scientific reports

OPEN

Morphological diversity variation of seed traits among 587 germplasm resources of *Medicago* Genus and 32 germplasm resources of *Trigonella* Genus

Bingcheng Han[1], Xueming Dong[1], Mingyu Li[1], Zhaoming Wang[2], Congcong Shi[1], Qiang Zhou[1], Zhipeng Liu[1] & Longfeng Yan[3]✉

Germplasm resources within the *Medicago* genus are highly regarded for their value as forage crops and their critical roles in nitrogen cycling, ecosystem restoration, and soil structure improvement. Therefore, understanding the diversity of seed morphology in this genus is essential for advancing its development and utilization. This study analyzed seed samples from 587 germplasm accessions representing 77 species within *Medicago* genus, as well as 32 accessions from 21 species within the closely related genus *Trigonella*. A statistical analysis was conducted on twelve traits, including seven quantitative traits-straight length (SL), straight width (SW), width-to-length ratio (WL), perimeter (PE), radicle length (RL), hilum length (HL), and 100-seed weight (SY)-and five qualitative traits, including seed coat condition, radicle characteristics, seed size, shape, and color. The results revealed that: (1) there was significant diversity ($P < 0.05$) in SL, SW, WL, PE, RL, HL, and SY across *Medicago* species; (2) principal component analysis of the 587 *Medicago* accessions identified SL, SW, PE, HL, RL, and SY as the primary contributors to morphological diversity; and (3) high-resolution images of seeds from various accessions were captured for future research. This study provides a solid foundation for the establishment of seed banks and the enhancement of germplasm resources through the systematic analysis of these morphological traits.

Keywords *Medicago* germplasm, Morphological diversity, Variance analysis, Principal component analysis

Abbreviations

SL	Seed length
SW	Seed width
WL	Width to length ratio
PE	Perimeter
HL	Hilum length
RL	Radicle length
SY	100-seed weight

The *Medicago* genus consists of approximately 90 species worldwide, predominantly found in Western Europe, Western Asia, Southeast Asia, and Africa. Most species within this genus are wild herbs, with many being significant as forage crops[1]. In China, *Medicago* species are primarily distributed across 14 provinces, especially in regions north of the Yangtze River and northwest of the Yellow River. Notable species, such as *Medicago sativa* L. (alfalfa), *Medicago falcata* L., and *Medicago polymorpha* L., are recognized for their high yield, palatability, and nutritional value[2]. Wild *Medicago* varieties in China, such as *M. edgeworthii* L., are notable for their strong cold tolerance. For example, *M. edgeworthii* thrives in high-altitude, cold regions like Tibet, Qinghai, and Yunnan,

[1]State Key Laboratory of Herbage Improvement and Grassland Agro-ecosystems, College of Pastoral Agriculture Science and Technology, Lanzhou University, Lanzhou 730020, China. [2]National Center of Pratacultural Technology Innovation (under preparation), Hohhot 010070, China. [3]Ministry of Education Key Laboratory of Cell Activities and Stress Adaptations, School of Life Sciences, Lanzhou University, Lanzhou 730000, People's Republic of China. ✉email: yanlf@lzu.edu.cn

often growing along hillsides[3,4]. Similarly, *M. archiducis-nicolai* L. is found on slopes and grasslands at elevations of 2000–4000 m in areas like Qinghai, Gansu, Ningxia, and Tibet. In contrast, *M. polymorpha* L. is less cold-tolerant and is primarily found in the warmer, more humid regions of Jiangsu, Zhejiang, Sichuan, Hubei, Hunan, and Guizhou[5,6]. As leguminous forage plants, *Medicago* species are not only valuable for their nitrogen-fixing abilities but also exhibit tolerance to cold temperatures, making them integral to agricultural ecosystems. They contribute significantly to nitrogen cycling, soil and water conservation, soil fertility enhancement, and soil structure improvement[7]. Among these species, *M. sativa* (alfalfa) is the most important forage crop[8,9]. Often referred to as the "king of forage," alfalfa is a perennial herbaceous plant rich in minerals such as calcium, phosphorus, potassium, magnesium, and sulfur, as well as essential vitamins[10,11]. Cultivated for over 2000 years in China, alfalfa is now grown globally[12]. In the United States, it ranks as the fourth most valuable crop after *Zea mays* L., *Glycine max* L., and *Triticum aestivum* L[13]. Its popularity stems from its high yield, rich protein content, good palatability, and adaptability. In recent years, with rapid economic development and rising living standards, demand for livestock products in China has surged, leading to a steady expansion of alfalfa cultivation areas[14].

In our experiment, the *Trigonella* genus, a related group to *Medicago*, was included for comparison. The *Trigonella* genus contains approximately 70 species globally, fewer than those in *Medicago*[15,16]. Most *Trigonella* species are found in regions along the Mediterranean coast, central Europe, northern Africa, southwest and central Asia, and Oceania[17,18]. In China, only nine species of *Trigonella* have been identified, with *Trigonella foenum-graecum* L. (fenugreek) being the most prominent. Fenugreek is an annual or biennial herb known for its diverse uses[19]. It is widely recognized for its medicinal properties, such as dispelling cold, alleviating dampness, and relieving pain[20]. Additionally, fenugreek seeds contain diosgenin, a steroidal compound that plays a critical role in the synthesis of hormone-based medicines, demonstrating significant pharmaceutical potential[21–23]. Furthermore, studies have shown that fenugreek contains 4-hydroxyisoleucine, which has demonstrated efficacy in managing Type II diabetes[24]. Fenugreek gum is also used in the oil refining process for preparing fracturing fluid and as a thickening agent in the food industry[25]. Moreover, processed fenugreek is used in perfumes, further expanding its commercial applications. As demand for fenugreek increases, research into its chemical composition, pharmacological properties, cultivation methods, and diverse uses continues to grow.

As a closely related genus to *Medicago*, *Trigonella* shares several similarities with *Medicago* species. Both are leguminous forage plants that are nutrient-rich and have valuable wild and cultivated varieties for forage use. Some species from both genera thrive in similar environments and exhibit comparable morphological characteristics, making them difficult to differentiate. Numerous studies have demonstrated a close genetic relationship between *Medicago* and *Trigonella*, with transitional groups blurring the lines between the two genera. According to available agronomy literature, the seed size difference between plants of the *Trigonella* genus and those of the *Medicago* (Alfalfa) genus varies depending on the species and environmental conditions, however, in general, *Medicago* species tend to have larger seeds than *Trigonella* species. Seeds of *Trigonella* plants are typically smaller in both diameter and mass. For instance, the seeds of *Trigonella* are approximately 1.5 millimeters in diameter, whereas seeds of *Medicago* sativa (alfalfa) typically range from 2 to 3 millimeters in diameter[26–28]. These intermediate species, often referred to as "*Medicago*-like" or "*Trigonella*-like," include species such as *Melilotoides ruthenica* L[29]. Despite these similarities, the taxonomic classification of these intermediate groups remains unclear, and no comprehensive study comparing the two genera has been conducted to date.

The study of morphological seed diversity focuses on both quantitative and qualitative traits, collectively referred to as phenotypic traits[30,31]. These traits reflect the morphological diversity of seeds and are of great value in biological research, plant breeding, and other fields. As research into morphological diversity has gained momentum, it has become an increasingly prominent topic both in China and globally[32]. Various experiments have been conducted to measure and analyze the phenotypic characteristics of germplasm accessions. For example, Alireza and Ali (2020) analyzed the phenotypic traits of Damask rose germplasm accessions, contributing to the selection of superior breeding lines[33]. Dong (2016) assessed phenotypic traits in 537 common vetch germplasm accessions, revealing significant diversity that aided taxonomic studies[34]. Similarly, Oral (2020) observed high morphological variability in Caribbean breadfruit germplasm, with quantitative and qualitative traits providing insights into these variations. This research identified superior germplasm accessions with increased morphological variability and phenotypic diversity, supporting the commercial production and utilization of breadfruit[35]. Thus, studying the morphological diversity of germplasm resources is of great importance. However, few studies have simultaneously analyzed such a wide range of species.

In this experiment, 587 *Medicago* and 32 *Trigonella* germplasm accessions were collected from germplasm banks in China and the United States. The large sample size, with multiple accessions per species, provided a robust dataset. High-resolution images of seeds were captured, and 12 seed characteristics-straight length (SL), straight width (SW), width-to-length ratio (WL), perimeter (PE), radicle length (RL), hilum length (HL), 100-seed weight (SY), seed size, seed coat condition, seed shape, seed color, and radicle characteristics-were statistically analyzed. The findings provide valuable data for future efforts to identify and evaluate seed morphology in *Medicago* and related genera.

Results

Analysis of quantitative and qualitative characters

A total of seven quantitative and five qualitative traits from 587 *Medicago* and 32 *Trigonella* germplasm resources were analyzed. Their accession numbers can be checked in Table S1. The results of the quantitative analysis are presented in Tables S2 and S3, while representative seeds for different quantitative traits are shown in Fig. 1.

For easier figure preparation, we have assigned numbers to the seeds of 77 *Medicago* species and 21 *Trigonella* species (Table 1). The results in Table 2 indicate that, among the 77 *Medicago* species, the average seed length (SL) was 2.64 mm, average seed width (SW) was 1.64 mm, average seed perimeter (PE) was 7.00 mm, average hilum length (HL) was 0.09 mm, average radicle length (RL) was 0.96 mm, and average 100-seed weight (SY)

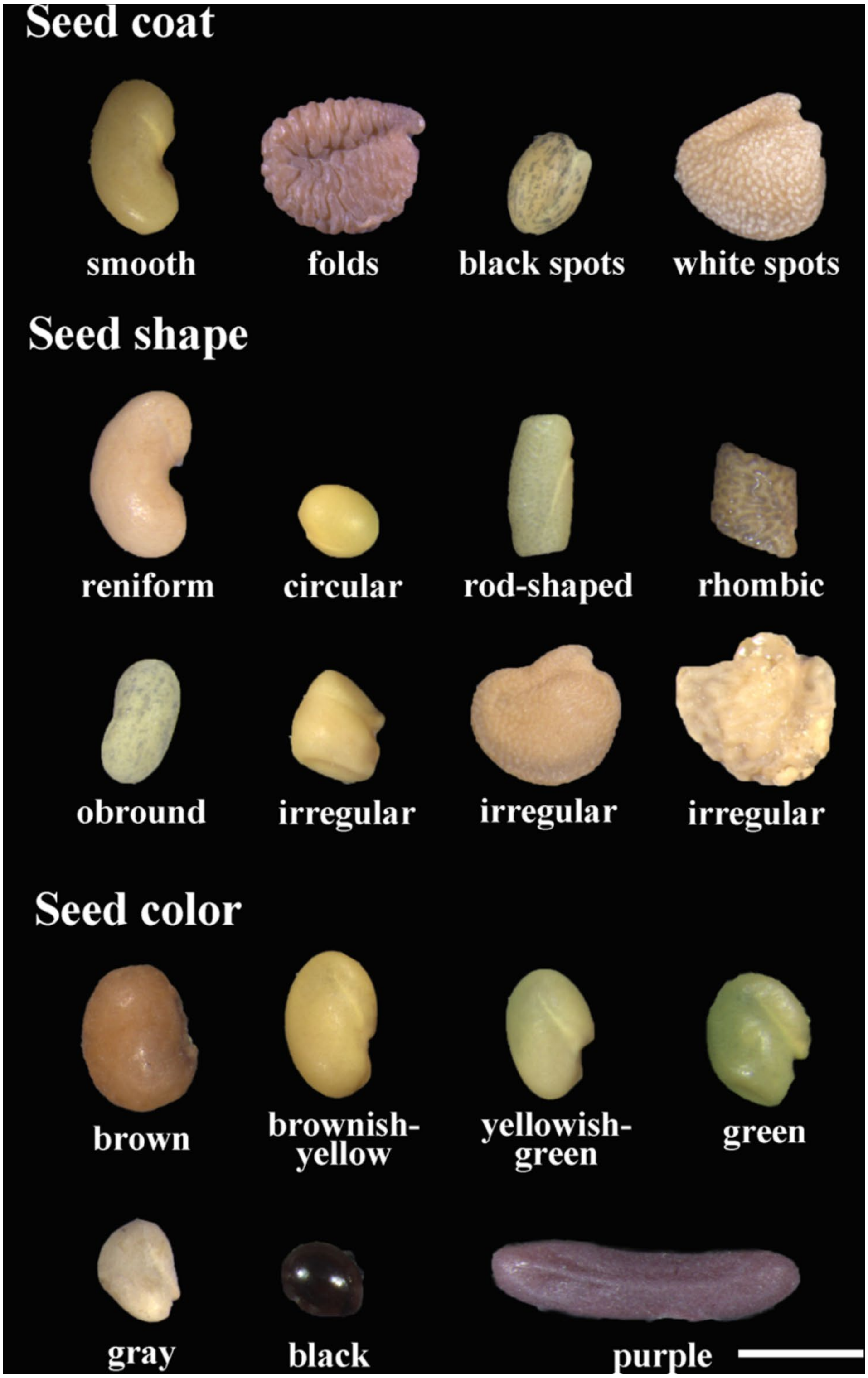

Fig. 1. Representative photos of germplasm accessions of different quantitative traits, including seed coat, seed color and seed shape. The white measuring scale represents 2000 μm.

was 0.31 g. The seed perimeter exhibited the largest standard deviation (1.82 mm), while the largest coefficient of variation (CV) was observed for 100-seed weight (98.69%), followed by hilum length. The smallest CV was found in the width-to-length ratio, at 20.90%.

Further analysis, as shown in Table 2, reveals variation within each species of the *Medicago* genus. For instance, the average SL of *Medicago sativa* (accession no. 62) was 2.14 mm, with a CV of 14.19%. The largest

Serial Number	Name	Amount	Serial Number	Name	Amount	Serial Number	Name	Amount	Serial Number	Name	Amount
1	*M. agropyreterum*	7	27	*M. heyniana*	2	53	*M. praecox*	2	1	*T. anguina*	2
2	*M. arabica*	3	28	*M. hybrida*	5	54	*M. prostrata*	3	2	*T. arabica*	2
3	*M. arborea*	4	29	*M. intertexta*	3	55	*M. quasifalcata*	8	3	*T. balansae*	2
4	*M. astroites*	3	30	*M. italica*	3	56	*M. radiata*	3	4	*T. caerulea*	2
5	*M. biflora*	1	31	*M. laciniata*	3	57	*M. rigidula*	3	5	*T. calliceras*	1
6	*M. blancheana*	3	32	*M. lanigera*	1	58	*M. rigiduloides*	3	6	*T. coelesyriaca*	1
7	*M. bonarotiana*	3	33	*M. lesinsii*	3	59	*M. rotata*	3	7	*T. coerulescens*	1
8	*M. borealis*	4	34	*M. littoralis*	4	60	*M. rugosa*	3	8	*T. corniculata*	2
9	*M. brachycarpa*	3	35	*M. lupulina*	11	61	*M. ruthenica*	28	9	*T. cretica*	2
10	*M. cancellata*	2	36	*M. marina*	2	62	*M. sativa*	67	10	*T. foenum-graecum*	2
11	*M. carstiensis*	1	37	*M. medicaginoides*	3	63	*M. sativa L.×M. ruthenica*	5	11	*T. glabra*	2
12	*M. ciliaris*	3	38	*M. minima*	6	64	*M. sauvagei*	2	12	*T. gladiata*	2
13	*M. coerulea*	8	39	*M. monantha*	3	65	*M. scutellata*	7	13	*T. gracilis*	1
14	*M. constricta*	3	40	*M. monspeliaca*	3	66	*M. secundiflora*	2	14	*T. kotschyi*	1
15	*M. coronata*	3	41	*M. murex*	2	67	*M. shepardii*	3	15	*T. macrorrhyncha*	1
16	*M. cretacea*	2	42	*M. muricoleptis*	3	68	*M. soleirolii*	3	16	*T. schlumbergeri*	1
17	*M. difalcata*	7	43	*M. noeana*	3	69	*M. subdicycla*	1	17	*T. spicata*	1
18	*M. disciformis*	3	44	*M. orbicularis*	3	70	*M. suffruticosa*	5	18	*T. spruneriana*	2
19	*M. doliata*	9	45	*M. orthoceras*	3	71	*M. tenoreana*	3	19	*T. stellata*	2
20	*M. edgeworthii*	2	46	*M. papillosa*	3	72	*M. tianschanica*	9	20	*T. suavissima*	1
21	*M. falcata*	102	47	*M. pironae*	3	73	*M. trautvetteri*	9	21	*Trigonella* spp.	1
22	*M. fischeriana*	3	48	*M. platycarpos*	3	74	*M. truncatula*	16			
23	*M. glandulosa*	8	49	*M. polyceratia*	3	75	*M. turbinata*	3			
24	*M. glutinosa*	6	50	*M. polychroa*	8	76	*M. varia*	72			
25	*M. granadensis*	3	51	*M. polymorpha*	21	77	*Medicago* spp.	6			
26	*M. hemicycla*	14	52	*M. popovii*	1						

Table 1. Germplasm code of *Medicago* and *Trigonella* Genus. For convenience, we prepared a table for each species in the *Medicago* and *Trigonella* genera. To describe some traits of species in the next part of this article, we used the serial number to represent them. Each copy contained at least 10 seeds in the same species. In this table, *Medicago* spp. and *Trigonella* spp. were both from the National Plant Germplasm System of the United States, and they were unidentified.

CV in *M. sativa* was observed for HL (32.28%). Comparing these results with those in Table S2, it is evident that there were generally larger CV values observed between species.

In Table S3, the results show that, among 21 different species in the *Trigonella* genus, average seed length was 2.31 mm, average seed width was 1.47 mm, average seed perimeter was 6.12 mm, average hilum length was 0.08 mm, average radicle length was 1.11 mm, and average 100-seed weight was 0.33 g. The largest value of standard deviation was seed perimeter, which was 1.94. The highest coefficient of variation was 100-seed weight, which was 112.98%, followed by hilum length and perimeter.

A coefficient of variation (CV) greater than 10% indicates a substantial difference between the samples. The CV for seed length, width, width-to-length ratio, perimeter, hilum length, radicle length, and 100-seed weight in *Medicago* germplasms ranged from 20.90 to 98.69%, with an average CV of 40.00%. For *Trigonella*, the variation coefficients for these traits ranged from 23.98 to 137.45%, with an average CV of 48.77%. The largest CV was observed for 100-seed weight in both genera. These results suggest that the CV of *Trigonella* was larger than that of *Medicago*, and that both genera exhibited substantial variation in their quantitative traits. Additionally, variance analysis revealed significant differences ($P < 0.05$) among traits, confirming that there were considerable variations in quantitative characteristics between the 587 *Medicago* and 32 *Trigonella* germplasms, reflecting rich morphological diversity.

The statistical results for the qualitative traits of *Medicago* germplasms are shown in Table 3. Seed coats were classified into four types: smooth, folded, and those with black or white spots. Most germplasms had smooth seed coats (83.48%), followed by folded seed coats (14.31%), with black and white spot patterns being rare. Seed shape was categorized into six types: reniform, obround, circular, irregular, rhombic, and rod-shaped. Reniform seeds were the most common, comprising 62.01% of the total. Circular, rod-shaped, and rhombic seed types were less frequent. Seed color was divided into five categories: brown, black, brownish yellow, gray, and yellowish green. Brownish yellow seeds were the most frequent (75.98%), followed by brown seeds (14.48%), with other colors occurring in smaller proportions. Radicle length was classified into two categories: long and medium,

Traits	Item	Serial number										
		1	2	3	4	5	6	7	8	9	10	11
SL/mm	Average	1.97	2.76	3.42	1.73	2.55	3.71	4.12	1.98	2.40	2.80	2.10
	CV	14.32%	9.47%	26.69%	16.30%	7.88%	10.43%	8.36%	13.75%	21.40%	8.11%	5.62%
SW/mm	Average	1.27	1.49	2.17	0.81	1.91	2.23	2.37	1.30	1.61	1.49	1.46
	CV	13.95%	10.54%	27.67%	13.53%	12.71%	11.41%	11.16%	11.64%	18.15%	10.62%	6.27%
PE/mm	Average	5.08	7.00	8.98	4.43	7.08	9.64	10.67	5.36	6.46	7.39	5.56
	CV	13.68%	9.76%	27.67%	10.67%	6.69%	10.76%	8.79%	11.87%	19.34%	9.17%	4.58%
HL/mm	Average	0.09	0.08	0.10	0.05	0.04	0.09	0.12	0.08	0.05	0.09	0.06
	CV	38.73%	30.54%	39.26%	39.72%	62.31%	40.72%	31.38%	28.38%	36.89%	27.89%	31.41%
RL/mm	Average	0.80	1.09	1.09	0.75	0.55	1.16	1.24	0.92	1.05	0.90	0.94
	CV	22.17%	18.20%	21.97%	20.70%	41.99%	17.50%	13.99%	9.73%	15.11%	21.35%	12.30%
		Serial Number										
		12	13	14	15	16	17	18	19	20	21	22
SL/mm	Average	4.50	1.49	3.78	1.83	2.88	1.98	3.02	3.72	1.61	1.92	1.82
	CV	6.16%	17.24%	14.45%	9.52%	9.98%	11.21%	9.33%	27.41%	9.53%	15.46%	18.95%
SW/mm	Average	2.73	0.98	2.23	0.98	2.33	1.29	1.70	2.21	1.10	1.29	0.83
	CV	5.38%	13.30%	15.06%	16.11%	10.50%	11.13%	10.24%	26.52%	10.24%	13.10%	12.22%
PE/mm	Average	11.79	3.85	10.00	4.88	8.22	5.21	7.67	9.86	4.32	5.04	4.42
	CV	5.61%	14.62%	13.29%	10.46%	8.73%	8.83%	9.66%	28.09%	8.71%	13.17%	15.45%
HL/mm	Average	0.15	0.04	0.08	0.04	0.07	0.08	0.06	0.09	0.03	0.08	0.04
	CV	26.30%	43.92%	38.75%	48.22%	49.31%	32.43%	43.36%	43.79%	38.05%	31.17%	42.59%
RL/mm	Average	1.30	0.53	0.78	0.70	1.24	0.89	1.27	0.99	0.84	0.80	0.82
	CV	24.94%	30.66%	38.48%	17.58%	19.23%	18.21%	16.16%	28.66%	14.13%	19.51%	25.48%
		Serial Number										
		23	24	25	26	27	28	29	30	31	32	33
SL/mm	Average	2.14	2.14	3.70	1.93	2.24	2.24	4.87	2.75	2.41	3.17	3.56
	CV	12.32%	10.63%	8.89%	15.32%	12.44%	12.44%	8.00%	16.18%	10.21%	7.64%	11.70%
SW/mm	Average	1.39	1.37	2.27	1.31	1.37	1.37	3.05	1.69	1.26	1.60	2.11
	CV	10.09%	10.93%	9.29%	13.33%	9.54%	9.54%	9.25%	15.68%	8.78%	10.85%	14.20%
PE/mm	Average	5.60	5.59	9.64	5.32	5.81	5.81	12.93	7.44	5.98	8.43	9.56
	CV	10.16%	9.31%	9.14%	13.46%	11.15%	11.15%	7.76%	17.50%	9.30%	8.05%	12.58%
HL/mm	Average	0.10	0.09	0.20	0.09	0.07	0.07	0.27	0.08	0.06	0.02	0.10
	CV	28.97%	27.16%	21.46%	28.97%	36.01%	36.01%	28.09%	30.86%	37.00%	39.24%	32.48%
RL/mm	Average	0.87	0.91	0.81	0.91	0.92	0.92	1.96	0.73	0.95	1.10	1.02
	CV	19.81%	17.86%	19.31%	16.64%	16.27%	16.27%	9.20%	22.24%	19.05%	14.43%	16.36%
		Serial Number										
		34	35	36	37	38	39	40	41	42	43	44
SL/mm	Average	2.82	1.64	2.58	1.94	1.98	2.07	1.53	3.57	3.43	3.10	2.55
	CV	12.80%	12.45%	20.25%	12.57%	11.87%	15.18%	12.90%	8.81%	10.63%	8.82%	12.10%
SW/mm	Average	1.62	1.22	1.59	0.92	1.17	1.16	0.90	2.19	2.11	1.83	2.18
	CV	12.44%	11.16%	16.16%	17.82%	18.06%	12.15%	15.02%	12.91%	11.60%	9.92%	14.62%
PE/mm	Average	7.30	4.44	6.99	4.80	5.16	5.26	3.86	9.98	8.93	7.99	7.45
	CV	12.87%	12.24%	19.59%	8.85%	13.18%	13.46%	12.18%	8.24%	10.54%	8.52%	12.58%
HL/mm	Average	0.08	0.05	0.09	0.04	0.06	0.04	0.04	0.12	0.15	0.09	0.10
	CV	28.96%	42.68%	28.40%	27.52%	34.94%	39.34%	33.25%	21.05%	28.87%	26.21%	31.27%
RL/mm	Average	0.85	0.67	0.80	0.86	0.73	0.95	0.68	1.01	0.77	0.90	1.29
	CV	19.93%	20.42%	20.47%	24.06%	18.65%	16.59%	18.21%	14.60%	55.69%	19.31%	17.63%
		Serial Number										
		45	46	47	48	49	50	51	52	53	54	55
SL/mm	Average	1.83	2.26	2.50	2.53	1.78	2.07	2.61	1.87	2.26	2.19	1.92
	CV	10.15%	10.95%	7.23%	8.29%	21.33%	10.58%	16.64%	6.67%	10.36%	8.53%	17.23%
SW/mm	Average	0.98	1.48	1.56	1.74	1.03	1.39	1.52	1.45	1.16	1.37	1.28
	CV	12.29%	13.35%	8.22%	9.38%	16.36%	9.41%	16.68%	7.02%	14.42%	9.03%	15.54%
PE/mm	Average	4.59	5.95	6.54	7.06	4.56	5.49	6.64	5.39	5.67	5.69	4.98
	CV	8.80%	11.03%	7.17%	8.65%	16.95%	8.01%	16.60%	6.12%	9.30%	7.66%	15.35%
Continued												

HL/mm	Average	0.05	0.10	0.11	0.09	0.05	0.10	0.10	0.06	0.08	0.08	0.09
	CV	35.96%	21.91%	26.95%	44.49%	44.07%	25.46%	34.65%	30.42%	25.27%	36.50%	37.99%
RL/mm	Average	0.78	0.94	0.88	1.46	0.81	0.90	0.86	0.94	0.69	0.80	0.81
	CV	16.35%	14.60%	14.70%	12.80%	30.29%	15.86%	20.77%	15.16%	20.14%	26.31%	19.66%
		Serial Number										
		56	57	58	59	60	61	62	63	64	65	66
SL/mm	Average	2.84	2.92	3.24	3.47	2.91	2.19	2.14	2.24	3.57	3.65	1.93
	CV	9.40%	11.31%	13.35%	11.72%	15.86%	15.68%	14.19%	10.53%	9.83%	49.89%	5.64%
SW/mm	Average	2.26	1.83	2.07	2.15	2.02	1.65	1.37	1.41	1.98	2.49	1.36
	CV	12.14%	11.12%	13.14%	10.71%	23.39%	14.31%	12.91%	10.43%	14.05%	43.63%	9.70%
PE/mm	Average	8.17	7.62	8.56	9.06	8.29	6.29	5.60	5.78	9.19	10.38	5.26
	CV	9.60%	11.18%	15.18%	11.39%	21.23%	14.07%	12.91%	8.55%	9.49%	49.89%	6.52%
HL/mm	Average	2.84	0.09	0.11	0.15	0.10	0.06	0.10	0.09	0.15	0.25	0.02
	CV	9.40%	27.83%	35.60%	22.40%	29.00%	74.96%	32.28%	34.72%	23.30%	21.86%	38.94%
RL/mm	Average	1.53	0.96	1.04	1.13	0.86	1.19	0.89	0.92	1.34	1.16	0.84
	CV	14.90%	15.03%	14.75%	17.36%	14.64%	22.50%	19.40%	16.85%	13.30%	20.56%	22.51%
		Serial Number										
		67	68	69	70	71	72	73	74	75	76	77
SL/mm	Average	2.92	3.60	2.20	2.76	2.46	2.09	1.96	3.18	3.22	2.13	2.28
	CV	9.67%	11.38%	9.67%	16.72%	10.28%	11.01%	14.41%	13.08%	29.57%	12.44%	18.71%
SW/mm	Average	1.84	2.25	1.35	1.66	1.36	1.23	1.30	1.80	1.96	1.37	1.47
	CV	11.38%	14.14%	8.70%	15.06%	11.30%	10.77%	13.79%	12.96%	25.05%	11.01%	12.76%
PE/mm	Average	7.83	9.79	5.66	7.10	6.20	5.23	5.21	8.05	8.89	5.46	5.92
	CV	9.80%	12.58%	7.53%	17.02%	9.94%	9.92%	12.37%	12.87%	30.23%	10.66%	17.57%
HL/mm	Average	0.10	0.14	0.10	0.11	0.09	0.09	0.10	0.09	0.08	0.09	0.08
	CV	31.08%	36.36%	22.91%	32.08%	33.75%	37.23%	32.23%	33.56%	38.88%	33.09%	46.84%
RL/mm	Average	1.00	1.13	0.94	1.20	1.00	0.78	0.88	0.93	1.17	0.85	0.91
	CV	15.14%	17.81%	17.78%	19.28%	13.01%	20.90%	19.45%	19.53%	13.47%	18.48%	19.76%

Table 2. Statistical analysis of the quantitative characters among *Medicago* species. Note: SL, straight length; SW, straight width; WL, width-to-length ratio; PE, perimeter; HL, hilum length; RL, radicle length; SY, 100-seed weight. Result was obtained in EXCEL 2016. For convenience, we used serial number (Table 1) to represent each *Medicago* genus species. Means among quantitative traits in different traits were significantly different at $P < 0.05$.

with medium radicles representing 72.91%, and long radicles accounting for 18.74%. Medium-sized seeds made up 84.33% of the total.

The statistical results for the qualitative traits of *Trigonella* are presented in Table S4. Seed coats were categorized into three types: smooth, folded, and those covered with black spots. The highest proportion of germplasms had smooth seed coats (50.00%), followed by those with black spots (34.38%), while folded seed coats were the least common (15.63%). Seed shape was classified into five types: obround, circular, rod-shaped, rhombic, and irregular. Obround seeds were most prevalent (34.38%), with circular and irregular seeds each comprising 25.00%. Seed color was classified into six categories: brownish yellow, purple, brown, gray, green, and yellowish green. Brownish yellow seeds were most common (56.25%), followed by yellowish green and brown seeds (15.63%). Medium-length radicles were the most frequent (46.88%), and medium-sized seeds represented 43.75% of the total.

Cluster analysis

Cluster analysis of quantitative characters of *Medicago* and *Trigonella* was carried out by using R language software. The results are shown in Fig. 2 and S1.

The 77 different *Medicago* species can be divided into four groups according to their quantitative traits. There are four different germplasms in *Medicago* surrounded by red frames in Fig. 2, and seeds are larger from the left to the right side.

The 23 different species of *Trigonella* species are roughly divided into two groups according to their quantitative traits that were measured before (Fig. S1). The result may not be accurate, cause the number of *Trigonella* germplasms was far less than that of *Medicago* germplasm.

Principal-component analysis

Principal component analysis (PCA) was conducted using the seed traits of *Medicago* and *Trigonella* germplasms, including seed length (SL), seed width (SW), width-to-length ratio (WL), radicle length (RL), hilum length (HL), and 100-seed weight (SY), to clearly illustrate the contribution of each trait to morphological seed diversity. The

Qualitative traits	Sort	No. of germplasm	Frequency(%)
Seed coat	Smooth	490	83.48
	Fold	84	14.31
	Black spots	12	2.04
	White spots	1	0.17
Seed shape	Reniform	364	62.01
	Circular	27	4.60
	Irregular	95	16.18
	Obround	83	14.14
	Rod-shaped	16	2.73
	Rhombic	2	0.34
Seed color	Brown	85	14.48
	Brownish yellow	446	75.98
	Gray	7	1.19
	Black	13	2.21
	Yellowish green	36	6.13
Situation of radicle	Long	110	18.74
	Medium	428	72.91
	Short	49	8.35
Seed seizes	Large	44	7.50
	Medium	495	84.33
	Short	48	8.18

Table 3. Statistical analysis of the qualitative characters of *Medicago* germplasm resources. Radicle situation means the ratio of radicle length to seed length. If radicle length exceeded 2/3 of seed width, the seed was classified as long radicle. If the radicle length was within 1/3 – 2/3 of the seed width, the seed was classified as medium radicle. The rest were classified as short radicle.

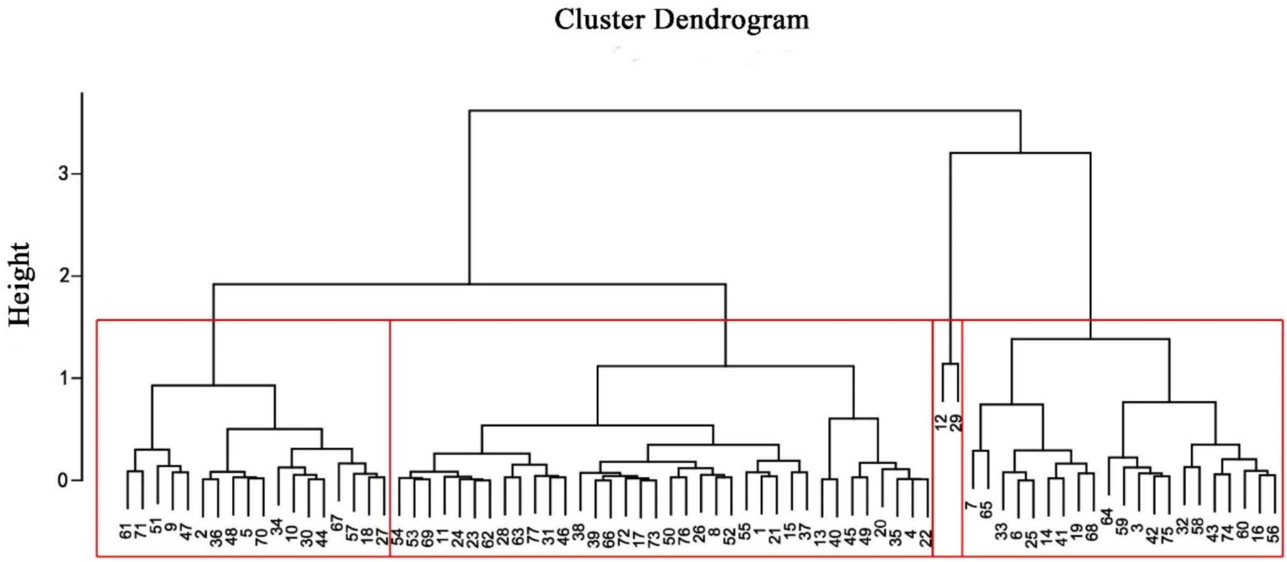

Fig. 2. Dendrogram of *Medicago* germplasm. The 77 different species of *Medicago* genus can be roughly divided into four groups according to their SL, SW, PE, and SY, signified by four parts surrounded by red lines.

correlation analysis (Table S5) revealed that most quantitative traits were positively correlated, except for WL, SL, perimeter (PE), and SY, which were negatively correlated. The strongest correlations were observed between SL and PE (0.990), and SW and PE (0.944).

In the PCA, components with eigenvalues greater than 1 were selected to ensure that each component provided meaningful interpretation and representativeness. Three principal components were extracted, explaining a cumulative variance of 85.64%. The loadings of each trait under the different principal components reflect the degree to which each trait contributes to the component. These results are summarized in Table S6.

From the PCA results, in the first principal component, SL, SW, HL, RL, and SY had the largest loadings, collectively representing the primary quantitative traits of *Medicago* germplasms. In the second principal component, the width-to-length ratio (WL) had the highest loading, supplementing the quantitative characteristics of seed morphology. The third principal component was dominated by seed coat, shape, color, and radicle traits, which are key qualitative traits of the germplasm.

The biplot (Fig. 3) shows 7 quantitative characters in 8 representative germplasms selected in Fig. 2, which were divided into four groups according to their SL, SW, PE, and SY. We randomly selected two different germplasms in each group to make PCA-plot, and they can be divided into four groups, too.

Multiple regression analysis and path analysis

Although we observed that seed length (SL), seed width (SW), hilum length (HL), radicle length (RL), and 100-seed weight (SY) had large loadings in the first principal component of the PCA, we sought to better understand how the other quantitative traits influence SY. The correlation results in Table S4 were not sufficient for this purpose. Therefore, we conducted path analysis, setting SY as the dependent variable (Y) and SL, SW, width-to-length ratio (WL), perimeter (PE), HL, and RL as independent variables, denoted as X1, X2, X3, X4, X5, and X6, respectively.

The path analysis results, presented in Table S7, show the following regression equation:

$Y=-0.517+0.219\times X1+0.235\times X2-0.138\times X3-0.022\times X4+1.351\times X5+0.31\times X6$.

From this equation, we observe that SY was positively correlated with SL, SW, HL, and RL. Specifically, for every 1 mm increase in SL, SY increased by 0.219 mm; for every 1 mm increase in SW, SY increased by 0.235 mm; for every 1 mm increase in HL, SY increased by 1.351 mm; and for every 1 mm increase in RL, SY increased by 0.31 mm. In contrast, PE had a negative influence on SY, decreasing by 0.022 mm for each unit increase in PE.

The results of the regression analysis, shown in Table S8, indicate that the *F*-statistic was 68.366 with a P-value of less than 0.05, suggesting that the regression model is statistically significant ($P<0.05$).

Boxplot of quantitative traits in *Medicago* and *Trigonella* species

A boxplot was constructed based on the numerical values of the quantitative traits of *Medicago* germplasms, with the results shown in Fig. 4.

From the statistical data presented in Fig. 4, it is evident that seed samples from the *Medicago* genus are generally larger than those from the *Trigonella* genus, particularly in terms of seed length, width, perimeter, and hilum length. However, when examining the width-to-length (WL) ratio, *Trigonella* seeds exhibit a larger proportion compared to *Medicago* seeds. This suggests that, while *Medicago* seeds may be larger in size overall, *Trigonella* seeds stand out for their distinctive seed shape in terms of the WL ratio. Additionally, there is little difference in radicle length between the two genera, making this trait less useful for distinguishing them. Furthermore, the variation or dispersion in the six quantitative traits measured for *Trigonella* seeds is more pronounced compared to *Medicago*. This greater variability may be partly due to the smaller sample size for *Trigonella* seeds, which could influence the reliability of the calculated averages and standard deviations for these traits.

According to statistical data (Fig. S2), the average straight length of the 77 *Medicago* species was 2.64 mm, indicated by the dotted line in the boxplot. In addition, the boxplot showed that the length of *M. scutellata* (no.65), *M. doliata* (no.19), and *M. intertexta* (no.29) were significantly different from the others. The seed length of *M. falcata* (no.21) was smaller than that of *M. sativa* (no.62).

According to the statistical data (Fig. S2A), the average seed width of *Medicago* germplasm was 1.64 mm. The boxplot showed that the widths of *M. scutellata* and *M. doliata* were much larger than the average, while the widths of *M. astroites* (no.4) and *M. fischeriana* (no.22) were smaller than the average.

According to the statistical data (Fig. S2B), the average perimeter of *Medicago* was 7.00 mm. The box plot showed that the perimeter of *M. scutellata* was much larger than the average. Besides, *M. turbinate* (no.75) had huge difference within species, because it had the longest bar in boxplot.

According to the statistical data (Fig. S2C), the average length of the whole hilum was 0.090 mm. It can be seen from the box plot that the hilum lengths of *M. scutellata* and *M. intertexta* were much larger than the average, while the hilum length of *M. biflora* (no.5) was significantly smaller than the overall average.

The average radicle length of *Medicago* germplasms was 0.960 mm. According to the boxplot (Fig. S2), the radicle length of *M. intertexta* (no.29) was much larger than the average.

Considering the length and completeness of this article, we also prepared several boxplots of different quantitative traits in the *Trigonella* genus, these pictures can be found in the supplementary materials (Fig. S3).

Discussion

The diversity observed in both qualitative and quantitative traits provide valuable insight into the morphological characteristics of plants. This morphological diversity reflects the genetic and phenotypic richness inherent in germplasm collections, which are essential resources for plant breeding efforts. As noted by Wonneberger (2017)[36], the variety within germplasm collections forms the foundation for successful breeding programs, offering a broad range of traits that can be selectively combined to achieve desirable crop characteristics. In this study, we analyzed 587 *Medicago* and 32 *Trigonella* germplasms sourced from Chinese and American germplasm banks. The selected samples were subjected to an extensive statistical analysis of 12 key morphological traits, including seed length (SL), seed width (SW), width-to-length ratio (WL), seed perimeter (PE), radicle length (RL), hypocotyl length (HL), seed yield (SY), seed coat characteristics, seed shape, seed color, seed size, and radicle condition. Our findings revealed substantial morphological diversity within both genera, highlighting the

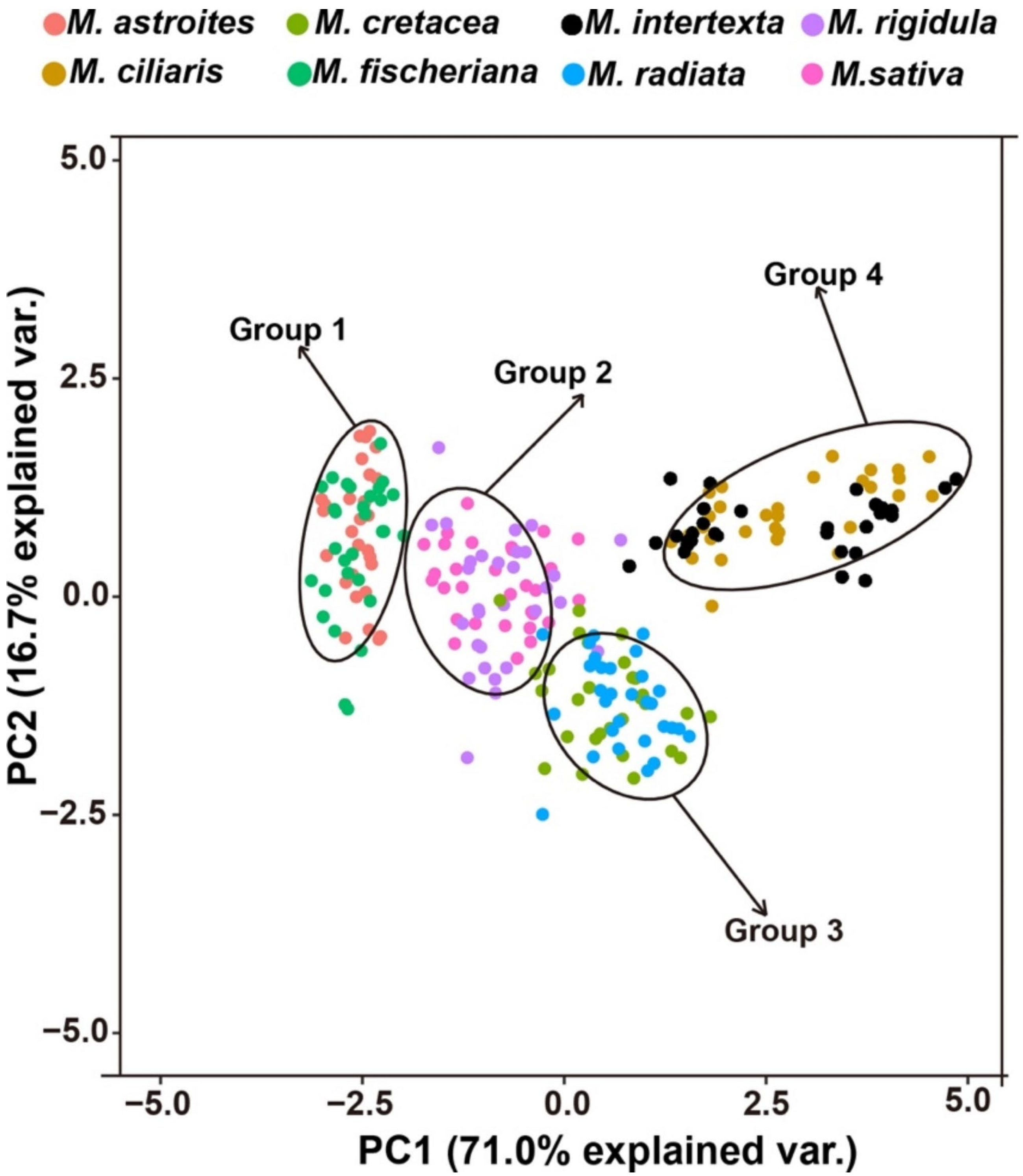

Fig. 3. PCA-plot of 7 quantitative characters in 8 representative germplasms selected in Fig. 2, which were divided into four groups according to their SL, SW, PE, and SY. We randomly selected two different germplasms in each group to make PCA-plot, and they can be divided into four groups, too. The result was the same as the diagram. Group 1 contained *M. astroites* and *M. fischeriana*; Group 2 contained *M.sativa* and *M. rigidula*; Group 3 contained *M. radiata* and *M. cretacea*; Group 4 contained *M. intertexta* and *M. ciliaris*.

wide range of phenotypic expressions present in these germplasm collections. This rich diversity underscores the potential of these germplasms to significantly contribute to breeding programs. By investigating and harnessing these unique traits, researchers and breeders can select for important attributes-such as disease resistance, drought tolerance, and improved yield-ultimately enhancing the genetic pool available for agricultural advancement.

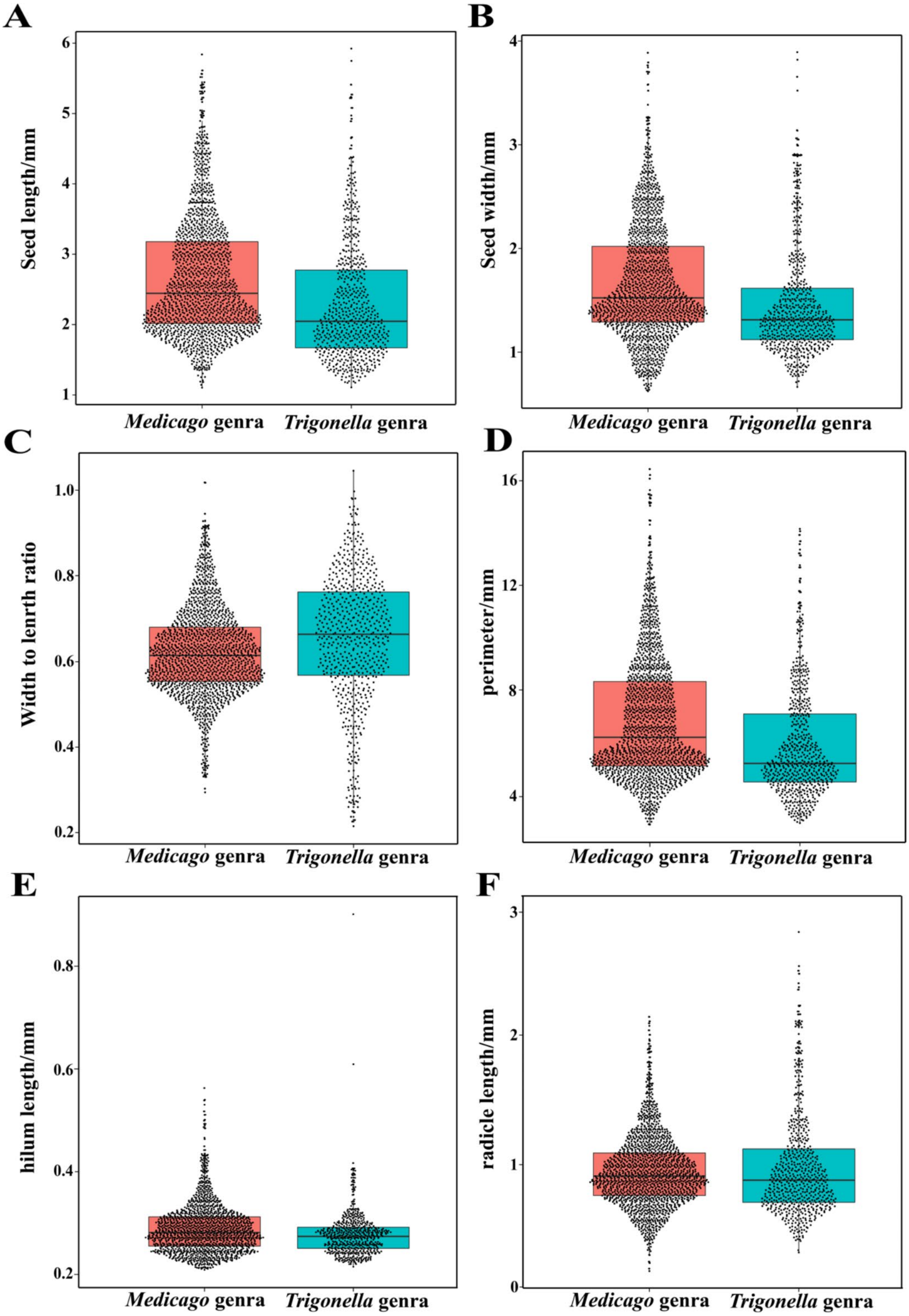

Fig. 4. Boxplot of 6 quantitative traits in *Mediacgo* and *Trigonella* genus. (**A**) Seed length, (**B**) seed width, (**C**) width to length ratio, (**D**) perimeter, (**E**) hilum kength, (**F**) radical length.

In our experiment, the variation range of seed length (SL) was 1.01 to 5.84 mm in the *Medicago* genus and 0.93 to 5.92 mm in the *Trigonella* genus. For seed width (SW), the variation range was 0.51 to 3.89 mm in *Medicago* species and 0.63 to 3.89 mm in *Trigonella* species. The variation range for seed yield (SY) in the two genera was 0.06 to 1.84 g for *Medicago* and 0.04 to 2.25 g for *Trigonella*. Notably, the coefficient of variation for 100-seed weight was the largest across all traits. Additionally, there were differences in the coefficient of variation for 100-seed weight within each species. For *Trigonella*, the coefficient of variation exceeded 100%, likely due to the smaller sample size of Trigonella germplasms, which may have resulted in a larger deviation.

Seed size plays a critical role in seed reproduction and seedling establishment. Research suggests that, given equal external energy, plants with smaller seeds tend to produce more seeds and exhibit stronger reproductive capacity. In contrast, larger seeds often show better stress resistance, particularly during the later stages of seedling establishment, when environmental stress factors are more pronounced[37]. This highlights the importance of seed morphology in influencing seedling development, survival, and establishment in varying environments. Thus, information on seed size and related traits in plant germplasm can be instrumental in variety selection, particularly for improving seed yield, seedling establishment, and survival rates[38]. In this experiment, the *Medicago* and *Trigonella* seeds were collected from 19 different countries, each exhibiting distinct qualitative and quantitative traits. By combining environmental factors with these morphological characteristics, we can offer more accurate data for seed identification and classification, facilitating more precise breeding and selection strategies based on the unique traits of each germplasm source.

The morphological diversity observed in plants is the result of the complex interaction between genes and the external environment. Phenotypic traits, which are influenced by gene expression and regulation, often provide valuable insights into the underlying genetic diversity. These phenotypic variations are not only a product of long-term evolutionary processes but also serve as a foundation for the continued adaptation and evolution of germplasms in specific environmental contexts[27]. Thus, studying morphological diversity is crucial for understanding the evolutionary trajectories of species or populations and offers important insights into their breeding potential[39]. In our study, we examined germplasms from the *Medicago* and *Trigonella* genera. Analysis of the data revealed that while these germplasms exhibit broad geographic distribution, notable differences exist between species within the same genus. These differences reflect varying levels of adaptability to environmental conditions, emphasizing the role of both genetic and environmental factors in shaping phenotypic diversity. Such variation underscores the importance of considering both genetic and environmental influences when studying species diversity and breeding potential.

In this experiment, we observed substantial morphological diversity among the different germplasms of *Medicago* and *Trigonella*, suggesting that each germplasm carries unique genotypes shaped by long-term evolution and adaptation to diverse environments. This diversity has significant implications for studying species evolution. The phenotypic variation observed within a genus is primarily due to intrageneric variation, which results from the complex interaction between genetic factors and environmental conditions. Phenotypic diversity often mirrors genotypic variation to a certain extent[40]. For example, *Medicago* species are distributed across the Mediterranean, southwestern and central Asia, and Africa. Over time, *Medicago* has evolved new genotypes to adapt to these varied environments, leading to the formation of distinct phenotypic traits, which have formed the basis for breeding programs.

While molecular markers, such as simple sequence repeats (SSR), have shown significant genetic similarity between *Trigonella* and *Medicago*[41], this experiment focused primarily on phenotypic traits. The cluster analysis revealed similarities between the two genera in key quantitative traits such as seed length, width, and perimeter. This suggests that distinguishing between *Medicago* and *Trigonella* based on these morphological traits alone can be challenging, highlighting the need for more comprehensive analysis, including molecular techniques, to better differentiate them.

Principal component analysis (PCA) has been applied to various crops, such as oats[42] and soybeans[43], to simplify the analysis and comprehensively capture the morphological characteristics of plant populations. In this experiment, PCA was performed on 587 *Medicago* germplasms, extracting three principal components that explained a cumulative 85.64% of the total variance. In the first principal component, traits like seed length (SL), seed width (SW), hilum length (HL), root length (RL), and seed yield (SY) had higher loadings, reflecting the overall quantitative characteristics of *Medicago* germplasms. The second principal component highlighted the seed width-to-length ratio, which served as a supplementary trait to the quantitative seed features and complemented the qualitative characteristics of the germplasm. In the third principal component, traits such as seed coat, seed shape, seed color, and radicle condition were more influential, providing additional insights into the qualitative aspects of the germplasm. These findings suggest that seed size can be an effective index for simplifying the seed collection process, making it more streamlined and accurate. By focusing on key traits revealed through PCA, researchers can optimize germplasm evaluation and improve the efficiency of field assessments.

Germplasm resources are crucial for advancing both taxonomic and breeding research, as they provide the genetic diversity essential for understanding species classification and for developing improved crop varieties[44–46]. However, current taxonomic research often encounters challenges due to incomplete species representation, limited genus coverage, and the underrepresentation of key species in certain contentious groups[47,48]. This gap in comprehensive germplasm data hampers the resolution of many taxonomic issues. For instance, in the *Medicago* genus (commonly known as alfalfa), several groups contain only a single species, despite their critical taxonomic importance[49]. These groups are underexplored in research, which increases the risk of overlooking valuable taxonomic insights. Therefore, a global collaborative effort among alfalfa taxonomists is essential to comprehensively collect and share germplasm resources[50]. Such efforts would strengthen the foundation for addressing these taxonomic gaps[51]. One example is the *Sporocarps* group within the *Medicago* genus, which comprises 35 species and represents one of the most complex and poorly understood evolutionary relationships in the genus[52]. The genetic diversity within this group is highly variable, and a more focused collection and preservation effort is needed to clarify its evolutionary history[53,54]. Moreover, the presence of species like *M. medicaginoides* and *M. orbicularis* in the wild remains uncertain and warrants further field investigations[55]. From a breeding perspective, there are three primary objectives that should guide research on alfalfa germplasm resources. First, there is an urgent need for the extensive collection of wild germplasm, particularly focusing on species closely related to cultivated alfalfa, including various subspecies, alfalfa complex variants, and other *Medicago* species[56,57]. Utilizing high-throughput sequencing technologies will enable researchers to identify and locate valuable genes in both cultivated and wild species[58]. These genetic resources can then be incorporated into molecular breeding programs to enhance the resilience and productivity of high-yielding alfalfa varieties, especially for herbivorous livestock farming[59,60]. Secondly, while 19 species of *Medicago* are known to have cultivated varieties, further domestication efforts are required to improve their yield and quality. Such efforts offer significant promise for breeding programs aimed at increasing the economic and nutritional value of alfalfa in animal husbandry[61,62]. Third, wild alfalfa varieties, especially those native to specific regions, exhibit strong adaptability to local climates and soils. By collecting and domesticating these local germplasm resources, breeders can select for traits such as higher seed yield and lower hard seed rates. This would not only improve local livestock production but also support ecological conservation initiatives[63]. Areas like the Qinghai-Tibet Plateau, Xinjiang, and Inner Mongolia, which are rich in wild alfalfa resources, offer great potential for such targeted breeding efforts. In these regions, breeding could enhance agricultural productivity and contribute to sustainable land management practices[64,65]. Collectively, these taxonomic and breeding initiatives underscore the critical importance of germplasm resources in addressing research gaps, improving crop resilience, and contributing to sustainable agricultural and environmental practices[66].

In this experiment, a substantial collection of *Medicago* and *Trigonella* germplasm accessions was analyzed. However, variations in the number of samples collected across different species introduced some inconsistencies that may have impacted the results. These discrepancies highlight the need for a more balanced representation of germplasms from each species to ensure robust and reliable findings. Looking forward, we aim to expand the germplasm collection to improve the comprehensiveness and accuracy of our dataset, which will enable more nuanced and representative analyses across a wider range of species. Additionally, the current analysis of morphological diversity was limited to 12 phenotypic traits. While these traits provided valuable insights, there remains significant potential to expand our understanding by investigating additional seed characteristics. Exploring a broader spectrum of morphological and genetic traits would allow for a more complete and detailed evaluation of *Medicago* and *Trigonella* seeds, thereby enriching the information available for their identification, characterization, and assessment in breeding programs. Future studies will thus aim to incorporate a wider array of phenotypic markers, enhancing our ability to identify and evaluate unique traits within these genera and supporting their use in agricultural innovation.

Materials and methods

Seed materials

In this study, a total of 619 germplasm accessions were evaluated. These germplasms consisted of 587 accessions in the *Medicago* genus, which contains 77 different species, and 32 accessions in the *Trigonella* genus, which contains 21 different species. These accessions were acquired from the National Animal Husbandry Station of China and the National Plant Germplasm System of the United States. They were from 19 countries, including China, the United States, Russia, German, Australia, and Canada. These seeds were collected in different years. For convenience, a code table was prepared (Table 1) to represent every germplasm in the *Medicago* and *Trigonella* genera. We selected representative seed photos which were obtained by the Discovery.V20 instrument produced by Zeiss, Germany (Fig. 5). We uploaded seed photos for each species, that can be found on the web: https://pan.baidu.com/s/19n8chQzNy6x2awv-n5YsNA, the password was "Medi", in a file titled "Seed Photos".

Morphological traits evaluation

In this study, 12 morphological traits, including 7 quantitative traits and 5 qualitative traits, were recorded. Qualitative traits were examined on the basis of seed coat, shape, color, size, and radicle situation. Seed coat was recorded for smooth, fold, and black and white spots. Seed color was recorded for brown, gray, black, green, purple, brownish yellow, and yellowish green. Seed size was recorded for large, medium, and small size. Seed shape was reniform, rhombic, circular, obround, rod-shaped, and irregular. The radicle situation was recorded for long, medium, and short radicles. If radicle length exceeded 2/3 of the width of the seed, the seed was classified as a long radicle. If radicle length was within 1/3 – 2/3 of seed width, the seed was classified as a medium radicle. The rest were classified as short radicles. To avoid the error of classification caused by different observers, the same observer was used in the statistics of a certain trait.

Quantitative traits were examined based on straight length (SL), straight width (SW), width-to- length ratio (WL), perimeter (PE), hilum length (HL), radicle length (RL), and 100-seed weight (SY). SL, SW, PE, HL, and RL were measured on Digimizer 4.2 software. WL was calculated on SPSS 21.0, and SY was measured by an electronic balance.

Statistical analysis

In this study, we used Discovery.V20 stereoscope produced by Zeiss, Germany to obtain seed images, which were produced in Digimizer 4.2 software. Then, we measured SL, SW, RL, and HL. The 100-seed weight was measured by an electronic balance with 0.01 g precision. Morphological evaluations were carried out using at least 20 repetitions of per accession.

The average values of the measured characters were used for analysis. The coefficients of variation (CV) were calculated as the variance index. Correlation between traits was determined using Pearson's correlation coefficient with SPSS software version 21.0 (SPSS Inc., Chicago, IL, USA, Norusis 1998). The relationship between accessions was assessed using principal-component analysis (PCA) with SPSS 21.0 software. The first and second components (PC1 and PC2) were used to create the scatter plot with SPSS 21.0 software. Cluster analysis dendrogram was made in R 3.6.3. In addition, data of quantitative traits were used to produce boxplots;

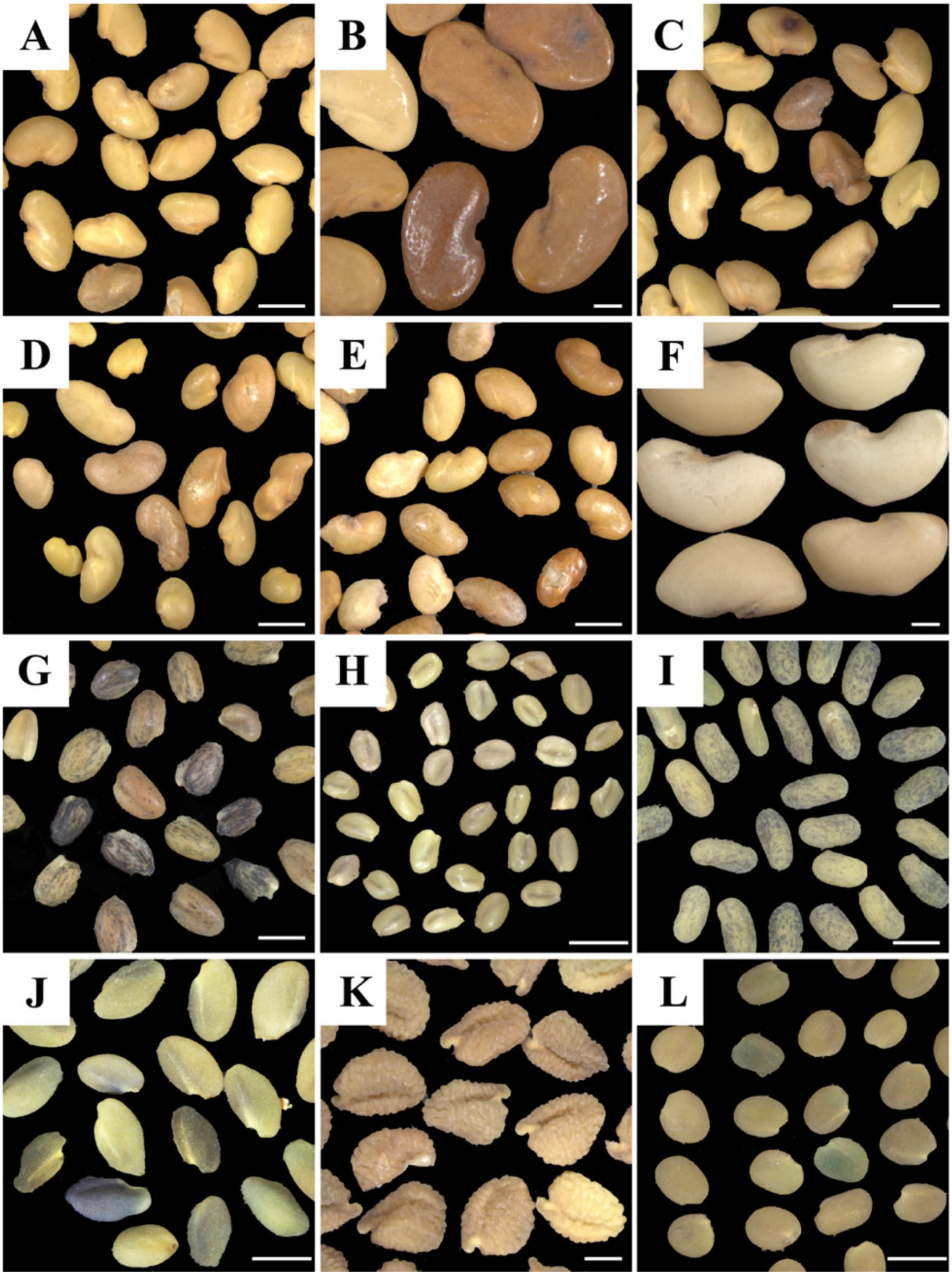

Fig. 5. Representative photos of germplasm accessions of *Medicago* and *Trigonella* genus. Each picture represents different germplasm: (**A**) *M. agropyreterum*; (**B**) *M. arborea*; (**C**) *M. borealis*; (**D**) *M. coerulea*; (**E**) *M. difalcata*; (**F**) *M. doliata*; (**G**) *T. suavissima*; (**H**) *T. stellata*; (**I**) *T. spruneriana*; (**J**) *T. spicata*; (**K**) *T. schlumbergeri*; (**L**) *Trigonella macrorrhyncha*. There were large differences among the received germplasm resources. The white measuring scale represents 2000 μm.

the tools that we used are on the following website: http://www.ehbio.com/ImageGP/index.php/Home/Index/Boxplot.html.

Conclusions.

The diversity of both qualitative and quantitative traits is essential for the morphological characterization of plants, and the genetic variability within germplasm forms the cornerstone of successful breeding programs. In this study, 587 *Medicago* and 32 *Trigonella* accessions were selected from germplasm banks in China and the United States. A statistical analysis of 12 traits-seed length (SL), seed width (SW), width-to-length ratio (WL), seed perimeter (PE), radicle length (RL), hilum length (HL), 100-seed weight (SY), and seed color-revealed significant morphological diversity within both genera.

While this study gathered a diverse range of germplasms, the number of accessions varied across species, which could affect the comprehensiveness of the findings. To strengthen the robustness of future research, increasing the number of accessions is recommended. Additionally, this analysis focused on 12 phenotypic traits, but a more in-depth exploration of additional seed characteristics would provide a more detailed and comprehensive understanding, aiding in the accurate identification and evaluation of *Medicago* and *Trigonella* seeds.

This research lays a valuable foundation for establishing seed banks and further expanding germplasm resources through a systematic examination of morphological traits.

Data availability

The data that support the findings of this study are openly available in Science Data Bank at https://www.scidb.cn/en/s/VJNnYv.

Received: 21 November 2024; Accepted: 16 January 2025

References

1. Reid, R., Konopka, J. & Rihana, J. R. Needs and priorities collection of annual medic germplasm. International Center for Agricultural Research in the dry areas. *Aleppo Syria* 253–270 (1989).
2. Jiang, J. P., Xiong, Y. C., Jia, Y., Li, F. M. & Xu, J. Z. Soil quality dynamics under successional alfalfa field in the semi-arid Loess Plateau of Northwestern China. *Arid Land. Res. Manage.* **21**, 287–303 (2007).
3. Wang, T. M., Xia, K. H. & Jin, C. The complete chloroplast genome sequences of a wild diploid alfalfa *Medicago edgeworthii* (*Leguminosae*). *Mitochond. DNA Part. B* **5**, 1683–1684 (2020).
4. Campbell, T. A. & Bauchan, G. R. Organelle based molecular analyses of the genetic relatedness of cultivated alfalfa (*Medicago sativa* L.) to *Medicago Edgeworthii* Sirjaev, and *Medicago ruthenica* (L.) Ledebour. *Euphytica* **125**, 51–58 (2002).
5. Badri, M., Ben, C., Mahjoub, A. & Abdelly, C. Morpho-phenological diversity among natural populations of *Medicago polymorpha* of different Tunisian ecological areas. *Afr. J. Biotechnol.* **15**, 1330–1338 (2016).
6. Kotula, L., Kwa, H. Y., Nichols, P. G. H. & Colmer, T. D. Tolerance and recovery of the annual pasture legumes *Melilotus siculus*, *Trifolium michelianum* and *Medicago polymorpha* to soil salinity, soil waterlogging and the combination of these stresses. *Plant Soil* **444**, 267–280 (2019).
7. Xiang, H. M., Zhang, Y., Wei, H., Zhang, J. E. & Zhao, B. L. Soil properties and carbon and nitrogen pools in a young hillside longan orchard after the introduction of leguminous plants and residues. *Peer J.* **6**, e5536 (2018).
8. Baral, R. et al. Exploring the impact of soil water variability and varietal diversity on alfalfa yield, nutritional quality, and farm profitability. *Grassland Res.* **4**, 266–278 (2023).
9. Anderson, A. W. et al. Salinity and cultivar effects on alfalfa forage yield and nutritive value in a. Mediterranean climate. *Grassl. Res.* **3**, 153–166 (2023).
10. Radovic, J., Sokolovic, D. & Markovic, J. Alfalfa-most important perennial forage legume in animal husbandry. *Biotechnol. Anim. Husb.* **25**, 465–475 (2009).
11. Jefferson, P. G. & Cutforth, H. W. Sward age and weather effects on alfalfa yield at a semi-arid location in southwestern Saskatchewan. *Can. J. Plant Sci.* **77**, 595–599 (1997).
12. Armstrong, J. M. Cytological studies in alfalfa polyploids. *Can. J. Bot.* **32**, 531–542 (1954).
13. Matt, A. Y., Michael, P. R., Jeffrey, A. C. & Paul, V. B. Alfalfa stands length and subsequent crop patterns in the upper Midwestern United States. *Agron. J.* **165**, 1697–1708 (2014).
14. Wen, Z. H., Duan, T. Y., Christensen, M. J. & Nan, Z. B. *MicrodoTabacinumacinum*, confirmed as a pathogen of alfalfa in Gansu Province, China. *Plant Dis.* **99**, 87–92 (2015).
15. Yin, M. et al. Genomic analysis of *Medicago ruthenica* provides insights into its tolerance to abiotic stress and demographic history. *Mol. Ecol. Resour.* **21**, 1641–1657 (2021).
16. Wang, T. et al. The genome of a wild *Medicago* species provides insights into the tolerant mechanisms of legume forage to environmental stress. *BMC Biol.* **19**, 96 (2021).
17. Lu, X. S. Status of germplasm classifications for genus Medicago. *Acta Agrestia Sin.* **17**, 680–685 (2009).
18. Xu, L. R., Chen, D. Z. & Zhu, X. Y. *Flora of China* (Science, 2010).
19. Spanda, C., Pooja, S. C. & Surendra, K. Review on fenugreek (*Trigonella foenum-graecum* L.) and its important secondary metabolite diosgenin. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca* **46**, 22–31 (2018).
20. Zaidi, S. N. F. & Masood, J. The protective effect of fenugreek seeds extract supplementation on glucose and lipid profile in thioacetamide induced liver damage in rats. *Pak. J. Pharm. Sci.* **33**, 2003–2008 (2020).
21. Vigh, S. et al. Analysis of phytoconstituent profile of fenugreek –*Trigonella foenuem-graecum* L. - seed extracts. *Studia Universitatis Babes Bolyai Chem.* **62**, 145–166 (2017).
22. Yadav, U. C. S. & Baquer, N. Z. Pharmacological effects of *Trigonella foenum-graecum* L. in health and disease. *Pharm. Biol.* **52**, 243–354 (2014).
23. Moradi, K. & Moradi, N. Physiological and pharmaceutical effects of fenugreek (*Trigonella foenum-graecum* L.) as a multipurpose and valuable medicinal plant. *Global J. Med. Plant. Res.* **1**, 199–206 (2013).
24. Allaoui, A., Gascon, S., Benomar, S., Quero, J. & Osada, J. Protein hydrolysates from fenugreek (*Trigonella foenum-graecum*) as nutraceutical molecules in colon cancer treatment. *Nutrients* **11**, 724 (2019).
25. Khorshidian, N., Asli, M. Y. & Arab, M. Fenugreek: potential applications as a functional food and nutraceutical. *Nutr. Food Sci. Res.* **3**, 5–16 (2016).
26. Small, E., Crompton, C. W. & Brookes, B. S. The taxonomic value of floral characters in tribe *Trigonelleae* (*Leguminosae*), with special reference to *Medicago*. *Can. J. Bot.* **59**, 1578–1598 (1981).

27. Soltis, S. & Soltis, E. Genetic variation in endemic and widespread plant species: examples from *Saxifragaceae* and *Polystichum* (*Dryopteridaceae*). *Aliso J. Syst. Evol. Bot.* **13**, 215–223 (1991).
28. Bena, G. Molecular phylogeny supports the morphologically based taxonomic transfer of the medicagoid *Trigonella* species to the genus *Medicago* L. *Plant Syst. Evol.* **229**, 217–236 (2001).
29. Elbatanouny, N. H., Castellano-Hinojosa, A. & Bedmar, E. J. The diversity of rhizobia modulating the *Medicago*, *Melilotus* and *Trigonella* inoculation group in Egypt is marked by the dominance of two genetic types. *Symbiosis* **67**, 3–10 (2015).
30. Wu, S. et al. Extensive genomic rearrangements mediated by repetitive sequences in plastomes of Medicago and its relatives. *BMC Plant Biol.* **21**, 421 (2021).
31. Chen, J. Y. et al. Phylogeny and species delimitation of Chinese *Medicago* (*Leguminosae*) and its relatives based on molecular and morphological evidence. *Front. Plant Sci.* **11**, 6197 (2021).
32. King, J. N., Yeh, F. C. & Heaman, J. C. Selection of growth and yield traits in controlled crosses of coastal Douglas-fir. *Silvae Genetica*. **37**, 158–164 (1988).
33. Alireza, K. & Ali, K. Morphological characterization of Damask rose (*Rosa×damascena* Herrm.) Germplasm to select superior accessions. *Genet. Resour. Crop Evol.* **67**, 1981–1997 (2019).
34. Dong, R., Jahufer, M. Z. Z., Dong, D. K., Wang, Y. R. & Liu, Z. P. Characterization of the morphological variation for seed traits among 537 germplasm accessions of common vetch (*Vicia sativa* L.) using digital image analysis. *New. Z. J. Agric. Res.* **59**, 422–435 (2016).
35. Oral, O. D., Laura, B. & Angela, T. A. Morphological diversity of breadfruit [*Artocarpus altilis* (Parkinson) Fosberg] in the Caribbean. *Sci. Hort.* 266 (2020).
36. Wonneberger, R., Ficke, A. & Lillemo, M. Identification of quantitative trait loci associated with resistance to net form net blotch in a collection of nordic barley germplasm. *Theor. Appl. Genet.* **130**, 2025–2043 (2017).
37. Aarssen, L. W. & Jordan, C. Between-species patterns of covariation in plant size, seed size and fecundity in monocarpic herbs. *Ecoscience* **8**, 471–477 (2001).
38. Odiaka, N. Morphological diversity among local germplasm of fluted pumpkin collected in Makurdi and Nigeria. *J. Food Agric. Environ.* **3**, 199–204 (2003).
39. Kelly, M. Adaptation to climate change through genetic accommodation and assimilation of plastic phenotypes. *Philosophical Trans. Royal Soc. Biol. Sci.* **374**, 1768 (2019).
40. Chandra, N. Rapid evolution of *Medicago polymorpha* during invasion shifts interactions with the soybean looper. *Ecol. Evol.* **9**, 20–25 (2019).
41. José, D., Esther, M. & Alvaro, P. History and current taxonomic status of genus *Agrobacterium*. *Syst. Appl. Microbiol.* **43**, 17–31 (2020).
42. Admassu-Yimer, B., Gordon, T. & Bonman, J. M. Development and validation of a quantitative PCR assay method of assessing relative resistance of oat (*Avena sativa*) to crown rust (*Puccinia coronata f.* sp. *avenae*). *Plant. Pathol.* **68**, 669–677 (2019).
43. Wang, G. F., Zhou, Q. Z., He, M. D., Zhong, X. B. & Tang, G. X. Wilting index and root morphological characteristics used as drought-tolerance variety selection at the seedling stage in soybean (*Glycine max* L). *Plant. Growth Regul.* **92**, 29–42 (2019).
44. Adhikair, L., Orville, M. L., Jonathan, M. & Missaoui, A. M. Dissecting key adaptation traits in the polyploid perennial medicago sativa using gbs-snp mapping. *Front. Plant. Sci.* **9**, 934 (2018).
45. Han, Y. et al. High density array for SNP genotyping and mapping in tetraploid alfalfa. In: (eds Sokolović, D., Huyghe, C. & Radović, J.) *Quantitative Traits Breeding for Multifunctional Grasslands and turf* 255–259 (Springer, Dordrech, Netherlands, 2014).
46. Shen, C. et al. The chromosome-level genome sequence of the autotetraploid alfalfa and resequencing of core germplasms provide genomic resources for alfalfa research. *Mol. Plant.* **13**, 1250–1261 (2020).
47. Wang, Z. et al. Development and characterization of simple sequence repeat (SSR) markers based on RNA-sequencing of *Medicago sativa* and in silico mapping onto the *M. truncatula* genome. *PLoS One*. **9**, e92029 (2014).
48. Riday, H., Brummer, E. C., Campbell, T. A., Luth, D. & Cazcarro, P. Comparisons of genetic and morphological distance with heterosis between *Medicago sativa* subsp. *sativa* and subsp. *falcata*. *Euphytica* **131**, 37–45 (2003).
49. Reyno, R., Khu, D. M., Monteros, M. J., Bouton, J. H. & Parrott, W. Evaluation of two transgenes for aluminum tolerance in alfalfa. *Crop Sci.* **53**, 1581–1588 (2013).
50. Richins, R. D., Scholthof, H. B. & Shepherd, R. J. Sequence of figwort mosaic virus DNA (caulimovirus group). *Nucl. Acids Res.* **15**, 8451–8466 (1987).
51. Liu, Z. H., Zhang, H. M., Li, G. L., Guo, X. L. & Chen, S. Y. Enhancement of salt tolerance in alfalfa transformed with the gene encoding for betaine aldehyde dehydrogenase. *Euphytica* **178**, 363–372 (2011).
52. McCaslin, M., Resisen, P. & Ho, J. New strategies for forage quality improvement in alfalfa. In *Proceedings of the 44th California alfalfa, forage, and grain symposium* (University of California Cooperative Extension, 2014).
53. Rogan, G. & Fitzpatrick, S. *Petition for the determi- nation of Nonregulated Status: Roundup Ready Alfalfa (Medicago sativa L.) Events J101 and J163* (USDA- APHIS, 2004).
54. Samac, D. A. & Smigocki, A. C. Expression of oryza- cystatin I and II in alfalfa increases resistance to the root-lesion nematode. *Phytopathology* **93**, 799–804 (2003).
55. Rosellini, D., Capomaccio, S., Ferradini, N., Savo Sardaro, M. L. & Nicolia, A. Non-antibiotic, efficient selection for alfalfa genetic engineering. *Plant. Cell. Rep.* **26**, 1035–1044 (2007).
56. Sun, M. et al. Ectopic expression of GsSRK in *Medicago sativa* reveals its involvement in plant architecture and salt stress responses. *Front. Plant. Sci.* **9**, 226 (2018).
57. Rosellini, D., Pezzotti, M. & Veronesi, F. Characterization of transgenic male sterility in alfalfa. *Euphytica* **118**, 313–319 (2001).
58. Wei, Z., Liu, Y., Lin, C., Wang, Y. & Cai, Q. Transformation of alfalfa chloroplasts and expression of green fluorescent protein in a forage crop. *Biotech. Nol Lett.* **33**, 2487–2494 (2011).
59. Yang, S., Gao, M., Xu, C., Gao, J. & Deshpande, S. Alfalfa benefits from *Medicago truncatula*: the RCT1 gene from *M. Truncatula* confers broad-spectrum resistance to anthracnose in alfalfa. *Proc. Natl. Acad. Sci. USA*. **105**, 12164–12169 (2008).
60. Wang, Z. Y. & Brummer, E. C. Is genetic engineering ever going to take off in forage, turf and bioenergy crop breeding? *Ann. Bot.* **110**, 1317–1325 (2012).
61. Ullah, A. H., Sethumadhavan, K., Mullaney, E. J., Ziegelhoffer, T. & Austin-Phillips, S. Cloned and expressed fungal phya gene in alfalfa produces a stable phytase. *Biochem. Biophys. Res. Commun.* **290**, 1343–1348 (2002).
62. Samac, D. & Temple, S. J. Development and utilization of transformation in *Medicago* species. In *Genetically Modified Crops: Their Development, Uses, and Risks* (eds Liang, G. & Skinner, D. Z.) 165–202 (Haworth, 2004).
63. Vail, A. W., Wang, P., Uefuji, H., Samac, D. A. & Vance, C. P. Biodegradation of atrazine by three transgenic grasses and alfalfa expressing a modified bacterial atrazine chlorohydrolase gene. *Transgenic Res.* **24**, 475–488 (2014).
64. Waltz, E. Industry exhales as USDA okays glyphosate-resistant alfalfa. *Nat. Biotechnol.* **29**, 179–181 (2011).
65. Teuber, L. R., van Deynze, A., Mueller, S., McCaslin, M. & Fitzpatrick, S. Gene flow in alfalfa under honeybee (*Apis mellifera*) pollination. In *Proceedings of the 39th North American Alfalfa Improvement Conference* (2004).
66. Tesfaye, M., Temple, S. J., Allan, D. L., Vance, C. P. & Samac, D. A. Overexpression of malate dehydrogenase in transgenic alfalfa enhances organic acid synthesis and confers tolerance to aluminum. *Plant. Physiol.* **127**, 1836–1844 (2001).

Author contributions

L. F. Yan and Z. P. Liu conceived and designed the experiment. B. C. Han, X.M. Dong and C. C. Shi performed the experiments. B. C. Han and X.M. Dong analyzed all the data. B. C. Han wrote this manuscript. Z. P. Liu revised the manuscript. All the authors read and approved the final manuscript.

Funding

National Center of Pratacultural Technology Innovation (under preparation)Special fund for innovation platform construction: CCPTZX2023N04; Leading Scientist Project of Gansu Province: 23ZDKA013; Gansu Provincial Science and Technology Major Projects: 23ZDNA009; the earmarked fund for CARS (CARS-34).

Declarations

Competing interests

The authors declare no competing interests.

Ethics approval and consent to participate

Field and laboratory studies were conducted by local legislation. This article does not contain any studies with human participants or animals and does not involve any endangered or protected species. The plant materials sampled, and experiments performed in this research complied with institutional, national, and international guidelines and legislation.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-87185-3.

Correspondence and requests for materials should be addressed to L.Y.