



Genetic diversity and population structure of *Robinia pseudoacacia* from six improved variety bases in China as revealed by simple sequence repeat markers

Qi Guo^{1,2} · Sen Cao¹ · Li Dong¹ · Xiuyu Li¹ · Jiangtao Zhang³ · Yuanshuai Zhang⁴ · Zijie Zhang¹ · Yuhan Sun¹ · Cui Long¹ · Yingming Fan¹ · Chao Han¹ · Peiyao Han¹ · Xiaodi Liu¹ · Yun Li¹

Received: 3 July 2020 / Accepted: 28 September 2020
© The Author(s) 2021

Abstract *Robinia pseudoacacia* is an important afforestation tree introduced to China in 1878. In the present study, we examined the genetic diversity among 687 strains representing four improved varieties and two secondary provenances, comprising 641 clones and 46 seedlings. Ninety-one simple sequence repeats (SSRs) were selected through segregation analysis and polymorphism characterization, and all sampled individuals were genotyped using well-characterized SSR markers. After excluding loci with non-neutral equilibrium, missing locus data and null alleles, we used 36 primer pairs to assess the genetic diversity of these

germplasm resources, revealing vast genetic differentiation among the samples, with an average of 8.352 alleles per locus and a mean Shannon's index of 1.302. At the population level, the partitioning of variability was assessed using analysis of molecular variance, which revealed 93% and 7% variation within and among collection sites, respectively. Four clusters were detected using STRUCTURE analysis, indicating a degree of genetic differentiation among the six populations. Insights into the genetic diversity and structure of *R. pseudoacacia* provide a theoretical basis for the conservation, breeding and sustainable development in China.

Qi Guo and Sen Cao contributed equally to this work.

Project funding: This work was supported by the Doctoral Scientific Research Starting Foundation of Henan University of Science and Technology (13480089), National Nature Science Foundation of China (31570677), the National Key R&D Program of China (2017YFD0600503) and the National Forestry and Grassland Administration of Science and Technology Development Center Project (2016007).

The online version is available at <http://www.springerlink.com>

Corresponding editor: Yanbo Hu

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11676-021-01356-2>.

✉ Yun Li
yunli@bjfu.edu.cn

¹ Beijing Advanced Innovation Center for Tree Breeding By Molecular Design, Engineering Technology Research Center of Black Locust of National Forestry and Grassland Administration, College of Biological Sciences and Technology, National Engineering Laboratory for Tree Breeding, Beijing Forestry University, Beijing 100083, P.R. China

Keywords *Robinia pseudoacacia* · Genetic diversity · Genetic differentiation · Microsatellite · Population structure · Black locust

Introduction

The persistence and evolutionary potential of species depend on genetic diversity (Linda et al. 2010), which increases as the population size increases because larger populations contain more potential diversity, while small populations tend to lose genetic diversity over time through random events (genetic drift processes) (Frankham et al. 2010; Leimu et al. 2006; Fant et al. 2014). As a result, populations with more

² College of Agriculture, Henan University of Science and Technology, Luoyang 471023, P.R. China

³ Henan Academy of Forestry, Zhengzhou 450008, P.R. China

⁴ Feixian County State Owned Daqingshan Forest Farm, Linyi 273400, P.R. China

individuals usually have greater genetic diversity (Basey et al. 2015). In addition, genetic diversity can alleviate rapid anthropogenic environmental changes (Hughes et al. 2004; Jump et al. 2009). The hereditary basis of a breeding population determines breeding quality and potential after long-term natural or artificial selection (Jin et al. 2016). Therefore, understanding the genetic diversity of plant resources provides opportunities for breeders to develop new and improved cultivars with desirable characteristics and is also important for the maintenance of germplasm collections (Govindaraj et al. 2015; Duan et al. 2017).

Microsatellites are ideal for identifying individuals and studying genetic diversity because of their ubiquity, reproducibility, high level of polymorphism, codominant inheritance, and high level of transferability (Nasim et al. 2020; Lv et al. 2020; Guan et al. 2019; Carletti et al. 2019). Therefore, SSRs have been used to study genetic diversity, generate genetic linkage maps, and “fingerprinting” many economically important trees (e.g., *Ginkgo biloba*, *Platanus orientalis*, *Cunninghamia lanceolata* (Lamb.) Hook, and *Ziziphus jujuba* Mill. var. *jujuba*) (Jin et al. 2016; Duan et al. 2017; Xu et al. 2016; Zhou et al. 2020) including *R. pseudoacacia* L. (Fabaceae; black locust) (Guo et al. 2017, 2018; Dong et al. 2019a, b).

Black locust, originating in North America, is an important deciduous tree, that has been planted extensively in 27 provinces in China (Tian et al. 2003; Qiu et al. 2010; Zhang et al. 2015) and has been naturalized in this country since 1878. It possesses numerous useful characteristics, including a high relative growth rate and the production of a large biomass of high-density wood that is easy to dry and process and combusts well, making it one of the most economically important tree species in China (Rédei et al. 2008; Kropf et al. 2010; Benesperi et al. 2012; Boring and Swank 1984). The plant is stress-resistant; it has evolved to withstand drought and saline-alkaline soils and can grow well in barren soil. Therefore, it is an ideal choice for afforestation in northwest China. At the same time, it is a strong nitrogen-fixing tree and thus has potential for improving biodiversity due to its root nodules, which contain symbiotic nitrogen-fixing bacteria and parasitic tissues that allow root uptake of nitrogenous compounds (Rice et al. 2004; Nicolescu et al. 2018; Xu et al. 2019), and thus enhances plant growth, even in less-fertile soils. Unsurprisingly, demand for these valuable trees has increased.

Since the 1980s, toward fully exploring and exploiting the potential genetic capacity of *R. pseudoacacia*, phenotypic determination methods have been used to evaluate black locust throughout China, including the dominant wood comparative method, standard land method, index method, statistical test method, and scoring method (Shu 1988; Gu et al. 1990; Zhang et al. 1990; Xun et al. 2009). The selection and collection of high-quality trees has led to

the development of a seed orchard and improvement of the genetic quality of seeds available for artificial afforestation. That work was carried out through the cooperation of three divisions of forestry departments (city, county, farm) to construct breeding archives and monitor disease periodically to prevent damage (Hunt et al. 2004; Li et al. 2014; Yin et al. 2014; Zhao et al. 2014). Subsequently, the selected high-quality trees were used to establish large-scale seed orchards using clones with provenance in four provinces (Shandong, Gansu, Liaoning, and Shanxi) in China. Ex situ conservation of germplasm resources not only provides the basis for reintroduction but also is an important supplemental measure for in situ conservation (Ramsay et al. 2000; Li et al. 2018a, b). These high-quality trees were used in traditional breeding research, including the creation of new cultivars based on hybridization, estimation of volume growth, development and anatomy of the floral nectary, and studies of black locust resistance to physiological and biochemical stresses (Jiang et al. 2015; Zhang et al. 2016; Li et al. 2018a, b; Han et al. 2019; Wang et al. 2019). In addition, in several studies, black locust genetic diversity has been evaluated using molecular markers (Yin et al. 2014; Mao et al. 2017; Li et al. 2019; Dong et al. 2019a, b); however, small sample sizes and different types of markers were used. Therefore, they cannot be used to assess whether high-quality trees are used effectively in different orchards.

In the present study, we used 36 SSR loci to obtain genotyping data of *R. pseudoacacia* germplasm resources from orchards in China by nondenaturing polyacrylamide gel electrophoresis (PAGE). Their genetic diversity and structure were evaluated and systematically described. In addition, gene flow among orchards was revealed, and migrants and admixed individuals were identified. This information may be useful for the genetic differentiation of black locust in China and aid the conservation and breeding of this species.

Materials and methods

Plant materials and DNA extraction

A total of 687 strains representing four varieties and two secondary provenances of black locust in China were collected from April to May of 2017. In total, 332, 151, 62 and 96 clones of known varieties were collected from Daqingshan Forest Farm, Fei County, Linyi, Shandong Province (35°23'30" N, 106°06'37" E); Jianshan Forest Farm, Tianshui, Gansu Province (34°42'52" N, 106°06'37" E); the state-owned Linghai Hongqi Forest, Jinzhou, Liaoning Province (41°16'07" N, 121°00'48" E); and a Ji County seed orchard, Linfen, Shanxi (36°11'39" N, 110°42'09" E), respectively (Fig. S1). Forty-six seedlings were gathered in Henan, which included Mengjin (34°49'18" N, 112°28'12" E) and Minquan

(34°43'59" N, 115°05'10" E) two secondary provenances, and each sample was divided equally. The tender leaves of the collected plants were placed in a self-sealing bag containing silica gel and stored at room temperature in dark, dry conditions for DNA extraction.

Total DNA was isolated from all black locust samples through drying and extraction with DP320 Plant Genomic DNA kits (Tiangen, Beijing, China). After extraction, the integrity of the genomic DNA was determined using 1% agarose gel, and the DNA purity and quantity were measured using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA). Each DNA sample was diluted to an equal concentration of 20 ng/μL with TB elution buffer (Tiangen), labeled quickly, and placed in a quick-freeze box for storage at -80 °C until they were used for SSR amplification.

Microsatellite marker screening and SSR genotyping

Ninety-one pairs of SSR primers were used to screen eight different phenotypes and sources of black locust (Guo et al. 2018). The SSR primers were synthesized by TSINGKE Biological Technology Co., Ltd. (Beijing, China). PCR amplification system and procedure were reported previously by Guo et al. (Guo et al. 2017). Briefly, the 20-μL reaction mixture included 2 μL (20 ng/μL) genomic DNA, 1 μL (10 μM) each reverse and forward primer, 10 μL 2×PCR Master Mix (blue) (TSINGKE, Beijing, China), and 6 μL ddH₂O. PCR amplification was carried out using a Bio-Rad T100 thermal cycler (Hercules, CA, USA), with the described amplification protocol by Guo et al. (2017) and Schuelke (Schuelke 2000). Subsequently, the PCR products were resolved through 6% non-denaturing PAGE and visualized by silver staining and ultraviolet light also described by Guo et al. (Guo et al. 2018). After screening, the SSR markers rated as excellent were used for the amplification and analysis of all 687 black locust clones. The same 20-μL reaction volume was used for population SSR amplification, and the PCR was carried out using the Touchdown program described by Guo et al. (Guo et al. 2018). The PCR products were then stored in a freezer at -20 °C until retrieved the same afternoon (Ruibio BioTech Co., Ltd., Beijing, China) and separated using an ABI 3730XL DNA capillary electrophoresis analyzer (Applied Biosystems, Foster, CA, USA).

Data analysis

Genetic diversity

Allele sizes of SSRs were converted into various formats for further analysis using Convert ver. 1.3.1 (Glau-bitz 2004). Estimation of parameters of genetic diversity

including the number of different alleles (N_a), effective number of alleles (N_e), observed heterozygosity (H_o), expected heterozygosity (H_e) and Shannon's information index (I) for each SSR locus in four populations were calculated using GenAlEx ver. 6.501 (Peakall and Smouse 2005; Rod and Smouse 2012). The inbreeding coefficient (F_{IS}), Wright's fixation index (F_{IT}), and the fixation index (F_{ST}) for each microsatellite locus and population were estimated using FSTAT ver. 2.9.3 (Goudet 2001). Polymorphism information content (PIC) and gene diversity (H) were calculated using PowerMarker ver. 3.25 (Liu and Muse 2005). Whether the population was at Hardy-Weinberg equilibrium (HWE) was assessed based on Markov chain iterations and a neutrality test for all loci using Arlequin ver. 3.5 (Excoffier and Lischer 2010). Subsequently, null alleles were detected using Micro-Checker ver. 2.2.3 (Oosterhout et al. 2004). GENECLASS ver. 1.4 was employed to detect samples with identical genotypes (Wilberg and Dreher 2004).

Differentiation analysis

An analysis of molecular variance (AMOVA) was performed in GenAlEx ver. 6.501 to partition the total genetic variation among and within populations, and the number of private alleles for each population (PA) and Nei's genetic distance were also estimated (Rod and Smouse 2012; Peakall and Smouse 2005). The pairwise genetic distances [$F_{ST}/(1 - F_{ST})$] (Rousset 1997) between pairs of provinces were also estimated.

Population structure

The Bayesian model-based clustering algorithm in the software package STRUCTURE ver. 2.2.2 was used to estimate the population genetic structure based on the screened SSR markers (Pritchard et al. 2000). In this study, for each value of K ($K = 1-20$), 10 independent runs were performed with a burn-in period of 100,000 iterations and 1,000,000 Markov chain Monte Carlo replications, with the remaining parameters set to their default values. The online program Structure Harvester (Earl and Vonholdt 2012) was used to determine the optimal value of K based on the ΔK model developed by Evanno et al. (Evanno et al. 2005). To verify the K values, the calculations were repeated 10 times from $K_{\min} = 1$ to $K_{\max} = 20$ using the default parameters in Maverick software (Verity and Nichols 2016). The programs CLUMPAK ver. 1.1.2 (Kopelman et al. 2015) and DISTRICT ver. 1.1 (Rosenberg 2003) were used to create the bar plot of the probability of membership based on the Q -matrix results.

Principal coordinates and Mantel's analysis

Principal coordinate analysis (PCoA) based on genotype data of SSR markers were used to examine differences among and within populations and geographic regions using GenAlEx ver. 6.501.

Mantel's test of the genetic and geographic distances in the clone populations was performed using GenAlEx ver. 6.501 (Peakall and Smouse 2005; Rod and Smouse 2012), with geographic distance depending on the measured latitude.

Results

Selection of SSR markers for black locust populations in China

Forty-eight SSR markers were successfully amplified and passed the primer-screening process. We analyzed the primers before evaluating the genetic diversity among 687 samples of black locust from five provinces. GENECAAP ver. 1.4 was used to estimate the paired genotypes of all individuals, and there were no pairs of samples with identical genotypes. Four markers had more than 20% missing locus data. Two loci deviated significantly from predicted results from a neutral equilibrium model based on the Ewen–Watterson test. Additionally, null alleles were detected for nine markers among the 48 SSR loci. Finally, 12 pairs of SSR markers were deleted, and the remaining 36 markers were used for further analysis of black locust clonal populations in China (Table S1).

Genetic diversity of black locust in six improved variety bases

A total of 587 alleles were detected at 36 SSR loci in 687 samples, with the number of alleles per locus ranging from 8 (at locus Rp-22) to 52 (Rops15), and an average of 16.306 alleles per locus. The number of effective alleles per marker ranged from 1.083 to 8.306, with an average of 3.351. H_o ranged from 0.038 (at locus Rp-44) to 0.821 (Rp109), averaging 0.551 over all loci. H_e ranged from 0.065 (at locus Rp-44) to 0.873 (Rops15), averaging 0.608 over all loci. The variation in PIC and H was consistent, with the lowest values at Rp-44 (0.071 and 0.072, respectively), the highest at Rops18 (0.891 and 0.897, respectively), and averages of 0.612 and 0.645 over all loci, respectively (Table 1). At the population level, Shandong had the largest numbers of different and private alleles (Table S2) among all populations, probably due to the larger sample size (Glaubitz 2004). Correspondingly, Mengjin and Minquan had the fewest individual alleles, and also the fewest different and private

alleles. The highest I (1.504), H (0.659), and PIC (0.628) values were observed in the Shandong population, indicating rich genetic diversity. The second most diverse population was that of Shanxi, with I , H , and PIC values 1.393, 0.627, and 0.593, respectively. The H_o of the total population was smaller than the H_e , and the F_{IS} was consistently positive, ranging from 0.053 (Liaoning) to 0.151 (Shandong), with an average of 0.091 (Table S3).

Population structure analysis

A Bayesian model implemented in STRUCTURE was used to assess black locust population structure. The highest ΔK value was detected at $K=4$, which represented the most probable number of clusters (Fig. 1). Furthermore, the K -value was confirmed using Maverick software, which also determined that $K=4$ was the optimal value, indicating that all individuals should be divided into four clusters in this study (Fig. 2). From the results of 10 independent runs in STRUCTURE with $K=4$, the major mode accurately produced identical patterns of individual assignments for each run (Fig. S2). Cluster 2 contained the largest number of individuals ($n=289$), followed by Cluster 3 ($n=182$) and Cluster 1 ($n=118$). Cluster 4 contained the fewest individuals ($n=98$). The average Q -values were similar and were calculated to be 0.841 (Cluster 1), 0.912 (Cluster 2), 0.859 (Cluster 3), and 0.917 (Cluster 4). Among them, individuals from Shandong were mainly distributed in Cluster 1 ($n=107$, 32.23%) and Cluster 3 ($n=179$, 53.92%), whereas individuals from Gansu ($n=144$, 95.36%), Liaoning ($n=61$, 98.39%), Mengjin ($n=17$, 73.91%), and Minquan ($n=21$, 91.30%) were mainly distributed in Cluster 2. By contrast, all individuals in Shanxi ($n=96$) were placed in Cluster 4, indicating the presence of a strong population structure (Table S4). In addition, the results of genetic structure analysis with multiple K values (Fig. 3) and PCoA based on unweighted genetic distances were also consistent (Fig. 4), indicating a degree of genetic differentiation among the six populations of black locust in China.

Population differentiation analysis

The degree of genetic differentiation (F_{ST}) between any two populations was calculated for all six populations (Table 2). All pairwise F_{ST} combinations were significant ($P < 0.05$), with an overall F_{ST} value of 0.028 (< 0.05). Among all populations, F_{ST} values for global and pairwise multilocus analysis ranged from 0.008 (Liaoning vs. Gansu) to 0.053 (Shanxi vs. Minquan).

AMOVA was used to evaluate components of the variance among and within populations, as well as among and within clusters. The results revealed that variation was low

Table 1 Analysis of the genetic diversity of black locus populations in China with 36 simple sequence repeat (SSR) loci

Locus	N_a	N_e	I	H_o	H_e	F_{IS}	F_{IT}	F_{ST}	N_m	PIC	H	P_{HWE}
Rp-01	5.833	2.923	1.211	0.619	0.650	0.049	0.071	0.023	10.542	0.631	0.686	0.234
Rp-06	6.333	3.908	1.451	0.737	0.743	0.008	0.016	0.009	28.616	0.703	0.748	0.355
Rp-10	5.333	2.210	1.018	0.538	0.538	0.001	0.021	0.020	12.143	0.457	0.507	0.154
Rp-12	6.667	3.720	1.436	0.760	0.728	-0.044	-0.032	0.012	20.882	0.714	0.751	0.239
Rp-13	5.167	1.507	0.710	0.267	0.334	0.199	0.204	0.007	37.499	0.319	0.335	0.259
Rp-15	4.667	1.642	0.720	0.409	0.386	-0.059	0.144	0.192	1.052	0.422	0.452	0.522
Rp-21	6.667	1.997	0.731	0.277	0.325	0.149	0.390	0.283	0.633	0.637	0.652	0.323
Rp-22	6.000	3.038	1.334	0.640	0.666	0.040	0.051	0.011	22.079	0.616	0.658	0.373
Rp-23	5.167	2.056	0.905	0.454	0.503	0.097	0.118	0.023	10.627	0.512	0.567	0.233
Rp-28	6.500	3.285	1.349	0.668	0.693	0.036	0.060	0.025	9.796	0.667	0.715	0.302
Rp-31	5.500	1.758	0.846	0.250	0.422	0.409	0.418	0.015	16.196	0.407	0.435	0.000
Rp-33	4.500	1.894	0.757	0.467	0.441	-0.057	0.009	0.062	3.755	0.395	0.458	0.269
Rp-35	4.667	2.430	1.006	0.696	0.587	-0.186	-0.162	0.020	12.466	0.529	0.608	0.042
Rp-36	5.333	2.965	1.188	0.601	0.651	0.078	0.095	0.019	12.738	0.653	0.706	0.254
Rp-37	7.167	3.301	1.404	0.493	0.695	0.291	0.305	0.020	12.285	0.683	0.720	0.022
Rp-38	11.500	4.517	1.771	0.774	0.777	0.004	0.028	0.023	10.423	0.776	0.801	0.186
Rp-41	7.167	2.045	1.022	0.490	0.488	-0.004	0.027	0.031	7.844	0.550	0.579	0.205
Rp-43	4.667	2.013	0.853	0.594	0.479	-0.241	-0.098	0.115	1.917	0.546	0.590	0.004
Rp-44	4.000	1.083	0.164	0.038	0.065	0.411	0.442	0.053	4.452	0.071	0.072	0.167
Rply2	8.000	3.654	1.442	0.718	0.720	0.004	0.026	0.023	10.817	0.739	0.771	0.188
Rply3	7.000	4.176	1.584	0.515	0.760	0.322	0.335	0.020	12.411	0.633	0.674	0.059
Rply5	7.833	3.070	1.359	0.666	0.663	-0.004	0.044	0.049	4.884	0.677	0.720	0.222
Rply15	7.167	2.654	1.259	0.603	0.610	0.011	0.061	0.050	4.704	0.659	0.686	0.031
Rply16	12.500	2.234	1.281	0.464	0.529	0.122	0.137	0.017	14.677	0.508	0.526	0.389
Rply22	6.333	2.754	1.140	0.557	0.593	0.061	0.260	0.212	0.931	0.721	0.757	0.300
Rply28	4.167	1.800	0.741	0.393	0.434	0.095	0.266	0.188	1.076	0.482	0.545	0.418
Rply33	6.167	1.423	0.613	0.164	0.282	0.418	0.433	0.026	9.294	0.237	0.248	0.129
Rply49	15.833	4.517	1.849	0.533	0.768	0.306	0.322	0.023	10.808	0.794	0.813	0.046
Rops08	8.167	2.854	1.374	0.572	0.643	0.109	0.125	0.018	13.438	0.831	0.847	0.247
Rops15	26.000	8.306	2.542	0.566	0.873	0.352	0.363	0.016	14.987	0.615	0.642	0.010
Rops18	17.667	7.295	2.248	0.598	0.856	0.301	0.313	0.017	14.765	0.891	0.897	0.000
Rp01B	10.500	4.235	1.798	0.779	0.750	-0.039	-0.023	0.015	16.072	0.881	0.891	0.393
Rp032	11.833	5.808	1.960	0.644	0.821	0.215	0.234	0.024	10.304	0.739	0.756	0.029
Rp106	6.333	3.833	1.429	0.726	0.738	0.017	0.025	0.008	29.701	0.819	0.837	0.353
Rp109	15.000	5.779	2.056	0.821	0.823	0.003	0.018	0.015	16.059	0.700	0.743	0.119
Rp200	17.333	7.958	2.323	0.759	0.870	0.128	0.143	0.017	14.491	0.822	0.837	0.299
Mean	8.352	3.351	1.302	0.551	0.608	0.100	0.144	0.047	12.093	0.612	0.645	0.205

N_a number of different alleles; N_e effective number of alleles; I Shannon's information index; H_o observed heterozygosity; H_e Expected heterozygosity; F_{IS} Inbreeding coefficient; F_{IT} Wright's fixation index; F_{ST} Fixation index; N_m gene flow; PPL percentage of polymorphic loci; H gene diversity; PIC polymorphism information content; P_{HWE} P -value for Hardy-Weinberg equilibrium

among populations (7%) and clusters (9%) and high within populations (93%) and clusters (91%) (Table 3).

Nei's unbiased genetic distance (D) and genetic consistency were determined between pairs of populations. Among all pairs, Liaoning and Gansu exhibited the smallest D (0.010) and the largest genetic identity (0.990). By contrast, Minquan and Shanxi exhibited the largest D (0.158) and smallest genetic identity (0.854) (Table 4). In addition, Mantel's test using data from the six populations revealed

that there was no correlation between genetic distance and geographic distance (Fig. 5).

Discussion

To establish suitable breeding, conservation, and management strategies, we need to collect and identify germplasm resources and study their genetic differences. To date, such

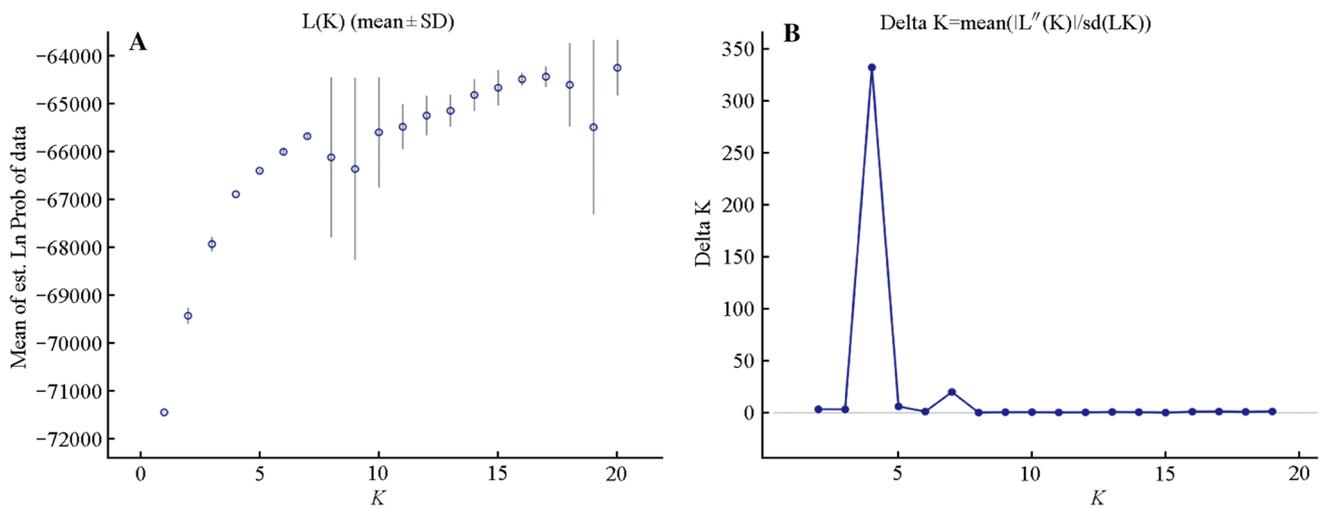


Fig. 1 Estimation of the optimal number of *Robinia pseudoacacia* populations performed using STRUCTURE software. **a** The mean Ln $P(D)$ was based on 10 repeats for each K -value. **b** Optimal K -value according to maximum ΔK

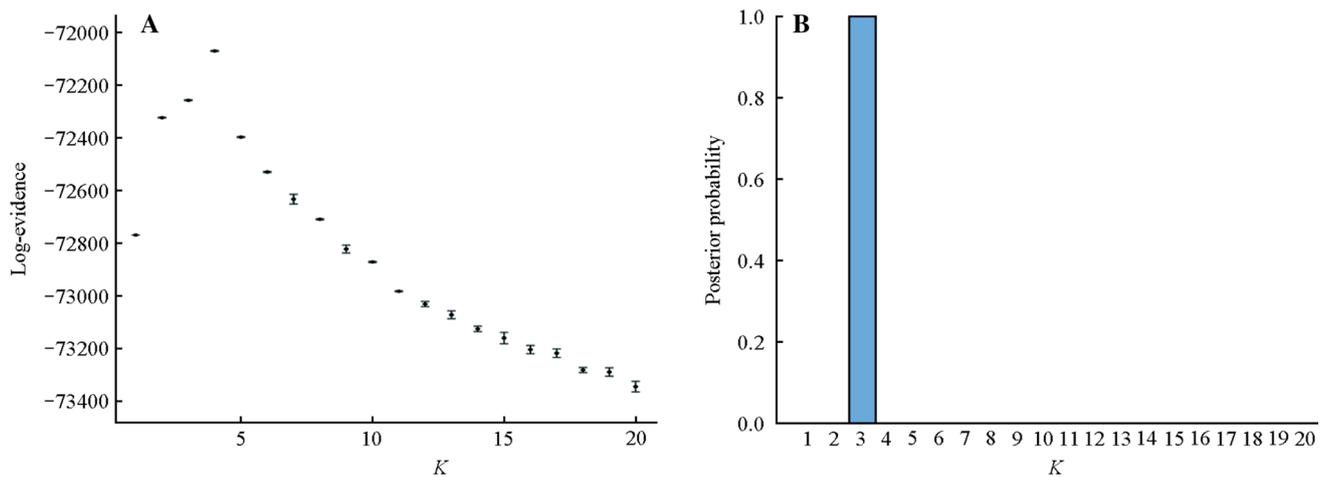


Fig. 2 Estimation of the optimal number of *R. pseudoacacia* populations performed using Maverick software. **a** Values of $\log(K)$ obtained from the Maverick analysis. **b** Values in linear space after normalizing to a sum of 1

research on the genetic diversity in *R. pseudoacacia* in China has been limited; in particular, the levels of genetic diversity among *R. pseudoacacia* populations and the types of genetic resources available has been unclear. For these reasons, 687 individuals representing four known high-quality varieties and two secondary provenances were collected, comprising 641 clones and 46 seedlings. A set of neutral SSR markers was obtained via a series of screening procedures and used to assess germplasm diversity. Additionally, individuals collected from different regions may have the same genotype, so GeneCap ver. 1.4 was used to analyze all strains. Of the 687 samples, no individuals shared the same genotype, thus indicating that all samples could be used for subsequent analysis.

In the present study, the average Shannon's information index value (I) within 687 black locust genotypes by 36 neutral SSR markers was 1.302. The value was higher than that of the *R. pseudoacacia* samples from 10 main planting districts in China using AFLP and ISSR markers (Huo et al. 2009; Sun et al. 2009). The higher I value may be related to use the different types of molecular markers and/or the sources of materials (Lu et al. 2020; Xiong et al. 2019). Sequentially, 12 EST-SSR markers to evaluate 123 black locust cultivars in China were compared with those in our study, showing lower values for N_a , N_e and I in our populations (Dong et al. 2019a, b). The high number of the above parameters observed in this study may be due to the simultaneous use of two types of SSR markers and the relatively

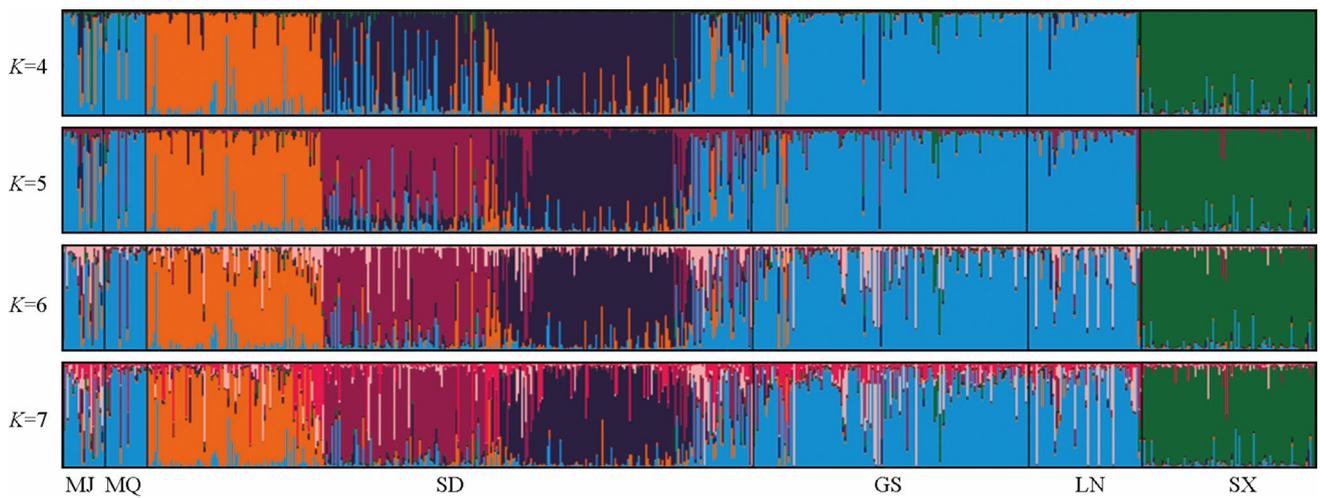


Fig. 3 Genetic structure of black locust populations in China ($K=4-7$). Each individual is shown as a vertical line partitioned into K colored segments whose length is proportional to the individual coefficients of membership in $K=4-K=7$ genetic clusters that rep-

resent the populations assessed. The populations from left to right in the above figure are MJ (Mengjin), MQ (Minquan), SD (Shandong), GS (Gansu), LN (Liaoning) and SX (Shanxi), respectively. And among them, MJ and MQ belong to HN (Henan Province)

Fig. 4 Principal coordinate analysis (PCoA) results for black locust geographic populations in China based on Euclidean distance. *SD* Shandong; *GS* Gansu; *LN* Liaoning; *SX* Shangxi; *MJ* Mengjin; *MQ* Minquan

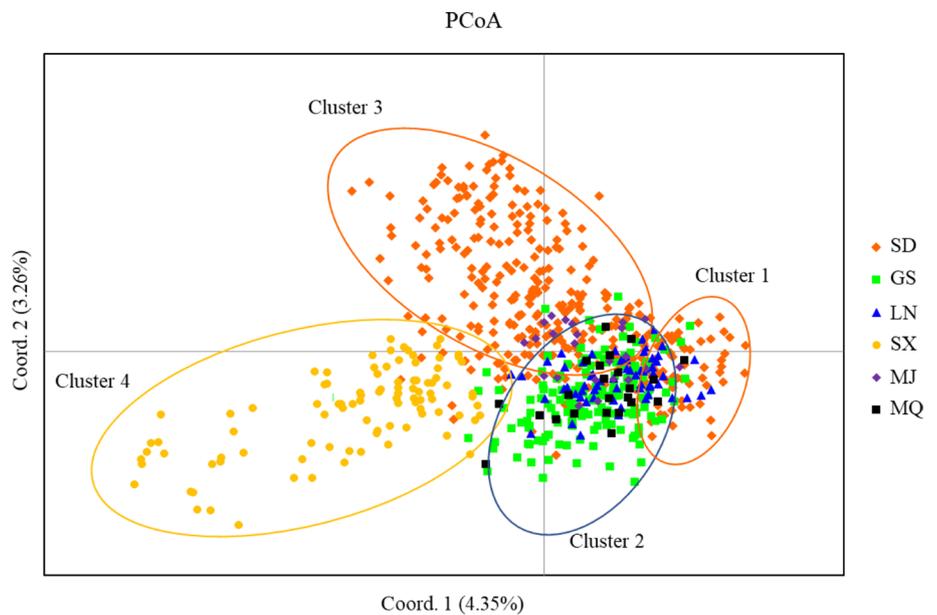


Table 2 Pairwise genetic distance (based on fixation index F_{ST} ; lower diagonal) between black locust populations (Pop.) in China and the corresponding significance tests (upper diagonal)

Pop	SD	GS	LN	SX	MJ	MQ
SD (Shandong)	1.000	*	*	*	*	*
GS (Gansu)	0.014	1.000	*	*	*	*
LN (Liaoning)	0.016	0.008	1.000	*	*	*
SX (Shanxi)	0.040	0.044	0.050	1.000	*	*
MJ (Mengjin)	0.019	0.021	0.016	0.052	1.000	*
MQ (Minquan)	0.019	0.013	0.013	0.053	0.022	1.000

* $P \leq 0.05$

Table 3 Results of analysis of molecular variance (AMOVA) among black locust populations (pops.) in China based on 36 paired primers

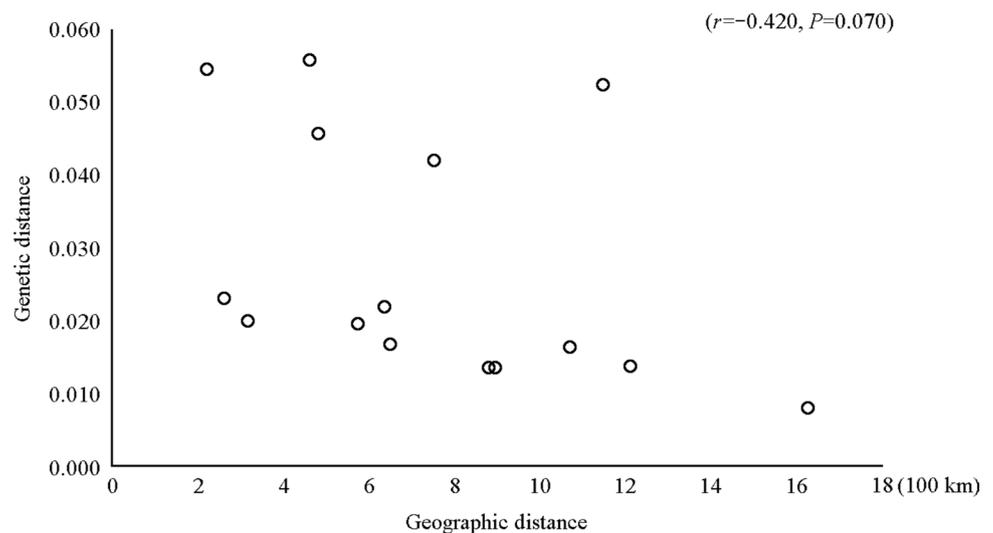
Source of variation	df	SS	MS	Est. Var	% variation	P
Among pops	5	1142.652	228.530	2.129	7	<0.01
Within pops	681	18 528.360	27.280	27.208	93	<0.01
Total	686	19 671.012		29.337	100	
Among clusters	3	1402.730	467.577	2.737	9	<0.01
Within clusters	683	18 268.281	26.747	26.747	91	<0.01
Total	686	19 671.012		29.484	100	

df degrees of freedom; SS sum of squares; MS mean square; Est. var. estimated variance

Table 4 Nei's unbiased genetic distance (lower diagonal) and genetic identity (upper diagonal) between black locust populations (Pops.) in China

Pops	SD	GS	LN	SX	MJ	MQ
SD	1.000	0.958	0.960	0.861	0.960	0.956
GS	0.043	1.000	0.990	0.865	0.960	0.982
LN	0.041	0.010	1.000	0.858	0.976	0.985
SX	0.150	0.145	0.154	1.000	0.857	0.854
MJ	0.041	0.041	0.025	0.154	1.000	0.969
MQ	0.045	0.019	0.015	0.158	0.032	1.000

SD Shandong; GS Gansu; LN Liaoning; SX Shanxi; MJ Mengjin; MQ Mengjin

Fig. 5 Relationship between genetic distance and geographic distance for all black locust populations in China

large sample size compared with a population growing at the geographic origin of black locust (Guo et al. 2018). In addition, values of genetic diversity indices including N_a , I , PIC, and H in the present study were higher than in those from previous research, which may be due to the large number of sequences analyzed. Species with higher genetic diversity may have stronger adaptability to a changing environment, and such adaptability may also reflect the ability of a particular genotype to tolerate multiple conditions (Schaal et al. 1991; Hawtin et al. 1996). Although black locust is an exotic tree species, the genetic diversity of this species' germplasm resources in China and that of native populations in the United States are similar. Furthermore, the analysis of

variance using SPSS ver. 24 revealed that there was no significant difference in the level of diversity between the two groups (Guo et al. 2018). For these reasons, black locust is widely distributed and exhibits strong adaptability in China (Table S5).

H_e is a critical measurement of genetic diversity. In this study, the average observed H_o and H_e were 0.551 and 0.608, respectively. Similar results were obtained by Lian et al. (2002) ($H_o=0.615$, $H_e=0.773$) and Mishima et al. (Mishima et al. 2009) ($H_o=0.661$, $H_e=0.739$) using SSR markers. In addition, at both the locus and population levels, we found that most H_e values were higher than the H_o values, indicating that heterozygote deficiency is possible.

Furthermore, the F_{IS} value was 0.100 at the locus level and 0.091 at the population level, suggesting that self-pollination or the Wahlund effect may occur in this species (Selkoe and Toonen 2006).

F_{ST} is usually used to measure the degree of genetic differentiation between populations. All pairwise comparisons of F_{ST} indicated that the differences were significant, and most (80%) were less than 0.05 (Table 3), implying that the genetic differentiation among populations was notable but the degree of differentiation was low, which is consistent with the findings of Guo et al. (Guo et al. 2018). Moreover, AMOVA results revealed that there was only 7% genetic variation among populations and 9% genetic variation among clusters. These results indicate that the observed variation was mainly attributable to within-population and within-cluster differences. The long growth cycle and outcrossing dominance of woody plants may lead to an increase in genetic diversity among individuals and a reduction in differences among populations and clusters (Xu et al. 2018). In addition, the analysis of Nei's unbiased genetic distance (D) and identity showed that there was genetic differentiation among populations, indicating that the variation among populations may not be related to their geographic distribution. Furthermore, the Mantel test indicated that in the populations sampled, genetic variation was not significantly correlated with geographic distance. These results are consistent with those previously reported (Guo et al. 2018; Sun et al. 2009; Hamrick et al. 1989; Yang et al. 2004).

In the population structure analysis using STRUCTURE and Maverick software, four clusters were identified within the 687 black locust samples. Moreover, genetic structure analysis using multiple K values and PCoA produced results consistent with those of the tests described above. Interestingly, we found that samples from Shandong Province were assigned mainly to two clusters (Clusters 1 and 3). Additionally, genetic diversity analysis demonstrated that the diversity and NA are generally higher in populations from Shandong Province. Samples from Gansu, Liaoning, Mengjin, and Minquan were mainly grouped into Cluster 3. Based on these results, we propose that individuals within the same cluster share a common origin (i.e., kinship), whereas their relationships with individuals in other clusters are more distant. Despite the loss of information about the geographic origin, introduction, and planting of the sampled trees, we could roughly determine the source of each individual within a group of black locust; for instance, the genetic relationship between trees from Shanxi and those from the other five populations is distant, whereas the genetic relationships of strains sampled in Gansu, Liaoning, Mengjin, and Minquan are closer. Taken together, it is the first broad analysis of the population genetic diversity and differences in an extensive collection of black locust germplasm resources sampled from multiple populations in China. Our results will help

with elucidating the genetic relationships among populations of this species that lack introduction information. In addition, this research also provides valuable genetic information for the breeding and conservation of black locust in China.

Conclusion

In this study, we conducted the first evaluation of black locust germplasm resources in China. We (1) identified a large number of germplasm resources representing several varieties in China, (2) found that the populations contain high genetic diversity, (3) observed that heterozygote deficiency existed in all populations, and (4) determined that geographic distance is not the main driver of black locust genetic structure. These results provide comprehensive and important information for the breeding of black locust resources in China.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Basey A, Fant JB, Kramer A (2015) Producing native plant materials for restoration: 10 rules to collect and maintain genetic diversity. *Native Plants J* 16:37–53
- Benesperi R, Giuliani C, Zanetti S, Gennai M, Lippi MM, Guidi T, Nascimbene J, Foggi B (2012) Forest plant diversity is threatened by *Robinia pseudoacacia* (black-locust) invasion. *Biodivers Conserv* 21(14):3555–3568
- Boring LR, Swank WT (1984) The role of black locust (*Robinia pseudoacacia*) in forest succession. *J Ecol* 72(3):749–766
- Carletti G, Cattivelli L, Vietto L, Nervo G (2021) Multiallelic and multilocus simple sequence repeats (SSRs) to assess the genetic diversity of a *Salix* spp germplasm collection. *J Forestry Res* 32(1):263–217
- Dong L, Sun YH, Zhao KQ, Zhang J, Li Y (2019a) Development and Application of EST-SSR Markers for DNA Fingerprinting and Genetic Diversity Analysis of the Main Cultivars of Black Locust (*Robinia pseudoacacia* L.) in China. *Forests* 10:644. <https://doi.org/10.3390/f10080644>
- Dong L, Zhang JT, Wen YZ, Niu DS, Tian SY, Sun YH, Li Y (2019b) Genetic Differences of Black Locust (*Robinia pseudoacacia*) with Genomic-SSR and EST-SSR Markers. *Chin Agric Sci Bull* 35(19):49–57
- Duan HJ, Cao S, Zheng HQ, Hu DH, Lin J, Cui BB, Lin HZ, Hu RY, Wu B, Sun YH (2017) Genetic characterization of Chinese fir

- from six provinces in Southern China and construction of a core collection. *Sci Rep* 7(1):13814
- Earl DA, Vonholdt BM (2012) Structure harvester: a website and program for visualizing structure output and implementing the evanno method. *Conserv Genet Resour* 4(2):359–361
- Evanno GS, Regnaut SJ, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol Ecol* 14(8):2611–2620
- Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Res* 10(3):564–567
- Fant JB, Havens K, Keller JM, Radosavljevic A, Yates ED (2014) The influence of contemporary and historic landscape features on the genetic structure of the sand dune endemic, *Cirsium picheri* (Asteraceae). *Heredity* 112(5):519–530
- Frankham R, Ballou J, Briscoe D (2010) Introduction to conservation genetics. Cambridge University Press, Cambridge
- Glaubitz JC (2004) Convert: a user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages. *Mol Ecol Notes* 4:309–310
- Goudet J (2001) FSTAT (version 2.9.3) : a program to estimate and test gene diversities and fixation indices, version 2.9.3 GOUDET J. <http://www2.unil.ch/popgen/softwares/fstat.htm>
- Govindaraj M, Vetriventhan M, Srinivasan M (2015) Importance of genetic diversity assessment in crop plants and its recent advances: an overview of its analytical perspectives. *Genet Res Int* 2015:1–14
- Gu WC, Wang QY, Zhou ZH, Liu DA (1990) Genetic difference and selection evaluation of secondary provenance of *Robinia Pseudoacacia*. *For Res* 1:70–75
- Guan CF, Zhang PX, Hu CQ, Chachar S, Riaz A, Wang R, Yang Y (2019) Genetic diversity, germplasm identification and population structure of *Diospyros kaki* Thunb. from different geographic regions in China using SSR markers. *Sci Hortic* 251:233–240
- Guo Q, Wang JX, Su LZ, Lv W, Sun YH, Li Y (2017) Development and evaluation of a novel set of EST-SSR markers based on transcriptome sequences of black locust (*Robinia pseudoacacia* L.). *Genes*. <https://doi.org/10.3390/genes8070177>
- Guo Q, Li XY, Yang SH, Yang ZH, Sun YH, Zhang JT, Cao S, Dong L, Uddin S, Li Y (2018) Evaluation of the genetic diversity and differentiation of black locust (*Robinia pseudoacacia* L.) based on genomic and expressed sequence tag-simple sequence repeats. *Int J Mol Sci* 19(9):2492
- Hamrick JL, Bongarten BC, Surlis SE (1989) Allozyme variation in black locust (*Robinia pseudoacacia*). *Can J for Res* 19(19):471–479
- Han CC, Xun SH, Zhang YS, Qiao Y, Dong YF, Zhong WG, Liu ZD, Li QH (2019) A New Cultivar of *Robinia pseudoacacia* “Lüman Qingshan” (in Chinese). *Acta Horticulturae Sinica* 46:1425–1426
- Hawtin G, Iwanaga M, Hodgkin T (1996) Genetic resources in breeding for adaptation. *Euphytica* 92(1–2):255–266
- Hughes AR, Stachowicz JJ, Tilman D (2004) Genetic diversity enhances the resistance of a seagrass ecosystem to disturbance. *Proc Natl Acad Sci* 101:8998–9002
- Hunt J, Bussière LF, Jennions MD, Brooks R (2004) What is genetic quality? *Trends Ecol Evol* 19(6):329–333
- Huo XM, Han HW, Zhang J, Yang MS (2009) Genetic diversity of *Robinia pseudoacacia* populations in China detected by AFLP markers. *Front Mech Eng China* 3(3):337–345
- Jiang LN, Xun SH, Sun BY, Zhang YS, Mao XH, Zhang ZH, Qiao YL (2015) Determination and preservation of the pollen vitality of *Robinia neo-mexicana* var. *Luxurians*. *J Shandong For Sci Technol* 45:10–15
- Jin YQ, Ma YP, Wang S, Hu XG, Huang LS, Li Y, Wang XR, Mao JF (2016) Genetic evaluation of the breeding population of a valuable reforestation conifer *Platycladus orientalis* (Cupressaceae). *Sci Rep* 6(1):34821
- Jump AS, Marchant R, Peñuelas J (2009) Environmental change and the option value of genetic diversity. *Trends Plant Sci* 14(1):51–58
- 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*. *Mol Ecol Resour* 15(5):1179–1191
- Kropf U, Korosec M, Bertonec J, Ogrinc N, Necemer M, Kump P, Golob T (2010) Determination of the geographical origin of Slovenian black locust, lime and chestnut honey. *Food Chem* 121(3):839–846
- Leimu R, Mutikainen P, Koricheva J, Fischer M (2006) How general are positive relationships between plant population size, fitness and genetic variation? *J Ecol* 94:942–952
- Li B, Zheng YQ, Lin FR, Li WY (2014) The state of utilization and sustainable management of forest genetic resources of China. *J Plant Gen Res* 15:1390–1394
- Li GX, Xun SH, Zhang YS, Xu GX, Zhang ZH, Wang YG, Zhong WG, Liu ZD, Dong YF, Li XL (2018) Estimation of volume growth and shape of new varieties of *Robinia Pseudoacacia* L. *J Shandong For Sci Technol* 48:15–18
- Li BJ, Wang JY, Liu ZJ, Zhuang XY, Huang JX (2018) Genetic diversity and ex situ conservation of *Loropetalum subcordatum*, an endangered species endemic to China. *BMC Genet* 19(1):12
- Li X, Guo Q, Dong L, Sun Y, Niu D, Liu J, Wang H, Li Y (2019) EST-SSR analysis of genetic diversity of *Robinia pseudoacacia* clones in Jixian County, Shanxi Province of northern China. *J Beijing For Univ* 41:39–48
- Linda L, Allendorf FW, Aroner LC, Baker CS, Gregovich DP, Hansen MM, Jackson JA, Kendall KC, McKelvey K, Neel MC, Olivieri I, Ryman N, Schwartz MK, Bull RS, Stetz JB, Tallmon DA, Taylor BL, Vojta CD, Waller DM, Waples RS (2010) Neglect of genetic diversity in implementation of the convention on biological diversity. *Conserv Biol* 24(1):86–88
- Liu K, Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics* 21:2128–2129
- Lu M, Zhang HS, An HM, Zhou W (2020) Genetic variation and population structure of *Rosa roxburghii* by EST-based and genomic SSR markers. *Pak J Bot* 52(4):1315–1322
- Lv JB, Li CG, Zhou CP, Chen JB, Li FG, Weng QJ, Li M, Wang YQ, Chen SK, Chen JC (2020) Genetic diversity analysis of a breeding population of *Eucalyptus cloeziana* F. Muell. (Myrtaceae) and extraction of a core germplasm collection using microsatellite markers. *Ind Crops Products*. <https://doi.org/10.1016/j.indcrop.2020.112157>
- Mao XH, Zheng YQ, Sun BY, Zhang YS, Han CC, Wei X, Xun SH (2017) Genetic Diversity and Fingerprints of *Robinia pseudoacacia* Clones Based on SSR Markers. *Scientia Silvae Sinicae* 53:80–89
- Mishima K, Hirao T, Urano S, Watanabe A, Takata K (2009) Isolation and characterization of microsatellite markers from *Robinia pseudoacacia* L. *Mol Ecol Resour* 9(3):850–852
- Nasim N, Sandeep IS, Ambika S, Das S, Panda MK, Acharya L, RamaRao VV, Sanghamitra N, Mohanty S (2020) Population genetic structure and diversity analysis in economically important *Pandanus odorifer* (Forssk.) Kuntze accessions employing ISSR and SSR markers. *Ind Crops Prod*. <https://doi.org/10.1016/j.indcrop.2019.111894>
- Nicolescu VN, Hernea C, Bakti B, Keserü Z, Antal B, Rédei K (2018) Black locust (*Robinia pseudoacacia* L.) as a multi-purpose tree species in Hungary and Romania: a review. *J For Res* 29:1449–1463

- Oosterhout CV, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes* 4:535–538
- Peakall R, Smouse PE (2005) genalex analysis in Excel. population genetic software for teaching and research. *Mol Ecol Notes* 6(1):288–295
- Pritchard JK, Stephens MJ, Donnelly PJ (2000) Inference of population structure using Multilocus genotype data. *Genetics* 155:945–959
- Qiu L, Zhang X, Cheng J, Yin X (2010) Effects of black locust (*Robinia pseudoacacia*) on soil properties in the loessial gully region of the Loess Plateau. *China Plant Soil* 332(1–2):207–217
- Ramsay MM, Jacscon AD, Porley R (2000) A pilot study for ex situ conservation of UK bryophytes. In: BGCI (ed) Eurogard 2000—II European botanic garden congress. EBGC, Las Palmas de Gran Canaria, pp 52–57
- Rédei K, Osváthbujtás Z, Veperdi I (2008) Black locust (*Robinia pseudoacacia* L.) Improvement in Hungary: a Review. *Acta Silv Lignaria Hung* 4(4):127–132
- Rice SK, Westerman B, Federici R (2004) Impacts of the Exotic, Nitrogen-Fixing Black Locust (*Robinia pseudoacacia*) on Nitrogen-Cycling in a Pine-Oak Ecosystem. *Plant Ecol* 174(1):97–107
- Rod P, Smouse PE (2012) GenAEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28:2537–2539
- Rosenberg NA (2003) distruct: a program for the graphical display of population structure. *Mol Ecol Resour* 4(1):137–138
- Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* 145(4):1219–1228
- Schaal BA, O’Kane SL, Rogstad SH (1991) DNA variation in plant populations. *Trends Ecol Evol* 6(10):329–333
- Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. *Nat Biotechnol* 18(2):233–234
- Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecol Lett* 9(5):615–629
- Shu CH (1988) Determination of excellent clones of *Robinia Pseudoacacia* L. *Shandong For Sci Technol* 2:46–50
- Sun F, Yang MS, Zhang J, Gu JT (2009) ISSR analysis of genetic diversity of *Robinia pseudoacacia* populations. *J Plant Gen Res* 10(01):91–96
- Tian CJ, He XY, Zhong Y, Chen JK (2003) Effect of inoculation with ecto- and arbuscular mycorrhizae and Rhizobium on the growth and nitrogen fixation by black locust, *Robinia pseudoacacia*. *New Forest* 25:125–131
- Verity R, Nichols RA (2016) Documentation for Maveric K software: Version 1.0. *Genetics*
- Wang Y, Xun Sh, Mao Xh, Han CC, Sun JW, Zhang YS, Dong YF (2019) Study on the development and anatomical of floral nectary of *Robinia pseudoacacia*. *J Arid Land Res Environ* 33:171–175
- Wilberg MJ, Dreher BP (2004) genecap: a program for analysis of multilocus genotype data for non-invasive sampling and capture-recapture population estimation. *Mol Ecol Resour* 4(4):783–785
- Xiong YL, Liu WH, Xiong Y, Yu QQ, Ma X, Lei X, Zhang XQ, Li DX (2019) Revelation of genetic diversity and structure of wild *Elymus excelsus* (Poaceae: Triticeae) collection from western China by SSR markers. *PeerJ* 7(2):e8038
- Xu CQ, Gao J, Du ZF, Li DK, Wang Z, Li YY, Pang XM (2016) Identifying the genetic diversity, genetic structure and a core collection of *Ziziphus jujuba* Mill. var. *jujuba* accessions using microsatellite markers. *Sci Rep* 6:31503. <https://doi.org/10.1038/srep31503>
- Xu XY, Zhou CP, Yong Z, Zhang WQ, Gan SM (2018) A novel set of 223 EST-SSR markers in *Casuarina* L. ex Adans.: polymorphisms, cross-species transferability, and utility for commercial clone genotyping. *Tree Genet Genomes* 14(2):30. <https://doi.org/10.1007/s11295-018-1246-0>
- Xu YD, Wang T, Li H, Ren CJ, Chen JW, Yang GH, Han XH, Feng YZ, Ren GX, Wang XJ (2019) Variations of soil nitrogen-fixing microorganism communities and nitrogen fractions in a *Robinia pseudoacacia* chronosequence on the Loess Plateau of China. *CATENA* 174:316–323
- Xun SH, Qiao YL, Zhang JT, Zhang LJ, Bai J, Dou YQ (2009) Research progress and development tactics on genetics breeding of *Robinia Pseudoacacia* L. in China. *J Shandong For Sci Technol* 39:92–96
- Yang MS, Heike H, Volker S (2004) Allozyme Variability of Provenance Populations of *Robinia pseudoacacia* from Middle Europe. *Acta Genet Sin* 31(12):1440–1447
- Yin D, Du N, Xu F, Wang RQ (2014) Research progress of Alien species of *Robinia pseudoacacia* L. in China. *J Shandong For Sci Technol* 44:92–99
- Zhang DL, Zhang ZF, Wang FQ (1990) Study on clonal selection of *Robinia Pseudoacacia* L. *Shandong For Sci Technol* 2:16–20
- Zhang JG, Guan JH, Shi WY, Yamanaka N, Du S (2015) Interannual variation in stand transpiration estimated by sap flow measurement in a semi-arid black locust plantation, Loess Plateau. *China Ecohydrol* 8(1):137–147
- Zhang GJ, Sun YH, Li Y (2016) A New variety of *Robinia pseudoacacia* ‘Beilinhuai 3.’ *Scientia Silvae Sinicae* 52:1425–1426
- Zhao H, Zong YC, Zheng YQ (2014) Current status of protection of forest genetic resources in China. *Hunan For Sci Technol* 41:80–85
- Zhou Q, Mu KM, Ni ZX, Liu XH, Li YG, Xu LA (2020) Analysis of genetic diversity of ancient *Ginkgo* populations using SSR markers. *Ind Crops Prod* 145:111942

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.