



## OPEN ACCESS

## EDITED BY

Jian Zeng,  
Sichuan Agricultural University, China

## REVIEWED BY

Sathish Kumar Ponniah,  
University of Arkansas at Pine Bluff,  
United States  
Jiban Shrestha,  
Nepal Agricultural Research Council, Nepal  
Rituraj Khound,  
University of Nebraska-Lincoln, United States

## \*CORRESPONDENCE

Minxuan Liu

✉ liuminxuan@caas.cn

Baili Feng

✉ fengbaili@nwsuaf.edu.cn

RECEIVED 18 June 2025

ACCEPTED 18 August 2025

PUBLISHED 08 September 2025

## CITATION

Lv Z, Yang Y, Hou H, Yang S, Cui Z, Zhang X,  
Li J, Yuan Y, Liu M and Feng B (2025)  
Genetic diversity analysis of proso millet  
(*Panicum miliaceum* L.) germplasm resources  
based on phenotypic traits and SSR markers.  
*Front. Plant Sci.* 16:1649200.  
doi: 10.3389/fpls.2025.1649200

## COPYRIGHT

© 2025 Lv, Yang, Hou, Yang, Cui, Zhang, Li,  
Yuan, Liu and Feng. This is an open-access  
article distributed under the terms of the  
Creative Commons Attribution License (CC BY).  
The use, distribution or reproduction in other  
forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication in  
this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# Genetic diversity analysis of proso millet (*Panicum miliaceum* L.) germplasm resources based on phenotypic traits and SSR markers

Ziyang Lv<sup>1</sup>, Yixuan Yang<sup>1</sup>, Hanghang Hou<sup>1</sup>, Shangkun Yang<sup>1</sup>,  
Zhijia Cui<sup>1</sup>, Xi Zhang<sup>1</sup>, Jing Li<sup>2</sup>, Yuhao Yuan<sup>3</sup>, Minxuan Liu<sup>4\*</sup>  
and Baili Feng<sup>1\*</sup>

<sup>1</sup>State Key Laboratory of Crop Stress Biology in Arid Areas, College of Agronomy, Northwest A&F University, Yangling, Shaanxi, China, <sup>2</sup>School of Life Sciences, Shaanxi Key Laboratory of Research and Utilization of Resource Plants on the Loess Plateau, Yan'an University, Yan'an, Shaanxi, China, <sup>3</sup>College of Agronomy, Henan Agricultural University, Zhengzhou, Henan, China, <sup>4</sup>Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing, China

**Introduction:** Germplasm resources are vital for food security and agricultural sustainability, providing the basis for seed industry innovation.

**Methods:** In the present study, 1,582 proso millet (*Panicum miliaceum* L.) germplasm resources, comprising landraces, cultivars, wild varieties, and foreign varieties, were sown in early June 2024 at the Yulin Minor Grain Comprehensive Experimental Demonstration Station. Subsequently, the genetic diversity was analyzed according to 14 agronomic, 5 yield, and 3 grain traits.

**Results:** Significant phenotypic diversity was observed: agronomic trait diversity indices ranged from 0.15 to 2.10, with straw weight per plant showing the highest variation coefficients (60.03%). Yield traits exhibited diversity indices of 1.95–2.08 and variation coefficients of 14.94–37.37%. Grain traits had diversity indices exceeding 2, with the lowest variation coefficients (5.22–6.61%). Principal component analysis identified 4 key components, with tiller number and panicle length having the highest loading, leading to the selection of 10 superior germplasms. Cluster analysis grouped 1,582 samples into 5 categories, from which 147 representative germplasms were chosen. Then, 80 SSR primers were designed, 15 of which revealed polymorphism, confirming high genetic variation among these 147 germplasms. Varieties from Northwest China and Loess Plateau region showed the greater diversity.

**Conclusion:** These findings provide a scientific foundation for the efficient utilization and breeding of proso millet germplasm resources.

## KEYWORDS

proso millet, germplasm resources, SSR marker, genetic diversity, genetic structure

## Highlights

- The diversity of 1,582 global proso millet germplasm resources was evaluated.
- Phenotypic and molecular data were combined to assess genetic diversity.
- SSR markers were developed to provide a foundation for genetic breeding.
- Fingerprint maps were constructed to facilitate digital management.

## Introduction

Proso millet (*Panicum miliaceum* L.) is one of the oldest domesticated crops worldwide, with its history of cultivation dating back to 10,000 years ago in China (Wang et al., 2016a, b; Shi et al., 2019; Jiang et al., 2024). Owing to its long cultivation history, extensive geographical spread, and prolonged domestication, proso millet germplasm resources are particularly rich and diverse (Wang et al., 2016a). Globally, approximately 29,000 germplasm resources have been conserved, and high variability exists among them (Vetriventhan et al., 2019). In China, a total of 9,885 germplasm resources have been documented, with 8,515 collected over a 22-year span from 1982 to 2003 (Wang et al., 2005). Of these 8,515 resources, 98% originate from 23 provinces of China, while 2% are from 14 other countries (Wang et al., 2016a). Its short growth period, drought resistance, and tolerance to poor soils makes it a pioneer crop suitable for disaster relief, famine preparation, and yield enhancement (Zhao, 2011), and it thus plays a pivotal role in enhancing food security and fostering agricultural diversification (Rajput and Santra, 2016; Zhang et al., 2019; Yang et al., 2023; Khound et al., 2024). Recent studies have advanced from core germplasm construction (Upadhyaya et al., 2011) and phenotypic-genetic analysis (Rajput and Santra, 2016) to integrated multi-trait evaluation (Vetriventhan and Upadhyaya, 2018, 2019), successfully identifying elite germplasm groups with high-yield potential, large-grain characteristics, and multi-nutrient enrichment. However, these breakthroughs remain constrained by substantial geographical bias, fragmented trait analysis, and homogenized germplasm coverage, critically impeding the in-depth exploration of genome-wide adaptive mechanisms of proso millet.

Simple sequence repeat (SSR) markers can be used to directly detect DNA molecular structure variation (Weber, 1990), reflecting differences with high sensitivity, stability, and ease of use. It is thus commonly used in genetic diversity analysis (Badiane et al., 2012; Wang et al., 2012; Rajput et al., 2014; Krishna et al., 2018; He et al., 2024), germplasm identification, and DNA fingerprinting. This technology is minimally impacted by environmental factors, ensuring stable and reliable results (Grover and Sharma, 2016), and it has been widely applied to crops, including foxtail millet (Li

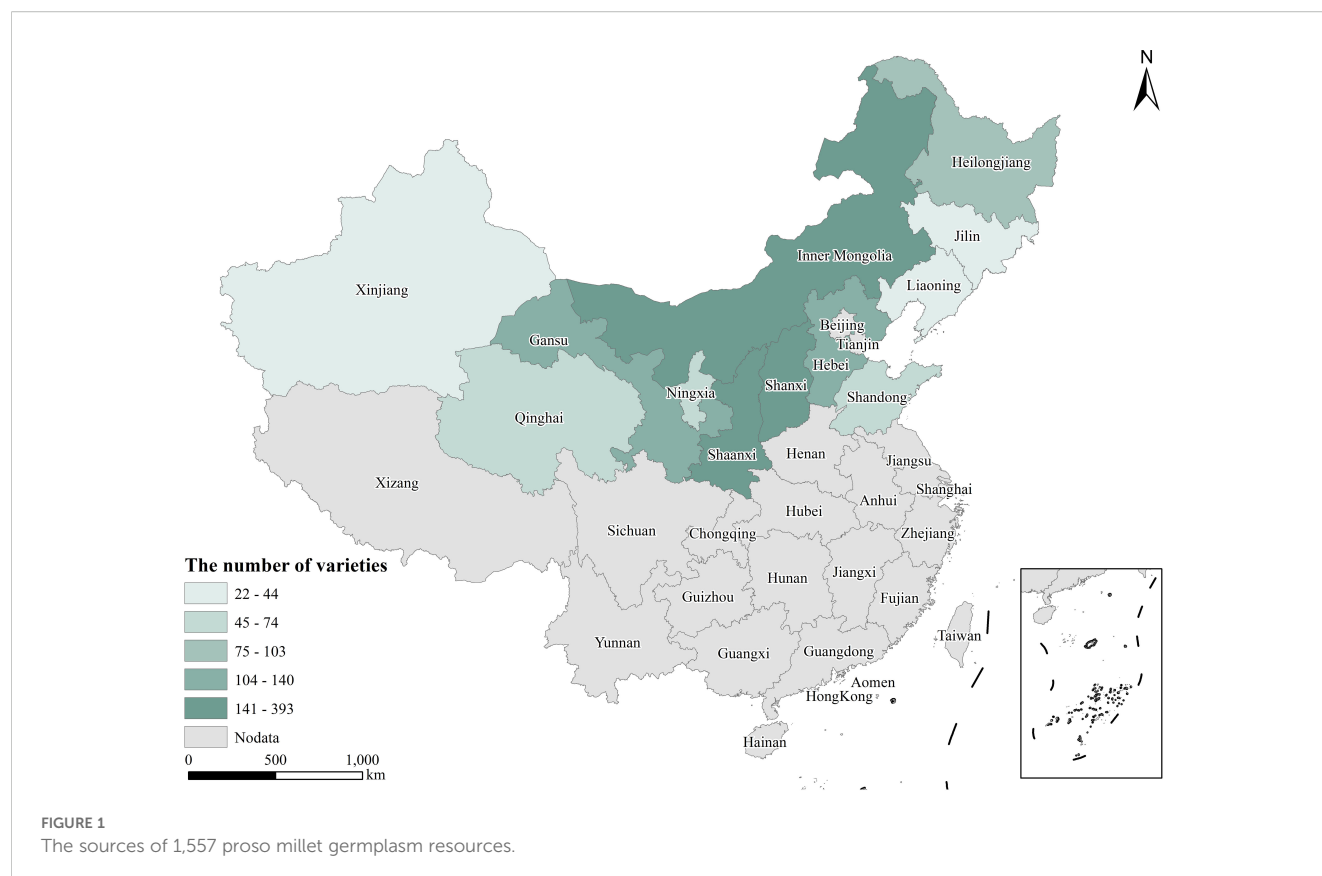
et al., 1998; Wang et al., 2012), maize (Lilia et al., 2008), soybean (Wen et al., 2009), sorghum (Li et al., 2010), and cowpea (Badiane et al., 2012). Rajput et al. (2014) analyzed 548 switchgrass SSR markers and identified 339 that could be applied to proso millet, with 254 exhibiting polymorphism. Hunt et al. (2011, 2018) conducted a genomic analysis of proso millet accessions collected from across the Eurasian continent, determining that genetic clustering aligned with their geographic origins and confirming the western Loess Plateau as the primary domestication center for proso millet. Liu et al. (2016) used 67 SSR markers to analyze 88 proso millet varieties, detecting 179 alleles and revealing rich genetic variation within ecological zones, along with complex intergroup relationships. DNA molecular identifiers (IDs) convert genetic differences into numerical codes, thereby assigning a distinct digital identifier to each variety (Tian et al., 2015). This method has been widely applied to sweet potatoes (Meng et al., 2018), potatoes (Duan et al., 2019), and mustard (Panigrahi et al., 2018) since its initial development. However, there have been no published studies reporting DNA molecular IDs for proso millet.

The present study initially utilized phenotypic traits to evaluate the diversity of 1,582 germplasms, comprising 1,516 landraces, 32 cultivars, 9 wild varieties and 25 foreign varieties. The 1,577 germplasms from China were specifically sourced from the following provinces: Inner Mongolia, Heilongjiang, Jilin, Liaoning, Hebei, Shandong, Gansu, Qinghai, Xinjiang, Shanxi, Shaanxi, and Ningxia. Based on the clustering analysis of phenotypic traits, the 1,582 germplasms were subdivided into five groups that were then ranked in descending order based on comprehensive F values and compared against qualitative traits. This approach resulted in the identification of 147 representative superior germplasms. These germplasms demonstrated outstanding overall phenotypic traits while representing a substantial proportion of the genetic diversity among all germplasms. Subsequently, genetic diversity analysis was performed on these 147 germplasms using 15 highly polymorphic SSR markers alongside the construction of their molecular IDs. This work provides an empirical foundation for the protection and utilization of proso millet germplasm.

## Materials and methods

### Material source

A total of 1,582 proso millet germplasm resources (including 1,516 landraces, 32 cultivars, 9 wild varieties from China, and 25 foreign varieties, with the accessions from China specifically sourced from Inner Mongolia, Heilongjiang, Jilin, Liaoning, Hebei, Shandong, Gansu, Qinghai, Xinjiang, Shanxi, Shaanxi, and Ningxia), were obtained from the Crop Science Institute of the Chinese Academy of Agricultural Sciences (Figure 1, Supplementary Tables S1, S2).



## Field phenotypic trait determination

### Field experiment design

The materials were sown in early June 2024 at the Yulin Minor Grain Comprehensive Experimental Demonstration Station, Shaanxi, China. Each variety was planted in three randomized plots (3 meters long  $\times$  2 meters wide) with 0.33-meter inter-plot spacing to ensure independence of samples. Soil fertility within the experimental field was uniform across all plots. Irrigation, fertilization, weeding, and other field management practices adhered to local large-scale farming protocols.

### Phenotypic trait determination

During the growth period, inflorescence color, inflorescence density, axis shape, branched spike length, projection on branch base, angle of divergence of branches, panicle type, and trichome, among other measurements, were observed and recorded. At maturity, measurements of stem height, main stem nodes, stem diameter, panicle length, tiller number, branching habit, panicle weight per plant, grain weight per plant, straw weight per plant, and shatter resistance were recorded. After harvest, thousand-grain weight, grain length, grain width, and grain length-to-width ratio were measured.

Inflorescence color was visually evaluated in each test plot, using the principle of maximum similarity for classification. Inflorescence density was assessed by observing the spikelet distribution on the branches. The straightness or curvature of the axis was evaluated visually according to similar criteria. Branch

spike length was measured by selecting three random plants from each plot; the length of the first branch on the main stem was measured with a ruler, and the average was used to determine the short, intermediate, or long type. Projection on branch base was examined for the presence of swollen joint-like structures and categorized according to presence, absence, or degree. The angle of divergence of branches was measured with a protractor on three randomly selected plants per plot, with the average determining the small, intermediate, or large type. Panicle type was qualitatively assessed based on morphology. Trichomes were observed to determine their length and density. Stem height was measured from the plant base to the panicle base on three randomly selected plants per plot, and these measurements were averaged. The number of main stem nodes above ground was visually counted for three plants per plot, and the average was then calculated. Stem diameter was measured using calipers at the maximum internode diameter of the basal main stem, with the average obtained from three plants. Panicle length was measured from the first branch node to the top of the plant and averaged across three plants per plot. Tiller number was counted for three randomly selected plants per plot and averaged. Branching habit was visually assessed for conditions between leaf axils on stem nodes. Panicle weight per plant was determined by cutting and weighing three panicles from the first branch and averaging their weights. Grain weight per plant was determined by threshing and weighing grain from three panicles, with the weights then averaged. Straw weight per plant was measured by removing three panicles, cutting off roots, and weighing the stems and leaves, with the weights then averaged.

Shatter resistance was assessed by allowing the panicle to drop freely three times, after which the shattering percentage was calculated, which was the basis for determining the shattering resistance type. Thousand-grain weight was calculated by cleaning, drying, and weighing two samples of 500 seeds with an error margin of 0.1 g. Grain length and width were measured using a seed analyzer, with values averaged from three measurements.

The assignment criteria for panicle type, inflorescence density, branched spike length, and shatter resistance were based on the methodology described by Zhang et al. (2019) and are detailed in Supplementary Table S3.

## DNA extraction and detection

Leaf samples were collected at the three-leaf stage, and plant DNA was extracted from them using a modified CTAB method. DNA quality was assessed by 1% agarose gel electrophoresis, while DNA concentration was determined using a nucleic acid detector (NanoDropOne).

## SSR primer design

SSR locus detection for the transcriptome unigenes of salt-tolerant varieties was performed using MISA software. The minimum repeat thresholds were set to 10, 6, 5, 5, and 5 for mono-, di-, tri-, tetra-, penta-, and hexa-nucleotide repeat sequences, respectively. The SSR sequences obtained were then utilized for primer design with Primer 3.0 software. Eighty SSR primers (Supplementary Table S4) were synthesized by Beijing TSINGKE Biological Co., Ltd.

## SSR primer population amplification and electrophoresis detection

Twenty varieties with significant phenotypic differences (Supplementary Table S5) were used for polymorphism screening. Fifteen primers (Supplementary Table S6) yielded clear amplification bands that revealed stable polymorphisms, making them suitable for subsequent genetic diversity analysis of 147 identified superior proso millet germplasm resources (Supplementary Tables S7, S8). PCR amplification was conducted using the EasyCycler PCR system, and PCR products were separated by 8% polyacrylamide gel electrophoresis, followed by silver staining for band detection.

## Statistical analysis

Microsoft Excel 2021 was employed to compile phenotypic data, while SPSS 27.0 and Origin 2025SR1 were employed to analyze and visualize phenotypic data. Electrophoresis results were manually scored to create a binary data matrix. Genetic

diversity parameters were calculated using PopGene 1.32 and PowerMarker 3.25, and cluster analyses were performed using MEGA 6.06. Population genetic structure and principal component analyses were conducted with Structure 2.34 and NTSYSpc 2.11. Based on SSR marker scoring results, data were sequentially encoded to construct DNA molecular IDs. QR code online technology (<https://cli.im/text>) facilitated the conversion of basic information into QR code DNA molecular IDs.

## Results

### Analysis of agronomic trait diversity in proso millet

The genetic diversity indices of 14 agronomic traits, 5 yield traits, and 3 grain traits were evaluated (Figure 2, Supplementary Table S9). The genetic diversity index was lowest for inflorescence color (0.15) and highest for stem diameter (2.10) among the 14 agronomic traits (comprising 10 qualitative and 4 quantitative traits). Notably, there were two traits with genetic diversity indices exceeding 2, both quantitative in nature. The differences in the frequency distribution of various qualitative traits are presented in Figure 2C. For the 1,582 germplasms, inflorescence color was predominantly green (96.5%), with purple accounting for merely 3.5%. The dense trichome type (59.5%) was the main trichome type. Regarding panicle type, the lateral type was predominant (52.6%), whereas scattered and compact types represented 23.6% and 23.8% of accessions, respectively. The main axis shape was primarily slightly bend (54.1%), with erect and bend types accounting for 23.3% and 22.6% of accessions, respectively. Inflorescence density was chiefly intermediate (41.7%) and slightly dense (37.4%). Most varieties lacked projections on branch bases (66.8%) and branching habits (72.6%), with minor angle deviations present in most varieties (60.6%). Short, intermediate, and long-branched spike length were evenly distributed, accounting for 32.7%, 32.7%, and 34.6% of accessions, respectively. Furthermore, the majority of varieties exhibited weak shatter resistance (44.4%), while strong shatter-resistant varieties were the least frequent (24.8%). As shown in Figure 2D, the average number of main stem nodes among the 1,582 germplasms was 6.34, with a variation coefficient of 16.8%. Stem height ranged from 33.67 to 236.67 cm, with an average of 129.57 cm, reflecting nearly a 7-fold disparity between the extrema, accompanied by a variation coefficient of 20.09%. The average stem diameter was 6.79 mm, with a variation coefficient of 22.12%. Straw weight per plant spanned from 1.53 to 110.10 g, averaging 32.15 g, with a 72-fold disparity between extrema, marked by the largest variation coefficient (60.03%).

### Analysis of yield trait diversity in proso millet

The genetic diversity indices of yield traits ranged from 1.95 to 2.08, with tiller number having the lowest index at 1.95, while grain

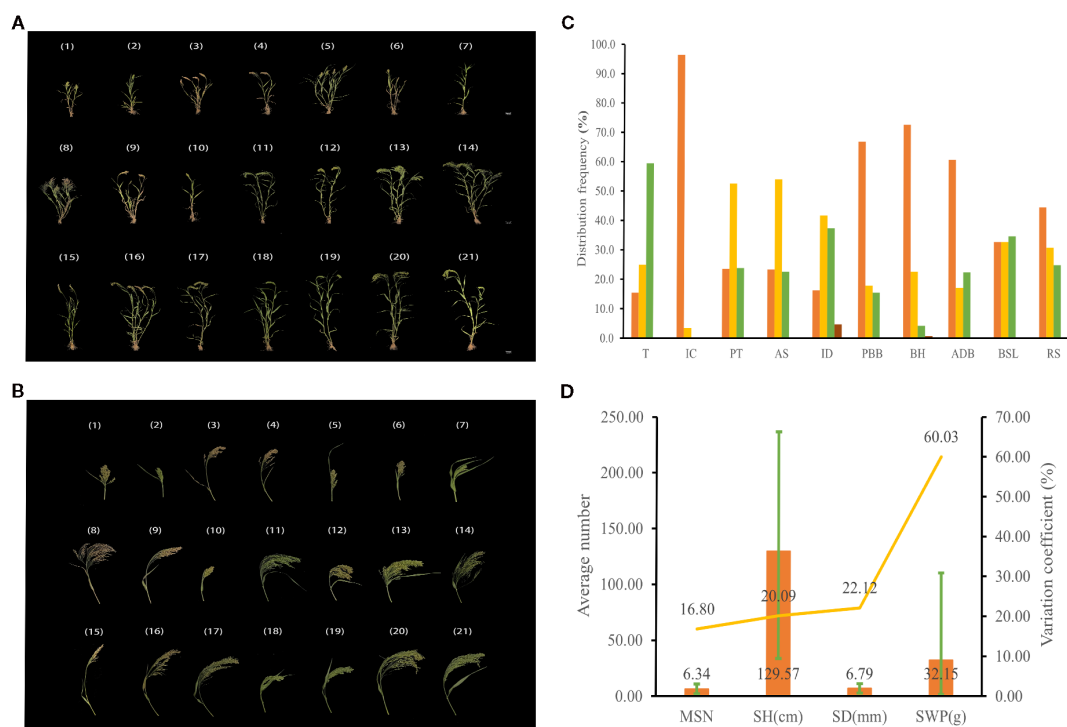


FIGURE 2

Agronomic trait diversity of proso millet. (a) Entire plant; (b) Panicle trait; (c, d) Agronomic trait. From 1 to 21 are Heidaetou, Shihuqianjingmi, Hongmi, Shuanglimizi, Linhedahuangmizi, Heitounianmizi, Zhouquhuangmizi, Helinxiaoqingmi, Dahuangmizi, Qianjingshu, Mizi, Hongmizi, Baimizi, Hongmizi, Gumi13, Bamenghongmizi, Linxidahongshu, Huangruanshu, Heimizi, Gancaohuangmizi, and Dalishu. These accessions display significant differences in agronomic traits and can effectively represent the diversity of germplasm resources. T, trichome; IC, inflorescence color; PT, panicle type; AS, axis shape; ID, inflorescence density; PBB, projection on branch base; BH, branching habit; ADB, angle of divergence of branches; BSL, branched spike length; RS, shatter resistance. Orange and yellow of inflorescence color denote green and purple; Orange, yellow, and green of panicle type correspond to scattered, lateral, and compact types, the sequence from orange to yellow to green (to brown) of remaining traits signifies a systematic enhancement in rank. MSN, main stem nodes; SH, stem height; SD, stem diameter; SWP, straw weight per plant. The bar chart illustrates the average values, the line graph depicts the coefficients of variation, and the error bars indicate the range.

weight per plant had the highest at 2.08 (Supplementary Table S9). The genetic diversity indices for panicle length, panicle weight per plant, and thousand-grain weight all exceeded 2. As shown in Figure 3, the average tiller number among 1,582 germplasms was 2.26, with a considerable variation coefficient (31.54%). The average panicle length was 37.05 cm, exhibiting a variation coefficient of 21.43%. Panicle weight per plant ranged between 0.88 and 28.83 g, averaging 9.97 g, with a 33-fold difference between extrema and a substantial variation coefficient (35.64%). Grain weight per plant varied from 0.61 to 19.11 g, averaging 7.20 g, displaying a 32-fold difference between extrema and a high variation coefficient (37.37%). The average thousand-grain weight was 7.09 g, with a variation coefficient of 14.94%. The above results collectively indicate there was substantial variation in yield traits among proso millet accessions, with considerable genetic diversity.

## Analysis of grain trait diversity in proso millet

The genetic diversity indices for grain length, grain width, and length-to-width ratio were 2.04, 2.02, and 2.00, respectively, indicating considerable genetic diversity in these grain traits

(Figure 4, Supplementary Table S9). As shown in Figure 4B, the average grain length and width among 1,582 accessions were 2.85 mm and 2.16 mm, respectively. Grain length exhibited the lowest variation coefficient at 5.22%, whereas grain width had a variation coefficient of 6.61%. The average grain length-to-width ratio was 1.33, with a variation coefficient of 5.48%. Genetic diversity indices for grain traits were higher than those of certain agronomic and yield traits, while their variation coefficients are comparatively lower than those of all agronomic and yield traits.

## Correlation analysis among traits of proso millet

Figure 5 presents the correlation analysis between four primary qualitative traits and eight representative quantitative traits. Among the qualitative traits, a highly significant positive correlation was observed between panicle type and inflorescence density ( $r = 0.50$ ), while a highly significant negative correlation was observed between panicle type and branched spike length ( $r = -0.24$ ). No correlations were detected among other qualitative traits. In examining the quantitative traits, the tiller number demonstrated a positive correlation with the main stem nodes ( $r = 0.07$ ) and a negative



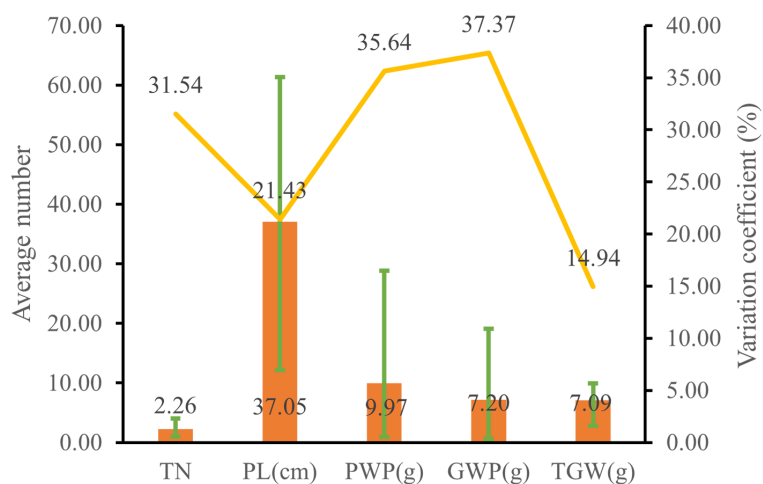


FIGURE 3

Yield trait diversity of proso millet. TN, tiller number; PL, panicle length; PWP, panicle weight per plant; GWP, grain weight per plant; TGW, thousand-grain weight. The bar chart illustrates the average values, the line graph depicts the coefficients of variation, and the error bars indicate the range.

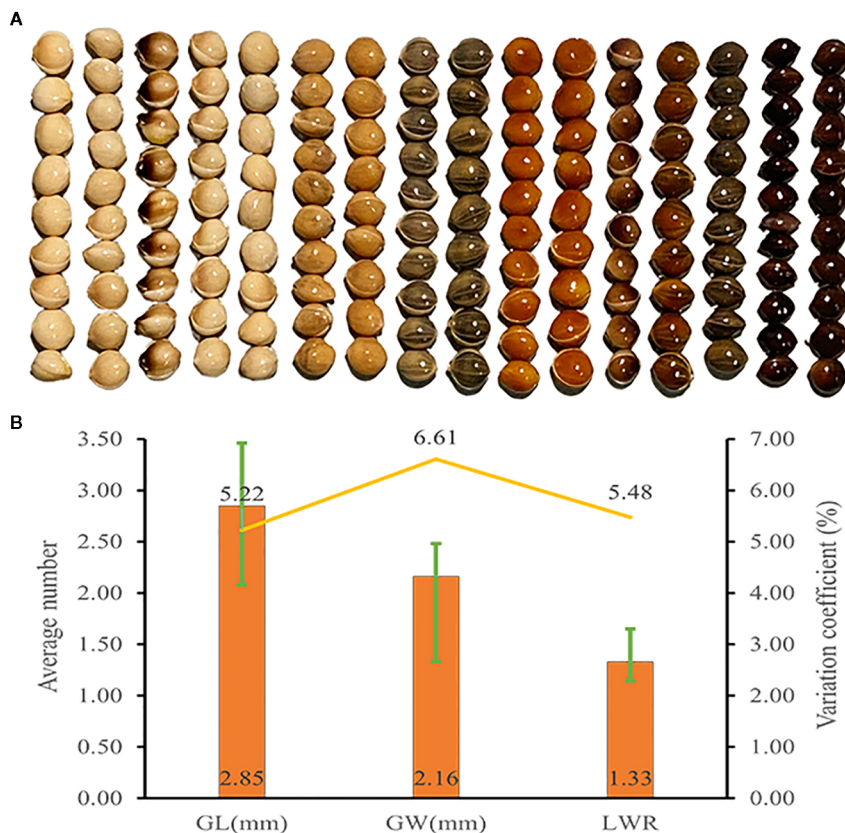


FIGURE 4

Grain trait diversity of proso millet. (A) Representative grain phenotype; (B) Grain trait. From left to right are Bairuanmi, Ziganbai, Hualianmi, Jiaojuruanmi, Zaershu, Gumi13, Huangmi, Jingyuanxiaqing, Fengshouma, Mizi, Hongshuzi, Xiaohuashu, Qingshuihexiangsemizi, Sezhaertuo, Heitounianmizi, Heimizi. These accessions display significant differences in grain traits and can effectively represent the diversity of germplasm resources. GL, grain length; GW, grain width; LWR, grain length to width ratio. The bar chart illustrates the average values, the line graph depicts the coefficients of variation, and the error bars indicate the range.

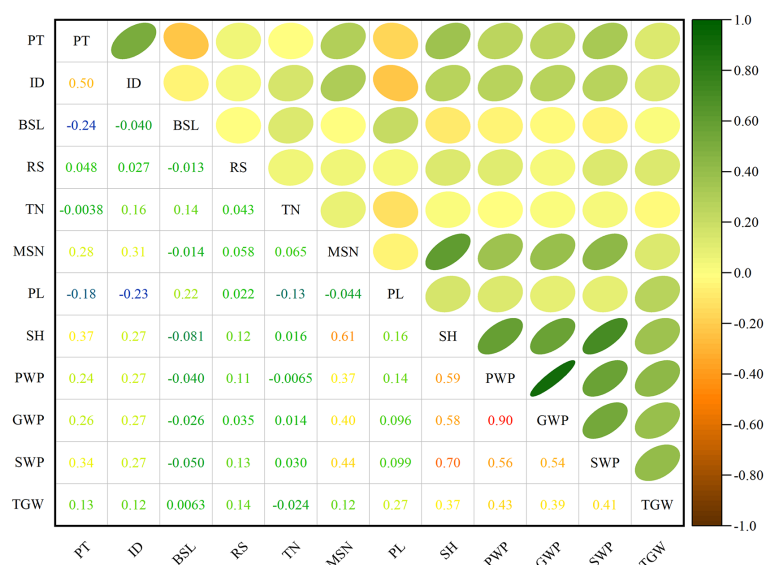


FIGURE 5

Correlation coefficients of various traits. PT, panicle type; ID, inflorescence density; BSL, branched spike length; RS, shatter resistance; TN, tiller number; MSN, main stem nodes; PL, panicle length; SH, stem height; PWP, panicle weight per plant; GWP, grain weight per plant; SWP, straw weight per plant; TGW, thousand-grain weight.

correlation with panicle length ( $r = -0.13$ ), yet showed no significant correlation with other traits. The number of main stem nodes was not correlated with panicle length but was highly significantly and positively associated with stem height ( $r = 0.61$ ), panicle weight per plant ( $r = 0.37$ ), grain weight per plant ( $r = 0.40$ ), straw weight per plant ( $r = 0.44$ ), and thousand-grain weight ( $r = 0.12$ ). Panicle length was not correlated with the number of main stem nodes, was negatively correlated with the tiller number ( $r = -0.13$ ), and was positively correlated with the stem height ( $r = 0.16$ ), panicle weight per plant ( $r = 0.14$ ), grain weight per plant ( $r = 0.10$ ), straw weight per plant ( $r = 0.10$ ), and thousand-grain weight ( $r = 0.27$ ). Stem

height, panicle weight per plant, grain weight per plant, straw weight per plant, and thousand-grain weight exhibited no correlations with tiller number. Stem height was highly significantly and positively associated with panicle weight per plant ( $r = 0.59$ ), grain weight per plant ( $r = 0.58$ ), straw weight per plant ( $r = 0.70$ ), and thousand-grain weight ( $r = 0.37$ ). Panicle weight per plant was highly significantly and positively associated with grain weight per plant ( $r = 0.91$ ), straw weight per plant ( $r = 0.56$ ), and thousand-grain weight ( $r = 0.43$ ). Grain weight per plant was highly significantly and positively associated with straw weight per plant ( $r = 0.54$ ) and thousand-grain weight ( $r = 0.39$ ). Straw

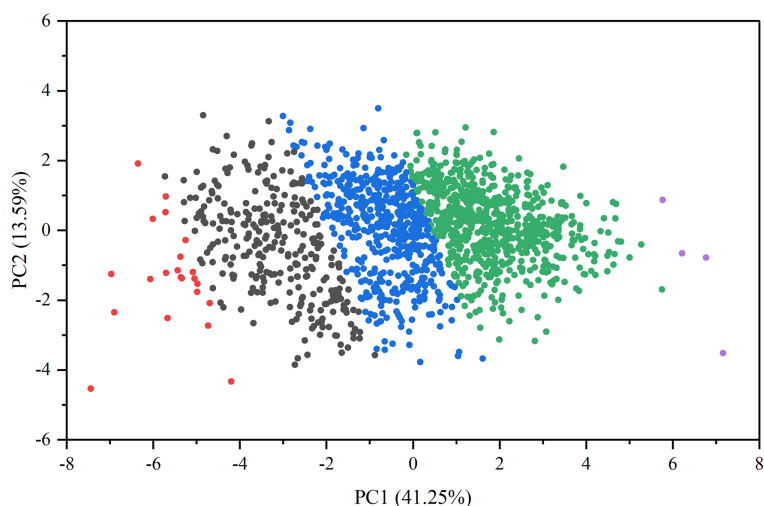


FIGURE 6

Principal component analysis of relative values of traits. The groups are color-coded as follows: red (Group 1), black (Group 2), blue (Group 3), green (Group 4), and purple (Group 5).

weight per plant was highly significantly and positively associated with thousand-grain weight ( $r = 0.41$ ). Among the relationships between qualitative and quantitative traits, panicle type was not correlated with the tiller number, but was negatively correlated with panicle length ( $r = -0.18$ ) and positively correlated with the number of main stem nodes ( $r = 0.28$ ), stem height ( $r = 0.37$ ), panicle weight per plant ( $r = 0.24$ ), grain weight per plant ( $r = 0.26$ ), straw weight per plant ( $r = 0.34$ ), and thousand-grain weight ( $r = 0.13$ ). Inflorescence density was negatively correlated with panicle length ( $r = -0.23$ ), and positively correlated with tiller number ( $r = 0.16$ ), the number of main stem nodes ( $r = 0.31$ ), stem height ( $r = 0.27$ ), panicle weight per plant ( $r = 0.27$ ), grain weight per plant ( $r = 0.27$ ), straw weight per plant ( $r = 0.27$ ), and thousand-grain weight ( $r = 0.13$ ). Branched spike length was highly significantly positively correlated with the tiller number ( $r = 0.16$ ) and panicle length ( $r = 0.22$ ) and highly significantly negatively correlated with stem height ( $r = -0.08$ ), while also exhibiting a significant negative correlation with straw weight per plant ( $r = -0.05$ ); however, it showed no significant correlations with the remaining traits. Shatter resistance exhibited no correlation the tiller number, panicle length, or grain weight per plant, yet was highly significantly positively correlated with the number of main stem nodes ( $r = 0.06$ ), stem height ( $r = 0.12$ ), panicle weight per plant ( $r = 0.11$ ), straw weight per plant ( $r = 0.13$ ), and thousand-grain weight ( $r = 0.14$ ).

## Principal component analysis and cluster analysis

Principal component analysis was employed to standardize the relative values of various traits, effectively eliminating factors with minimal impact but significant interference, to facilitate a comprehensive evaluation of proso millet germplasms (Supplementary Table S10). The first principal component accounted for 41.25% of the variance, the second for 13.59%, the third for 9.66%, and the fourth for 9.27%, with these four components collectively explaining 73.78% of the variance. The eigenvalues of these four principal components each exceeded 1, confirming their suitability for a comprehensive evaluation of proso millet germplasm. The highest loading values for the first principal component were associated with stem height, panicle weight per plant, and grain weight per plant. The second principal component was most positively correlated with thousand-grain weight, indicating that this trait heavily contributed to the component. The third principal component displayed the highest positive loadings for grain length and grain length-to-width ratio, suggesting it mainly reflects the overall condition of the grain. The fourth principal component exhibited higher absolute values for the tiller number and panicle length, showing its dependency on these traits.

Based on the score coefficient matrix of the four principal components, the comprehensive score ( $F$ ) for each variety was calculated to rank proso millet germplasms by excellence (Supplementary Table S11). A higher  $F$  value indicates a better overall phenotype for that variety. The top 10 performing varieties

were Hongruanmi, Hongruanmi, Laohuangmi, Getami, Ruantiaozhoumi, Zhuyeqing, Limashu, Mizi, Nuomizi, and Hongnianmi. Among these, Ruantiaozhoumi and Limashu are from Shanxi province; the remaining eight high-performing varieties are all from Shaanxi province.

Cluster analysis of the 1,582 germplasms identified five distinct groups (Figure 6). The first group includes 22 varieties, including Etou and Sansui, mainly originating from Heilongjiang, accounting for 1.4% of accessions. The second group comprises 292 varieties, including Zaohongmi and Anchunwei, primarily sourced from Heilongjiang and Inner Mongolia, along with most foreign varieties, representing 18.5% of the germplasms. The third group consists of 528 varieties, including Daqihuangmizi and Taolibai, largely obtained from Shanxi and Inner Mongolia, comprising 33.4% of accessions. The fourth group contains 736 varieties, including Ruantiaozhoumi and Ziganmi (shu), predominantly sourced from Shanxi and Shaanxi, followed by Gansu and Inner Mongolia, constituting 46.4% of accessions. The fifth group consists of only four varieties, namely Hongruanmi, Hongruanmi, Getami, and Laohuangmi, all originating from Shaanxi, accounting for just 0.3% of accessions. Based on the clustering analysis of phenotypic traits, the germplasms within the five groups were ranked in descending order of comprehensive  $F$  values and compared against qualitative traits. This approach resulted in the identification of 147 representative superior germplasms, which were systematically selected for subsequent genetic diversity analysis.

## Genetic diversity analysis of proso millet

### Polymorphic SSR primer selection

Twenty varieties with significant differences in phenotypic traits were subjected to polymorphism screening using 80 pairs of SSR primers. All primers successfully amplified bands, resulting in an effective amplification rate of 100%. Among these primers, 15 pairs demonstrated polymorphism, with a development efficiency of 18.75%. Ultimately, these 15 primer pairs were subsequently utilized for genetic diversity analysis of the 147 superior germplasms.

### Genetic diversity analysis using SSR markers

Utilizing 15 pairs of polymorphic SSR primers, the genetic diversity of 147 proso millet germplasms was analyzed (Table 1). Across these 15 loci, a total of 30 observed alleles were identified, with each locus exhibiting two alleles, averaging 2.00. The effective allelic number range was between 1.03 (BLF21) and 1.96 (BLF79), with an average of 1.28. The Shannon diversity index ranged from 0.07 (BLF21) to 0.68 (BLF79), with an average of 0.33. Observed heterozygosity varied from 0 (BLF4, BLF47, BLF8, BLF9, BLF51, BLF52, BLF58, BLF59, BLF61, BLF27, BLF80) to 0.73 (BLF79), with an average of 0.06. Expected heterozygosity ranged from 0.03 (BLF21) to 0.49 (BLF79), averaging 0.19. Nei's expected heterozygosity ranged from 0.03 (BLF21) to 0.49 (BLF79), averaging 0.19. The polymorphism information content spanned 0.17 (BLF51) to 0.59 (BLF79), averaging 0.39.



TABLE 1 Genetic parameters from the 15 polymorphic SSR markers used in the study.

Primer	Na	Ne	I	Ho	He	Nei	PIC
BLF-41	2	1.3478	0.4265	0.0145	0.2590	0.2580	0.3116
BLF-4	2	1.1912	0.2979	0.0000	0.1612	0.1605	0.3367
BLF-47	2	1.2800	0.3768	0.0000	0.2199	0.2188	0.4619
BLF-8	2	1.3081	0.3983	0.0000	0.2373	0.2355	0.4476
BLF-9	2	1.1133	0.2095	0.0000	0.1023	0.1017	0.4077
BLF-51	2	1.0606	0.1327	0.0000	0.0573	0.0571	0.1771
BLF-52	2	1.4764	0.5035	0.0000	0.3245	0.3227	0.5118
BLF-58	2	1.0533	0.1205	0.0000	0.0509	0.0506	0.3945
BLF-59	2	1.1492	0.2530	0.0000	0.1303	0.1298	0.2898
BLF-61	2	1.1995	0.3061	0.0000	0.1669	0.1663	0.2071
BLF-21	2	1.0267	0.0701	0.0263	0.0261	0.0260	0.3046
BLF-66	2	1.1730	0.2791	0.1603	0.1480	0.1475	0.2882
BLF-27	2	1.7676	0.6259	0.0000	0.4366	0.4342	0.5613
BLF-79	2	1.9625	0.6836	0.7340	0.4931	0.4904	0.5874
BLF-80	2	1.1150	0.2117	0.0000	0.1041	0.1031	0.3843
Mean	2	1.2816	0.3263	0.0623	0.1945	0.1935	0.3781
St.Dev	0	0.2685	0.1788	0.1903	0.1373	0.1366	

Na, number of alleles; Ne, number of effective alleles; I, Shannon's diversity index; Ho, observed heterozygosity; He, expected heterozygosity; Nei, Nei's expected heterozygosity; PIC, polymorphism information content.

Genetic diversity and similarity analysis from different regions

The analysis of genetic diversity parameters of proso millet germplasms across different regions (Table 2) revealed observed allelic variability from 1.33 to 1.87, highest in the Northwest China region and lowest in the Inner Mongolia region. Effective allelic variability ranged from 1.20 to 1.32, and was highest in the Loess Plateau region and lowest in the Northeast region. The highest observed heterozygosity was recorded in the Northeast region (0.08), and the lowest value was observed among germplasms from abroad (0.01). The highest expected heterozygosity was in the Loess Plateau region (0.20), and the lowest was in the Inner Mongolia region (0.13). Nei's expected heterozygosity was highest

TABLE 2 Parameters of genetic diversity in six populations.

Population	Na	Ne	I	Ho	He	Nei	PIC
IMR	1.3333±0.4880	1.2138±0.3517	0.1849±0.2801	0.0667±0.2582	0.1325±0.2050	0.1245±0.1930	0.2877
NER	1.4000±0.5071	1.1961±0.2944	0.1948±0.2578	0.0794±0.2317	0.1365±0.1877	0.1253±0.1715	0.3652
NR	1.6667±0.4880	1.2888±0.3386	0.2850±0.2570	0.0672±0.1742	0.1859±0.1857	0.1808±0.1800	0.3471
NWR	1.8667±0.3519	1.2727±0.2502	0.3131±0.2014	0.0539±0.1687	0.1915±0.1441	0.1887±0.1420	0.3371
LPR	1.7333±0.4577	1.3177±0.3627	0.3025±0.2621	0.0713±0.2271	0.1984±0.1942	0.1932±0.1890	0.3573
Abroad	1.3571±0.4972	1.2816±0.4323	0.2158±0.3129	0.0143±0.0535	0.1789±0.2677	0.1503±0.2226	0.2345

IMR, Inner Mongolia region; NER, Northeast region; NR, Northern region; NWR, Northwest region; LPR, Loess Plateau region. The same as below.

TABLE 3 Nei's genetic identity and genetic distance from different populations.

Population	IMR	NER	NR	NWR	LPR	Abroad
IMR	-	0.9776	0.9590	0.9865	0.9636	0.8204
NER	0.0227	-	0.9617	0.9901	0.9853	0.8138
NR	0.0418	0.0391	-	0.9781	0.9751	0.8218
NWR	0.0136	0.0099	0.0221	-	0.9866	0.8123
LPR	0.0371	0.0148	0.0252	0.0135	-	0.8529
Abroad	0.1980	0.2060	0.1963	0.2079	0.1591	-

Nei's genetic identity (above diagonal) and genetic distance (below diagonal).

in the Loess Plateau region (0.19) and lowest in the Inner Mongolia region (0.12). In terms of Shannon diversity index and polymorphism information content, varieties from Northwest China and Loess Plateau region exhibited higher values, while foreign varieties showed the lowest value, indicating greater genetic diversity in the Northwest and Loess Plateau region.

Analysis of genetic distance and genetic consistency parameters of proso millet germplasms from different regions (Table 3) showed that genetic distance ranged from 0.01 to 0.21, averaging 0.08, and genetic consistency ranged from 0.81 to 0.99, averaging 0.92. The genetic distance was greatest between Northwest region and foreign accessions (0.21), with the lowest genetic consistency (0.81), while accession from the Northeast and the Northwest showed the closest genetic distance (0.01) and highest genetic consistency (0.99). These findings indicate that greater geographical separation is associated with greater genetic distance, lower genetic consistency, and more distant kinship; conversely, closer geographical proximity is correlated with closer genetic distance, higher genetic consistency, and nearer kinship.

Cluster analysis based on UPGMA

Cluster analysis of 147 proso millet germplasms was conducted by unweighted pair group method with arithmetic mean analysis (UPGMA), which indicated that the germplasms could be divided into four groups (Figure 7A, Table 4). Group 1 encompasses one variety, sourced from Heilongjiang, representative of the Northeast region. Group 2 includes one variety, sourced from Shandong,

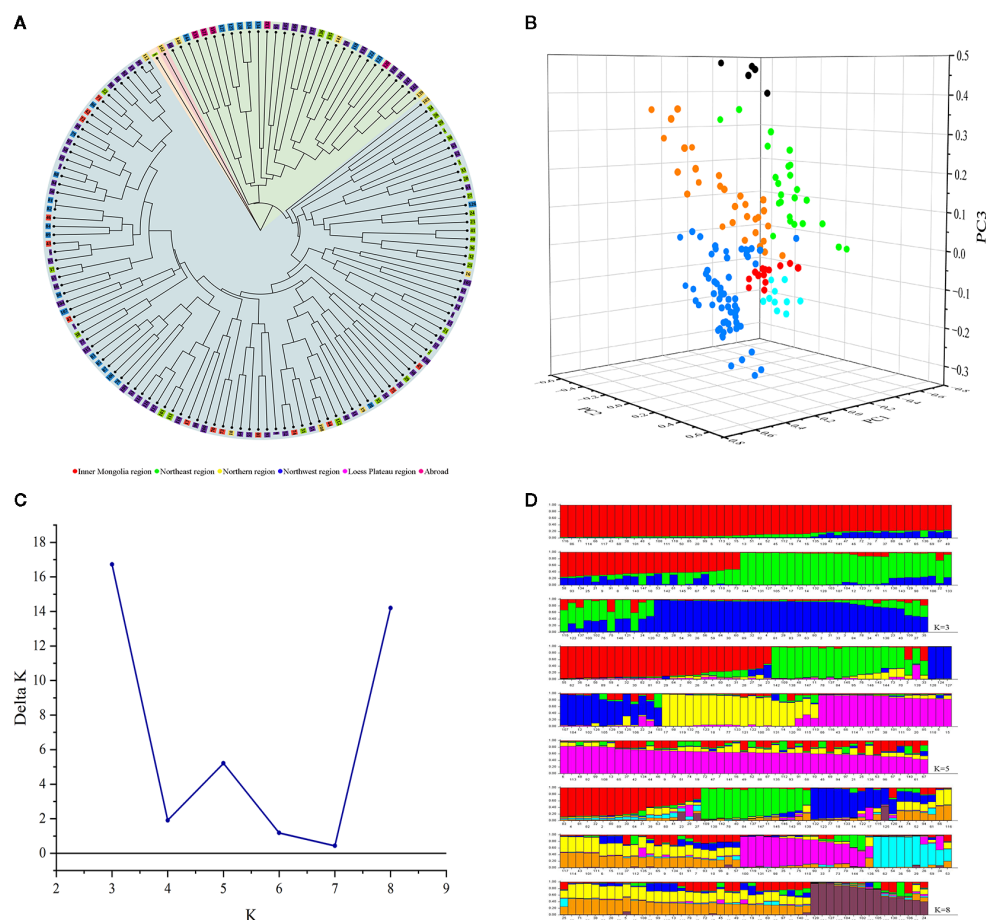


FIGURE 7

(A) Cluster analysis of 147 proso millet germplasms based on UPGMA. Numbers represent serial number of accessions. (B) Principal component analysis of proso millet germplasms. The groups are color-coded as follows: red (Inner Mongolia region), yellow (Northeast region), aqua (Northern region), green (Northwest region), blue (Loess Plateau region) and black (Abroad). (C) Population modeling of 147 proso millet germplasms based on structure. (D) Population genetic structure graph of 147 proso millet germplasms based on structure. Numbers in the horizontal axis represent serial number of accessions.

representative of the Northern region. Group 3 comprises 31 varieties with 8 from Shanxi, 6 from Qinghai, 3 from Gansu, 3 from Ningxia, 3 from Hebei, 2 from Heilongjiang, 1 from Shaanxi, and 5 from abroad, mainly from the Loess Plateau and Northwest regions. Group 4 contains 114 varieties, with contributions from Shanxi (22), Shaanxi (22), Heilongjiang (19), Inner Mongolia (14), Gansu (9), Jilin (7), Ningxia (7), Xinjiang (5), Hebei (4), Qinghai (2), Shandong (2), and Liaoning (1), primarily from the Loess Plateau, Northeast, and Northwest regions. Additionally, Figure 7A and Table 4 collectively illustrate that all foreign varieties are consolidated within Group 3.

## Principal component analysis of genetic differences

Three-dimensional principal component analysis (PCA) was conducted on the 147 proso millet superior germplasms, as shown in Figure 7B. The first three principal components, denoted as PC1, PC2, and PC3, accounted for 21.37%, 18.87%, and 17.39% of the total

variance, respectively, with a cumulative explanation of variance reaching 57.63%. The 147 germplasms were thus classified into six groups: 14 from the Inner Mongolia region, 30 from the Northeast region, 10 from the Northern region, 25 from the Northwest region, 63 from the Loess Plateau region, and 5 from abroad, with each group corresponding to their geographical sources.

## Population genetic structure analysis based on structure

The genetic structure of 147 proso millet accessions was analyzed using Structure. Peaks in the number of genetic populations ( $K$ ) were identified at  $K = 3$ ,  $K = 5$ , and  $K = 8$  (Figure 7C), prompting analysis under both models. For  $K = 3$ , the germplasms were categorized into three groups as illustrated in Figure 7D. Group 1, marked in red in Figure 7D (73 varieties), predominantly sourced from the Loess Plateau, and Inner Mongolia regions. This includes varieties from Shanxi (19), Shaanxi (13), and Inner Mongolia (12). Group 2, marked in green in Figure 7D (39

**TABLE 4** Group distribution of 147 proso millet germplasms based on UPGMA.

Province	Group1	Group2	Group3	Group4
Inner Mongolia				14
Heilongjiang	1		2	19
Jilin				7
Liaoning				1
Hebei			3	4
Shandong		1		2
Gansu			3	9
Qinghai			6	2
Xinjiang				5
Shanxi			8	22
Shaanxi			1	22
Ningxia			3	7
Abroad			5	
Total	1	1	31	114

varieties), primarily consists of varieties mainly originating from the Loess Plateau, Northwest regions, and abroad (5), including varieties specifically from Qinghai (7), Shanxi (6), Shaanxi (6), and abroad (5). Group 3, marked in blue in [Figure 7D](#) (35 varieties), encompasses varieties mainly from the Loess Plateau, and Northeast regions, including Shanxi (5), Shaanxi (4), and Heilongjiang (14).

For  $K = 5$ , the germplasms were divided into five groups. Group 1, marked in red in [Figure 7D](#) (27 varieties), consists of varieties predominantly from the Northeast, and Loess Plateau regions, comprising Heilongjiang (11), Shanxi (5), and Shaanxi (3). Group 2, marked in green in [Figure 7D](#) (20 varieties), includes varieties mostly sourced from the Loess Plateau region, featuring Shanxi (6), Ningxia (3), and Shaanxi (2). Group 3, marked in blue in [Figure 7D](#) (16 varieties), encompasses varieties primarily from the Northwest region, specifically Qinghai (5). Group 4, marked in yellow in [Figure 7D](#) (20 varieties), primarily comprises varieties mainly from the Loess Plateau region, and abroad, including Shaanxi (4), Ningxia (3), and abroad (5). Group 5, marked in purple in [Figure 7D](#) (64 varieties), contains varieties primarily from the Loess Plateau, and Inner Mongolia regions, comprising Shanxi (15), Shaanxi (12), and Inner Mongolia (10).

For  $K = 8$ , the germplasms were categorized into eight groups as illustrated in [Figure 7D](#). Group 1, marked in red in [Figure 7D](#) (18 varieties), predominantly sourced from the Northeast, and Northwest regions. This includes varieties from Heilongjiang (8), and Gansu (4). Group 2, marked in green in [Figure 7D](#) (14 varieties), primarily consists of varieties mainly originating from the Loess Plateau region, including varieties specifically from Shanxi (6), Shaanxi (2), and Ningxia (2). Group 3, marked in blue in [Figure 7D](#) (16 varieties), encompasses varieties mainly from the Loess Plateau, Inner Mongolia, and Northwest regions, including

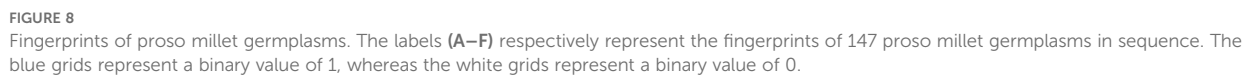
Shanxi (3), Ningxia (3), Qinghai (3), and Inner Mongolia (3). Group 4, marked in yellow in [Figure 7D](#) (25 varieties), mainly comprises varieties from the Loess Plateau, and Northeast regions, with varieties specifically from Shanxi (6), Shaanxi (6), Ningxia (3), Heilongjiang (3), and Jilin (3). Group 5, marked in purple in [Figure 7D](#) (17 varieties), contains varieties primarily from the Loess Plateau, and Northeast regions, comprising Shaanxi (3), and Heilongjiang (5). Group 6, marked in aqua in [Figure 7D](#) (11 varieties), consists of varieties predominantly from the Northeast, and Loess Plateau regions, comprising Heilongjiang (4), and Shanxi (5). Group 7, marked in golden in [Figure 7D](#) (31 varieties), includes varieties mostly sourced from the Loess Plateau, Northwest, and Inner Mongolia regions, featuring Shanxi (8), Xiangjiang (4), Shaanxi (3), and Inner Mongolia (6). Finally, Group 8, marked in crimson in [Figure 7D](#) (15 varieties), primarily includes varieties from the Loess Plateau, and Northwest regions, specifically Qinghai (4), Shanxi (3), and Shaanxi (2).

## Construction of DNA fingerprints and molecular IDs

To facilitate the conservation of the genetic resources of proso millet varieties, fingerprint maps were constructed for all 147 evaluated superior varieties on the basis of the SSR marker electrophoresis results ([Figure 8](#)). The electrophoresis results were digitally encoded, and following the principle of distinguishing the maximum number of varieties with the minimum number of primers, all amplification outcomes of the assayed varieties were combined in the following order: BLF41, BLF4, BLF47, BLF8, BLF9, BLF52, BLF58, BLF21, BLF66, BLF-27, BLF-79, and BLF80. This procedure produced a distinct 24-character string for each variety ([Supplementary Table S12](#)). These strings of DNA molecular IDs were then augmented with additional information such as Name, Unicode, Source, and Classification to create QR code DNA molecular IDs using online QR code generation technology. QR code DNA molecular IDs for varieties 1, 2, 3, and 4 are presented in [Figure 9](#), while those for the remaining 143 varieties are detailed in [Supplementary Figure S1](#).

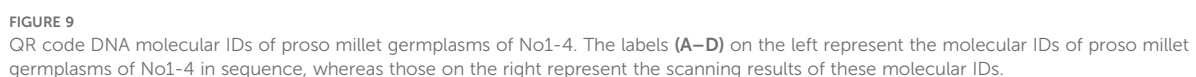
## Discussion

In proso millet, morpho-agronomic traits, which are important for its genetic improvement, include plant height, panicle length, thousand-grain weight, and grain weight per plant ([Rajput et al., 2013; 2024](#)). In several cereal crops, many of such traits demonstrated high genetic variability and were used in breeding programs for genetic improvement ([Gaju et al., 2009; Upadhyaya et al., 2011; Pucher et al., 2015](#)). In the present study, the variation coefficients for these various traits ranged from 5.22 to 60.03% ([Figures 2, 3, 4](#)), and the diversity index for these traits ranged from 0.15 to 2.10 ([Supplementary Table S9](#)), with larger values noted for panicle length, stem height, stem diameter, panicle weight per plant, grain weight per plant, and thousand-grain weight, all exceeding 2, in agreement with the findings of [Zhang et al. \(2019\)](#). The variation



Correlation analysis between various morpho-agronomic traits is necessary for direct and indirect selection and genetic improvement of crops, which may help breeders to identify traits that are quickly and easily measurable, making them very useful in

selections (Rajput et al., 2024). The present study revealed significant or highly significant positive and negative correlations between multiple traits (Figure 5). As a physiological process, plant height is highly correlated with the number of nodes; an increase in the number of nodes corresponds to an increase in plant height (Figure 5), aligning with the findings of Zhang et al. (2019) and Rajput et al. (2024). Notably, this study found no correlation between the tiller number and grain weight per plant or





thousand-grain weight (Figure 5), diverging from the findings of Guo et al. (2025); this discrepancy is likely attributable to environmental factors, as this present research was confined to a single year and location. Principal component analysis reduced 12 indicators into four principal components, with a cumulative contribution rate of 73.78% (Supplementary Table S10), effectively describing trait variation. Major traits contributing to this variation include tiller number, panicle length, and thousand-grain weight, consistent with the findings of Yazdizadeh et al. (2020). Clustering by principal component analysis involves the decomposition of phenotypic data into ordered, unrelated, orthogonal principal components based on the Euclidean distance matrix (Odong et al., 2013). Simply put, PCA-based clustering improves the extraction of genetically relevant information from many quantitative variables to a limited number of variables (Rajput et al., 2024). In the present study, cluster analysis divided the 1,582 germplasms into five categories, finding that proso millet from each of the different ecological zones may be grouped together (Figure 6).

It is inevitable that the present study still has certain limitations, as it relies on data from a single year and location and thus may not fully account for environmental variability or long-term trends. Nonetheless, our findings are consistent with previous studies (Zhang et al., 2019; Rajput et al., 2024; Yazdizadeh et al., 2020), which effectively confirms the reliability of the present results. Future investigations incorporating diverse temporal and spatial scales will improve the study's applicability and relevance.

The Shannon diversity index ( $I$ ) and polymorphism information content (PIC) comprehensively reflect population genetic diversity levels (Liu and Muse, 2005). Generally, primers with PIC values less than 0.25 indicate low polymorphism, PIC values greater than or equal to 0.25 and less than 0.5 indicate moderate polymorphism, and PIC values greater than or equal to 0.5 indicate high polymorphism (Sharma et al., 2021). Most primers used in this study demonstrated moderate to high polymorphism (Table 1), making them suitable for genetic diversity analysis. This study assessed genetic differences in 147 proso millet germplasms, finding Shannon diversity indices ranging from 0.07 to 0.68, with an average of 0.33 (Table 1), which was lower than those reported by Yazdizadeh et al. (2020); Liu et al. (2017, 2016), Hunt et al. (2011), and Cho et al. (2010). PIC values in this study ranged from 0.17 to 0.59, averaging 0.38 (Table 1), higher than those reported by Hunt et al. (2011); Cho et al. (2010), and Liu et al. (2016), but lower than those reported by Hu et al. (2009), and Yazdizadeh et al. (2020), possibly owing to differences in experimental materials or SSR marker types used. With respect to accession source, the Shannon diversity index and PIC values for varieties from the Northwest, and Loess Plateau regions were higher than those from other regions (Table 2), indicating greater genetic diversity in these areas, consistent with the findings of Hu et al. (2009). Furthermore, the genetic distance between varieties from the Northeast and Northwest regions was closest, with the highest genetic

consistency (Table 3), consistent with a positive correlation between geographic origin and genetic distance (Hu et al., 2009), contrasting with the findings of Hou et al. (2017); this discrepancy may arise from the differences in marker types used or the specific properties of the markers selected.

Cluster analysis based on UPGMA is an effective method for evaluating the genetic diversity and background of crop germplasm resources, which are essential for the effective conservation and utilization of these resources (Rajput et al., 2024). The cluster analysis performed in this study revealed that all five foreign varieties were classified into the third group (Figure 7A, and Table 4), suggesting that the grouping may be associated with geographical origin. However, occasionally exceptions can occur in groupings based on UPGMA, as has been observed across major cereal crops, and principal component analysis can provide a reasonable explanation for this discrepancy (Van Inghelandt et al., 2010; Courtois et al., 2012; Bonman et al., 2015). In the present study, UPGMA clustering categorized varieties from Northeast China into three groups (Figure 7A), whereas principal component analysis classified them into a single group (Figure 7B). The individual cases appearing in the cluster analysis groups are all classified into appropriate groups in PCA, reasonably explaining the aforementioned classification issues.

## Conclusion

Proso millet germplasm resources exhibit substantial phenotypic and genetic variation, with the greatest variation coefficients in yield traits. A total of 10 high-performing germplasms were identified, which can meet the needs of germplasm resource development and genetic breeding. Additionally, 15 polymorphic markers were developed, enabling the precise evaluation of 147 proso millet accessions from both regions within China and foreign sources. Proso millet germplasm resources originating from China were observed to have abundant genetic diversity, generally higher than those from abroad. Meanwhile, efforts must be intensified to collect wild varieties and introduce foreign varieties to enhance the genetic diversity among existing resources, thereby promoting research on the genetic evolution and improvement of proso millet. In addition, this study will be followed by long-term research across various locations to ensure the validity and reliability of the present results.

## Supporting information

The following supporting information is available online: Supplementary Tables S1–S12, Supplementary Data S1–S2, and Supplementary Figures S1–S4.



## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/[Supplementary Material](#).

## Author contributions

ZL: Formal Analysis, Methodology, Writing – original draft, Visualization. YXY: Software, Investigation, Writing – original draft, Data curation. HH: Writing – original draft, Data curation, Investigation, Project administration. SY: Investigation, Writing – original draft, Data curation. ZC: Investigation, Data curation, Writing – original draft. XZ: Writing – original draft, Investigation, Data curation. JL: Writing – review & editing, Validation. YHY: Validation, Writing – review & editing, Conceptualization. ML: Supervision, Writing – review & editing, Resources. BF: Writing – review & editing, Funding acquisition, Supervision.

## Funding

The author(s) declare financial support was received for the research and/or publication of this article. This work was financially supported by the China Agriculture Research System of MOF and MARA (CARS-06-A26), the Minor Grain Crops Research and Development System of Shaanxi Province (NYKJ-2021-YL(XN) 40), the “Two-chain” Fusion Crop Breeding Key Project of Shaanxi (2021-LLRH-07).

## References

- Badiane, F. A., Gowda, B. S., Ciss, N., Diouf, D., Sadio, O., and Timko, M. P. (2012). Genetic relationship of cowpea (*Vigna unguiculata*) varieties from Senegal based on SSR markers. *Genet. Mol. Res.* 11, 292–304. doi: 10.4238/2012.february.8.4
- Bonman, J. M., Babiker, E. M., Cuesta-Marcos, A., Esvelt-Klos, K., Brown-Guedira, G., Chao, S., et al. (2015). Genetic diversity among wheat accessions from the USDA national small grains collection. *Crop Sci.* 55, 1243–1253. doi: 10.2135/cropsci2014.09.0621
- Cho, Y., Chung, J. W., Lee, G. A., Ma, K. H., Dixit, A., Gwag, J. G., et al. (2010). Development and characterization of twenty-five new polymorphic microsatellite markers in proso millet (*Panicum miliaceum* L.). *Genes Genomics* 3, 267–273. doi: 10.1007/s13258-010-0007-8
- Courtois, B., Frouin, J., Greco, R., Bruschi, G., Droc, G., Hamelin, C., et al. (2012). Genetic diversity and population structure in a European collection of rice. *Crop Sci.* 52, 1663–1675. doi: 10.2135/cropsci2011.11.0588
- Duan, Y., Liu, J., Xu, J., Bian, C., Duan, S., Pang, W., et al. (2019). DNA fingerprinting and genetic diversity analysis with simple sequence repeat markers of 217 potato cultivars (*Solanum tuberosum* L.) in China. *Am. J. Potato Res.* 96, 21–32. doi: 10.1007/s12230-018-9685-6
- Gaju, O., Reynolds, M. P., Sparkes, D. L., and Foulkes, M. J. (2009). Relationships between large-spike phenotype, grain number, and yield potential in spring wheat. *Crop Sci.* 49, 961–973. doi: 10.2135/cropsci2008.05.0285
- Grover, A., and Sharma, P. C. (2016). Development and use of molecular markers: past and present. *Crit. Rev. Biotechnol.* 36, 290–302. doi: 10.3109/07388551.2014.959891
- Guo, S. Q., Han, F., Liu, M. X., Han, H. B., Dong, K. J., Yang, J. X., et al. (2025). A genome-wide association study reveals the genetic architecture of 19 agronomic traits in broomcorn millet (*Panicum miliaceum* L.). *Theor. Appl. Genet.* 138, 89–89. doi: 10.1007/s00122-025-04870-7
- He, Q., Wang, C. C., Zhang, J., Liang, H. K., Lu, Z. F., Xie, K., et al. (2024). A complete reference genome assembly for foxtail millet and *Setaria*-DB, a comprehensive database for *Setaria*. *Mol. Plant* 17, 219–222. doi: 10.1016/j.molp.2023.12.017
- Hou, S. Y., Sun, Z. X., Li, Y. S., Wang, Y. J., Ling, H. B., Xing, G. F., et al. (2017). Transcriptomic analysis, genic SSR development, and genetic diversity of proso millet (*Panicum miliaceum*; Poaceae). *Appl. Plant Sci.* 5, 1600137. doi: 10.3732/app.1600137
- Hu, X. Y., Wang, J. F., Lu, P., and Zhang, H. S. (2009). Assessment of genetic diversity in broomcorn millet (*Panicum miliaceum* L.) using SSR markers. *J. Genet. Genomics* 36, 491–500. doi: 10.1016/S1673-8527(08)60139-3
- Hunt, H. V., Campana, M. G., Lawes, M. C., Park, Y. J., Bower, M. A., Howe, C. J., et al. (2011). Genetic diversity and phylogeography of broomcorn millet (*Panicum miliaceum* L.) across Eurasia. *Mol. Ecol.* 20, 4756–4771. doi: 10.1111/j.1365-294X.2011.05318.x
- Hunt, H. V., Rudzinski, A., Jiang, H. E., and Wang, R. Y. (2018). Genetic evidence for a western Chinese origin of broomcorn millet (*Panicum miliaceum* L.). *Holocene* 28, 1968–1978. doi: 10.1177/0959683618798116
- Jiang, Y. M., Dong, L., Li, H. Q., Liu, Y. N., Wang, X. D., and Liu, G. Q. (2024). Genetic linkage map construction and QTL analysis for plant height in proso millet (*Panicum miliaceum* L.). *Theor. And Appl. Genet.* 137, 78–78. doi: 10.1007/s00122-024-04576-2

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2025.1649200/full#supplementary-material>

- Khound, R., Rajput, S. G., Schnable, J. C., Vetriventhan, M., and Santra, D. K. (2024). Genome-wide association study reveals marker-trait associations for major agronomic traits in proso millet (*Panicum miliaceum* L.). *Planta* 260, 44. doi: 10.1007/s00425-024-04465-4
- Krishna, T. A., Maharajam, T., David, R. H. A., Ramakrishnan, M., Ceasar, S. A., and Duraipandiyan, V. (2018). Microsatellite markers of finger millet (*Eleusine coracana* (L.) Gaertn) and foxtail millet (*Setaria italica* (L.) Beauv) provide resources for cross-genome transferability and genetic diversity analyses in other millets. *Biocatalysis Agric. Biotechnol.* 16, 493–501. doi: 10.1016/j.bcab.2018.09.009
- Li, Y., Jia, J., Wang, Y., and Wu, S. (1998). Intraspecific and interspecific variation in *Setaria* revealed by RAPD analysis. *Genet. Resour. Crop Evol.* 45, 279–285. doi: 10.1023/A:1008600123509
- Li, R., Zhang, H., Zhou, X., Guan, Y., Yao, F., Song, G., et al. (2010). Genetic diversity in Chinese sorghum landraces revealed by chloroplast simple sequence repeats. *Genet. Resour. Crop Evol.* 57, 1–15. doi: 10.1007/s10722-009-9446-y
- Lilia, E., Abderrazek, J., Stefania, G., Salvatore, C., Maria, V. K. S., Hisashi, T., et al. (2008). Genetic diversity and association analysis for salinity tolerance, heading date and plant height of barley germplasm using simple sequence repeat markers. *J. Integr. Plant Biol.* 50, 1004–1014. doi: 10.1111/j.1744-7909.2008.00670.x
- Liu, K. J., and Muse, S. V. (2005). PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics* 21, 2128–2128. doi: 10.1093/bioinformatics/bti282
- Liu, M. X., Xu, Y., He, J. H., Zhang, S., Wang, Y. Y., and Lu, P. (2016). Genetic diversity and population structure of broomcorn millet (*Panicum miliaceum* L.) cultivars and landraces in China based on microsatellite markers. *Int. J. Mol. Sci.* 17, 370. doi: 10.3390/ijms17030370
- Liu, M. X., Xu, Y., Yang, T. T., Qiao, Z. J., Wang, R. Y., Wang, Y. Y., et al. (2017). Development of species-specific microsatellite markers for broomcorn millet (*Panicum miliaceum* L.) via high-throughput sequencing. *Adv. Crop Sci. Technol.* 5, 1–10. doi: 10.4172/2329-8863.1000297
- Meng, Y., Zhao, N., Li, H., Zhai, H., He, S., and Liu, Q. (2018). SSR fingerprinting of 203 sweet potato (*Ipomoea batatas* (L.) Lam) varieties. *J. Integr. Agric.* 17, 86–93. doi: 10.1016/S2095-3119(17)61687-3
- Odong, T. L., Van Heerwaarden, J., Van Hintum, T. J. L., Van Eeuwijk, F. A., and Jansen, J. (2013). Improving hierarchical clustering of genotypic data via principal component analysis. *Crop Sci.* 53, 1546–1554. doi: 10.2135/cropsci2012.04.0215
- Panigrahi, P., Panigrahi, K. K., and Bhattacharya, S. (2018). SSR marker based DNA fingerprinting and diversity studies in mustard (*Brassica juncea*). *Electronic J. Plant Breed.* 9, 25–37. doi: 10.5958/0975-928X.2018.00004.2
- Pucher, A., Sy, O., Angarawai, I. I., Gondah, J., Zangre, R., Ouedraogo, M., et al. (2015). Agro-morphological characterization of West and Central African pearl millet accessions. *Crop Sci.* 55, 737–748. doi: 10.2135/cropsci2014.06.0450
- Rajput, G. S., Khound, R., and Santra, K. D. (2024). Phenotypic diversity for morpho-agronomic traits in the US proso millet (*Panicum miliaceum* L.) core collection. *Crop Sci.* 64, 2599–2611. doi: 10.1002/csc2.21245
- Rajput, S. G., Plyler-Harveson, T., and Santra, D. K. (2014). Development and characterization of SSR markers in proso millet based on switchgrass genomics. *Am. J. Plant Sci.* 5, 175–186. doi: 10.4236/ajps.2014.51023
- Rajput, G. S., and Santra, K. D. (2016). Evaluation of genetic diversity of proso millet germplasm available in the United States using simple-sequence repeat markers. *Crop Sci.* 56, 2401–2409. doi: 10.2135/cropsci2015.10.0644
- Rajput, S. G., Santra, D. K., and Schnable, J. (2013). Mapping QTLs for morpho-agronomic traits in proso millet (*Panicum miliaceum* L.). *Mol. Breed.* 36, 37. doi: 10.1007/s11032-016-0460-4
- Sharma, P., Mehta, G., Shefail, Muthusamy, S. K., Singh, S. K., and Singh, G. P. (2021). Development and validation of heart-responsive candidate gene and miRNA gene based SSR markers to analysis genetic diversity in wheat for heat tolerance breeding. *Mol. Biol. Rep.* 48, 381–393. doi: 10.1007/s11033-020-06059-1
- Shi, J. P., Ma, X. X., Zhang, J. H., Zhou, Y. S., Liu, M. X., Huang, L. L., et al. (2019). Chromosome conformation capture resolved near complete genome assembly of broomcorn millet. *Nat. Commun.* 10, 464. doi: 10.1038/s41467-018-07876-6
- Tian, H. L., Wang, F. G., Zhao, J. R., Yi, H. M., Wang, L., Wang, R., et al. (2015). Development of maize SNP3072, a high-throughput compatible SNP array, for DNA fingerprinting identification of Chinese maize varieties. *Mol. Breed.* 35, 136. doi: 10.1007/s11032-015-0335-0
- Upadhyaya, D. H., Sharma, S., Gowda, L. L. C., Reddy, G., and Singh, S. (2011). Developing proso millet (*Panicum miliaceum* L.) core collection using geographic and morpho-agronomic data. *Crop Pasture Sci.* 62, 383–389. doi: 10.1071/CP10294
- Van Inghelandt, D., Melchinger, A. E., Lebreton, C., and Stich, B. (2010). Population structure and genetic diversity in a commercial maize breeding program assessed with SSR and SNP markers. *Theor. Appl. Genet.* 120, 1289–1299. doi: 10.1007/s00122-009-1256-2
- Vetriventhan, M., Azevedo, V. C. R., Upadhyaya, H. D., and Naresh, D. (2019). Variability in the global proso millet (*Panicum miliaceum* L.) germplasm collection conserved at the ICRISAT genebank. *Agriculture* 9, 112. doi: 10.3390/agriculture9050112
- Vetriventhan, M., and Upadhyaya, D. H. (2018). Diversity and trait-specific sources for productivity and nutritional traits in the global proso millet (*Panicum miliaceum* L.) germplasm collection. *Crop J.* 6, 451–463. doi: 10.1016/j.cj.2018.04.002
- Wang, R., Hunt, H. V., Qiao, Z., Wang, L., and Han, Y. (2016a). Diversity and cultivation of broomcorn millet (*Panicum miliaceum* L.) in China: a review. *Economic Bot.* 70, 332–342. doi: 10.1007/s12231-016-9357-8
- Wang, C. F., Jia, G. Q., Zhi, H., Niu, Z. G., Chai, Y., Li, W., et al. (2012). Genetic diversity and population structure of Chinese foxtail millet [*Setaria italica* (L.) Beauv.] landraces. *G3* 2, 769–777. doi: 10.1534/g3.112.002907
- Wang, X., Wang, H., Liu, S., Ferjani, A., Li, J., Yan, J., et al. (2016b). Genetic variation in *ZmVPP1* contributes to drought tolerance in maize seedlings. *Nat. Genet.* 48, 1233–1241. doi: 10.1038/ng.3636
- Wang, L., Wang, X. Y., Wen, Q. F., and Wu, B. E. (2005). Research and utilization of proso millet germplasm resource in China. *J. Plant Genet. Resour.* 6, 474–477. doi: 10.3969/j.issn.1672-1810.2005.04.025
- Weber, J. L. (1990). Informativeness of human (dC-dA)<sub>n</sub>(dG-dT)<sub>n</sub> polymorphism. *Genomics* 7, 524–530. doi: 10.1016/0888-7543(90)90195-z
- Wen, Z. X., Ding, Y. L., Zhao, T. J., and Gai, J. Y. (2009). Genetic diversity and peculiarity of annual wild soybean (*G. soja* Sieb. et Zucc.) from various eco-regions in China. *Theor. Appl. Genet.* 119, 371–381. doi: 10.1007/s00122-009-1045-y
- Yang, Q. H., Yuan, Y. H., Liu, J. J., Han, M. R., Li, J., Jin, F., et al. (2023). Transcriptome analysis reveals new insights in the starch biosynthesis of non-waxy and waxy broomcorn millet (*Panicum miliaceum* L.). *Int. J. Biol. Macromolecules* 230, 123155. doi: 10.1016/j.ijbiomac.2023.123155
- Yazdizadeh, M., Fahmideh, L., Mohammadi-Nejad, G., Solouki, M., and Nakhoda, B. (2020). Association analysis between agronomic traits and AFLP markers in a wide germplasm of proso millet (*Panicum miliaceum* L.) under normal and salinity stress conditions. *BMC Plant Biol.* 20, 427. doi: 10.1186/s12870-020-02639-2
- Zhang, D. Z., Panhwar, R. B., Liu, J. J., Gong, X. W., Liang, J. B., Liu, M. X., et al. (2019). Morphological diversity and correlation analysis of phenotypes and quality traits of proso millet (*Panicum miliaceum* L.) core collections. *J. Integr. Agric.* 18, 958–969. doi: 10.1016/S2095-3119(18)61997-5
- Zhao, Z. (2011). New archaeobotanic data for the study of the origins of agriculture in China. *Curr. Anthropology* 52, S295–S306. doi: 10.1086/659308