Variation in protein and amino acids in global collection of pearl millet (*Pennisetum glaucum*) germplasm

Satbeer Singh [a,b], Devvart Yadav [c], Manfred Beckmann [d], Aavula Naveen [e], Prakash I. Gangashetty [f], Luis A.J. Mur [d], Rattan S. Yadav [a,*]

[a] *Institute of Biological Environmental and Rural Sciences (IBERS), Aberystwyth University, Aberystwyth SY23 3EE, United Kingdom*
[b] *Division of Agrotechnology, Council of Scientific and Industrial Research (CSIR), Institute of Himalayan Bioresource Technology, Palampur, Himachal Pradesh 176061, India*
[c] *Chaudhary Charan Singh (CCS) Haryana Agricultural University, Hisar 125004, India*
[d] *Department of Life Sciences, Aberystwyth University, Aberystwyth SY23 3FL, United Kingdom*
[e] *Indian Agricultural Research Institute (IARI), Pusa, New Delhi 110012, India*
[f] *International Crops Research Institute for the Semi-Arid Tropics, Patancheru, India*

ARTICLE INFO

ABSTRACT

Pearl millet is a major source of daily protein intake in south Asia and sub-Saharan Africa. Despite considerable importance, the extent of variation in protein and amino acids in pearl millet global germplasm is unknown. The present study assessed 165 genotypes from within the Pearl Millet Inbred Germplasm Association Panel (PMi-GAP), that includes breeding lines, landraces and improved cultivars randomly drawn from a core collection from 23 countries, for protein content and 18 amino acids. The results showed considerable variation for protein content (10.06 – 20.31 %) and amino acids in PMiGAP. Diverse patterns were observed across the geographical distribution and clustered the germplasm into 7 clusters, with one cluster ("2") containing most of the superior properties. Most amino acid levels were positively correlated but these were negatively correlated with protein content. A set of twelve genotypes was identified having higher protein with better amino acid compositions. These superior genotypes could directly feed into global and regional pearl millet improvement programs to counter hidden hunger in developing countries. We propose that these findings can be combined with the starches, lipids, antioxidants, micronutrients, and other healthful traits for which the PMiGAP resource has been extensively studied.

1. Introduction

According to the estimates of the Food and Agriculture Organization (FAO) of the United Nations, more than 843 million people worldwide are hungry, and ~ 2.5 billion are suffering from nutrient deficiencies. In particular, protein malnutrition affects about one billion people worldwide (Wu et al., 2014) and is particularly prevalent in sub–Saharan Africa and south Asia. Pearl millet (*Pennisetum glaucum*) is a dominant crop in many parts of sub Saharan Africa and India (Yadav & Rai, 2013). Indeed, due to its ability to withstand drought and harsh environments, it is the mainstay of a large population in Africa and Asia. Pearl millet has high carbohydrate levels, high dietary fibre (11.3 g/100 g, most of which is insoluble), greater α-amylase activity, low glycaemic index (<50), and is gluten-free (Saleh et al., 2013; Satyavathi et al., 2021).

These characteristics make pearl millet an ideal model crop for use in functional food industries.

Protein and amino acid contents are the major considerations in assessing the relative qualities in any food grain (Shewry, 2007). The available literature shows that total protein content of pearl millet ranges between 8.07–18.15 % (Tomar et al., 2021) of grain dry weight which is higher than maize and rice and comparable with wheat and sorghum (Tomicic et al., 2022), but less than with legumes (Anitha et al., 2020). The amino acid profiles of world's major cereals i.e., pearl millets, sorghum, maize, barley, wheat and rice are almost similar except for lysine, tyrosine, leucine, isoleucine and tryptophan (Ejeta et al., 1987; Hassan et al., 2021). Pearl millet had an average amino acids content of 0.67 %, however, the average levels of non-essential amino acids was higher (1.22 %). Moreover, the essential amino acids,

* Corresponding author.
  E-mail address: rsy@aber.ac.uk (R.S. Yadav).

methionine (0.31 %), histidine (0.39 %), and lysine (0.43 %) are lower than with leucine (1.7 %) and phenylalanine (0.71 %) (Kapustin et al., 2023). The levels of the limiting essential amino acid - lysine - are comparatively higher than normally found in sorghum and maize, while comparable to high-lysine sorghum and opaque 2 maize lines (Anitha et al., 2020; Hassan et al., 2021; Tomar et al., 2021). Lysine levels vary between 0.43 and 0.55 g/100 g of protein in pearl millet (Verma et al., 2018). The isoleucine/leucine ratio in pearl millet is lower than that in sorghum and maize but is comparable to the ratios found in small grains (wheat, barley, and rice) (Ejeta et al., 1987). The favourable amino acid balance with higher essential amino acids makes pearl millet a nutritious source of calories and protein for human consumption.

This stated, although compared to other cereals, pearl millet possesses better protein and amino acids in their grains (Hassan et al., 2021), no systematic efforts have been made to study genetic variations of such traits at germplasm level. Detailed study of such characteristics in the germplasm would allow genetic improvement of both protein quantity and quality in future pearl millet varieties, linked to yield and adaptive traits.

Successful plant breeding programmes exploit the variability under the trait of interest (Langyan et al., 2022). From 1950s to date, the genetic diversity of pearl millet has been successfully used to increase its grain yield by 176 %. Thus, pearl millet productivity has increased from 305 to 1132 kg per hectare over these years (Singh et al., 2018; Yadav & Rai, 2013). However, whilst there have been studies examining protein and amino acids in major cereals like wheat (Jiang et al., 2008; Urosevic et al., 2023), maize (Kahriman et al., 2020; Mahan et al., 2014) and rice (Choi et al., 1990; Kamara et al., 2010), this is not the case for pearl millet germplasm. Combining good protein and amino acid levels together with high slowly digestible and resistant starches found in pearl millet (Yadav et al. 2022) will further improve the glycaemic index for pearl millet to improve the control of obesity and type-2 diabetes in humans.

Recently a pearl millet inbred germplasm association panel (PMi-GAP) has been assembled from more than 20 countries worldwide (Sehgal et al., 2015; Varshney et al., 2017; Yadav et al., 2022). Assessment of the natural variation for protein and amino acid composition in this globally recognized resource offers an ideal opportunity to exploit these traits in wider breeding programmes. The present study aimed to 1) to access the genetic variation for protein and amino acids in the PMiGAP, 2) to study patterns of recorded diversity over the genetic constitution and geographical distributions of material, 3) to identify the association among the protein content and amino acid compositions and 4) Identify superior genotypes that possess higher protein and amino acid concentration for use as donors in breeding programmes.

2. Materials and methods

2.1. Plant materials

A set of 165 genotypes was randomly drawn from the PMiGAP population of 340 genotypes (Sehgal et al., 2015). These 165 genotypes included breeding lines (n = 20), improved cultivars (n = 10) and landraces (n = 135) from 23 countries (Table S1). Seeds of the PMiGAP genotypes were obtained from International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru (India) as described in earlier studies (Ramya et al., 2018; Upadhyaya et al., 2008; Yadav et al., 2022).

2.2. Estimation of total protein content

Grain samples were analysed for total nitrogen (N) by a rapid combustion method using a LECO FP-528 analyser (LECO Corp., St. Joseph, MI, USA). Nitrogen combustion was performed by following manufacturer's recommended temperatures in the tubes (850 °C) and reduction heater (750 °C). Two analytical-grade reference materials (minimum

assay 99.9 %), L-lysine-HCl (15.32 % N) and L-tryptophan (13.71 % N), as well as soy flour (8.55 % N) (Merck Life Science UK Limited, Dorset, UK) were used to monitor nitrogen recovery. Ethylene diamine tetraacetic acid (EDTA) (9.56 ± 0.04 % N) was used for nitrogen calibration using the Dumas method. The nitrogen-to-protein factor N × 6.25 was used to calculate protein content (Krul, 2019). Protein was expressed in percent of total grain weight.

2.3. Estimation of amino acids except tryptophan

A chloroform methanol mixture (2:1, v-v) was used to extract and remove lipids. The remaining samples after the lipid extraction were air-dried and subsequently hydrolysed according to Boulos et al. (2020). Briefly, for acid hydrolysis 50 mg were treated with hydrochloric acid at 110 °C for 24 hours. Directly after cooling the samples were neutralised with sodium hydroxide and dried in a speed-vacuum concentrator. A solution of Ethanol: Water: Triethylamine: Phenyl-isothiocyanate was added to derivatize the samples. After incubation, this was dried down again and the resulting pellet was re-suspended in a sodium phosphate buffer with 5 % acetonitrile and filtered prior to being transferred into liquid chromatography vials. A dilution series of Amino Acid Standard H standard (Thermo Fisher Scientific, Carlsbad, CA, USA) was run alongside the samples to allow accurate identifications and quantification. Samples were run on a C18 analytical Pico-Tag amino acid analysis column (3.9 × 150 mm) at 38°C on a Vanquish HPLC equipped with a diode array detector set at 254 nm. Separation was achieved by a gradient of eluent A (150 mM sodium acetate, 0.05 % NEt3 and 6 % ACN, adjusted to pH 6.4 with 10 % acetic acid) and eluent B (acetonitrile: water, 6:4) starting at 0 %B, increasing to 20 %B within 5.5 min, linearly to 46 %B at 10 min, changing to 100 % B until 10.5 min, 2 min isocratic, and back to 0 %B within 0.5 min following re-equilibration with eluent A for 20 min.

2.4. Tryptophan analysis

For tryptophan analysis, 50 mg of the air-dried sample obtained after lipid extraction was hydrolysed in sodium hydroxide at 110 °C for 20 hours (Boulos et al., 2020). Directly after cooling the samples were neutralised with hydrochloric acid. Samples were then centrifuged to remove particulates and the supernatant diluted 1/10 in water before being transferred to LC vials. A dilution series of Tryptophan was run alongside the samples to allow accurate identifications and quantification. Samples were run on the same C18 column at 30°C on a Vanquish HPLC with fluorescence detector (Thermo Fisher Scientific, Carlsbad, CA, USA). The mobile phase used was isocratic sodium acetate buffer with 10 % acetonitrile. An excitation wavelength of 280 nm and emission wavelength of 340 nm was used throughout the 10-minute run time.

2.5. Statistical analysis

Analysis of variance and means were computed on replicated data using statistical software OPSTAT (Sheoran et al., 1998). To allow considerations of genetic background, geographical diversity and different global breeding programs, outliers were not removed for the further analysis. Means were used to perform Hierarchical analyses based on Euclidean distance based clustering for both traits and genotypes using 'pheatmap' function in Rstudio (Posit Team, 2023). The number of clusters (k) were verified with within cluster sum of squares and silhouette width using 'wss' and 'silhouette' functions in Rstudio. In heatmap the values were scaled in columns. Principal component analysis and Pearson`s correlation analysis was carried out using 'prcomp' (ggbiplot) and 'corr' (metan) functions in Rstudio, respectively (Posit Team, 2023).

Table 1
Summary statistics of total protein content (%) and amino acids (g/100 g protein) over all the genotypes.

Parameters	No. of samples	Mean	Minimum	Maximum	Standard deviation	Coefficient of variation	*F*-value
Total protein	165	15.49	10.06	20.31	1.85	11.92	55.58**
Histidine (His)	165	0.42	0.09	1.11	0.18	17.2	12.89**
Isoleucine (Ile)	165	1.22	0.47	2.15	0.34	8.43	25.78**
Leucine (Leu)	165	2.89	1.59	4.81	0.68	5.78	12.28**
Lysine (Lys)	165	0.22	0.03	0.76	0.12	18.22	27.36**
Methionine (Met)	165	0.21	0.05	0.56	0.11	22.13	10.87**
Phenylalanine (Phe)	165	1.11	0.56	1.94	0.26	10.62	11.97**
Threonine (Thr)	165	1.02	0.56	1.66	0.19	5.32	27.29**
Tryptophan (Trp)	165	1.58	0.00	4.00	0.55	12.22	8.21**
Valine (Val)	165	1.67	0.84	2.74	0.39	10.57	11.52**
Alanine (Ala)	165	2.17	1.20	3.26	0.45	2.83	137.70**
Arginine (Arg)	165	1.33	0.14	2.38	0.43	3.23	5.91**
Asparatate (Asp)	165	0.68	0.05	2.51	0.64	9.3	211.76**
Cysteine (Cys)	165	0.18	0.03	0.48	0.09	64.23	1.058[NS]
Glutamic Acid (Glu)	165	2.12	0.21	7.20	1.44	5.84	6.22**
Glycine (Gly)	165	1.00	0.55	1.63	0.23	7.42	17.67**
Proline (Pro)	165	1.69	1.06	2.79	0.36	2.9	131.58**
Serine (Ser)	165	1.13	0.22	1.99	0.35	4.85	89.03**
Tyrosine (Tyr)	165	0.61	0.23	1.07	0.14	9.9	12.38**

**: significant at $p<0.01$, NS: non-significant
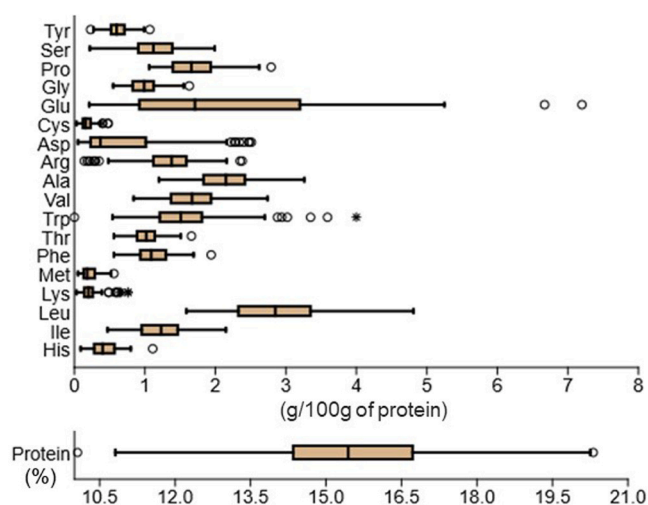
Fig. 1. Distribution of protein and amino acids across the germplasm studied.

3. Results and discussion

3.1. Protein and amino acids content per se in the germplasm

Total protein content and eighteen amino acids in 165 PMiGAP genotypes were analysed for significant differences by ANOVA. There was large significant variation in total protein content and all the amino acids except cysteine (Table 1). The total protein content varied from 10.06 % (PMiGAP088) to 20.31 % (PMiGAP168) over all the genotypes with an average of 15.49 % (Table 1). The genotypes PMiGAP168, PMiGAP311 and PMiGAP167 had highest protein content (>20 %) among all the studied genotypes (Table S1). On an average, the amino acids varied from 0.03 g/100 g of protein (cysteine) to 7.20 g/100 g of protein (Glutamic acid). The essential leucine (1.34 – 4.81 g/100 g of protein) and tryptophan (0.54 – 4.03 g/100 g of protein); and non-essential glutamic acid (0.21 – 7.20 g/100 g of protein) were the most abundant amino acids. However, essential amino acid such as lysine (0.03 – 0.76 g/100 g of protein) and methionine (0.05 – 0.56 g/100 g of protein) with non-essential amino acid cysteine (0.03 – 0.48 g/100 g of protein) were at the lowest levels. The observed range in protein content from 10.06 % to 20.31 % is interestingly higher than previously reports in pearl millet (Anitha et al., 2020; Ejeta et al., 1987) as well as in other major cereals like wheat, maize, rice and sorghum (Hajas et al., 2018;

Mulugeta et al., 2023; Shewry, 2007; Tomicic et al., 2022; Witten et al., 2020). This was because of the diverse germplasm used in the current study and indicated the importance of the PMiGAP in breeding novel varieties to meet human and animal dietary needs. The landraces such as PMiGAP168 (20.31 %), PMiGAP311 (20.31 %) and PMiGAP167 (20.31 %), and breeding lines PMiGAP346 (18.13 %) and PMiGAP038 (18.88 %), having high protein content, could straightaway be used as parents in hybrid and breeding line development programs of pearl millet for improve protein content.

The genotypes PMiGAP313, PMiGAP217, PMiGAP058 and PMiGAP150 had highest lysine content (Table S1). Further, PMiGAP307, PMiGAP317, PMiGAP229 and PMiGAP009 were highest for tryptophan, and PMiGAP133, PMiGAP025 and PMiGAP128 had highest methionine. Three genotypes *i.e.*, PMiGAP244, PMiGAP081 and PMiGAP167 had the highest concentrations for more than ten amino acids *viz* cysteine, histidine, aspartate, glutamic acid, tyrosine, phenylalanine, threonine, serine, proline, alanine, and leucine; and also had the higher protein contents. Considering all genotypes, glutamic acid and leucine had the greater degree of variation, whilst cysteine, lysine and methionine were less variable. (Fig. 1). Despite the cereals being important in human diet, they exhibited lower levels of lysine, tryptophan, leucine and other essential amino acids than found in legumes (Anitha et al., 2020). In this study, we included assessment of landraces and inbred lines that possess inbreeding depression however, most amino acid levels were comparable to the previous study on the released varieties 'Dhanshakthi' and "Proagro9444" (Anitha et al., 2020). However, higher amino acid content could be bred using those genotypes with better protein/amino acid traits identified here in this study. Such future improvement programs could especially exploit the potential of genotypes PMiGAP313, PMiGAP217, PMiGAP058 and PMiGAP150 for higher lysine content; PMiGAP307, PMiGAP317, PMiGAP229 and PMiGAP009 for higher tryptophan; and PMiGAP133, PMiGAP025 and PMiGAP128 for higher methionine. Comparatively, the three genotypes PMiGAP244, PMiGAP081 and PMiGAP167 from within the PMiGAP recorded higher protein, cysteine, histidine, aspartate, glutamic acid, tyrosine, phenylalanine, threonine, serine, proline, alanine and leucine confirming that they have great genetic potential for improving majority of the amino acids in pearl millet.

3.2. Diversity patterns of essential and non-essential amino acids across the germplasm

To study the genetic relatedness and grouping patterns of genotypes as well as amino acids, hierarchical clustering was carried out based on
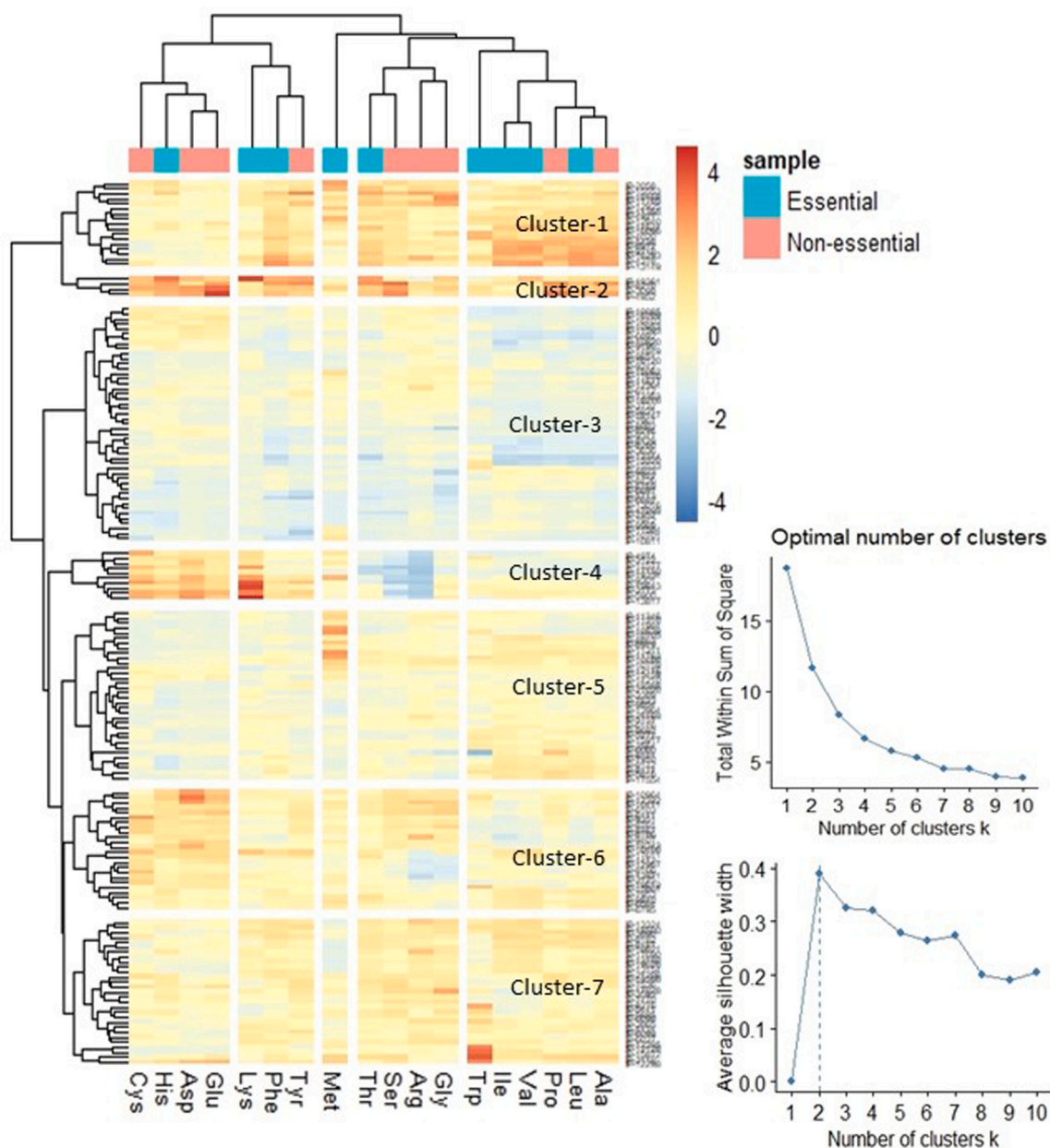
Fig. 2. Heatmap on clustering pattern of all the genotypes and amino acids.

Euclidian distances (Fig. 2). The 165 genotypes could be grouped into seven clusters. Cluster-2 comprised four genotypes (3 landraces and one breeding line) which were comparatively higher for most of the amino acids except methionine, isoleucine, and tryptophan than other genotypes/clusters. Cluster-1 contained 17 genotypes comprising 15 landraces from 12 different countries, one breeding line from Senegal and one improved cultivar from Nigeria. This cluster-1 showed above average performance for most of the amino acids. Cluster-4 comprised 10 genotypes (8 landraces, one improved cultivar and one breeding line) and, comparatively, had higher values for lysine, glutamic acid, aspartate, histidine, and cysteine, while lowest values for arginine and serine. Furthermore, cluster-6 and −7 comprised 24 and 29 genotypes, respectively and had average levels for almost all the amino acids. However, cluster-3 was the largest cluster (n = 47 genotypes) and had lowest performances for almost all the amino acids followed by cluster-5 (n =34). Surprisingly, most of the breeding lines (n=12), and half of the improved cultivars (n = 5), grouped in either cluster-3 or cluster-5, which had lower levels of almost all the amino acids.

When considering relative essential and non-essential amino acid levels these variated in almost all groups, with no clear partitioning was observed for the types of amino acids. The one exception was methionine.

3.3. Principal component analysis

Principal component analysis (PCA) was used to assess the possible relationship of protein and amino acid contents with geographical distribution (countries), genetic background and clustering patterns. The first two principal components (PC) cumulatively explained 64.0 % of the total variation (Fig. 3). Amino acids were positively correlated and explained 44.4 % of the total variation on PC1 (Fig. 3a). However, variation in protein content was totally opposite to all amino acids on PC1 (Fig. 3a). Considering the breeding lines, improved cultivars, and landraces (Fig. 3b), these overlapped in both PC1 and PC2. This indicated that amino acid and protein related variables were not improved in previous breeding programs as they maintained wide ranging
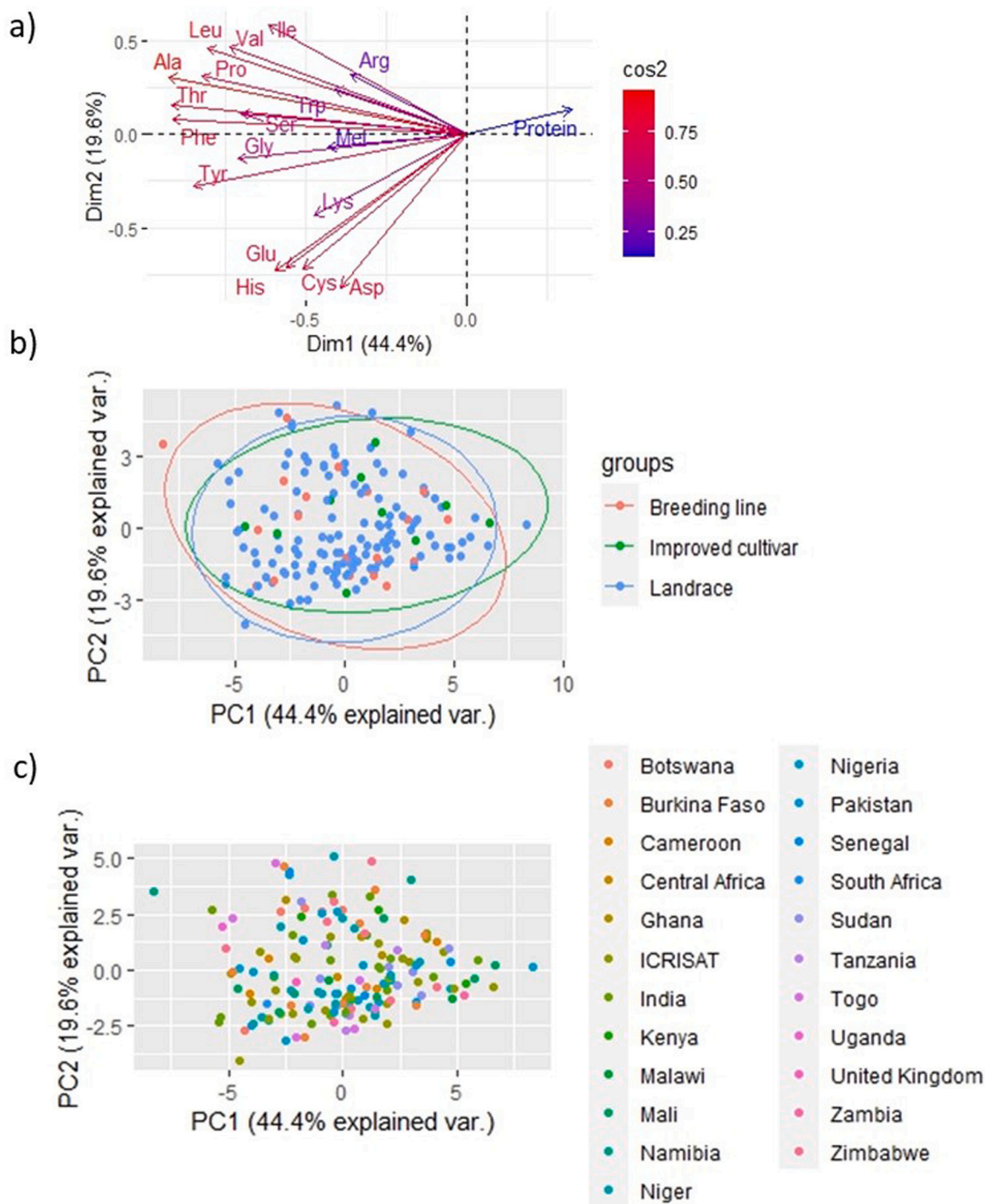
Fig. 3. Principal component analysis biplots of a) variables, b) genotype groups based on genetic background and c) geographical distribution on first two axes.

variation for these traits.

Surprisingly, the PCA conducted in the present study indicated that the amino acid/ total protein content could not differentiate between the landraces, breeding lines and improved cultivars. This was contrary to earlier reports involving pearl millet (Ramya et al., 2018; Singh and Gupta, 2019) and other crops (Crossa et al., 1994; Wang et al., 2014) that showed a structured grouping of the germplasm. This may be due to the reason that the PMiGAP panel is formed by including diverse representative lines from within the core pooling of diverse set of 1000 lines based on SSR markers (Sehgal et al., 2015). This could have countered any genetically advancement for yield related traits in breeding lines and improved cultivars that could have led to a structured

grouping pattern (Crossa et al., 1994). The fact that such structure was not observed, implies that protein and amino acid levels might not be important grain yield contributing traits. Similarly, the geographical distribution of the germplasm taken from more than 20 different countries confirmed substantial variation for all the amino acids (Fig. 3c). This variability however did not reflect the geographical origin of the material. Most of the germplasm material used in our study was of Indian and African origin, which showed high heterotic combinations in breeding history of this highly cross-pollinated crop (Singh et al., 2018). This reflects that the variability recorded with the Togo and other African types could be used with Indian and ICRISAT bred material to improve protein-related traits (Patil et al., 2021).
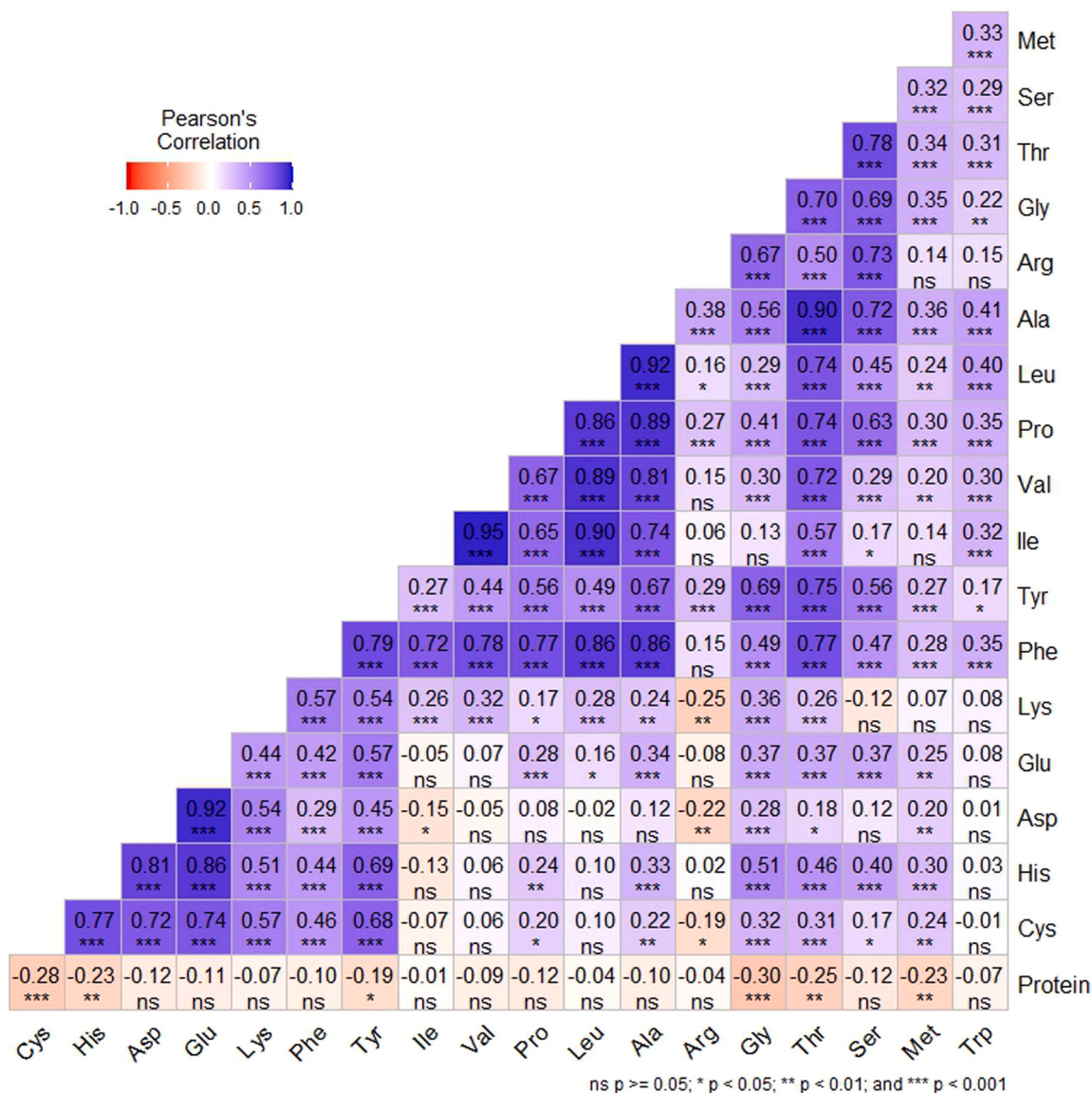
Fig. 4. Association among protein content and different amino acids.

3.4. Association among protein content and different amino acids

Direct and indirect selection on their genetic bases is aided by finding associations between protein and amino acid levels. Here Pearson's correlation coefficients were used to investigate if there were any significant associations in the PMiGAP genotypes. Total protein content was negatively correlated with almost all the amino acids (Fig. 4). The relationships are different for protein features of their structural classes in different genotypes, that may explain their negative correlation with amino acids individually. Such results are in agreement with the previous reports in cereals (Siddiqi et al., 2020; Tomar et al., 2021; Zhang et al., 2016) but not in legumes (Shi et al., 2022). This suggested that enhancement of total protein would not always lead to higher contents of targeted amino acid (Siddiqi et al., 2020). To tackle such negative association between desirable traits, multi-parent populations and backcross introgression populations would need developing to break strong negative correlations and for combining the favourable alleles of the quality traits. However, the amino acids exhibiting non-significant correlations (e.g., tryptophan, lysine, leucine, isoleucine, arginine, valine, proline, glutamic acid and aspartic acid) with protein could be genetically improved through the selection of discrete lines. Further, all

the amino acids were positively correlated, except arginine and Isoleucine, which would allow some amino acid levels to be improved in a single selection program. However, the associations of two important amino acids methionine and lysine with all other amino acids were comparatively low. Also, glutamic acid, asparagine, histidine, and cysteine had comparatively lower correlations with almost all other amino acids. This was in agreement to other reports suggesting that the strong association among amino acids leads to their direct selection in a genetic improvement program (Jiang and Katuuramu, 2021).

3.5. Identification of superior genotypes from within the PMiGAP for protein and amino acids

Pearl millet in general is considered as one of the major sources of protein in the human diet of Indian and African populations (Saleh et al., 2013; Yadav and Rai, 2013) but such considerations are based on a limited number of genotypes studied. Even though protein and amino acid content were negatively correlated in our study, twelve genotypes in the PMiGAP were identified which had a combination of higher total protein content, and higher levels of non/essential amino acids and other important amino acids (tryptophan, lysine, methionine,

Table 2

Subset of superior genotypes within PMiGAP for higher protein (%) and important amino acids (g/100 g of protein).

Superior genotypes	Total protein content	Total of nine essential amino acids	Total of nine non-essential amino acids	Important essential amino acids				
				Trypto-phan	Lysine	Methio-nine	Iso-leucine	Leucine
PMiGAP016	17.13	13.33	11.80	2.59	0.29	0.18	1.61	3.82
PMiGAP071	17.75	14.17	12.11	1.75	0.25	0.17	1.92	4.31
PMiGAP081	16.94	12.97	20.40	2.10	0.30	0.35	1.15	3.81
PMiGAP155	16.50	12.91	12.73	1.98	0.30	0.33	1.61	3.45
PMiGAP167	20.00	11.81	16.05	1.59	0.25	0.20	1.28	3.48
PMiGAP263	16.31	12.67	11.34	2.19	0.28	0.12	1.62	3.34
PMiGAP271	16.63	12.57	13.35	1.74	0.30	0.24	1.29	3.46
PMiGAP281	17.13	11.83	11.77	1.79	0.29	0.09	1.52	3.33
PMiGAP306	15.81	12.03	14.42	1.72	0.60	0.19	1.42	3.13
PMiGAP311	20.25	12.35	10.94	1.95	0.30	0.15	1.53	3.80
PMiGAP331	15.50	14.09	14.77	2.26	0.23	0.35	1.71	4.03
PMiGAP332	16.63	14.34	13.17	2.70	0.27	0.15	1.74	4.33

isoleucine, and leucine) (Table 2). Of these twelve genotypes, PMi-GAP167 and PMiGAP311 had the highest protein content (>20 %) and possessed higher amino acids than the population mean. PMiGAP081 had highest total essential (12.97 g/100 g of protein) and non-essential amino acids (20.40 g/100 g of protein) with protein content (16.94 %) higher than the population mean. Out of these twelve superior genotypes, two were of Indian origin (PMiGAP016 and PMIGAP311), one bred at ICRISAT (PMiGAP167), and the other nine were of African origin. These genotypes could be used as parents in population improvement programs, parents in hybrid breeding programs and a genetic resource for further niche area research for nutritional traits.

4. Conclusion

The present study identified wide genetic variability for protein and amino acid composition within PMiGAP to further develop it as a global resource available to improve yield and as well as to further improve health-associated traits in pearl millet. The extent of protein content recorded in the present study is much higher than reported in any pearl millet study so far confirming the extent of diversity captured in the PMiGAP and highlighting its suitability in genetic and breeding studies. The wide genetic diversity of traits reported from across the countries would favour in selecting parents from different heterotic patterns of African and Indian regions for yield and protein traits improvement for use in breeding varieties with regional and global adaptations. The genotype PMiGAP167 in have highest protein and essential amino acids like leucine, isoleucine, methionine and tryptophan. Like this, a total of twelve genotypes (PMiGAP016, PMiGAP071, PMiGAP081, PMiGAP155, PMiGAP167, PMiGAP263, PMiGAP271, PMiGAP281, PMiGAP306, PMiGAP311, PMiGAP331, PMiGAP332) were identified having superior protein and amino acid compositions in this study which can be utilised in regional and global breeding programme of pearl millet.

CRediT authorship contribution statement

Luis A. J. Mur: Writing – review & editing, Methodology, Formal analysis. **Prakash I Gangashetty:** Writing – review & editing, Resources. **Aavula Naveen:** Writing – review & editing. **Manfred Beckmann:** Methodology, Formal analysis, Data curation. **Devvart Yadav:** Writing – review & editing. **Satbeer Singh:** Writing – review & editing, Writing – original draft, Formal analysis. **Rattan S. Yadav:** Writing – review & editing, Supervision, Conceptualization.

Declaration of Competing Interest

Authors declare no conflict of interest.

Data Availability

Data will be made available on request.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jfca.2024.106557.

References

Anitha, S., Govindaraj, M., Kane-Potaka, J., 2020. Balanced amino acid and higher micronutrients in millets complements legumes for improved human dietary nutrition. Cereal Chem. 97, 74–84. https://doi.org/10.1002/cche.10227.

Boulos, S., Tännler, A., Nyström, L., 2020. Nitrogen-to-protein conversion factors for edible insects on the Swiss Market: T. molitor, A. domesticus, and L. migratoria. Front. Nutr. 7, e00089 https://doi.org/10.3389/fnut.2020.00089.

Choi, H.C., Cho, S.Y., Kim, K.H., 1990. Varietal diffrence and environmental variation in protein content and/or amino acid composition of rice seed. Korean J. Crop Sci. 35, 379–386.

Crossa, J., Taba, S., Eberhart, S.A., Bretting, P., Vencovsky, R., 1994. Practical considerations for maintaining germplasm in maize. Theor. Appl. Genet. 89, 89–95. https://doi.org/10.1007/BF00226988.

Ejeta, G., Hassen, M.M., Mertz, E.T., 1987. In vitro digestibility and amino acid composition of pearl millet (Pennisetum typhoides) and other cereals. Proc. Natl. Acad. Sci. USA 84, 6016–6019. https://doi.org/10.1073/pnas.84.17.6016.

Hajas, L., Scherf, K.A., Török, K., Bugyi, Z., Schall, E., Poms, R.E., Koehler, P., Tömösközi, S., 2018. Variation in protein composition among wheat (Triticum aestivum L.) cultivars to identify cultivars suitable as reference material for wheat gluten analysis. Food Chem. 267, 387–394. https://doi.org/10.1016/j.foodchem.2017.05.005.

Hassan, Z.M., Sebola, N.A., Mabelebele, M., 2021. The nutritional use of millet grain for food and feed: a review. Agric. Food Secur. 10, 16. https://doi.org/10.1186/s40066-020-00282-6.

Jiang, G., Katuuramu, D.N., 2021. Comparison of seed fatty and amino acids in edamame dried using two <scp>oven-drying</scp> methods and mature soybeans. J. Sci. Food Agric. 101, 1515–1522. https://doi.org/10.1002/jsfa.10766.

Jiang, X. ling, Tian, J. chun, Hao, Z., Zhang, W. dong, 2008. Protein content and amino acid composition in grains of wheat-related species. Agric. Sci. China 7, 272–279. https://doi.org/10.1016/S1671-2927(08)60066-8.

Kahriman, F., Akbulut, Ş., Düz, G., Songur, U., Egesel, C.Ö., 2020. Screening of genetic variability in Turkish maize landraces for protein and starch related traits Fatih. Maydica 65, 9.

Kamara, J.S., Konishi, S., Sasanuma, T., Abe, T., 2010. Variation in free amino acid profile among some rice (Oryza sativa L.) cultivars. Breed. Sci. 60, 46–54. https://doi.org/10.1270/jsbbs.60.46.

Kapustin, S.I., Volodin, A.B., Kapustin, A.S., Samokish, N.V., 2023. Amino acid composition of pearl millet grain and herbage. Scientific papers series management, economic engineering in agriculture and rural. Development 23, 455–460.

Krul, E.S., 2019. Calculation of nitrogen-to-protein conversion factors: a review with a focus on soy protein. J. Am. Oil Chem. Soc. 96, 339–364. https://doi.org/10.1002/aocs.12196.

Langyan, S., Bhardwaj, R., Kumari, J., Jacob, S.R., Bisht, I.S., Pandravada, S.R., Singh, A., Singh, P.B., Dar, Z.A., Kumar, A., Rana, J.C., 2022. Nutritional diversity in native germplasm of maize collected from three different fragile ecosystems of India. Front. Nutr. 9, e812599 https://doi.org/10.3389/fnut.2022.812599.

Mahan, A.L., Murray, S.C., Crosby, K., Scott, M.P., 2014. Quality protein maize germplasm characterized for amino acid profiles and endosperm opacity. Crop Sci. 54, 863–872. https://doi.org/10.2135/cropsci2012.11.0658.

Mulugeta, B., Tesfaye, K., Ortiz, R., Geleta, M., Haileselassie, T., Hammenhag, C., Hailu, F., Johansson, E., 2023. Unlocking the genetic potential of Ethiopian durum wheat landraces with high protein quality: Sources to be used in future breeding for pasta production. Food Energy Secur. 13, e511 https://doi.org/10.1002/fes3.511.

Patil, K.S., Mungra, K.D., Danam, S., Vemula, A.K., Das, R.R., Rathore, A., Gupta, S.K., 2021. Heterotic pools in African and Asian origin populations of pearl millet [Pennisetum glaucum (L.) R. Br.]. Sci. Rep. 11, e12197 https://doi.org/10.1038/s41598-021-91568-7.

Posit Team, 2023. RStudio: Integrated Development Environment for R. Posit Software, PBC, Boston, MA. ⟨http://www.posit.co/⟩.

Ramya, A.R., Ahamed M, L., Satyavathi, C.T., Rathore, A., Katiyar, P., Raj, A.G.B., Kumar, S., Gupta, R., Mahendrakar, M.D., Yadav, R.S., Srivastava, R.K., 2018. Towards defining heterotic gene pools in pearl millet [pennisetum glaucum (L.) R. Br. Front. Plant Sci. 8, e01934 https://doi.org/10.3389/fpls.2017.01934.

Saleh, A.S.M., Zhang, Q., Chen, J., Shen, Q., 2013. Millet grains: nutritional quality, processing, and potential health benefits. Compr. Rev. Food Sci. Food Saf. 12, 281–295. https://doi.org/10.1111/1541-4337.12012.

Satyavathi, C.T., Ambawat, S., Khandelwal, V., Srivastava, R.K., 2021. Pearl millet: a climate-resilient nutricereal for mitigating hidden hunger and provide nutritional security. Front. Plant Sci. 12, e659938 https://doi.org/10.3389/fpls.2021.659938.

Sehgal, D., Skot, L., Singh, R., Srivastava, R.K., Das, S.P., Taunk, J., Sharma, P.C., Pal, R., Raj, B., Hash, C.T., Yadav, R.S., 2015. Exploring potential of pearl millet germplasm association panel for association mapping of drought tolerance traits. PLoS ONE 10, e0122165. https://doi.org/10.1371/journal.pone.0122165.

Sheoran, O.P., Tonk, D.S., Kaushik, L.S., Hasija, R.C., Pannu, R.S., 1998. "Statistical software package for agricultural research workers," in Recent Adv. Inf. Theory Stat. Comput. Appl., eds D.S. Hooda, R.C. Hasija.

Shewry, P.R., 2007. Improving the protein content and composition of cereal grain. J. Cereal Sci. 46, 239–250. https://doi.org/10.1016/j.jcs.2007.06.006.

Shi, D., Hang, J., Neufeld, J., Zhao, S., House, J.D., 2022. Estimation of crude protein and amino acid contents in whole, ground and defatted ground soybeans by different types of near-infrared (NIR) reflectance spectroscopy. J. Food Compos. Anal. 111, e104601 https://doi.org/10.1016/j.jfca.2022.104601.

Siddiqi, R.A., Singh, T.P., Rani, M., Sogi, D.S., Bhat, M.A., 2020. Diversity in grain, flour, amino acid composition, protein profiling, and proportion of total flour proteins of different wheat cultivars of North India. Front. Nutr. 7, e00141 https://doi.org/10.3389/fnut.2020.00141.

Singh, S., Gupta, S.K., 2019. Formation of heterotic pools and understanding relationship between molecular divergence and heterosis in pearl millet [Pennisetum glaucum (L.) R. Br.]. PLoS ONE 14, 1–18. https://doi.org/10.1371/JOURNAL.PONE.0207463.

Singh, S., Gupta, S.K., Thudi, M., Das, R.R., Vemula, A., Garg, V., Varshney, R.K., Rathore, A., Pahuja, S.K., Yadav, D.V., 2018. Genetic Diversity Patterns and Heterosis Prediction Based on SSRs and SNPs in Hybrid Parents of Pearl Millet. Crop Sci. 58, 2379–2390. https://doi.org/10.2135/cropsci2018.03.0163.

Tomar, M., Bhardwaj, R., Kumar, M., Pal Singh, S., Krishnan, V., Kansal, R., Verma, R., Yadav, V.K., Dahuja, A., Ahlawat, S.P., Rana, J.C., Bollinedi, H., Ranjan Kumar, R., Goswami, S., T, V., Satyavathi, C.T., Praveen, S., Sachdev, A., 2021. Nutritional composition patterns and application of multivariate analysis to evaluate indigenous Pearl millet ((Pennisetum glaucum (L.) R. Br.) germplasm. J. Food Compos. Anal. 103, e104086 https://doi.org/10.1016/j.jfca.2021.104086.

Tomicic, Z., Pezo, L., Spasevski, N., Lazarević, J., Čabarkapa, I., Tomičić, R., 2022. Diversity of amino acids composition in cereals. Food Feed Res. 49, 11–22. https://doi.org/10.5937/ffr0-34322.

Upadhyaya, H.D., Reddy, K.N., Sastry, D.S.R., 2008. Regeneration guidelines: pearl millet, in: Dulloo, M.E., Thormann, I., Jorge, M.A., Hanson, J. (Eds.), Crop Specific Regeneration Guidelines [CD-ROM]. CGIAR System-wide Genetic Resource Programme, Rome, Italy., p. 9.

Urosevic, D., Knežević, D., Đurić, N., Matković Stojšin, M., Kandić, V., Mićanović, D., Stojiljković, J., Zečević, V., 2023. Assessing the potential of old and modern serbian wheat genotypes: yield components and nutritional profiles in a comprehensive study. Agronomy 13, e2426. https://doi.org/10.3390/agronomy13092426.

Varshney, R.K., Shi, C., Thudi, M., Mariac, C., Wallace, J., Qi, P., Zhang, H., Zhao, Y., Wang, X., Rathore, A., Srivastava, R.K., Chitikineni, A., Fan, G., Bajaj, P., Punnuri, S., Gupta, S.K., Wang, H., Jiang, Y., Couderc, M., Katta, M.A.K., Paudel, D.R., Mungra, K.D., Chen, W., Harris-Shultz, K.R., Garg, V., Desai, N., Doddamani, D., Kane, N.A., Conner, J.A., Ghatak, A., Chaturvedi, P., Subramaniam, S., Yadav, O.P., Berthouly-Salazar, C., Hamidou, F., Wang, Jianping, Liang, X., Clotault, J., Upadhyaya, H.D., Cubry, P., Rhoné, B., Gueye, M.C., Sunkar, R., Dupuy, C., Sparvoli, F., Cheng, S., Mahala, R.S., Singh, B., Yadav, R.S., Lyons, E., Datta, S.K., Hash, C.T., Devos, K.M., Buckler, E., Bennetzen, J.L., Paterson, A.H., Ozias-Akins, P., Grando, S., Wang, Jun, Mohapatra, T., Weckwerth, W., Reif, J.C., Liu, X., Vigouroux, Y., Xu, X., 2017. Pearl millet genome sequence provides a resource to improve agronomic traits in arid environments. Nat. Biotechnol. 35, 969–976. https://doi.org/10.1038/nbt.3943.

Verma, J., Kewat, R., Singh, P., Singh, R., Brijesh, Kumar, A., Mishra, K., 2018. Biochemical composition of pearl millet germplasm collected from Eastern UP. J. Pharmacogn. Phytochem. 7, 1888–1889.

Wang, K., Qiu, F., Dela Paz, M.A., Zhuang, J., Xie, F., 2014. Genetic diversity and structure of improved indica rice germplasm. Plant Genet. Resour. 12, 248–254. https://doi.org/10.1017/S1479262113000579.

Witten, S., Böhm, H., Aulrich, K., 2020. Effect of variety and environment on the contents of crude nutrients and amino acids in organically produced cereal and legume grains. Org. Agric. 10, 199–219. https://doi.org/10.1007/s13165-019-00261-7.

Wu, G., Fanzo, J., Miller, D.D., Pingali, P., Post, M., Steiner, J.L., Thalacker-Mercer, A.E., 2014. Production and supply of high-quality food protein for human consumption: sustainability, challenges, and innovations. Ann. N. Y. Acad. Sci. 1321, 1–19. https://doi.org/10.1111/nyas.12500.

Yadav, C.B., Srivastava, R.K., Beynon, S., Englyst, K., Gangashetty, P.I., Yadav, R.S., 2022. Genetic variability and genome-wide marker association studies for starch traits contributing to low glycaemic index in pearl millet. Food Energy Secur. 11, e341 https://doi.org/10.1002/fes3.341.

Yadav, O.P., Rai, K.N., 2013. Genetic improvement of pearl millet in India. Agric. Res. 2, 275–292. https://doi.org/10.1007/s40003-013-0089-z.

Zhang, M., Ma, D., Wang, C., Zhao, H., Zhu, Y., Guo, T., 2016. Responses of amino acid composition to nitrogen application in high- and low-protein wheat cultivars at two planting environments. Crop Sci. 56, 1277–1287. https://doi.org/10.2135/cropsci2015.08.0504.