



OPEN ACCESS

EDITED BY
Cailong Xu,
Chinese Academy of Agricultural
Sciences (CAAS), China

REVIEWED BY
Santosh Rajput,
Dryland Genetics, Inc., United States
Lorena Armando,
National University of the South,
Argentina

*CORRESPONDENCE
Feng Guojun
✉ fengguojun@126.com
Nadeem Bhanbhro
✉ nadm428@nenu.edu.cn

[†]These authors have contributed equally
to this work

RECEIVED 23 January 2026
REVISED 11 April 2026
ACCEPTED 23 April 2026
PUBLISHED 19 May 2026

CITATION
Yun Z, Jun L, Hong-Jin W, Yang Z,
Xiangwei H, Kuerban Z, Hui W, Baoyi Y,
Guojun F and Bhanbhro N (2026)
Phenotypic variation and genetic
diversity assessment of proso millet
(*Panicum miliaceum* L.) germplasm
resources in northern Xinjiang.
Front. Plant Sci. 17:1792533.
doi: 10.3389/fpls.2026.1792533

COPYRIGHT
© 2026 Yun, Jun, Hong-Jin, Yang,
Xiangwei, Kuerban, Hui, Baoyi, Guojun
and Bhanbhro. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/)
(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.

Phenotypic variation and genetic diversity assessment of proso millet (*Panicum miliaceum* L.) germplasm resources in northern Xinjiang

Zhao Yun^{1†}, Liu Jun^{1†}, Wang Hong-Jin^{1†}, Zhao Yang²,
Hu Xiangwei¹, Zaituniguli Kuerban¹, Wang Hui¹, Yang Baoyi¹,
Feng Guojun^{1*} and Nadeem Bhanbhro^{3,4,5*}

¹Crop Research Institute, Xinjiang Academy of Agricultural Sciences, Urumqi, Xinjiang, China, ²Seed Industry Development Center of Xinjiang Uygur Autonomous Region, Urumqi, Xinjiang, China, ³Sanya Institute, Hainan Academy of Agricultural Sciences, Sanya, China, ⁴National Nanfan Research Institute, Chinese Academy of Agricultural Sciences, Sanya, China, ⁵Biotechnology Research Institute, Chinese Academy of Agricultural Sciences, Beijing, China

Characterizing genetic diversity in proso millet (*Panicum miliaceum* L.) is essential for crop improvement and breeding programs. This study evaluated the genetic diversity, phenotypic variation, and population structure of 46 proso millet accessions using five qualitative and twelve quantitative traits across two growing seasons (2023–2024) in northern Xinjiang. Grain color showed the highest genetic diversity index ($H' = 2.064$), while panicle type showed the lowest ($H' = 1.383$). Green panicle color (84.78%), long panicle branch length (63.04%), lateral panicle type (71.74%), and yellow grain color (86.96%) were predominant. Grain yield (Y) and grain weight per panicle (GWP) exhibited the highest coefficients of variation (15.00% and 14.61%, respectively) and substantial genetic diversity indices (Y: $H' = 2.11$; GWP: $H' = 2.12$), indicating considerable selection potential. Grain yield was significantly positively correlated with grain weight per panicle ($r = 0.901$, $P < 0.01$) and main panicle length ($r = 0.863$, $P < 0.01$). Combined analysis of variance revealed highly significant genotypic effects for all agronomic traits, while genotype $\times$ environment interaction effects were non-significant for all traits, indicating stable genotype rankings across years. Broad-sense heritability estimates were very high for yield-related traits, with grain yield ($H^2 = 0.998$) and grain weight per panicle ($H^2 = 0.996$) demonstrating strong genetic control. Stepwise regression analysis identified grain weight per panicle and main panicle length as the primary yield predictors ($R^2 = 0.91$). Principal component analysis showed PC1 and PC2 explained 70.13% of total variation, representing vegetative vigor and yield capacity, respectively. Cluster analysis identified three groups with distinct trait profiles suitable for dual-purpose breeding, early-maturing cultivar development, and forage breeding, respectively. The evaluated germplasm possesses substantial genetic diversity, providing valuable resources for breeding parent selection in northern Xinjiang proso millet improvement programs.

KEYWORDS

genetic diversity, genotype $\times$ environment interaction, germplasm evaluation, phenotypic variation, proso millet (*Panicum miliaceum* L.), trait-based breeding

1 Introduction

Proso millet (*Panicum miliaceum* L.), a member of the Poaceae family, has maintained its agricultural significance across diverse geographical regions as a climate-resilient cereal crop (Stevens et al., 2021; Dai et al., 2021). Archaeological evidence places its domestication approximately 10,000 years ago in northern China (He et al., 2021), and traditional cultivars such as ‘Taromi’ remain integral to local farming systems in Xinjiang, where the crop constitutes a staple among the Kazakh population (Zhang et al., 2021). The crop’s exceptional drought tolerance, short growing season (60–90 days), and minimal water requirements (300–400 mm annually) make it strategically important for sustainable agriculture in semi-arid and arid environments (James et al., 2025; Zhao et al., 2025; Loni et al., 2023), which provides comparative context from foxtail millet, a closely related *Setaria italica* species sharing similar arid-zone adaptations). These attributes are particularly relevant in the context of global climate change and increasing food security concerns in marginal lands where conventional cereals struggle to thrive (Bhanbhro et al., 2024; 2025).

The adaptability of proso millet is evidenced by its recent expansion into marginal environments such as northern Xinjiang (Pavithra and Rawat, 2024; Luo et al., 2022), where the planting area reached approximately 13,333 ha, with average yields of 3.75 t·ha⁻¹. The successful establishment of a complete local value chain, encompassing cultivation, processing, and marketing, underscores the crop potential to contribute to food security and climate-resilient agriculture, as discussed in this chapter (Pavithra and Rawat, 2024; Nandini et al., 2026). This growth occurs within a challenging agricultural environment characterized by extreme continental climate conditions, including significant temperature fluctuations, limited precipitation, soil salinity, and relatively short frost-free periods (Han et al., 2020; Bhanbhro et al., 2020; Wimalasiri et al., 2023). Despite the crop’s inherent adaptability, the industry advancement is constrained by critical deficiencies in its genetic foundation: inadequate germplasm characterization, incomplete collection efforts, and insufficient evaluation protocols (AbdElgawad et al., 2020; Aibai et al., 2024). These limitations hinder targeted varietal improvement and sustainable development of the regional proso millet sector.

Germplasm resources encompassing landraces, improved varieties, wild relatives and genetic stocks constitute the fundamental basis for crop improvement and genetic research (Zhang et al., 2021). They serve as essential repositories of genetic diversity, harboring valuable alleles for key traits such as drought tolerance, disease resistance, early maturity and enhanced grain quality (Zhao et al., 2025). The systematic evaluation and characterization of these collections are critical prerequisites for understanding genetic potential, identifying superior genotypes, and developing climate-resilient cultivars (Mukami et al., 2019; Rivero et al., 2022). In northern Xinjiang, such assessment is particularly important due to the region’s distinctive agro-climatic pressures (Ma et al., 2025a; 2025b) and the urgent need to identify genotypes optimized for local agricultural systems. Furthermore, genotype × environment (G×E) interactions play a critical role in determining crop performance, making location-specific evaluation indispensable for maximizing agricultural productivity (Pei et al., 2023; Rizwan et al., 2025).

Recently studies on proso millet germplasm from Xinjiang and neighboring regions have reported substantial phenotypic variation, with coefficients of variation ranging from 12.37% to 91.40% and genetic diversity indices of 0.410 to 0.809 (Lv et al., 2025). Principal component analysis has revealed that the first four components account for 76.64% of total phenotypic variation, with yield per plant, stem diameter, effective panicle number, and leaf width identified as primary sources of variability. Cluster analysis based on morphological data has revealed distinct genetic differentiation patterns correlating with geographical origin in related millet species (Darkwa et al., 2020). However, critical knowledge gaps remain regarding the adaptive mechanisms, phenotypic stability, and precise yield-formation relationships of proso millet genotypes under northern Xinjiang conditions.

This study addresses these gaps through integrated evaluation of 46 cultivated and introduced proso millet accessions adapted to northern Xinjiang conditions using a multidisciplinary approach combining field-based agronomic characterization with multivariate statistical analysis. The specific objectives were to: (1) evaluate genetic diversity and phenotypic variation across qualitative and quantitative traits; (2) analyze genotype × environment (G×E) interactions and estimate broad-sense heritability for key agronomic traits; (3) develop optimized regression models linking grain yield with contributing traits; and (4) identify superior germplasm with desirable trait combinations for targeted breeding applications. The outcomes will provide empirical foundations for conservation strategies, germplasm selection, parental line identification, and breeding program design in support of the regional proso millet industry (Allier et al., 2020).

2 Materials and methods

2.1 Experimental materials

This study evaluated 46 proso millet accessions collected from 11 provinces and autonomous regions of China (Inner Mongolia, Gansu, Shanxi, Ningxia, Hebei, Xinjiang, Heilongjiang, and Shaanxi, Liaoning, Henan, Jilin; Figure 1). Accessions were selected to represent diverse ecological zones across major proso millet-producing regions of China (Supplementary Table 1). A total of 17 phenotypic traits were evaluated: five qualitative traits (inflorescence color, panicle branch length, inflorescence density, panicle type, and grain color) and 12 quantitative traits (growth period, number of productive tillers, main stem height, main panicle length, number of main stem nodes, main stem diameter, flag leaf length, flag leaf width, number of leaves on the main stem, grain weight per panicle, thousand-grain weight, and grain yield).

2.2 Experimental design

The field trial was conducted over two consecutive growing seasons (2023 and 2024) in Qitai County, Xinjiang, China (44.02°N, 89.57°E). The experiment employed a randomized complete block design (RCBD) with three biological replicates. Each experimental plot measured 8 m² (5 m × 1.6 m) and was planted using a locally adapted plastic film mulching system with hole-seeding at 40 cm

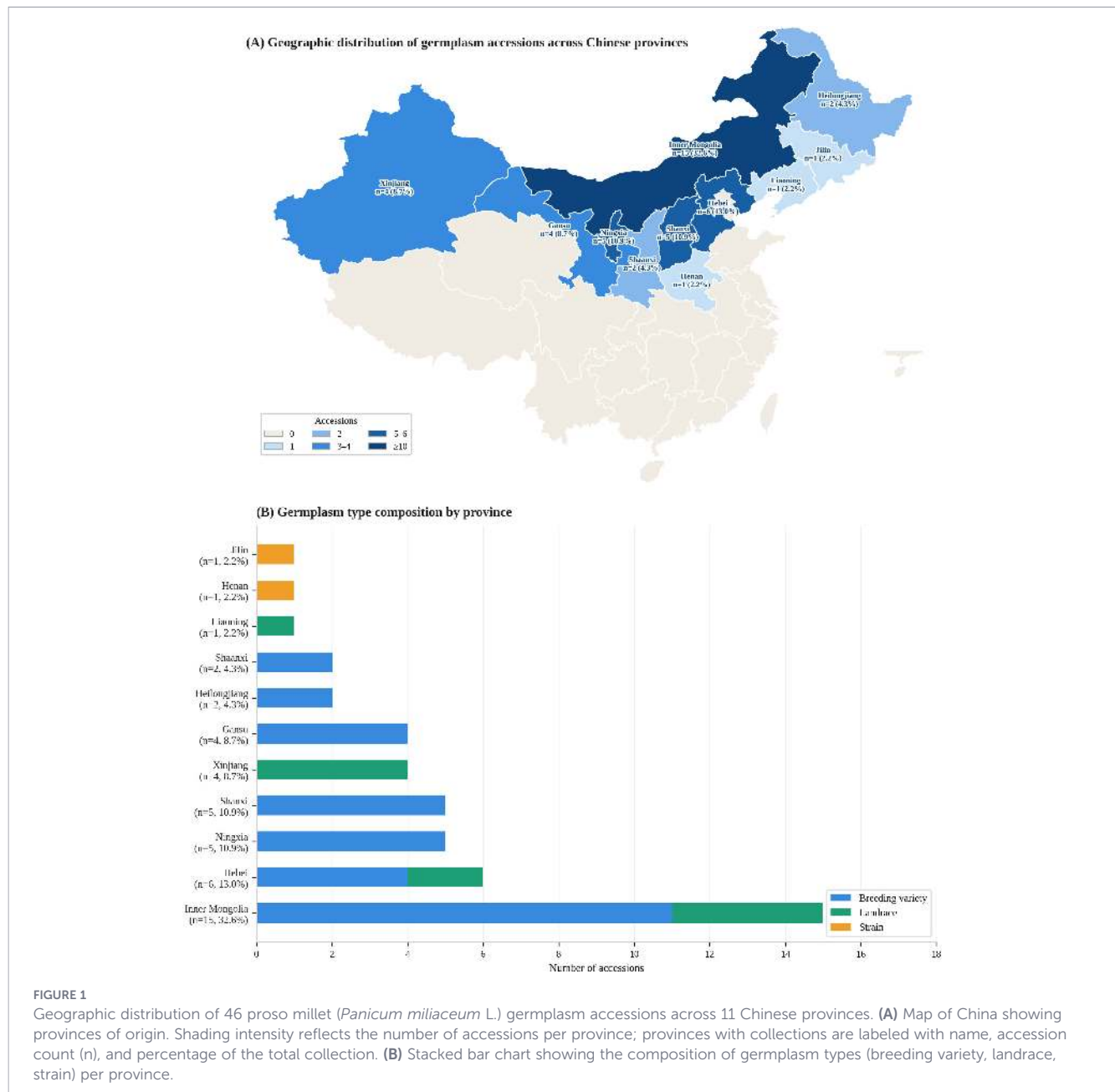


FIGURE 1

Geographic distribution of 46 proso millet (*Panicum miliaceum* L.) germplasm accessions across 11 Chinese provinces. (A) Map of China showing provinces of origin. Shading intensity reflects the number of accessions per province; provinces with collections are labeled with name, accession count (n), and percentage of the total collection. (B) Stacked bar chart showing the composition of germplasm types (breeding variety, landrace, strain) per province.

row spacing, 10 cm hill spacing and a seeding rate of approximately 6.7 g per plot (equivalent to 8.4 kg ha⁻¹), calculated based on a target plant density of 375,000 hills ha⁻¹ (40 cm × 10 cm spacing) with 4 seeds per hill and the mean 1000-grain weight of 8.41 g of the accessions evaluated. This seed rate, although lower than conventional broadcast seeding rates, is appropriate for the hole-seeding plastic film mulching system used in this study, which achieves higher germination efficiency per hill than broadcast methods (Palchetti et al., 2023). Diammonium phosphate (150 kg ha⁻¹) was applied as basal fertilizer. All plots received uniform irrigation four times during the growth cycle, and standardized field management practices were maintained throughout. Border rows surrounded the experimental area to minimize edge effects.

2.3 Trait evaluation methods

2.3.1 Qualitative trait assessment

Five qualitative traits were visually evaluated following the descriptors and data standards for common millet germplasm (Ren et al., 2024): inflorescence color (green/purple), panicle branch length (referring to the length of primary branches arising from the main panicle axis, classified as short: <5 cm, medium: 5–10 cm, and long: >10 cm), inflorescence density (sparse/medium/moderately dense/dense), panicle type (lateral/lax/compact), and grain color (yellow/red/brown/white/gray/variegated). Ten representative plants per plot were assessed for each trait.

2.3.2 Quantitative trait measurement

Twelve quantitative traits were measured following standard descriptors (Anyaocha et al., 2023; Zhao et al., 2025), with ten representative plants per plot evaluated for each trait. Growth period (GP, days) was recorded as the number of days from seedling emergence to physiological maturity. Number of productive tillers (NPT, tillers plant⁻¹) was counted at maturity, defined as all tillers bearing a panicle. Main stem height (MSH, cm) was measured from the soil surface to the base of the main panicle at maturity. Main panicle length (MPL, cm) was measured from the basal node of the main panicle to its tip. Number of main stem nodes (NMSN) was counted on the main culm at maturity. Main stem diameter (MSD, mm) was measured at the first internode above the ground using digital calipers. At the heading stage, flag leaf length (LL, cm) was measured from the ligule to the leaf tip, and flag leaf width (LW, mm) was measured at the widest point of the flag leaf blade using digital calipers. Number of leaves on the main stem (NL) was counted at maturity. Grain weight per panicle (GWP, g) was determined by threshing and weighing grains from the main panicle using a precision balance. Thousand-grain weight (GW, g) was determined by counting 1,000 clean, air-dried grains and recording their weight using a precision balance. Grain yield (Y, kg-ha⁻¹) was calculated from plot harvest weight adjusted to 13% moisture content, then converted to a per-hectare basis. All measurements were performed by the same trained personnel using calibrated instruments to minimize inter-observer variation.

2.4 Data analysis

2.4.1 Genetic diversity analysis

Genetic diversity of qualitative traits was assessed using the Shannon-Weaver diversity index: $H' = -\sum(p_i \times \ln p_i)$, where p_i represents the phenotypic frequency of category i (Gebre et al., 2025). Quantitative trait variation was evaluated using the coefficient of variation: $CV = (\text{standard deviation}/\text{mean}) \times 100\%$ (Pélabon et al., 2020).

2.4.2 Multivariate statistical analysis

Principal component analysis (PCA) was performed using SPSS software (Version 26.0, IBM Corp.) (Elhaik, 2022; Greenacre et al., 2022; Varraso et al., 2012). Comprehensive evaluation scores (F-values) for each accession were calculated based on PCA results using weighted principal component contributions. Pearson correlation analysis and hierarchical cluster analysis (Ward's method with squared Euclidean distance) were conducted using Origin 2024 software.

2.4.3 Combined analysis of variance and heritability

A mixed-effects model was used to analyze the two-year data, with year (environment) treated as a random effect and genotype as a fixed effect. Combined analysis of variance (ANOVA) was

performed to partition phenotypic variance into genotypic (σ^2_g), genotype $\times$ environment interaction (σ^2_{ge}), and residual error (σ^2_e) components. Broad-sense heritability was estimated as: $H^2 = \sigma^2_g / (\sigma^2_g + \sigma^2_{ge}/e + \sigma^2_e/r)$, where e is the number of environments and r is the number of replicates (Yu et al., 2022). Multiple comparisons were performed using the LSD method ($p < 0.05$).

2.4.4 Regression analysis

Stepwise multiple regression analysis was performed using SPSS 26.0 to identify the most significant predictors of grain yield. Grain yield was defined as the dependent variable and the twelve quantitative traits were defined as independent variables. Variables were entered and removed at significance thresholds of $p \leq 0.05$ and $p \geq 0.10$, respectively. Model fit was assessed using the coefficient of determination (R^2), adjusted R^2 , and the standard error of the estimate.

3 Results

3.1 Genetic diversity of qualitative traits

The genetic diversity of five qualitative traits across 46 proso millet accessions is presented in Figure 2. Grain color exhibited the highest genetic diversity index ($H' = 2.064$), indicating broad variation and stable genetic expression for this trait, followed by panicle branch length ($H' = 2.045$). Panicle type displayed the lowest diversity index ($H' = 1.383$). The genetic diversity indices for the five qualitative traits ranked as follows: grain color > panicle branch length > inflorescence density > inflorescence color > panicle type.

3.2 Distribution patterns of qualitative traits

The distribution patterns and frequency analysis of the five qualitative traits revealed distinct variation within the germplasm collection (Figure 3). Inflorescence color exhibited a clear predominance of green types (84.78%), with purple types representing a minority variant (15.22%). Purple inflorescence pigmentation was observed in the same accessions across both growing seasons (2023 and 2024), confirming that this trait is under stable genetic control in this germplasm panel. Although anthocyanin expression can be influenced by environmental factors such as day length and nitrogen availability (Zhang et al., 2026), no switching between pigmentation classes was observed across years in this study. Panicle branch length showed a skewed distribution toward longer types (63.04%), followed by medium (28.26%) and short (8.70%) types. Inflorescence density was predominantly medium (71.74%), with slightly dense types (15.22%), sparse types (8.70%), and dense types being uncommon (4.35%). Panicle type was dominated by lateral panicles (71.74%), with lax panicles as a secondary type (23.91%) and compact panicles being relatively uncommon (4.35%) (Table 1). Grain color was strongly dominated by yellow types (86.96%), with red and brown grain types repre-

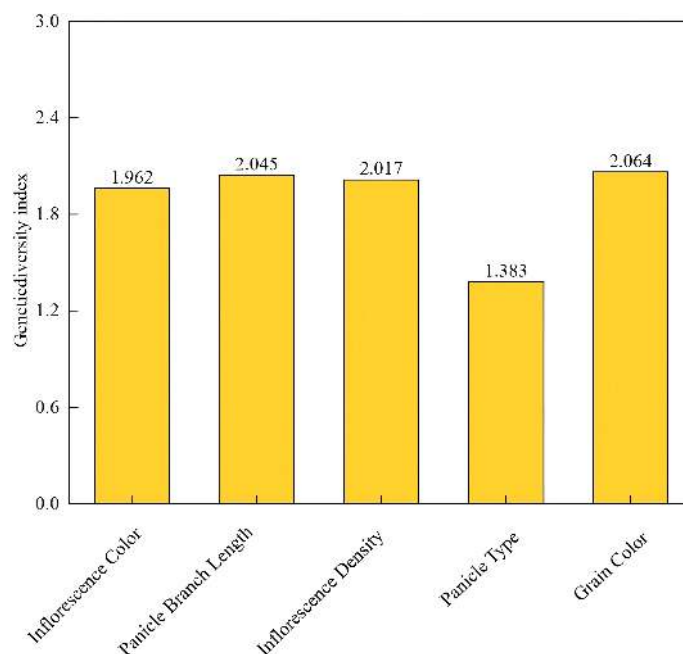


FIGURE 2

Shannon-Weaver genetic diversity index (H') for five qualitative traits in proso millet germplasm. Grain color exhibited the highest diversity ($H' = 2.064$), while panicle type showed the lowest ($H' = 1.383$).

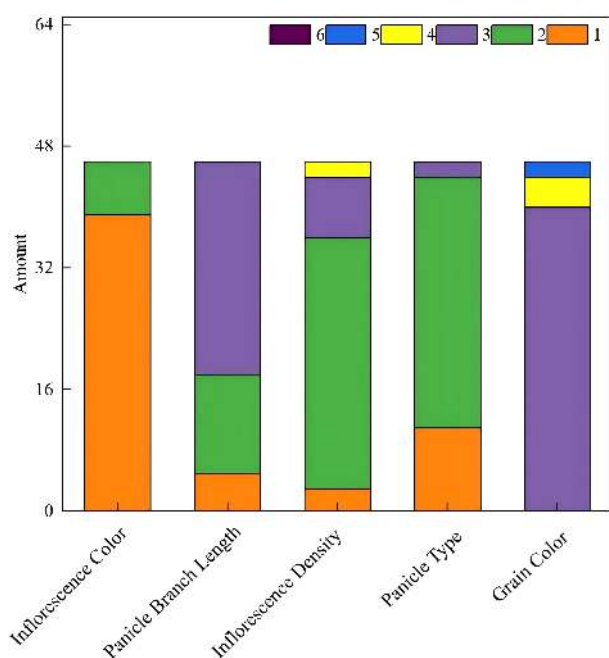


FIGURE 3

Frequency distribution of five qualitative traits in proso millet germplasm. Inflorescence Color: 1. Green, 2. Purple; Panicle Branch Length: 1. Short, 2. Medium, 3. Long; Inflorescence Density: 1. Sparse, 2. Medium, 3. Moderately dense, 4. Dense; Panicle Type: 1. Lax panicle, 2. Lateral panicle, 3. Compact panicle; Grain Color: 1. White, 2. Gray, 3. Yellow, 4. Red, 5. Brown, 6. Variegated.

senting 8.70% and 4.34%, respectively; white, gray, and variegated grain types were absent from this collection. These distribution patterns provide valuable baseline information for germplasm management and targeted breeding strategies.

3.3 Variation and diversity of quantitative traits

Analysis of 12 quantitative traits revealed considerable variation in both coefficient of variation (CV) and Shannon-Weaver diversity index (H') across the 46 accessions (Table 2). Leaf width (LW) exhibited the highest CV (24.14%), indicating substantial variation in this trait among accessions. Grain yield (Y) showed the second highest CV (15.00%), followed by number of main stem nodes (NMSN, CV = 14.85%), main stem diameter (MSD, CV = 14.81%), number of leaves (NL, CV = 14.67%), and grain weight per panicle (GWP, CV = 14.61%), with GWP ranging from 18.96 to 36.88 g. Main stem height (MSH) showed a CV of 13.88%, and main panicle length (MPL) a CV of 13.83%. Moderate variation was observed for leaf length (LL, CV = 10.04%), thousand-grain weight (GW, CV = 10.05%), and number of productive tillers (NPT, CV = 8.92%), while growth period (GP) showed the lowest CV (8.23%), reflecting relatively stable phenological development across accessions.

Shannon-Weaver diversity indices ranged from 1.788 (MSH) to 2.155 (LL). Leaf length demonstrated the highest H' value (2.155), followed by thousand-grain weight ($H' = 2.126$), grain weight per panicle ($H' = 2.122$), leaf width and grain yield (both $H' = 2.11$). Main panicle length showed intermediate diversity ($H' = 2.023$), while main stem diameter ($H' = 1.984$), growth period ($H' = 1.956$), number of main stem nodes ($H' = 1.865$), number of leaves ($H' = 1.849$), number of productive tillers ($H' = 1.815$), and main stem height ($H' = 1.788$) showed comparatively lower diversity levels. It should be noted that the Shannon-Weaver diversity index (H') is non-linear and sensitive to the number of phenotypic classes; therefore, comparisons of H' values across traits with differing numbers of categories should be interpreted with appropriate

TABLE 1 Classification and coding system for five qualitative traits in proso millet germplasm.

Trait	Descriptive criteria
Inflorescence Color	Green = 1, Purple = 2
Panicle Branch Length	Short = 1, Middle = 2, Long = 3
Inflorescence Density	Sparse = 1, Middle = 2, Slightly dense = 3, Dense = 4
Panicle Type	Scattered = 1, Side = 2, Dense = 3
Grain Color	White = 1, Grey = 2, Yellow = 3, Red = 4, Brown = 5, Broken = 6

caution. Furthermore, since approximately 59% of the 46 accessions originate from only three provinces (Inner Mongolia, Gansu, and Xinjiang), the observed diversity indices reflect the genetic composition of this specific collection rather than the full breadth of Chinese proso millet diversity. Accordingly, conclusions regarding diversity levels should be considered within the context of this geographically structured sample.

3.4 Phenotypic distribution of quantitative traits

Phenotypic variation analysis conducted on 12 agronomic traits across the 46 accessions indicated continuous variation for all traits with wide ranges and substantial phenotypic diversity (Figure 4). Growth period ranged from 75 to 110 days, with 78.26% of accessions concentrated between 85 and 105 days; a few late-maturing genotypes extended up to 110 days, indicating the presence of photoperiod-sensitive materials that may be valuable for specific agro-ecological zones. Main stem height showed a positively skewed distribution, with the majority of accessions classified as tall-stature types, reflecting the predominantly tall-growing ecotype characteristic of northern Chinese proso millet germplasm. Main panicle length ranged from 24 to 36 cm, with 82.61% of accessions distributed between 22 and 36 cm and 17.39%

classified as long-panicle types (≥ 36 cm); given the strong positive correlation of main panicle length with grain yield, these long-panicle accessions represent priority targets for yield-oriented breeding. Main stem diameter ranged from 4.5 to 9.0 mm, with thick-stem types (≥ 9.0 mm) accounting for 4.35% of the collection, representing valuable resources for lodging-resistant cultivar development. The number of productive tillers was dominated by the 3-tiller type (73.91%), a structure favoring assimilate concentration in the main panicle. High-yielding accessions ($\geq 4,500$ kg·ha⁻¹) accounted for 13.04% of the collection, indicating significant potential for further yield enhancement through targeted crossing and genotype aggregation strategies.

3.5 Correlation analysis of quantitative traits

Correlation analysis of key agronomic traits revealed that grain yield (Y) was highly significantly positively correlated with grain weight per panicle (GWP) and main panicle length (MPL) ($r = 0.938$ and 0.863 , respectively; $P < 0.01$), confirming that large-panicle characteristics constitute the primary yield-formation pathway in this germplasm under the prevailing field conditions of this study (Figure 5). Grain weight per panicle also showed a strong positive correlation with thousand-grain weight (GW) ($r = 0.847$, $P < 0.01$), suggesting that sink capacity expansion in this germplasm operates through concurrent increases in both spikelet number per panicle and individual grain mass, two complementary components that breeders can exploit simultaneously. Importantly, thousand-grain weight was also significantly positively correlated with grain yield ($r = 0.847$, $P < 0.01$), reinforcing its value as a secondary selection criterion for yield improvement. Main stem height (MSH) exhibited strong positive correlations with main stem node number (NMSN) and main stem diameter (MSD) ($r \geq 0.884$, $P < 0.01$) but showed a weak positive correlation with number of productive tillers (NPT) ($r = 0.195$, not significant), reflecting a plant architecture that prioritizes main stem growth over tillering—an adaptive strategy for concentrating photoassimilates in the main panicle under semi-

TABLE 2 Mean values, ranges, amplitude, coefficients of variation (CV), and Shannon–Weaver diversity indices (H') for 12 quantitative traits in 46 proso millet germplasm accessions (combined 2023–2024).

Trait	Mean $\pm$ SD	Range	Amplitude	CV (%)	H'
Growth Period (GP/d)	98.61 $\pm$ 8.11	76.67–116.00	39.33	8.23	1.956
No. of Productive Tillers (NPT)	3.36 $\pm$ 0.30	2.80–4.20	1.4	8.92	1.815
Main Stem Height (MSH/cm)	188.33 $\pm$ 26.14	114.35–218.71	104.36	13.88	1.788
Main Panicle Length (MPL/cm)	37.72 $\pm$ 5.22	25.26–46.19	20.93	13.83	2.023
No. of Main Stem Nodes (NMSN)	7.61 $\pm$ 1.13	4.12–9.46	5.34	14.85	1.865
Main Stem Diameter (MSD/mm)	7.10 $\pm$ 1.05	4.03–8.60	4.58	14.81	1.984
Leaf Length (LL/cm)	36.33 $\pm$ 3.65	28.00–41.37	13.38	10.04	2.155
Leaf Width (LW/cm)	1.75 $\pm$ 0.42	0.80–2.72	1.92	24.14	2.11
No. of Leaves (NL)	8.16 $\pm$ 1.20	4.45–9.93	5.48	14.67	1.849
Grain Weight per Panicle (GWP/g)	28.11 $\pm$ 4.11	18.96–36.88	17.92	14.61	2.122
1000-Grain Weight (GW/g)	8.19 $\pm$ 0.82	6.18–9.70	3.52	10.05	2.126
Yield (Y/kg·hm ⁻²)	4143.67 $\pm$ 621.75	2840.24–5513.26	2673.02	15	2.11

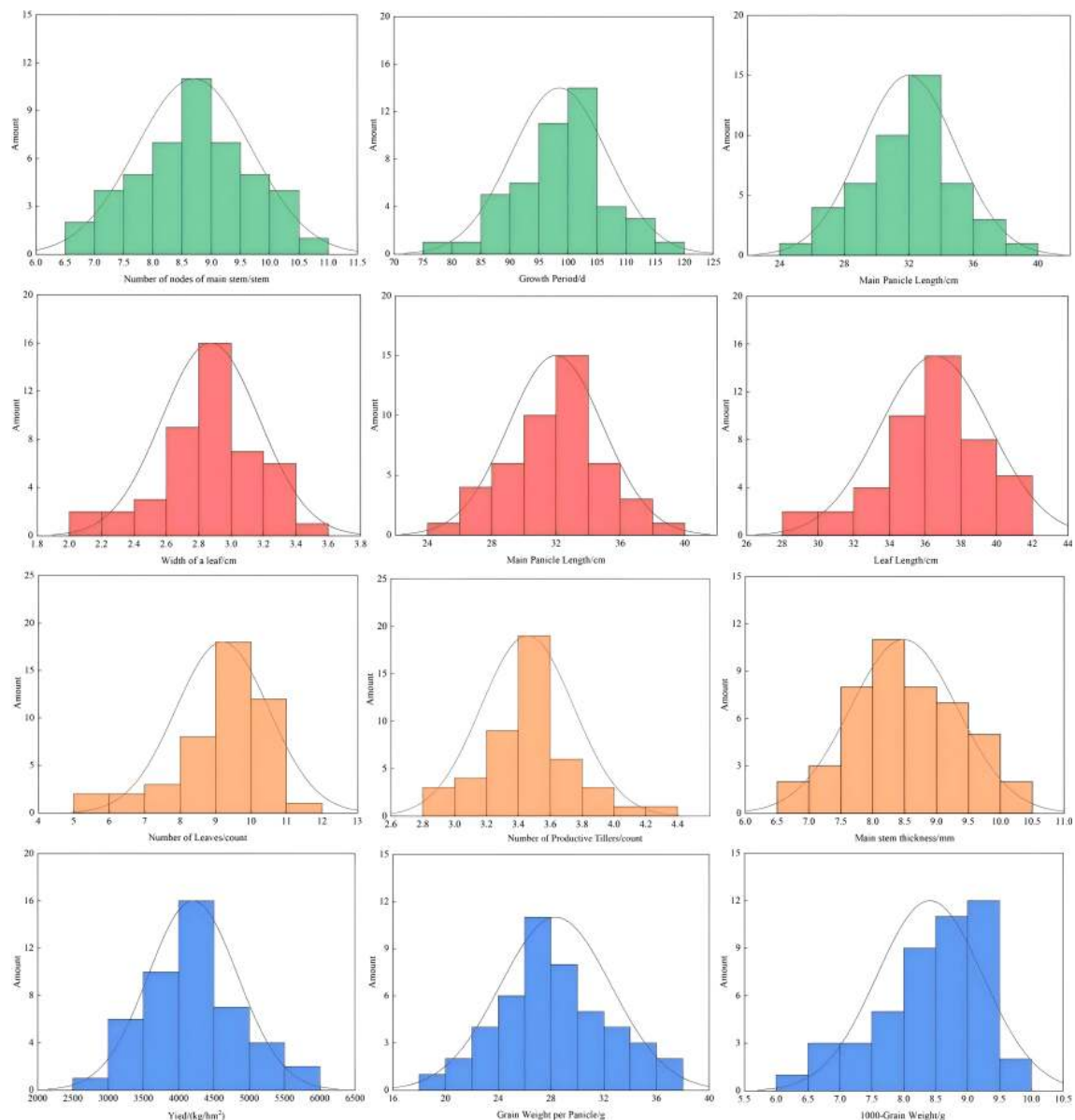


FIGURE 4

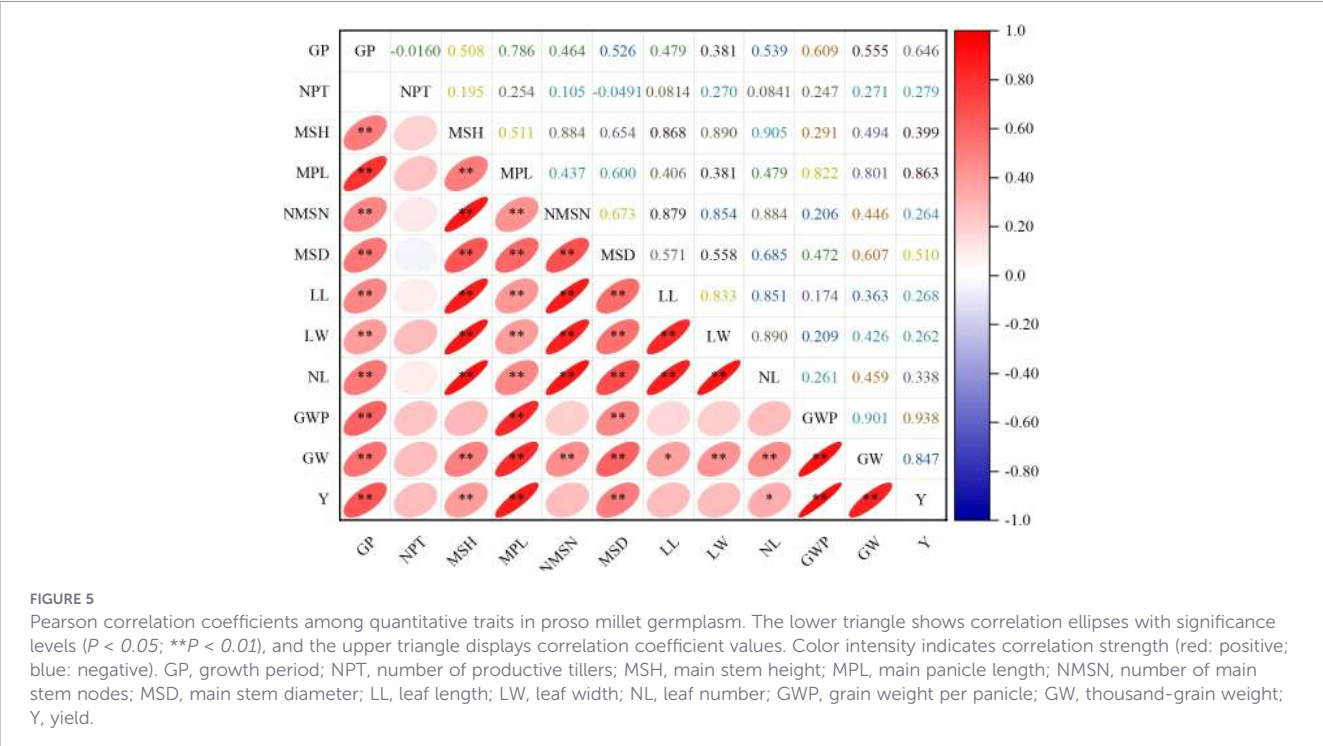
Frequency distribution and phenotypic variation of 12 quantitative traits in proso millet germplasm. Histograms show the distribution patterns with normal distribution curves overlaid. Traits displayed include: growth period (GP, days), main stem height (MSH, cm), main panicle length (MPL, cm), number of main stem nodes (NMSN), main stem diameter (MSD, mm), leaf length (LL, cm), leaf width (LW, mm), number of productive tillers (NPT), number of leaves (NL), grain weight per panicle (GWP, g), thousand-grain weight (GW, g), and grain yield (Y, kg·ha⁻¹).

arid conditions. Leaf length (LL) was highly significantly positively correlated with leaf width (LW) ($r = 0.833$, $P < 0.01$) and showed moderate positive correlations with main panicle length and grain weight per panicle ($r \approx 0.36$ – 0.41 , $P < 0.05$), indicating a modest contribution of flag leaf area to grain filling capacity. Growth period (GP) exhibited weak correlations with most yield traits ($r < 0.20$), suggesting that early-maturing genotypes can be selected without substantial yield penalty, a particularly valuable finding for breeding short-season adapted cultivars.

3.6 Principal component analysis

Principal component analysis (PCA) was applied to the 12 quantitative traits to identify the major axes of phenotypic variation

among the 46 accessions (Table 3). The first four principal components collectively explained 86.69% of total phenotypic variation. PC1 accounted for the largest proportion of variance (45.97%), with high positive loadings for main stem height (0.400), number of main stem nodes (0.403), main stem diameter (0.388), and number of leaves (0.395), representing a “Vegetative Vigor” factor that captures variation in overall plant architecture and biomass accumulation. PC2 contributed 24.16% of variance and was dominated by grain weight per panicle (0.570), thousand-grain weight (0.550), and grain yield (0.560), clearly representing a “Yield Capacity and Grain Size” axis that distinguishes high-yielding accessions based on sink strength and grain filling efficiency. PC3 accounted for 9.12% of variance, with high positive loadings for leaf length (0.578) and main panicle length (0.503), representing a “Panicle and Leaf



Architecture” factor that reflects morphological adaptation related to light interception and photosynthetic capacity. PC4 contributed 7.44% of variance and was dominated by growth period (0.837), representing a “Phenological Maturity” axis that separates early- from late-maturing genotypes. Together, PC1 and PC2 explained 70.13% of total variation, confirming that vegetative vigor and yield components are the two dominant dimensions of diversity in this collection.

TABLE 3 Principal component analysis (PCA) of 12 quantitative traits in 46 proso millet germplasm accessions: component loadings, eigenvalues, and variance contribution rates (combined 2023–2024 data).

Trait	PV1	PV2	PV3	PV4
Growth Period	0.131	0.188	0.271	0.837
No. of Productive Tillers	-0.322	-0.019	0.403	-0.064
Main Stem Height	0.400	0.016	0.120	-0.173
Main Panicle Length	0.314	-0.105	0.503	0.051
No. of Main Stem Nodes	0.403	-0.001	-0.194	-0.116
Main Stem Diameter	0.388	0.039	-0.199	0.056
Leaf Length	0.290	-0.055	0.578	-0.268
Leaf Width	0.259	-0.081	-0.218	0.345
Number of Leaves	0.395	0.008	-0.171	-0.155
Grain Weight per Panicle	0.016	0.570	-0.023	-0.044
1000-Grain Weight	0.009	0.550	0.091	-0.164
Yield	0.000	0.560	-0.012	-0.042
Eigenvalue	5.639	2.964	1.119	0.912
Variance contribution (%)	45.969	24.160	9.123	7.435
Cumulative contribution (%)	45.969	70.129	79.252	86.687

Comprehensive evaluation scores (F-values) based on weighted PC contributions were calculated to rank the 46 accessions by overall agronomic performance (Table 4). The top-ranked accession was Yishu 13 (F = 1.785), followed by Chishu 5 (F = 1.404), Yanshu 11 (F = 1.304), Gumi 23 (F = 1.288), and Yimi 21 (F = 1.267). Among the top 10 ranked accessions, the majority originated from Inner Mongolia and Xinjiang, with several accessions from Gansu and Ningxia also represented. These top-performing genotypes combine superior yield potential with favorable vegetative architecture, making them priority candidates for direct utilization in breeding programs. Conversely, the lowest-ranked accessions included Habahe 2 (F = -3.799), Jinshu 10 (F = -2.871), and Tuoli 1 (F = -1.407), which were characterized by reduced yield components and shorter growth periods, reflecting alternative adaptive strategies suited to specific agro-ecological niches.

3.7 Cluster analysis

Hierarchical cluster analysis (UPGMA method, squared Euclidean distance) based on the 12 quantitative traits classified the 46 accessions into three phenotypically and functionally distinct groups at a Euclidean distance threshold of 10 (Figure 6).

Group I contained 10 accessions (21.74%) representing the elite high-yield cluster, with mean grain weight per panicle (34.26 g), thousand-grain weight (9.23 g), and grain yield (5,112.90 kg·ha⁻¹) all significantly exceeding the overall population averages of 28.09 g, 8.18 g, and 4,143.94 kg·ha⁻¹, respectively. These materials predominantly originated from Inner Mongolia and are the most suitable candidates as elite parental lines for dual-purpose breeding programs targeting simultaneous grain yield and forage biomass improvement.

Group II comprised 10 accessions (21.74%) distinguished by markedly shorter growth periods (76–90 days) and reduced plant

TABLE 4 Comprehensive evaluation scores (F-values) and ranking of all 46 proso millet germplasm accessions based on weighted principal component analysis of 12 quantitative traits (combined 2023–2024 data).

Rank	No.	Variety	F-value	Rank	No.	Variety	F-value
1	21	Yishu 13	1.785	24	6	Ningmi 10	0.176
2	28	Chishu 5	1.404	25	15	Yimi 11	0.127
3	29	Yanshu 11	1.304	26	13	Shanmi 2	0.121
4	4	Gumi 23	1.288	27	37	Longmi 19	0.109
5	18	Yimi 21	1.267	28	41	N1255-2-3	0.033
6	24	Neimi 5	1.259	29	31	Jinshu 8	0.032
7	35	Longmi 15	1.258	30	46	Jishu 8	-0.001
8	19	Yishu 10	1.094	31	44	Jishu 4	-0.245
9	17	Yimi 16	1.091	32	39	N1524-3	-0.249
10	22	Yishu 15	1.056	33	7	Ningmi 17	-0.252
11	20	Yishu 11	1.035	34	42	Jishu 9	-0.273
12	23	Yixuanhuangmi	0.923	35	25	Chimi 3	-0.282
13	16	Yimi 12	0.922	36	34	Longmi 14	-0.344
14	12	Shanmi 1	0.896	37	26	Chishu 9	-0.429
15	5	Gumi 24	0.805	38	3	Gumi 22	-0.464
16	30	Yanshu 13	0.745	39	1	Qishu 1	-1.038
17	14	Yimi 5	0.656	40	38	N1278-1-1	-1.131
18	45	Jishu 7	0.544	41	2	Qishu 3	-1.187
19	43	Jishu 3	0.536	42	10	Tuoli 1	-1.407
20	40	N1191-2-1-3	0.350	43	33	Jinshu 10	-2.871
21	36	Longmi 17	0.341	44	8	Habahe 2	-3.799
22	11	Yilizao 1	0.318	45	32	Jinshu 9	-3.824
23	27	Chishu 3	0.271	46	9	Habahe 1	-3.951

height (105.74–139.32 cm), together with narrower stem diameters and compact panicle architecture. These early-maturing, semi-dwarf materials primarily originated from Xinjiang landraces and exhibited lower mean grain yields than Group I but demonstrated the most stable performance across years, attributed to their short-season phenology that avoids late-season heat and water stress. Their compact architecture confers excellent lodging resistance and mechanical harvestability, traits of direct practical value for modern proso millet production systems in northern Xinjiang.

Group III was the largest cluster, containing 26 accessions (56.52%) characterized by superior vegetative development, including the highest mean values for main stem height (150.88 cm), leaf length (25.10 cm), and leaf width (27.79 mm), as well as high total above-ground biomass. Despite modest grain yields, these materials produced the greatest vegetative biomass among the three groups, making them particularly suitable as foundation material for forage variety breeding and as sources of alleles governing vegetative vigor and stress tolerance. These accessions predominantly originated from Gansu, Ningxia, Shaanxi, Shanxi, and Hebei, reflecting the broad geographical provenance of the collection and the ecological diversity of their source environments. The clear separation of the three groups along both yield-component and phenological axes confirms that the germplasm collection encompasses complementary breeding pools that can be strategically combined, particularly

Group I × Group II crosses to pyramid high yield with early maturity and lodging resistance.

3.8 Genotype × environment interaction analysis

Combined analysis of variance across the two growing seasons (2023 and 2024) revealed highly significant effects of genotype (G) for all twelve agronomic traits ($P < 0.001$), confirming substantial genetic variation in the germplasm collection (Table 5). The environmental main effect (year) was significant ($P < 0.05$) for leaf width, grain weight per panicle, thousand-grain weight, and yield, reflecting year-to-year climatic variation at the experimental site. Crucially, genotype × environment (G×E) interaction effects were non-significant for all traits, indicating that genotype rankings remained stable across years and that the expression of these agronomic traits was largely consistent under the irrigated conditions of northern Xinjiang.

Broad-sense heritability estimates were very high for most traits. Grain weight per panicle exhibited the highest heritability ($H^2 = 0.996$), followed by grain yield ($H^2 = 0.998$) and thousand-grain weight ($H^2 = 0.979$), indicating that these traits are predominantly under genetic control and are highly reliable targets for phenotypic selection. High heritability was also estimated for growth period ($H^2 = 0.999$), main stem height ($H^2 = 0.941$), main panicle length ($H^2 = 0.875$) and main

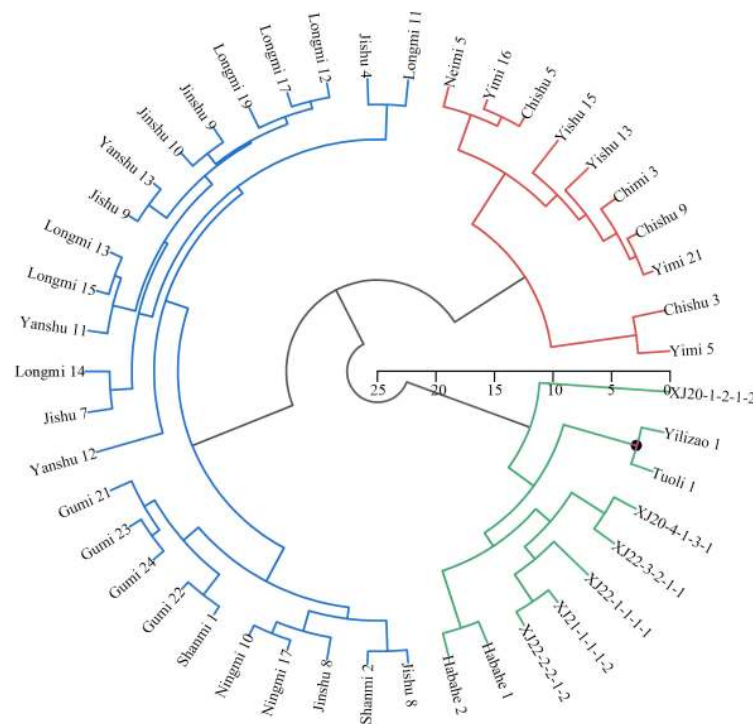


FIGURE 6

Hierarchical cluster analysis of proso millet germplasm based on quantitative trait variation. UPGMA clustering with euclidean distance grouped 46 accessions into three clusters: Group I (21.74%, red branches) with superior yield traits; Group II (21.74%, blue branches) with early maturity and lodging resistance; and Group III (56.52%, green branches) with potential for forage breeding.

stem diameter ($H^2 = 0.910$) (Supplementary Table 2). These heritability estimates support the use of mean performance across environments as a reliable selection criterion for yield-related traits in northern Xinjiang conditions. However, it is important to note that the trial was conducted under irrigated conditions (four uniform irrigations during the growth cycle); therefore, these heritability estimates may not be directly applicable to dryland or rainfed production environments where water stress could amplify G×E interaction effects and reduce the proportion of variance attributable to genotype.

Inter-annual comparison revealed that 2024 was characterized by higher average temperatures and lower precipitation during the grain-filling period compared to 2023, resulting in a mean yield reduction of approximately 6.8% in 2024 relative to 2023. Despite this environmental variation, Group I accessions maintained superior yield performance across both seasons, confirming their stability and suitability as elite parental materials. Several accessions from Group II exhibited notably improved relative performance in 2024, suggesting greater adaptation to warmer, drier growing conditions consistent with their short-season, drought-tolerant phenotype.

3.9 Regression analysis of yield-contributing traits

Stepwise multiple regression analysis identified the key agronomic traits contributing to grain yield variation. Grain weight per panicle (GWP) entered the model first as the single most significant predictor ($R^2 = 0.880$, $P < 0.001$), followed by main panicle length (MPL), which significantly improved model fit ($R^2 = 0.910$, $P <$

0.001). No additional variables met the entry criterion ($P \leq 0.05$), indicating that GWP and MPL together account for the majority of yield variation in this germplasm. The final regression equation was:

$$Y = -2148.3 + 98.6 \times \text{GWP} + 52.4 \times \text{MPL}$$

where Y is grain yield ($\text{kg}\cdot\text{ha}^{-1}$), GWP is grain weight per panicle (g), and MPL is main panicle length (cm). The model demonstrated high predictive accuracy (adjusted $R^2 = 0.906$; standard error of estimate = $187.3 \text{ kg}\cdot\text{ha}^{-1}$), indicating that approximately 91% of grain yield variation among accessions can be reliably predicted from these two panicle-related traits alone.

Standardized regression coefficients confirmed that GWP ($\beta = 0.621$) exerted a substantially greater direct effect on yield than MPL ($\beta = 0.298$), consistent with the higher bivariate correlation between Y and GWP and confirming grain weight per panicle as the dominant yield determinant. The practical implication of this model is that breeders can apply indirect selection for yield through direct measurement of GWP and MPL—traits that are easier to measure early in the season and with higher precision than plot-based yield harvests.

4 Discussion

4.1 Genetic diversity patterns

The 46 proso millet accessions evaluated in this study exhibited substantial phenotypic diversity across both qualitative and

TABLE 5 Analysis of Variance (ANOVA) for 12 quantitative traits of 46 proso millet germplasm accessions evaluated in 2023 and 2024..

Trait	Source of Variation	df	SS	MS	F	Sig.	H ²
Growth Period (d)	Genotype (G)	45	17,765.739	394.794	1214.3	***	0.999
	Year (E)	1	0.000	0.000	0	ns	
	G×E	45	0.000	0.000	0	ns	
	Rep(E)	4	5.478	1.370	4.21	**	
	Error	180	58.522	0.325	—		
	Total	275	17,829.739		—		
No. of Productive Tillers	Genotype (G)	45	28.040	0.623	4.47	***	0.817
	Year (E)	1	0.000	0.000	0	ns	
	G×E	45	0.000	0.000	0	ns	
	Rep(E)	4	0.654	0.163	1.17	ns	
	Error	180	25.096	0.139	—		
	Total	275	53.790		—		
Main Stem Height (cm)	Genotype (G)	45	176,110.056	3,913.557	15095.61	***	0.941
	Year (E)	1	423.761	423.761	0.12	ns	
	G×E	45	11.666	0.259	0	ns	
	Rep(E)	4	14,434.587	3,608.647	14.65	***	
	Error	180	44,341.492	246.342	—		
	Total	275	235,321.562		—		
Main Panicle Length (cm)	Genotype (G)	45	6,838.225	151.961	12629.3	***	0.875
	Year (E)	1	16.952	16.952	0.06	ns	
	G×E	45	0.541	0.012	0	ns	
	Rep(E)	4	1,061.740	265.435	12.27	***	
	Error	180	3,893.134	21.629	—		
	Total	275	11,810.593		—		
No. of Main Stem Nodes	Genotype (G)	45	318.126	7.069	978.17	***	0.91
	Year (E)	1	16.800	16.800	2.59	ns	
	G×E	45	0.325	0.007	0.01	ns	
	Rep(E)	4	25.902	6.475	9.25	***	
	Error	180	126.033	0.700	—		
	Total	275	487.186		—		
Main Stem Diameter (mm)	Genotype (G)	45	326.277	7.251	517.01	***	0.91
	Year (E)	1	17.139	17.139	4.73	ns	
	G×E	45	0.631	0.014	0.02	ns	
	Rep(E)	4	14.501	3.625	5.05	***	
	Error	180	129.257	0.718	—		
	Total	275	487.805		—		
Leaf Length (cm)	Genotype (G)	45	3,217.881	71.508	6922.97	***	0.822
	Year (E)	1	16.337	16.337	0.06	ns	
	G×E	45	0.465	0.010	0	ns	
	Rep(E)	4	1,012.596	253.149	16.4	***	
	Error	180	2,778.512	15.436	—		
	Total	275	7,025.791		—		
Leaf Width (cm)	Genotype (G)	45	32.088	0.713	40.48	***	0.861
	Year (E)	1	13.346	13.346	39.86	**	
	G×E	45	0.793	0.018	0.16	ns	

(Continued)

TABLE 5 Continued

Trait	Source of Variation	df	SS	MS	F	Sig.	H ²
	Rep(E)	4	1.339	0.335	2.99	*	
	Error	180	20.130	0.112	—		
	Total	275	67.696		—		
No. of Leaves	Genotype (G)	45	371.770	8.262	662.07	***	0.926
	Year (E)	1	17.840	17.840	7.52	ns	
	G×E	45	0.562	0.012	0.02	ns	
	Rep(E)	4	9.486	2.371	3.58	**	
	Error	180	119.398	0.663	—		
	Total	275	519.055		—		
Grain Weight per Panicle (g)	Genotype (G)	45	4,636.919	103.043	1615.36	***	0.996
	Year (E)	1	17.023	17.023	1813.35	***	
	G×E	45	2.871	0.064	0.15	ns	
	Rep(E)	4	0.038	0.009	0.02	ns	
	Error	180	74.153	0.412	—		
	Total	275	4,731.002		—		
1000-Grain Weight (g)	Genotype (G)	45	186.649	4.148	105.59	***	0.979
	Year (E)	1	16.360	16.360	681.13	***	
	G×E	45	1.768	0.039	0.45	ns	
	Rep(E)	4	0.096	0.024	0.28	ns	
	Error	180	15.571	0.087	—		
	Total	275	220.444		—		
Yield (kg·hm ⁻²)	Genotype (G)	45	474,240.068	10,538.668	35758.28	***	0.998
	Year (E)	1	155.881	155.881	19.68	*	
	G×E	45	13.262	0.295	0.02	ns	
	Rep(E)	4	31.684	7.921	0.47	ns	
	Error	180	3,055.969	16.978	—		
	Total	275	477,496.864		—		

G, Genotype; E, Year (Environment); G×E, Genotype × Environment interaction; Rep (E), Replication nested within environment. σ^2_g , genotypic variance; σ^2_{ge} = G×E interaction variance; σ^2_e , residual (error) variance. H^2 = broad-sense heritability on a genotype-mean basis: $H^2 = \sigma^2_g / (\sigma^2_g + \sigma^2_{ge}/e + \sigma^2_e/re)$, where $e = 2$ years, $r = 3$ replications. Significance codes: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; ns = not significant ($p > 0.05$). F-value for Genotype (G) is tested against G×E mean square; F for Year (E) is tested against Rep (E) mean square.

quantitative traits, demonstrating the breadth of variation available within Chinese proso millet germplasm for breeding applications. The higher Shannon diversity indices for grain color ($H' = 2.064$) and panicle branch length ($H' = 2.045$) relative to panicle type ($H' = 1.383$; Figure 2) suggest that these traits have experienced less directional selection pressure and retain broader allelic variation, consistent with previous reports on proso millet germplasm diversity in China (Sharma and Niranjana, 2018; Zhang et al., 2023). The predominance of yellow grain color (86.96%; Figure 3) aligns with distribution patterns documented in northern Chinese germplasm collections (Nandini et al., 2026), reflecting both strong consumer market preferences and the ecological adaptation of cultivated types to regional agricultural systems. Similarly, the high frequency of lateral panicle types (71.74%) and long panicle branches (63.04%; Figure 3) is consistent with the morphological adaptation of high-latitude proso millet germplasm (Vetriventhan et al., 2019), where maximizing panicle surface area supports efficient grain filling under the relatively short growing seasons of northern China.

Among quantitative traits, leaf width exhibited the highest coefficient of variation (CV = 24.14%; Table 2), substantially exceeding all other traits including yield-related characters. This high variability likely reflects the sensitivity of leaf width to both genotypic differences and microenvironmental variation during vegetative growth, and suggests that leaf width may be a particularly responsive indicator of genotype × environment interactions in this germplasm. Grain yield and grain weight per panicle also showed considerable variation (CV = 15.00% and 14.61%, respectively), corroborating findings from He et al. (2015) and Lv et al. (2025) and confirming that meaningful yield improvement through phenotypic selection remains feasible within this collection. In contrast to the CV pattern, leaf length demonstrated the highest Shannon diversity index of all quantitative traits ($H' = 2.155$; Table 2), suggesting that while the absolute range of leaf length variation may be moderate, accessions are broadly distributed across phenotypic classes, a pattern indicative of diverse genetic backgrounds rather than clustering around a single adaptive optimum. The relatively stable

growth period across accessions ($CV = 8.23\%$) reflects the phenological adaptation of these cultivated varieties to the specific photoperiod and thermal conditions of their respective regions of origin, consistent with observations in other temperate cereal crops (Ma et al., 2020). Taken together, these diversity patterns indicate that the evaluated accessions capture a broad range of phenotypic variation suitable for identifying superior genotypes for targeted breeding programs.

4.2 Trait correlations and yield formation

The strong positive correlations between grain yield and both grain weight per panicle ($r = 0.901$, $P < 0.01$) and main panicle length ($r = 0.863$, $P < 0.01$) (Figure 5), confirm the 'large panicle, high yield' strategy as the dominant yield-formation pathway in this germplasm, and demonstrate that panicle-based traits are the most actionable selection criteria for yield improvement under northern Xinjiang conditions. This finding aligns with Lv et al. (2025), who identified panicle length as a key trait with the highest loading in the first principal component in a phenotypic diversity study of proso millet germplasm from Xinjiang and neighboring regions, and is consistent with yield component analyses reported for other small-grain cereals (Du et al., 2020; Zhang et al., 2025; Zhou et al., 2025), where panicle architecture is the primary determinant of sink capacity. These relationships are further supported by the stepwise regression model, which explained 91% of yield variation using only two panicle traits (GWP and MPL), providing a highly efficient and practically deployable indirect selection strategy for breeding programs (Wei et al., 2024). The strong positive correlation between thousand-grain weight and grain yield ($r = 0.847$, $P < 0.01$) (Figure 5) further supports a dual strategy of simultaneously improving panicle size and individual grain filling to maximize yield gains. Main stem height exhibited strong positive correlations with main stem node number and main stem diameter ($r \geq 0.884$, $P < 0.01$) but showed a weak positive correlation with productive tillers ($r = 0.195$, not significant) (Figure 5), reflecting a plant architecture that prioritizes main stem growth without a pronounced trade-off with tillering—an adaptive strategy for concentrating photoassimilates in the main panicle under semi-arid, high-irradiance conditions where source-limited grain filling is a primary constraint. This pattern aligns with reports documenting resource allocation strategies in drought-stressed cereals (Stevens et al., 2021).

4.3 Genotype $\times$ environment interactions and heritability

The non-significant genotype $\times$ environment (G $\times$ E) interaction effects detected for all agronomic traits (Table 5) indicate that genotype rankings remained stable across the two growing seasons (2023 and 2024), highlighting the consistency of phenotypic expression under the irrigated conditions of northern Xinjiang. This stability is particularly valuable for breeding programs, as it suggests that selection decisions based on single-year trials can be reliably applied across environments, provided irrigation remains consistent.

The very high broad-sense heritability estimates for yield-related traits, grain weight per panicle ($H^2 = 0.996$), grain yield ($H^2 = 0.998$) and thousand-grain weight ($H^2 = 0.979$), indicate that these traits are predominantly under genetic control and are highly reliable targets for phenotypic selection. High heritability was also observed for growth period ($H^2 = 0.999$), main stem height ($H^2 = 0.941$), main panicle length ($H^2 = 0.875$), and main stem diameter ($H^2 = 0.910$). These values substantially exceed the moderate heritability estimates commonly reported in rainfed cereal trials (Savin et al., 2022), reflecting the favorable irrigated conditions that minimized environmental noise and maximized genetic expression. The consistent superiority of Group I accessions across environments further validates their potential as stable, high-yielding parental materials for northern Xinjiang breeding programs. It should be noted, however, that all trials were conducted under irrigated conditions with four standardized irrigation events; thus, the reported heritability values reflect the genetic contribution to phenotypic variation in the absence of water deficit. Genotype rankings and heritability estimates may differ under dryland or rainfed conditions where G $\times$ E interactions driven by water availability could be substantially larger. Future multi-location, multi-environment trials including dryland settings are recommended to validate the stability of these findings for broader application.

4.4 Germplasm structure and breeding implications

The cluster analysis (Figure 6) revealed three phenotypically and geographically distinct groups that reflect the ecological differentiation of Chinese proso millet germplasm resources across diverse agro-climatic zones, consistent with the geographic structure of genetic diversity documented by Wang et al. (2021) and Lv et al. (2025).

Group I materials originating predominantly from Inner Mongolia demonstrated superior yield components grain yield, grain weight per panicle, and thousand-grain weight all significantly exceeding population means representing the highest-priority candidates for use as elite parental lines in yield-focused breeding programs.

Group II materials from Xinjiang, characterized by shorter growth periods (76–90 days) and semi-dwarf stature (105.74–139.32 cm), are ideally suited for early-maturing cultivar development adapted to high-latitude or double-cropping production systems (Mude et al., 2020). Their compact, mechanically harvestable plant architecture and stable cross-year performance make them particularly valuable for developing commercial varieties compatible with large-scale mechanized production in Xinjiang.

Group III materials, originating from diverse southern and central northern provinces (Gansu, Ningxia, Shaanxi, Shanxi, and Hebei) and characterized by vigorous vegetative development, offer valuable germplasm resources for forage variety breeding, biomass-oriented programs, and as a source of alleles governing stress-adaptive vegetative traits.

Strategic crossing between groups, particularly Group I $\times$ Group II combinations to pyramid high yield with early maturity and lodging resistance, and Group I $\times$ Group III combinations to introduce biomass-associated traits into high-yield backgrounds represents the

most promising near-term avenue for variety development in northern Xinjiang. To further validate and accelerate these breeding strategies, future research should incorporate molecular marker characterization, including SSR and SNP-based genomic tools, to complement phenotypic evaluation, identify QTL associated with key yield traits, facilitate marker-assisted selection, and broaden the genetic base through targeted exotic germplasm introduction (Tyagi et al., 2024, 2025; Lv et al., 2025; Wang et al., 2017; Yang et al., 2025; Yin et al., 2024). Multi-environment trials under dryland and irrigated conditions across additional locations in northern Xinjiang are also strongly recommended to assess genotype stability and refine selection criteria for broader regional deployment.

5 Conclusion

This study systematically evaluated 46 proso millet accessions sourced from seven provinces across two growing seasons (2023–2024) in northern Xinjiang, revealing substantial and actionable genetic diversity at both qualitative and quantitative trait levels. Among qualitative traits, grain color ($H' = 2.064$) and panicle branch length ($H' = 2.045$) exhibited the highest genetic diversity, while panicle type showed the lowest ($H' = 1.383$). Among quantitative traits, grain yield ($CV = 15.00\%$) and grain weight per panicle ($CV = 14.61\%$, $H' = 2.122$) showed the greatest variation, confirming substantial potential for yield improvement through targeted phenotypic selection. The diversity indices for these complex traits should be interpreted in the context of the non-linear nature of the Shannon–Weaver index and the geographic structure of the collection, wherein approximately 60% of accessions originate from three provinces (Inner Mongolia, Gansu, and Xinjiang).

Combined analysis of variance confirmed highly significant genotypic effects for all agronomic traits, while genotype $\times$ environment interaction effects were non-significant for all traits, indicating stable genotype rankings across the two growing seasons under irrigated conditions. Very high broad-sense heritability estimates were observed for yield-related traits grain weight per panicle ($H^2 = 0.996$), grain yield ($H^2 = 0.998$) and thousand-grain weight ($H^2 = 0.979$), demonstrating that these traits are predominantly under genetic control and are highly reliable targets for phenotypic selection under northern Xinjiang conditions. It should be noted, however, that these estimates reflect performance under irrigated management and may not fully extend to dryland or rainfed environments.

Stepwise regression analysis established a highly predictive yield model ($R^2 = 0.91$) using grain weight per panicle ($\beta = 0.621$) and main panicle length ($\beta = 0.298$) as the sole predictors, providing a practical, early-season indirect selection criterion that reduces reliance on laborious plot-level yield harvests.

Cluster analysis identified three functionally distinct germplasm groups: Group I (21.74%; predominantly from Inner Mongolia) with superior yield performance (mean 5,112.90 kg·ha⁻¹), suitable as elite parental lines for dual-purpose breeding; Group II (21.74%; predominantly Xinjiang landraces) with early maturity (76–90 days), compact architecture (105.74–139.32 cm) and stable cross-environment performance, ideal for developing short-season

cultivars adapted to mechanized harvesting; and Group III (56.52%; predominantly from Gansu, Ningxia, Shaanxi, Shanxi, and Hebei) with superior vegetative biomass, valuable for forage breeding and as a repository of stress-adaptive alleles. Strategic crossing among groups, particularly Group I $\times$ Group II-offers the most promising near-term route to combining high yield potential with early maturity and lodging resistance. These findings provide an empirical foundation and prioritized germplasm resources for evidence-based parent selection and targeted breeding program design in support of sustainable and productive proso millet cultivation in northern Xinjiang.

Data availability statement

The original contributions presented in the study are included in the article/[Supplementary Material](#). Further inquiries can be directed to the corresponding authors.

Author contributions

NB: Writing – review & editing. ZYu: Writing – original draft. LJ: Writing – review & editing. ZYa: Methodology, Writing – original draft. HX: Software, Writing – original draft. ZK: Data curation, Writing – original draft. WH: Writing – original draft, Formal analysis. WH-J: Formal analysis, Writing – original draft. YB: Data curation, Writing – original draft. FG: Writing – review & editing, Supervision.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This study was supported by Mining of Excellent Genes and New Variety Breeding of Xinjiang Characteristic Coarse Cereals (2025B02016-1) and Urumqi Comprehensive Experimental Station of the National Millet and Sorghum Industry Technology System (CARS-06-B30).

Conflict of interest

The author(s) declared that this work 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) declared that generative AI was not 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.2026.1792533/full#supplementary-material>

References

- Abdelgawad, H., Avramova, V., Baggerman, G., Van Raemdonck, G., Valkenborg, D., Van Ostade, X., et al. (2020). Starch biosynthesis contributes to the maintenance of photosynthesis and leaf growth under drought stress in maize. *Plant Cell Environ.* 43, 2254–2271. doi: 10.1111/pce.13813
- Aibai, A., Chen, X., Santosh, M., Wu, Y., Deng, X., Wang, Y., et al. (2024). Petrology, geochronology, geochemistry, whole-rock Sr-Nd and zircon Lu-Hf isotopes of the Habahe Intrusion in the Chinese Altai: implications for petrogenesis and tectono-magmatic significance. *Lithos* 478–479, 107646. doi: 10.1016/j.lithos.2024.107646
- Allier, A., Teysseire, S., Lehermeier, C., Moreau, L., and Charcosset, A. (2020). Optimized breeding strategies to harness genetic resources with different performance levels. *BMC Genomics* 21, 349. doi: 10.1186/s12864-020-6756-0
- Anyaocha, C. O., Arogundade, O., Akinkunmi, O. Y., Salaudeen, M. T., Omoloye, A. A., Ospitan, A. O., et al. (2023). Agro-morphological characterization of okra (*Abelmoschus* species) for biotic stress and yield contributing traits under rainfed conditions in South Western Nigeria. *Plant Physiol. Rep.* 28, 92–106. doi: 10.1007/s40502-023-00716-w
- Bhanbhro, N., Wang, H.-J., Bakhsh, Q., Basit, M. F., Song, W., Ullah, U., et al. (2025). TaSnRK2.1-2D contributes to drought tolerance by modulating ROS production in wheat. *Plant Cell Environ.* 48, 6440–6443. doi: 10.1111/pce.15243
- Bhanbhro, N., Wang, H.-J., Yang, H., Xu, X.-J., Jakhar, A. M., Shalmani, A., et al. (2024). Revisiting the molecular mechanisms and adaptive strategies associated with drought stress tolerance in common wheat (*Triticum aestivum* L.). *Plant Stress* 11, 100298. doi: 10.1016/j.stress.2023.100298
- Bhanbhro, N., Xiao, B., Han, L., Amar, M. H., Xu, X., Li, B., et al. (2020). Adaptive strategy of allohexaploid wheat to long-term salinity stress. *BMC Plant Biol.* 20, 210. doi: 10.1186/s12870-020-02423-2
- Dai, J., Cai, X., Jin, J., Ge, W., Huang, Y., Wu, W., et al. (2021). Earliest arrival of millet in the South China coast dating back to 5,500 years ago. *J. Archaeol. Sci.* 129, 105356. doi: 10.1016/j.jas.2021.105356
- Darkwa, K., Agre, P., Olanmi, B., Iseki, K., Matsumoto, R., Powell, A., et al. (2020). Comparative assessment of genetic diversity matrices and clustering methods in white Guinea yam (*Dioscorea rotundata*) based on morphological and molecular markers. *Sci. Rep.* 10, 13191. doi: 10.1038/s41598-020-69925-9
- Du, Y.-L., Xi, Y., Cui, T., Anten, N. P. R., Weiner, J., Li, X., et al. (2020). Yield components, reproductive allometry and the tradeoff between grain yield and yield stability in dryland spring wheat. *Field Crops Res.* 257, 107930. doi: 10.1016/j.fcr.2020.107930
- Elhaik, E. (2022). Principal component analyses (PCA)-based findings in population genetic studies are highly biased and must be reevaluated. *Sci. Rep.* 12, 14683. doi: 10.1038/s41598-022-14395-4
- Gebre, W., Mekbib, F., Tirfessa, A., and Bekele, A. (2025). Phenotypic diversity of sorghum [*Sorghum bicolor* (L.) Moench] genotypes based on qualitative and quantitative traits. *Heliyon* 11, e42378. doi: 10.1016/j.heliyon.2025.e42378
- Greenacre, M., Groenen, P. J. F., Hastie, T., D'Enza, A. I., Markos, A., and Tuzhilina, E. (2022). Principal component analysis. *Nat. Rev. Methods Primers* 2, 100. doi: 10.1038/s43586-022-00184-w
- Han, L., Xiao, C., Xiao, B., Wang, M., Liu, J., Bhanbhro, N., et al. (2020). Proteomic profiling sheds light on alkali tolerance of common wheat (*Triticum aestivum* L.). *Plant Physiol. Biochem.* 138, 58–64. doi: 10.1016/j.plaphy.2019.02.024
- He, C., Liu, Z., Wu, J., Pan, X., Fang, Z., Li, J., et al. (2021). Future global urban water scarcity and potential solutions. *Nat. Commun.* 12, 4667. doi: 10.1038/s41467-021-25026-3
- He, L., Zhang, B., Wang, X., Li, H., and Han, Y. (2015). Foxtail millet: nutritional and eating quality, and prospects for genetic improvement. *Front. Agric. Sci. Eng.* 2, 124–133. doi: 10.15302/J-FASE-2015054
- James, M., Kambham, S., Ceasar, S. A., and Ignacimuthu, S. (2025). “Proso millet,” in *Millet: crops for climate resilience and for food and nutritional security*. Eds. S. A. Ceasar, S. Penna, C. W. P. Carvalho and S. M. Jain (Springer, Singapore), 107–138. doi: 10.1007/978-981-95-1256-0_6
- Loni, F., Ismaili, A., Nakhoda, B., Ramandi, H. D., and Shobbar, Z. (2023). The genomic regions and candidate genes associated with drought tolerance and yield-related traits in foxtail millet: an integrative meta-analysis approach. *Plant Growth Regul.* 101, 169–185. doi: 10.1007/s10725-023-01010-3
- Luo, Y., Liu, C., Dang, K., Gong, X., and Feng, B. (2022). Cultivar sensitivity of broomcorn millet (*Panicum miliaceum* L.) to nitrogen availability is associated with differences in photosynthetic physiology and nitrogen uptake. *Plant Physiol. Biochem.* 182, 90–103. doi: 10.1016/j.plaphy.2022.04.002
- Lv, Z., Yang, Y., Hou, H., Yang, S., Cui, Z., Zhang, X., et al. (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
- Ma, K., Jia, Z., Wen, X., Chen, X., Lu, H., and Chen, F. (2025a). Genetic gain of yield and agricultural phenotype in foxtail millet from 1970s to 2020s in China. *Eur. J. Agron.* 170, 127780. doi: 10.1016/j.eja.2025.127780
- Ma, K., Yuan, X., Jia, Z., Lu, H., Chen, X., Wen, X., et al. (2025b). Changes in the grain quality of foxtail millet released in China from the 1970s to the 2020s. *Food Res. Int.* 209, 116316. doi: 10.1016/j.foodres.2025.116316
- Mude, L. N., Mondam, M., Gujjula, V., Jinka, S., Pinjari, O. B., Yellodu Adi Reddy, N., et al. (2020). Morpho-physiological and biochemical changes in finger millet [*Eleusine coracana* (L.) Gaertn.] under drought stress. *Physiol. Mol. Biol. Plants* 26, 2151–2171. doi: 10.1007/s12298-020-00909-9
- Mukami, A., Ngetich, A., Mweu, C., Oduor, R. O., Muthangya, M., and Mbinda, W. M. (2019). Differential characterization of physiological and biochemical responses during drought stress in finger millet varieties. *Physiol. Mol. Biol. Plants* 25, 837–846. doi: 10.1007/s12298-019-00679-z
- Nandini, C., Chandra Joshi, D., Mani, V., Nandini, B., Niranjan Kumar, B., and Meenakshi, J. (2026). “Proso millet (*Panicum miliaceum*): the prodigious millet for diverse uses,” in *Fasting superfoods*. Eds. A. Kumar, J. K. Tiwari, D. Chandra Joshi and R. K. Varshney (Springer, Singapore). doi: 10.1007/978-981-95-2762-5_8
- Palchetti, E., Moretta, M., Calamai, A., Mancini, M., Dell'Acqua, M., Brilli, L., et al. (2023). Effects of nitrogen fertilization and plant density on proso millet (*Panicum miliaceum* L.) growth and yield under Mediterranean pedoclimatic conditions. *Agriculture* 13, 1657. doi: 10.3390/agriculture13091657
- Pavithra, T., and Rawat, S. (2024). “Proso millet: biology, functional potential and sustainable utilization,” in *Millet: the multi-cereal paradigm for food sustainability*. Ed. M. Thakur (Springer, Cham). doi: 10.1007/978-3-031-64237-1_11
- Pei, D., Song, S., Kang, J., Zhang, C., Wang, J., Dong, T., et al. (2023). Characterization of simple sequence repeat (SSR) markers mined in whole grape genomes. *Genes* 14, 663. doi: 10.3390/genes14030663
- Pélabon, C., Hilde, C. H., Einum, S., and Gamelon, M. (2020). On the use of the coefficient of variation to quantify and compare trait variation. *Evol. Lett.* 4, 180–188. doi: 10.1002/evl3.171
- Ren, J., Yu, X., Wang, X., Wang, Y., Xin, X., Wang, R., et al. (2024). The genetic diversity assessment of broomcorn millet (*Panicum miliaceum*) and the construction of a mini-core collection. *Agronomy* 14, 2226. doi: 10.3390/agronomy14102226

- Rivero, R. M., Mittler, R., Blumwald, E., and Zandalinas, S. I. (2022). Developing climate-resilient crops: improving plant tolerance to stress combination. *Plant J.* 109, 373–389. doi: 10.1111/tpj.15483
- Rizwan, M., Rani, R., Rafi, A., and Shimelis, H. (2025). “Genotype $\times$ environment interaction: molecular basis and environmental adaptation of demand-led crop varieties,” in *Plant breeding 2050*. Eds. P. M. Priyadarshan and R. Ortiz (Springer, Singapore), 163–183. doi: 10.1007/978-981-95-0583-8_8
- Savin, R., Cossani, C. M., Dahan, R., Ayad, J. Y., Albrizio, R., Todorovic, M., et al. (2022). Intensifying cereal management in dryland Mediterranean agriculture: rainfed wheat and barley responses to nitrogen fertilisation. *Eur. J. Agron.* 137, 126518. doi: 10.1016/j.eja.2022.126518
- Sharma, N., and Niranjana, K. (2018). Foxtail millet: properties, processing, health benefits, and uses. *Food Rev. Int.* 34, 329–363. doi: 10.1080/87559129.2017.1290103
- Stevens, C. J., Shelach-Lavi, G., Zhang, H., Teng, M., and Fuller, D. Q. (2021). A model for the domestication of *Panicum miliaceum* (common, proso or broomcorn millet) in China. *Veg. Hist. Archaeobot.* 30, 21–33. doi: 10.1007/s00334-020-00804-z
- Tyagi, A., Mir, Z. A., Almalki, M. A., Deshmukh, R., and Ali, S. (2024). Genomics-assisted breeding: a powerful breeding approach for improving plant growth and stress resilience. *Agronomy* 14, 1128. doi: 10.3390/agronomy14061128
- Tyagi, V., Pragma, Chandra, P., Cahoon, R. S., and Lohwasser, U. (2025). “Plant germplasm introduction and exchange activities,” in *Textbook of plant genetic resources*. Eds. K. Tripathi, V. Gupta, G. Krishnappa, J. Kumari and G. P. Singh (Springer, Singapore), 71–96. doi: 10.1007/978-981-96-5024-8_4
- Varraso, R., Garcia-Aymerich, J., Monier, F., Le Moual, N., De Batlle, J., Miranda, G., et al. (2012). Assessment of dietary patterns in nutritional epidemiology: principal component analysis compared with confirmatory factor analysis. *Am. J. Clin. Nutr.* 96, 1079–1092. doi: 10.3945/ajcn.112.038109
- 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
- Wang, C., Hu, S., Gardner, C., and Lübberstedt, T. (2017). Emerging avenues for utilization of exotic germplasm. *Trends Plant Sci.* 22, 624–637. doi: 10.1016/j.tplants.2017.04.002
- Wang, X.-H., Lee, M.-C., Choi, Y.-M., Kim, S.-H., Han, S., Desta, K. T., et al. (2021). Phylogeography and antioxidant activity of proso millet (*Panicum miliaceum* L.). *Plants* 10, 2112. doi: 10.3390/plants10102112
- Wei, Z., Zhang, Y., and Jin, W. (2024). Yield gap analysis of super high-yielding rice (>15 t ha⁻¹) in two ecological regions. *Agriculture* 14, 491. doi: 10.3390/agriculture14030491
- Wimalasiri, E. M., Ashfold, M. J., Jahanshahi, E., Walker, S., Azam-Ali, S. N., and Karunaratne, A. S. (2023). Agro-climatic sensitivity analysis for sustainable crop diversification; the case of proso millet (*Panicum miliaceum* L.). *PLoS One* 18, e0283298. doi: 10.1371/journal.pone.0283298
- Yang, P., Lu, L., Condric, A., Muni, G. A., Scranton, S., Xu, S., et al. (2025). Innovative approaches for engineering the seed microbiome to enhance crop performance. *Seeds* 4, 24. doi: 10.3390/seeds4020024
- Yin, M., Tong, X., Yang, J., Cheng, Y., Zhou, P., Li, G., et al. (2024). Dissecting the genetic basis of yield traits and validation of a novel quantitative trait locus for grain width and weight in rice. *Plants* 13, 770. doi: 10.3390/plants13060770
- Yu, Z., Guindani, M., Grieco, S. F., Chen, L., Holmes, T. C., and Xu, X. (2022). Beyond t test and ANOVA: applications of mixed-effects models for more rigorous statistical analysis in neuroscience research. *Neuron* 110, 21–35. doi: 10.1016/j.neuron.2021.10.030
- Zhang, F., Ning, C., Scott, A., Fu, Q., Björn, R., Li, W., et al. (2021). The genomic origins of the Bronze Age Tarim Basin mummies. *Nature* 599, 256–261. doi: 10.1038/s41586-021-04052-7
- Zhang, L., Ma, K., Zhao, X., Zhang, X., Jia, Z., Wen, X., et al. (2023). Development of a comprehensive quality evaluation system for foxtail millet from different ecological regions. *Foods* 12, 2545. doi: 10.3390/foods12132545
- Zhang, Q., Sun, J., Wang, L., Chen, J., Ke, J., and Wu, L. (2025). Effects of nitrogen application at different panicle development stages on the panicle structure and grain yield in hybrid indica rice cultivars. *Agronomy* 15, 595. doi: 10.3390/agronomy15030595
- Zhang, Q., Zhang, D., Yu, K., Xia, Y., Zhang, J., Zhao, K., et al. (2021). Frequent germplasm exchanges drive the high genetic diversity of Chinese-cultivated common apricot germplasm. *Hortic. Res.* 8, 215. doi: 10.1038/s41438-021-00650-8
- Zhang, X., Zhao, L., Xu, X., Zhu, J., Qu, S., and Wang, S. (2026). The influence of environmental factors on anthocyanin biosynthesis in fruits. *Plant Sci.* 363, 112866. doi: 10.1016/j.plantsci.2025.112866
- Zhao, Y., Liu, J., Kuerban, Z., Wang, H., Yang, B., Wang, H.-J., et al. (2025). Drought tolerance evaluation and classification of foxtail millet core germplasms using comprehensive tolerance indices. *Life* 15, 1485. doi: 10.3390/life15091485
- Zhou, N., Sun, T., Zhang, Y., Shi, Q., Zhou, Y., Xiong, Q., et al. (2025). Comparative analysis of yield and grain-filling characteristics of conventional rice with different panicle types in response to nitrogen fertilization. *Agronomy* 15, 1858. doi: 10.3390/agronomy15081858