Received: 10 May 2023 | Revised: 25 September 2023 | Accepted: 8 October 2023

DOI: 10.1002/fes3.511

Food and Energy Security Open Access WILEY

ORIGINAL ARTICLE

Unlocking the genetic potential of Ethiopian durum wheat landraces with high protein quality: Sources to be used in future breeding for pasta production

Behailu Mulugeta[1,2,3] | **Kassahun Tesfaye**[1,4] | **Rodomiro Ortiz**[2] | **Mulatu Geleta**[2] | **Teklehaimanot Haileselassie**[1] | **Cecilia Hammenhag**[2] | **Faris Hailu**[5] | **Eva Johansson**[2]

[1]Institute of Biotechnology, Addis Ababa University, Addis Ababa, Ethiopia

[2]Department of Plant Breeding, Swedish University of Agricultural Sciences, Alnarp, Sweden

[3]Sinana Agricultural Research Center, Bale-Robe, Ethiopia

[4]Bio and Emerging Technology Institute, Addis Ababa, Ethiopia

[5]Department of Biology and Biotechnology, Wollo University, Dessie, Ethiopia

Correspondence
Behailu Mulugeta, Institute of Biotechnology, Addis Ababa University, Addis Ababa, Ethiopia.
Email: behailu.mulugeta@slu.se and behailu.mulugeta30@gmail.com

Funding information
This work was funded by the Swedish International Development Cooperation Agency (Sida) grant awarded to Addis Ababa University and the Swedish University of Agricultural Sciences for a bilateral capacity#x2010;building program in biotechnology

Abstract

The content and composition of the grain storage proteins in wheat determine to a high extent its end-use quality for pasta and bread production. This study aimed to evaluate the content and composition of the grain storage proteins in Ethiopian landraces and cultivars to contribute to future breeding toward improved pasta quality. Thus, 116 landraces and 34 cultivars originating from Ethiopia were grown in three locations, and the protein parameters were analyzed using size exclusion-high performance liquid chromatography (SE-HPLC). A considerable variation in the amount of the analyzed protein parameters was found. The genotypes, environments, and interactions contributed significantly ($p < 0.001$) to the differences obtained. The broad-sense heritability was high (0.75–0.98) for all protein parameters except for unextractable small monomeric protein (uSMP). Using the principal component analysis (PCA) to evaluate the impact of protein parameters and using either PCA or unweighted pair group method with arithmetic mean (UPGMA) to assess the impact of the genetic composition, the cultivar group was found to form a separate cluster. This indicates that durum wheat improvement in Ethiopia has relied on exotic materials, which might result from a narrow genetic base. Unlike most landraces, most released cultivars showed a high and stable gluten strength across environments. Two landraces, G057 and G107, were found genetically distinct from the released cultivars but with high and stable gluten. The two selected landraces might be of extremely high value for future use in durum wheat breeding programs, as they might be adapted to wide-ranging Ethiopian growing conditions, they might carry genes of relevance to withstand abiotic and biotic stresses, and they seem to hold essential protein properties, which might result in high-quality grains for industrial processes.

KEYWORDS

durum wheat, grain storage protein, HPLC, landraces, polymeric protein

1 | INTRODUCTION

Durum wheat (*Triticum turgidum*, L. var. *durum* Desf.) is worldwide the second most produced and consumed type of wheat (Johansson et al., 2020; Shewry & Hey, 2015). The genome size of durum wheat is 12 Mbp, and the species is a disomic polyploid (2n = 4x = 28, AABB) (Maccaferri et al., 2019), which is primarily utilized for pasta products (Alvarez & Guzmán, 2018; Johansson et al., 2020). In Ethiopia, durum wheat has also been used to prepare traditional food, which includes "difo-dabo" (bread made from a thick fermented dough), "kitta" (thin unleavened bread), "injera" (a flat and spongy bread made from a fermented dough), "kinche" (boiled coarse-ground grains), "nifro" (boiled whole grains), "kollo" (roasted grains), "dabo-kollo" (small roasted pieces of unleavened dough), as well as traditional drinks, such as "tella" (fermented filtered or unfiltered drinks) and "areke" (a strong liquor) (Badebo et al., 2009; Tsegaye & Berg, 2007). The unique technological and nutritional features of durum wheat are a result of a high protein content, amber color, carotenoid pigments, vitreous grains, and semolina quality (Kadkol & Sissons, 2016; Mastrangelo & Cattivelli, 2021). Furthermore, the hardness of the durum wheat kernels is highly important, impacting the milling quality and enabling a high yield of semolina, resulting in high-quality pasta (Samaan et al., 2006).

Due to the importance of durum wheat's technological and nutritional quality, the compositions of wheat storage proteins are known to correlate with several technological traits of wheat (Johansson et al., 2020; Malik et al., 2013; Rajnincová et al., 2018). The ability of the proteins to form polymers in the grain and during processing is essential for the end-use quality of wheat (Ceresino et al., 2020; Lafiandra & Shewry, 2022; Markgren et al., 2022). Previous studies have shown that the amount and size distribution of polymeric and monomeric proteins in the wheat grain, as determined by SE-HPLC (Helguera et al., 2020; Shewry & Lafiandra, 2022), have an effect on the end-use quality of the wheat (Gupta et al., 1993; Johansson et al., 2005, 2020; Malik, 2012). The two parameters determined by SE-HPLC that are mostly related to the end-use quality of wheat are: total sodium dodecyl sulfate (SDS)-extractable proteins (TOTE) and percentage of total unextractable polymeric protein in total polymeric protein (%UPP). Previous studies have shown that TOTE correlates with total grain protein concentration, whereas %UPP correlates with gluten strength (Johansson et al., 2004, 2013; Labuschagne et al., 2004; Malik et al., 2011). Both parameters are highly affected by genotype, environment, and genotype-by-environment interactions (Gulia & Khatkar, 2015; Husenov et al., 2015; Johansson et al., 2013, 2020; Malik et al., 2011; Vida et al., 2014).

Wheat landraces are vital genetic resources that may contain desirable genes with an impact on grain quality, which in turn might result in novel products satisfying the preferences of consumers and manufacturers (Adhikari et al., 2022; Requena-Ramírez et al., 2021). Ethiopia is the major center of durum wheat genetic diversity (Harlan, 1969; Negisho et al., 2021; Savage et al., 1994). Previous studies have shown that the Ethiopian durum wheat gene pool is genetically distinct from the durum wheat gene pools of the Mediterranean region and other parts of the world (Alemu et al., 2020; Dejene et al., 2016; Kabbaj et al., 2017; Kidane et al., 2019). More than 7000 accessions of durum wheat landraces have been conserved ex situ at the Ethiopian Biodiversity Institute (EBI) gene bank and are available for research and use in breeding programs. Traditionally, durum wheat breeding in Ethiopia has focused on improving grain yield, wide adaptation, and host plant resistance to fungi diseases, insects, and weeds (Badebo et al., 2009; Tesemma & Bechere, 1998). Only a limited number of studies have focused on determining the amount and size distribution of the storage proteins in durum wheat grown in the country. This is despite the fact that some novel alleles of specific proteins have been found in local durum wheat germplasm (Hailegiorgis et al., 2020; Hailu et al., 2006). Also, the amount and size distribution of polymeric and monomeric proteins utilizing SE-HPLC have only been evaluated in a few studies that utilized a small number of local genotypes (Hailu et al., 2016; Labuschagne et al., 2004). Overall, research that targets the evaluation of a sufficient representation of the country's durum wheat gene pool for protein content and composition is lacking. This is particularly true for determining grain protein composition in terms of end-use quality and finding germplasm for use in breeding programs for developing novel cultivars with improved end-use qualities comparable with those from other parts of the world. Hence, this study aimed at (i) evaluating a large set of Ethiopian durum wheat landraces and cultivars to determine the amount and size distribution of monomeric and polymeric proteins using SE-HPLC; (ii) evaluating genotype-by-environmental interactions and population structure in relation to the variation in the amount and size distribution of monomeric and polymeric protein; (iii) evaluating the relationship between genetic diversity in the material, the protein composition, and gluten strength of the material; and (iv) identifying germplasm suitable for use in Ethiopian durum wheat breeding programs targeting end-use qualities.

2 | MATERIALS AND METHODS

2.1 | Plant material

A total of 116 landraces and 34 released cultivars (jointly designated as genotypes for simplicity hereafter) of Ethiopian durum wheat (Table S1) were used for the

present study. The 150 genotypes were sown at three selected experimental sites within well-known durum wheat cultivation areas in Ethiopia (Sinana, Kulumsa, and Chefe Donsa), following an alpha lattice design (Patterson & Williams, 1976). However, one of the genotypes was excluded from the analyses due to its poor agronomic performance at the three sites. The landrace genotypes were selected from the durum wheat accessions conserved at the EBI gene bank and are a subset of accessions described in Mulugeta et al. (2022). These landrace genotypes were grouped into ten populations (Table 1) based on their geographical regions of origin (collection sites) and genetic background. The cultivars used in this study were developed by national and regional research centers and registered as cultivars by the Ethiopian Ministry of Agriculture (MoA) after their distinctness, uniformity, and stability (DUS) were confirmed. Most of the selected cultivars are currently widely cultivated in Ethiopia. The cultivars were treated as a single population for simplicity; hence, the entire set of genotypes was grouped into 11 geographical regions of origin-based populations (Table 1). The genotypes were coded as G001 to G149 from now and onward.

2.2 | Field phenotyping

2.2.1 | Phenotypic traits

Eight phenotypic traits described previously (Mulugeta et al., 2022) for the genotypes were used in this study to evaluate their relationship with the protein composition parameters evaluated here. The phenotypic traits included were morphophenological traits (days to heading and maturity), plant architecture traits (plant height (cm), spike length (cm), and number of tillers per plant), grain yield ($t\,ha^{-1}$), and yield components (numbers of spikelets per spike and thousand kernel weight (gm)), all of which as described in the wheat descriptors (IBPGR, 1985).

2.3 | Grain storage protein composition analysis

2.3.1 | Flour preparation and protein extraction

Grains of each genotype (10 grams) harvested from all of the three field trial sites were separately sampled and milled using an Ultra Centrifugal Mill ZM 200 RESTCH at 16000 rpm with a distance sieve size of 0.5 mm. Following milling, the whole grain flour of each sample was adequately mixed and freeze-dried. Each flour sample was then further sampled in triplicate as described by Hu et al. (2017) and separately weighed for protein extraction, followed by determining the amount of polymeric proteins and their size distributions. Hence, each durum wheat genotype is represented by nine flour samples (samples from the three sites and their triplicates).

The protein extraction was performed using a two-step procedure described by Gupta et al. (1993) with modifications described in Johansson et al. (2001, 2008). For the protein extraction, 16.5 mg of each flour sample was

TABLE 1 Description of the geographical regions of origin and the corresponding number of genotypes, as well as the ranges of altitudes and geographical positions of the sampling sites of the durum wheat populations used in this study.

Population code	The geographical region of origin	NGP	Altitudinal range (m.a.s.l)	Geographical positions (UTM)	
				Latitudinal range	Longitudinal range
AB[a]	Arsi-Bale	16	1660–3100	07°00′00″ – 08°35′00″	39°58′00″ – 40°43′00″
ANSH[a]	A/North Shewa	8	2260–3020	08°48′00″ – 09°49′00″	39°12′00″ – 39°42′00″
EH[a]	East Hararge	8	2040–2415	09°16′00″ – 09 °46′00″	41°41′00″ – 42°07′00″
ESH[a]	East Shewa	29	1675–2680	08°20′00″ – 09 °47′00″	38°49′00″ – 39°16′00″
GOJ[a]	Gojam	8	2050–2610	10°18′00″ – 11°05′00″	37°29′00″ – 38°12′00″
GOWO[a]	Gonder-Wollo	11	1790–2960	10°25′00″ – 12 °55′00″	37°29′00″ – 39°45′00″
ONSH[a]	O/North Shewa	20	2400–2933	09°03′00″ – 09°58′00″	38°04′00″ – 39°20′00″
SAJ[a]	South and Jimma	5	2000–2650	06°34′00″ – 09 °50′00″	36°28′00″ – 38°23′00″
Tig[a]	Tigray	5	1960–2687	13°30′00″ – 14°10′00″	38°29′00″ – 39°33′00″
WSH[a]	West Shewa	6	1772–2567	08°51′00″ – 09°04′00″	37°51′00″ – 38°52′00″
MVar[b]	Modern Cultivars	34			

Abbreviations: m.a.s.l, meter above sea level; NGP, number of genotypes in the population, UTM, Universal Transverse Mercator coordinate system.

[a]Landraces.

[b]Cultivars.

added to a 1.0 mL solution of 0.5% (w/v) SDS and phosphate buffer (pH 6.9) in a 1.5 mL Eppendorf tube and mixed by vortexing for 10 s. Then, the samples were shaken for 5 min at 2000 rpm using an ICA VIBRAX VXR basic shaker and centrifuged for 30 min at 10,000 rpm using Thermo Fisher Scientific SORVALL LEGEND MICRO 17 Centrifuge to obtain the protein-containing supernatant. The supernatant containing the SDS-extractable proteins was then transferred to 1.5 mL vials to run SE-HPLC. To extract SDS-unextractable proteins, the pellet remaining in the Eppendorf tube was resuspended in 1.0 mL of 0.5% (w/v) SDS and phosphate buffer (pH 6.9) and thereafter sonicated using a MSE Soniprep 150 ultrasonic disintegrator at an amplitude of 5 for 45 s. The samples were then centrifuged for 30 min at 10000 rpm to collect the supernatant, which was then transferred to 1.5 mL vials to run SE-HPLC.

2.3.2 | SE-HPLC separation and integration

The amount and size distribution of the monomeric and polymeric proteins were determined using SE-HPLC, according to Johansson et al. (2005). The SE-HPLC was carried out on a Waters HPLC system (Milford, NH, USA) with a Phenomenex BIOSEP SEC-4000 column. The eluent consisted of 50% (v/v) of each acetonitrile and ultrapure water comprising 0.1% (v/v) trifluoroacetic acid (TFA), and a flow rate of 0.2 mL/min for 30 min was used (Johansson et al., 2008). The proteins were quantified by measuring UV absorbance at 210 nm. After separation, the chromatograms were divided into four different parts based on the time scale (Figure 1), and the area under the chromatogram was used to calculate large polymeric proteins (LPP), small polymeric proteins (SPP), large monomeric proteins (LMP), and small monomeric proteins (SMP) (Kuktaite et al., 2004). Following previously used approaches (Gupta et al., 1993; Johansson et al., 2005), TOTE, TOTU, %UPP, total protein content (TPC) and %LargeUPP were calculated as follows:

a. Total SDS-extractable proteins (TOTE), which is the total area under the chromatogram of SDS-soluble polymeric and monomeric proteins:

$$\text{TOTE} = \text{eLPP} + \text{eSPP} + \text{eLMP} + \text{eSMP}$$

b. Total SDS-unextractable proteins (TOTU), which is the total area under the chromatogram of SDS-unextractable polymeric and monomeric proteins:

$$\text{TOTU} = \text{uLPP} + \text{uSPP} + \text{uLMP} + \text{uSMP}$$

c. The percentage of large unextractable polymeric proteins in the total large polymeric proteins (%LargeUPP), which is the SDS-insoluble large polymeric proteins in the total large polymeric proteins:

$$\%\,\text{LargeUPP} = \frac{\text{uLPP}}{(\text{uLPP} + \text{eLPP})}\, x\, 100$$

d. The percentage of total unextractable polymeric protein (%UPP) in total polymeric proteins is the proportion of total SDS-unextractable polymeric proteins in the total polymeric protein extracts of the two steps of extraction:

$$\%\,\text{UPP} = \frac{\text{uLPP} + \text{uSPP}}{\text{eLPP} + \text{eSPP} + \text{uLPP} + \text{uSPP}}\, x\, 100$$

e. The total protein content (TPC = TOTU+TOTE) was determined by adding the total SDS-extractable proteins (TOTE) and total SDS-unextractable proteins (TOTU), and TPC here refers to the total protein concentration as determined by HPLC rather than NIR.

2.4 | Molecular characterization

2.4.1 | Genotyping and quality control

To genotype the Ethiopian germplasm, young leaf tissue was sampled and shipped to TraitGenetics (Gmbh) for DNA extraction. All samples were analyzed with a new high-density Illumina Infinium 25K wheat SNP array developed at Trait Genetics (Gmbh). For the quality control, strict filtering was applied by removing the SNP markers with a missing value above 5% and minor allele frequency below 5% using TASSEL v 5.2.67 software (Bradbury et al., 2007). Similarly, samples with low quality were checked across all the markers, no samples were found to have a missing rate per sample greater than 1%, and all genotyped samples were used in the analysis. After stringent quality control, 8139 SNPs were used for further analysis.

2.5 | Statistical analysis

Analysis of variance (ANOVA), principal component analysis (PCA), and Spearman rank correlation (r) analysis were carried out using the R packages lme4 with lmer () function (Bates et al., 2016), factoextra with fviz_pca_biplot () function (Kassambara & Mundt, 2020), and psych with corPlot () function (Revelle, 2019), respectively, to estimate the levels of similarities and differences between the genotypes. Before conducting PCA, the data for each protein component were standardized to a variance of the unit and zero means. The ANOVA results were used to

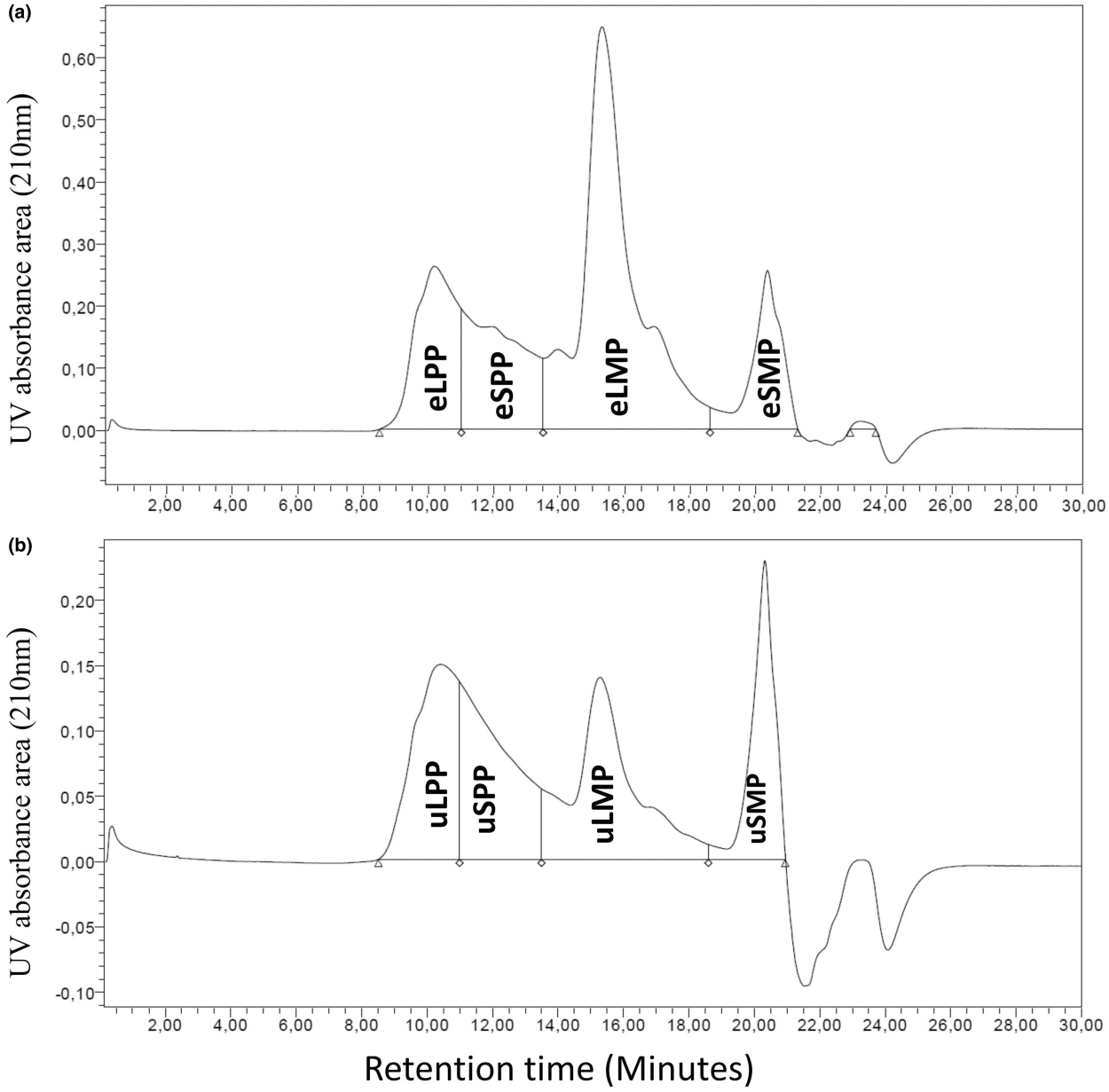

FIGURE 1 SE-HPLC chromatograms of genotype G011 displaying (a) SDS-soluble proteins divided into four parts comprising extractable large and small polymeric proteins (eLPP and eSPP), and extractable large and small monomeric proteins (eLMP and eSMP), and (b) SDS-insoluble proteins consisting of insoluble large and small polymeric proteins (uLPP and uSPP), and insoluble large and small monomeric proteins (uLMP and uSMP).

estimate heritability in a broad sense (H^2) as (Gonçalves-Vidigal et al., 2008):

$$H^2 = \frac{\sigma_g^2}{\left[\sigma_g^2 + \left(\sigma_{gl}^2/l\right) + \left(\sigma_e^2/lr\right)\right]}$$

where σ_g^2 is genotypic variance, σ_{gl}^2 is genotype-by-environment interaction variance, σ_e^2 is environmental variance, l is the number of environments, and r is the number of replications.

Additive main effect and multiplicative interaction (AMMI) stability value (ASV) and AMMI rank stability value (rASV) were computed as described in Mulugeta et al. (2022) to evaluate the stability of the genotypes across the test sites in terms of their protein content and composition.

Based on the genotypic data, cluster analysis (unweighted pair group method with arithmetic mean (UPGMA)) and PCA were carried out to identify genetic relationships between the genotypes. Pairwise Nei's standard genetic distance between the genotypes

(Nei, 1972) was computed and used for cluster analysis using Power Marker v.3.25 (Liu & Muse, 2005). The generated tree was viewed with MEGA version x (Kumar et al., 2018).

3 | RESULTS

3.1 | Genotype performance and variation

The combined analysis of variance revealed highly significant ($p < 0.001$) variation among the 149 genotypes for all protein factors evaluated (Table 2). Generally, the cultivars had a significantly higher TOTU than the landraces, while the landraces had a higher TOTE than the cultivars (Table S1). TOTE, which is known to correlate with grain protein concentration, varied two-fold among the genotypes, with the lowest level recorded for G125, G126, G122, G137, G119, G144, G130, and G145, and the highest level for G081, G032, G010, G080, G039, G060, G101, and G88. More than half of the genotypes (55%) had a higher amount of TOTE, on average, than the overall mean TOTE amount across genotypes and test environments. TOTU varied ten-fold with the lowest level recorded for G094, G010, G032, G102, G025, G098, G060, G001, G024, and G005, and the highest level for G147, G125, G122, G146, G137, G141, G128, G142, G126, G107, and G131. On average, 42% of the genotypes had a higher amount of TOTU than the overall mean TOTU amount across all genotypes and test environments. For %LargeUPP, 36% of the genotypes had a higher percentage, on average, than the overall mean of 21.7% across all genotypes and test environments. A 12-fold variation was found for %UPP, with the lowest level recorded for genotypes G085, G010, G032, G094, G060, G025, G098, G001, and G107 and the highest level recorded for G125, G147, G057, G122, G137, G114, G146, G142, G130, and G144 (Table S1). The estimated broad-sense heritability was generally high (0.87–0.99) for all evaluated protein fractions except for uSMP ($H^2 = 0.32$; Table 2).

3.2 | Effect of cultivation environmentsandgenotype-by-environment interaction on protein quality

The combined ANOVA revealed a highly significant ($p < 0.001$) impact of cultivation environments and genotype-by-environment interaction (GEI) on all protein

TABLE 2 Mean squares of the combined analysis of variance and broad-sense heritability for different protein characteristics of the durum wheat genotypes determined through SE-HPLC analysis.

	Source of variation				
Protein characteristics	**G (DF = 148)**	**L (DF = 2)**	**G × L (DF = 296)**	**Pooled error (DF = 894)**	H^2
SDS-extractable proteins					
eLPP (10^{12})	85.0	130.0	28.5	1.10	0.90
eLPP (10^{12})	84.0	340.0	15.0	1.10	0.94
eLMP (10^{13})	54.0	283.0	16.5	1.12	0.91
eSMP (10^{12})	9.4	115.0	6.5	0.50	0.98
SDS-unextractable proteins					
uLPP (10^{12})	80.0	172.0	10.3	0.45	0.96
uSPP (10^{12})	64.7	579.0	65.0	0.50	0.75
uLMP (10^{13})	49.0	4150.0	20.9	1.80	0.87
uSMP (10^{11})	41.0	998.0	25.7	0.63	0.32
TOTE (10^{13})	134.0	4307.0	38.4	2.85	0.91
TOTU (10^{13})	53.3	1059.0	8.9	0.70	0.99
%LargeUPP	673.5	720.3	62.2	3.80	0.97
%UPP	624.0	1281.9	48.5	3.45	0.97

Note: All main effects were highly significant ($p < 0.001$) for measured protein components.

Abbreviations: %LargeUPP, percentage of large unextractable polymeric protein in total large polymeric protein; %UPP, percentage of total unextractable polymeric protein in total polymeric proteins; DF, degrees of freedom; eLMP and eSMP, extractable large and small monomeric proteins, respectively; eLPP and eSPP, extractable large and small polymeric proteins, respectively; G × L, genotype × environment interaction; G, genotype; H^2, broad-sense heritability; L, location; TOTE, total SDS-extractable protein; TOTU, total SDS-unextractable protein; uLMP and uSMP, unextractable large and small monomeric protein, respectively; uLPP and uSPP, unextractable large and small polymeric protein, respectively.

factors analyzed (Table 2). Consequently, the performance of the genotypes differed across test sites, with varying protein concentration (TOTE) and gluten strength (%UPP) in the genotypes at specific test sites.

3.3 | Genotype stability for SE-HPLC-based protein factors

The evaluated genotypes differed largely in AMMI stability value (ASV) and rank ASV (rASV) across the test environments for SE-HPLC-based protein factors. For TOTE, genotypes G137, G139, G132, G127, G081, G010, G036, G019, G077, G108, and G078 were stable across environments. In contrast to this, the genotypes G025, G022, G093, G095, G006, G011, G107, G104, G040, G026, G052, and G143 had low stability across environments for this trait (Table 3). For TOTU, %Large UPP, and %UPP, genotypes G137, G146, G122, G121, G128, G107, G125, G147, G138, G119, G126, G142, G057, and G141 were stable across environments, while genotypes G045, G022, G053, G019, G123, G122, G119, G118, G062, G026, G141, and G143 showed low stability across the environments for these traits (Table 3). Based on the overall average performance, for TOTE, G081 was the winner genotype at Chefe donsa and Kulumsa, while genotype G032 won at Sinana. For TOTU, %Large UPP, and %UPP, G147, G125, G045, and G057 were the winner genotypes at all three sites.

3.4 | Correlations between protein parameters and phenotypic traits

Spearman rank correlation analyses (Figure 2) revealed a significant ($p<0.001$) negative correlation for TOTE with %UPP ($r=-0.70$), TOTU ($r=-0.53$), thousand kernel weight ($r=-0.25$), and grain yield ($r=-0.21$). Furthermore, TOTE correlated significantly ($p<0.001$) and positively with TPC ($r=0.78$), days to heading ($r=0.48$), days to maturity ($r=0.23$), and plant height ($r=0.39$, Figure 2). TOTU revealed positive significant associations with %UPP ($r=0.94$) and thousand kernel weight ($r=0.27$). In contrast, TOTU correlated significantly ($p<0.01$) and negatively with days to heading ($r=-0.60$), days to maturity ($r=-0.24$), and plant height ($r=-0.40$). %UPP showed significant positive correlations with thousand kernel weight ($r=0.29$) (Figure 2) and highly significant negative correlations with days to heading ($r=-0.62$), days to maturity ($r=-0.27$), and plant height ($r=-0.43$, respectively). Moreover, a highly significant positive correlation was obtained between TPC and TOTE ($r=0.78$).

3.5 | Variation and relatedness among genotypes

A PCA using SE-HPLC-based protein factors revealed that the main part of the variation was clearly explained along the first principal component (PC1; 64.61%), while the second principal component (PC2) only contributed 34.61% of the total variation (Figure 3a). The variation explained within PC1 was mainly contributed by TPC and TOTE protein parameters. However, TOTE, TOTU, %UPP, and TPC were traits to contribute more to PC2 (Figure 3b). The PCA based on protein factors grouped the genotypes into four main clusters (Figure 3a) by differentiating the genotypes based on their gluten strength. Cluster 2 consisted of genotypes (all cultivars plus three landraces (G045, G057, and G108)) with high %UPP and %LargeUPP (Figure 3a), while clusters 1, 3, and 4 contained genotypes particularly landraces having the lowest gluten strength. The highest TOTE was found in genotypes with clusters 1 and cluster 4 (mainly landraces).

A PCA based on genotypic (SNP) data produced comparable results to the PCA based on SE-HPLC-based protein factors, with PC1 describing the majority of the variation (67%) and PC2 contributing 11.5%. In addition, the genotypes were again clearly distinguishable as four clusters, with the majority of the cultivars in cluster 4 (Figure 4). In both the PCA analyses, the landraces did not follow their geographical origin and were scattered across clusters 1, 3, and 4 (Figure 3a; Figure 4).

A standard Nei's genetic distance-based UPGMA cluster analysis (using SNP-based genotypic data) following the average linkage algorithm resulted in the grouping of 149 genotypes into five major clusters. The modern cultivars were all grouped in cluster II (CI – II), while the genotypes from the same geographical origin-based populations (inner labels keys) were inconsistently distributed across the UPGMA clusters (Figure 5). All genotypes with the highest gluten strength (PCA cluster 2, outer labels' key), using PCA grouping based on HPLC protein factors, were found in the UPGMA cluster II, except forG057 and G045, which belong to Cl-V and are a landrace from Tigray and North Shewa, respectively (Figure 5). In the Cl-II, mainly consisting of modern cultivars, only G069, with the weakest gluten strength performance (cluster 2 outer labels' key), was found (Figure 5).

Except for G057, G118, G134, and G149, no genotypes with strong gluten performance were found in the UPGMA clusters beside Cl-II, although genotypes with relatively high gluten strength (cluster 1 outer labels' key) were found in Cl-IV and Cl-V (Figure 5). No consistency was found among the geographical origin-based populations and these genotypes with relatively high gluten strength.

TABLE 3 The 10% top and bottom genotypes for the various protein factors and their corresponding stability values.

TOTE				TOTU				%large UPP				%UPP			
Mean (x10^8)	ASV	rASV	G	Mean (x10^7)	ASV	rASV	G	Mean	ASV	rASV	G	Mean	ASV	rASV	G
Top 10% highly stable genotypes (top-down)															
1.10	0.29	1	G137	2.85	0.06	1	G079	12.82	0.10	1	G090	16.62	0.05	1	G038
1.16	0.31	2	G139	4.01	0.08	2	G070	16.33	0.12	2	G038	18.74	0.07	2	G068
1.34	0.33	3	G105	3.17	0.08	3	G068	20.50	0.12	3	G087	23.77	0.09	3	G047
1.33	0.34	4	G019	3.45	0.11	4	G056	22.60	0.15	4	G051	12.84	0.13	4	G001
1.48	0.34	5	G078	4.13	0.11	5	G105	16.66	0.15	5	G036	20.62	0.14	5	G056
1.29	0.43	6	G127	3.73	0.15	6	G047	16.17	0.16	6	G148	20.76	0.17	6	G087
1.36	0.50	7	G036	3.42	0.19	7	G004	11.70	0.17	7	G001	12.11	0.19	7	G098
1.02	0.50	8	G125	3.02	0.22	8	G075	22.51	0.18	8	G047	23.24	0.23	8	G055
1.24	0.51	9	G132	3.52	0.24	9	G087	21.21	0.20	9	G091	24.54	0.23	9	G051
1.35	0.52	10	G077	4.26	0.24	10	G072	11.54	0.22	10	G098	17.67	0.26	10	G148
1.29	0.56	11	G047	3.76	0.24	11	G055	17.54	0.24	11	G068	9.10	0.29	11	G094
1.59	0.56	12	G010	4.71	0.24	12	G138	14.96	0.26	12	G008	14.17	0.32	12	G067
1.77	0.59	13	G081	3.88	0.25	13	G108	45.26	0.27	13	G057	18.59	0.33	13	G029
1.26	0.60	14	G108	2.89	0.26	14	G067	14.79	0.29	14	G099	8.97	0.34	14	G085
1.47	0.60	15	G056	3.30	0.27	15	G015	7.55	0.30	15	G032	22.45	0.37	15	G101
Bottom 10% unstable genotypes (bottom-up)															
1.43	2.33	135	G073	4.48	2.09	135	G117	35.47	1.99	135	G128	17.47	2.26	135	G062
1.27	2.38	136	G034	4.36	2.16	136	G120	39.27	2.06	136	G142	20.81	2.32	136	G143
1.46	2.42	137	G107	3.99	2.19	137	G039	39.54	2.07	137	G146	30.04	2.36	137	G140
1.44	2.53	138	G052	4.92	2.27	138	G142	12.94	2.16	138	G002	21.67	2.38	138	G053
1.46	2.61	139	G086	5.34	2.29	139	G146	16.36	2.18	139	G062	23.43	2.42	139	G019
1.29	2.63	140	G011	3.61	2.46	140	G052	26.83	2.19	140	G135	30.07	2.46	140	G135
1.24	2.79	141	G026	3.69	2.55	141	G019	26.52	2.25	141	G136	22.24	2.49	141	G116
1.41	2.81	142	G143	4.68	2.61	142	G119	42.21	2.40	142	G119	37.07	2.49	142	G144
1.36	2.84	143	G040	5.69	2.62	143	G125	30.90	2.41	143	G123	15.83	2.58	143	G115
1.38	2.93	144	G104	3.83	2.85	144	G053	38.79	2.42	144	G141	41.03	2.62	144	G119
1.23	2.99	145	G006	4.84	3.01	145	G130	37.98	2.73	145	G130	38.31	2.70	145	G141
1.44	3.08	146	G095	4.63	3.01	146	G140	37.12	2.80	146	G144	33.36	3.12	146	G121
1.35	3.13	147	G093	4.79	3.12	147	G045	42.94	3.06	147	G122	38.42	3.17	147	G130
1.45	3.42	148	G022	5.16	4.06	148	G141	31.15	3.27	148	G121	43.31	3.27	148	G122
1.49	3.54	149	G025	3.79	4.09	149	G118	26.27	3.82	149	G118	26.20	4.82	149	G118

Abbreviations: %Large UPP, percentage of large unextractable polymeric protein in total large polymeric protein; %UPP, percentage of total unextractable polymeric protein in total polymeric protein; ASV, AMMI stability value; G, genotype number; rASV, rank ASV; TOTE, total SDS-extractable protein; TOTU, total SDS-unextractable protein.

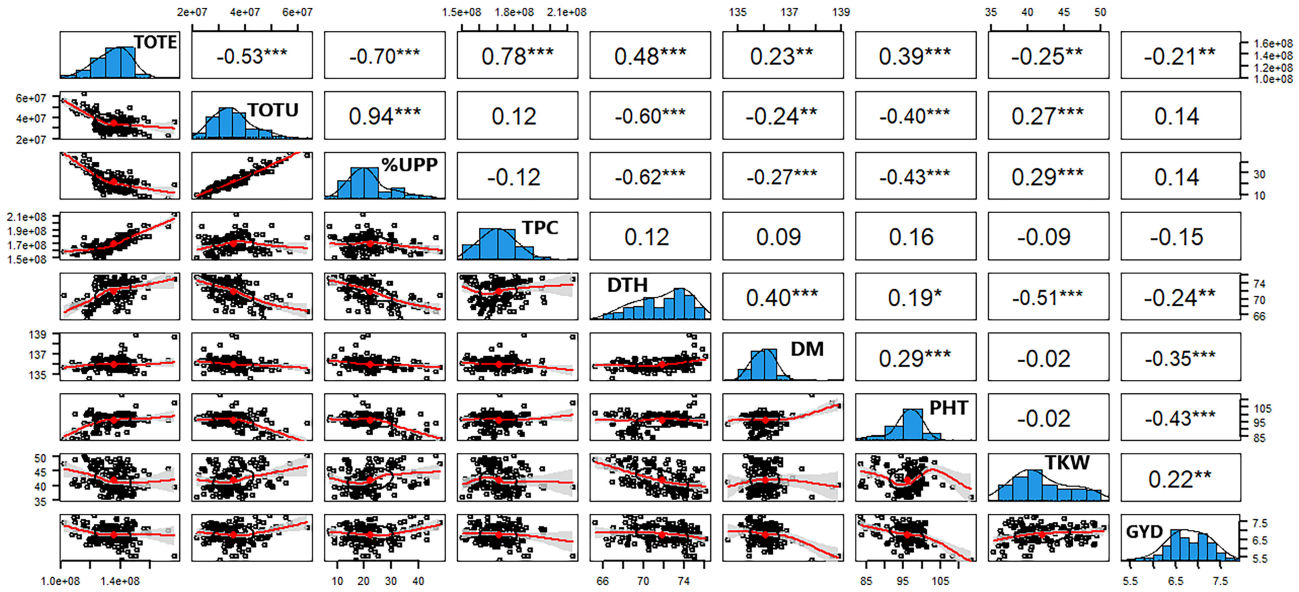

FIGURE 2 Spearman rank correlation for protein fraction parameters and some phenotypic traits. %UPP, percentage of total unextractable polymeric protein; DM, days to maturity; DTH, days to heading; GYD, grain yield; PHT, plant height; TKW, thousand kernel weight; TOTE, total SDS-extractable proteins; TOTU, total SDS-unextractable proteins and TPC, total protein content.

4 | DISCUSSION

The present study, evaluating the genetic diversity using genotyping by SNP array and variation in amount and size distribution of the wheat grain polymeric protein, clearly pinpointed the considerable variation in Ethiopian durum wheat genotypes. Interestingly, all the evaluated cultivars were more closely genetically related than they were with the landraces, the latter not closely related to each other based on geographical origin. Furthermore, despite the fact that all cultivars did not belong to the group with the highest gluten strength, the cultivars were generally found with higher gluten strength than the landraces. Gluten strength is highly relevant for the pasta and baking industries (Johansson et al., 2013). The present study thereby indicates that most of the landraces used in this study have low gluten strength compared to cultivars to be of interest to be incorporated in a breeding program designed for high-quality industrial production. However, some potential candidate landraces were identified as of interest for breeding high-quality wheat for industrial applications, specifically, the genotypes G057, G108, G045, and some genotypes with medium gluten strength in clusters IV and V from the UPGMA analyses, where especially G107 with high stability in gluten strength across environments is of interest.

The present study evaluated variation in grain protein composition in a large set of Ethiopian wheat genotypes using HPLC, and similarly to other studies, genotypic (Francki et al., 2009; Hailu et al., 2016; Husenov et al., 2021; Johansson et al., 2003; Tsilo et al., 2013), environmental (Hailu et al., 2016; Johansson, 2002; Kuktaite et al., 2004; Labuschagne et al., 2004; Lindeque et al., 2018; Ohm et al., 2017; Tsilo et al., 2013), and genotype-by-environment interaction (Ohm et al., 2017; Tsilo et al., 2010, 2013) effects were evident. However, the range of variation was exceptional in the present study for the protein factors evaluated: TOTE (7.9×10^7 (G040) to 2.08×10^8 (G032)), TOTU (1.6×10^7 (G010) to 7.8×10^8 (G147)), %LargeUPP (4.9% (G085) to 61.2% (G125)), and %UPP (4.3% (G010) to 56.9% (G125)). The highest values of TOTU, %LargeUPP, and %UPP were found in the Ethiopian cultivars, while the highest values for TOTE were found in Ethiopian landraces. Previous studies have shown that TOTE positively correlates to wheat grain protein concentration (Johansson et al., 2004, 2013; Labuschagne & Aucamp, 2004; Malik et al., 2011). Similarly, %UPP has been positively correlated to gluten strength (Husenov et al., 2021; Johansson et al., 2020; Malik et al., 2011, 2013). Previous studies on less wide collections of Ethiopian wheat than were used in the present study have also reported a high variation in the amount and size distribution of the proteins in Ethiopian genotypes (Hailu et al., 2016; Labuschagne et al., 2004).

Unlike previous studies (Malik et al., 2013), where days to heading and maturity have been positively correlated to %UPP and negatively correlated to TOTE in Swedish bread wheat, the opposite relationships prevailed in the Ethiopian durum wheat. A delayed crop maturation time of wheat grown in Sweden is generally connected to a more extended period of biomass accumulation in the plants, which is then transported to the grains and accumulated as starch (Johansson et al., 2013). Increases in

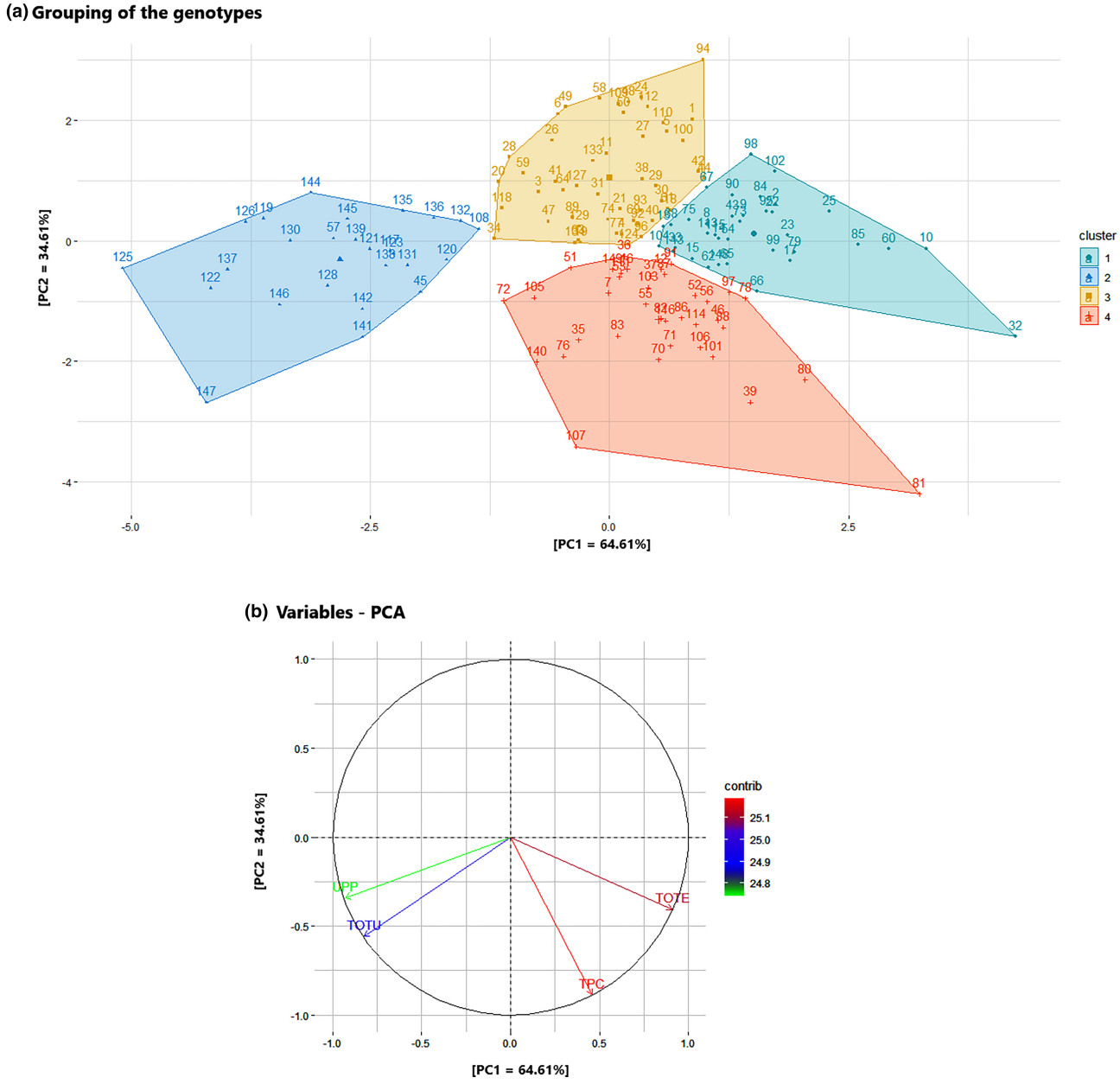

FIGURE 3 (a) PCA biplot showing the grouping patterns of genotypes based on SE-HPLC-based protein factors, and (b) contribution of protein factors traits for principal components 1 (PC1) and 2 (PC2). Genotypes are coded as 1 to 149 by removing the preceding G (e.g., G001 is coded as 1 and G149 as 149).

starch accumulation to the grain result in decreased grain protein concentration (TOTE) and increases in %UPP as a secondary result (Johansson et al., 2013). Two major environmental differences are evident for the production of wheat in Ethiopia as related to Sweden: (i) The cultivation conditions are drier in Ethiopia than in Sweden, where the conditions typically are humid, and (ii) the wheat might be subjected to heat stress conditions during grain filling in Ethiopia which seldom happens in Sweden. Thus, the dry conditions may affect biomass accumulation, so starch accumulation is not higher in the grain despite delayed crop maturation. Furthermore, heat stress during grain maturation may decrease %UPP due to the formation of heat shock proteins, as reported in, for example, Australia (Blumenthal et al., 1998). Thus, the two mentioned differences in growing conditions may be the causes of the lack of similarity between the effects of maturity data with grain protein content and composition.

In this study, a high broad-sense heritability ($H^2 > 90\%$) was obtained for all the protein traits evaluated (except for uSMP), which indicates that these traits are highly heritable, which has also been reported by others (Tsilo et al., 2013). Consequently, opportunities are available to transfer desirable genotypes and develop cultivars with

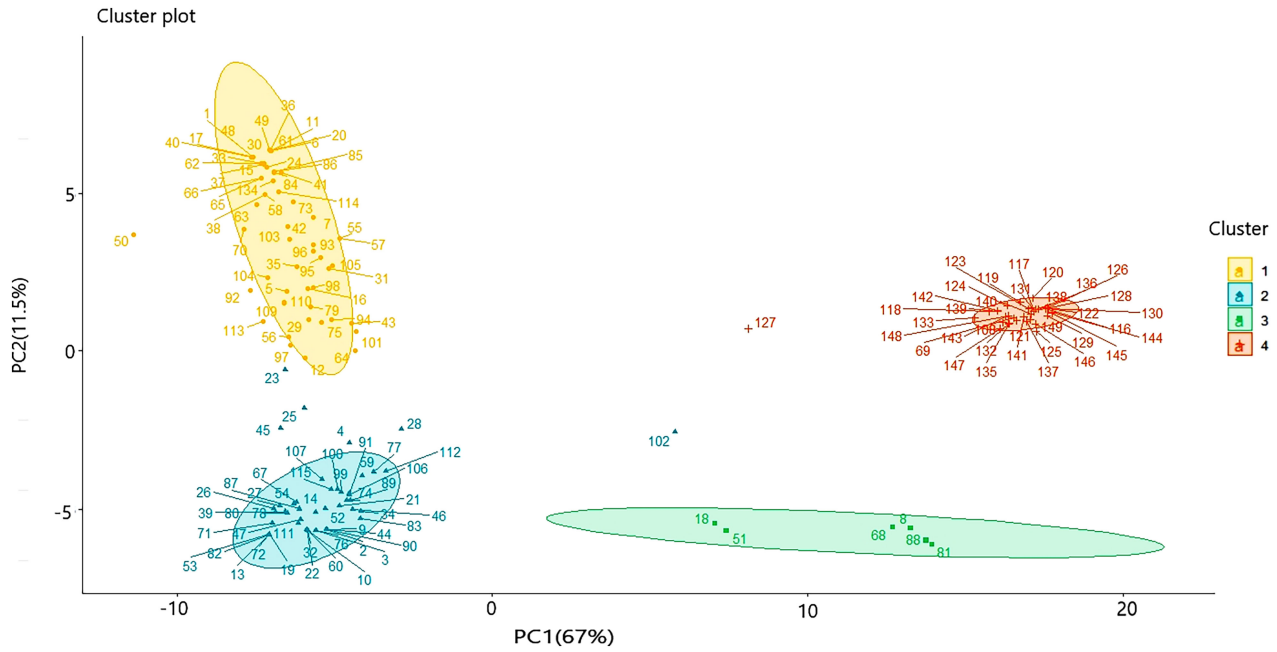

FIGURE 4 PCoA based on 8139 SNPs data illustrating the grouping pattern of the genotypes along the first two PCs. Genotypes are coded as 1 to 149 by removing the preceding G in this figure (e.g., G001 is coded as 1, and G149 is coded as 149).

suitable grain protein concentration and composition. Despite the high level of heritability, differences in stability across environments were reported in this study for the amount and size distribution of the protein. However, the genotypes found most stable across environments for factors indicating high gluten strength, that is, TOTU, %LargeUPP, and %UPP (G137, G146, G122, G121, G128, G107, G125, G147, G138, G119, G126, G142, G057, and G141), were all cultivars that belonged to Cl-II, except for the landraces G057 (Cl-V) and G107 (Cl-IV). Furthermore, these stable cultivars across environments also all belonged to cluster 4 (genotypes with the highest gluten strength) except for G121 (cluster 1), G107 (cluster 1), and G138 (cluster 3). The two landraces with high (G057) and relatively high (G107) gluten strength, which was stable across environments, are of particular interest to be used in breeding programs to develop high-quality Ethiopian durum wheat cultivars.

The present study used both genetic information and protein data to group the genotypes evaluated in the present study. Grouping based on both pieces of information clearly clustered most cultivars into one cluster. This is interesting from two perspectives: (i) The importance of a high gluten strength in durum wheat cultivars is obvious and must have been a selection criterion (although it may not be clearly understood), as most of the cultivars were grouped in the same cluster based on this criterion, and (ii) limited genetic variation was found among the cultivars as they all clustered together based on their genetic information. Also, previous studies have shown cultivars being clustered together and differently from landraces based on genetic information (Asmamaw et al., 2019; Balfourier et al., 2018; Baloch et al., 2017; Kabbaj et al., 2017; Mahboubi et al., 2020; Negisho et al., 2021). Thus, there is a clear need to incorporate novel lines into Ethiopian durum wheat breeding programs to broaden the genetic base, which is essential to secure yield through breeding for resistance/tolerance to abiotic and biotic stresses. Also, previous studies have reported that plant breeding of cultivars contributes to a decrease in genetic variation (Louwaars, 2018; Rauf & Silva, 2010; Sansaloni et al., 2020). In order to broaden the genetic base of Ethiopian durum wheat cultivars, landraces are extremely important. This study highlights the opportunities to use landraces with high and stable gluten strength, for example, G057 and G107, in the Ethiopian durum wheat breeding programs for high quality.

The lack of grouping of the landraces based on their geographical regions of origin may result from events of demographical factors such as gene flow through germplasm accompanying human movements during new settlements, trade routes, and through a step-by-step seed exchange among farmers. The separate grouping of modern cultivars indicates the involvement of exotic germplasm in the national durum wheat breeding programs to develop improved cultivars. Consequently, the allelic diversity available within local landraces has probably been underutilized for developing modern durum wheat cultivars. The present study searched opportunities to use landraces in durum wheat breeding for quality. However,

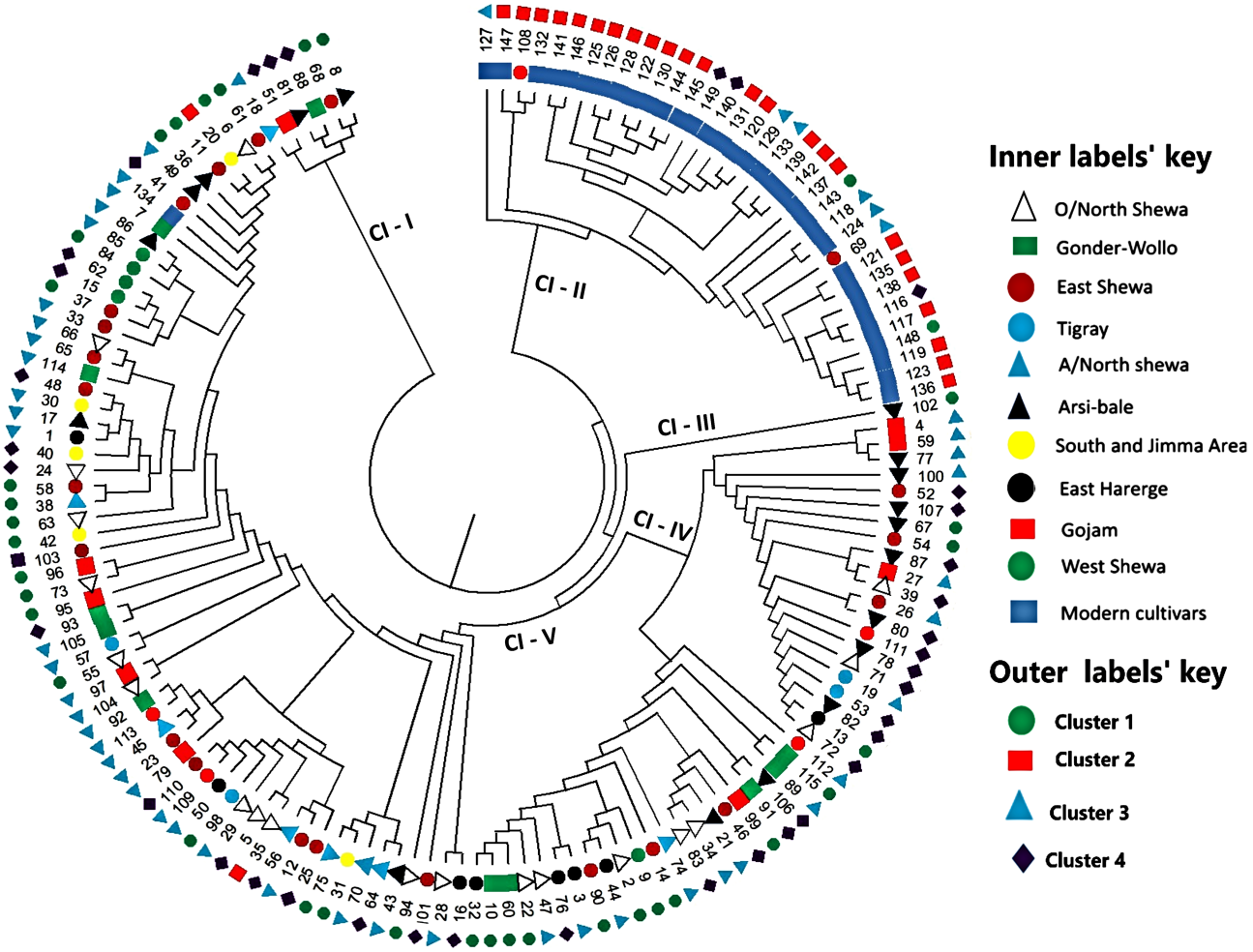

FIGURE 5 Unweighted pair group method with arithmetic Mean (UPGMA) dendrogram constructed based on Nei's standard genetic distance depicting the relationship between the 149 genotypes as well as the distribution of the genotypes across the five clusters (Cl-I to Cl-V) in terms of the eleven geographical origin-based populations and the four protein profile-based PCA clusters. For easy space usage, genotypes are coded as 1 to 149 by removing the preceding G (e.g., G001 is coded as 1 and G149 as 149).

genes for other traits, such as disease resistance and abiotic stress tolerance, might also be available in the durum wheat landrace lines used in the present study, which need to be further evaluated. With novel genomic methods evolving, the transfer of genetic material from landraces or alien introgression lines has become more precise, reliable, and faster (Johansson et al., 2021).

5 | CONCLUSIONS

Protein quality is one of the most critical factors in determining the premium quality of durum wheat cultivars. Ethiopian durum wheat landraces vary widely in grain protein content and composition. However, only a few landraces show protein quality, primarily concerning gluten strength, comparable to released cultivars in Ethiopia. Among the 116 landraces evaluated here, only four (G057, G045, G108, and G107) were found to have high and stable gluten strength across environments. The four identified landraces were found genetically distant from the Ethiopian-released cultivars, although with exciting protein quality attributes. Therefore, these four landraces are of specific importance to be included in future breeding programs in Ethiopia. The robust sole grouping of released cultivars in Ethiopia indicates a need to broaden the genetic base of cultivars to secure significant resistance and tolerance genes in the material to protect against future biotic and abiotic stresses.

AUTHOR CONTRIBUTIONS

Behailu Mulugeta contributed to conceptualization, data curation, formal analysis, investigation, methodology, software, validation, visualization, roles/writing original draft, and writing—review and editing. **Kassahun Tesfaye** contributed to conceptualization, funding acquisition, investigation, methodology, project administration, resources, supervision, and writing—review and

editing. **Rodomiro Ortiz** contributed to conceptualization, funding acquisition, investigation, methodology, project administration, resources, supervision, and writing—review and editing. **Mulatu Geleta** contributed to conceptualization, funding acquisition, methodology, project administration, resources, supervision, visualization, and writing—review and editing. **Teklehaimanot Haileselassie** contributed to conceptualization, funding acquisition, methodology, project administration, supervision, and writing—review and editing. **Cecilia Hammenhag** contributed to conceptualization, methodology, resources, supervision, and writing—review and editing. **Faris Hailu** contributed to supervision and writing—review and editing. **Eva Johansson** contributed to conceptualization, methodology, project administration, resources, supervision, and writing—review and editing.

ACKNOWLEDGMENTS

The authors would like to thank the Institute of Biotechnology, Addis Ababa University (AAU), for the technical support received during the fieldwork and the Swedish University of Agricultural Science (SLU) for the laboratory work support. The authors are grateful to the Sinana Agricultural Research Center (SARC), Kulumsa Agricultural Research Center (KARC), and Debrezeit Agricultural Research Center (DZARC) for providing their experimental sites and equipment for the field trials. The unreserved support from the technical staff of SARC, KARC, and DZARC during the field trial is highly appreciated. We would also like to thank Ethiopia Biodiversity Institute, DZARC, and SARC for providing durum wheat germplasm.

FUNDING INFORMATION

This work was funded by the Swedish International Development Cooperation Agency (Sida) grant awarded to Addis Ababa University and the Swedish University of Agricultural Sciences for a bilateral capacity-building program in biotechnology. The funders played no role in the study's design, data collection, analysis, decision to publish, or manuscript preparation.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no competing financial interests that could have appeared to influence the work reported in this paper.

DATA AVAILABILITY STATEMENT

Data will be made available on request.

ORCID

Behailu Mulugeta https://orcid.org/0000-0002-6370-940X
Mulatu Geleta https://orcid.org/0000-0002-2293-1178

REFERENCES

Adhikari, S., Kumari, J., Jacob, S. R., Prasad, P., Gangwar, O. P., Lata, C., Thakur, R., Singh, A. K., Bansal, R., Kumar, S., Bhardwaj, S. C., & Kumar, S. (2022). Landraces-potential treasure for sustainable wheat improvement. *Genetic Resources and Crop Evolution*, *69*, 499–523. https://doi.org/10.1007/s10722-021-01310-5

Alemu, A., Feyissa, T., Letta, T., & Abeyo, B. (2020). Genetic diversity and population structure analysis based on the high density SNP markers in Ethiopian durum wheat (Triticum turgidum ssp. durum). *BMC Genetics*, *21*, 18. https://doi.org/10.1186/s12863-020-0825-x

Alvarez, J. B., & Guzmán, C. (2018). Interspecific and intergeneric hybridization as a source of variation for wheat grain quality improvement. *Theoretical and Applied Genetics*, *131*(2), 225–251. https://doi.org/10.1007/s00122-017-3042-x

Asmamaw, M., Keneni, G., & Tesfaye, K. (2019). Genetic diversity of Ethiopian durum wheat (Triticum durum Desf.) landrace collections as reveled by SSR markers. *Advances in Crop Science and Technology.*, *07*, 1–9. https://doi.org/10.4172/2329-8863.1000413

Badebo, A., Gelalcha, S., Ammar, K., Nachit, M., & Abdalla, O. (2009). Overview of durum wheat research in Ethiopia: Challenges and prospects.

Balfourier, F., Bouchet, S., Robert, S., De Oliveira, R., Rimbert, H., Kitt, J., Choulet, F., & Paux, E. (2018). Worldwide phylogeography and history of wheat genetic diversity. BioRxiv https://doi.org/10.1101/477455

Baloch, F. S., Alsaleh, A., Shahid, M. Q., Çiftçi, V., Sáenz De Miera, L. E., Aasim, M., Nadeem, M. A., Aktaş, H., Özkan, H., & Hatipoğlu, R. (2017). A whole genome DArTseq and SNP analysis for genetic diversity assessment in durum wheat from central fertile crescent. *PLoS One*, *12*, e0167821. https://doi.org/10.1371/journal.pone.0167821

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2016). Fitting linear mixed-effects models using lme4 (R package version 1.1-12). *Statistical Software*.

Blumenthal, C., Stone, P. J., Gras, P. W., Bekes, F., Clarke, B., Barlow, E. W. R., Appels, R., & Wrigley, C. W. (1998). Heat-shock protein 70 and dough-quality changes resulting from heat stress during grain filling in wheat. *Cereal Chemistry*, *75*, 43–50. https://doi.org/10.1094/CCHEM.1998.75.1.43

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics*, *23*(19), 2633–2635. https://doi.org/10.1093/bioinformatics/btm308

Ceresino, E. B., Kuktaite, R., Hedenqvist, M. S., Sato, H. H., & Johansson, E. (2020). Processing conditions and transglutaminase sources to "drive" the wheat gluten dough quality. *Innovative Food Science & Emerging Technologies*, *65*, 102439. https://doi.org/10.1016/J.IFSET.2020.102439

Dejene, K. M., & Mario, E. (2016). Revisiting the ignored Ethiopian durum wheat (Triticum turgidum var. durum) landraces for genetic diversity exploitation in future wheat breeding programs. *Journal of Plant Breeding and Crop Science.*, *8*, 45–59. https://doi.org/10.5897/jpbcs2015.0542

Francki, M. G., Walker, E., Crawford, A. C., Broughton, S., Ohm, H. W., Barclay, I., Wilson, R. E., & McLean, R. (2009). Comparison of genetic and cytogenetic maps of hexaploid wheat (Triticum aestivum L.) using SSR and DArT markers. *Molecular Genetics and Genomics*, *281*, 181–191. https://doi.org/10.1007/s00438-008-0403-9

Gonçalves-Vidigal, M. C., Mora, F., Bignotto, T. S., Munhoz, R. E. F., & De Souza, L. D. (2008). Heritability of quantitative traits in segregating common bean families using a Bayesian approach. *Euphytica*, *164*(2), 551–560. https://doi.org/10.1007/s10681-008-9758-6

Gulia, N., & Khatkar, B. S. (2015). Quantitative and qualitative assessment of wheat gluten proteins and their contribution to instant noodle quality. *International Journal of Food Properties*, *18*, 1648–1663. https://doi.org/10.1080/10942912.2013.805765

Gupta, R. B., Khan, K., & Macritchie, F. (1993). Biochemical basis of flour properties in bread wheats. i. Effects of variation in the quantity and size distribution of polymeric protein. *Journal of Cereal Science*, *18*, 23–41. https://doi.org/10.1006/jcrs.1993.1031

Hailegiorgis, D., Mekonnen, F., Hailu, F., Lee, C. A., & Yun, S. J. (2020). Composition and molecular weight distribution of albumin and globulin protein isolates from durum wheat genotypes. *American Journal of Plant Sciences*, *11*, 137–147. https://doi.org/10.4236/ajps.2020.112011

Hailu, F., Johansson, E., Merker, A., Belay, G., Singh, H., & Zeleke, H. (2006). Composition of and variation in high- and low-molecular weight glutenin subunits, and omega gliadins in Ethiopian tetraploid wheat germplasm. *Plant Genetic Resources*, *4*, 134–143. https://doi.org/10.1079/pgr2006110

Hailu, F., Labuschagne, M., Van Biljon, A., Hovmalm, H. P., & Johansson, E. (2016). Quality assessment with HPLC in released varieties of tetraploid (Triticum durum Desf.) wheat from Ethiopia and Spain. *Cereal Research Communications*, *44*(4), 617–627. https://doi.org/10.1556/0806.44.2016.040

Harlan, J. R. (1969). Ethiopia: A center of diversity. *Economic Botany*, *23*, 309–314. https://doi.org/10.1007/BF02860676

Helguera, M., Abugalieva, A., Battenfield, S., Békés, F., Branlard, G., Cuniberti, M., Hüsken, A., Johansson, E., Morris, C. F., Nurit, E., Sissons, M., & Vazquez, D. (2020). Grain quality in breeding. In *Wheat Quality for Improving Processing and Human Health*. Springer. https://doi.org/10.1007/978-3-030-34163-3_12

Hu, Y., Wang, L., & Li, Z. (2017). Modification of protein structure and dough rheological properties of wheat flour through superheated steam treatment. *Journal of Cereal Science*, *76*, 222–228. https://doi.org/10.1016/j.jcs.2017.06.013

Husenov, B., Asaad, S., Muminjanov, H., Garkava-Gustavsson, L., & Johansson, E. (2021). Sustainable wheat production and food security of domestic wheat in Tajikistan: Implications of seed health and protein quality. *International Journal of Environmental Research and Public Health*, *18*(11), 1–20. https://doi.org/10.3390/ijerph18115751

Husenov, B., Makhkamov, M., Garkava-Gustavsson, L., Muminjanov, H., & Johansson, E. (2015). Breeding for wheat quality to assure food security of a staple crop: The case study of Tajikistan. *Agriculture & Food Security*, *4*(1), 1–8. https://doi.org/10.1186/s40066-015-0029-1

IBPGR. (1985). *Descriptors Wheat* (p. 12).

Johansson, E. (2002). Effect of two wheat genotypes and Swedish environment on falling number, amylase activities, and protein concentration and composition. *Euphytica*, *126*(1), 143–149. https://doi.org/10.1023/A:1019646916905

Johansson, E., Branlard, G., Cuniberti, M., Flagella, Z., Hüsken, A., Nurit, E., Peña, R. J., Sissons, M., & Vazquez, D. (2020). Genotypic and environmental effects on wheat technological and nutritional quality. In *Wheat Quality for Improving Processing and Human Health*. Springer. https://doi.org/10.1007/978-3-030-34163-3_8

Johansson, E., Kuktaite, R., Andersson, A., & Prieto-Linde, M. L. (2005). Protein polymer build-up during wheat grain development: Influences of temperature and nitrogen timing. *Journal of the Science of Food and Agriculture*, *85*(3), 473–479. https://doi.org/10.1002/jsfa.2006

Johansson, E., Malik, A. H., Hussain, A., Rasheed, F., Newson, W. R., Plivelic, T., Hedenqvist, M. S., Gällstedt, M., & Kuktaite, R. (2013). Wheat gluten polymer structures: The impact of genotype, environment, and processing on their functionality in various applications. *Cereal Chemistry*, *90*(4), 367–376. https://doi.org/10.1094/CCHEM-08-12-0105-FI

Johansson, E., Prieto-Linde, M. L., Andersson, A., Gissén, C., Malik, A., & Kuktaite, R. (2008). Protein polymer accumulation during grain development and relations to quality: Influences of cultivar and environment. *The 11th international wheat genetics symposium proceedings*, 5–7. http://hdl.handle.net/2123/3343

Johansson, E., Prieto-Linde, M. L., & Jönsson, J. O. (2001). Effects of wheat cultivar and nitrogen application on storage protein composition and breadmaking quality. *Cereal Chemistry*, *78*(1), 19–25. https://doi.org/10.1094/CCHEM.2001.78.1.19

Johansson, E., Prieto-Linde, M. L., & Larsson, H. (2021). Locally adapted and organically grown landrace and ancient spring cereals—A unique source of minerals in the human diet. *Food*, *10*(2), 393. https://doi.org/10.3390/foods10020393

Johansson, E., Prieto-Linde, M. L., & Svensson, G. (2004). Influence of nitrogen application rate and timing on grain protein composition and gluten strength in Swedish wheat cultivars. *Journal of Plant Nutrition and Soil Science*, *167*(3), 345–350. https://doi.org/10.1002/jpln.200320332

Johansson, E., Prieto-Linde, M. L., Svensson, G., & Jönsson, J. Ö. (2003). Influences of cultivar, cultivation year and fertilizer rate on amount of protein groups and amount and size distribution of mono- and polymeric proteins in wheat. *Journal of Agricultural Science*, *140*(3), 275–284. https://doi.org/10.1017/S0021859603003162

Kabbaj, H., Sall, A. T., Al-Abdallat, A., Geleta, M., Amri, A., Filali-Maltouf, A., Belkadi, B., Ortiz, R., & Bassi, F. M. (2017). Genetic diversity within a global panel of durum wheat (Triticum durum) landraces and modern germplasm reveals the history of alleles exchange. *Frontiers in Plant Science*, *8*, 1–13. https://doi.org/10.3389/fpls.2017.01277

Kadkol, G. P., & Sissons, M. (2016). Durum wheat overview. In *Reference module in food science* (2nd ed.). Elsevier Ltd. https://doi.org/10.1016/b978-0-08-100596-5.00024-x

Kassambara, A., & Mundt, F. (2020). Factoextra: Extract and visualize the results of multivariate data analyses. Package Version 1.0.7. R Package Version.

Kidane, Y. G., Gesesse, C. A., Hailemariam, B. N., Desta, E. A., Mengistu, D. K., Fadda, C., Mario Enrico, P., & Dell'Acqua, M. (2019). A large nested association mapping population for breeding and quantitative trait locus mapping in Ethiopian

durum wheat. *Plant Biotechnology Journal*, *17*, 1380–1393. https://doi.org/10.1111/pbi.13062

Kuktaite, R., Larsson, H., & Johansson, E. (2004). Variation in protein composition of wheat flour and its relationship to dough mixing behaviour. *Journal of Cereal Science*, *40*(1), 31–39. https://doi.org/10.1016/j.jcs.2004.04.007

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, *35*(6), 1547–1549. https://doi.org/10.1093/molbev/msy096

Labuschagne, M. T., & Aucamp, U. (2004). The use of size exclusion high performance liquid chromatography (SE-HPLC) for wheat quality prediction in South Africa. *South African Journal of Plant and Soil*, *21*(1), 8–12. https://doi.org/10.1080/02571862.2004.10635015

Labuschagne, M. T., Koen, E., & Dessalegn, T. (2004). Use of size-exclusion high-performance liquid chromatography for wheat quality prediction in Ethiopia. *Cereal Chemistry*, *81*(4), 533–537. https://doi.org/10.1094/CCHEM.2004.81.4.533

Lafiandra, D., & Shewry, P. R. (2022). Wheat glutenin polymers 2. The role of wheat glutenin subunits in polymer formation and dough quality. *Journal of Cereal Science*, *106*, 1–11. https://doi.org/10.1016/j.jcs.2022.103487

Lindeque, R. C., van Biljon, A., & Labuschagne, M. T. (2018). Defining associations between grain yield and protein quantity and quality in wheat from the three primary production regions of South Africa. *Journal of Cereal Science*, *79*, 294–302. https://doi.org/10.1016/j.jcs.2017.11.012

Liu, K., & Muse, S. V. (2005). PowerMaker: An integrated analysis environment for genetic maker analysis. *Bioinformatics*, *21*(9), 2128–2129. https://doi.org/10.1093/bioinformatics/bti282

Louwaars, N. P. (2018). Plant breeding and diversity: A troubled relationship? *Euphytica*, *214*, 114. https://doi.org/10.1007/s10681-018-2192-5

Maccaferri, M., Harris, N. S., Twardziok, S. O., Pasam, R. K., Gundlach, H., Spannagl, M., Ormanbekova, D., Lux, T., Prade, V. M., Milner, S. G., Himmelbach, A., Mascher, M., Bagnaresi, P., Faccioli, P., Cozzi, P., Lauria, M., Lazzari, B., Stella, A., Manconi, A., … Cattivelli, L. (2019). Durum wheat genome highlights past domestication signatures and future improvement targets. *Nature Genetics*, *51*(5), 885–895. https://doi.org/10.1038/s41588-019-0381-3

Mahboubi, M., Mehrabi, R., Naji, A. M., & Talebi, R. (2020). Whole-genome diversity, population structure and linkage disequilibrium analysis of globally diverse wheat genotypes using genotyping-by-sequencing DArTseq platform. *3 Biotech*, *10*, 48. https://doi.org/10.1007/s13205-019-2014-z

Malik, A. H. (2012). Governing grain protein concentration and composition in wheat and barley.

Malik, A. H., Kuktaite, R., & Johansson, E. (2013). Combined effect of genetic and environmental factors on the accumulation of proteins in the wheat grain and their relationship to bread-making quality. *Journal of Cereal Science*, *57*(2), 170–174. https://doi.org/10.1016/j.jcs.2012.09.017

Malik, A. H., Prieto-Linde, M. L., Kuktaite, R., Andersson, A., & Johansson, E. (2011). Individual and interactive effects of cultivar maturation time, nitrogen regime and temperature level on accumulation of wheat grain proteins. *Journal of the Science of Food and Agriculture*, *91*(12), 2192–2200. https://doi.org/10.1002/jsfa.4439

Markgren, J., Rasheed, F., Hedenqvist, M. S., Skepö, M., & Johansson, E. (2022). Clustering and cross-linking of the wheat storage protein α-gliadin: A combined experimental and theoretical approach. *International Journal of Biological Macromolecules*, *211*, 592–615. https://doi.org/10.1016/J.IJBIOMAC.2022.05.032

Mastrangelo, A. M., & Cattivelli, L. (2021). What makes bread and durum wheat different? *Trends in Plant Science*, *26*, 677–684. https://doi.org/10.1016/j.tplants.2021.01.004

Mulugeta, B., Tesfaye, K., Geleta, M., Johansson, E., Hailesilassie, T., Hammenhag, C., Hailu, F., & Ortiz, R. (2022). Multivariate analyses of Ethiopian durum wheat revealed stable and high yielding genotypes. *PLoS One*, *17*, 1–21. https://doi.org/10.1371/journal.pone.0273008

Negisho, K., Shibru, S., Pillen, K., Ordon, F., & Wehner, G. (2021). Genetic diversity of Ethiopian durum wheat landraces. *PLoS One*, *16*, 1–15. https://doi.org/10.1371/journal.pone.0247016

Nei, M. (1972). Genetic distance between populations. *The American Naturalist*, *106*(949), 283–292.

Ohm, J. B., Manthey, F., & Elias, E. M. (2017). Variation and correlation of protein molecular weight distribution and semolina quality parameters for durum genotypes grown in North Dakota. *Cereal Chemistry*, *94*, 780–788. https://doi.org/10.1094/CCHEM-07-16-0189-R

Patterson, H. D., & Williams, E. R. (1976). A new class of resolvable incomplete block designs. *Biometrika*, *63*, 83–92. https://doi.org/10.1093/biomet/63.1.83

Rajnincová, D., Gálová, Z., Petrovičová, L., & Chňapek, M. (2018). Comparison of nutritional and technological quality of winter wheat (Triticum aestivum L.) and hybrid wheat (Triticum aestivum L. x triticum spelta L.). *Journal of Central European Agriculture*, *19*, 437–452. https://doi.org/10.5513/JCEA01/19.2.2146

Rauf, S., & Silva, J. D. (2010). Consequences of plant breeding on genetic diversity. *International journal of plant breeding and genetics*, *4*(1), 1–21.

Requena-Ramírez, M. D., Hornero-Méndez, D., Rodríguez-Suárez, C., & Atienza, S. G. (2021). Durum wheat (Triticum durum l.) landraces reveal potential for the improvement of grain carotenoid esterification in breeding programs. *Food*, *10*, 1–10. https://doi.org/10.3390/foods10040757

Revelle, W. (2019). psych: Procedures for psychological, psychometric, and personality research (R package version 1.9.4) [Computer software]. *October*.

Samaan, J., El-Khayat, G. H., Manthey, F. A., Fuller, M. P., & Brennan, C. S. (2006). Durum wheat quality: II. The relationship of kernel physicochemical composition to semolina quality and end product utilisation. *International Journal of Food Science and Technology*, *41*(SUPPL. 2), 47–55. https://doi.org/10.1111/j.1365-2621.2006.01313.x

Sansaloni, C., Franco, J., Santos, B., Percival-Alwyn, L., Singh, S., Petroli, C., Campos, J., Dreher, K., Payne, T., Marshall, D., Kilian, B., Milne, I., Raubach, S., Shaw, P., Stephen, G., Carling, J., Pierre, C. S., Burgueño, J., Crosa, J., … Pixley, K. (2020). Diversity analysis of 80,000 wheat accessions reveals consequences and opportunities of selection footprints. *Nature Communications*, *11*, 4572. https://doi.org/10.1038/s41467-020-18404-w

Savage, M., Vavilov, N. I., & Love, D. (1994). Origin and geography of cultivated plants. *Geographical Review*, *84*, 231. https://doi.org/10.2307/215338

Shewry, P. R., & Hey, S. J. (2015). The contribution of wheat to human diet and health. *Food and Energy Security*, *4*, 178–202. https://doi.org/10.1002/FES3.64

Shewry, P. R., & Lafiandra, D. (2022). Wheat glutenin polymers 1. Structure, assembly and properties. *Journal of Cereal Science*, *106*, 103486. https://doi.org/10.1016/J.JCS.2022.103486

Tesemma, T., & Bechere, E. (1998). Developing elite durum wheat landrace selections (composites) for Ethiopian peasant farm use: Raising productivity while keeping diversity alive. *Euphytica*, *163296*, 323–328.

Tsegaye, B., & Berg, T. (2007). Utilization of durum wheat landraces in east Shewa, Central Ethiopia: Are home uses an incentive for on-farm conservation? *Agriculture and Human Values*, *24*, 219–230. https://doi.org/10.1007/s10460-006-9055-8

Tsilo, T. J., Hareland, G. A., Ohm, J. B., & Anderson, J. A. (2013). Polymeric proteins and their association with grain yield in hard red spring wheat lines. *Euphytica*, *194*(2), 187–196. https://doi.org/10.1007/s10681-013-0912-4

Tsilo, T. J., Ohm, J. B., Hareland, G. A., & Anderson, J. A. (2010). Association of size-exclusion HPLC of endosperm proteins with dough mixing and breadmaking characteristics in a re-combinant inbred population of hard red spring wheat. *Cereal Chemistry*, *87*(2), 104–111. https://doi.org/10.1094/CCHEM-87-2-0104

Vida, G., Szunics, L., Veisz, O., Bedő, Z., Láng, L., Árendás, T., Bónis, P., & Rakszegi, M. (2014). Effect of genotypic, meteorological and agronomic factors on the gluten index of winter durum wheat. *Euphytica*, *197*, 61–71. https://doi.org/10.1007/s10681-013-1052-6

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Mulugeta, B., Tesfaye, K., Ortiz, R., Geleta, M., Haileselassie, T., Hammenhag, C., Hailu, F., & Johansson, E. (2024). Unlocking the genetic potential of Ethiopian durum wheat landraces with high protein quality: Sources to be used in future breeding for pasta production. *Food and Energy Security*, *13*, e511. https://doi.org/10.1002/fes3.511