

# Assessment and phenotypic identification of millet germplasm (*Setaria italica* [L.]) in Liaoning, China

**Xin-tong LI**

GuiZhou University of Finance and Economics

**Wei-Feng HE**

Liaoning Academy of Agricultural Science

**Hong-hao WANG**

Liaoning Academy of Agricultural Science

**Min XU** (✉ [shumin690101@163.com](mailto:shumin690101@163.com))

Liaoning Academy of Agricultural Science

---

## Article

### Keywords:

**Posted Date:** October 17th, 2023

**DOI:** <https://doi.org/10.21203/rs.3.rs-3434734/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

**Additional Declarations:** No competing interests reported.

---

# Abstract

Millet (*Setaria italica* [L.]) originated in China and is one of the oldest cultivated crops in the world. The current study aims to assess the present millet germplasms in Liaoning Province scientifically, in order to provide a theoretical basis for the collection, preservation, innovation, and use of millet germplasm resources. This study was conducted from 2018 to 2020. A total of 105 millet germplasm collections were selected from the germplasm resource bank of the Liaoning Academy of Agricultural Sciences. Thirty-one agronomic traits, including 19 qualitative traits and 12 quantitative traits, were observed and recorded. Data were comprehensively evaluated via multivariate analyses, including diversity analysis, correlation analysis, principal component analysis, grey relational analysis, and cluster analysis. Results show that millet resources have abundant diversity, traits represent ear size and ear-bearing capacity were found to be closely related to yield. In high-yield breeding, the resources with upthrow seedling leaves, more tillers, more branches, large ears, and long growth periods should be selected as cross parents. Meanwhile, one early-maturing and lower plant resource, one early-maturing and small-grain resource, one later-maturing and higher plant and big-ear resource, and one later-maturing and big-grain resource were screened for excellent millet breeding.

## 1. Introduction

Millet (*Setaria italica* [L.]) is a diploid ( $2n = 2X = 18$ ) species of grass (*Gramineae*, *Setaria*). It is one of the oldest cultivated crops in the world, with 5000 to 8000 years of cultivation history<sup>[1]</sup>. Millet, which originated in China, has the characteristics of drought resistance, water saving, high light efficiency, easy storage, and both grain and grass usage<sup>[2]</sup>. The millet shell, also called “small grain” is rich in balanced nutrients<sup>[3]</sup>, including abundant proteins<sup>[4]</sup> and vitamins<sup>[5]</sup> thus, it is often the first diet choice for new mothers or patients in recovery.

Research on germplasm resources is the basis of any crop study. According to statistics, more than 40,000 millet resources are found worldwide. Meanwhile, China has the most abundant germplasm resources, with more than 26,000 resources stored in the National Crop Germplasm Bank<sup>[4]</sup>, thus accounting for 70% of the world's resource holdings<sup>[6]</sup>. In order to scientifically evaluate and use millet germplasm resources, previous researchers have conducted research from many perspectives and have made excellent progress. Most of these studies involve genetic diversity analysis of phenotypic traits. For example, Wang *et al.*<sup>[7]</sup> comprehensively evaluated 15 phenotypic traits in 878 millet resources around the world; eight traits, including leaf sheath color, ear length, seed color, and grain color, were identified as main indicators for phenotype identification. Tian *et al.*<sup>[8]</sup> studied the genetic diversity of 482 millet varieties in Henan and Shandong provinces and found that the diversity level of millet breeding varieties was significantly lower than those of local varieties; this result indicates that certain traits were much more concerned on the breeding process. Li *et al.*<sup>[9]</sup> found that among the multiple phenotypic traits of millet varieties, only some traits, including 1000-seed weight and leaf sheath color, did not differ significantly among different ecological areas, whereas the correlation between varieties and ecological areas was significantly. Molecular biotechnology has also been used in the study of millet germplasm resources. For example, Yang *et al.*<sup>[10]</sup> and Schontz *et al.* found rich genetic diversity of millet resources from different regions. Jia *et al.*<sup>[12]</sup> performed genome resequencing to complete the diversity analysis of 916 millet core germplasms, clarifying the geographical distribution of millet genetic resources, all these results are of great significance for the study of other germplasm resources such as millet and bristle grass.

Millet is a light-and-temperature-sensitive crop<sup>[3]</sup>. According to Liu *et al.*<sup>[5]</sup>, the dominant millet-producing areas are steadily spread in north and east part of China. Being important millet-producing area in northeast China, millet produced in Liaoning is famous for its golden-red grain color, nice taste, and good quality. In this experiment, 105 millet resource materials stored in the Germplasm Bank of Liaoning Academy of Agricultural Sciences (LAAS) were selected, based on 19 quality and 12 quantitative agronomy traits. the selected resources were analyzed using multiple analysis

methods and comprehensive assessments. Compared with molecular markers of millet resources, more simpler and intuitive indicators are needed to evaluate and select resource materials in actual production. Therefore, we carried out this experiment to fully understanding basic situation of millet germplasm resources in Liaoning, select agronomy traits as intuitive indicators, screen special resources, to promote the local millet scientific research and industrial development.

## 2. Results and analysis

### 2.1 Phenotypic diversity analysis of millet germplasm resources

As shown in Table 1, among the 19 tested quality traits, 10 exhibited all the features found in Descriptors and Data Standard for Millet (*S. italica* [L.]) <sup>[13]</sup>, including Leaf color of seedling (LCS), branch habit (BH), and bristle length (BL). The remaining 9 traits exhibited most of the features, including leaf sheath color (LSC), ear shape (ES) and bristle color (BC). By comparing the frequency distribution of phenotypes of different traits, it was found that two traits, tiller habit (TH) and branch habit (BH), showed weak, medium and strong characteristics with uniform distribution, and DI were higher. The phenotypic distribution of the other 17 traits showed obvious tendencies, among which grain color (GC) and seed color (SC) were mainly yellow, and DI were lower.

Table 1 Phenotypic diversity of 19 quality traits

[Table 1 near here].

Phenotypic diversity analysis of 12 quantitative traits was showed in Table 2. The coefficient of variation CV represents the degree of dispersion of traits, CV of the 12 traits was ranged from 6.34% to 43.77%, the highest CV was obtained from stem number per plant (SNP), followed by the main stem nod number (MSN), seed weight per plant (SWP), and ear weight per plant (EWP) indicates high degree of variation and high potential for genetic improvement; the CV of peduncle length (PL) and growing period (GP) were close or less than 10%, indicates lower dispersion and relatively stable performance among the varieties. The Shannon diversity index (DI) represents the distribution of trait performance, that of the 12 traits ranged between 1.0566 and 2.0428, and those of main stem length (MSL), ear weight per plant (EWP), seed weight per plant (SWP), and grass weight per plant (GWP) were close to or greater than 2.0, indicates the four traits easy to vary according to external conditions and are relatively rich in phenotype. Meanwhile, the DI of stem number per plant (SNP) was the lowest, close to 1.0, indicating a simple genetic basis and relatively stable performance.

Table 2 Phenotypic diversity of 12 quantitative traits

[Table2 near here]

### 2.2 Correlation analysis of millet germplasm resources

Correlation analysis aims to study the relationship between two or more random variables of equal status. In this experiment, correlation analysis was performed based on the 12 quantitative traits (Table 3). EWP showed positive correlations with SWP extremely significant, while EWP and SWP showed positive correlated with MEL, MED and SNP extremely significant, positive correlated with MSL, GP extremely significant or significant respectively, negative correlation with MSN extremely significant. In addition, GP exhibited extremely significant positive correlations with SNP, MSL, and MEL as well as significant positive correlations with PL and MSD. GWP showed extremely significant positive correlations with MSN and significant positive correlations with MSD. By contrast, the correlation of TSW with all traits was not significant. For the quality traits, BH, TH, SD, SH, and LR were correlated with yield. For growth status, BH, TH

were correlated with SLA. For pigment deposition, SC and GC, STC and AC and PGC performed certain correlation, respectively.

[Figure 1 near here]

Note:  $r_{0.05}=0.1918$   $r_{0.01}=0.2504$

**Fig. 1:** Correlation analysis of 12 quantitative traits

**2.3 Principal component analysis (PCA) of quantitative traits**

Principal component analysis is a comprehensive transformation of multiple trait indexes and dimensionality reduction into several principal components. Each principal component is a relatively independent indicator system, and there is no correlation between each principal component and the numerical value is intuitive and easy to analyze<sup>[14]</sup>. The PCA results of 12 quantitative traits from 105 millet germplasms are shown in Table 4. Six principal components with eigen values larger than 1 or close to 1 were selected, with cumulative contribution rate of 74.94%, thus covering most of the genetic information of millet germplasms. The eigen values of each component were then compared. For the first principal component, which is the GWP factor, the relationships between eigenvalues show that as MEL increased, MSL, TSW, and GP increased too, but GWP decreased. The second principal component is the MSD factor, the relationship shows that GP and TSW increased with increasing MSD. The third principal component is the SNP factor, the relationship indicates that EWP and SWP increased as SNP increased. The fourth principal component is the PL factor, the relationship shows that as PL increased, GP, SWP, and EWP also increased. The fifth principal component is the TSW factor, the relationship indicates that as MSL increased, MEL and GWP also increased, but TSW decreased. The sixth principal component is the MED factor, and the relationship indicates that as MED increased, MSL and MEL decreased. Overall, PCA allowed further interpretation of the correlations between quantitative traits, with the focus on growth period and TSW (1000-seed weight).

**Table 3** Principal components and eigen value of 12 quantitative traits

[Table 3 near here]

**2.4 Grey relational analysis (GRA) of quantitative traits and yield**

Grey correlation analysis refers to the quantitative description and comparison of the development of a system, to measure the degree of correlation between factors according to the degree of similarity or difference of the development trend between factors. Took SWP as the reference column after standardized processing, the correlation between the main quantitative traits and yield was then analyzed by GRA (Table 5). Among the quantitative traits tested, the larger impact on SWP was from EWP, SNP, MEL, and MED these four traits directly represent the ear-bearing capacity of the plant. The coefficients of three traits, namely, GP, MSL, and PL were close; these three traits mainly represent the duration of the growth period. MSD, GWP, and MSN represent the vegetative growth status of the plant, and their restriction on the yield are decreased in turn.

**Table 4** Grey relational analysis of seed yield and 11 quantitative traits

[Table 4 near here]

**2.5 System cluster analysis of all traits**

Cluster analysis reflects the genetic differences of different varieties, and the clustering of resources with similar traits into one group plays a guiding role in selecting parents in hybridization<sup>[15]</sup>. Selecting resources from different groups

with large genetic distance for hybridization is easy to obtain large genetic variation<sup>[16,17]</sup>.

Based on all 31 traits with different performance (12 quantitative traits + 19 quality traits) of millet resources, SCA was conducted via standardized data transformation - Euclidean distance-deviation square sum method. All materials were divided into six groups at Euclidean distance = 17.09 (Fig. 2).

[Figure 2 near here]

**Fig. 2** System cluster analysis of *Setaria italica* (L.) germplasm

By comparing the quantitative traits of different groups, the results showed that: the first group consisted of 11 samples, which were early-maturing, higher plant, and nod-rich. The second group consisted of 30 samples, which were recorded as early-maturing, less-tiller, higher-TSW, and lower-yield. The third group represented 33 accessions, which were later-maturing, higher plant, higher-TSW, and higher-yield. The fourth group contained six samples, which were later-maturing, higher plant, longer-ear, and high-yielding types. The fifth group contained five samples, which were rich-tiller, draft, shorter-ear, and low-yielding types. The sixth group contained 20 samples, which were rich-tiller, thinner-stem with thick ear, and lower-TSW.

The quality traits of each group were compared too. The results showed own in Table 6: the LCS and BL of the first group exhibited characteristics that were not found in other groups. All the traits of the second group showed common characteristics. The characteristics of the third group were abundant and showed a red SC, which was not found in other groups. The fourth group has a small number of resources, which are relatively simple and concentrated in trait phenotypes; however, a relatively rare purple BC appeared. The fifth group has only five resources, but the trait phenotypes are relatively scattered, GC, SC, and SH all showed characteristics that were not found in other groups. The sixth group showed abundant characteristics, with LR and ES exhibiting characteristics not found in other groups.

**Table 5** System Cluster Analysis of 12 quantitative traits

[Table 5 near here]

Further comparison of the geographical distribution of resources in each group, it showed that among the groups, materials from Liaoning have certain advantages, particularly in Group 4, where six materials were all from Liaoning. group 5 is evenly distributed from Liaoning and Jilin. Materials in groups 3 and 6 were mainly from Liaoning, Beijing, and inner mongolia. Materials in group 2 is relatively scattered and covers almost all geographical origins.

**2.6 K-mean clustering analysis**

K-Mean clustering analysis was further conducted, the theoretical distribution intervals and central values of the 12 quantitative traits are shown in Table 7. For the resources studied, the performance of each trait mainly concentrated near the central value, with more distribution around the minimum value and less around the maximum value.

**Table 6** K-mean cluster analysis of 12 quantitative traits

[Table 6 near here]

Based on the results, resource materials with specific performances in maturity, plant height, branching, ear, and grain size were screened out. After a comprehensive consideration of the different traits, special resources were selected. These resources are one early-mature and draft and small-ear and small-seed resource from Taiyuan, Shanxi Province, named Dungu No.1; one small-seed and early-mature resource from Shijiazhuang, Hebei Province, named Xiao-li-xiang;

one later-mature and taller and large-ear resources from Fuxin, Liaoning Province, named Basen Shengu; and one later-mature and low-yield and large-seed resource from Chifeng, Inner Mongolia, named Yuhuanggu No.1.

### 3. DISCUSSION

All 31 traits examined in this experiment showed rich diversity in both qualitative and quantitative traits. In terms of 19 quality traits, the Shannon DI values of traits representing plant color such as grain color (GC), seed color (SC), leaf color of seedling (LSC), etc., were lower, infer that the pigmentation of local resources had an obvious tendency, mainly green leaves and yellow grains. DI values of traits representing ear characteristics, such as ear compactness (EC), spike density (SD), bristle length (BL), etc., were in the middle, the ear of local resources is mainly sparse and loose, and easy to drop grain. DI of traits representing plant structure, such as tiller habit (TH), branch habit (BH), blooming leaf attitude (BLA), etc., were higher, eaves are raised upward, the bristle are short, the peduncle shape is curved, the ear shape is mainly cylindrical and spindle, and the branching and tiller show various characteristics. For the 12 quantitative traits examined, the DI values were distributed between 1.0566 and 2.0428, and CV ranged from 6.34% to 45.84%. The DI and CV of growing period (GP) and 1000-seed weight (TSW) are small, indicating strong adaptability through long-term selection, and had little potential for genetic improvement this result is similar to that obtained by Gao *et al.*<sup>[18]</sup> on mung bean. Yield traits such as ear weight per plant (EWP), seed weight per plant (SWP), and grass weight per plant (GWP), which represent plant growth status, all have higher CV and DI, indicating these traits were regulated by multiple quantitative genes, susceptible to external conditions, and has great potential for genetic improvement. This conclusion is consistent with the research results of Wang *et al.*<sup>[19]</sup> and Wang *et al.*<sup>[20]</sup>.

Results of the correlation analysis of the quality and quantitative traits show that ear size, stem number per plant (SNP), and growing period (GP) were closely correlated with grain yield. No direct correlation was found between traits represented plant growth status and grain yield. These results are basically consistent with those of Jia *et al.*<sup>[21]</sup>, but differ from those of Yan *et al.*<sup>[22]</sup>, the differences are considered to be caused by differences in the test sites, sampling method, and measurement index, etc. Moreover, branch habit (BH), tiller habit (TH), spike density (SD), shattering habit (SH), and lodging resistance (LR) showed positive correlation with yield, whereas BH and TH were correlated with seedling leaf attitude (SLA). Therefore, high yield lines are expected to be seedling leaf upthrow, more tillers and branches, large ears, composed ear, and long growth periods. Since the correlation of quality traits were completed by assigning values on the basis of their phenotypes, where human factors have significant influence, the analysis results provided a reference in the work only.

Six principal components of the phenotype traits of millet resources were extracted via PCA. These components are the grass weight per plant (GWP), main stem diameter (MSD), stem number per plant (SNP), peduncle length (PL), 1000-seed weight (TSW), and main ear diameter (MED) factors. The first and second components represent the vegetative growth capacity of the plant. The third, fifth and sixth components are indirectly correlated with plant yield. The fourth component is correlated with the growth period of the plant. Compared with correlation analysis, PCA further interpreted the correlation between quantitative traits of plants, and paid attention to the growth period and 1000-grain weight traits. Gray relational analysis (GRA) of 12 quantitative traits based on GWP showed that the yield was mainly limited by ear size and ear-bearing capacity, followed by growth period traits. Overall, the results of PCA, GRA, and correlation analysis are consistent. It is worth noting that TSW does not directly affect the yield, it is speculated that millet has the characteristics of seed shatter, and some of seeds were not fully mature on harvest too. These results are similar to the those obtained by Jia *et al.*<sup>[23]</sup> on Tartary buckwheat.

Using systematic cluster analysis, all materials were divided into six groups at Euclidean distance = 17.09, each group showed different phenotypic characteristics. During the breeding process, materials with large genetic differences and

outstanding performance can be selected as cross mating groups, to obtain the maximum genetic variant<sup>[16]</sup>. At the same time, materials from different origins show certain distribution tendencies in each group. According to the previous experience on adzpea<sup>[24]</sup>, cotton<sup>[25]</sup>, and other crops, the characteristics of different groups have a certain correlation with the natural climate conditions of the original source areas, which can be categorized by cluster analysis. The results of this experiment partially reflected this feature, the reason why the geographical distribution tendency among groups is not obvious is due to the frequent resources introduction between regions and blood mixed.

## 4. Materials and Methods

### 4.1 Experiment materials

A total of 105 millet resources that are being preserved in the Germplasm Bank of Liaoning Academy of Agricultural Sciences were selected. The resources are from Liaoning (52), Beijing (20), Hebei (8), Jilin (9), Inner Mongolia (9), Heilongjiang (2), and Shanxi (5), about half are from Liaoning, the rest from surrounding provinces that north of the Yellow River, can represent the basic characteristics of millet resources in northeast of China. Fig. 1 shows the origins of the resources used in this study.

[Figure 3 near here]

Fig. 3 Geographical origin of millet (*Setaria italica* [L.]) germplasm

### 4.2 Experiment design

The experiment was conducted at the experimental site of Industrial Crops Institute of Liaoning (Liaoyang City, Liaoning Province) from 2018 to 2020 for three growing seasons. The soil is sandy loam, with 1.97% organic matter, 0.08% total nitrogen, 73.4 mg/kg alkali-hydrolyzable nitrogen, 23.6 mg/kg available phosphorus, and 247.5 mg/kg available potassium in the top soil. When sowing, all resources were arranged in sequence, each resource was planted in four rows with 3-4 cm plant spacing, row length is 5 m with 50 cm space, the plot area was 10 square meters. The sowing time was around May 11<sup>th</sup>, and harvesting was done around September 25<sup>th</sup>. Three-compound fertilizer (N:P:K=15%:15%:15%) was applied disposably as the base fertilizer. Ploughing and weeding were conducted thrice during growth period.

Two points were selected from each plot, and 5 consecutive plants with uniform growth situation were selected from each point as samples, quality and quantitative traits were investigated following the Descriptors and Data Standard for Millet (*S. italica* [L.]) compiled by Lu<sup>[13]</sup>. A total of 12 quantitative traits, which include growing period, main stem length, main ear length, and grain weight per plant, and 19 quality traits, which include seedling color, branch habit, ear shape, and grain color, were selected for statistical analysis. The mean value of the 10 plants and the inter-annual mean value were obtained.

### 4.3 Data processing

EXCEL 2007 and DPS 9.5 [26] were used for data collection and analysis. Based on all traits, phenotypic diversity analysis and systematic cluster analysis was conducted, the Shannon diversity index was analyzed by referring to the method of Wang *et al.* [27]. Based on quantitative traits, correlation analyses, principal component analyses, grey relational analysis and K-mean cluster analysis were performed.

Before systematic cluster analysis and grey relational analysis, all data were standardized using the following formula:  $x'_{ij} = (x_{ij} - \bar{x}_j) / S_j$ , thereinto:  $\bar{x}_j = \sum_{i=1}^n x_{ij} / n$ ,  $S_j = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}$ . After standardization, a new data series is

obtained with dimension 1, mean value 0 and variance 1.

## Declarations

### Author Contribution Statement

LI X. T., HE W.F., WANG H.H. performed experiments.

HE W.F. supplied germplasm resources materials.

WANG H.H. analyzed data.

LI X.T. wrote the paper.

XU M. planned and supervised the research.

### Acknowledgments:

We wish to express our sincere gratitude to philosophy and Social Sciences Planning Project Guizhou Grant number:21GZYB62 for financially supported to this study.

### Disclosure statement:

No potential competing or conflict of interest was reported by the author(s).

### Statement

All of the 105 Millet (*Setaria italica* [L.]) resources were selected from the Germplasm Bank of Liaoning Academy of Agricultural Sciences with special permission. The use of it complies with relevant institutional, national, and international guidelines and legislation.

#### **Data availability statement:**

The data that support this study will be shared upon reasonable request to the corresponding author.

## **References**

1. Lu, H. Y. et al. Earliest domestication of common millet (*Panicum miliaceum*) in East Asia extended to 10,000 years ago. *PNAS*. 106(18),7367–372(2009).
2. Yang, X. Y. et al. Early millet use in northern China. *Proceedings of the National Academy of Sciences of the United States of America*. 109,3726–3730 (2012).
3. Diao, X. M. Chinese Industry and Technical System of Foxtail Millet. *China Agricultural Sciences and Technology Press*. 20–30. (Beijing, 2011)
4. Liu, X. Geographic Distribution and Germplasm-rich Region of Cereals in China. *Journal of Plant Genetic Resources*. 10(1),1–8. (2009)
5. Liu, J. A.Spatiotemporal variation and regional advantages of foxtail millet production in recent 30 years in China.*Scientia Agricultura Sinica*. 52(11),1883–1894 (2019).
6. Doust, A., Kellogg, E. A., Devos, K. M., Bennetzen, J. L. Foxtail millet: A sequence-driven grass model system. *Plant Physiology*.149(1),137–141(2009).
7. Wang, H. G. Phenotypic Diversity Evaluations of Foxtail Millet Core Collections. *Acta Agronomica Sinica*. 42(1), 19–30. (2016)
8. Tian, B.H. Genetic diversity of landrace and improves cultivars in foxtail millet. *Journal of Plant Genetic Resources*. 11(2), 224–228. (2010)
9. Li, Y., Wu, S. Z., Cao, Y. S., Zhang, X. Z. A phenotypic diversity analysis of foxtail millet (*Setaria italica* (L.) P. Beauv.) landraces of Chinese origin. *Genetic Resources and Crop Evolution*. 43(4), 377–384(1996).
10. Yang, T. Y. Genetic variation among varieties of foxtail millet from different regions in China based on RAPD markers. *Acta Botanica Boreali-Occidentelia Sinica*. 23(5), 765–770. (2003)
11. Schontz, D., Rether, B. Genetic variability in foxtail millet, *Setaria italica* (L.) P. Beauv. - RFLP using a heterologous rDNA probe. *Plant Breeding*. 117(3),231–234(1998).
12. Jia, G. A haplotype map of genomicvariations and genome-wide association studies of agronomic traits in foxtail millet (*Setaria italica*). *Nature Genetics*. 45, 957–961(2013).
13. Lu, P. Descriptors and data standard for foxtial millet [*Setaria italica* (L.) Beauv.] [M]. *China Agriculture Press*. (Beijing, 2006).
14. Shi, J.J. Genetic diversity analysis and comprehensive evaluation of introduced *Avena sativa* L. germplasm resources in the western area of Heilongjiang. *Journal of Southern Agriculture*. 50(3), 515–523 (2019).
15. Bakhsh, A., Iqbal, S.M., Rahman, M.U. Javaid, A. Use of RAPD markers in comparison with agro-morphological traits for estimation of diversity among chickpea genotypes. *International Journal of Agriculture and Biology*. 19, 427–431(2017).
16. Santhosha, H.M., Varalakshmi, B., Narase Gowda, N.C. Genetic diversity in early cauliflower (*Brassica oleracea* var. botrytis L.) germplasm. *Journal of Horticultural Sciences*.6(1), 21–24(2011).

17. Zhu, S.Y. Principal component analysis and cluster analysis for main morphological characteristics of cauliflower inbred lines. *Journal of Plant Genetic Resources*. 13(1),77–82(2012).
18. Gao, Y. Q., Li, S.T., Shang, Q. B., Liu, H., Xu, D.X. Screening and Evaluation of Germplasm Resources of Mung Bean. *Seed*. 39(08), 66–69 (2020).
19. Wang, X.J., Qi, X.S., Wang, X. R., Su, J. Y. Analysis of genetic diversity of local germplasm resources for foxtail millet in Gansu. *Agricultural Research in the Arid Areas*. 27(6), 129 ~ 133,153 (2009).
20. Wang, Y. L. Evaluation on phenotypic variation of rice bean germplasm collected from Hunan Province. *Journal of Plant Genetic Resources*. 22(2), 317–327(2021).
21. Jia, X.P., Yuan, X.L., Lu, P., Fan, B.Y., Huang, H., Dai, L.F. Grey Incidence Analysis and Comprehensive Evaluation of 71 Millet Germplasm Resources in China. *Seed*. 36(9), 63 ~ 66 (2017).
22. Yan, F. Grey Correlation Degree Analysis and Comprehensive Evaluation of Serotinous Millet Germplasm Resources in Heilongjiang Province. *Heilongjiang Agricultural Sciences*. (7), 30 ~ 33 (2010).
23. Jia, R. L. Genetic Diversity Analysis and Comprehensive Assessment of Agronomic Traits of 64 Tartary Buckwheat Germplasms. *Crops*. (3),19–27(2021)
24. Pu, S.J., Jin, W. L., Zhao, B., Wang, Z. L. Study on Adzuki Bean Variety Sources of North China VII The Genetic Parameters and Correlation Analysis of Agronomic Traits of Adzuki Bean in Beijing Environment. *Journal of Beijing University of Agriculture*. 18(3).174–177 (2003).
25. Xu, M., Hu, Y. S., Li, J. L., Jin, L. L., Wang, Z. S. Clustering and Correlation Analysis of Earlier-Mature Cotton Innovation Germplasm based on Biological Traits. *Crops*.1,25–31 (2017).
26. Tang, Q. Y. DPS data processing system-Experimental Design, Statistical Analysis and data Mining (Second edition). *Science Press*, China. (Beijing, 2010)
27. Wang, Y.L. Evaluation on phenotypic variation of rice bean germplasm collected from Hunan Province. *Journal of Plant Genetic Resources*. 22(2), 317–327(2021).

## Tables

**Table 1 Phenotypic diversity of 19 quality traits**

| Trait                      | Shannon diversity index DI | Distribution Frequency |        |        |        |        |        |        |
|----------------------------|----------------------------|------------------------|--------|--------|--------|--------|--------|--------|
|                            |                            | 1                      | 2      | 3      | 4      | 5      | 7      | 9      |
| Leaf sheath color LSC      | 0.5257                     | 0.7810                 | -      | 0.2190 | -      | -      | -      | -      |
| Leaf color of seedling LCS | 0.2925                     | 0.9143                 | 0.0857 | -      | -      | -      | -      | -      |
| Bristle color BC           | 0.5324                     | 0.0762                 | 0.8476 | 0.0762 | -      | -      | -      | -      |
| Bristle length BL          | 0.7438                     | 0.0476                 | 0.7524 | 0.1810 | 0.0190 | -      | -      | -      |
| Protecting glume color PGC | 0.3673                     | 0.0095                 | 0.8952 | 0.0952 | -      | -      | -      | -      |
| Stigma color STC           | 0.6483                     | 0.2190                 | 0.7524 | 0.0286 | -      | -      | -      | -      |
| Anther color AC            | 0.6830                     | 0.1524                 | 0.0762 | 0.7714 | -      | -      | -      | -      |
| Seed color SC              | 0.1075                     | -                      | 0.9810 | -      | 0.0095 | 0.0095 |        |        |
| Grain color GC             | 0.0943                     | -                      | 0.9810 | -      | 0.0190 | -      | -      | -      |
| Seedling leaf attitude SLA | 0.7825                     | 0.0381                 | 0.6286 | 0.3333 | -      | -      | -      | -      |
| Blooming leaf attitude BLA | 0.9655                     | 0.6667                 | 0.1143 | 0.0476 | 0.1714 | -      | -      | -      |
| Tiller habit TH            | 1.0566                     | 0.3524                 | 0.4381 | 0.2095 | -      | -      | -      | -      |
| Branch habit BH            | 1.0543                     | 0.3429                 | 0.4476 | 0.2095 | -      | -      | -      | -      |
| Peduncle shape PS          | 0.9367                     | -                      | 0.2190 | 0.6095 | 0.1714 | -      | -      | -      |
| Ear compactness EC         | 0.5004                     | 0.0286                 | 0.8476 | 0.1238 | 0.0000 | -      | -      | -      |
| Spike density SD           | 0.5799                     | 0.7333                 | -      | 0.2667 | 0.0000 | -      | -      | -      |
| Ear shape ES               | 1.1206                     | 0.0381                 | 0.4476 | 0.3810 | 0.1333 | -      | -      | -      |
| Shattering habit SH        | 0.7091                     | -                      | 0.3714 | 0.6190 | 0.0095 | -      | -      | -      |
| Lodging resistance LR      | 1.0316                     | -                      | -      | 0.4762 | -      | 0.4000 | 0.1048 | 0.0190 |
| Mean                       | 0.6701                     | 0.3143                 | 0.5554 | 0.2905 | 0.0593 | 0.2048 | 0.1048 | 0.0190 |

Note: Arabic numbers (1–9) in the first line of the table refer to the phenotypic trait according to Descriptors and Data Standard for Millet [*S. italica* (L.)] (Lu, 2006).

**Table 2 Phenotypic diversity of 12 quantitative traits**

| Traits              |                              | Min V. | Max V. | Mean V.   | SD       | CV %  | Shannon diversity index DI |
|---------------------|------------------------------|--------|--------|-----------|----------|-------|----------------------------|
| Growing period GP d |                              | 101.0  | 131.0  | 118.809 5 | 7.535 9  | 6.34  | 1.696 6                    |
| Stem                | Stem number per plant SNP    | 1.0    | 3.0    | 1.771 4   | 0.775 3  | 43.77 | 1.056 6                    |
|                     | Main Stem Length MSL cm      | 90.0   | 170.0  | 133.847 6 | 19.025 2 | 14.21 | 2.042 8                    |
|                     | Main stem diameter MSD mm    | 15.0   | 41.0   | 25.895 2  | 4.235 7  | 16.36 | 1.850 7                    |
|                     | Main stem nod number MSN     | 0.3    | 1.0    | 0.522 4   | 0.150 6  | 28.83 | 1.517 6                    |
| Ear                 | Peduncle length PL cm        | 9.0    | 16.0   | 13.115 4  | 1.430 0  | 10.90 | 1.707 4                    |
|                     | Main ear length MEL mm       | 16.0   | 40.0   | 22.847 6  | 4.018 7  | 17.59 | 1.874 1                    |
|                     | Main ear diameter MED mm     | 0.7    | 4.2    | 2.635 2   | 0.503 6  | 19.11 | 1.764 9                    |
| Yield               | Grass weight per plant GWP g | 9.2    | 119.8  | 47.040 0  | 21.563 6 | 45.84 | 1.951 3                    |
|                     | 1000-seed weight TSW g       | 2.0    | 3.9    | 2.688 6   | 0.350 1  | 13.02 | 1.578 4                    |
|                     | Ear weight per plant EWP g   | 11.5   | 41.1   | 22.511 4  | 5.948 4  | 26.42 | 1.996 7                    |
|                     | Seed weight per plant SWP g  | 9.3    | 27.3   | 15.514 3  | 4.104 3  | 26.45 | 1.957 0                    |

**Table 3 Correlation analysis of 12 quantitative traits**

| Correlation Coefficient | SNP          | MSL          | PL          | MSD          | MSN           | MEL          | MED          | GP           | GWP         | TSW         | EWP          |
|-------------------------|--------------|--------------|-------------|--------------|---------------|--------------|--------------|--------------|-------------|-------------|--------------|
| SNP                     | 1            |              |             |              |               |              |              |              |             |             |              |
| MSL cm                  | 0.023<br>7   | 1            |             |              |               |              |              |              |             |             |              |
| PL cm                   | 0.083<br>4   | 0.428<br>4** | 1           |              |               |              |              |              |             |             |              |
| MSD cm                  | 0.196<br>5   | -0.150<br>6  | 0.082<br>8  | 1            |               |              |              |              |             |             |              |
| MSN                     | -0.169<br>3  | 0.088<br>1   | 0.017<br>7  | -0.075<br>9  | 1             |              |              |              |             |             |              |
| MEL mm                  | 0.223<br>3*  | 0.325<br>5** | 0.194<br>5* | 0.246<br>3*  | -0.215<br>1   | 1            |              |              |             |             |              |
| MED mm                  | 0.262<br>2** | -0.008<br>1  | -0.007<br>3 | 0.180<br>9   | -0.093<br>5   | 0.005<br>1   | 1            |              |             |             |              |
| GP d                    | 0.410<br>5** | 0.302<br>8** | 0.199<br>7* | 0.097<br>8   | -0.095<br>6   | 0.400<br>0** | 0.192<br>0*  | 1            |             |             |              |
| GWP g                   | -0.010<br>7  | 0.081<br>1   | 0.075<br>9  | -0.194<br>5* | 0.287<br>2**  | 0.042<br>7   | -0.105<br>8  | 0.055<br>8   | 1           |             |              |
| TSW g                   | -0.190<br>4  | -0.044<br>6  | -0.027<br>4 | -0.017<br>0  | 0.052<br>5    | -0.028<br>6  | -0.045<br>1  | -0.024<br>2  | -0.050<br>8 | 1           |              |
| EWP g                   | 0.395<br>5** | 0.292<br>7** | 0.182<br>1  | 0.093<br>1   | -0.284<br>6** | 0.461<br>2** | 0.279<br>9** | 0.330<br>3** | -0.160<br>9 | -0.095<br>5 | 1            |
| SWP g                   | 0.253<br>3** | 0.214<br>7 * | 0.177<br>7  | 0.003<br>5   | -0.267<br>2** | 0.376<br>2** | 0.257<br>3** | 0.193<br>8*  | -0.092<br>1 | -0.083<br>9 | 0.826<br>6** |

Note:  $r_{0.05}=0.1918$   $r_{0.01}=0.2504$

**Table 4 Principal components analysis and eigen value of 12 quantitative traits**

| Traits            |        | 1st Principle<br>Component | 2nd Principle<br>Component | 3rd Principle<br>Component | 4th Principle<br>Component | 5th Principle<br>Component | 6th Principle<br>Component |
|-------------------|--------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|
| GP d              |        | 0.441 0                    | 0.521 7                    | 0.099 1                    | 0.651 3                    | -0.081 1                   | -0.163 0                   |
| Stem              | SNP    | -0.067 9                   | -0.247 0                   | 0.946 3                    | 0.070 9                    | 0.052 6                    | -0.015 8                   |
|                   | MSL cm | 0.606 6                    | -0.220 3                   | -0.170 2                   | -0.070 3                   | 0.366 3                    | 0.487 6                    |
|                   | MSD cm | -0.278 7                   | 0.732 4                    | 0.113 6                    | -0.169 5                   | 0.237 3                    | -0.447 6                   |
|                   | MSN    | 0.486 8                    | -0.433 0                   | -0.053 9                   | 0.211 8                    | 0.125 2                    | -0.049 3                   |
| Ear               | PL cm  | 0.164 1                    | -0.058 7                   | -0.114 4                   | 0.717 9                    | 0.106 0                    | -0.194 2                   |
|                   | MEL mm | 0.713 2                    | 0.198 1                    | 0.046 6                    | -0.188 3                   | 0.346 9                    | -0.289 0                   |
|                   | MED mm | 0.032 7                    | 0.724 0                    | 0.195 2                    | 0.071 0                    | 0.191 3                    | 0.542 1                    |
| Yield             | GWP g  | -0.751 5                   | 0.278 5                    | 0.049 6                    | 0.169 5                    | 0.242 2                    | 0.227 1                    |
|                   | TSW g  | 0.538 6                    | 0.487 7                    | 0.207 1                    | -0.016 2                   | -0.477 2                   | 0.253 2                    |
|                   | EWP g  | -0.254 7                   | -0.353 7                   | 0.678 0                    | 0.367 0                    | 0.055 0                    | 0.105 8                    |
|                   | SWP g  | -0.361 8                   | -0.008 1                   | -0.716 6                   | 0.404 3                    | -0.011 9                   | 0.122 4                    |
| Eigen Value       |        | 3.114 9                    | 1.663 4                    | 1.209 1                    | 1.114 5                    | 0.970 8                    | 0.919 9                    |
| Contribution Rate |        | 25.957 3                   | 13.861 3                   | 10.076 0                   | 9.287 4                    | 8.090 0                    | 7.666 2                    |
| Cumulative C.R.   |        | 25.957 3                   | 39.818 6                   | 49.894 6                   | 59.182 0                   | 67.272 0                   | 74.938 2                   |
| Factor Weight     |        | 34.638 5                   | 18.497 4                   | 13.445 5                   | 12.393 5                   | 10.795 5                   | 10.229 5                   |

**Table 5 Grey relational analysis of seed yield and 11 quantitative traits**

| Trait  | Interacting<br>coefficient | Correlation<br>sequence | Trait  | Interacting<br>coefficient | Correlation<br>sequence |
|--------|----------------------------|-------------------------|--------|----------------------------|-------------------------|
| EWP g  | 0.5987                     | 1                       | PL cm  | 0.4168                     | 7                       |
| SNP    | 0.4674                     | 2                       | MSD cm | 0.3915                     | 8                       |
| MEL mm | 0.4575                     | 3                       | GWP g  | 0.3808                     | 9                       |
| MED mm | 0.4332                     | 4                       | TSW g  | 0.3714                     | 10                      |
| GP d   | 0.4169                     | 5                       | MSN    | 0.3476                     | 11                      |
| MSL cm | 0.4168                     | 6                       |        |                            |                         |

**Table 6 System Cluster Analysis of 12 quantitative traits**

| Traits |        | 1 <sup>st</sup> Group | 2 <sup>nd</sup> Group | 3 <sup>rd</sup> Group | 4 <sup>th</sup> Group | 5 <sup>th</sup> Group | 6 <sup>th</sup> Group |
|--------|--------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| GP d   |        | 117.0000              | 113.4333              | 121.4848              | 123.5000              | 118.4000              | 122.1500              |
| Stem   | SNP    | 1.1818                | 1.1333                | 1.9697                | 2.3333                | 2.4000                | 2.4000                |
|        | MSL cm | 143.2727              | 129.2333              | 136.3333              | 143.1667              | 99.8000               | 137.2000              |
|        | MSD cm | 0.4909                | 0.4967                | 0.4955                | 0.7500                | 0.8400                | 0.4750                |
|        | MSN    | 13.7273               | 13.7333               | 12.9091               | 13.0000               | 11.6000               | 12.5500               |
| Ear    | PL cm  | 26.9091               | 25.0000               | 25.5758               | 31.5000               | 26.0000               | 25.5000               |
|        | MEL mm | 21.6364               | 20.6333               | 23.8485               | 33.3333               | 20.0000               | 22.7500               |
|        | MED mm | 49.9636               | 51.8867               | 47.6697               | 58.5167               | 25.9800               | 38.9450               |
| Yield  | GWP g  | 49.9636               | 51.8867               | 47.6697               | 58.5167               | 25.9800               | 38.9450               |
|        | TSW g  | 2.5455                | 2.8100                | 2.7545                | 2.5833                | 2.6200                | 2.5250                |
|        | EWP g  | 18.5909               | 17.7100               | 25.3758               | 27.3667               | 19.7200               | 26.3850               |
|        | SWP g  | 14.3455               | 12.9267               | 17.1697               | 17.9333               | 13.9000               | 16.9850               |

**Table 7 K-mean cluster analysis of 12 quantitative traits**

| Traits |        | Min V. | Frequency | Center V. | Frequency | Max V. | Frequency |
|--------|--------|--------|-----------|-----------|-----------|--------|-----------|
| GP d   |        | 102.21 | 0.1333    | 120.08    | 0.7143    | 127.38 | 0.1524    |
| Stem   | SNP    | 1.00   | 0.3524    | 2.00      | 0.4381    | 3.00   | 0.2095    |
|        | MSL cm | 106.42 | 0.2286    | 131.20    | 0.3810    | 152.49 | 0.3905    |
|        | MSD cm | 0.45   | 0.7333    | 0.66      | 0.2095    | 0.97   | 0.0571    |
|        | MSN    | 10.63  | 0.1524    | 13.21     | 0.6952    | 15.13  | 0.1524    |
| Ear    | PL cm  | 19.63  | 0.1524    | 26.01     | 0.7429    | 34.18  | 0.1048    |
|        | MEL cm | 19.92  | 0.4952    | 24.62     | 0.4476    | 34.33  | 0.0571    |
|        | MED cm | 0.70   | 0.0095    | 2.44      | 0.7619    | 3.36   | 0.2286    |
| Yield  | GWP g  | 29.71  | 0.5619    | 50.79     | 0.3238    | 82.23  | 0.2000    |
|        | TSW g  | 2.43   | 0.5333    | 2.93      | 0.4286    | 3.63   | 0.0381    |
|        | EWP g  | 16.80  | 0.4286    | 24.75     | 0.4000    | 31.56  | 0.1714    |
|        | SWP g  | 11.75  | 0.4381    | 16.96     | 0.4095    | 22.43  | 0.1524    |

## Figures



**Figure 1**

Geographical origin of millet (*Setaria italica* [L.]) germplasm

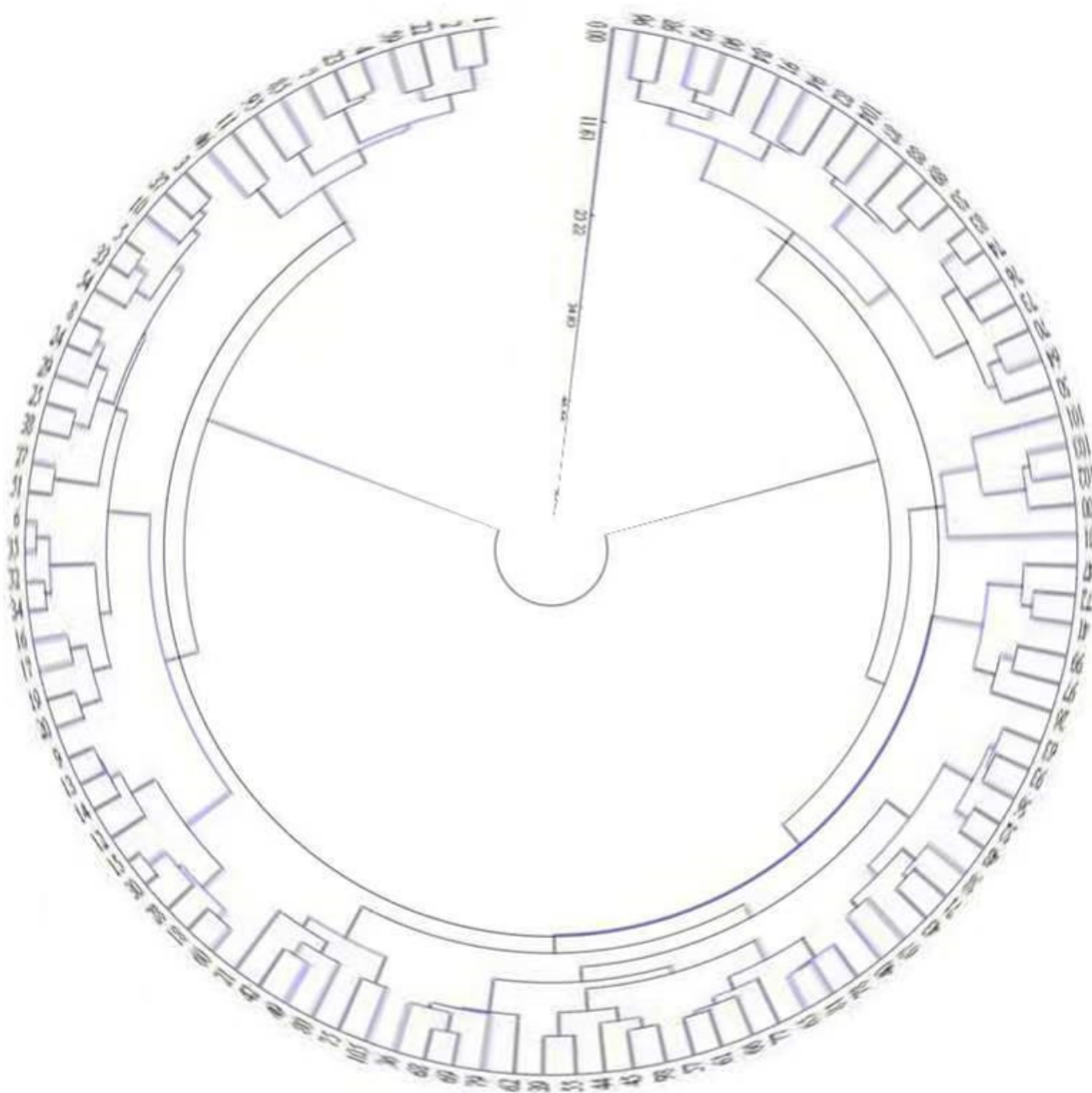


**Figure 2**

System cluster analysis of *Setaria italica* (L.) germplasm

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [GraphicalAbstract.jpg](#)