Legume Science

WILEY

Legume Science Open Access

ORIGINAL ARTICLE OPEN ACCESS

Phenotypic Diversity and Nutritional Profiling of a Mung Bean (*Vigna radiata* L.) Germplasm Collection for Functional Food Breeding

Rong Cao[1] | Qiannan Song[1] | Liangliang Hu[1] | Tianxiao Chen[1] | Suhua Wang[1] | Lixia Wang[1] | Yang Yao[1] | Xuzhen Cheng[1] | Jindong Fu[1] | Xinbo Guo[2] | Honglin Chen[1]

[1]Northeast Asia Agricultural Science and Technology Innovation Center/Laboratory of Quality Evaluation and Nutrition Health of Agro-products, Ministry of Agriculture and Rural Affairs, Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing, China | [2]School of Food Science and Engineering, South China University of Technology, Guangzhou, China

Correspondence: Jindong Fu (fujindong@caas.cn) | Xinbo Guo (guoxinbo@scut.edu.cn) | Honglin Chen (chenhonglin@caas.cn)

Received: 30 September 2025 | **Revised:** 16 March 2026 | **Accepted:** 6 April 2026

Keywords: correlation analysis | cluster analysis | Mung bean (*Vigna radiata* L.) | nutritional quality | principal component analysis (PCA)

ABSTRACT

Mung bean (*Vigna radiata* L.) is a nutritionally valuable legume; however, the lack of a comprehensive evaluation of quality traits across diverse germplasm hinders targeted breeding efforts. To identify superior genetic resources, this study assessed key nutritional quality traits, including crude protein, starch, fat, total polyphenols, total flavonoids, caffeic acid, vitexin, and isovitexin, across 370 diverse mung bean accessions. Phenotypic data were evaluated utilizing correlation analysis, principal component analysis (PCA), and cluster analysis. Significant phenotypic variation was observed, with coefficients of variation ranging from 5.05% to 37.41% and Shannon's diversity indices between 0.10 and 0.59. Key correlations revealed a negative relationship between protein and starch, alongside strong positive associations among total polyphenols, flavonoids, caffeic acid, vitexin, and isovitexin. Furthermore, cluster analysis stratified the germplasm into five distinct groups exhibiting significant differences in quality traits. A comprehensive evaluation successfully identified elite accessions characterized by high overall nutritional profiles. These findings provide a strategic roadmap for targeted mung bean improvement, offering pre-selected elite germplasm and clear trait–association insights to accelerate the development of nutritionally superior and functionally specialized cultivars.

1 | Introduction

Mung bean (*Vigna radiata* (L.) Wilczek) is a vital global legume recognized for its rapid growth, drought tolerance, and impressive nutritional profile (Chen et al. 2015; Mehandi et al. 2019). Cultivated extensively across Asia, Africa, and South America, it serves as a staple pulse crop (Nair and Schreinemachers 2020; Pataczek et al. 2018). As a primary center of origin for mung bean, China maintains substantial production with approximately 273,570 ha cultivated and 389,300 tons produced in 2022 (Li et al. 2017; National Bureau of Statistics of China 2022). Valued for its high protein, complex carbohydrates, and low-fat content, mung bean significantly contributes to dietary regimes and helps combat protein–energy malnutrition (Dahiya et al. 2015; Hou et al. 2019). Its nutritional value is further enhanced by a favorable amino acid composition (Du et al. 2018; Huang et al. 2024) and the presence of bioactive compounds, such as polyphenols and flavonoids, which exhibit antioxidant, hypoglycemic, and antihypertensive properties (Deepika et al. 2023). Additionally, its functional properties, including

Rong Cao and Qiannan Song contributed equally to this study.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *Legume Science* published by Wiley Periodicals LLC.

moisture retention and emulsification, support the development of novel value-added food products (Jeong and Cho 2024; Wang et al. 2022).

Despite this established nutritional significance, the escalating global demand for plant-based proteins and functional foods underscores the critical need to systematically characterize natural variation within mung bean genetic resources. Building on previous assessments of molecular diversity and population structure (Chen et al. 2015), this characterization is essential to advance functional food-oriented breeding. However, current evaluation practices in mung bean breeding remain fragmented, often confined to narrow metabolite profiling across limited samples (Shen et al. 2025), mineral-focused assessments (Wu et al. 2020), or yield-centric genetic analyses (Yang et al. 2025). These studies lack the integration of core macronutrients (protein, starch, fat) with functionally critical bioactive phytochemicals (e.g., vitexin, isovitexin, anthocyanins) into a unified, breeding-actionable framework. This fragmentation directly impedes breeders' ability to identify elite germplasm possessing superior or synergistic trait combinations essential for functional food development. Traditional single-trait selection approaches prove inadequate for optimizing multiple nutritional objectives simultaneously within mung bean breeding programs, particularly where unfavorable correlations (e.g., protein–starch trade-offs) complicate advancement. Although multi-trait evaluation strategies offer a promising pathway for developing nutritionally balanced cultivars, their systematic deployment across large-scale mung bean germplasm resources remains unestablished. Consequently, a critical gap persists in mung bean breeding: the absence of a scalable, phenotype-driven framework to efficiently translate germplasm diversity into actionable selection criteria. Bridging this gap is essential to empower breeders to develop next-generation cultivars that meet evolving consumer expectations for nutrient-dense, health-promoting plant-based foods. Therefore, this study aimed to systematically evaluate the phenotypic diversity of key macronutrients and phytochemicals across a diverse collection of 370 mung bean accessions, as well as to elucidate the complex interrelationships among these quality traits. Ultimately, this research seeks to identify elite germplasm with outstanding nutritional profiles for direct use in breeding programs and to propose targeted hybridization strategies for functional food-oriented breeding.

2 | Materials and Methods

2.1 | Plant Materials and Field Trials

A diverse panel of 370 mung bean accessions was used in this study, comprising 351 domestic and 19 foreign accessions from 29 regions. The germplasm was sourced from the Chinese National Genebank at the Institute of Crop Sciences, Chinese Academy of Agricultural Sciences (ICS-CAAS). The accessions exhibited a broad spectrum of seed colors and predominantly consisted of breeding lines and landraces from China, alongside a smaller proportion of wild species (Figure 1). Field trials were conducted during the summer of 2023 in Beijing, China (39.9127° N, 116.3972° E), under standard agronomic practices. This single-location, single-season design serves as a primary platform for screening and identifying elite germplasm, with subsequent multi-environment trials (MET) required to validate trait stability across diverse agroecologies. After reaching physiological maturity, pods were harvested, sun-dried, and manually threshed. Cleaned seeds were ground into a fine powder (60-mesh) and stored at 4°C prior to analysis.

2.2 | Determination of Crude Protein, Starch, and Fat Content

2.2.1 | Crude Protein Content (CPC)

The CPC was determined using the Kjeldahl method, following a previously described method (Multescu et al. 2024). Briefly, 0.5 g of prepared mung bean powder was digested with 0.4 g

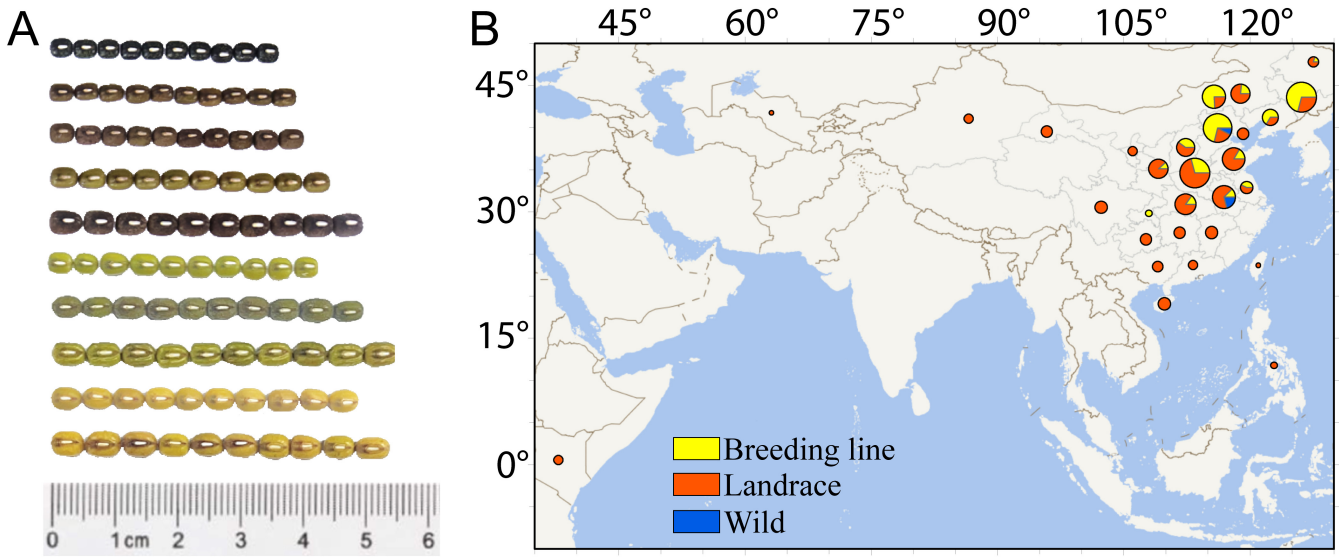


FIGURE 1 | Morphological and geographical diversity of the mung bean collection. (A) Different types of mung bean germplasm; (B) Geographical origin of the mung bean accessions and the proportion of different germplasm types.

$CuSO_4$, 6 g K_2SO_4, and 10 mL H_2SO_4 at 420°C for 1 h until a clear green solution was obtained. Post-digestion, the protein content was assessed using an automatic Kjeldahl nitrogen analyzer (K9860, Hanon Advanced Technology Group Co. Ltd., China). Each sample was analyzed in three biological replicates, with the average value reported.

2.2.2 | Crude Starch Content (CSC)

The CSC was analyzed according to a previous method (Umnajkitikorn et al. 2024). Mung bean powder was dissolved in sodium hydroxide and gelatinized in a boiling water bath. The gelatinized sample was then transferred to a volumetric flask, mixed with deionized water, acetic acid, and an iodine solution, and made up to volume. Absorbance was measured at 610 nm against a blank reference. The amylose content was calculated using a standard curve prepared from an amylose-amylopectin mixture.

2.2.3 | Crude Fat Content (CFC)

CFC was determined via Soxhlet extraction following a reported procedure (Kan et al. 2018). Approximately 2 g of mung bean powder was placed in a filter paper cartridge and extracted continuously with petroleum ether (30°C–60°C boiling range) using a Soxhlet apparatus. The extraction was considered complete when no oil residue was detected upon evaporation of a solvent drop. The extract was collected, dried, cooled, and weighed until a constant weight was achieved. All measurements were performed in triplicate.

2.3 | Determination of Total Polyphenol and Flavonoid Content

The mung bean samples were dried at 60°C for 10 h and ground into a fine powder. For extraction, 2 g of the powder was mixed with 20 mL of 70% ethanol (v/v) and shaken at 60°C and 250 rpm for 2 h. After centrifugation at 7000 rpm for 8 min, the supernatant was collected. The residue underwent a second extraction under identical conditions and the combined supernatants were retained for subsequent analysis. All extraction procedures were repeated three times.

Total polyphenol content (TPC) was determined using the Folin–Ciocalteu method, and total flavonoid content (TFC) was assessed via an aluminum chloride colorimetric assay (Siriparu et al. 2022). For TPC determination, the extract was reacted with Folin–Ciocalteu reagent, followed by the addition of a sodium carbonate solution. The mixture was then incubated in the dark for 2 h and the absorbance was measured at 765 nm. TPC was quantified based on a gallic acid standard curve and expressed as mg of gallic acid equivalents per gram of dry weight (mg GAE/g DW). For TFC determination, the extracts were sequentially treated with sodium nitrite, aluminum chloride, and sodium hydroxide. Following incubation, absorbance was measured at 510 nm. TFC was calculated using a rutin standard curve and expressed as mg of rutin equivalents per gram of dry weight (mg RE/g DW). All spectrophotometric measurements were performed using a SpectraMax Plus 384-well microplate reader.

2.4 | UPLC-MS Analysis of Individual Phenolic Compounds

The contents of caffeic acid, vitexin, and isovitexin were quantified using a UPLC system coupled to a triple quadrupole mass spectrometer (TQD) (Waters, Milford, MA, United States), following the gradient elution program (Yao et al. 2012). The analysis was performed on a Waters ACQUITY UPLC BEH Shield RP18 column (2.1 × 100 mm; 1.7 μm). The mobile phases consisted of methanol (A) and a 0.1% (v/v) formic acid aqueous solution (B). The flow rate was 0.25 mL/min, the injection volume was 10 μL, and the column and autosampler temperatures were maintained at 35°C and 10°C, respectively. Compounds were quantified using external standard calibration curves.

2.5 | Data Analysis

Phenotypic data analysis of the quality traits was conducted utilizing Excel 2010 and R v4.4.3 software (R Core Team 2024). Descriptive statistics were used to calculate means, standard deviations, standard errors, and coefficients of variation (CVs). Pearson's correlation coefficients were calculated to determine relationships between quality traits. Duncan's Multiple Range Test was employed for post hoc mean comparisons following ANOVA. The GGally (Schloerke et al. 2024), FactoMineR (Lê et al. 2008), factoextra (Kassambara and Mundt 2020), ggplot2 (Wickham 2016), and ggtree (Yu et al. 2017) packages in RStudio were utilized to calculate correlation coefficients and to generate scatter plots, scree plots, principal component analysis (PCA) diagrams, and cluster analysis diagrams. PCA was performed to reduce the data dimensionality and identify the main components contributing to the variation in quality traits (Ringnér 2008). Components with a cumulative contribution rate of more than 80% were considered significant. Hierarchical cluster analysis was conducted using Ward's method and squared Euclidean distance to group accessions based on their quality traits, enabling the identification of distinct clusters with similar trait profiles. The calculation of membership degree was determined using the following formula: $U\left(X_{ij}\right) = \left(X_{i,j} - X_{i,\min}\right) / \left(X_{i,\max} - X_{i,\min}\right)$. Where U represents the membership degree; $X_{i,j}$ denotes the phenotype of the i-th trait of the j-th accession ($i, j = 1, 2, ..., n$); and $X_{i,\min}$ and $X_{i,\max}$ are the minimum and maximum values of the i-th trait, respectively. This method normalizes data to a 0–1 scale to eliminate unit differences, ensuring that each trait contributes equally to the comprehensive score (D-value), which is a standard and reproducible method for germplasm evaluation.

3 | Results

3.1 | Phenotypic Variation in Morphological and Nutritional Quality Traits

The 370 mung bean accessions displayed significant phenotypic diversity across both morphological (Table S1) and nutritional traits (Table 1). Morphologically, most accessions were of

TABLE 1 | Descriptive statistics analysis of the nutritional quality traits across the mung bean germplasm.

Traits	Max	Min	Mean	SD	SE	CV (%)
Crude protein content (%)	30.94	22.07	25.99	1.51	0.08	5.81
Crude starch content (%)	53.96	42.48	48.19	2.43	0.13	5.05
Crude fat content (%)	1.41	0.28	0.81	0.20	0.01	24.12
Total polyphenol content (mg GAE/g DW)	2.89	0.94	1.66	0.32	0.02	19.36
Total flavonoid content (mg RE/g DW)	9.22	2.13	4.69	1.33	0.07	28.28
Caffeic acid content (μg/g DW)	46.17	6.11	17.70	6.01	0.31	33.97
Vitexin content (μg/g DW)	209.63	35.39	91.71	28.06	1.46	30.59
Isovitexin content (μg/g DW)	209.01	27.43	79.80	29.85	1.55	37.41

Abbreviations: CV, coefficient of variation; SD, standard deviation; SE, standard error.

domestic origin (94.9%), predominantly exhibiting green (90.3%) and smooth (71.4%) seeds of medium size (53.2%). The germplasm comprised mainly landraces (61.9%) and breeding lines (35.9%).

A comprehensive assessment of nutritional quality revealed substantial variation across all eight measured traits. CVs ranged from 5.05% for crude starch to 37.41% for isovitexin. Notably, primary metabolites (crude protein and crude starch) showed lower variability (CVs of 5.81% and 5.05%, respectively) compared with secondary metabolites. Secondary metabolites, including total polyphenols (TPC; CV = 19.36%), total flavonoids (TFC; CV = 28.28%), and their derivatives like isovitexin (CV = 37.41%), exhibited considerably higher variation. This pronounced variability in phytochemicals suggests that their biosynthetic pathways are more genetically plastic and environmentally responsive, as previously demonstrated by the significant impact of environmental factors such as photoperiod on polyphenol biosynthesis and antioxidant capacity (Cheng et al. 2022). This indicates significant potential for targeted genetic improvement. On average, the germplasm contained 25.99% crude protein, 48.19% crude starch, and 4.69 mg RE/g DW TFC (Table 2), providing a robust genetic foundation for mung bean enhancement. Accessions with notably high or low values for specific traits are detailed in Table S1.

3.2 | Trait Distribution Patterns and Interrelationships

Frequency distribution analysis revealed distinct patterns among metabolite classes (Figure 2). Primary metabolites (protein, starch, and fat) generally exhibited approximately normal distributions, consistent with their essential roles under stabilizing selection. In contrast, all secondary metabolites showed right-skewed distributions, characterized by a prevalence of accessions with low-to-moderate content and a long tail of high-content outliers. This asymmetry highlights the potential to identify rare, elite genotypes with superior nutraceutical properties. Correlation analysis further elucidated the complex biochemical relationships among quality traits (Figure 3). A significant negative correlation was observed between crude protein and crude starch ($r = -0.37$, $p < 0.001$), indicating a fundamental metabolic trade-off between nitrogen and carbon allocation during seed development. Conversely, strong positive correlations were found within the phenolic compound groups: between TPC and TFC ($r = 0.82$, $p < 0.001$), and between vitexin and isovitexin ($r = 0.97$, $p < 0.001$). These robust positive associations suggest co-regulated biosynthetic pathways, offering a clear opportunity for breeders to simultaneously select for multiple antioxidant compounds.

3.3 | Influence of Germplasm Origin and Phenotype on Nutritional Quality

To identify phenotype-biochemistry linkages, nutritional quality was analyzed based on predefined germplasm groupings (Table 2, Figure 4). Variation in nutritional profiles was strongly associated with specific morphological traits and domestication status. Seed coat phenotype was a significant determinant of nutritional content. Specifically, brown-seeded accessions were characterized by the highest levels of crude protein, fat, and phenolic compounds compared to other color groups. Furthermore, the presence of seed coat texture influenced primary metabolite distribution; hairy-seeded accessions contained significantly higher protein but lower starch content than those with smooth coats. In contrast, seed size did not show any statistically significant correlation with the measured quality traits.

The domestication degree also played a critical role in biochemical differentiation. Wild accessions exhibited significantly higher levels of crude protein, fat, TPC, and TFC compared with domesticated landraces and breeding lines. This suggests that wild relatives have retained higher concentrations of both primary nutrients and defense-related secondary metabolites, likely due to the absence of selection pressure for high yield that typically characterizes the domestication process.

Regarding the geographical origin of the germplasm, its influence on nutritional quality was found to be highly specific. While most nutritional parameters remained consistent across different regions, geographic origin exerted a significant effect on caffeic acid content. Specifically, accessions originating from international sources displayed significantly higher caffeic acid levels than those collected within domestic regions.

TABLE 2 | Differences in nutritional quality traits among mung bean accessions based on distinct morphological categories and origin.

Traits	Category	Crude protein content	Crude starch content	Crude fat content	Total polyphenol content	Total flavonoid content	Caffeic acid content	Vitexin content	Isovitexin content
Seed coat color	Yellow	25.22 ± 0.37b	48.84 ± 0.64ab	0.75 ± 0.05bc	1.50 ± 0.08c	4.13 ± 0.31c	19.44 ± 1.15a	101.18 ± 4.87a	89.83 ± 5.36a
	Green	25.95 ± 0.08b	48.18 ± 0.13ab	0.81 ± 0.01bc	1.63 ± 0.01c	4.58 ± 0.06bc	17.70 ± 0.33a	92.12 ± 1.57a	80.19 ± 1.67ab
	Brown	29.26 ± 0.48a	50.06 ± 0.74a	1.19 ± 0.06a	2.62 ± 0.05a	8.45 ± 0.35a	14.64 ± 1.93a	69.02 ± 7.99a	54.22 ± 6.46b
	Dark green	25.66 ± 0.54b	46.39 ± 0.80b	0.64 ± 0.08c	2.06 ± 0.32b	5.96 ± 1.36b	21.11 ± 3.27a	93.67 ± 7.21a	77.63 ± 8.62ab
	Black	26.48 ± 0.40b	47.58 ± 0.68ab	0.91 ± 0.07b	2.11 ± 0.13b	5.99 ± 0.44b	15.68 ± 1.29a	78.74 ± 5.84a	70.11 ± 6.19ab
Seed size	Small	26.21 ± 0.22a	48.10 ± 0.34a	0.84 ± 0.03a	1.68 ± 0.06a	4.83 ± 0.25a	18.54 ± 0.79a	91.42 ± 4.50a	78.99 ± 4.75a
	Medium	25.79 ± 0.11a	48.28 ± 0.19a	0.81 ± 0.01a	1.62 ± 0.02a	4.53 ± 0.08a	17.62 ± 0.40a	93.35 ± 1.95a	81.65 ± 2.11a
	Large	26.20 ± 0.13a	48.09 ± 0.19a	0.80 ± 0.02a	1.72 ± 0.03a	4.88 ± 0.11a	17.43 ± 0.62a	89.11 ± 2.45a	77.09 ± 2.55a
Seed coat texture	Smooth	25.75 ± 0.09b	48.45 ± 0.14a	0.77 ± 0.01b	1.71 ± 0.02a	4.90 ± 0.07a	17.56 ± 0.37a	91.17 ± 1.68a	79.71 ± 1.81a
	Hairy	26.58 ± 0.16a	47.55 ± 0.27b	0.93 ± 0.02a	1.54 ± 0.04b	4.16 ± 0.14b	18.04 ± 0.57a	93.07 ± 2.94a	80.02 ± 3.02a
Germplasm type	Breeding line	25.63 ± 0.12b	48.15 ± 0.16a	0.82 ± 0.02b	1.76 ± 0.02b	5.13 ± 0.10b	17.51 ± 0.52a	89.5 ± 2.20a	78.55 ± 2.39a
	Landrace	26.11 ± 0.10b	48.17 ± 0.18a	0.80 ± 0.01b	1.57 ± 0.02b	4.29 ± 0.07b	17.72 ± 0.40a	93.35 ± 1.96a	80.99 ± 2.07a
	Wild	28.18 ± 0.67a	49.56 ± 1.06a	1.03 ± 0.11a	2.61 ± 0.08a	8.69 ± 0.19a	20.28 ± 2.65a	81.61 ± 7.37a	66.60 ± 7.12a
Geographic origin	China	26.00 ± 0.08a	48.23 ± 0.13a	0.81 ± 0.01a	1.66 ± 0.02a	4.69 ± 0.07a	17.52 ± 0.32b	91.09 ± 1.49a	79.02 ± 1.57a
	Other countries	25.71 ± 0.41a	47.59 ± 0.74a	0.86 ± 0.04a	1.70 ± 0.07a	4.71 ± 0.25a	21.03 ± 1.14a	103.18 ± 6.84a	94.16 ± 8.14a

Note: Data are expressed as means ± standard error. Different lowercase letters within the same column indicate significant differences at $p < 0.05$.

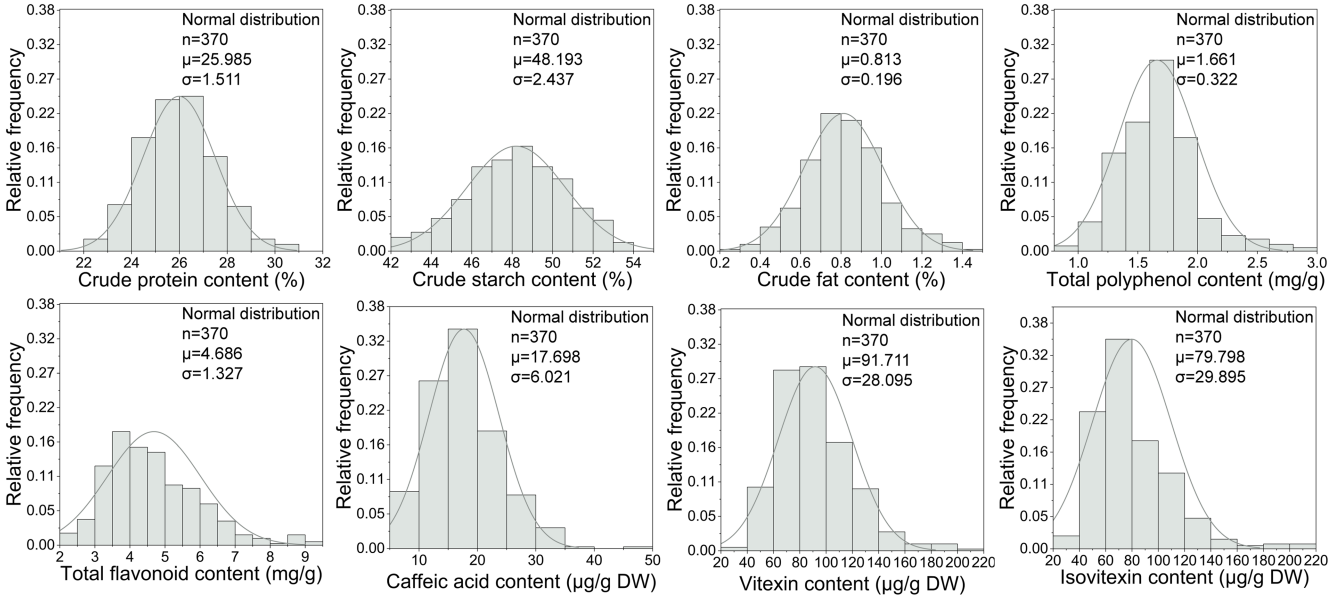


FIGURE 2 | Histograms illustrating the frequency distribution of the eight evaluated nutritional quality traits.

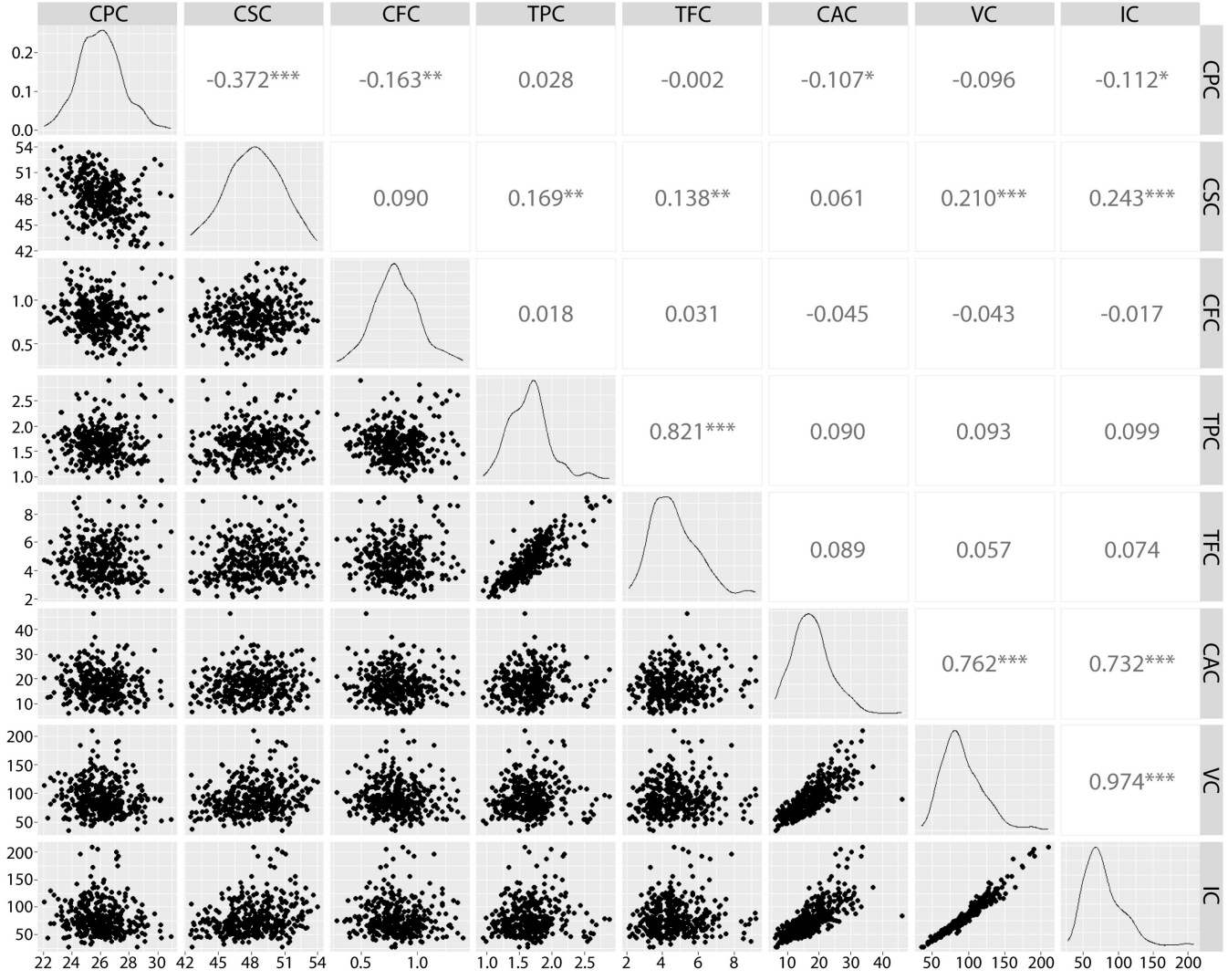


FIGURE 3 | Correlation matrix, distribution plots, and scatter plots evaluating the eight quality traits in *V. radiata*. Asterisks indicate statistical significance: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. Abbreviations: CAC, caffeic acid content; CFC, crude fat content; CPC, crude protein content; CSC, crude starch content; IC, isovitexin content; TFC, total flavonoid content; TPC, total polyphenol content; VC, vitexin content.

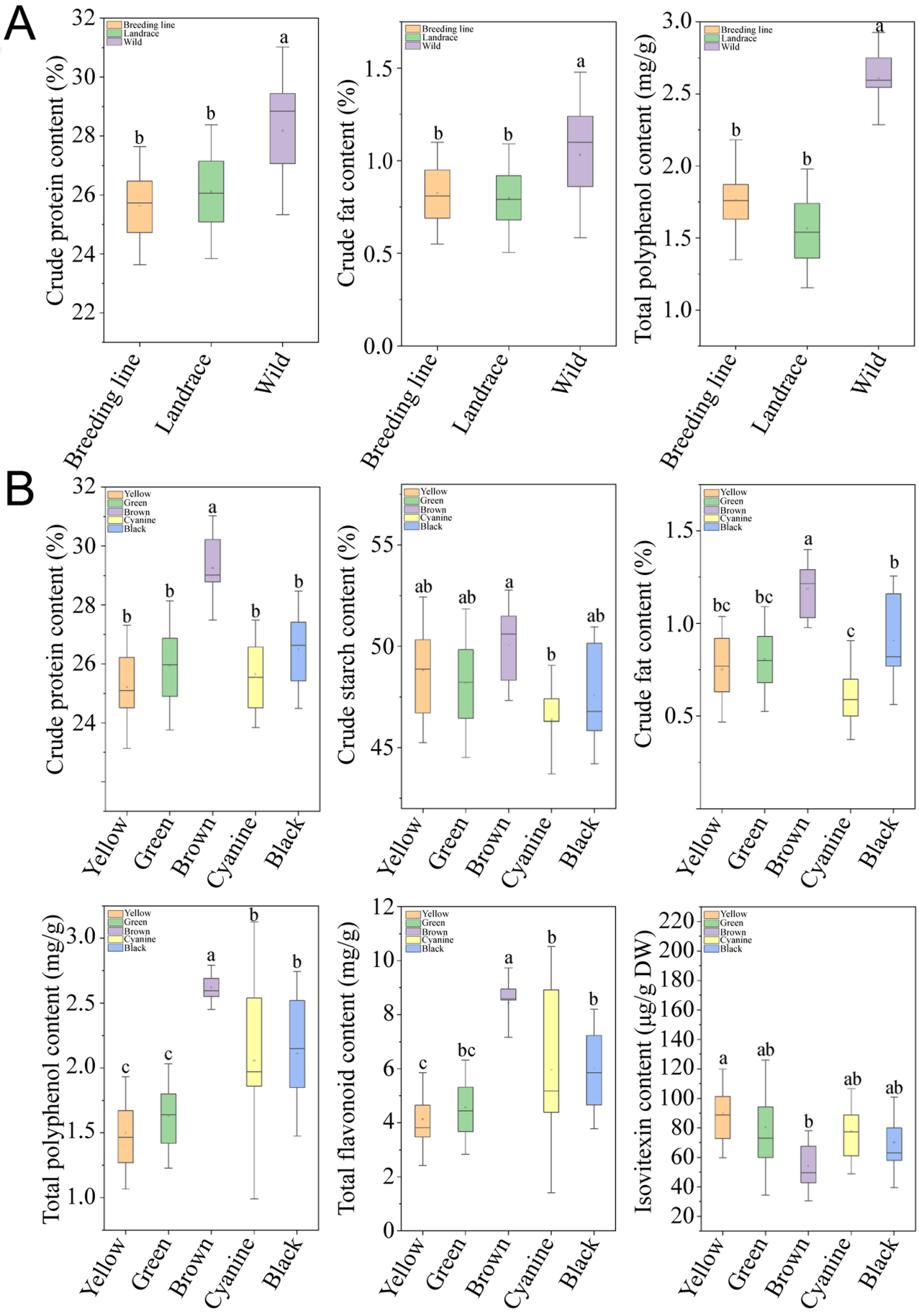


FIGURE 4 | Legend on next page.

FIGURE 4 | Box plots illustrating significant differences in the eight nutritional traits among mung bean germplasm groups. (A) Content of nutritional traits across mung beans of different germplasm types; (B) Content of nutritional traits across mung beans of different seed coat colors. Groups with significantly different means are indicated by different lowercase letters.

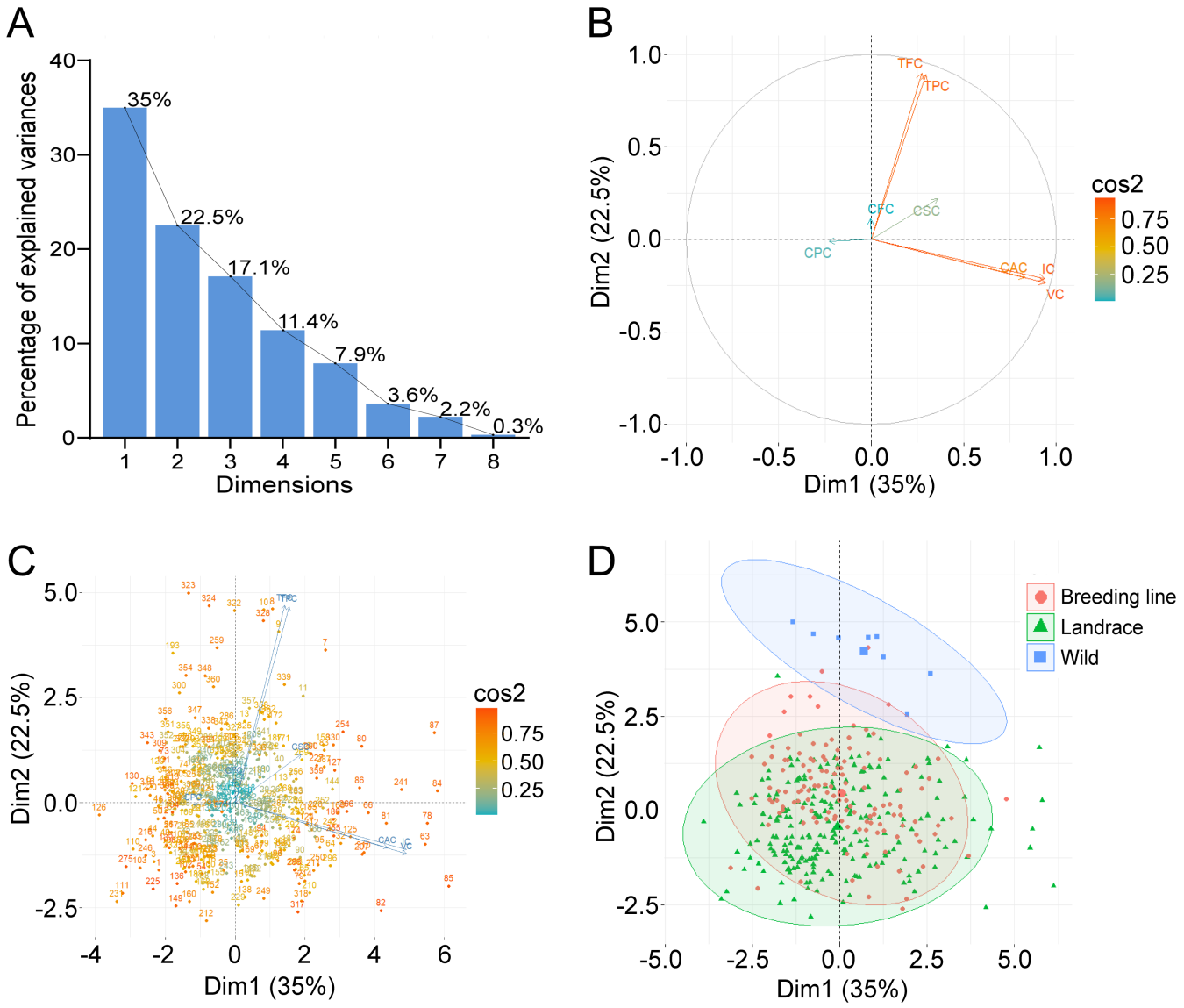


FIGURE 5 | Multivariate analysis of the evaluated quality traits. (A) Scree plot illustrating the variance explained by the principal components as a percentage contribution of each component. (B) PCA loading plot detailing the contribution and correlation of each variable to the principal components. (C) PCA biplot illustrating the distribution of quality traits among the 370 accessions based on the loadings of the eight nutritional traits. (D) Grouping of breeding lines, landraces and wild species of *V. radiata* based on their nutritional profiles.

3.4 | PCA-Defined Orthogonal Metabolic Axes in Nutritional Trait Diversification

PCA reduced the dimensionality of the quality trait data, with the initial four principal components (PCs) collectively accounting for 86.05% of the total variation among the accessions (Table 4). Their respective eigenvalues were 2.80, 1.80, 1.37, and 0.91, explaining 34.99%, 22.51%, 17.14%, and 11.41% of the total variance. The scree plot (Figure 5A) visually confirmed that these initial dimensions captured the majority of the variation.

The PCA loading plot (Figure 5B) effectively delineated complex trait relationships. PC1 (34.99% variance) was primarily driven by strong positive loadings for vitexin (0.94), isovitexin (0.94), and caffeic acid (0.83), interpretable as the "Antioxidant Glycoside Axis". PC2 (22.51% variance) was characterized by high positive loadings for TFC (0.90) and TPC (0.89), which clustered tightly, representing the "Core Phenolic Axis". PC3 (17.14% variance) was primarily driven by a strong positive loading for crude protein (0.78) and a strong negative loading for crude starch (−0.64). PC4 (11.41% variance) was dominated by a

high positive loading for crude fat (0.84). Together, PC3 and PC4 represent the "Primary Metabolite Axes", effectively separating crude protein from starch and isolating crude fat. This orthogonal separation of trait groups reinforces findings from the correlation analysis, visually confirming that the primary drivers of diversity are linked antioxidant compounds (PC1 and PC2) and a distinct trade-off between protein and starch (PC3). The PCA biplot (Figure 5C) illustrates how these trait loadings project onto the first two PCs, clearly separating accessions based on their combined phenolic and primary metabolite profiles. The individual PCA plot (Figure 5D) revealed considerable overlap between breeding lines and landraces. However, landraces (green dots) showed greater dispersion across the plot than the more tightly grouped breeding lines (orange dots), suggesting broader trait diversity within the landrace collection.

3.5 | Cluster Analysis and Characterization of Germplasm Groups

Cluster analysis classified the 370 mung bean accessions into five functionally distinct groups based on their quality trait profiles (Figure 6A). The unique characteristics of each cluster are detailed below and visualized in Figure 6B. Group I ($n = 101$) had the highest median CPC but relatively low starch and the lowest median concentrations of TPC and TFC. This profile is valuable for protein-focused nutritional enhancement. Group II ($n = 72$), characterized by the lowest median values for CFC, caffeic acid (CAC), vitexin (VC), and isovitexin (IC), represents a baseline or lower-metabolite profile for these compounds. Group III ($n = 93$), comprising primarily breeding lines and landraces, had the highest median CSC with moderate levels of other compounds, making it ideal for high-starch yield applications. Group IV ($n = 36$) displayed the highest median levels of CFC, TPC, and TFC. The box plots show a particularly wide range of variation for phenolic and flavonoid content, suggesting diverse genotypes for these traits. This group is a key resource for developing novel ingredients high in both fat and phenolics. Group V ($n = 68$) exhibited the lowest median CPC but was exceptionally rich in bioactive compounds, with the highest median concentrations of caffeic acid (CAC), vitexin (VC), and isovitexin (IC), significantly exceeding those of all other groups. This "Bioactive Elite" group is highly promising for functional food development and breeding programs aimed at enhancing antioxidant properties. Collectively, this clustering provides a clear roadmap for breeders by systematically categorizing the germplasm into functionally distinct pools, including "high protein" (Group I), "high starch" (Group III), "high fat and phenolic" (Group IV), and "bioactive elite" (Group V). This transforms the diverse collection into a well-characterized and readily usable resource for the targeted genetic improvement of mung bean.

3.6 | Screening and Identification of Elite Germplasm Resources

A comprehensive quality score was calculated for each accession using the membership function method (Table S1). *D*-scores ranged from 0.20 to 0.69 across the 370 accessions. The top 10 accessions based on this comprehensive evaluation were C01521, C01494, AHW-18, AHW-19, C01014, C04487, C01477, HBW106-11, C01497, and AHW-03. These elite accessions comprised four wild accessions (AHW-18, AHW-19, HBW106-11, and AHW-03), five landraces (C01521, C01494, C01014, C01477, and C01497), and one breeding line (C04487). Notably, among these elite accessions, wild accessions (AHW-18, AHW-19, HBW106-11, and AHW-03) demonstrated superior performance, suggesting that the outstanding traits of wild relatives hold significant potential as valuable genetic resources for enhancing mung bean quality.

4 | Discussion

Mung bean (*V. radiata* L.) is an increasingly recognized pulse crop valued for its high nutritional density and diverse bioactive compounds (Sehrawat et al. 2023). Comprehending and evaluating the genetic diversity of its quality traits is paramount

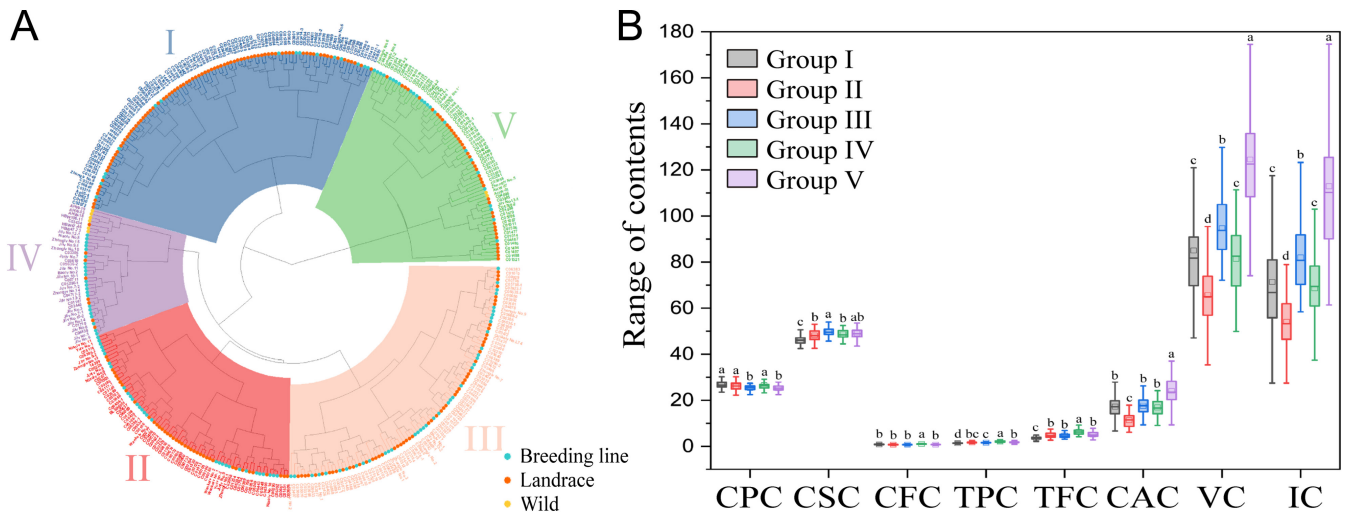


FIGURE 6 | Cluster analysis of the 370 mung bean accessions. (A) Circular dendrogram illustrating the clustering of the 370 mung bean accessions; (B) Box plots displaying the mean, standard errors, maximum, and minimum values, alongside the differences between five distinct groups across the eight quantitative traits. Different lowercase letters indicate significant differences between groups ($p < 0.05$).

for developing new cultivars tailored for specific food processing applications and enhanced nutritional benefits (Zhao et al. 2022). Our comprehensive study contributes significantly to this understanding by characterizing a large, diverse mung bean germplasm collection, identifying elite resources, and outlining strategies for targeted quality improvement. The extensive phenotypic diversity observed in our collection, particularly for secondary metabolites, corroborates previous reports on mung bean and other pulses (Lazaridi et al. 2017; Manjunatha et al. 2023; Bhagyawant et al. 2019). This significant genetic variation underscores the untapped genetic potential in mung bean germplasm, particularly in the application of functional foods, and helps to predict breeding strategies (Bhadkaria et al. 2020). The predominantly Chinese origin of our germplasm highlights China's role as a vital center of diversity, affirming its importance as a resource for novel trait discovery. Such broad variation enables the selection of accessions with specific high-value traits, such as exceptionally high protein content for protein-rich food ingredients, or high levels of polyphenols and flavonoids for enhanced antioxidant properties. Notably, the variability in secondary metabolites far exceeded that of primary metabolites like protein and starch, suggesting that the biosynthetic pathways governing these health-promoting phytochemicals are highly plastic and offer significant potential for targeted enhancement.

Our analysis confirmed a fundamental metabolic trade-off, evidenced by a significant negative correlation between crude protein and crude starch, which is a well-documented phenomenon across various legume species reflecting carbon-nitrogen allocation dynamics during seed filling (Bueno et al. 2018; Zhou et al. 2023). This inherent physiological bottleneck presents a breeding challenge; however, we identified rare accessions, particularly among wild relatives and landraces, that concurrently accumulate elevated protein and starch. This provides empirical evidence that the protein-starch trade-off is not absolute in mung bean germplasm, redirecting focus toward leveraging these "outlier" genotypes for developing nutritionally balanced cultivars. Moreover, the strong positive correlations within phenolic groups reveal coordinated biosynthetic regulation. The near-perfect correlation between vitexin and isovitexin indicates shared enzymatic control and a fixed ratio of carbon flux into C-glycosylflavones. Similarly, the tight coupling of TPC and TFC suggests co-expression of core phenylpropanoid pathway components under coordinated metabolic control. Interestingly, the weak but significant positive correlations observed between crude starch and key phenolics (vitexin and isovitexin) may reflect a unique developmental synchrony where active starch deposition coincides with a transient upregulation of the photosynthetic carbon pool utilized for flavonoid glycosylation. Critically, vitexin content showed no significant correlation with protein content, confirming its metabolic uncoupling from the protein-starch dynamics and suggesting that antioxidant capacity can be improved independently of protein levels, which aligns with recent findings on the specific metabolic regulation of vitexin-like compounds during mung bean seed development (Zhao et al. 2025).

The elevated levels of secondary metabolites in wild accessions align with established theories of plant domestication. Wild relatives, facing greater environmental and predatory pressures, often retain robust defense-related compounds, whereas domesticated crops are typically selected for yield and palatability, sometimes at the expense of these phytochemicals (Zhang et al. 2019). Our findings specifically highlight that brown-seeded and wild accessions possess the highest levels of crude protein, fat, and phenolic compounds. This association between darker seed coat color and higher phenolic content is supported by research in other crops (Marathe et al. 2011; Segev et al. 2010; Slavin et al. 2009), where pigmentation regulation is intricately linked to the phenylpropanoid pathway. Furthermore, we observed that accessions with hairy seed coats contained higher protein but lower starch than those with smooth coats, while seed size did not significantly influence quality traits, providing practical morphological markers for initial germplasm screening.

Our hierarchical cluster analysis successfully classified the 370 accessions into five functionally distinct groups, providing a practical framework for germplasm utilization. This systematic categorization allows for the identification of specific breeding pools, such as the "High Protein" Group I and the "Bioactive Elite" Group V. Such a clear demarcation directly informs strategic hybridization efforts, enabling breeders to pyramid desirable alleles from different groups to develop nutritionally complete cultivars. The top 10 elite accessions identified (e.g., C01521, C01494) are immediately valuable as parents in breeding programs or for the development of specialized food products. The presence of wild types or landraces with wild genetic signatures among our high-ranking accessions underscores the significant potential of exotic germplasm for quality improvement (von Wettberg et al. 2018).

While the present study employed a single-location, single-season field trial to minimize environmental noise and precisely resolve inherent genetic variation in core nutritional traits, which is a deliberate design choice for initial-stage germplasm characterization, we acknowledge that genotype-by-environment (G×E) interactions may influence trait stability. This limitation is inherent to the foundational phase of screening, where establishing a robust phenotypic baseline precedes multi-environment validation. Critically, the elite accessions and functional groups identified herein serve as a prioritized subset for immediate advancement into China's national mung bean breeding network for multi-location, multi-year trials across key production regions. Future research should focus on validating the stability of these quality traits across diverse agro-ecological zones and expanding the germplasm pool with international accessions. By bridging this comprehensive phenotypic characterization with systematic multi-environment validation and metabolic profiling, this roadmap empowers breeders to develop next-generation cultivars. Furthermore, integrating these evaluations with high-resolution genomic resources, as in recent precision breeding advancements in closely related species such as adzuki bean (Hu et al. 2026), will be crucial for meeting the urgent global demand for nutrient-dense, health-promoting plant-based foods. In conclusion, this study comprehensively characterized the diversity of nutritional and phytochemical traits across 370 mung bean accessions, revealing key metabolic relationships and identifying elite germplasm with superior quality profiles. Wild relatives (AHW-18, AHW-19, HBW106-11, and AHW-03) demonstrated outstanding performance, highlighting their value as genetic resources for

breeding. Furthermore, to fully leverage these findings, future research should prioritize MET to verify the stability of these quality traits across diverse agro-ecological zones. Additionally, genome-wide association studies (GWAS) are recommended to dissect the genetic basis of the accumulation of key bioactives, such as vitexin and isovitexin. These subsequent steps will effectively translate the metabolic synergies observed herein into actionable tools for targeted breeding.

Author Contributions

Rong Cao: investigation, validation, formal analysis, writing – original draft. **Qiannan Song:** methodology, data curation. **Liangliang Hu:** software, formal analysis. **Tianxiao Chen:** methodology, data curation. **Suhua Wang:** resources, investigation. **Lixia Wang:** resources, investigation. **Yang Yao:** methodology, investigation. **Xuzhen Cheng:** resources, investigation. **Jindong Fu:** methodology, validation, writing – review and editing. **Xinbo Guo:** methodology, data curation, writing – review and editing. **Honglin Chen:** conceptualization, funding acquisition, supervision, project administration, resources, investigation, writing – review and editing.

Acknowledgments

The authors thank Dr. Li Wu, Dr. Longguo Jin, and Dr. Lin Guo (ICS-CAAS) for technical assistance. This research was funded by the Northeast Asia Agricultural Science and Technology Innovation Program of CAAS, the China Agricultural Research System (CARS-08), and the Technology Innovation Program of CAAS.

Funding

This research was supported by the Northeast Asia Agricultural Science and Technology Innovation Program of CAAS, the China Agricultural Research System (CARS-08), and the Technology Innovation Program of CAAS.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Bhadkaria, A., N. Gupta, D. Narvekar, et al. 2020. "ISSR-PCR Approach as a Means of Studying Genetic Variation in Moth Bean (*Vigna aconitifolia* (Jacq.) Maréchal)." *Biocatalysis and Agricultural Biotechnology* 30: 101827. https://doi.org/10.1016/j.bcab.2020.101827.

Bhagyawant, S., A. Bhadkaria, D. Narvekar, et al. 2019. "Multivariate Biochemical Characterization of Rice Bean (*Vigna umbellata*) Seeds for Nutritional Enhancement." *Biocatalysis and Agricultural Biotechnology* 20, no. 5: 101193. https://doi.org/10.1016/j.bcab.2019.101193.

Bueno, R., L. Borges, P. God, et al. 2018. "Quantification of Anti-Nutritional Factors and Their Correlations With Protein and Oil in Soybeans." *Anais da Academia Brasileira de Ciências* 90, no. 1: 205–217. https://doi.org/10.1590/0001-3765201820140465.

Chen, H., L. Qiao, L. Wang, S. Wang, M. W. Blair, and X. Cheng. 2015. "Assessment of Genetic Diversity and Population Structure of Mung Bean (*Vigna radiata*) Germplasm Using EST-Based and Genomic SSR Markers." *Gene* 566, no. 2: 175–183. https://doi.org/10.1016/j.gene.2015.04.043.

Cheng, Y., N. Xiang, X. Cheng, H. Chen, and X. Guo. 2022. "Effect of Photoperiod on Polyphenol Biosynthesis and Cellular Antioxidant Capacity in Mung Bean (*Vigna radiata*) Sprouts." *Food Research International* 159: 111626. https://doi.org/10.1016/j.foodres.2022.111626.

Dahiya, P., A. Linnemann, M. Van Boekel, et al. 2015. "Mung Bean: Technological and Nutritional Potential." *Critical Reviews in Food Science and Nutrition* 55, no. 5: 670–688. https://doi.org/10.1080/10408398.2012.671202.

Deepika, D., K. Tripathi, G. Abhishek, et al. 2023. "Identification of BCMV Resistant Germplasm in Mung Bean (*Vigna radiata*) Using Serological and Molecular Diagnostics." *Indian Journal of Agricultural Sciences* 93, no. 1: 98–101. https://doi.org/10.56093/ijas.v93i1.123053.

Du, M., J. Xie, B. Gong, et al. 2018. "Extraction, Physicochemical Characteristics and Functional Properties of Mung Bean Protein." *Food Hydrocolloids* 76: 131–140. https://doi.org/10.1016/j.foodhyd.2017.01.003.

Hou, D., L. Yousaf, Y. Xue, et al. 2019. "Mung Bean (*Vigna radiata* L.): Bioactive Polyphenols, Polysaccharides, Peptides, and Health Benefits." *Nutrients* 11, no. 6: 1238. https://doi.org/10.3390/nu11061238.

Hu, L., X. Wang, Y. Zhang, et al. 2026. "High-Resolution Genomic Resources for Trait Mapping and Precision Breeding for Adzuki Bean (*Vigna angularis*)." *Advanced Science* 13, no. 4: e07157. https://doi.org/10.1002/advs.202507157.

Huang, Z., Y. Li, M. Fan, H. Qian, and L. Wang. 2024. "Recent Advances in Mung Bean Protein: From Structure, Function to Application." *International Journal of Biological Macromolecules* 273: 133210. https://doi.org/10.1016/j.ijbiomac.2024.133210.

Jeong, M., and S. Cho. 2024. "Effect of pH-Shifting on the Water Holding Capacity and Gelation Properties of Mung Bean Protein Isolate." *Food Research International* 177: 113912. https://doi.org/10.1016/j.foodres.2023.113912.

Kan, L., S. Nie, J. Hu, et al. 2018. "Comparative Study on the Chemical Composition, Anthocyanins, Tocopherols and Carotenoids of Selected Legumes." *Food Chemistry* 260: 317–326. https://doi.org/10.1016/j.foodchem.2018.03.148.

Kassambara, A., and F. Mundt. 2020. "factoextra: Extract and Visualize the Results of Multivariate Data Analyses." https://CRAN.R-project.org/package=factoextra.

Lazaridi, E., G. Ntatsi, J. Fernández, et al. 2017. "Phenotypic Diversity and Evaluation of Fresh Pods of Cowpea Landraces From Southern Europe." *Journal of the Science of Food and Agriculture* 97, no. 13: 4326–4333. https://doi.org/10.1002/jsfa.8249.

Lê, S., J. Josse, and F. Husson. 2008. "FactoMineR: A Package for Multivariate Analysis." *Journal of Statistical Software* 25, no. 1: 1–18. https://doi.org/10.18637/jss.v025.i01.

Li, L., T. Yang, R. Liu, B. Redden, F. Maalouf, and X. Zong. 2017. "Food Legume Production in China." *Crop Journal* 5, no. 2: 115–126. https://doi.org/10.1016/j.cj.2016.06.001.

Manjunatha, P., M. Aski, G. Mishra, et al. 2023. "Genome-Wide Association Studies for Phenological and Agronomic Traits in Mung Bean (*Vigna radiata* L. Wilczek)." *Frontiers in Plant Science* 14: 1209288. https://doi.org/10.3389/fpls.2023.1209288.

Marathe, S., V. Rajalakshmi, S. Jamdar, and A. Sharma. 2011. "Comparative Study on Antioxidant Activity of Different Varieties of Commonly Consumed Legumes in India." *Food and Chemical Toxicology* 49, no. 9: 2005–2012. https://doi.org/10.1016/j.fct.2011.04.039.

Mehandi, S., S. M. Quatadah, S. P. Mishra, I. P. Singh, N. Praveen, and N. Dwivedi. 2019. "Mung Bean (*Vigna radiata* L. Wilczek): Retrospect and Prospects." In *Legume Crops - Characterization and Breeding for*

Improved Food Security, pp. Ch., 5. IntechOpen. https://doi.org/10.5772/intechopen.85657.

Multescu, M., A. Culetu, and I. Susman. 2024. "Screening of the Nutritional Properties, Bioactive Components, and Antioxidant Properties in Legumes." *Foods* 13, no. 22: 3528. https://doi.org/10.3390/foods13223528.

Nair, R., and P. Schreinemachers. 2020. "Global Status and Economic Importance of Mung Bean." In *The Mung Bean Genome. Compendium of Plant Genomes*. Springer. https://doi.org/10.1007/978-3-030-20008-4_1.

National Bureau of Statistics of China. 2022. "China Statistical Yearbook—2022." National Bureau of Statistics of China Retrieved from https://data.stats.gov.cn/english.

Pataczek, L., Z. Zahir, M. Maqshoof Ahmad, et al. 2018. "Beans With Benefits—The Role of Mung Bean (*Vigna radiata*) in a Changing Environment." *American Journal of Plant Sciences* 9, no. 7: 1577–1600. https://doi.org/10.4236/ajps.2018.97115.

R Core Team. 2024. "R: A Language and Environment for Statistical Computing." R Foundation for Statistical Computing. https://www.R-project.org.

Ringnér, M. 2008. "What is Principal Component Analysis?" *Nature Biotechnology* 26, no. 3: 303–304. https://doi.org/10.1038/nbt0308-303.

Schloerke, B., D. Cook, J. Larmarange, et al. 2024. "GGally: Extension to 'ggplot2'." https://CRAN.R-project.org/package=GGally.

Segev, A., H. Badani, Y. Kapulnik, I. Shomer, M. Oren-Shamir, and S. Galili. 2010. "Determination of Polyphenols, Flavonoids, and Antioxidant Capacity in Colored Chickpea (*Cicer arietinum* L.)." *Journal of Food Science* 75, no. 2: S115–S119. https://doi.org/10.1111/j.1750-3841.2009.01477.x.

Sehrawat, N., M. Yadav, S. Kumar, et al. 2023. "Mung Bean as a Potent Emerging Functional Food Having Anticancer Therapeutic Potential: Mechanistic Insight and Recent Updates." *Biotechnology and Applied Biochemistry* 70, no. 6: 2002–2016. https://doi.org/10.1002/bab.2505.

Shen, X., H. Shen, F. Wang, et al. 2025. "Comparative Non-Targeted Metabolomics Reveals Distinct Metabolic Profiles and Functional Traits in Six Mung Bean (*Vigna radiata*) Varieties." *PLoS ONE* 20, no. 7: e0327962. https://doi.org/10.1371/journal.pone.0327962.

Siriparu, P., P. Panyatip, T. Pota, et al. 2022. "Effect of Germination and Illumination on Melatonin and Its Metabolites, Phenolic Content, and Antioxidant Activity in Mung Bean Sprouts." *Plants* 11, no. 21: 2990. https://doi.org/10.3390/plants11212990.

Slavin, M., W. Kenworthy, and L. Yu. 2009. "Antioxidant Properties, Phytochemical Composition, and Antiproliferative Activity of Maryland-Grown Soybeans With Colored Seed Coats." *Journal of Agricultural and Food Chemistry* 57, no. 23: 11174–11185. https://doi.org/10.1021/jf902609n.

Umnajkitikorn, K., P. Boonchuen, R. Senavongse, et al. 2024. "Transcriptomics and Starch Biosynthesis Analysis in Leaves and Developing Seeds of Mung Bean Provide a Basis for Genetic Engineering of Starch Composition and Seed Quality." *Frontiers in Plant Science* 15: 1332150. https://doi.org/10.3389/fpls.2024.1332150.

von Wettberg, E., P. Chang, F. Başdemir, et al. 2018. "Ecology and Genomics of an Important Crop Wild Relative as a Prelude to Agricultural Innovation." *Nature Communications* 9, no. 1: 649. https://doi.org/10.1038/s41467-018-02867-z.

Wang, Y., J. Zhao, S. Zhang, et al. 2022. "Structural and Rheological Properties of Mung Bean Protein Emulsion as a Liquid Egg Substitute: The Effect of pH Shifting and Calcium." *Food Hydrocolloids* 126: 107485. https://doi.org/10.1016/j.foodhyd.2022.107485.

Wickham, H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag. https://ggplot2.tidyverse.org.

Wu, X., A. S. M. F. Islam, N. Limpot, et al. 2020. "Genome-Wide SNP Identification and Association Mapping for Seed Mineral Concentration in Mung Bean (*Vigna radiata* L.)." *Frontiers in Genetics* 11: 656. https://doi.org/10.3389/fgene.2020.00656.

Yang, F., J. Qiao, B. Li, et al. 2025. "Comprehensive Genetic Dissection of Yield-Related Traits Utilizing Quantitative Trait Loci Sequencing Approach in Mung Bean." *Journal of Plant Physiology* 312: 154552. https://doi.org/10.1016/j.jplph.2025.154552.

Yao, Y., X. Cheng, L. Wang, et al. 2012. "Major Phenolic Compounds, Antioxidant Capacity and Antidiabetic Potential of Rice Bean (*Vigna umbellata* L.) in China." *International Journal of Molecular Sciences* 13, no. 3: 2707–2716. https://doi.org/10.3390/ijms13032707.

Yu, G., D. Smith, H. Zhu, et al. 2017. "ggtree: An R Package for Visualization and Annotation of Phylogenetic Trees With Their Covariates and Other Associated Data." *Methods in Ecology and Evolution* 8, no. 1: 28–36. https://doi.org/10.1111/2041-210X.12628.

Zhang, H., F. Yasmin, and B. Song. 2019. "Neglected Treasures in the Wild—Legume Wild Relatives in Food Security and Human Health." *Current Opinion in Plant Biology* 49: 17–26. https://doi.org/10.1016/j.pbi.2019.04.004.

Zhao, T., X. Meng, C. Chen, L. Wang, X. Cheng, and W. Xue. 2022. "Agronomic Traits, Fresh Food Processing Characteristics and Sensory Quality of 26 Mung Bean (*Vigna radiata* L.) Cultivars (Fabaceae) in China." *Foods* 11, no. 12: 1687. https://doi.org/10.3390/foods11121687.

Zhao, Y., N. Xiang, Y. Cheng, H. Chen, and X. Guo. 2025. "Metabolic Regulation of Vitexin-Like Compound Accumulation During Mung Bean (*Vigna radiata*) Seed Development." *Journal of the Science of Food and Agriculture* 106, no. 1: 81–91. https://doi.org/10.1002/jsfa.70141.

Zhou, H., Y. Liu, B. Mu, F. Wang, N. Feng, and D. Zheng. 2023. "Nitrogen Limitation Affects Carbon and Nitrogen Metabolism in Mung Bean (*Vigna radiata* L.)." *Journal of Plant Physiology* 290: 154105. https://doi.org/10.1016/j.jplph.2023.154105.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Table S1:** Comprehensive morphological, nutritional, and geographical data of the 370 mung bean accessions.