doi.org/10.1007/s00267-012-9817-6>
- Groninger JW, Ruffner CM (2010) Hearts, minds, and trees: forestry's role in operation enduring freedom. *Journal of Forestry* 108(3): 141–147.
- Hardy OJ, Vekemans X (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes* 2(4): 618–620. <https://doi.org/10.1046/j.1471-8286.2002.00305.x>
- Hill WG (1981) Estimation of effective population size from data on linkage disequilibrium. *Genetics Research* 38(3): 209–216. <https://doi.org/10.1017/S0016672300020553>
- Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24(11): 1403–1405. <https://doi.org/10.1093/bioinformatics/btn129>
- Jombart T, Devillard S, Dufour A-B, Pontier D (2008) Revealing cryptic spatial patterns in genetic variability by a new multivariate method. *Heredity* 101(1): 92–103. <https://doi.org/10.1038/hdy.2008.34>
- Kant A, Dutt V, Sharma DR (2006a) Genetic variability in phenotypic characters of *Pinus gerardiana*. *Indian Forester* 132(6): 681–690. <https://doi.org/10.36808/if/2006/v132i6/4218>
- Kant A, Pattanayak D, Chakrabarti SK, Sharma R, Thakur M, Sharma DR (2006b) RAPD analysis of genetic variability in *Pinus gerardiana* Wall. in Kinnaur (Himachal Pradesh). *Indian Journal of Biotechnology* 5: 62–67.
- Khan H, Akbar M, Zaman M, Hyder S, Khan M, Nafees M, Raza G, Begum F, Hussain S, Khan S (2015) Diameter size class distributions of *Pinus gerardiana* Wall. Ex D. Don from Gohar Abad Valley district Diaper, Gilgit-Baltistan. *Journal of Biodiversity and Environmental Sciences* 6: 50–56.
- Khurram S, Shalizi MN (2016) Traditional and alternative techniques of Chilgoza Pine (*Pinus gerardiana* Wall.ex D.Don) nut harvesting and processing in Afghanistan. *International Journal of Bioassays* 5(11): 5011–5015. <https://doi.org/10.21746/ijbio.2016.11.004>
- Kitamura K, Nakanishi A, Lian C, Goto S (2018) Distinctions in fine-scale spatial genetic structure between growth stages of

- Picea jezoensis* Carr. *Frontiers in Genetics* 9 : 490. <https://doi.org/10.3389/fgene.2018.00490>
- Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I (2015) Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources* 15(5): 1179–1191. <https://doi.org/10.1111/1755-0998.12387>
- Kubisiak TL, Nelson CD, Staton ME, Zhebentyayeva T, Smith C, Olukolu BA, Fang G-C, Hebard FV, Anagnostakis S, Wheeler N (2013) A transcriptome-based genetic map of Chinese chestnut (*Castanea mollissima*) and identification of regions of segmental homology with peach (*Prunus persica*). *Tree Genetics & Genomes* 9(2): 557–571. <https://doi.org/10.1007/s11295-012-0579-3>
- Kumar R, Shamet G, Chaturvedi O, Avasthe R, Singh C (2013) Ecology of chilgoza pine (*Pinus gerardiana* Wall) in dry temperate forests of North West Himalaya. *Ecology, Environment and Conservation* 19(4): 1063–1066.
- Kumar R, Shamet GS, Pandey A, Kakade V, Dinesh D, Sharma N, Gupta D (2016) Impact of anthropogenic disturbances on ecology of *Pinus gerardiana* Wall in Indian Himalaya: a review. *Agricultural Research & Technology: Open Access Journal* 1(3): 555561. <https://doi.org/10.19080/ARTOAJ.2016.01.555561>
- Łabiszak B, Zaborowska J, Wójkiewicz B, Wachowiak W (2021) Molecular and paleo-climatic data uncover the impact of an ancient bottleneck on the demographic history and contemporary genetic structure of endangered *Pinus uliginosa*. *Journal of Systematics and Evolution* 59(3): 596–610. <https://doi.org/10.1111/jse.12573>
- Lanner RM (1990) Biology, taxonomy, evolution, and geography of stone pines of the world. General Technical Report INT (USA) 270: 14–24.
- Leonarduzzi C, Spanu I, Labriola M, González-Martínez SC, Piotti A, Vendramin GG (2016) Development and characterization of three highly informative EST-SSR multiplexes for *Pinus halepensis* Mill. and their transferability to other Mediterranean pines. *Plant Molecular Biology Reporter* 34(5): 993–1002. <https://doi.org/10.1007/s11105-016-0980-4>
- Li X, Liu X, Wei J, Li Y, Tigabu M, Zhao X (2020) Development and transferability of EST-SSR markers for *Pinus koraiensis* from cold-stressed transcriptome through Illumina sequencing. *Genes* 11(5): 500. <https://doi.org/10.3390/genes11050500>
- Liewlaksaneeyanawin C, Ritland CE, El-Kassaby YA, Ritland K (2004) Single-copy, species-transferable microsatellite markers developed from loblolly pine ESTs. *Theoretical and Applied Genetics* 109(2): 361–369. <https://doi.org/10.1007/s00122-004-1635-7>
- Loiselle BA, Sork VL, Nason J, Graham C (1995) Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *American Journal of Botany* 82(11): 1420–1425. <https://doi.org/10.2307/2445869>
- Luikart G, Cornuet J-M (1998) Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. *Conservation Biology* 12(1): 228–237. <https://doi.org/10.1111/j.1523-1739.1998.96388.x>
- Madge S, Spencer AJ, Kirwan GM (2020) Kashmir nutcracker (*Nucifraga multipunctata*). In: Kirwan GM, Spencer AJ, Rodewald PG, Keeney BK, Billerman SM (Eds) *Birds of the World*. Cornell Lab of Ornithology, Ithaca, NY. <https://doi.org/10.2173/bow.sponut1.01.1>
- MAIL (2019) Forest department of Ministry of Agriculture, Irrigation and Livestock. <https://www.mail.gov.af/en> [accessed 25.11.2021]
- Malik AR, Shamet GS, Butola JS, Bhat GM, Mir AA, Nabi G (2013) Standardization of seed storage conditions in chilgoza pine (*Pinus gerardiana* Wall.): an endangered pine of Hind Kush Himalaya. *Trees* 27(5): 1497–1501. <https://doi.org/10.1007/s00468-013-0868-y>
- Martin LJ, Blossey B, Ellis E (2012) Mapping where ecologists work: biases in the global distribution of terrestrial ecological observations. *Frontiers in Ecology and the Environment* 10(4): 195–201. <https://doi.org/10.1890/110154>
- Masumy SA (2006) *Dendrology, a study book for Afghanistan*. Faculty of Agriculture, Agha Khan Foundation and USDA, Kabul.
- Milad M, Schaich H, Bürgi M, Konold W (2011) Climate change and nature conservation in Central European forests: a review of consequences, concepts and challenges. *Forest Ecology and Management* 261(4): 829–843. <https://doi.org/10.1016/j.foreco.2010.10.038>
- Mosca E, Di Pierro EA, Budde KB, Neale DB, González-Martínez SC (2018) Environmental effects on fine-scale spatial genetic structure in four Alpine keystone forest tree species. *Molecular Ecology* 27(3): 647–658. <https://doi.org/10.1111/mec.14469>
- Naydenov KD, Naydenov MK, Tremblay F, Alexandrov A, Aubin-Fournier LD (2011) Patterns of genetic diversity that result from bottlenecks in Scots Pine and the implications for local genetic conservation and management practices in Bulgaria. *New Forests* 42(2): 179–193. <https://doi.org/10.1007/s11056-010-9245-5>
- Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89(3): 583–590. <https://doi.org/10.1093/genetics/89.3.583>
- Neuschulz EL, Merges D, Bollmann K, Gugerli F, Böhning-Gaese K (2018) Biotic interactions and seed deposition rather than abiotic factors determine recruitment at elevational range limits of an alpine tree. *Journal of Ecology* 106(3): 948–959. <https://doi.org/10.1111/1365-2745.12818>
- Ni Z, Zhou P, Xu M, Xu L-A (2018) Development and characterization of chloroplast microsatellite markers for *Pinus massoniana* and their application in *Pinus* (Pinaceae) species. *Journal of Genetics* 97(1): 53–59. <https://doi.org/10.1007/s12041-018-0931-y>
- Oline DK, Mitton JB, Grant MC (2000) Population and subspecific genetic differentiation in the foxtail pine (*Pinus balfouriana*). *Evolution* 54(5): 1813–1819. <https://doi.org/10.1111/j.0014-3820.2000.tb00725.x>
- Pautasso M (2009) Geographical genetics and the conservation of forest trees. *Perspectives in Plant Ecology, Evolution and Systematics* 11(3): 157–189. <https://doi.org/10.1016/j.ppees.2009.01.003>

- Peakall R, Smouse PE (2012) GenAEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28(19): 2537–2539. <https://doi.org/10.1093/bioinformatics/bts460>
- Petit RJ, Hampe A (2006) Some evolutionary consequences of being a tree. *Annual Review of Ecology, Evolution, and Systematics* 37: 187–214. <https://doi.org/10.1146/annurev.ecolsys.37.091305.110215>
- Piry S, Luikart G, Cornuet J-M (1999) Computer note. BOTTLENECK: a computer program for detecting recent reductions in the effective size using allele frequency data. *Journal of Heredity* 90(4): 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(2): 945–959. <https://doi.org/10.1093/genetics/155.2.945>
- R Core Team (2021) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. <https://www.R-project.org/> [accessed 20.02.2022]
- Rajora OP, Zinck JWR (2021) Genetic diversity, structure and effective population size of old-growth vs. second-growth populations of keystone and long-lived conifer, eastern white pine (*Pinus strobus*): conservation value and climate adaptation potential. *Frontiers in Genetics* 12. <https://doi.org/10.3389/fgene.2021.650299>
- Ranot M, Shrama R (2015) Genetic variations studies on the different morphological characters of chilgoza pine (*Pinus gerardiana* Wall.). *International Journal of Science and Research* 4(1): 191–193. <https://doi.org/10.36106/ijsr>
- Rico Y, Wagner H (2016) Reduced fine-scale spatial genetic structure in grazed populations of *Dianthus carthusianorum*. *Heredity* 117: 367–374. <https://doi.org/10.1038/hdy.2016.45>
- Saba DS (2001) Afghanistan: Environmental degradation in a fragile ecological setting. *The International Journal of Sustainable Development & World Ecology* 8(4): 279–289. <https://doi.org/10.1080/13504500109470086>
- Saeed M, Thanos C (2006) The effect of seed coat removal on seed germination of *Pinus gerardiana* Wallich ex D. Don. Chilgoza pine. *Journal of Applied and Emerging Sciences* 1(2): 174–177.
- Safari S, Yousefi M, Khawari A, Sajjadi M, Nazari ML, Alipour A, Salehi MH, Mousavi Y (2017) Effect of Afghan chehelgoza (*Pinus gerardiana* L.) on sperm parameters of male rats. *Journal of Food and Nutritional Disorders* 6: 5. <https://doi.org/10.4172/2324-9323.1000237>
- Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology* 18(2): 233–234. <https://doi.org/10.1038/72708>
- Shalizi M, Khurram S, Mayer E, Hurmat A (2016) Afghanistan chilgoza pine forests: current status, anthropogenic pressure, trends in regeneration and management. Assistance in Building Afghanistan by Developing Enterprises (ABADE) Program and United States Agency for International Development (USAID), Kabul.
- Shalizi MN, Khurram S, Groninger JW, Ruffner CM, Burney OT (2018) Indigenous knowledge and stand characteristics of a threatened tree species in a highly insecure area: Chilgoza pine in Afghanistan. *Forest Ecology and Management* 413: 1–8. <https://doi.org/10.1016/j.foreco.2018.01.053>
- Siddiqui MF, Ahmed M, Wahab M, Khan N, Khan MU, Nazim K, Hussain SS (2009) Phytosociology of *Pinus roxburghii* Sargent (chir pine) in lesser Himalayan and Hindu Kush range of Pakistan. *Pakistan Journal of Botany* 41(5): 2357–2369.
- Singh G, Kumar D, Dash AK (2021) *Pinus gerardiana* Wallich ex. D. Don. – A review. *Phytomedicine Plus* 1(2): 100024. <https://doi.org/10.1016/j.phyplu.2021.100024>
- Singh RV, Khanduri DC, Lal K (1973) Chilgoza pine (*Pinus gerardiana*) regeneration in Himachal Pradesh. *Indian Forester* 99(3): 126–133.
- Șofletea N, Mihai G, Ciocîrlan E, Curtu AL (2020) Genetic diversity and spatial genetic structure in isolated Scots Pine (*Pinus sylvestris* L.) populations native to eastern and southern Carpathians. *Forests* 11(10): 1047. <https://doi.org/10.3390/f11101047>
- Song Y-G, Petitpierre B, Deng M, Wu J-P, Kozłowski G (2019) Predicting climate change impacts on the threatened *Quercus arbutifolia* in montane cloud forests in southern China and Vietnam: conservation implications. *Forest Ecology and Management* 444: 269–279. <https://doi.org/10.1016/j.foreco.2019.04.028>
- Thakur NS, Sharma S, Gupta R, Gupta A (2014) Studies on drying and storage of chilgoza (*Pinus gerardiana*) nuts. *Journal of Food Science and Technology* 51(9): 2092–2098. <https://doi.org/10.1007/s13197-012-0692-1>
- Thomas P (2019) *Pinus gerardiana*, from the website, Threatened Conifers of The World. <https://threatenedconifers.rbge.org.uk/conifers/pinus-gerardiana> [accessed 19.02.2022]
- Van Oosterhout C, Hutchinson WF, Wills DP, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4(3): 535–538. <https://doi.org/10.1111/j.1471-8286.2004.00684.x>
- Vander Wall SB (2003) Effects of seed size of wind-dispersed pines (*Pinus*) on secondary seed dispersal and the caching behavior of rodents. *Oikos* 100(1): 25–34. <https://doi.org/10.1034/j.1600-0706.2003.11973.x>
- Vekemans X, Hardy OJ (2004) New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology* 13(4): 921–935. <https://doi.org/10.1046/j.1365-294x.2004.02076.x>
- Vendramin GG, Lelli L, Rossi P, Morgante M (1996) A set of primers for the amplification of 20 chloroplast microsatellites in Pinaceae. *Molecular Ecology* 5(4): 595–598. <https://doi.org/10.1111/j.1365-294x.1996.tb00353.x>
- Wang J, Santiago E, Caballero A (2016) Prediction and estimation of effective population size. *Heredity* 117(4): 193–206. <https://doi.org/10.1038/hdy.2016.43>
- Waples RS (2006) A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conservation Genetics* 7(2): 167–184. <https://doi.org/10.1007/s10592-005-9100-y>
- Waples RS, Do C (2010) Linkage disequilibrium estimates of contemporary N_e using highly variable genetic markers:

- a largely untapped resource for applied conservation and evolution. *Evolutionary Applications* 3(3): 244–262. <https://doi.org/10.1111/j.1752-4571.2009.00104.x>
- Wells PV (1983) Paleobiogeography of montane islands in the Great Basin since the last glaciopluvial. *Ecological Monographs* 53(4): 341–382. <https://doi.org/10.1093/ecol/53.4.341>
- Williamson-Natesan EG (2005) Comparison of methods for detecting bottlenecks from microsatellite loci. *Conservation Genetics* 6(4): 551–562. <https://doi.org/10.1007/s10592-005-9009-5>
- Wingard J, Karlstetter M, Simms A, Johnson M (2010) Eastern forests program – Timber trade survey. Wildlife Conservation Society, Kabul.
- Xiang X-Y, Zhang Z-X, Duan R-Y, Zhang X-P, Wu G-L (2015) Genetic diversity and structure of *Pinus dabeshanensis* revealed by expressed sequence tag-simple sequence repeat (EST-SSR) markers. *Biochemical Systematics and Ecology* 61: 70–77. <https://doi.org/10.1016/j.bse.2015.06.001>
- Yang Y, Wang M, Liu Z, Zhu J, Yan M, Li Z (2016) Nucleotide polymorphism and phylogeographic history of an endangered conifer species *Pinus bungeana*. *Biochemical Systematics and Ecology* 64: 89–96. <https://doi.org/10.1016/j.bse.2015.11.016>
- Zhang D, Xia T, Yan M, Dai X, Xu J, Li S, Yin T (2014) Genetic introgression and species boundary of two geographically overlapping pine species revealed by molecular markers. *PLoS ONE* 9(6): e101106. <https://doi.org/10.1371/journal.pone.0101106>
- Zhou HP, Chen J (2010) Spatial genetic structure in an understorey dioecious fig species: the roles of seed rain, seed and pollen-mediated gene flow, and local selection. *Journal of Ecology* 98(5): 1168–1177. <https://doi.org/10.1111/j.1365-2745.2010.01683.x>

SUPPLEMENTARY FILE

Supplementary file 1

DBH histogram of all *P. gerardiana* samples (Fig. S1); list of EST-SSRs, genomic SSRs, and chloroplast SSRs (Table S1); transferability information of genetic markers from related pine species to *P. gerardiana* (Table S2); genetic variation over seven loci for each subpopulation and DBH class (Table S3); genetic variation over eight loci for each marker in the old and young cohorts (Tables S4, S5); scatterplot of 191 individuals heterozygosity plotted against the DBH (Fig. S2); PCoA based on genetic distances (Fig. S3); structure plots of all samples (Fig. S4); the plot of mean likelihood L(K) and variance per K value from STRUCTURE (Fig. S5); the plot of delta K (Fig. S6); sPCA plots for young (Fig. S7) and old cohort (Fig. S8).

Link: <https://doi.org/10.5091/plecevo.95754.suppl1>