Journal of Agriculture and Food Research 24 (2025) 102403

Contents lists available at ScienceDirect

Journal of Agriculture and Food Research

journal homepage: www.sciencedirect.com/journal/journal-of-agriculture-and-food-research

Diversities in seed storage protein composition in the Panel of Central European landraces and modern wheat genotypes

Gyöngyvér Gell [a,1], András Cseh [b,*,1], Christakis George Florides [c], Marianna Rakszegi [d], Zsófia Birinyi [a,e], Dalma Nagy-Réder [a,e], Ádám Horváth [b], Ádám D. Horváth [b], Ferenc Békés [f,2], Ildikó Karsai [b,2]

[a] Department of Biological Resources, Agricultural Institute, HUN-REN Centre for Agricultural Research, Martonvásár, Hungary
[b] Molecular Breeding Department, Agricultural Institute, HUN-REN Centre for Agricultural Research, Martonvásár, Hungary
[c] Murdoch University, College of Science, Health, Engineering and Education, Perth WA, Australia, Perth, WA, 6150, Australia
[d] Cereal Breeding Department, Agricultural Institute, HUN-REN Centre for Agricultural Research, Martonvásár, Hungary
[e] Doctoral School of Biology, Institute of Biology, ELTE Eötvös Loránd University, Budapest, Hungary
[f] FBFD PTY Ltd., Sydney, Australia

ARTICLE INFO

Keywords:
Wheat
Storage protein diversity
HPLC
MALDI-TOF
SNP array

ABSTRACT

This study aims to collect information about the diversity of seed storage proteins in Central European landraces and compare it to modern cultivars. The quantitative comparison of the protein composition is also monitored to demonstrate the alteration of the glutenin-to-gliadin ratio and the size distribution of the polymeric glutenins in landraces, old, and modern cultivars. In this study, 197 Central European winter wheat landraces and 67 modern winter wheat cultivars with different genetic backgrounds were examined. Employing a 20K SNP array, we explored the available genetic diversity within a collection of Central European bread wheat landraces and compared these landraces to modern wheat cultivars. For proteomic characterization of the analyzed population SE- and RP-HPLC, and MALDI-TOF analyses were carried out in order to identify new genetic sources for breeding. Comparative results were made accessible through interactive, user-friendly PCA plots, allowing efficient data exploration. Significant differences in protein content and composition were observed between landraces and modern wheat cultivars. A comprehensive analysis of the Zeleny index, grain gluten, and protein content, supported by population structure analysis using storage protein subunit markers, revealed previously untapped variation for the improvement of protein content and end-use quality in wheat breeding.

1. Introduction

Winter wheat (*Triticum aestivum* L., BBAADD, 2n = 6x = 42) has historically served as a major food source for humanity and currently accounts for 15 % of the global daily calorie intake [1]. Population genetic studies have provided substantial evidence that bread wheat landraces exhibit enhanced genetic variation, identifying them as valuable reservoirs of various alleles that hold potential for modern breeding strategies [2–4]. The conservation and exploration of a diverse genetic background within these populations can significantly enhance wheat improvement programs.

The primary aim of wheat breeding since its beginning has been to develop new varieties with higher yields to satisfy the needs of the growing population in the world. Modern agriculture (since the 1950s) has focused on crossbreeding different wheat and grass species, thus generating new genetic combinations of increased yield. The impact of breeding on grain yields of the wheat varieties released during the twentieth century has been extensively studied, and a general genetic determination has been observed [5,6]. Furthermore, a negative correlation was observed between the protein concentration and the grain yield, according to several studies [7–12].

The lower protein content of the high-yielding varieties had to be compensated by an altered protein composition to keep or even improve the functional properties of the new wheat varieties. Despite the high

* Corresponding author.
E-mail address: cseh.andras@atk.hun-ren.hu (A. Cseh).
[1] Gyöngyvér Gell and András Cseh contributed equally to this work.
[2] Ildikó Karsai and Ferenc Békés contributed equally to this work.

https://doi.org/10.1016/j.jafr.2025.102403
Received 21 February 2025; Received in revised form 5 September 2025; Accepted 27 September 2025
Available online 6 October 2025

number of investigations on the yield-protein relationship [13–16], there are only a few reports on how the alteration of the protein composition could compensate the lower amount of protein. Based on the data available, two main trends are typical for altering the protein composition to improve dough strength and extensibility: these are the increase of the glutenin-to-gliadin ratio and the selection of the lines containing superior high molecular weight (HMW) glutenin alleles.

Comparing old and modern Italian durum wheat varieties, De Santis et al. [17] reported that the higher gluten index of modern cultivars was correlated with elevated glutenin-to-gliadin ratio, and increased expression level of B-type low molecular weight (LMW)-GS. Regarding the glutenin-to-gliadin ratio and alteration of HMW-GS, Izadi-Darbandi et al. [18] reported similar results by comparing historic and modern Iranian wheat cultivars. In 60 German winter wheat cultivars first registered between 1891 and 2010, the protein and the gliadin content showed a decreasing trend, while the glutenin content increased significantly, without any change in the albumin/globulin ratio and the gluten content. Riaz et al. [12] investigated the general trend in the changes of the glutenin and the gliadin composition comparing 78 Australian historic and modern wheat varieties. Both the protein and the gliadin content decreased in their experiment, while the glutenin content and thus the glutenin–to-gliadin ratio increased. To improve dough elasticity and strength, varieties with superior HMW glutenin alleles have been selected significantly shifting up the ratio of the larger glutenin polymers.

Significant variation of the HMW-GS alleles of hexaploid wheat landraces and old varieties was reported in several studies [19–29]. A high-throughput genotyped set of Spanish wheat landraces was screened for quality-related genes [4], and new puroindoline and HMW-GS alleles were identified [30]. Furthermore, the landraces showed a wider genetic diversity and higher variation in the HMW-GS composition than modern varieties. Allelic variation at the Glu-1 and Glu-3 loci were evaluated in Xinjiang wheat landraces and historical varieties [31] with SDS-PAGE and allele-specific molecular markers. High diversity was found at the Glu-1, Glu-A3, and Glu-B3 loci and new alleles were identified at Glu-A3 and Glu-B3 loci.

All of the above-mentioned studies support the conclusion that several alleles present in landraces and old cultivars are absent in modern cultivars. One of these identified ‘archaic’ high molecular weight glutenin alleles, which contained the mutation of the 1Ax2* gene and was called $1Ax2^{*B}$ gene, was remarkably effective in improving the bread-making quality of wheat [22]. This result also supports the idea that the genetic diversity of glutenins decreased during the development of modern breeding practices. Similar conclusions were made by Branlard et al. [32], summarizing numerous reports dealing with the genetic diversity of durum wheat landraces, and historic and modern lines from different countries.

A large number of cultivars and lines (1060) bred in the twentieth century were examined in the work of Metakovsky et al. [33], which led to the development of a new and improved version of the catalog of alleles at the six Gli loci of common wheat. Less documentation is available on the genetic polymorphism of gliadin alleles in landraces. In the study of Ruiz et al. [34,35], Spanish landraces of common wheat were investigated, and the Gli-A1 was found to be the most polymorphic, but no consistent trend was found in the gliadin composition by comparing old and modern Australian cultivars [12,36].

The comparison of gliadins of primitive wheat relatives, old and modern wheat cultivars became the target of intensive research in the last 10 years, which suggests that the effect of the selection on gliadin composition is rather limited in wheat breeding [17,37–42].

The prevalence of various forms of cereal-related health disorders has increased in the last decades and the general public opinion claims without any scientific or practical evidence that modern wheat breeding is responsible for this expanding trend. However, numerous studies like those of Ribeiro et al. [43] and Malalgoda [44], clearly demonstrated that each historic and modern bread wheat and ancient relatives of wheat, such as emmer, einkorn, and spelt contain variable amounts of coeliac and allergenic epitopes. Spontaneous and conscious selection in wheat breeding did not target the nutritive or health-related attributes in the 20th century and thus the yield- and functionality-driven breeding did not result in any changes of these characteristics of genotypes. Several scientific investigations [45–48] provide supporting evidence that modern wheat breeding is not responsible for the increasing trend in the spread of the various forms of cereal-related health disorders.

The aim of this study is to collect information on the genetic diversity and the variation of the storage protein allele composition in Central European wheat landraces and to compare them to modern wheat cultivars in order to find new genetic sources for breeding. The genetic diversity resulting the variation in the quantitative traits of proteins and the related bread-making quality traits were analyzed. Varieties were ranked based on bread-making quality, as assessed by three years of NIR data, and landraces demonstrating superior potential were subsequently selected as new crossing partners for breeding programs.

2. Materials and methods

2.1. Plant material

The plant material analyzed in this study comprised of 197 Central European winter wheat landraces and 67 modern winter wheat cultivars. The landraces were originated from six, former and/or present Central European countries (Austria, Bulgaria, former Czechoslovakia, Hungary, Romania, former Yugoslavia), while the modern cultivars were originated from ten countries worldwide (Austria, Bulgaria, former Czechoslovakia, France, Germany, Hungary, Romania, Switzerland, USA, former Yugoslavia). The landraces were part of the Gene Bank collection of the Centre for Plant Diversity (NÖDIK) in Tápiószele, Hungary, collected during 1950–60. The modern wheat cultivars used in this study were registered between 1970 and 2015 and formed part of the Genebank of HUN-REN ATK. Further details regarding the characteristics of the accessions are available in Supplementary Table 1.

2.2. Field conditions

All samples were sown in the field at the HUN-REN Centre for Agricultural Research in Martonvásár, Hungary (Latitude: 47° 21′ N, Longitude: 18° 49’ E, Altitude: 150 m) over three years (2017, 2019, and 2020). A Completely Randomized Design was employed, in which six rows of each genotype were sown in 0.4 × 2 m plots, with a row spacing of 20 cm. The experiment was conducted on haplic chernozem soil characterized by average levels of nitrogen (N), phosphorus (P_2O_5), and potassium (K_2O). Nutrient supply was ensured by applying 60 kg N ha^{-1}, 60 kg P_2O_5 ha^{-1}, and 60 kg K_2O ha^{-1} in the fall, followed by an additional 60 kg N ha^{-1} at the beginning of spring. Standard plant protection protocols were applied throughout the experiment to avoid the effects of diseases and pests. Weather conditions of growing seasons are summarized in Supplementary Table 2. After harvesting in the three years (2018, 2020, and 2021), the samples were stored in a dry, cooled warehouse.

2.3. SNP genotyping

DNA was extracted from 6-week-old seedlings by using the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) by the manufacturer’s instructions. SNP genotyping was conducted by TraitGenetics GmbH (http://www.traitgenetics.com/en/) employing a 20K Illumina SNP chip. SNPs displaying a missing value rate exceeding 10 % and heterozygosity exceeding 5 % were excluded from subsequent analysis, resulting in a final set of 15,784 high-quality SNP markers.

2.4. Protein content and breadmaking quality traits

Grinding of grains was carried out with a Retsch MM 400 ball mill (Retsch GmbH, Haan, Germany) in all three years of the experiment. The protein-, gluten-, moisture-content, water absorption, and Zeleny sedimentation of the wholemeal samples were determined using the Perten Inframatic 8611 instrument [49].

2.5. Protein composition

2.5.1. Size exclusion high-performance liquid chromatography (SE-HPLC)

For the 2018 harvest sample set, SE-HPLC was used to determine the glutenin, gliadin, and albumin + globulin contents, as well as the unextractable polymeric protein percentage (UPP%) using the method of Gupta et al. [50], applying a two-step extraction procedure using a modified method of Batey et al. [51], as described by Larroque and Békés [52]. Analysis has been carried out with three technical replicate injections from two biological replicate extracts on samples harvested in 2018.

2.5.2. Reversed-phase high performance liquid chromatography (RP-HPLC)

For the 2018 harvest sample set, Gliadin proteins in the samples have been characterized by the method of Larroque et al. [53]. Each sample was sequentially injected three times for technical replication. RP-HPLC peak areas (expressed in arbitrary units, AU) under the chromatograms were used to calculate gliadin amounts. Retention times for ω1,2 gliadins were between 15 and 20, for ω5 20–30, for α-gliadins 30–40, and for γ-gliadins 40–55 min. The amounts of ω1,2-, ω5-, α- and γ-gliadin [% of the sample] were calculated from the data derived from the RP-HPLC separation of gliadin proteins, multiplied by the gliadin content of the total protein, determined by the SE-HPLC separation and multiplied by the protein content of the samples harvested in 2018.

2.5.3. Protein profiling by MALDI-TOF MS analysis

A four-step extraction procedure was used in this study, isolating salt-soluble proteins, gliadin, LMW, and HMW glutenin subunits for MALDI-TOF analysis [54]. Mass spectra were acquired in positive ion mode and averaged over 100 laser shots to enhance the signal-to-noise (S/N) ratio. Calibration was performed using bovine serum albumin (66,463 Da) as an external standard. The acquisition parameters were as follows: mass range of 60,000–110,000 Da, sampling rate of 1.00 GS/s, laser power set to 85 %, laser frequency at 80 Hz, and a detector gain of 13.3x. The methodology of this sample preparation as well as the instrument settings for MALDI-TOF analyses of the protein classes are detailed in Supplementary Table 3.

2.6. Analysis of population diversity

A hierarchical co-clustering dendrogram and heatmap were generated based on the analysis of the Zeleny index, gluten, and total protein content, as well as the clustering of protein, gliadin, and glutenin content. A clustergram was used to group wheat lines according to the similarity of the evaluated traits and the strength of their intercorrelations. All statistical analyses were conducted using R software (version 3.5). The Dendextend R package was used to analyze with Ward's method for clustering [55]. The SNP calls of exceptional quality for 264 wheat lines were merged to create a matrix, from which an unrooted dendrogram was generated using the Genome Association and Prediction Integrated Tool (GAPIT 3.0) package in R [56], employing the neighbor-joining method (NJ) [57]. Principal component analysis (PCA) was employed to condense the patterns of variation observed among the wheat accessions. The analysis was carried out with GAPIT 3.0 with default parameters. The 3D PCA plot was created in Plotly [58]. The data of wheat protein profiling has been analyzed to identify inter-cultivar variations. We used peaks with different retention times as markers and labeled '1' if the peak was presented and '0' if not. The quantities of the protein groups were not analyzed during the diversity analysis. Neighbor-Joining and PCA analyses were used to summarize patterns of variations among the wheat accession. The analysis was carried out by GAPIT 3.0 package in R with default parameters [56] and PCA plots were generated using Plotly in R [58].We examined the similarity of different wheat lines by comparing groups based on unrooted NJ trees of protein profiling. If a line belonged to a group, we marked it with '1', and if not, we marked it with '0'. The resulting data matrix was analyzed using hierarchical co-cluster analysis in the same manner as described above with the help of the Dendextend R package [59].

2.7. Statistical analysis

Descriptive statistical evaluations—including the computation of mean values, standard deviations, and coefficients of variation—were carried out using built-in functions in Microsoft Excel, based on data obtained from replicate MALDI-TOF and HPLC analyses. Prior to ANOVA, homoscedasticity was evaluated using Levene's test. In cases where Levene's test indicated significant heteroscedasticity, ANOVA with Welch's correction was applied. Statistically significant differences among sample groups were assessed by analysis of variance (ANOVA), followed by pairwise comparisons using Student's t-tests. These inferential statistical procedures were conducted using the STATISTICA software package (version 13.5.0.17; TIBCO Software Inc.). To assess the effects of genotype, environment, and their interaction (genotype × environment), a two-way multifactorial ANOVA was conducted.

3. Results

3.1. General description of MALDI-TOF and RP-HPLC-based databases

In wheat samples harvested during the 2018 season, four classes of allelic proteins—high molecular weight glutenins (HMW-GS), low molecular weight glutenins (LMW-GS), gliadins, and soluble proteins—were identified via MALDI-TOF analysis across 264 accessions. The same flour-derived gliadin extracts were also subjected to RP-HPLC analysis. Table 1 presents a brief overview of the general characteristics of these protein classes, while the detailed compositional data are provided in Supplementary Table 4.

The 264 genotypes contained 28938 polypeptide peaks in the 4 MALDI-TOF protein classes, the largest number of which was separated in the soluble protein fraction (n = 12113), followed by the gliadins, the LMW GS, and the HMW GS (8581, 7617, and 1318), respectively. Individual samples contained 4 or 5 HMW GS-, 16–45 LMW-GS-, 16–48 gliadin-, and 37–51 soluble protein-alleles. The molecular mass intervals for the four protein classes were the following: HMW-GS: 67473–88128, LMW-GS: 25023–43620, gliadins: 20040–50612, soluble proteins: 7858–39576.

The number of appearances of a given individual protein peak in the samples varied significantly. There were proteins in each of the four classes, which were present in one sample only (for example Dx4 in the HMW GS group appeared ones), while certain protein alleles appeared in a high number of samples (in HMW GS, LMW GS, gliadin, and soluble classes they appeared in 235, 221, 197 and 205 samples, respectively).

The RP-HPLC-based separation of the gliadins resulted in 12470 peaks altogether in the 264 samples from which 219 different individual peaks were distinguished, with retention times falling between 4.156 and 43.145 min interval. A large variation in the expression level of the same alleles (peaks with the same retention time) was also detected in different samples. The variation in the protein composition of the landraces and the modern cultivars was compared, and it was found that the number of protein peaks of modern cultivars significantly decreased compared to that of the landraces. The extent of the decrease was 32.4 %, 32.7 %, and 27.53 % for the LMW GS, gliadins, and soluble proteins,

Table 1
Characteristics of the different protein classes in a set of 264 wheat genotypes based on MALDI-TOF and RP-HPLC analysis.

Method	Protein class	Number of peaks differentiated	Total number of peaks in the population	Number of peaks in a sample				Polypeptide			
								Molecular mass [kDa]		Prevalence	
				mean	stDev	min	max	min	max	min	max
MALDI-TOF	HMW GS	21	1283	5	0.34	4	5	67473	88128	1 (Dx4)	235 (Bx7)
	LMW GS	191	7496	28	4.53	16	45	25023	43620	1	221
	Gliadin	251	8357	32	6.34	16	48	20040	50612	1	197
	Soluble proteins	247	11802	45	3.67	37	51	7858	39576	1	205
	Combined	710	28938	110	8.98	77	134				

Method	Protein class	Number of peaks differentiated	Total number of peaks in the population	Number of peaks in a sample				Polypeptide			
								Retention time (min)		Prevalence	
				mean	stDev	min	max	min	max	min	max
RP-HPLC	Gliadin	219	12470	48	3.88	35	56	4.156	43.145	1	247

respectively. Meanwhile, the variation in the number of polypeptides of the individual samples also decreased when landraces and modern cultivars were compared: from 16 to 35 LMW GS proteins to 24–45, from 13 to 46 gliadins to 23–48, and from 27 to 50 soluble proteins to 33–51.

The relative amounts of the individual gliadin polypeptides (% of the total gliadin) varied between 0.029 and 7.27 %, the amount of a peak with a given retention time in the different samples varied on average by 64 %.

3.2. Yield and protein content of samples of the entire population

The higher yield potential of modern cultivars was observed in a two-year field experiment, and the differences in average yield between 2020 and 2021 were smaller in the landraces (Table 2).

The average protein content of modern wheat varieties was significantly lower than that of Central European landraces. Consistent with previous studies [12–16], a declining trend in protein content was observed with more recent release years of modern cultivars (Fig. 1).

3.3. Protein composition and functional properties of samples

For the samples of the 2018 harvest the distribution of the three protein classes (glutenins, gliadins, and soluble proteins) was determined by SE-HPLC, while the distribution of the ω1,2-, ω5-, α/β-, and γ-gliadins related to the total gliadin proteins was measured by RP-HPLC (Supplementary Table 4).

In this study, NIR was employed to determine protein content, wet gluten content, and two key functional quality traits—Zeleny sedimentation index and water absorption capacity—which are closely related to grain protein levels and composition. Corresponding data are presented in Supplementary Table 5. Mean values for the entire population, as well as for landraces and modern cultivars separately, are provided in Supplementary Table 6.

The comparison of the HMW GS allelic composition of landraces and modern cultivars (Supplementary Table 7) showed, that the prevalence of Glu1Aa, Glu1Bb, Glu1Bc, Glu1Be, and Glu1Bh alleles decreased, while that of Glu1Ab and Glu1Ac increased with the progress of modern wheat breeding. Meanwhile, the Glu1Bk completely disappeared, and the Glu1Bal allele appeared only in the modern cultivars. In the case of the Glu1D loci, the 'c' allele disappeared, the prevalence of the 'a' allele decreased, while that of the 'd' allele increased. Considering the contribution of the different HMW-GS alleles to the functional properties of wheat, it can be concluded that changes in the allelic composition improved gluten strength and dough stability, resulting in higher values in modern cultivars due to selection in wheat breeding. The enhanced overall functionality was further evidenced by the Payne score, calculated as the sum of the weighting factors assigned to the Glu-1A, Glu-1B, and Glu-1D alleles in the sample, according to the method proposed by Payne [60]. The mean of the Payne score was almost 10 % higher in modern varieties than in the landraces (8.88 and 8.05, respectively).

3.4. Population structure of the wheat collection, using SNP markers

The winter wheat accessions were categorized by their historical origin (landraces vs. modern cultivars) to facilitate subsequent comparisons between genotypic data and protein markers (Supplementary Table 8). Neighbor-Joining (NJ) analyses and Principal Component Analysis (PCA) were conducted to delineate the population structure of the wheat collection, using a set of 15,784 SNP markers and 929 protein markers.

Based on the three-year average values of the Zeleny index, gluten and protein content measured by Near-InfraRed (NIR) spectroscopy, genotypes were ranked using one-way ANOVA (Supplementary Table 9). To visualize the 20 top-performing lines relative to the percentage-based standard deviation of the mean values of modern cultivars, a hierarchical co-clustering analysis was performed (Fig. 2).

The top ten entries based on the Zeleny index—including nine landraces (RMF61, RMF101, RMF103, RMF107, RMF110, RMF117, RMF143, RMF164, and RMF165) and one modern cultivar (RMF258)—were subsequently subjected to detailed analysis using SNP and protein

Table 2
Yield and protein content of samples of the entire population, the landraces and modern cultivars.

	n	2020						2021					
		Yield			Protein content			Yield			Protein content		
		kg/parcel			g/100g			kg/parcel			g/100g		
Landraces + Cultivars	264	0.99	±	0.35	16.43	±	1.44	1.07	±	0.29	15.43	±	1.61
Landraces	197	0.89	±	0.22	16.85	±	1.19	0.95	±	0.18	16.07	±	1.13
Cultivars	67	1.3	±	0.45	15.19	±	1.41	1.42	±	0.28	13.56	±	1.34

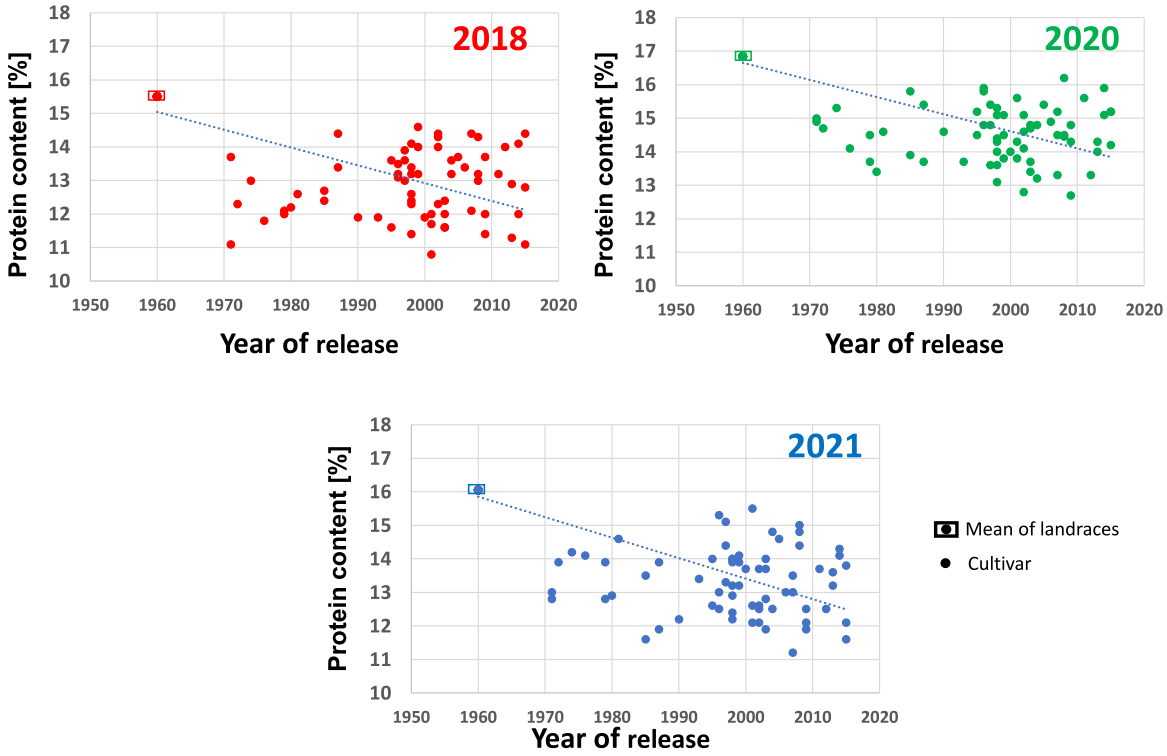

Fig. 1. Protein content of modern cultivars and landraces by year of release. The framed dot indicates the mean protein content of landraces analyzed in this study, while individual dots represent the protein contents of modern cultivars included in the analysis.

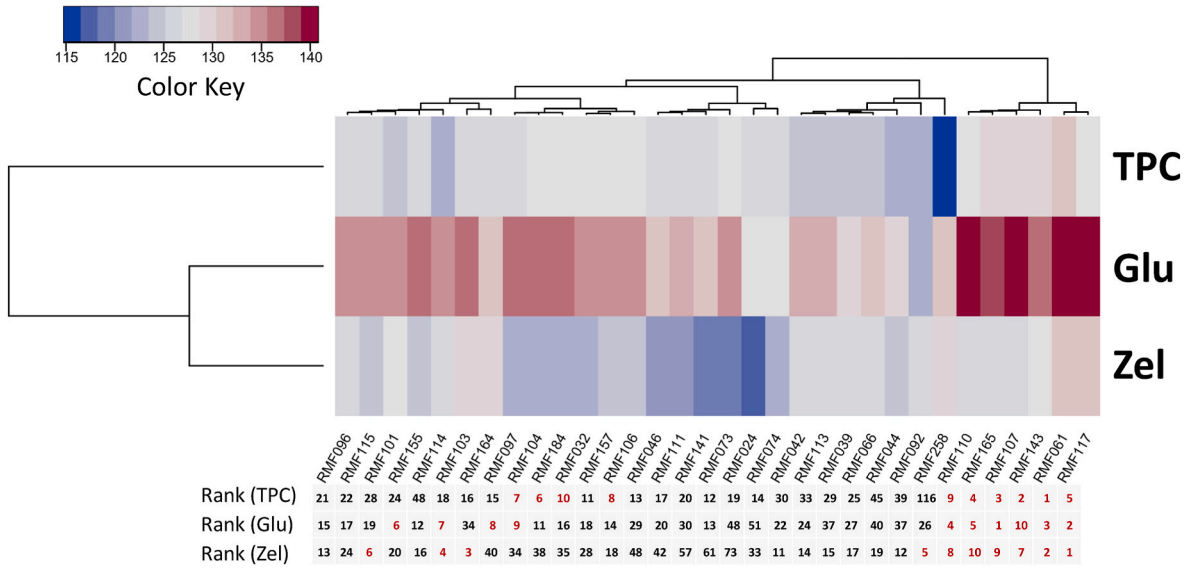

Fig. 2. Clustering multivariate analysis using Ward's method to differentiate 32 winter wheat genotypes based on the phenotypical markers (Total Protein Content (TPC), Glutenin content (Glu), and Zeleny sedimentation value (Zel)) by using hierarchical co-clustering dendrogram and heatmap; cluster rows were obtained at trait level, whereas the cluster columns were recorded at genotype level. The color key represents the percentage-based standard deviation of the mean values for modern cultivars, ranging from dark blue (low) to red (high), with colors intensifying as values increase. The numbers in the "Rank" rows represent the ranking of cultivars based on each respective phenotypic trait, with lower numbers indicating more favorable values. Each trait was ranked independently (See also Supplementary Table 9). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

marker-based classifications.

The Neighbor-Joining (NJ) analysis based on the SNP markers divided the population into three main groups (blue, green and red), and these groups effectively illustrate the historical source and geographical origins of the accessions (Fig. 3).

The SNP-3 (blue) group comprises modern varieties and those local landraces that likely participated in breeding programs. The SNP-2 (green) group is the largest and most diverse group, containing landraces from all countries investigated; however, they likely did not participate in crossing programs for the production of modern cultivars. Genetically, the SNP-1 (red) group is the most distant, and this group primarily consists of Hungarian landraces, as only four of the 37

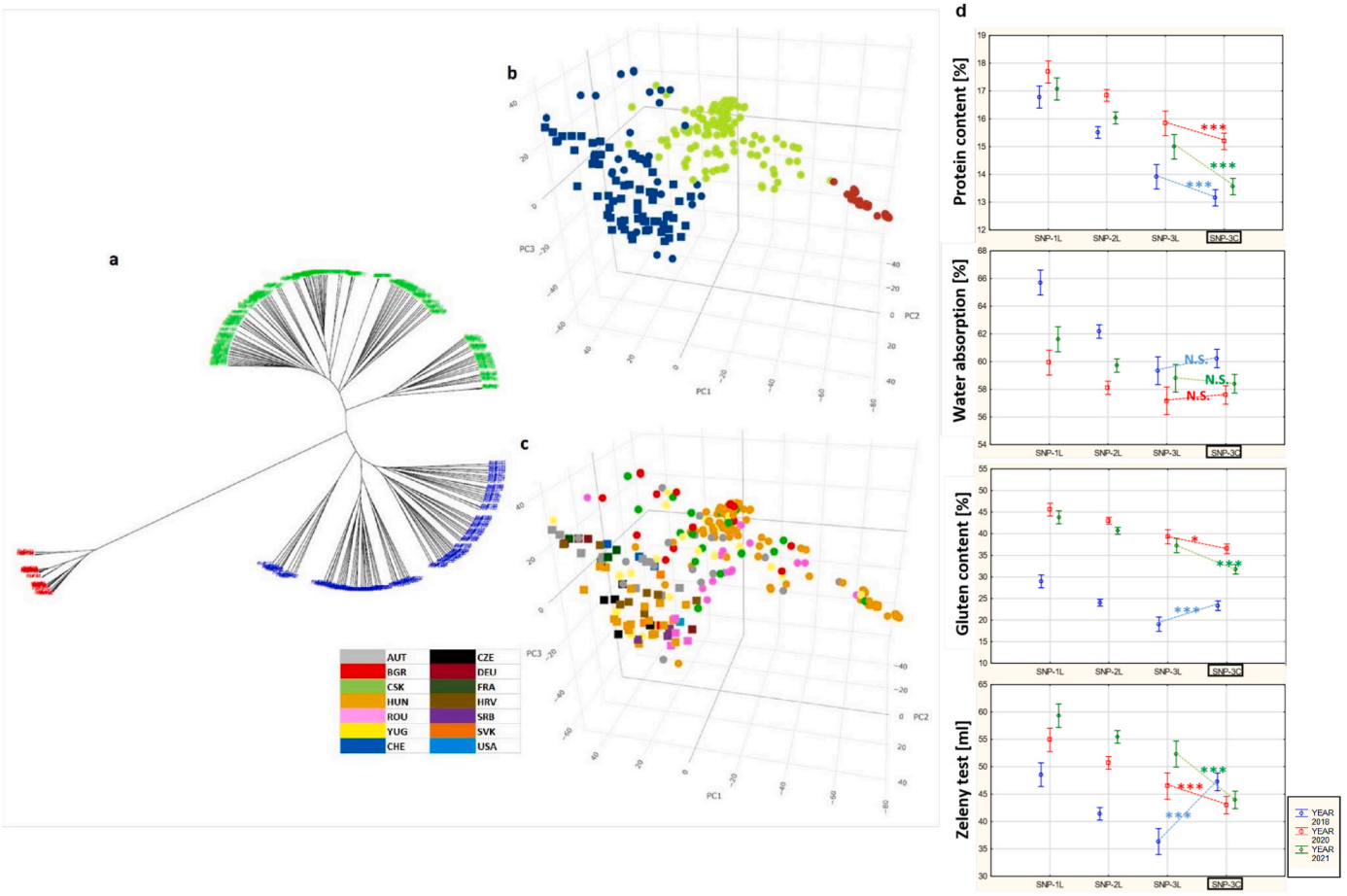

Fig. 3. Genetic structuring of wheat accessions. a) An unrooted Neighbor-Joining (NJ) tree of wheat accessions. The tree branches are colored based on genetic grouping. Group SNP-1 (red) encompasses a total of 37 landraces, there are 130 landraces within Group SNP-2 (green), and Group SNP-3 (blue) consists of 67 modern cultivars and 30 landraces. b) The 3D PCA plot showing the relationship between the accessions belonging to the three groups (blue, red, and green) and modern varieties indicated by squares (See also Supplementary Fig. 1: Interactive 3D PCA plot). c) The 3D PCA plot illustrates the relationship among accessions belonging to different countries and modern varieties indicated by squares (See also Supplementary Fig. 2: Interactive 3D PCA plot with geographical origins). d) Statistical comparison of protein content, water absorption, gluten content, and Zeleny test results among landraces and modern cultivars in the SNP marker groups. Statistically significant differences between landraces and cultivars within the SNP-3 group are indicated with asterisks (***). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

accessions originate from other countries. Based on SNP marker data, five of the nine landraces exhibiting the highest Zeleny sedimentation index were classified into the SNP-1 group, three into the SNP-2 group, and one landrace was grouped together with the single modern cultivar in the SNP-3 group.

The PCA analysis captures the primary sources of variation within the genomic data, revealing distinctions among the sampled regions and illustrating variation within groups of accessions. Consequently, it quantifies the extent of variation between populations that can be attributed to the three principal component axes. In this context, the first and third principal component axes primarily accounted for the genomic variation present in the modern varieties, while the second and third principal component axes predominantly explained the variation observed in the landraces, in accordance with the NJ tree (Fig. 3b). Mostly, the Austrian and former Yugoslavian landraces grouped together with the modern varieties in the left corner of the PCA plot (Fig. 3c), these landraces might have had a greater impact on the genetic structure of the modern wheat cultivars. Protein and gluten content, water absorption, and Zeleny test results across the SNP groups are presented in Fig. 3d. The protein content in samples belonging to the SNP-3 group (blue) was significantly lower than in the other two groups.

A set of statistical comparisons of protein and quality attributes of samples in the three main groups divided by Neighbor-Joining (NJ) analysis has been carried out using ANOVA (Table 3).

Comparing the mean values of each sample in the SNP-1, SNP-2, and SNP-3 groups from the year 2018, the gluten protein composition of the three groups is quite similar. However, when separating the landraces and modern cultivars in SNP-3, significant differences can be observed among the landraces in this group. Glutenin level is significantly lower, while gliadin content is significantly higher in landraces in SNP- 3, resulting in significant differences in the glutenin-to-gliadin ratio and UPP% (Supplementary Fig. 3). The remarkable similarity of these two protein parameters with the Zeleny test results in the 3 SNP groups illustrates the effects of altering protein composition by the selection process in breeding on the functional properties of the modern cultivars. Interestingly, only the SNP-3 group includes modern cultivars, despite SNP-3L landraces displaying the lowest protein content and yield compared to the other landrace groups (Fig. 4).

3.5. *Population structure of the wheat collection by using seed proteins as markers*

The separated proteins by using the MALDI-TOF (HMW GS, LMW GS, gliadin, and soluble proteins) as well as the RP-HPLC method (gliadins) (Supplementary Table 4) was used as protein markers to characterize the genetic diversity of the wheat collection from another point of view.

Table 3
Statistical comparison of protein and functional parameters of samples in the SNP marker groups.

		Landraces + Cultivars					Landraces					Cultivars					Comparing Landraces with Cultivars
		N	mean		stdev		SNP group	N	mean		stdev		SNP group	N	mean		stdev	
Protein content [%]	SNP-1	111	17.18	±	0.13	**C**	SNP-1L	111	17.18	±	0.13	**C**						
	SNP-2	390	16.13	±	0.07	**B**	SNP-2L	390	16.13	±	0.07	**B**						
	SNP-3	291	14.26	±	0.08	**A**	SNP-3L	90	14.92	±	0.15	**A**	SNP-3C	201	13.97	±	0.10	19.73
	F for SNP groups	64.76					95.77											
	F for Environment (Year)	85.48					47.71						37.83					
	F for SNP groups x Environment (Year)	3.59					1.84											
Gluten content [%]	SNP-1	111	39.44	±	0.86	**C**	SNP-1L	111	39.44	±	0.86	**C**						
	SNP-2	390	35.89	±	0.46	**B**	SNP-2L	390	35.89	±	0.46	**B**						
	SNP-3	291	30.95	±	0.53	**A**	SNP-3L	90	31.86	±	0.95	**A**	SNP-3C	201	30.54	±	0.64	1.390
	F for SNP groups	156.84					82.78											
	F for Environment (Year)	730.01					793.44						89.34					
	F for SNP groups x Environment	8.84					2.15											
Water absorption [%]	SNP-1	111	62.41	±	0.31	**C**	SNP-1L	111	62.41	±	0.31	**C**						
	SNP-2	390	60.00	±	0.16	**B**	SNP-2L	390	60.00	±	0.16	**B**						
	SNP-3	291	58.64	±	0.19	**A**	SNP-3L	90	58.44	±	0.34	**A**	SNP-3C	201	58.74	±	0.23	0.95
	F for SNP groups	74.5					57.8											
	F for Environment	111.6					73.9						13.09					
	F for SNP groups x Environment	5.7					4.9											
Zeleny test [ml]	SNP-1	111	54.26	±	0.79	**C**	SNP-1L	111	54.26	±	0.79	**C**						
	SNP-2	390	49.19	±	0.42	**B**	SNP-2L	390	49.19	±	0.42	**B**						
	SNP-3	291	44.83	±	0.49	**A**	SNP-3L	90	45.04	±	0.88	**A**	SNP-3C	201	44.73	±	0.59	0.087
	F for SNP groups	78.72					53.42											
	F for Environment	85.57					153.64						6.16					
	F for SNP groups x Environment	20.56					1.69											

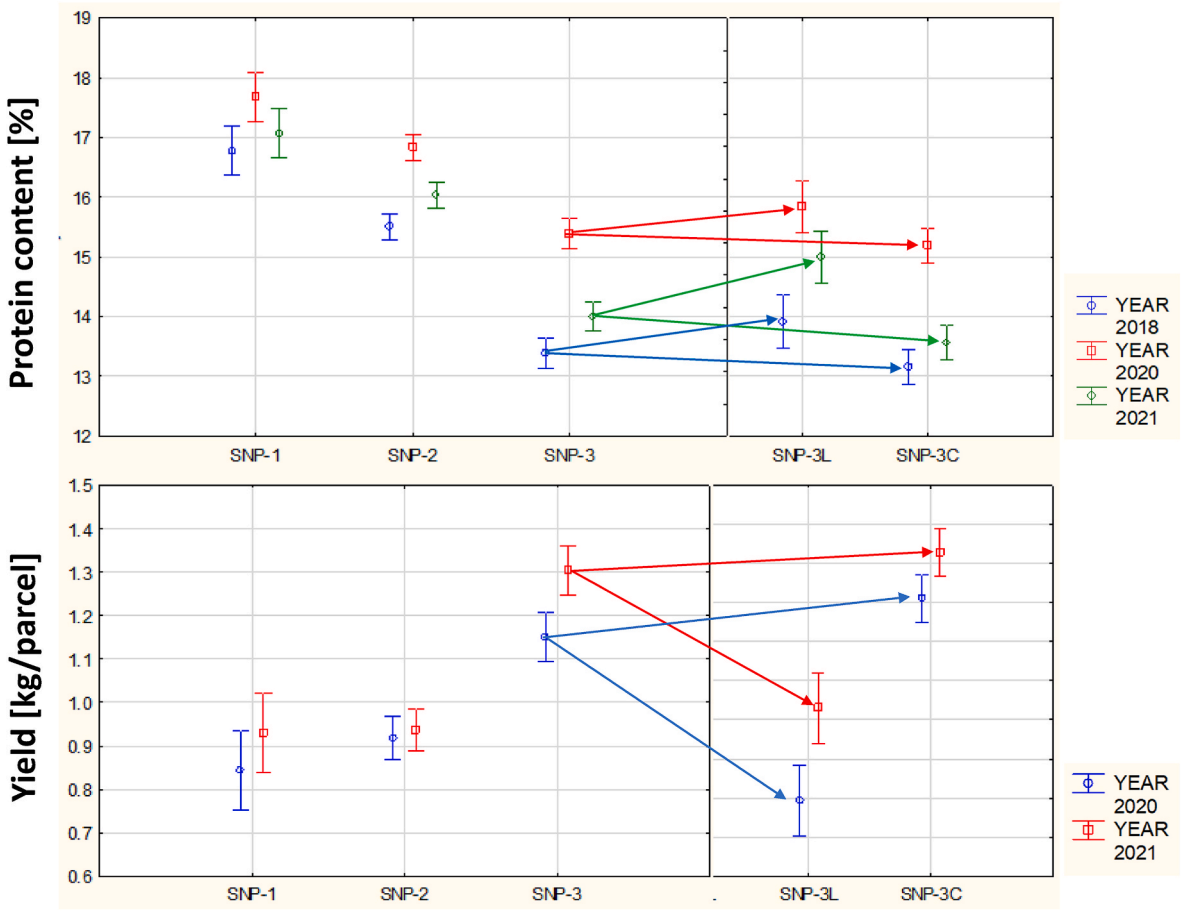

Fig. 4. Comparisons of the mean protein content and yield data in the three SNP groups.

Taking into consideration of different combinations of the protein markers in the evaluation, different sample groups were formed (Supplementary Table 8). Each sample group has been assigned a three-digit code: HMW for HMW glutenin markers; LMW for LMW glutenin markers; GLU for combined HMW and LMW glutenin markers; GLI for gliadin markers; SIK for combined HMW and LMW glutenin as well as gliadin markers; SOL for soluble protein markers ALL for all protein markers. These sample groups also differed based on their protein composition and functional parameters (Supplementary Tables 4 and 5) as approved by the statistical analysis (Supplementary Table 10). In this statistical analysis, the treatment of data was similar to that applied in the SNP marker analysis. Thus the statistical analysis compared the means of protein and functional data of i) each sample in the groups (Landraces + Modern cultivar, L + C), ii) the landraces (L) in the group, iii) the modern cultivars (C) in the group and iv) comparing data of landraces with those of the modern cultivars in the same group.

3.5.1. *Application of glutenin markers*

An unrooted Neighbor-Joining (NJ) tree (Fig. 5) shows the population structure of wheat accessions based on MALDI-TOF MS HMW glutenin subunit markers (HMW grouping). This tree distinguishes two main groups of accessions, each represented by a different color.Due to the limited number of markers, not every accession could be differentiated; the figure only depicts the lines that differ from each other.

The two groups (HMW-1 and HMW-2 (encompassing landraces and cultivars)) were compared based on the mean values of the protein traits and the functional parameters, and a significant difference was observed between them ($F = 18.55$, 16.16, 6.57, and 6.86 for protein and gluten content, water absorption, and the Zeleny test, respectively (Supplementary Table 10-HMW). Separating the landraces and cultivars, no significant differences were found between the means of the HMW-1L and HMW-2L groups, as well as between the means of the HMW-1C and HMW-2C groups. Comparing the data of landraces with the modern cultivars in the same groups (HMW-1L compared with HMW-1C and HMW-2L with HMW2C), large, significant differences have been found within each group in the protein content ($F = 168.22$), gluten content ($F = 20.56$), water absorption ($F = 20.32$), and Zeleny test results ($F = 24.29$).

The data presented in Supplementary Table 10-HMW-A indicate that harvest year significantly influenced the measured parameters. Compared to the landraces, protein content consistently declined in cultivars across all harvest years. Gluten content showed a significant reduction in the 2020 and 2021 samples, while no statistically significant decrease was observed in the 2018 samples. Notably, reductions in water absorption were significant only in the 2018 HMW-1C samples. The Zeleny test data showed the strongest effects of growing conditions: in contrast to the landrace data, there were notable declines in cultivars in the 2020 and 2021 samples whereas a notable rise was observed in 2018. These alterations of Zeleny values in the 2018 samples correspond nicely to the protein composition findings: as indicated in Supplementary Table 10-HMW-B, in this harvest year, the glutenin-to-gliadin ratio and UPP% values of the samples are significantly higher in the modern cultivars compared to the values observed in the landraces of the same HMW group. All of these differences found in the protein composition resulted in differences in the functional parameters; thus, modern cultivars were found to have better functional properties than landraces in both HMW groups (red and blue), and this difference was greater within the HMW-2 (blue) group than in the HMW-group 1 (red) group (Fig. 5).

Based on the MALDI-TOF MS HMW markers, five of the nine landraces with the highest Zeleny index rankings were assigned to the HMW-

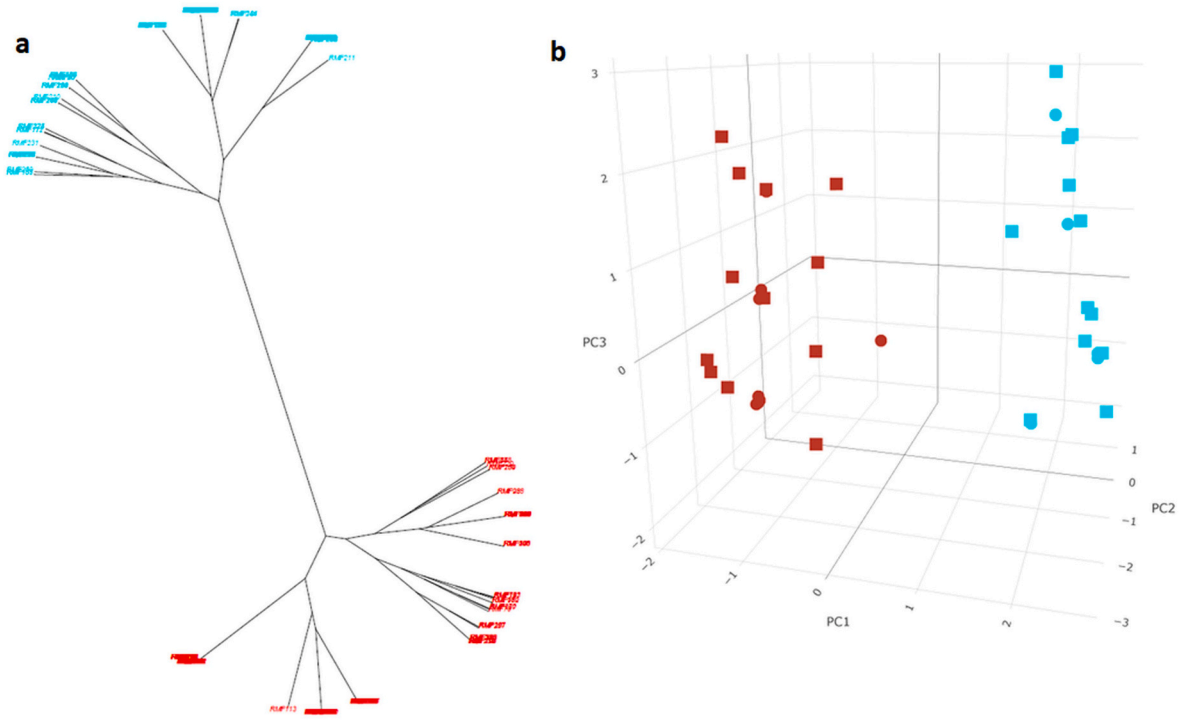

Fig. 5. *Population structure of wheat accessions based on MALDI-TOF MS HMW glutenin subunit markers. a) An unrooted Neighbor-Joining (NJ) tree of wheat accessions. The two branches are colored based on the population structure. HMW 1 (red) group consists of 130 landraces and 28 modern cultivars; there are 67 landraces and 39 modern cultivars within the HMW 2 (blue) group. b) The 3D PCA plot showing the relationship between the accessions belonging to the two groups (red and blue) and modern varieties indicated by squares. The low number of markers did not allow for the differentiation of every accession; the figure represents only the lines differing from each other (See also* Supplementary Fig. 4*: Interactive 3D PCA plot: MALDI-TOF MS HMW glutenin).* (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

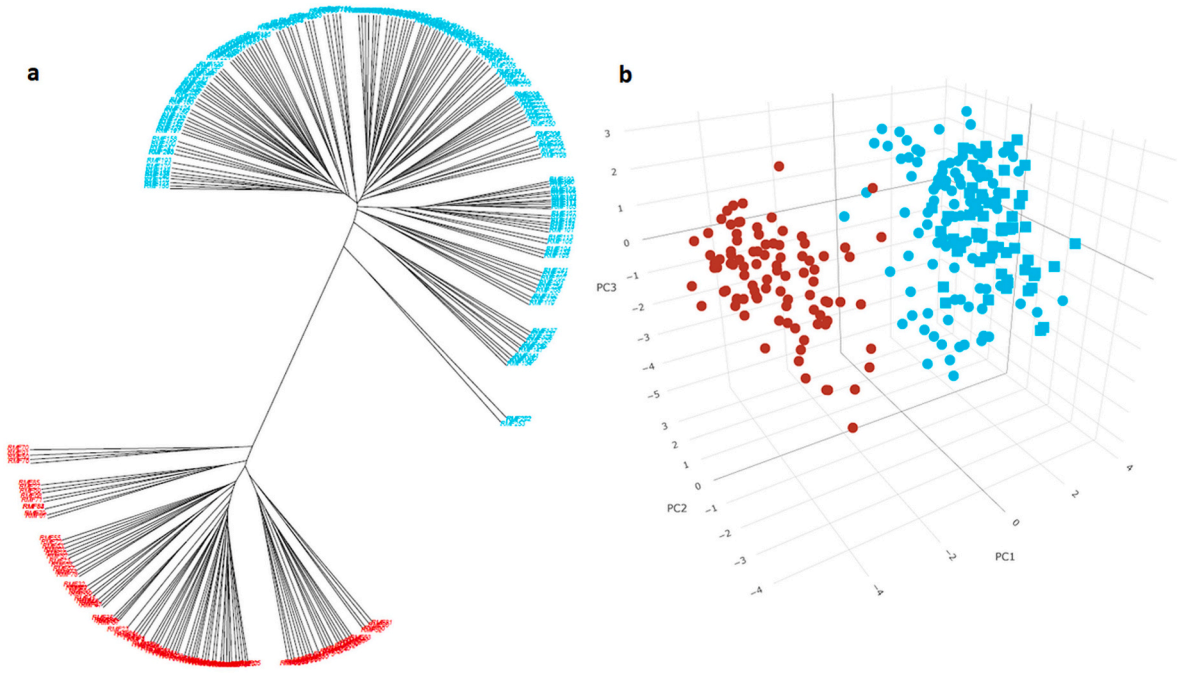

Fig. 6. *Population structure of wheat accessions based on MALDI-TOF MS LMW glutenin subunit markers. a) An unrooted Neighbor-Joining (NJ) tree of wheat accessions. The two branches (LMW-1 and LMW-2) are colored according to the population structure. LMW-1 group 1 (red) consists of 96 landraces; there are 101 landraces and 67 modern cultivars within the LMW-2 (blue). b) The 3D PCA plot showing the relationship between the accessions belonging to the two groups (red and blue) and modern varieties indicated by squares (See also* Supplementary Fig. 5*: Interactive 3D PCA plot).* (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

1 group. The remaining four landraces, along with the single modern cultivar, were classified into the HMW-2 group, consistent with the grouping of most modern cultivars. The unrooted Neighbor-Joining (NJ) tree of wheat accessions and the 3D PCA plot based on the MALDI-TOF MS LMW glutenin subunit markers (Fig. 6a and b) formed two groups (LMW-1 and LMW-2). The two groups (branches) are colored according to the population structure. Furthermore, the 3D PCA plot shows the relationship between the accessions belonging to the landraces (dots) and modern varieties (squares). The red LMW-1 group comprises 96 landraces, whereas the blue LMW-2 group includes 101 landraces and 67 modern cultivars. Of the top ten accessions ranked by the Zeleny index, only one landrace (RMF61) belongs to the LMW-1 group. The remaining eight landraces, together with the modern cultivar (RMF258), are classified within the LMW-2 group.

A comparative analysis of F values revealed a pronounced masking effect from modern cultivars within the LMW-1 group. Specifically, the LMW-1L subset (landraces only) exhibited substantially lower F values—1.81, 4.21, 0.58, and 0.50 for protein content, gluten content, water absorption, and Zeleny sedimentation index, respectively—compared to the combined group of landraces and cultivars, which showed F values of 39.23, 85.36, 17.60, and 33.21 (Supplementary Table 10-LMW).

Similarly to the above-described observations for the HMW groups, cultivars (in LMW-2C) from the 2018 harvest had significantly higher Zeleny sedimentation index than the landraces (LMW-2L), while significant drops have been found in the corresponding data of the 2020 and 2021 samples.

The population structure of wheat accessions based on the combined use of MALDI-TOF MS HMW and LMW glutenin subunit markers (Fig. 6a), consists of five groups. The red GLU-1 group encompasses a total of 67 landraces. The light green GLU-2 group is formed by 29 landraces. The green GLU-3 group includes 27 modern cultivars and 63 landraces. The blue GLU-4 group consists of 37 modern cultivars and 38 landraces. The purple GLU-5 group includes only three modern wheat cultivars.

The distribution of modern cultivars and landraces in the five groups was determined by the combined use of HMW and LMW GS markers (GLU grouping). GLU-1 and GLU-2 (red and light green) did not include modern varieties, while group GLU-5 (purple) included only 3 cultivars. Additionally, 27 cultivars from group 3 and 37 out of 39 cultivars from GLU-2 (light green) of the HMW GS marker-based structure formed GLU-4 (blue) in the structure based on the HMW + LMW GS markers.

The modern cultivars with good baking quality were located in the GLU-3 (green) and GLU-4 (blue) groups (Fig. 7), - with significantly lower protein content – masking the comparisons of the five overall (landraces and cultivars) groups. The GLU-3 group comprises five landraces ranked among the top performers based on the Zeleny index. The GLU-1 group contains only one of the Zeleny-selected landraces (RMF61), while the HMW-4 group includes four landraces along with the single modern cultivar. Significant differences were however detected between the 5 GLU groups when their corresponding landraces or cultivars were compared, but mostly in the landraces comparisons, while no significant differences were found between the GLU-3C, GLU-4C, and GLU-5C groups for each protein and quality parameter (Supplementary Table 10-GLU). When comparing L to C within the same GLU group, the difference between them (C vs L) was only significant for the GLU-3 and GLU-4 groups. Similarly to the observations found for the HMW and LMW groupings, the extent and the directions of these alterations differ for the four investigated parameters and for the samples derived from different harvesting seasons (Supplementary Table 10-GLU).

3.5.2. *Application of gliadin markers*

Population structures of wheat accessions were displayed separately based on the MALDI-TOF MS Gliadin protein (Fig. 8) or the RP-HPLC Gliadin protein (Fig. 9) composition, and on their combined data (Fig. 10), using NJ and PCA analysis (GLI grouping, Fig. 8, GLX grouping, Fig. 9, and GLY grouping, Fig. 10, respectively). More than 80 % of the modern cultivars formed a distinct cluster in all three evaluations (Fig. 8: GLI-4 [purple], Fig. 9: GLX-2 [green], Fig. 10: GLY-2 [green]). The mean values of the functional parameters for these groups were significantly different from those of the other groups, which

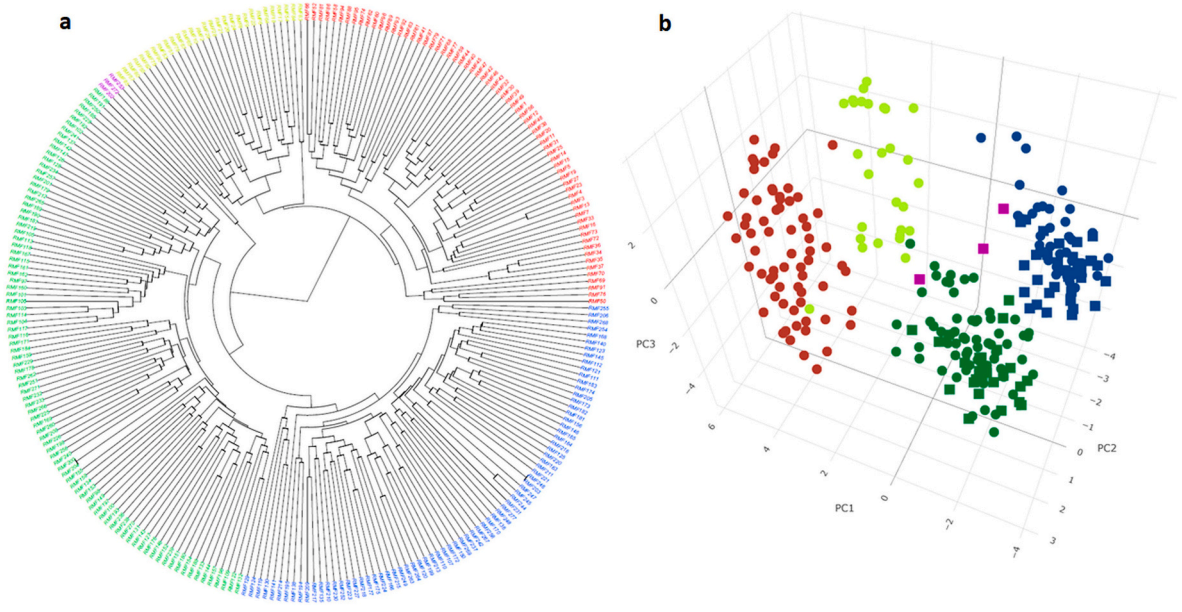

Fig. 7. Population structure of wheat accessions based on MALDI-TOF MS LMW and MALDI-TOF MS HMW glutenin subunit markers (GLU grouping). a) Circular Neighbor-Joining (NJ) tree of wheat accessions. GLU-1 (red) group encompasses a total of 67 landraces, GLU-2 (light green) group is formed by 29 landraces, GLU-3 (green) group includes 27 modern cultivars and 63 landraces, HMW-4 (blue) group consists of 37 modern cultivars and 38 landraces, while the GLU-5 (purple) group only includes three modern wheat cultivars b). The 3D PCA plot showing the relationship between the accessions belonging to the five groups (red, blue, light green, green and purple) and modern varieties indicated by squares (See also Supplementary Fig. 6: Interactive 3D PCA plot). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

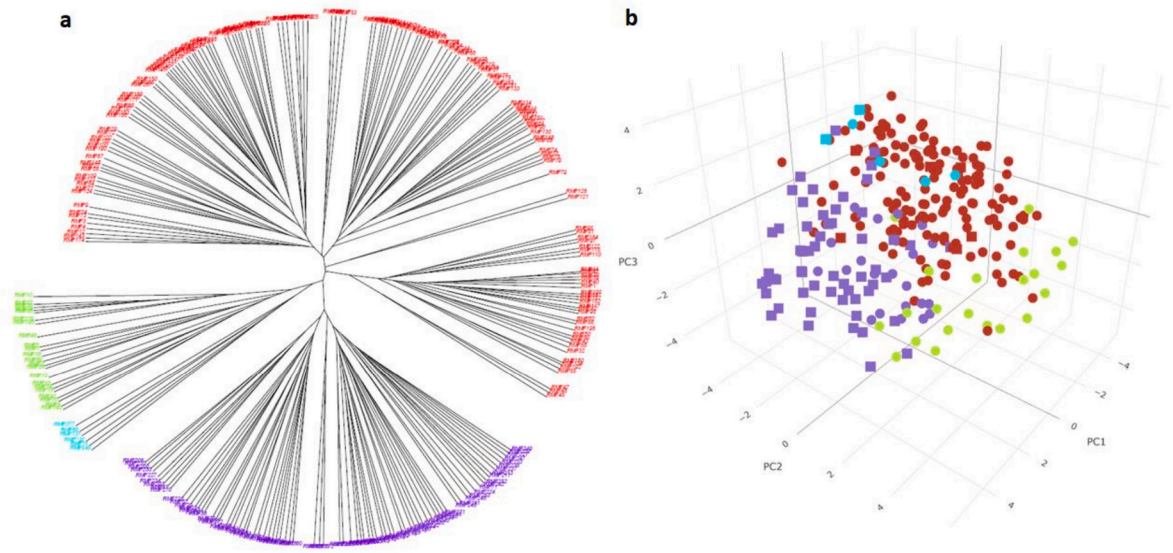

Fig. 8. Population structure of wheat accessions based on MALDI-TOF MS Gliadin protein markers (GLI groups). a) An unrooted Neighbor-Joining (NJ) tree of wheat accessions. The four branches are colored according to the population structure. GLI- 1 (red) group comprises 148 landraces and 8 modern cultivars; GLI-2 (green) group encompasses a total of 22 landraces; GLI-3 (blue) group comprises four landraces and two modern cultivars and GLI-4 (purple) group includes 23 landraces and 57 modern cultivars b) The 3D PCA plot showing the relationship between the accessions belonging to the four GLI-1, GLI-2, GLI-3 and GLI-4 groups (red, green, blue and purple, respectively) and modern varieties indicated by squares (See also Supplementary Fig. 7: Interactive 3D PCA plot). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

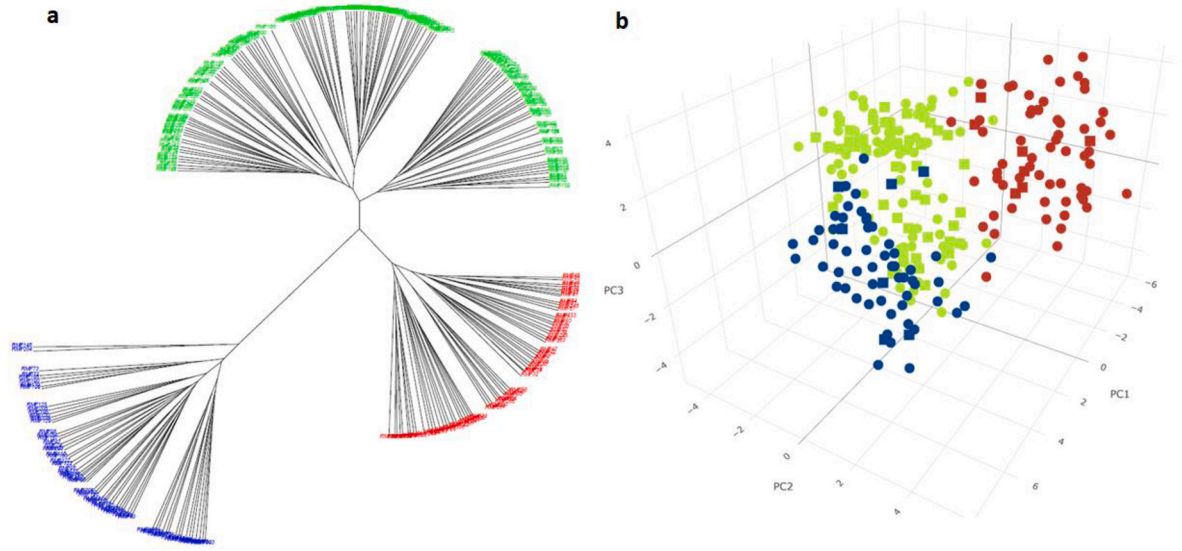

Fig. 9. Population structure of wheat accessions based on RP-HPLC Gliadin protein markers (GLX groups. a). An unrooted Neighbor-Joining (NJ) tree of wheat accessions. The tree branches are colored according to the population structure. GLX-1 (red) group encompasses 61 landraces and 7 modern cultivars, there are 79 landraces and 53 modern cultivars within GLX-2 (green) group, and GLX-3 (blue) group consists of 57 landraces and 7 modern cultivars. b) The 3D PCA plot showing the relationship between the accessions belonging to the GLX-1, GLX-2 and GLX-3 groups (blue, green and red, respectively) and modern varieties indicated by squares (See also Supplementary Fig. 8: Interactive 3D PCA plot). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

exhibited lower protein contents (Supplementary Table 10). The cultivars with the best quality parameters were grouped separately, belonging mostly to the green GLX-2 and GLY-2 groups, which were based on the RP-HPLC or on the combined MALDI-TOF MS plus RP-HPLC Gliadin protein markers (Supplementary Table 10-GLI). According to MALDI-TOF MS gliadin protein marker profiling, all landraces selected based on the Zeleny index were classified into the GLI-2 group, whereas the single modern cultivar was assigned to the GLI-4 group.

Based on RP-HPLC gliadin marker classifications, two of the landraces selected by Zeleny index ranking were assigned to the GLX-1 group. Six landraces, along with the modern cultivar, were grouped into GLX-2, while one landrace was classified within the GLX-3 group.

Mean values of NIR-based data from three harvest years were compared for landraces + cultivars, landraces, and cultivars among the different GLI map groups (Supplementary Table 10-GLI). For all the parameters, F values obtained from landraces + cultivars were significantly higher than those calculated from landraces alone. Among landrace groups, both protein contents and Zeleny test values were found to be the lowest in the GLI-4L group. In the cultivar comparison, the protein content, and the Zeleny test data of GLI-1C was significantly

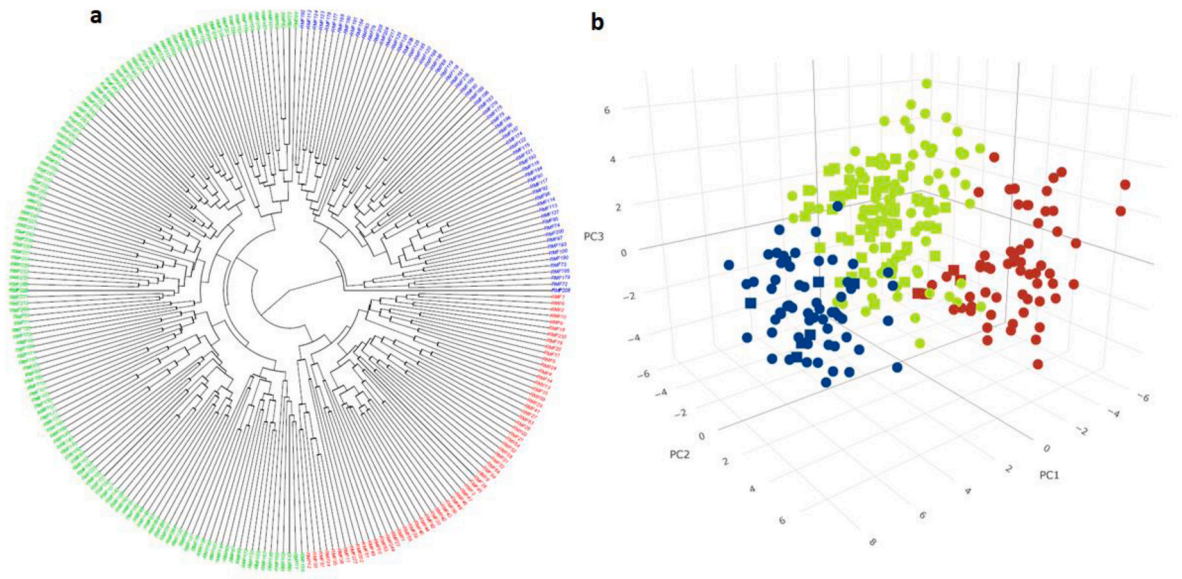

Fig. 10. Population structure of wheat accessions based on combining the MALDI-TOF MS Gliadin protein and RP-HPLC Gliadin protein markers (GLY groups). a) Circular Neighbor-Joining (NJ) tree of wheat accessions. The tree branches are colored according to the population structure. GLY-1 (red) group encompasses 59 landraces and 4 modern cultivars, GLY-2 (green) group consists of 80 landraces and 56 modern cultivars, and there are 58 landraces and 7 modern cultivars within GLY-3 (blue) group. b) The 3D PCA plot showing the relationship between the accessions belonging to the GLY1, GLY-2 and GLY-3 groups (blue, green and red, respectively) and modern varieties indicated by squares (See also Supplementary Fig. 9: Interactive 3D PCA plot). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

lower than that in GLI-3C.

The overall comparisons between the landraces and cultivars in the same group (comparison of GLI-4L with GLI-4C) utilizing data from three harvest years revealed, that landraces had higher Zeleny test results than the cultivars. However, as shown in the figure of Supplement Table 10-GLI, cultivar samples from 2018 had significantly higher Zeleny test values than landraces cultivated in the same year, while the relationship reversed in 2020 and 2021. The improved sedimentation characteristics (measured by Zeleny test) in the 2018 cultivars correlate well with the glutenin-to-gliadin ratio and with the UPP% of these samples (Supplement Table 10-GLI**).** Based on combined MALDI-TOF MS and RP-HPLC gliadin protein marker data, only one of the nine landraces selected by Zeleny index ranking (RMF117) was classified into the GLY-3 group, while the remaining eight landraces, along with the modern cultivar, were assigned to the GLY-2 group.

3.5.3. Application of salt-soluble protein markers

Using the NJ clustering method, the salt-soluble protein markers were divided the accessions into two major groups (Fig. 11). The red (**SOL-1**) group consists of 57 landraces, the blue (**SOL-2**) group contains 140 landraces and 67 cultivars. Supplementary Table 10-SOL shows a statistical comparison of the SOL groups: there is a substantial difference

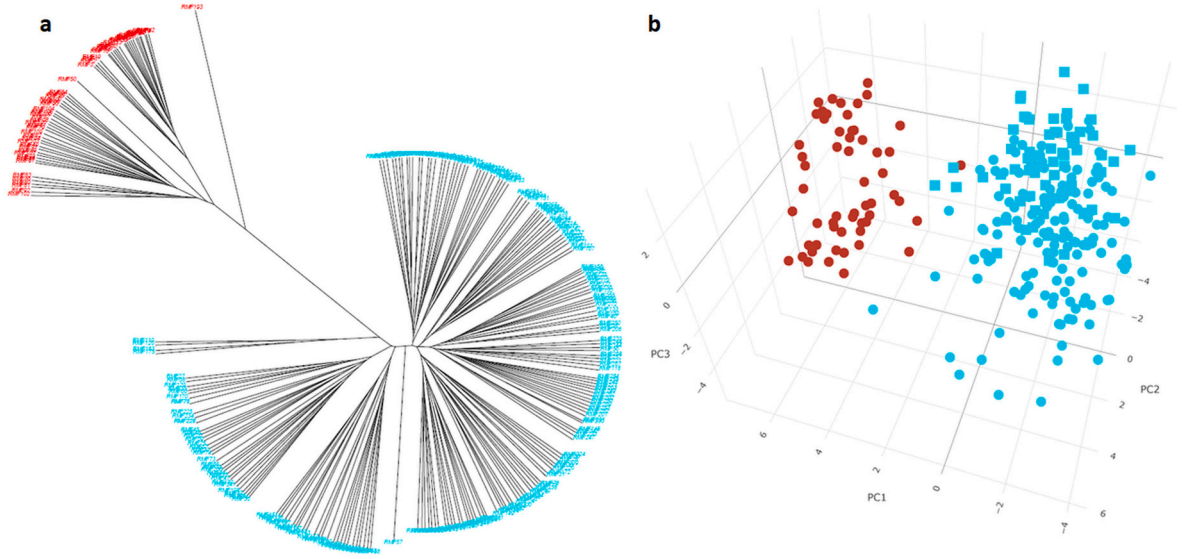

Fig. 11. Population structure of wheat accessions based on Salt-soluble protein markers. a) An unrooted Neighbor-Joining (NJ) tree of wheat accessions. The two branches are colored according to the population structure. SOL- 1 (red) group consists of 57 landraces; there are 140 landraces and 67 modern cultivars within the SOL-2 (blue) group. b) The 3D PCA plot showing the relationship between the accessions belonging to the two groups (red and blue) and modern varieties indicated by squares (See also Supplementary Fig. 10: Interactive 3D PCA plot). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

in protein content, as well as a modest but not significant difference in Zeleny test results between the **SOL-1L** and **SOL-2L** groups. The classification based on salt-soluble protein markers does not fully align with the Zeleny index-based ranking. Approximately half of the landraces were assigned to the SOL-1 group, while the remaining landraces were grouped into SOL-2.

3.6. Identification of similar lines based on hierarchical co-clustering analysis

The hierarchical co-clustering dendrogram and heatmap (Fig. 12) show the distribution of the 264 winter wheat genotypes according to the groups obtained from the NJ trees of protein profiling. The clustering multivariate analysis was performed by using the Ward's method. The cluster rows display the groups obtained from the NJ trees of protein profiling and SNP genotyping, whereas the cluster columns represent the wheat accessions. On the heat map, red color indicates that the corresponding accession line is included in the group obtained from the NJ trees of protein profiling, while the yellow color indicates that the wheat accession is not included in the group. The rows numbered from 1 to 16 are represented the following protein and SNP genotyping groups: **1**: Salt-soluble proteins Group 2 (SOL-2), **2**: MALDI-TOF MS LMW glutenin subunit Group 2 (LMW-2), **3**: RP-HPLC Gliadin Group 2 (GLX-2), **4**: SNP genotyping Group 3 (SNP-3), **5**: MALDI-TOF MS Gliadin Group 4 (GLI-4), **6**: MALDI-TOF MS HMW glutenin subunit Group 2 (HMW-2), **7**: Salt-soluble proteins Group 1 (SOL-1), **8**: SNP genotyping Group 1 (SNP-1), **9**: MALDI-TOF MS Gliadin Group 3 (HLI-3), **10**: MALDI-TOF MS Gliadin Group 2 (GLI-2), **11**: RP-HPLC Gliadin Group 3 (GLX-3), **12**: MALDI-TOF MS LMW glutenin subunit Group 1 (LMW-1), **13**: RP-HPLC Gliadin Group 1, **14**: SNP genotyping Group 2 (SNP-2), **15**: MALDI-TOF MS Gliadin Group 1 (GLI-1), **16**: MALDI-TOF MS HMW glutenin subunit Group 1 (HMW-1). Further details regarding the groups are available in Supplementary Table 8. The modern cultivars are located in the first six rows (1–6 groups). The dendrogram on the top of the figure (Fig. 12) divided the population into three main groups, which are from the left to the right: **(I)** Landraces similar to the modern cultivars; **(II)** Modern cultivars – divided into two subgroups by the HMW glutenin Group 2 - and **(III).** Landraces are strongly different from the modern cultivars. The list of the lines in these three groups is listed in Supplementary Table 1 and 8

The statistical comparison of the three dendrogram groups is illustrated in Table 4 and Supplementary Table 11. Each parameter indicated a highly significant improvement from the two landrace groups (II.-group), and the extent of the differences between the means for group II-group III was greater than for those for group II-group I.

The three groups represent: (I) landraces similar to modern cultivars, (II) modern cultivars, and (III) landraces that differ substantially from modern cultivars.

3.7. The impact of HMW-, LMW subunits, and gliadin markers evaluated with MALDI TOF MS and RP-HPLC on quality parameters

The NJ clustering approach classified the population into three groups (SIK-1, SIK-2, and SIK-3) by combining the MALDI-TOF MS HMW, MALDI-TOF MS LMW, MALDI-TOF MS Gliadin, and RP-HPLC Gliadin protein markers (Fig. 13). Modern wheat varieties exhibited a limited number of new variations due to the predominant use of the local landraces in breeding programs. Only one group (GLU-4), containing only 3 cultivar samples, was found where cultivars were located without any landrace in the same group (Fig. 7). The PCA grouping (Figs. 5b-11b) clearly illustrates this pattern, with certain landraces clustering alongside current cultivars.

With one exception, all landraces selected based on the Zeleny index, together with the modern cultivar, were classified into the SIK-2 (green) group. The remaining landrace, RMF117, was assigned to the SIK-3 (blue) group.

Statistical investigation of SIK groups is shown in Supplementary Table 10-SIK: 2-ways ANOVA has been used to characterize the NIR-based four functional parameters (protein and gluten content, water absorption and Zeleny test results) of samples grown 2018, 2020 and 2021 (Supplementary Table 10-SIK-a). The results of one-way ANOVA investigation for the protein composition of the 2018 samples is shown in the Supplementary Table 10-SIK: b) Comparison of landraces + landraces, (SIK-1, SIK-2 and SIK-3); c) only landraces, (SIK-1L, SIK-2L, and SIK-3L); and d) only cultivars (SIK-2C, SIK-3C) groups in samples of the 2018 harvest season. F – F value and p - probability calculated by ANOVA, different capital letters designate significant differences between the decades (one-way ANOVA and Tukey's test; $p < 0.05$).

SIK-2 (green) and **SIK-3** (blue) groups contain landraces with very similar protein composition and functionality, while there are significantly higher amounts of ω and γ gliadin and significantly lower amounts of α/β gliadin in the **SIK-1** (red) group**.** So, the differences found by comparing the parameters of the entire population were caused by the presence of the modern cultivars in groups 2 and 3, but not in

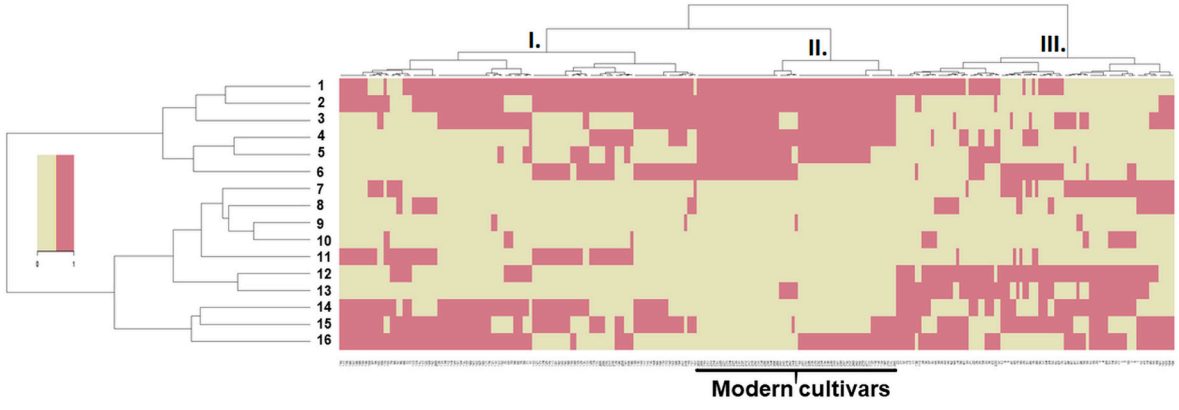

Fig. 12. The hierarchical co-clustering dendrogram and heatmap showing the distribution of the 264 winter wheat genotypes based on the groups obtained from the NJ trees of protein profiling, The rows numbered from 1 to 16 represent the following protein and SNP genotyping groups: **1**: Salt-soluble proteins Group 2 (SOL-2), **2**: MALDI-TOF MS LMW glutenin subunit Group 2 (LMW-2), **3**: RP-HPLC Gliadin Group 2 (GLX-2), **4**: SNP genotyping Group 3 (SNP-3), **5**: MALDI-TOF MS Gliadin Group 4 (GLI-4), **6**: MALDI-TOF MS HMW glutenin subunit Group 2 (HMW-2), **7**: Salt-soluble proteins Group 1 (SOL-1), **8**: SNP genotyping Group 1 (SNP-1), **9**: MALDI-TOF MS Gliadin Group 3 (HLI-3), **10**: MALDI-TOF MS Gliadin Group 2 (GLI-2), **11**: RP-HPLC Gliadin Group 3 (GLX-3), **12**: MALDI-TOF MS LMW glutenin subunit Group 1 (LMW-1), **13**: RP-HPLC Gliadin Group 1, **14**: SNP genotyping Group 2 (SNP-2), **15**: MALDI-TOF MS Gliadin Group 1 (GLI-1), **16**: MALDI-TOF MS HMW glutenin subunit Group 1 (HMW-1). Red cells indicate that the corresponding accession line is included in the group. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Table 4
Analysis of variance comparing the functional characteristics of samples from the 3 harvest seasons in the three dendrogram groups separated on the hierarchical co-clustering dendrogram (Fig. 12).

Year	Groups	n	Protein content			Gluten content			Water Absorption			Zeleny test		
2018	I	113	15.26	±	0.13	23.96	±	0.48	62.21	±	0.28	42.17	±	0.68
	II	63	13.16	±	0.18	23.04	±	0.65	59.67	±	0.38	45.88	±	0.92
	III	88	15.68	±	0.15	24.67	±	0.54	62.91	±	0.31	43.01	±	0.77
2020	**I**	113	16.80	±	0.13	42.69	±	0.48	58.15	±	0.28	50.33	±	0.68
	II	63	15.27	±	0.18	36.71	±	0.65	57.41	±	0.38	43.03	±	0.92
	III	88	16.75	±	0.15	42.66	±	0.54	58.59	±	0.31	51.02	±	0.77
2021	**I**	113	15.91	±	0.13	40.19	±	0.48	59.89	±	0.28	55.00	±	0.68
	II	63	13.63	±	0.18	32.08	±	0.65	58.33	±	0.38	44.35	±	0.92
	III	88	16.08	±	0.15	40.69	±	0.54	59.95	±	0.31	55.64	±	0.77

	F	**p**	**F**	**p**	**F**	**p**	**F**	**p**
YEAR	81.39	0.00	753.86	0.00	91.10	0.00	75.18	0.00
Heatmap group	152.72	0.00	72.70	0.00	27.50	0.00	35.30	0.00
YEAR * heatmap group	3.34	0.01	11.88	0.00	2.90	0.02	25.04	0.00

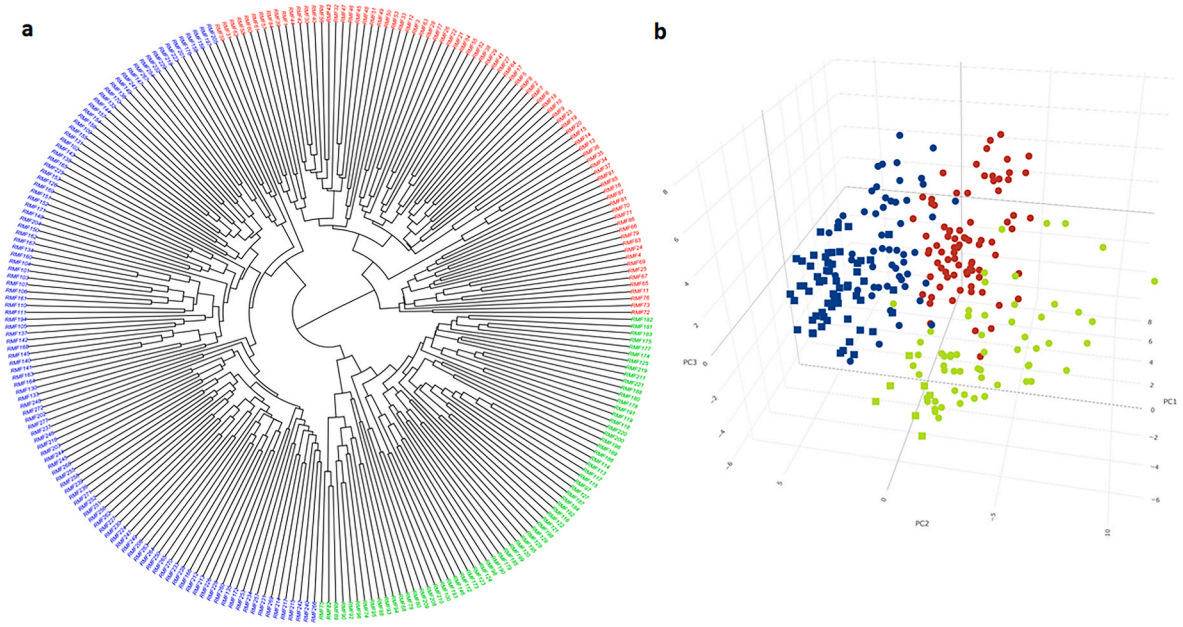

Fig. 13. Population structure of wheat accessions based on combining the MALDI TOF MS LMW, MALDI TOF MS HMW, MALDI TOF MS Gliadin, RP HPLC gliadin protein markers (SIK groups). a) Circular Neighbor-Joining (NJ) tree of wheat accessions. The tree branches are colored according to the population structure. SIK-1 (red) group encompasses 82 landraces,SIK-2 (green) group consists of 60 landraces and 7 modern cultivars, and there are 55 landraces and 60 modern cultivars within SIK-3 (blue) group. b) The 3D PCA plot showing the relationship between the accessions belonging to the SIK-1, SIK-2 and SIK-3 groups (red, green and blue respectively) and modern varieties indicated by squares (See also Supplementary Fig. 11: Interactive 3D PCA plot). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

group 1.

In the case of **SIK-2** (green) and **SIK-3** (blue) groups, all parameters of modern cultivars are significantly different from those of landraces within the same group. Comparing the parameters of cultivars from groups 2 and 3 reveals slightly, non-significant changes, despite the fact that the lines in the two groups come from distinct gene pools.

4. Discussion

4.1. The effect of the selection based on protein composition during modern breeding

Wheat cultivation must intensify, become more sustainable, and achieve greater resilience towards pests and climate change simultaneously. Bread wheat is traditionally one of the main food sources for humankind, and it must also meet the requirements of healthy nutrition and the expectations of the baking industry. During the 'Green Revolution', wheat breeding primarily concentrated on creating high-yielding, disease-resistant wheat varieties incorporating dwarfing genes, which led to a reduction in genetic diversity among modern elite varieties [3,61]. However, our heritage left behind a wealth of variation present in landraces, including the protein content and composition, which affect the bread-making quality. The Central European wheat landrace collection, comprising 197 accessions from six countries, was gathered between 1950 and 1960 and preserved in the Gene Bank collection of the Centre for Plant Diversity (NÖDIK) in Tápiószele, Hungary. The quantitative analysis of certain key functional properties and the diversity of protein markers detected in the present study identified a group of genotypes that conserved allelic diversity absent in modern varieties. Near-infrared (NIR) spectroscopy is a well-established and suitable method for screening basic quality attributes, particularly in large-scale populations [62]. The Zeleny index is a widely used parameter in wheat breeding programs for predicting overall baking quality, typically determined from ground seed or flour samples [63].

Hrušková and Faměra [64] found a statistically significant relationship ($p < 0.01$) between NIR-predicted and empirically measured values of protein content and Zeleny sedimentation.

Separating the population of investigated samples into 197 landraces and 67 modern varieties allowed important observations to be made on the effects of the breeding process on the diversity of the different protein classes (Fig. 12). The reduced protein content was compensated by a variable protein composition to keep or even to improve the functional properties of the new wheat varieties.

Selection and breeding resulted in increased glutenin levels and decreased gliadin levels, leading to a higher glutenin-to-gliadin ratio (Supplementary Fig. 12). Additionally, a shift in the size distribution of polymeric glutenin proteins towards higher molecular weight fractions was observed, as characterized by the UPP% based on HPLC analysis from 2018 (Supplementary Table 6). Both changes contributed to improved dough strength and stability, as well as increased dough development time, in agreement with previous findings on wheat landrace collections [65–69]. Meanwhile, the water absorption, another important functional parameter strongly related to protein content, showed a significant decrease in parallel with the reduction of the protein content.

Selection strategies during wheat breeding through the 20th century did not prioritize nutritive or health-related attributes [46]. As a result of the selection process primarily aiming at yield improvement and the preservation of functional properties, significant changes occurred in the gliadin composition, directly related to the amount of allergenic and coeliac epitopes of the samples. We found significantly lower α/β gliadin levels in the modern cultivars than in the landraces (Supplementary Table 6). This reduction (expressed as the percentage of the dry matter) resulted from the reduced total protein content, the lower total amount of gliadins and changes in the distribution of the gliadin subclasses within the total gliadin content. The relative amount of ω1,2 and ω5 gliadins did not change significantly, but an increase was observed in the γ gliadins content. The α/β gliadin group is potentially the most nutritionally harmful protein class for the sensitive individuals; thus, the reduction in their quantity indicates that breeding resulted in the development of less harmful germplasms. These findings contradict the hypothesis that human activities—particularly wheat breeding—may have contributed to the increased prevalence of coeliac disease (CD) in the latter half of the twentieth century, as suggested by van den Broeck et al. [70] and de Lorgil et al. [71]. In general, the protein composition and functional properties have changed during modern breeding, with a decrease in protein content and an increase in grain yield. In agreement with the research conducted by several groups [5–8,10,11,36] our results approved the reciprocal relationship between the protein content and the grain yield. On the other hand, we showed a declining trend between the release year of the cultivars and their protein content (Fig. 1). The protein content was lower in varieties appearing before the 2000s, and lines with higher protein content could only be detected within varieties bred thereafter. This indicated a renewed effort to increase protein content during the selection processes of the years 2000s.

We estimated the population structure based on SNP marker analysis and principal component analysis supported the NJ tree, which divided the collection into three main groups associated with ancestral origin (Fig. 3). The statistical comparison of the protein and quality attributes of the three main groups (Table 3) suggest that the protein composition of group 3 (blue) landraces could have been the basis of their selection for wheat breeding, and this is in agreement with the population structure analysis, as well. The efficiency of selection was characterized by the p-values (significance level comparing the landraces and modern cultivars in group 3). Significant differences were observed between landraces and modern cultivars for each but one (namely γ gliadin) parameter. The variation in the protein composition resulted in higher glutenin content, lower gliadin content, increased glutenin-to-gliadin ratio, and UPP%, improved Zeleny index value even better water absorption and as a side effect, the amounts of α/β and ω5 gliadins decreased. Therefore, modern cultivars have been developed with improved functionality and less harmful-component content [12,36].

4.2. The year effect modifies the quality parameters differentially in landraces and modern wheat cultivars

The amounts of precipitation in the 2017–2018, 2019–2020 and 2020–2021 harvest years were 596 mm, 545 mm and 393 mm, respectively. Furthermore, the precipitation distribution was unfavorable during the 2019–2020 and 2020–2021 growing seasons relative to 2017–2018. In the 2019–2020 period precipitation totaled 153 mm between March and June, while 109 mm was recorded in 2020–2021 and 274 mm during the 2017–2018 harvest year. This indicates a 44 % reduction in 2019–2020 and 60 % in 2020–2021 compared to 2017–2018, which was coupled with elevated global radiation values (Supplementary Table 2). The resulting abiotic stress was better tolerated by landraces than by modern cultivars as indicated by the quality parameters. Traditional wheat landraces often retain a unique ability to withstand both biotic and abiotic stresses, contributing to consistent yields and moderate productivity in low-input agricultural systems [72,73]. The *Rht-B1b* and *Rht-D1b* loci, which cause dwarfism and are present in most modern cultivars, are among the most important factors in the yield increase achieved during the 'Green Revolution'. Short-straw wheat varieties tend to underperform compared to taller ones when faced with drought or elevated temperatures [74,75]. Field experiments revealed that semi-dwarf wheat cultivars develop shorter and shallower root systems compared to tall cultivars. Under drought conditions the amount of available water and nitrogen decreases in the upper soil layers, which can significantly impact both yield and quality parameters [76]. Our experiments confirmed that in the two dry years (2020, 2021), the landraces were able to maintain higher protein and gluten content with average yields. In the 2018 year showing normal precipitation the modern cultivars performed better compared to the closely related SNP-3L landrace group in terms of gluten content and Zeleny test results (Figs. 3 and 4). In summary, the unique protein subunit composition of landraces from the SNP-1 group and their genetic distance from modern cultivars makes them a desirable breeding material providing a high bread-making quality stable under abiotic stress conditions (Fig. 3).

4.3. Population structure of the wheat collection by using protein markers

The MALDI-TOF separation results of HMW GS, LMW GS, gliadin, and soluble proteins, along with the RP-HPLC separation results of gliadins (Supplementary Table 4, Figs. 4, Fig. 5, Fig. 6, Figs. 7, Fig. 8, Fig. 9), have been used as protein markers for characterizing the population structure of the wheat collection in this study.

The limited number of HMW glutenin subunit markers divided the population into two groups (Fig. 4). Comparing the data of the protein composition led to differences in the functional parameters in both groups: modern cultivars exhibited superior functional properties compared to landraces, with the differences being more pronounced in group 2 (blue) than in group 1 (red) (Supplementary Table 7). The HMW glutenin allelic composition gives a direct explanation for this difference: each of the modern cultivars in group 1 ($n = 28$) contains the *Glu1Da* (2 + 12), while 38 of the 39 cultivars in group 2 has the *Glu1Dd* (5 + 10) allele (Supplementary Table 4). As it was proven by *in vitro* incorporation studies by Anderson et al. [77] the extra cysteine in the N-terminal region of the subunit *Glu1D5* is responsible for the superior contribution of this subunit determining dough strength and stability through its effect on the shift of the size distribution of the polymeric glutenin to higher values.

Using the NJ clustering method, the LMW glutenin subunit markers divided the accessions into two major groups (Fig. 5, Supplementary Table 10-LMW). Modern cultivars grown in 2018 contain significantly less proteins, gliadins, α/β gliadins, and significantly more glutenins, larger glutenin-to-gliadin ratio, and significantly higher UPP% referring

to the improvement in functionality achieved by breeding [12,36,47].

Combining the MALDI-TOF MS HMW and LMW glutenin subunit markers the NJ clustering method divided the population into five groups (Fig. 6a). This distribution separates the cultivars containing the Glu1Da (2 + 12) allele from those with the Glu1Dd (5 + 10) allele. Consequently, modern cultivars with better quality are separated in Group 3 (green) from those with poorer quality in Group 4 (blue) (Supplementary Table 7). The complex analysis of the MALDI-TOF MS HMW and LMW glutenin subunit markers can support breeders in selecting crossing partners to improve quality parameters. The methods of Eagles et al. [78] and the Protein Scoring System (PSS) developed by Békés et al. [79] describe the contribution of HMW-GS and LMW-GS alleles both individually and in pair-wise interactions on dough strength and extensibility. The results point to the significant impact of the allelic interactions on the quality. Because of the large contribution of the allele–allele interactions, the different allelic combinations, rather than the individual glutenin alleles need to be controlled by breeders when selecting wheat for a desired performance [79].

The gliadins, as constituents of the gluten, are essential for maintaining dough elasticity and stickiness and proper dough rising [80,81]. Since gliadins have a significant impact on the baking quality, their markers could help to identify favorable lines. In this study, we compared the baking quality parameters of groups identified by the NJ clustering method based on MALDI-TOF MS Gliadin protein, RP-HPLC Gliadin protein, and the combined use of MALDI-TOF MS Gliadin protein and RP-HPLC Gliadin protein markers (Figs. 8–10). The majority of modern varieties (approx. 80 %) were grouped identically in all three classifications, which was also observed for landraces. Based on these observations it could be concluded that 7 modern cultivars in Group 3 (blue) seem to be superior among the 65 cultivars of the whole population with respect to functional properties (Fig. 10). The population structure of wheat accessions based on salt-soluble protein markers (Fig. 11) clearly indicated that salt-soluble proteins – not being involved directly in determining dough rheology parameters – also play a role in the overall quality of wheat lines (Supplementary Table 10). It was shown before, that the water absorption is mostly dependent on the protein and pentosane content of the flour and starch damage, however, certain soluble proteins also alter it [82]. Consequently, the effect of the lower protein content in modern cultivars could be compensated by the altered soluble protein composition and the pentosan level. Soluble protein composition of landraces included in Group 1 (red) was significantly different from landraces located in Group 2 (blue) The Group 2 (blue) landraces containing 47 individual salt-soluble protein markers represent good sources of candidates in breeding programs aiming for quality improvement.

To develop premium quality wheat and satisfy the customers' diverse demands, optimally selected wheat storage protein allele composition is a necessary but not a sufficient condition. Instead, a complex integrated approach is needed. In this approach, the effect of different genes and gene products is not investigated individually, but in a process where the interactive effects – inhibitions and synergisms – could also be realized, quantified, and finally utilized. It also means that in the case of bread-making, next to keeping the emphasis on gluten proteins, the contribution of non-gluten components, such as soluble proteins, non-starchy carbohydrates, lipids, and starch has to be taken into account [83].

The hierarchical co-clustering analysis was used to demonstrate the similarity between the groups created by NJ clustering (Fig. 12). In this analysis, most of the modern cultivars were grouped together, which indicates, that each type of protein marker supplemented by SNP markers effectively distinguishes the different lines.

Combining the analyzed gluten protein markers with the NJ clustering method divided the population into three groups (Fig. 13) where modern cultivars were distributed within the green and the blue groups (Group 2 and 3). The landraces exhibiting the highest Zeleny index rankings also showed the highest protein and gluten contents within the analyzed population (Fig. 2d–Supplementary Table 9). Based on high molecular weight (HMW) allelic composition, two of these landraces had a Payne score of 10 (RMF164, RMF165), six had a score of 8, and one landrace scored 6 (RMF117) (Supplementary Table 4). Based on the evaluated quality attributes, several landraces were identified as having superior bread-baking potential compared to modern varieties, including RMF061 (DUNAV) and RMF117 (TISZAORS). Notably, RMF117 exhibited a Payne score of 6, while RMF061 achieved a score of 8. When integrating allelic composition data with Near-Infrared Spectroscopy (NIR)-based quality parameters, RMF107 (ECS) and RMF110 (MARCALTO) emerged as particularly promising. Both landraces achieved a Payne score of 8 and ranked among the top ten for Zeleny sedimentation index, gluten content, and protein content. Additionally, these accessions belong to SNP-1 group, the most genetically distinct cluster from modern cultivars, suggesting ancient genetic origins and unique quality potential. RMF165 (CLUJ-722) further exemplifies the exceptional quality potential of landraces. With a Payne score of 10—the highest among evaluated accessions—and ranking within the top ten for Zeleny sedimentation index, gluten content, and protein content, RMF165 stands out as an elite candidate for bread-making applications. Although it belongs to the SNP-2 group, also comprises exclusively landraces and no modern cultivars. These findings suggest that the inclusion of such landraces in breeding programs may contribute to the development of wheat cultivars with enhanced technological properties.

The combined analysis incorporating both gluten protein markers and salt-soluble protein markers is presented in Supplementary Fig. 13. Integrating the analyzed protein markers with the Neighbor-Joining (NJ) clustering method divided the population into six distinct groups, with modern cultivars clustering within the blue and purple groups (Groups 5 and 6). Notably, the ALL-5 group includes four landraces selected based on their high Zeleny index values, along with the single modern cultivar. The ALL-3 group contained one landrace (RMF61), while the ALL-6 group included another landrace (RMF117). The remaining three landraces were classified into the ALL-2 group. These results are consistent with the patterns observed for salt-soluble protein markers (Fig. 11), suggesting that salt-soluble seed proteins exhibit weak correlation with quality parameters—likely due to their limited contribution to overall grain quality. Consequently, they appear to be unsuitable as reliable markers for quality assessment. Similar patterns were observed in the PCA analyses, where the first (PC1), second (PC2), and third (PC3) component axes provided limited explanation at the regional level (Supplementary Fig. 13b and 13c), but it was informative when the chronological origin of the collection was considered.

Among the various protein markers evaluated, the MALDI-TOF MS LMW, MALDI-TOF MS Gliadin, RP-HPLC Gliadin, the combined MALDI-TOF MS Gliadin + RP-HPLC Gliadin, and the integrated MALDI-TOF MS LMW + HMW + Gliadin + RP-HPLC Gliadin markers demonstrated the strongest relevance with quality parameters, indicating their superior effectiveness for quality prediction. The landraces included in the present study showed significant diversity and their comparison with modern cultivars (Supplementary Table 10-ALL) revealed changes in the protein subunit composition defined by breeding purposes. Our study demonstrated the potential of landraces to utilize useful protein subunit diversity that is currently missing in breeding programs.

5. Conclusion

Landrace collections are acknowledged as valuable sources for breeding endeavors to improve the essential agricultural traits of wheat. Their substantial allelic diversity positions them as appealing candidates for augmenting the genetic gain of wheat. The characterization of the Central European bread wheat landrace collection revealed a higher average protein content for the landraces than for the modern varieties; however, the modern varieties were characterized by higher glutenin and lower gliadin content. Gluten protein markers effectively identified landraces with superior quality traits. In contrast, salt-soluble protein

markers fragmented population structure and showed weak association with key quality parameters, limiting their utility for selection or classification.This is probably the result of the selection for favorable baking quality during breeding, which selection also greatly narrowed the richness of protein diversity.

Our results revealed the diversity of the Central European landrace variety collection at the protein level, which may pave the way to the launch of new, targeted breeding programs.

CRediT authorship contribution statement

Gyöngyvér Gell: Writing – original draft, Visualization, Formal analysis, Data curation, Conceptualization. **András Cseh:** Writing – original draft, Visualization, Funding acquisition, Formal analysis, Conceptualization. **Christakis George Florides:** Supervision, Methodology, Formal analysis, Data curation. **Marianna Rakszegi:** Writing – review & editing, Formal analysis, Data curation, Christakis George Florides, Methodology, Formal analysis, Data curation. **Zsófia Birinyi:** Visualization, Methodology, Data curation. **Dalma Nagy-Réder:** Visualization, Formal analysis, Data curation. **Ádám Horváth:** Visualization. **Ádám D. Horváth:** Visualization, Project administration. **Ferenc Békés:** Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal analysis, Data curation, Conceptualization. **Ildikó Karsai:** Writing – review & editing, Validation, Supervision, Conceptualization.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT-4o model (OpenAI, 2025) in order to improve the readability and language of the manuscript. After using this tool, the authors reviewed and edited the content as needed and takes full responsibility for the content of the published article.

Declaration of competing interest

This declaration of interest statement must be uploaded as a separate file in the submission system. Please ensure that this is done.

Acknowledgements

This research was funded by the National Research, Development and Innovation Fund (TKP2021-NKTA-06, NKFI-K-142934, NKFI-FK-142170, Hungarian National Laboratory Program, Agribiotechnology project (grant number RRF-2.3.1-21-2022-00007) and contributed to Eötvös Loránd Research Network, project No. SA-25/2021. AC is supported by János Bolyai Research Scholarships of the Hungarian Academy of Sciences (BO/00416/23/4).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jafr.2025.102403.

Data availability

Data will be made available on request.

References

[1] F. Balfourier, S. Bouchet, S. Robert, R. De Oliveira, H. Rimbert, J. Kitt, F. Choulet, , International Wheat Genome Sequencing Consortium, BreedWheat Consortium, E. Paux, Worldwide phylogeography and history of wheat genetic diversity, Sci. Adv. 5 (2019) eaav0536, https://doi.org/10.1126/sciadv.aav0536.

[2] C. Sansaloni, J. Franco, B. Santos, L. Percival-Alwyn, S. Singh, C. Petroli, J. Campos, K. Dreher, T. Payne, D. Marshall, B. Kilian, I. Milne, S. Raubach, P. Shaw, G. Stephen, J. Carling, C.S. Pierre, J. Burgueño, J. Crosa, H. Li, C. Guzman, Z. Kehel, A. Amri, A. Kilian, P. Wenzl, C. Uauy, M. Banziger, M. Caccamo, K. Pixley, Diversity analysis of 80,000 wheat accessions reveals consequences and opportunities of selection footprints, Nat. Commun. 11 (2020) 4572, https://doi.org/10.1038/s41467-020-18404-w.

[3] A. Cseh, P. Poczai, T. Kiss, K. Balla, Z. Berki, Á. Horváth, C. Kuti, I. Karsai, Exploring the legacy of Central European historical winter wheat landraces, Sci. Rep. 11 (2021) 23915, https://doi.org/10.1038/s41598-021-03261-4.

[4] L. Pascual, M. Ruiz, M. López-Fernández, H. Pérez-Pena, E. Benavente, J. F. Vázquez, C. Sansaloni, P. Giraldo, Genomic analysis of Spanish wheat landraces reveals their variability and potential for breeding, BMC Genom. 21 (2020) 122, https://doi.org/10.1186/s12864-020-6536-x.

[5] V. Mohler, G. Schweizer, L. Hartl, Bread-making quality and grain yield in German winter wheat. II. Marker-trait associations, 61. Tagung Der Vereinigung der Pfanzenzuechter und Saatgutkaufeute Oesterreichs 2010, 2011, pp. 29–31.

[6] J.D. Sherman, D. Nash, S.P. Lanninng, J.M. Martin, N.K. Blake, C.F. Morris, L. E. Talbert, Genetics and end-use quality differences between modern and historical spring wheat, Crop Sci. 54 (2014) 1972–1980, https://doi.org/10.2135/cropsci2013.11.0749.

[7] N.W. Simmonds, The relation between yield and protein in cereal grain, J. Sci. Food Agric. 67 (1995) 309–315, https://doi.org/10.1002/jsfa.2740670306.

[8] F.X. Oury, C. Godin, Yield and grain protein concentration in bread wheat: how to use the negative relationship between the two characters to identify favourable genotypes? Euphytica 157 (2007) 45–57, https://doi.org/10.1007/s10681-007-9395-5.

[9] H.F. Gomez-Becerra, H. Erdem, A. Yazici, Y. Tutus, B. Torun, L. Ozturk, I. Cakmak, Grain concentrations of protein and mineral nutrients in a large collection of spelt wheat grown under different environments, J. Cereal. Sci. 52 (2010) 342349, https://doi.org/10.1016/j.jcs.2010.05.003.

[10] M. Oberforster, M. Werteker, Inverse and non-inverse relations between grain yield and quality in the Austrian cultivars of wheat, barley and rye, Tagung der Vereinigung der Pflanzenzuechter und Saatgutkaufleute Oesterreichs 61 (2011) 2010, 9–17.

[11] D. Pronin, A. Borner, H. Weber, K.A. Scherf, Wheat (Triticum aestivum L.) breeding from 1891 to 2010 contributed to increasing yield and glutenin contents but decreasing protein and gliadin contents, J. Agric. Food Chem. 68 (2020) 13247–13256, https://doi.org/10.1021/acs.jafc.0c02815.

[12] Q. Riaz, D. Pleming, F. Békés, C.G. Florides, R. Eastwood, C. Blanchard, Changes in protein content, composition and dough quality: outcomes of Australian wheat breeding on varieties released from 1860 to 2015, Cereal Res. Commun. (2024) 1–15, https://doi.org/10.1007/s42976-024-00608-1.

[13] M. Sanchez-Garcia, F. Álvaro, A. Peremarti, J.A. Martín-Sánchez, C. Royo, Changes in bread-making quality attributes of bread wheat varieties cultivated in Spain during the 20th century, Eur. J. Agron. 63 (2015) 79–88, https://doi.org/10.1016/j.eja.2014.11.006.

[14] U. Goutam, R. Tiwari, R.K. Gupta, S. Kukreja, A. Chaudhury, Allelic variations of functional markers for high molecular weight glutenin genes in Indian wheat (Triticum aestivum L.) cultivars and their correlation with bread loaf volume, Indian J. Plant Physiol. 20 (2015) 97–102, https://doi.org/10.1007/s40502-015-0141-z.

[15] H. Fufa, P. Baenziger, B. Beecher, R. Graybosch, K. Eskridge, L. Nelson, Genetic improvement trends in agronomic performances and end-use quality characteristics among hard red winter wheat cultivars in Nebraska, Euphytica 144 (2005) 187–198, https://doi.org/10.1007/s10681-005-5811-x.

[16] M. Malalgoda, J.B. Ohm, S. Meinhardt, S. Simsek, Association between gluten protein composition and breadmaking quality characteristics in historical and modern spring wheat, Cereal Chem. 95 (2018) 226–238, https://doi.org/10.1002/cche.10014.

[17] M.A. De Santis, M.M. Giuliani, L. Giuzio, P. De Vita, A. Lovegrove, P.R. Shewry, Z. Flagella, Differences in gluten protein composition between old and modern durum wheat genotypes in relation to 20th century breeding in Italy, Eur. J. Agron. 87 (2017) 19–29, https://doi.org/10.1016/j.eja.2017.04.003.

[18] A. Izadi-Darbandi, B. Yazdi-Samadi, A.A. Shanejat-boushehri, Allelic variations in Glu-1 and Glu-3 loci of historical and modern Iranian bread wheat (Triticum aestivum L.) cultivars, J. Genet. 89 (2010) 193–199, https://doi.org/10.1007/s12041-010-0025-y.

[19] E. Gregová, V. Tisová, J. Kraic, Genetic variability at the Glu-1 loci in old and modern wheats Triticum aestivum L. cultivated in Slovakia, Genet. Resour. Crop Evol. 44 (1997) 301–306.

[20] E. Gregová, J. Hermuth, J. Kraic, L. Dotlačil, Protein heterogeneity in European wheat landraces and obsolete cultivars, Genet. Resour. Crop Evol. 46 (1999) 521–528.

[21] E. Gregová, J. Hermuth, J. Kraic, L. Dotlačil, Protein heterogeneity in European wheat landraces and obsolete cultivars: additional information II, Genet. Resour. Crop Evol. 53 (2006) 867–871.

[22] A. Juhász, O.R. Larroque, L. Tamás, S.L.K. Hsam, F.J. Zeller, F. Békés, Z. Bedő, Bánkúti 1201 - an old Hungarian wheat variety with special storage protein composition, Theor. Appl. Genet. 107 (2003) 697–704, https://doi.org/10.1007/s00122-003-1292-2.

[23] X. An, Q. Li, Y. Yan, Y. Xiao, S.L.K. Hsam, F.J. Zeller, Genetic diversity of European spelt wheat (Triticum aestivum ssp spelta L. em. Thell.) revealed by glutenin subunit variations at the Glu-1 and Glu-3 loci, Euphytica 146 (2005) 193–201, https://doi.org/10.1007/s10681-005-9002-6.

[24] W. Zheng, Y. Peng, J. Ma, R. Appels, D. Sun, W. Ma, High frequency of abnormal high molecular weight glutenin alleles in Chinese wheat landraces of the Yangtze-

River region, J. Cereal. Sci. 54 (2011) 401–408, https://doi.org/10.1016/j.jcs.2011.08.004.
[25] S. Lee, Y.-M. Choi, M.-C. Lee, D.Y. Hyun, S. Oh, Y. Jung, Geographical comparison of genetic diversity in Asian landrace wheat (Triticum aestivum L.) germplasm based on high-molecular- weight glutenin subunits, Genet. Resour. Crop Evol. 65 (2018) 1591–1602, https://doi.org/10.1007/s10722-018-0633-6.
[26] S. Goel, M. Yadav, K. Singh, R.S. Jaat, N.K. Singh, Exploring diverse wheat germplasm for novel alleles in HMW-GS for bread quality improvement, J. Food Sci. Technol. Mys. 55 (2018) 3257–3262, https://doi.org/10.1007/s13197-018-3259-y.
[27] H. Nakamura, Genetic diversity of high-molecular-weight glutenin subunit compositions in landraces of hexaploid wheat from Japan, Euphytica 120 (2001) 227–234.
[28] S.W. Cho, S.K. Roy, J.-B. Chun, K. Cho, C.S. Park, Overexpression of the Bx7 high molecular weight glutenin subunit on the Glu-B1 locus in a Korean wheat landrace, Plant Biotechnol. Rep. 11 (2017) 97–105, https://doi.org/10.1007/s11816-017-0434-y.
[29] S.W. Cho, K. Cho, G. Bang, C.S. Park, Molecular profiling of a y-type high molecular weight glutenin subunit at Glu-D1 locus from a North Korean landrace wheat Triticum aestivum L, Plant Biotechnol. Rep. 12 (2018) 139–148, https://doi.org/10.1007/s11816-018-0479-6.
[30] M. López-Fernández, L. Pascual, I. Faci, M. Fernández, M. Ruiz, E. Benavente, P. Giraldo, Exploring the end-use quality potential of a collection of Spanish bread wheat landraces, Plants 10 (2021) 620, https://doi.org/10.3390/plants10040620.
[31] S. Dai, D. Xu, Y. Yan, Z. Wen, J. Zhang, H. Chen, Z. Lu, H. Li, H. Cong, Y. Wei, Y. Zheng, Z. Yan, Characterization of high- and low-molecular-weight glutenin subunits from Chinese Xinjiang wheat landraces and historical varieties, Food Sci. Technol. 57 (2020) 3823–3835, https://doi.org/10.1007/s13197-020-04414-5.
[32] G. Branlard, P. Giraldo, Z. He, G. Igrejas, T.M. Ikeda, M. Janni, M.Y. Labuschagne, D. Wang, B. Wentzel, K. Zhang, Contribution of genetic resources to grain storage protein composition and wheat quality, in: G. Igrejas, T. Ikeda, C. Guzmán (Eds.), Wheat Quality for Improving Processing and Human Health, Springer, Cham, 2020, pp. 39–72, https://doi.org/10.1007/978-3-030-34163-3_4.
[33] E.V. Metakovsky, V. Melnik, M. Rodriguez-Quijano, V. Upelniek, J.M. Carrillo, A catalog of gliadin alleles: polymorphism of 20th-century common wheat germplasm, Crop J. 6 (2018) 628–641, https://doi.org/10.1016/j.cj.2018.02.003.
[34] M. Ruiz, E.V. Metakovsky, M. Rodriguez-Quijano, J.F. Vazquez, J.M. Carrillo, Assessment of storage protein variation in relation to some morphological characters in a sample of Spanish landraces of common wheat Triticum aestivum L. ssp aestivum, Genet. Resour. Crop Evol. 49 (2002) 371–382, https://doi.org/10.1023/A:1020626430815.
[35] M. Ruiz, E.V. Metakovsky, M. Rodriguez-Quijano, J.F. Vazquez, J.M. Carrillo, Polymorphism, variation and genetic identity of Spanish common wheat germplasm based on gliadin alleles, Field Crops Res. 79 (2002) 185–196, https://doi.org/10.1016/S0378-4290(02)00139-9.
[36] Q. Riaz, Changes in Australian Wheat Quality over the Last 150 Years, Doctoral Thesis 2020.
[37] D. Pronin, A. Börner, K.A. Scherf, Old and modern wheat (Triticum aestivum L.) cultivars and their potential to elicit celiac disease, Food Chem. 339 (2021) 127952, https://doi.org/10.1016/j.foodchem.2020.127952.
[38] S. Geisslitz, C.F.H. Longin, K.A. Scherf, P. Koehler, Comparative study on gluten protein composition of ancient (Einkorn, emmer and spelt) and modern wheat species (durum and common wheat), Foods 8 (2019) 409, https://doi.org/10.3390/foods8090409.
[39] E. Metakovsky, V.A. Melnik, L. Pascual, C.V. Wrigley, Gliadin genotypes worldwide for spring wheats (Triticum aestivum L.) 1. Genetic diversity and grain-quality gliadin alleles during the 20th century, J. Cereal. Sci. 87 (2019) 172–177, https://doi.org/10.1016/j.jcs.2019.03.008.
[40] E. Metakovsky, V.A. Melnik, L. Pascual, C.V. Wrigley, Gliadin genotypes worldwide for spring wheats (Triticum aestivum L.) 1. Genetic diversity Gliadin genotypes worldwide for spring wheats (Triticum aestivum L.) 2. Strong differentiation of polymorphism between countries and regions of origin, J. Cereal. Sci. 87 (2019) 311–317, https://doi.org/10.1016/j.jcs.2019.04.015.
[41] M. Schopf, K.A. Scherf, Wheat cultivar and species influence variability of gluten ELISA analyses based on polyclonal and monoclonal antibodies R5 and G12, J. Cereal. Sci. 83 (2018) 32–41, https://doi.org/10.1016/j.jcs.2018.07.005.
[42] Y. Li, R. Xin, D. Zhang, S. Li, Molecular characterization of α-gliadin genes from common wheat cultivar Zhengmai 004 and their role in quality and celiac disease, Crop J. 2 (2014) 10–21, https://doi.org/10.1016/j.cj.2013.11.003.
[43] M. Ribeiro, M. Rodriguez-Quijano, F.M. Nunes, J.M. Carrillo, G. Branlard, G. Igrejas, New insights into wheat toxicity: breeding did not seem to contribute to a prevalence of potential celiac disease's immunostimulatory epitopes, Food Chem. 213 (2016) 8–18, https://doi.org/10.1016/j.foodchem.2016.06.043.
[44] M.M.G. Malalgoda, Investigation of Protein Composition in Historic and Modern Hard Red Spring Wheat Cultivars, Nort Dakota Fargo, 2016. PhD thesis.
[45] P.R. Shewry, Do ancient types of wheat have health benefits compared with modern bread wheat? J. Cereal. Sci. 79 (2018) 469–476, https://doi.org/10.1016/j.jcs.2017.11.010.
[46] D.A.I. Suter, F. Békés, Who is to blame for the increasing prevalence of dietary sensitivity to wheat? Cereal Res. Commun. 49 (2021) 1–19, https://doi.org/10.1007/s42976-020-00114-0.
[47] F. Brouns, S. Geisslitz, C. Guzman, T.M. Ikeda, A. Arzani, G. Latella, S. Simsek, M. Colomba, A. Gregorini, V. Zevallos, V. Lullien-Pellerin, D. Jonkers, P.R. Shewry, Do ancient wheats contain less gluten than modern bread wheat, in favour of better health? Nutr. Bull. 47 (2022) 157–167, https://doi.org/10.1111/nbu.12551.
[48] F. Brouns, S. Geisslitz, P.S. Shewry, Is bread bad for health? J. Cereal. Sci. 105 (2022) 103447 https://doi.org/10.1016/j.jcs.2022.103447.
[49] International Association for Cereal Science and Technology ICC 202, Procedure for near infrared (NIR) Reflec-Tance analysis of ground wheat and milled wheat products, Int. Asso. Cereal Sci. Technol. (1995). Vienna, Austria.
[50] R.B. Gupta, K. Khan, F. MacRitchie, Biochemical basis of flour properties in bread wheats. I. Effects of variation in the quantity and size distribution of polymeric protein, Cereal Sci 18 (1993) 23–41, https://doi.org/10.1006/jcrs.1993.1031.
[51] I. Batey, R. Gupta, F. MacRitchie, Use of size-exclusion high-performance liquid chromatography in the study of wheat flour proteins: an improved chromatographic procedure, Cereal Chem. 68 (1991) 207–209.
[52] O.R. Larroque, F. Békés, Rapid molecular size distribution analysis of wheat storage protein, Cereal Chem. 77 (2000) 451–454, https://doi.org/10.1094/CCHEM.2000.77.4.451.
[53] O.R. Larroque, F. Békés, C.W. Wrigley, W.G. Rathmell, Analysis of gluten proteins in grain and flour blends by RP-HPLC, in: P.R. Shewry, A.S. Tatham (Eds.), Wheat Gluten Proceedings 7th International Workshop, Royal Society of Chemistry, 2000, pp. 136–139, https://doi.org/10.1039/9781847552372-00136.
[54] C.G. Florides, Predicting Allergenicity of Wheat Genotypes, PhD Thesis, Charles Sturt University, Wagga Wagga, Australia, 2019, 2019.
[55] J.H. Ward Jr., Hierarchical grouping to optimize an objective function, JASA. 58 (1963) 236–244.
[56] J. Wang, Z. Zhang, GAPIT version 3: boosting power and accuracy for genomic association and prediction, Genom. Proteom. Bioinform. bioRxiv. (2021), https://doi.org/10.1016/j.gpb.2021.08.005.
[57] N. Saitou, M. Nei, The neighbor-joining method: a new method for reconstructing phylogenetic trees, Mol. Biol. Evol. 4 (1987) 406–425, https://doi.org/10.1093/oxfordjournals.molbev.a040454.
[58] C. Sievert, Interactive Web-based Data Visualization with R, Plotly and Shiny, Chapman and Hall/CRC, 2020 isbn: 9781138331457, https://plotly-r.com.
[59] T. Galili, Dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering, Bioinformatics 31 (2015) 3718–3720, https://doi.org/10.1093/bioinformatics/btv428.
[60] P.I. Payne, M.A. Nightingale, A.F. Krattiger, L.M. Holt, The relationship between HMW glutenin subunit composition and the breadmaking quality of british-grown wheat varieties, J. Sci. Food Agric. 40 (1987) 51–65.
[61] M.O. Winfield, A.M. Allen, A.J. Burridge, G.L. Barker, H.R. Benbow, P. A. Wilkinson, J. Coghill, C. Waterfall, A. Davassi, G. Scopes, A. Pirani, T. Webster, F. Brew, C. Bloor, J. King, C. West, S. Griffiths, I. King, A.R. Bentley, K.J. Edwards, High-density SNP genotyping array for hexaploid wheat and its secondary and tertiary gene pool, Plant Biotechnol. J. 14 (2016) 1195–1206, https://doi.org/10.1111/pbi.12485.
[62] C. Schuster, J. Huen, K.A. Scherf, Prediction of wheat gluten composition via near-infrared spectroscopy, Curr. Res. Food Sci. 6 (2023) 100471, https://doi.org/10.1016/j.crfs.2023.100471.
[63] I.K. Thomsen, M.F. Samson, M. Carcea, V. Narducci, The influence of long-term inputs of catch crops and cereal straw on yield, protein composition and technological quality of a spring and a winter wheat, Int. J. Food Sci. Technol. 46 (2011) 216–220, https://doi.org/10.1111/j.1365-2621.2010.02467.x.
[64] M. Hruśková, O. Famĕra, Prediction of wheat and flour Zeleny sedimentation value using NIR technique, Czech J. Food Sci. 21 (2003) 91–96, https://doi.org/10.17221/3482-CJFS.
[65] M.M.R. Malalgoda, J.B. Ohm, S. Meinhardt, S. Chao, S. Simsek, Cluster analysis of historical and modern hard red spring wheat cultivars based on parentage and HPLC analysis of gluten-forming protein, Cereal Chem. 94 (2017) 560–567, https://doi.org/10.1094/CCHEM-08-16-0223-R.
[66] C.W. Wrigley, F. Békés, Processing Quality requirements for wheat and other ceeral grains, in: R.L. Benech-Arnold, A. Rodolfo (Eds.), Seed Physiology: Applications to Agriculture Haworth, Press, Inc., Binghamton, USA, 2003, pp. 143–187.
[67] C.W. Wrigley, F. Békés, W. Bushuk, Chapter 1. Gluten: a balance of gliadin and glutenin, in: C.W. Wrigley, F. Békés, W. Bushuk (Eds.), Gliadin and Glutenin. the Unique Balance of Wheat Quality, AACCI Press, St Paul, Min., USA, 2006, pp. 3–33.
[68] T.J. Tsilo, J.B. Ohm, A. Gary, G.A. Hareland, J.A. Anderson, Association of size-exclusion HPLC of endosperm proteins with dough mixing and breadmaking characteristics in a recombinant inbred population of hard red spring wheat, Cereal Chem. 87 (2010) 104–111, https://doi.org/10.1094/CCHEM-87-2-0104.
[69] S. Simsek, J.B. Ohm, V. Cariou, M. Mergoum, Effect of flour polymeric proteins on dough thermal properties and breadmaking characteristics for hard red spring wheat genotypes, J. Cereal. Sci. 68 (2016) 164–171, https://doi.org/10.1016/j.jcs.2016.01.003.
[70] H.C. van den Broeck, H.C. de Jong, E.M.J. Salentijn, L. Dekking, D. Bosch, R. J. Hamer, L.J.W.J. Gilissen, I.M. van der Meer, M.J.M. Smulders, Presence of celiac disease epitopes in modern and old hexaploid wheat varieties: wheat breeding may have contributed to increased prevalence of celiac disease, Theor. Appl. Genet. 121 (2010) 1527–1539, https://doi.org/10.1007/s00122-010-1408-4.
[71] M. de Lorgeril, P. Salen, Gluten and wheat intolerance today: are modern wheat strains involved? Int. J. Food Sci. Nutr. 65 (2014) 577–581, https://doi.org/10.3109/09637486.2014.886185.
[72] A.C. Zeven, Landraces: a review of definitions and classifications, Euphytica 104 (1998) 127–139.
[73] M.S. Lopes, I. El-Basyoni, P.S. Baenziger, S. Singh, C. Royo, K. Ozbek, H. Aktas, E. Ozer, F. Ozdemir, A. Manickavelu, T. Ban, P. Vikram, Exploiting genetic diversity from landraces in wheat breeding for adaptation to climate change, J. Exp. Bot. 66 (2015) 3477–3486, https://doi.org/10.1093/jxb/erv122.

[74] J.D. Butler, P.F. Byrne, V. Mohammadi, P.L. Chapman, S.D. Haley, Agronomic performance of *Rht* alleles in a spring wheat population across a range of moisture levels, Crop Sci. 45 (2005) 939–947.
[75] A. Cseh, A. Lenykó-Thegze, D. Makai, F. Szabados, K.Á. Hamow, Z. Gulyás, T. Kiss, I. Karsai, B. Moncsek, E. Mihók, A. Sepsi, Meiotic instability and irregular chromosome pairing underpin heat-induced infertility in bread wheat carrying the Rht-B1b or Rht-D1b green revolution genes, New Phytol. 241 (2024) 180–196, https://doi.org/10.1111/nph.19256.
[76] V.O. Sadras, P.T. Hayman, D. Rodriguez, M. Monjardino, M. Bielich, M. Unkovich, E. Wang, Interactions between water and nitrogen in Australian cropping systems: physiological, agronomic, economic, breeding and modelling perspectives, Crop Pasture Sci. 67 (2016) 1019–1053, https://doi.org/10.1071/CP16027.
[77] O.D. Anderson, F. Békés, R. D'Ovidio, Effects of specific domains of high-molecular-weight glutenin subunits' on dough properties by an in vitro assay, J. Cereal. Sci. 54 (2011) 280–287, https://doi.org/10.1016/j.jcs.2011.04.005.
[78] H.A. Eagles, G.J. Hollamby, N.N. Gororo, R.F. Eastwood, Estimation and utilisation of glutenin gene effects from the analysis of unbalanced data from wheat breeding programs, Aust. J. Agric. Res. 53 (2002) 1047–1057, https://doi.org/10.1071/AR01074.
[79] F. Békés, S. Kemény, M.K. Morell, An integrated approach to predicting end-product quality of wheat, Eur. J. Agron. 25 (2006) 155–162, https://doi.org/10.1016/j.eja.2006.04.008.
[80] G.P. Branlard, E.V. Metakovsky, Chapter 4 some gli alleles related to common wheat dough quality, in: C.W. Wrigley, F. Békés, W. Bushuk (Eds.), Gliadin and Glutenin: the Unique Balance of Wheat Quality, AACCI Press, St Paul, 2006, pp. 115–139, https://doi.org/10.1094/9781891127519.006.
[81] E.V. Metakovsky, R.A. Graybosch, Chapter 3 gliadin alleles in wheat: identification and applications, in: C.W. Wrigley, F. Békés, W. Bushuk (Eds.), Gliadin and Glutenin: the Unique Balance of Wheat Quality, AACCI Press, St Paul, 2006, pp. 85–114. ISBN: 978-1-891127-51-9.
[82] M. Rakszegi, G. Balázs, F. Békés, A. Harasztos, A. Kovács, L. Láng, Z. Bedő, S. Tömösközi, Modelling water absorption of wheat flour by taking into consideration of the soluble protein and arabinoxylan components, Cereal Res. Commun. 42 (2014) 629–639, https://doi.org/10.1556/CRC.2014.0007.
[83] S.J. Millar, J. Snape, J. Ward, P.R. Shewry, P. Belton, K. Boniface, R. Summers, Investigating Wheat Functionality Through Breeding and End Use (FQS 23) HGCA, 2008. Project Report No. 429.