Field Crops Research 284 (2022) 108585

Contents lists available at ScienceDirect

Field Crops Research

journal homepage: www.elsevier.com/locate/fcr

Effects of glutenins (*Glu-1* and *Glu-3*) allelic variation on dough properties and bread-making quality of CIMMYT bread wheat breeding lines

Carlos Guzmán [a,*], Jose Crossa [b], Suchismita Mondal [b], Velu Govindan [b], Julio Huerta [c], Leonardo Crespo-Herrera [b], Mateo Vargas [b], Ravi P. Singh [b], Maria Itria Ibba [b]

[a] Departamento de Genética, Escuela Técnica Superior de Ingeniería Agronómica y de Montes, Edificio Gregor Mendel, Campus de Rabanales, Universidad de Córdoba, CeiA3, ES-14071 Córdoba, Spain
[b] Global Wheat Program, International Maize and Wheat Improvement Center (CIMMYT), Apdo Postal 6-641, Mexico D.F., Mexico
[c] Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias (INIFAP), Campo Experimental Valle de México, Mexico

ARTICLE INFO

Keywords:
Bread wheat
Grain quality
Glutenins
Bread-making
Dough strength
Dough extensibility

ABSTRACT

Wheat dough characteristics and end-use quality are strongly influenced by the amount and specific composition of the glutenins, the major components of gluten. Such proteins are divided into high-molecular-weight glutenins, encoded by the *Glu-A1*, *Glu-B1* and *Glu-D1* loci; and low-molecular-weight glutenins, encoded by the *Glu-A3*, *Glu-B3* and *Glu-D3* loci. Allelic variation at each of these loci has been associated with changes in wheat functionality. However, most of the studies conducted so far included a relatively limited number of genotypes. Also for this reason, it is still unclear which locus contributes more to dough characteristics and how important are the interactions between the glutenin loci. To try to answer these questions, the quality data of 4623 grain samples derived from 2550 genotypes and generated across 10 years at the CIMMYT bread wheat breeding program, was used to estimate the effect of the glutenin loci and their interactions on gluten quality and bread-making potential. Gluten strength was the trait more strongly influenced by glutenin variations, with the *Glu-B1*, *Glu-D1* and *Glu-B3* loci having the greatest effect. Among the glutenin alleles, *Glu-A1a*, *Glu-A1b*, *Glu-B1al*, *Glu-B1i*, *Glu-B1f*, *Glu-D1d*, *Glu-A3b*, *Glu-A3d*, *Glu-A3f*, *Glu-B3c* and *Glu-B3d* were associated in general with greater gluten strength, good extensibility and higher bread loaf volume. Differently, alleles *Glu-A1c*, *Glu-B1a*, *Glu-B1d*, *Glu-D1a*, *Glu-A3e* and *Glu-B3j* were associated with an overall poor quality. Glutenin interactions were significantly associated with most of the analyzed quality traits even if their influence was often lower compared to the effect of the single glutenin loci. This is probably the largest study ever done on the effects of the glutenins on wheat quality. The results obtained confirm the importance of such proteins on wheat quality variation and corroborate the usefulness of determining the glutenin profile to improve the selection efficiency for wheat quality in breeding programs.

1. Introduction

Common or bread wheat (*Triticum aestivum* ssp. L., 2 n = 6x = 42, AABBDD) is one of the most important crops globally, covering around 215 million hectares (FAOSTAT, 2019). The great success of this crop is mostly due to its wide adaptability to different environments (it is grown at almost every latitude of the planet) (Feldman, 1995) and to the properties of its grains and flours, which allow wheat to be used for the production of hundreds of different products such as leavened and flat breads, noodles, cookies and cakes (Peña-Bautista et al., 2017). The quality of these products depends mainly on gluten, the protein network that is formed when flour is mechanically mixed with water and which is responsible for the development of a unique visco-elastic dough (Wrigley et al., 2006).

Gluten and dough characteristics can vary widely, especially in terms of strength and extensibility, and these variations modulate wheat suitability for the different end-uses. For example, weak and extensible doughs are preferred to produce cookies, whereas strong and non-tenacious doughs are better suited for bread-making processes in general. Changes in dough properties mostly depend on the amount and specific composition of the gluten-forming proteins. These proteins are the most abundant in wheat grain and are divided into two groups: the

* Corresponding author.
E-mail address: carlos.guzman@uco.es (C. Guzmán).

https://doi.org/10.1016/j.fcr.2022.108585
Received 23 February 2022; Received in revised form 23 May 2022; Accepted 26 May 2022
Available online 3 June 2022
0378-4290/© 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

gliadins, which constitute the gluten monomeric fraction mainly associated with dough extensibility; and the glutenins, which constitute the gluten polymeric fraction mainly responsible for dough elasticity. Based on their molecular weight, the glutenins are further divided into high and low molecular weight glutenins (HMWGs and LMWGs, respectively).

The HMWGs are encoded by the genes at the *Glu-A1, Glu-B1* and *Glu-D1* loci which are located on the long arms of chromosomes 1A, 1B, and 1D, respectively (Payne, 1987). The genes located at each of these loci, encode for one or two protein subunits which can be relatively easily characterized through a sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) analysis. Up to now, tenths of *Glu-1* alleles have been identified in the wide wheat gene pool (Mcintosh et al., 2020). The LMWGs are more abundant and complex than the HMWGs. These proteins are in fact encoded by a multigene family located at the *Glu-A3, Glu-B3* and *Glu-D3* loci, on the short arms of chromosomes 1A, 1B, and 1D, respectively, and are strongly linked with the *Gli-1* loci that encode the γ and ω-gliadins (Singh and Shepherd, 1985). Traditionally, the difficulty in separating and properly identifying the LMWGs alleles led to the conclusion that this group of proteins plays a secondary role compared to the HMWGs in the modulation of dough properties. In the last decades however, thanks to the improvements on protein separation techniques (Singh et al., 1991) and the use of molecular tools (Dreisigacker et al., 2020; Ibba et al., 2017a; Liu et al., 2010), the identification of LMWGs alleles has improved, allowing the proliferation of studies focused on analysing the effect of the different LMWGs alleles on processing and end-use quality. The effects of HMWGs on gluten and product quality have also been the topic of numerous studies (see Wrigley et al., 2006, for a review of the topic). These studies have used different sets of wheat materials and approaches to assess glutenins effects on quality traits and extend from the characterization of a set of cultivars from a specific region for both phenotypic traits and glutenins (Branlard et al., 2001; Hernández-Espinosa et al., 2019; Peltonen et al., 1993) to the development of sets of recombinant-inbred lines (RILs) and near-isogenic lines (NILs) populations varying in specific glutenins (Bonafede et al., 2015; Carrillo et al., 1990) or the generation of transgenic lines with special glutenins combinations (Blechl et al., 2007).

All these studies have led to the general agreement that both the HMWGs and LMWGs polymorphisms influence dough and end-use quality. Questions such as which alleles are the best contributors of each trait remain poorly answered because there are substantial discrepancies among the vast number of published studies. Nevertheless, other questions such as which locus contributes more to each dough or product characteristic and how important are the interactions across the different glutenin loci to define the dough traits, are far from being solved. These are fundamental issues for breeding programs focused on the development of new productive wheat varieties with a quality profile suitable for the end-use requirements of the target regions. The lack of a consistent solution to these points could be partially attributed to the use of a limited number of genotypes in the above-mentioned studies due to the high cost and time necessary to perform dough and end-use quality tests. To solve these limitations, a possibility is to use the data sets generated routinely in breeding programs when selection for quality traits is done during each breeding cycle. This allows to assemble data from a large number of genotypes that can be used to estimate the individual effects of each glutenin gene and alleles but also the effect of the interactions among the different glutenin loci (Eagles et al., 2002).

The International Maize and Wheat Improvement Center (CIMMYT) is the global leader in publicly-funded wheat breeding and research. CIMMYT runs one of the largest wheat breeding programs aimed to develop high-yielding and stress tolerant wheat cultivars that produce grain with desirable processing and end-use quality. To achieve this goal, thousands of breeding lines are tested every year in the field for their agronomic performance, and in the laboratory for their grain quality traits and bread-making potential. The best few hundreds of these lines are selected for inclusion in the crossing block and are used as new parents to make crosses. In order to better design the crosses, these lines are also analysed for their glutenin composition at both the *Glu-1* and *Glu-3* loci (Guzmán et al., 2019) which allowed to develop a sufficiently large dataset that could now be exploited to estimate both the main effects and the interactions of the glutenin loci.

For these reasons, the main objective of this study was to use the vast set of grain quality data generated over ten years by the CIMMYT wheat breeding program, in order to: 1-Define the glutenin loci that more strongly affect each quality trait; 2-Determine the effect of each glutenin allele on wheat quality; 3-Establish the effect of the interactions among different glutenin loci on the analysed traits; and 4-Identify the best combinations of alleles for specific dough characteristics.

2. Materials and methods

2.1. Plant materials

A total of 2550 genotypes were used in this study for which complete quality data was available. These genotypes were advanced breeding lines (generations F_5-F_7) selected from CIMMYT spring bread wheat breeding programs planted in yield trials during the period 2011–2020. The yield trials were conducted at the Campo Experimental Norman Ernest Borlaug (CENEB) station in Ciudad Obregón, Sonora, México, under full irrigation. All trials were planted in late November and harvested at the end of April-beginning of May. A total of 300 kg N was applied, which included the pre-planting nitrogen application. Herbicides and insecticides were used as needed to keep trials free from weeds and aphids. The growing cycles in this location are generally characterized by the absence of precipitation during the wheat growing season and with maximum temperatures reaching 31–32 °C in March and April, the grain filling time.

2.2. Grain quality analysis

Most of the genotypes included in the study (2550) were analysed for diverse quality traits in two consecutives cropping cycles during the period 2011–2020. Because of this a total of 4623 grain samples were used. Around 1 Kg of harvested grain was used to perform the quality analysis of these breeding lines. Thousand kernel weight (g) and test weight (kg/hl) were obtained using the digital image system SeedCount SC5000 (Next Instruments, Australia). Grain protein content (%), hardness (PSI, %) and moisture content were determined by near-infrared spectroscopy (NIR Systems 6500, Foss Denmark) calibrated based on official AACC methods 39–10.01, 55–30.01 and 46–11.02, respectively (AACC, 2010). Since 2016 grain hardness was determined using a SKCS (Perten, Sweden). Grain samples were tempered by adding water levels for use in tempering hard, medium-hard and soft wheat before milling, according to the official AACC method 26–95.01 (AACC, 2010). All samples were milled into flour using a Brabender Quadrumat Senior mill (Germany) and experimental flour yield (%) was recorded. Flour protein (%) and moisture content (%) were determined by near-infrared spectroscopy (NIR Systems 6500, Foss Denmark), calibrated as per official AACC methods 46–11.02 and 39–11.01, respectively (AACC, 2010). Additionally, 35 g flour samples were tested in a mixograph (National Mfg. Co.) to obtain optimum dough mixing time and %Torque × min according to AACC method 54–40.02 (AACC, 2010). Gluten extensibility (alveograph L), tenacity (alveograph P), elasticity or strength (alveograph W) and tenacity/extensibility ratio (alveograph P/L) were determined according to the Alveograph manufacturer's instructions (Chopin, France), using 60 g flour samples according to AACC method 54–30.02 (AACC, 2010). The bread-making process was carried out using the direct dough method with 100 g of flour (AACC method 10–09.01). Bread loaf volume (LV) was determined by rapeseed displacement using a volume-meter. Since 2013, the amounts of water added to the mixograph, alveograph and baking were determined by near-infrared spectroscopy (Antaris FT-NR analyzer,

Thermo Fisher Scientific, USA), calibrated according to Guzmán et al. (2015), and were variable according to the flour sample properties.

2.3. Glutenins composition

Few grains of each genotype were used for the determination of glutenins composition based on SDS-PAGE as described by Maryami et al. (2020). The glutenins subunits were named following the nomenclature systems developed by Jackson et al. (1996) and Branlard et al. (2003).

2.4. Statistical analysis

The statistical analyses were performed with SAS® OnDemand for Academics (SAS Institute, Cary, NC, U.S.A.). The correlation analysis was done with PROC CORR whereas the ANOVA analysis were done with PROC GLM using in all cases flour protein as covariate. The function LSmeans in PROC GLM was used to calculate the least square means of the analysed data and to determine significant differences between the values using the Fisher's protected LSD at the $\alpha = 0.05$ significance level. The regression model was selected by using all glutenin loci, the year and flour protein content as response variables. The selection was performed with the forward (step-wise) options in PROC GLMSELECT. The main criteria for model selection were the R^2 and adjusted R^2 values.

3. Results

3.1. Quality traits variation

During the period 2011–2020, 4623 grain samples of 2550 breeding lines derived from the CIMMYT spring bread wheat breeding program, were analysed for grain quality traits and glutenin composition (Supplementary Table 1). Most of these genotypes were analysed across two consecutive years and, overall, exhibited great variability for all traits (Table 1). In all cases, both the genotype and the year had a highly significant effect on the observed trait variation (Supplementary Table 2) which mainly followed a normal distribution (data not shown). Most of the breeding lines had well filled and medium-large grains with medium to low protein content, which is typical of irrigated high-yielding environments. Grain hardness average values corresponded to those of hard grains, although some soft grain genotypes were also present (3.1% of the total grain samples evaluated). Gluten and dough properties exhibited very large variation, with traits such as alveograph W and P/L having very wide ranges of values. Bread loaf volume was also very variable, with lines exhibiting very poor to excellent bread-making quality.

When analysing the relationships among the different quality traits (Fig. 1), several significant correlations were found, including positive associations between protein content, gluten and end-use quality traits. As expected, strong correlations were identified between traits measuring the same gluten characteristics such as MIXTQ and ALVW (both indicative of gluten strength). Bread loaf volume was also significantly correlated with all the parameters indicative of gluten quality except alveograph P (gluten tenacity). Alveograph L (gluten extensibility) and SDSS (overall gluten quality) were the two parameters more highly correlated with bread-making quality (r values of 0.65 and 0.45, respectively).

3.2. Glutenin alleles frequencies

A total of 34 glutenin alleles were identified in the 2550 breeding lines. Polymorphism was detected at each glutenin locus: 3 alleles for *Glu-A1,* 7 alleles for *Glu-B1,* 2 alleles for *Glu-D1*, 6 for *Glu-A3*, 9 for *Glu-B3,* and 7 for *Glu-D3*.

At the *Glu-A1* locus, specifically, alleles *Glu-A1b* and *Glu-A1a* were the most frequent whereas the null allele (*Glu-A1c*) was found in only a

Table 1
Average and standard deviation of the quality traits analyzed in the data set across different years.

Trait		Combined	2011–12	2012–13	2013–14	2014–15	2015–16	2016–17	2017–18	2018–19	2019–20
TW (kg/hL)	Avg	81,3	82,2	81,6	81,6	80,9	82,7	79,8	80,6	80,9	81,5
	Stdev	1,4	1,1	1,1	1,0	1,1	1,0	1,3	1,2	1,1	1,1
TKW (g)	Avg	47,6	48,3	46,5	46,3	45,4	48,6	47,8	49,0	49,2	50,4
	Stdev	3,9	3,4	3,7	4,0	3,4	3,8	3,0	3,7	3,7	3,5
PSI (%)	Avg	43,2	40,4	42,8	43,9	46,2	41,5	–	–	–	–
	Stdev	4	5,6	3,7	3,6	2,5	3,8	–	–	–	–
SKCS HI	Avg	65,3	–	–	–	–	–	67,9	63,8	63,8	65,2
	Stdev	7,6	–	–	–	–	–	5,8	8,3	8,9	6,1
FYIELD (%)	Avg	69,1	68,8	69,2	71,0	70,6	71,1	67,7	67,6	66,7	66,8
	Stdev	2,6	2,4	2,2	2,2	1,8	2,1	1,8	1,8	1,9	2,1
GPRO (%)	Avg	12,3	11,3	11,8	12,0	12,3	12,2	12,7	12,7	12,6	13,0
	Stdev	0,8	0,6	0,6	0,6	0,6	0,6	0,6	0,7	0,8	0,8
FPRO (%)	Avg	10,5	9,5	10,1	10,6	10,8	10,5	10,3	10,6	10,8	11,3
	Stdev	0,7	0,5	0,5	0,6	0,6	0,6	0,5	0,6	0,6	0,8
SDSS (mL)	Avg	14,4	13,9	14,4	13,5	15,1	15,9	13,7	14,5	15,3	13,9
	Stdev	2,4	1,9	2,5	2,4	2,3	2,2	1,8	2,2	2,4	2,0
MIXTIME (min)	Avg	3,1	3,2	3,5	3,0	3,3	3,0	3,3	2,9	3,1	3,0
	Stdev	0,8	0,7	0,9	0,7	0,8	0,6	0,7	0,7	0,7	0,8
MIXTQ (%Tq*min)	Avg	122,1	118,4	125,3	112,1	130,3	119,3	129,4	120,9	124,0	117,1
	Stdev	29,2	24,9	32,8	26,9	31,0	25,1	27,3	26,5	28,1	30,0
ALVW ($J*10^{-4}$)	Avg	259,2	271,2	295,2	243,0	273,0	277,2	258,2	245,3	226,7	226,6
	Stdev	85,2	70,6	91,0	81,1	85,3	84,8	79,1	75,2	76,5	77,6
ALVP (mm)	Avg	93	–	–	–	79,2	96,7	99,9	99,9	89,2	86,3
	Stdev	24,1	–	–	–	16,8	24,4	23,9	24,8	24,1	21,0
ALVL (mm)	Avg	86,9	–	–	–	105,5	91,7	82,3	79,9	84,4	86,7
	Stdev	18	–	–	–	17,3	15,3	15,4	15,3	16,2	18,4
ALVPL	Avg	1,1	1,0	1,0	0,9	0,8	1,1	1,3	1,3	1,1	1,1
	Stdev	0,5	0,3	0,4	0,4	0,2	0,4	0,5	0,6	0,5	0,4
LOFVOL (mL)	Avg	781,9	736,2	809,8	814,6	816,7	790,6	736,4	727,4	768,1	761,7
	Stdev	65,5	64,0	58,0	62,5	54,8	55,8	54,0	53,4	47,3	51,9

TW, test weight; TKW, thousand kernel weight; SKCS, grain hardness by Single Kernel Characterization System; PSI, grain hardness by Particle Size Index; GPRO, grain protein content; FPRO, flour protein content; SDSS, SDS-Sedimentation volume; MIXTIME, mixograph optimum mixing time; MIXTQ, mixograph midline peak integral; ALVW, alveograph W; ALVP, alveograph P; ALVL, alveograph L; ALVP/L, alveograph PL; LOFVOL, bread loaf volume. Avg, average; Stdev, standard deviation.

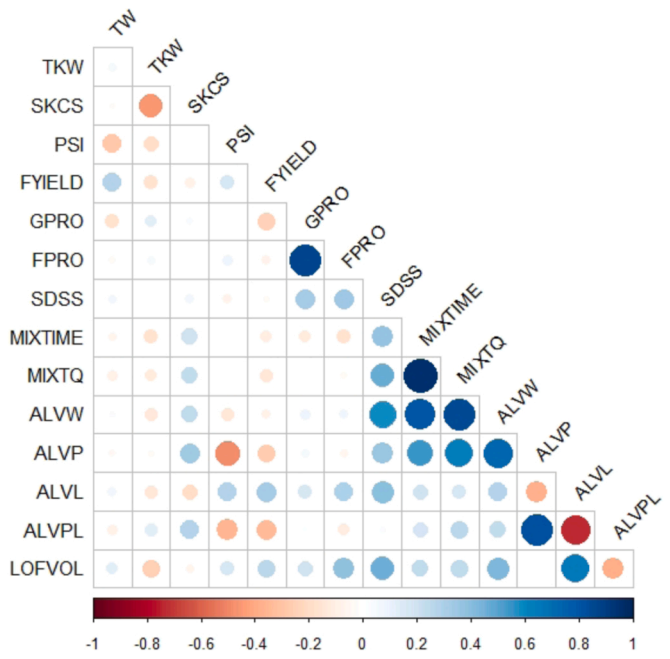

Fig. 1. Pearson phenotypic correlations among the different grain quality traits. The larger the circle the higher the Pearson coefficient value. Circles in blue color indicate direct co-relationships and circles in red color indicate inverse correlations. TW, test weight; TKW, thousand kernel weight; SKCS, grain hardness by Single Kernel Characterization System; PSI, grain hardness by Particle Size Index; FYIELD, Flour yield; GPRO, grain protein content; FPRO, flour protein content; SDSS, SDS-Sedimentation volume; MIXTIME, mixograph optimum mixing time; MIXTQ, mixograph torque. ALVW, alveograph W; ALVP, alveograph P; ALVL, alveograph L; ALVPL, alveograph P/L; LOFVOL, bread loaf volume.

limited number of lines across the 9 years (8.9–0%). At the *Glu-B1* locus, *Glu-B1c* was the most frequent allele, followed by allele *Glu-B1i*. No clear trends could be detected for any *Glu-B1* allele except for allele *Glu-B1a* which was clearly negatively selected along the years. At the *Glu-D1* locus, allele *Glu-D1d* was predominant across all years (86.5–99.6%) whereas the number of lines possessing allele *Glu-D1a* appeared to diminish with the years reaching a minimum frequency of 0.4% during the 2018–19 cycle. Interestingly, in the cycle 2019–20, the number of lines with this allele was unexpectedly higher (13.5%) suggesting a deliberate use of more parents carrying the allele *Glu-D1a* (Fig. 2). Among the *Glu-3* loci, a less clear selection pattern could be identified. At the *Glu-A3* locus for example, most of the lines appeared to have either the *Glu-A3c* or the *Glu-A3b* alleles whose frequencies did not considerably change across the years. At the *Glu-B3* locus instead, the two most frequent alleles were *Glu-B3h* (average frequency of 51.3%) and *Glu-B3b* (average frequency 28.8%). Alleles *Glu-B3c*, *Glu-B3g?* and *Glu-B3j* were the least represented. Finally, at the *Glu-D3* locus, most of the lines possessed allele *Glu-D3b* (65.9–86.3%) whereas the rest of the alleles were only marginally represented (Fig. 2).

3.3. *Effect of the individual glutenin loci*

The contribution of each individual glutenin locus on different quality traits was investigated using six analyses of variance each including either one of the *Glu* loci, Year as cofactor and FPRO as co-variate. In general, both the glutenin loci, the environmental conditions (year), and protein content, were significantly associated with variations of all the quality parameters with only few exceptions (Supplementary Table 3). Among the glutenin loci, *Glu-B1* and *Glu-B3* had the highest effect on the variations for gluten, dough and end-use quality traits, whereas *Glu-A1* and *Glu-D3* were found to have, on average, the lowest impact. As expected, the *Glu-D1* locus had a strong impact on gluten strength (MIXTIME, MIXTQ and ALVW) but, interestingly, its contribution to either SDSS, gluten extensibility and bread loaf volume was minimal (Table 2, R^2 comparisons).

Within each *Glu-1* locus, significant differences were found among the different alleles for most of the traits (Fig. 3, Supplementary Table 4). At the *Glu-A1* locus specifically, alleles *Glu-A1a* (subunit 1) and *Glu-A1b* (subunit 2*) were associated with stronger gluten, lower extensibility and higher loaf volumes whereas lines with the *Glu-A1c* allele had in general a poorer quality profile. At the *Glu-B1* locus, alleles

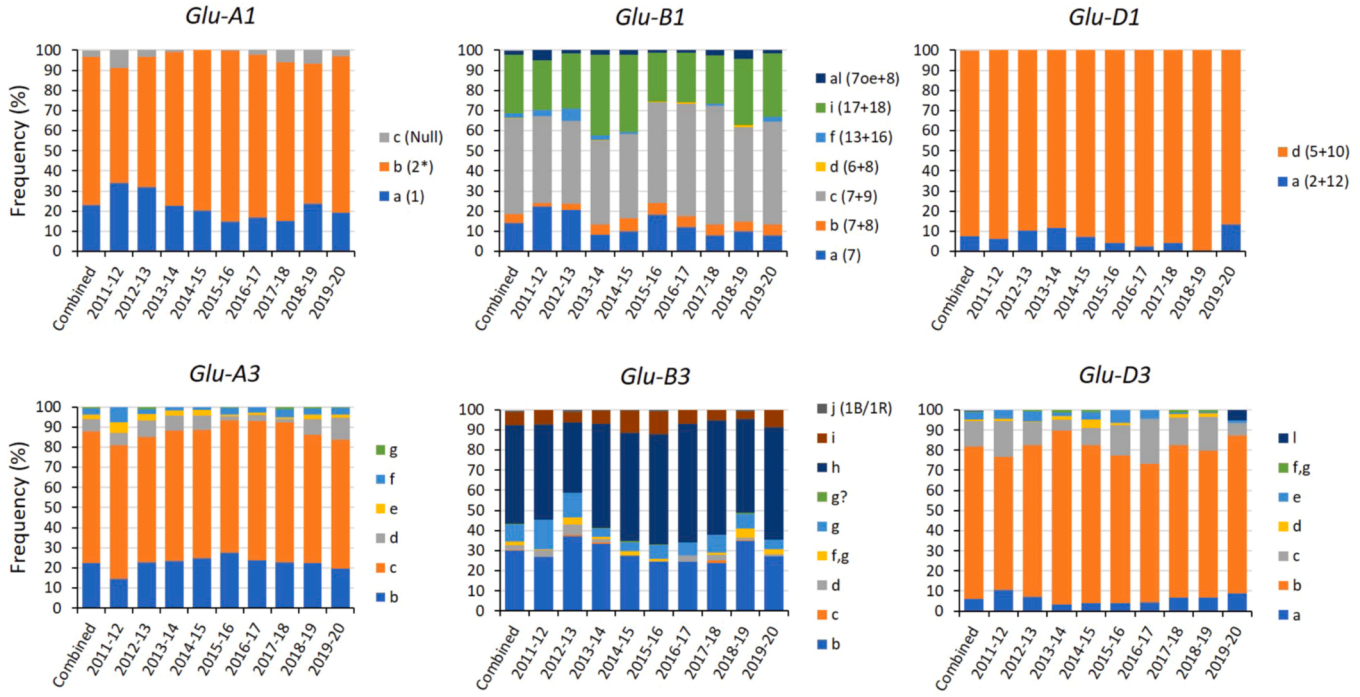

Fig. 2. Frequency of the *Glu-1* and *Glu-3* alleles identified in the analysed lines grown across 9 different agronomic cycles.

Table 2
Comparison of the R^2 values associated with the models including the year as cofactor and flour protein as cofactor and/or the glutenin loci.

Factor(s) in the model[a]	R^2							
	SDSS	MIXTIME	MIXTQ	ALVW	ALVP	ALVL	ALVPL	LOFVOL
No Glutenins	0,28	0,07	0,04	0,12	0,09	0,26	0,14	0,47
Glu-A1	0,29	0,07	0,04	0,13	0,10	0,26	0,15	0,48
Glu-B1	0,34	0,14	0,14	0,22	0,15	0,27	0,16	0,48
Glu-D1	0,28	0,20	0,15	0,18	0,11	0,27	0,14	0,47
Glu-A3	0,32	0,10	0,08	0,16	0,12	0,27	0,15	0,48
Glu-B3	0,33	0,20	0,18	0,27	0,16	0,30	0,15	0,51
Glu-D3	0,30	0,08	0,05	0,14	0,13	0,26	0,15	0,47
Glu-1 No Interactions	0,36	0,29	0,27	0,29	0,17	0,28	0,16	0,50
Glu-1 Interactions	0,38	0,30	0,28	0,31	0,19	0,30	0,18	0,51
Glu-3 No Interactions	0,38	0,23	0,21	0,30	0,22	0,31	0,18	0,52
Glu-3 Interactions	0,41	0,28	0,26	0,34	0,27	0,34	0,22	0,55
Glu-1 and *Glu-3* No Interactions	0,46	0,45	0,44	0,47	0,29	0,33	0,20	0,54

SDSS, SDS-Sedimentation volume; MIXTIME, mixograph optimum mixing time; MIXTQ, mixograph midline peak integral; ALVW, alveograph W; ALVP, alveograph P; ALVL, alveograph L; ALVP/L, alveograph PL; LOFVOL, bread loaf volume.

[a] All the models include year as cofactor and flour protein as covariate

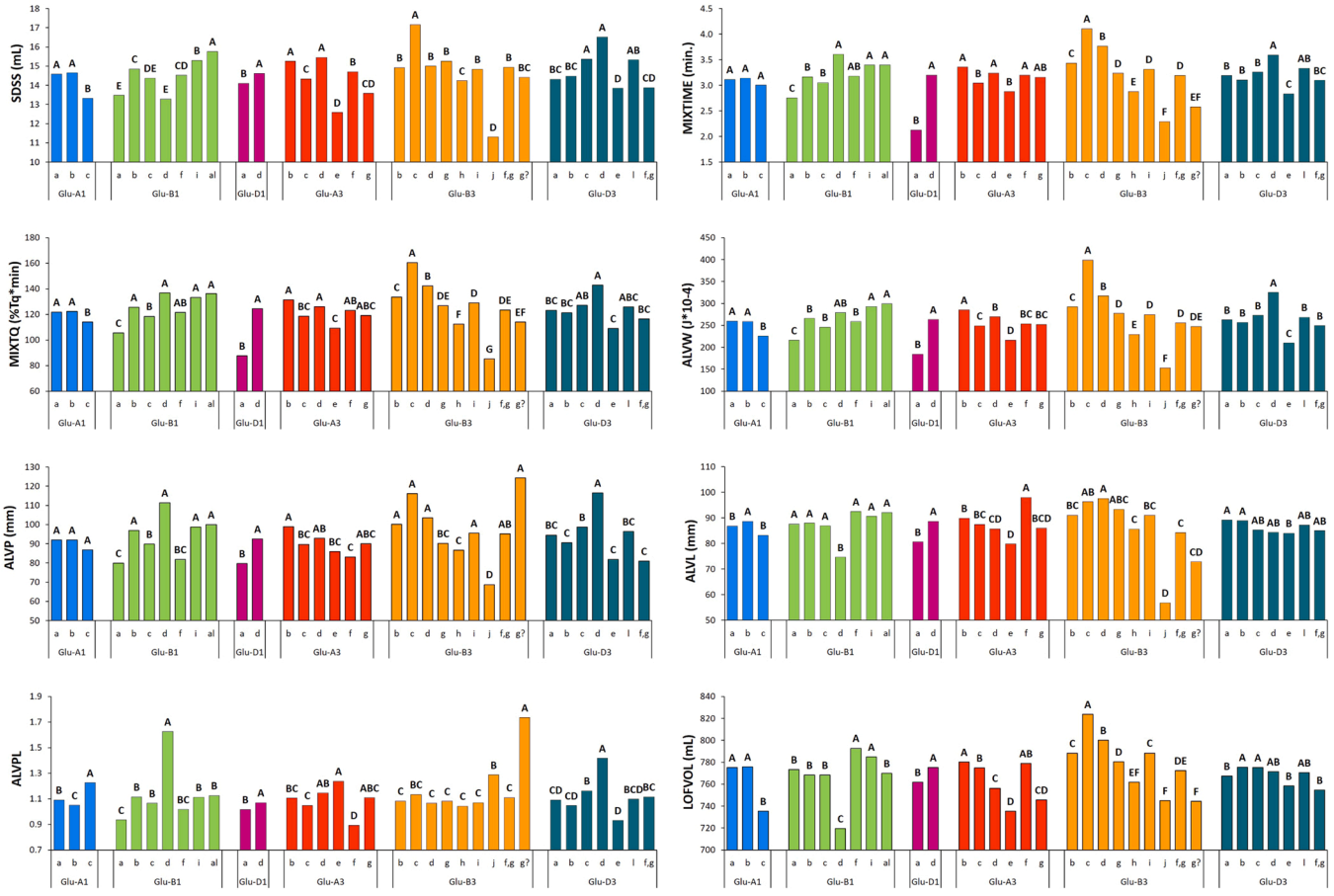

Fig. 3. LS mean values of the quality traits indicative of gluten and end-use quality of the genotypes carrying specific *Glu-1* and *Glu-3* alleles. Within each locus, capital letters on the top of the bars identify the different groups based on LSD test. SDSS, SDS-Sedimentation volume; MIXTIME, mixograph optimum mixing time; MIXTQ, mixograph torque. ALVW, alveograph W; ALVP, alveograph P; ALVL, alveograph L; ALVPL, alveograph P/L; LOFVOL, bread loaf volume.

Glu-B1i (subunits 17 +18) and *Glu-B1al* (subunits 7^{OE}+8) were generally associated with greater gluten strength (SDS-sedimentation, mixograph mixing time, alveograph W, alveograph P) and extensibility (alveograph L and P/L) whereas allele *Glu-B1f* (subunits 13 +16) was linked with medium gluten strength, low tenacity and with the highest gluten extensibility and bread loaf volumes (792.8 mL). Alleles *Glu-B1b* (subunits 7 +8) and *Glu-B1d* (subunits 7 +9), were associated with average values for all the analysed quality traits. Allele *Glu-B1a* (subunit 7) and allele *Glu-B1d* (subunits 6 +8) instead, were associated with the lowest gluten strength and bread loaf volumes, respectively. Allele *Glu-B1d* was very infrequent in the dataset. At the *Glu-D1* locus, allele *d* (subunits 5 +10) was associated with remarkably higher gluten strength than allele *a* (subunit 2 +12) and with slightly better bread-making quality. No clear differences in gluten extensibility could be identified between the *Glu-D1a* and *Glu-D1d* alleles (Table 3, allele effect).

Significant differences were also detected between the allelic variants of the *Glu-3* loci (Fig. 3, Supplementary Table 4). At the *Glu-A3*

Table 3
Comparison of the effect of the glutenin alleles on SDS-Sedimentation volume, gluten strength, gluten extensibility and bread loaf volume.

Locus	SDS-Sedimentation Volume	Gluten Strength[a]	Gluten Extensibility[b]	Bread loaf volume
Glu-A1	1 = 2 * > Null	1 = 2 * > Null	2 * > 1 ≥ Null	1 = 2 * > Null
Glu-B1	7oe+ 8 > 17 + 18 > 7 + 8 ≥ 13 + 16 ≥ 7 + 9 ≥ 7 = 6 + 8	7oe+ 8 = 17 + 18 = 6 + 8 ≥ 13 + 16 = 7 + 8 ≥ 7 + 9 > 7	7 ≥ 13 + 16 ≥ 7 + 8 = 7 + 9 = 17 + 18 = 70e+ 8 > 6 + 8	13 + 16 = 17 + 18 > 7 = 7oe+ 8 = 7 + 9 = 7 + 8 > 6 + 8
Glu-D1	5 + 10 > 2 + 12	5 + 10 > 2 + 12	5 + 10 = 2 + 12	5 + 10 > 2 + 12
Glu-A3	d = b > f > c ≥ g ≥ e	b > d ≥ f = g ≥ c > e	f > c ≥ b ≥ g ≥ d ≥ e	b ≥ f ≥ c > d ≥ g ≥ e
Glu-B3	c > g = d = f,g = b = i ≥ g? ≥ h > j	c > d > b > g = i = f,g ≥ g? ≥ h > j	d ≥ c = g ≥ b = i ≥ h = f,g > j ≥ g?	c > d > i = b > g ≥ f,g ≥ h ≥ j = g?
Glu-D3	d = c ≥ l ≥ b = a ≥ f,g ≥ e	d > c = l > a = b ≥ f,g ≥ e	a = b ≥ e ≥ l ≥ f,g ≥ c ≥ d	b = c ≥ d = l ≥ a = e = f,g

[a] The classification for gluten strength was based on the combined results obtained for mixograph peak time, mixograph midline peak integral and alveograph W.
[b] The classification for gluten extensibility was based on the results obtained for alveograph L and alveograph P/L.

locus, alleles *Glu-A3b* and *Glu-A3d* were associated with the highest gluten strength and, in the case of allele *Glu-A3b*, also with the highest bread loaf volumes. Allele *Glu-A3e* instead, was associated with the poorest quality across all different quality traits, followed by allele *Glu-A3g*. Alleles *Glu-A3c* and *Glu-A3f* both exhibited medium gluten strength and, in the case of allele *Glu-A3f*, also high gluten extensibility and bread loaf volume. At the *Glu-B3* locus, alleles *Glu-B3c* and *Glu-B3j* were associated with the highest and lowest values, respectively, for most of the traits. Both alleles however, were very infrequent in the dataset. All the other *Glu-B3* alleles were associated with medium gluten strength and extensibility and with average loaf volume values ranging from 744.3 to 800.1 mL. Similar to the *Glu-B3* locus, also two infrequent alleles at the *Glu-D3* locus were associated with the highest (*Glu-D3d*) and lowest (*Glu-D3e*) values of the different gluten and end-use quality traits. The rest of the alleles instead, exhibited average gluten properties and bread-baking quality. The only exceptions were alleles *Glu-D3b* and *Glu-D3c* which were associated with the highest bread loaf volumes (775.6 mL and 775.1 mL, respectively) (Table 3, allele effect).

3.4. Effects of the HMWGs

To understand the effect of the HMWGs on the variation of the analysed quality traits, two further ANOVA were conducted: one including all the *Glu-1* loci and another one including all the *Glu-1* loci with their interactions (Supplementary Table 5). Compared to the statistical models including only single glutenin loci, the models with all the *Glu-1* loci were able to better explain the variation of both the gluten properties and bread-baking quality. This was especially true for the mixograph parameters (MIXTIME and MIXTQ) for which a minimum R^2 increase of 0.10 could be observed. Interestingly however, the effect of the glutenin interactions appeared to be negligible and only a moderate increase (~ 0.02) in the R^2 values was achieved when model included the two-way and three-way interactions among the *Glu-1* loci (Table 2 R^2). Nevertheless, some interactions still appeared to significantly influence the analysed quality traits such as the *Glu-A1*Glu-B1* interaction which explained a greater amount of variation of bread loaf volume compared to the *Glu-A1* and *Glu-B1* loci alone (Supplementary Table 5).

Apart from the ANOVA analyses, the genotypes were grouped based on their *Glu-1* composition obtaining 19 different allele combinations represented by at least 15 genotypes. For each of these combinations, the mean and standard deviation values of the alveograph W, alveograph P/L and bread loaf volume were calculated with the aim of identifying the "best" *Glu-1* combinations associated with good gluten and bread making quality (Supplementary Fig. 1). As expected, the standard deviation was high across all the three quality traits and within each HMWGs combination group. However, some trends in the HMWGs combinations could still be identified. Two of the three best combinations for gluten strength included the allele *Glu-B1al* (7oe+8). All the combinations associated with greater gluten strength for example (average ALVW > 250) always included the *Glu-D1d* allele (subunits 5 +10). However, the presence of the *Glu-D1d* allele was also associated with the allelic combination that had the poorest gluten extensibility and loaf volume (*Glu-A1c/Glu-B1c/Glu-D1d*, N = 67). On the other hand, two (*Glu-A1a/Glu-B1i/Glu-D1a*, N = 23; *Glu-A1a/Glu-B1f/Glu-D1a*, N = 22) of the three best HMW combinations for bread loaf volume (average LOFVOL > 830 mL) and gluten extensibility (average ALVPL < 1), had the allele *Glu-D1a* (subunits 2 +12). The third combination associated with the highest loaf volume values (*Glu-A1a/Glu-B1f/Glu-D1d*, N = 59) exhibited medium-high alveograph W values and medium-low alveograph P/L values. The most frequent combination of HMWGs alleles (*Glu-A1b/Glu-B1c/Glu-D1d,* N = 1527) exhibited intermediate values for all the analysed quality traits, whereas the second most frequent combination (*Glu-A1b/Glu-B1i/Glu-D1d,* N = 884) was associated with strong gluten and medium-high bread-making quality.

3.5. Effects of the LMWGs

The effect of the LMWGs variation on gluten properties and bread-making quality was investigated by performing two additional ANOVA which included either all the three *Glu-3* loci alone, or all the three *Glu-3* loci and their interactions. In general, both these models were able to explain a greater amount of variation of all the analysed quality traits compared with the statistical models including only the single glutenin loci (Table 2 R^2). Also, for most of the traits, the models including all the three *Glu-3* loci were slightly more powerful than the models including all the *Glu-1* loci (with or without the interactions). The only exceptions were the two mixograph parameters (MIXTIME and MIXTQ) whose variation could be better explained by the HMWGs (Table 2 R^2). Differently from the HMWGs however, the effect of the interactions between the LMW-GSs loci was moderate and most of the two-way and three-way interactions were significantly associated with the analysed traits (Supplementary Table 6 on *Glu-3* interactions). The only exception was alveograph L for which only the triple interaction (*Glu-A3* *Glu-B3* *Glu-D3*) was found to be significant.

The mean and standard deviation values of the different groups of genotypes having the same LMWGs combination and a significant number of lines (N > 15), were also examined (Supplementary Fig. 2). Among these combinations, the one with alleles *Glu-A3b/Glu-B3i/Glu-D3c* (N = 37) was the best in terms of gluten strength, with average ALVW values > 350. This group however, was also associated with the highest ALVP/L values (> 1.5) which is indicative of tenacious gluten, and with average bread loaf volumes. Interestingly, the second-best combination for gluten strength (*Glu-A3c/Glu-B3d/Glu-D3b,* N = 77) had medium-low alveograph P/L and high loaf volume values. As for the HMWGs combinations, the LMWGs group associated with the highest loaf volume was also associated with the greatest extensibility (average ALVP/L of 0.73) and medium gluten strength. The most common LMWGs allele combination found among the genotypes used for the study (*Glu-A3c/Glu-B3h/Glu-D3b,* N = 1159) was associated with medium-low gluten strength and loaf volume, and balanced gluten (alveograph P/L ~1). The second most frequent combination (*Glu-A3c/Glu-B3b/Glu-D3b,* N = 673) was different from the first one only for the presence of the *Glu-B3b* allele instead of the allele *Glu-B3h*. Genotypes with this LMWGS combination exhibited medium-high gluten strength and loaf volume and balanced gluten.

3.6. Effects of the HMWGs and LMWGs

Finally, the effect of both the HMWGs and LMWGs on gluten strength, gluten extensibility and breadmaking quality, was evaluated by using a statistical model including all the *Glu-1* and *Glu-3* loci with no interactions. As expected, this model was able to better explain the variation of all quality traits compared to all the other models. Interestingly however, the increase in the % of variation explained, was higher for the traits associated with gluten strength (MIXTIME, MIXTW and ALVW) than for the other analysed traits (Table 2 R^2). Compared with the effect of the year and flour protein content, the glutenin loci had in general a lower effect on the analysed quality traits. The only exceptions were the two mixograph parameters and ALVW which appeared to be more genetically controlled. Indeed, the greatest amount of variation of these traits was caused by changes at the *Glu-B1*, *Glu-D1* and *Glu-B3* loci which explained somewhere between 7% and 16% of their variation (Supplementary Table 7). Bread loaf volume in contrast, was almost completely controlled by the environment (year) and protein content, which explained together more than 40% of the observed variation. Among the glutenin loci, *Glu-B3*, *Glu-A3* and *Glu-A1* were the loci with a greater impact on bread loaf volume, explaining respectively 4%, 1% and 1% of its variation (Supplementary Table 7).

When the HMWGs and LMWGs compositions were considered together a total of 482 combinations was found. More than half of these combinations (273) were represented by three or less genotypes. Again, we focused on those combinations that had more than 15 genotypes (Supplementary Table 8). The combination *Glu-A1b/Glu-B1i/Glu-D1d/Glu-A3b/Glu-B3i/Glu-D3c* (N = 21) showed the highest mean values for alveograph W and P/L and medium loaf volume. All the alleles of this combination were associated in the first analysis of the individual glutenins effects with the highest or second highest values for ALVW. Other two combinations with high gluten strength and high loaf volume were *Glu-A1a/Glu-B1i/Glu-D1d/Glu-A3c/Glu-B3g/Glu-D3b* (N = 30, alveograph W = 353 $J*10^{-4}$, loaf volume = 845 mL) and *Glu-A1a/Glu-B1f/Glu-D1d/Glu-A3c/Glu-B3d/Glu-D3b* (N = 36, alveograph W = 338 $J*10^{-4}$, loaf volume = 873 mL). The alveograph P/L mean values of these glutenins combinations were typical of balanced gluten (1 and 0.9, respectively). Another interesting combination was *Glu-A1b/Glu-B1a/Glu-D1d/Glu-A3b/Glu-B3b/Glu-D3b*, which was present in 51 genotypes and had mean ALVW, ALVPL and LOFVOL values of 287 $J* 10^{-4}$, 0.9, and 841 mL, respectively. The most common combinations were *Glu-A1b/Glu-B1c/Glu-D1d/Glu-A3c/Glu-B3h/Glu-D3b*(N = 577, alveograph W = 227 $J*10^{-4}$ and P/L=1.1, loaf volume=758 mL) and *Glu-A1b/Glu-B1i/Glu-D1d/Glu-A3c/Glu-B3b/Glu-D3b* (N = 209, alveograph W=331 $J*10^{-4}$ and P/L=1, loaf volume=820 mL). These two combinations had a strongly different quality profile but only differed in the *Glu-B1* and *Glu-B3* loci.

3.7. Model selection

In order to confirm the results obtained in the ANOVA analysis and to identify the factors (glutenin loci, flour protein content or environment) that should be prioritized when selecting for gluten and bread making quality, a regression model was selected for each of the analysed traits by using a stepwise selection approach (Supplementary Table 9). According to the obtained results, both SDSS, MIXTIME, MIXTQ and ALVW were best modeled by using all the independent variables. Interestingly however, the greatest amount of variation of SDSS was explained by flour protein content (12%) and year (16%) followed by the glutenin loci. On the other side, both the mixograph and the ALVW parameter were mostly explained by the *Glu-D1*, *Glu-B3* and *Glu-B1* loci confirming that variations at these glutenin loci are indeed the most influential factors affecting gluten strength. The ALVP, which is also somehow associated with gluten strength, was best modeled by the year and the glutenin loci which all together were able to explain ~ 30% of its variation. Dough extensibility instead, represented by both the alveograph parameter P/L and L, could be only poorly predicted by the selected factors and, in the case of ALVL, only the year, flour protein content, *Glu-B3* and *Glu-D1* loci appeared to significantly influence this trait explaining all together 30% of its variation. Bread loaf volume finally, was best modeled by all the factors apart from *Glu-D3*. As expected however, year and flour protein content were the first factors to enter the model explaining the greatest amount of variation, whereas the addition to the model of the glutenin loci only contributed to a minimum increase (< 5%) of the variation in bread loaf volume (Fig. 4).

4. Discussion

Glutenins are the major components of the wheat seed storage proteins and constitute the major structure of gluten, a macropolymer responsible for the unique visco-elastic properties of wheat dough. Allelic variation of both the HMWGS and LMWGS has been long associated with changes in dough rheological properties and bread-making quality and, up to now, several alleles associated with specific wheat quality characteristics have been identified. However, given the high cost of conducting a full quality characterization of a wheat line, most of the studies conducted so far have analyzed the effect of the glutenin alleles only on a relatively limited number of lines (typically less than 300). These limitations led to often inconsistent results and made it hard to unequivocally establish which glutenin locus or glutenin loci combinations more strongly influences gluten properties. Furthermore, most

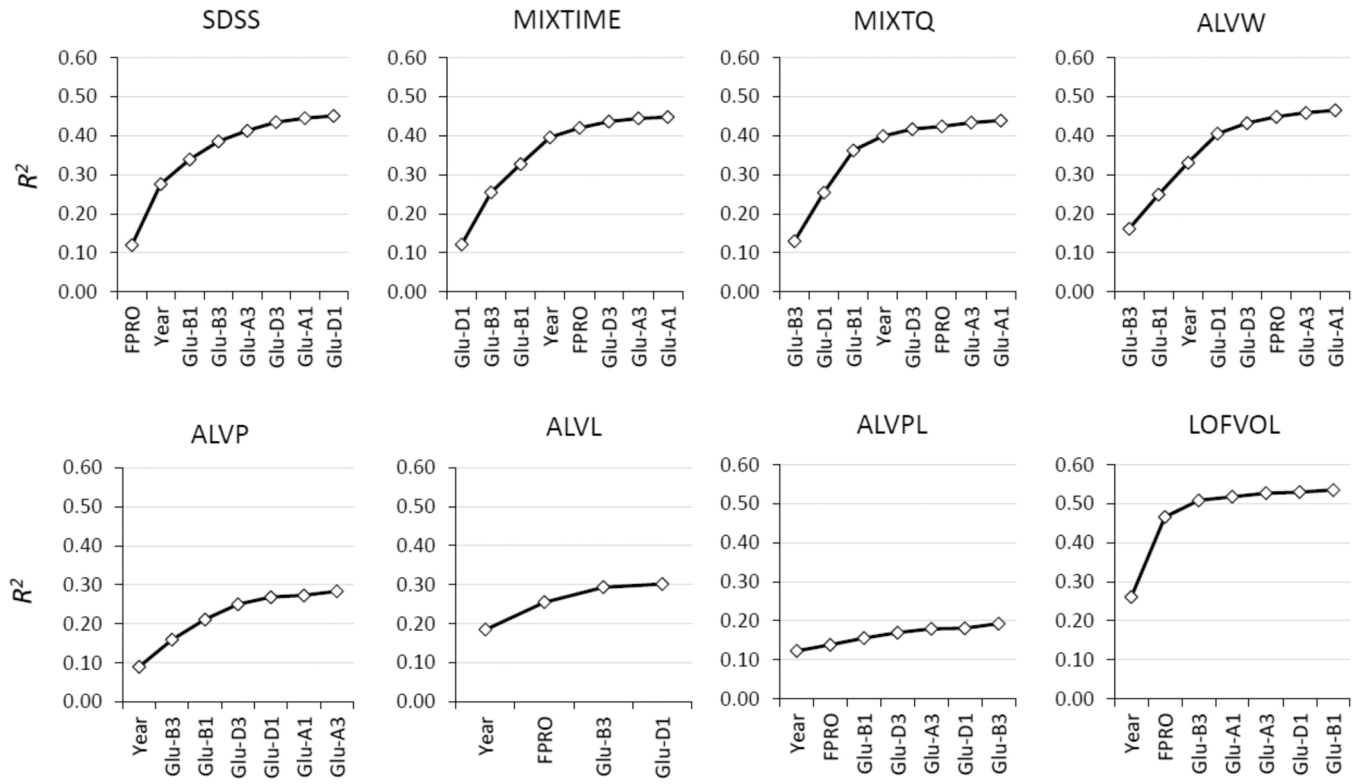

Fig. 4. Parameters selected for multiple regression to model gluten and bread-making quality. For each trait are reported the parameters that are gradually admitted in the model and their relative contribution to explain the analysed trait (R^2). SDSS, SDS-Sedimentation volume; MIXTIME, mixograph optimum mixing time; MIXTQ, mixograph torque. ALVW, alveograph W; ALVP, alveograph P; ALVL, alveograph L; ALVPL, alveograph P/L; LOFVOL, bread loaf volume.

of the results already reported in the literature were obtained using experimental lines whose results cannot often be applied to specific breeding programs. For these reasons, in the present study we utilized the quality data generated on 2550 lines across 10 different years at the CIMMYT bread wheat breeding program, to estimate the effect of each glutenin locus on gluten quality and bread-making potential. To the best of our knowledge, this is the first study investigating the effect of the glutenin loci on such a large number of lines and using such a complete set of wheat quality data.

Based on the results obtained, in general, all the glutenin loci appeared to significantly influence the analyzed gluten and end-use quality traits. However, depending on the trait, the effect of each locus changed. The SDS-sedimentation volume for example, was influenced by the *Glu-B1* > *Glu-B3* > *Glu-A3* > *Glu-D3* > *Glu-A1* > *Glu-D1* loci. Similar results were also obtained by Liu et al. (2005) and He et al. (2005), who also found that the *Glu-B3* and *Glu-B1* loci were the most influential for such trait. Gluten strength, as indicated by mixograph peak time, mixograph midline peak integral and alveograph W, was mostly influenced by the *Glu-B3*, *Glu-D1* and *Glu-B1* loci, with the *Glu-D3*, *Glu-A3* and *Glu-A1* loci only marginally contributing to its variation. These results are in agreement with those reported by Gupta and MacRitchie (1994), He et al. (2005) and Liu et al. (2005), who also found that gluten strength was mostly influenced by allelic variation of the *Glu-D1*, *Glu-B1* and *Glu-B3* loci. Interestingly, the *Glu-D1* contributed to less than 1% of the SDSS variation but played a fundamental role in the variation of the analyzed mixograph parameters and ALVW explaining on average 10% of their variation. The SDS-sedimentation volume analysis is a widely used method to estimate the overall gluten quality and bread-making quality of a wheat sample during the early stages of selection. In fact, this test showed a very high correlation with LOFVOL in our study (r = 0.45). However, even if a positive correlation exists between SDS-sedimentation volume and the other analyzed gluten strength parameters (MIXTIM, MIXTQ and ALVW), variation of SDSS is also strongly influenced by changes in protein content compared to the mixograph of ALVW parameters (Fig. 1) (Clarke et al., 2010). These observations, together with the inherent differences of the chemistry behind such methods, could explain the different contributions that the glutenin loci, and *Glu-D1* specifically, have on SDSS compared with the MIXTIM, MIXTQ and ALVW traits. Gluten extensibility was mostly determined by variations in the environments and protein content, and allelic variation of the glutenin loci was able to only explain a limited amount of variation of this trait (< 10%). Similar results were also obtained by Gupta and MacRitchie (1994) and Branlard et al. (2001) who found that dough extensibility, analyzed either through the alveograph (ALVL) or the extensograph, was mainly influenced by protein content or other unknown genetic or non-genetic factors. Nevertheless, at constant protein content, the different glutenin alleles still significantly influenced this trait (Gupta and MacRitchie, 1994) suggesting that selection of specific glutenin alleles is still relevant for the improvement of dough extensibility. Among the glutenin loci that more strongly influenced this trait, *Glu-B3* had the largest effect and consistently appeared to have a significant impact on dough extensibility in both this and previous studies (Branlard et al., 2001; He et al., 2005). Similar to dough extensibility, also variations in bread loaf volume were mostly determined by the environment and the protein content. Even if moderately however, all the glutenin loci significantly influenced this trait with the *Glu-B3* locus being the most influential followed by the *Glu-A3*, *Glu-A1*, *Glu-B1*, *Glu-D1* and *Glu-D3*, respectively. Interestingly, the *Glu-D1* locus only had limited effect on bread loaf volume explaining less than 1% of its variation. These results are in contrast with those reported by Payne (1987) where the *Glu-D1* locus appeared to be the second most influential factor for bread loaf volume contributing to more than 23% of its variation. Variations in the protein content, environmental conditions and glutenin composition at the other *Glu-1* and *Glu-3* loci, might explain these apparent inconsistencies.

When analyzing the effect of each allele on the gluten and bread making quality, the "best" HMWGS alleles associated with greater gluten strength, good extensibility and higher bread loaf volume were *Glu-A1a* (subunit 1), *Glu-A1b* (subunit 2 *), *Glu-B1al* (subunits 7^{OE}+8), *Glu-B1i* (subunits 17 +18), *Glu-B1f* (13 +16) and *Glu-D1d* (subunits 5 +10) whereas the alleles associated with the lower overall quality were *Glu-A1c* (Null), *Glu-B1a* (subunit 7), *Glu-B1d* (subunits 6 +8) and *Glu-D1a* (subunits 2 +12). These results are in agreement with previous studies conducted on different sets of germplasm (Branlard et al., 2001; Eagles et al., 2002; He et al., 2005; Hernández et al., 2013; Liu et al., 2005; Rathan et al., 2020; Ribeiro et al., 2017) suggesting that despite the different genetic background, the effect of the HMWGS alleles is constant and specific alleles can be selected to effectively improve this trait. Based on the HMWGS allele frequencies identified in the CIMMYT breeding program, it is clear that a strong selection towards alleles associated with greater gluten strength at the *Glu-A1* and *Glu-D1* loci has already been underway with *Glu-A1c* and *Glu-D1a* alleles being present at very low frequencies. A previous study carried out 20 years ago with CIMMYT germplasm (Trethowan et al., 2001) found frequencies of *Glu-A1c* and *Glu-D1a* alleles of 9% and 37%, respectively, which confirm that hypothesis. The average values of ALVW in the full irrigation environment of that study were of 152 and 165 J* 10^{-4}, while in the current study the average ALVW value across all samples was of 227 J* 10^{-4}, which reflect the improvement of this parameter over time. Similarly, average values of LOFVOL in the previous study were of 687 and 693 mL while we had 767 mL, which indicates the progress of the program for this important trait too. Unfortunately, Trethowan et al. (2001) did not analyze the LMWGs composition and, therefore, comparisons on how the frequencies of the LMWGS alleles have evolved cannot be made.

At the *Glu-B1* locus instead, it would be advisable to reduce the frequency of the alleles associated with low gluten strength (*Glu-B1a* specifically) and increase the frequency of other alleles such as the *Glu-B1f* allele (subunits 13 +16), associated with medium gluten strength but high loaf volume and dough extensibility. Similarly, for *Glu-A1,* sources of active Ay subunits could be used in the breeding program to enhance the contribution of this locus to processing and end-use quality traits (Roy et al., 2020). Even if the association between greater gluten strength and the *GluD1d* allele (subunits 5 +10) is clear and widely reported in this and previous studies (Branlard et al., 2001; Eagles et al., 2002; Hernández et al., 2013; Liu et al., 2005; Park et al., 2011; Rathan et al., 2020), elimination of the *Glu-D1a* allele is not advisable. Despite the lower gluten strength in fact, this allele has been associated with greater loaf volumes and dough extensibility (Kiszonas et al., 2021; Mohan and Gupta, 2015). In this study indeed, two of the three best HMWGS allele combinations that exhibited the highest bread loaf volume included the *Glu-D1a* allele.

Among the LMWGS alleles, *Glu-A3b*, *Glu-A3d*, *Glu-A3f*, *Glu-B3c* and *Glu-B3d* were associated with greater gluten strength, extensibility and bread loaf volume whereas alleles *Glu-A3e* and *Glu-B3j* were associated with the lowest score for most of the analyzed traits. Also in this case, the results agree with those previously published. In most cases in fact, alleles *Glu-A3b, Glu-A3d* and *Glu-A3f* were associated with higher SDS-sedimentation volume (He et al., 2005; Liu et al., 2005), greater gluten strength (Branlard et al., 2001; He et al., 2005; Liu et al., 2005) and higher bread loaf volumes (Ibba et al., 2017b). Allele *Glu-A3e* instead, described in the literature as a null allele (Ibba et al., 2017a) is consistently associated with lower gluten and end-use quality (Bonafede et al., 2015; Branlard et al., 2001; Eagles et al., 2002; He et al., 2005; Liu et al., 2005; Park et al., 2011; Ribeiro et al., 2017) confirming the results obtained in the present study. Similarly, the *Glu-B3c* and *Glu-B3d* alleles were found to be associated with an overall greater gluten quality in several previous studies (Branlard et al., 2001; He et al., 2005; Liu et al., 2005; Park et al., 2011) and allele *Glu-B3j*, indicative of the 1B/1 R translocation (Ibba et al., 2017a), is associated with a severe reduction of the overall gluten quality (Branlard et al., 2001; Eagles et al., 2002; He et al., 2005; Liu et al., 2005; Rathan et al., 2020). As also reported in previous studies, the effect of the *Glu-D3* locus was minimal for most of the analyzed quality traits. When analyzing the LMWGS allele frequencies of the samples analyzed in the current study, it seems that no clear selection of specific *Glu-3* alleles occurred. The only exception is allele *Glu-B3j* which is almost absent from the CIMMYT bread wheat breeding program germplasm. In the future, in order to improve the overall gluten and end-use quality of the CIMMYT bread wheat samples, it would be advisable to increase the number of lines with alleles *Glu-A3b*, *Glu-A3d, Glu-A3f, Glu-B3b* and *Glu-B3d* while reducing or eliminating the lines with the *Glu-A3* null allele (*Glu-A3e*). Reduction of the samples with the *Glu-D3e* alleles could also contribute to improving the overall gluten strength of a line. This allele in fact, even if it is not as influential as the other LMWGS alleles, was constantly associated with lower gluten strength as indicated by both the mixograph or ALVW.

Due to the high number of lines used in this study, several HMWGs and LMWGs allele combinations (482) were detected. Many of them had very little presence in the whole population, but other were quite well represented which allowed to take some conclusions about what alleles should be combined to get a specific profile in terms of gluten properties and bread-making quality. The most common combination was *Glu-A1b/Glu-B1c/Glu-D1d/Glu-A3c/Glu-B3h/Glu-D3b* (N = 577, alveograph W = 227 J*10–4 and P/L=1.1, loaf volume=758 mL), which represents well the average CIMMYT quality, with medium gluten strength, balanced gluten, and medium pan bread-making quality. This type of quality profile should work reasonably well for flat and artisan breads, main products in important CIMMYT target regions such as South Asia. For this type of products, some more gluten extensibility would be desired to get top quality, and based on the results obtained, combinations *Glu-A1a/Glu-B1i/Glu-D1d/Glu-A3c/Glu-B3i/Glu-D3b* (N = 36, alveograph W = 284 J*10–4 and P/L=0.8, loaf volume=830 mL) and *Glu-A1b/Glu-B1a/Glu-D1d/Glu-A3c/Glu-B3b/Glu-D3b* (N = 153, alveograph W = 244 J*10–4 and P/L=0.8, loaf volume=811 mL) might be ideal. For mechanized bread-making, higher levels of gluten strength and capacity to expand are necessary as represented for several combinations such as *Glu-A1a/Glu-B1i/Glu-D1d/Glu-A3c/Glu-B3g/Glu-D3b* (N = 30, alveograph W = 353 J*10^{-4}, P/L = 1, loaf volume = 845 mL), *Glu-A1a/Glu-B1f/Glu-D1d/Glu-A3c/Glu-B3d/Glu-D3b* (N = 36, alveograph W = 338 J*10^{-4}, P/L = 0,9, loaf volume = 873 mL), and *Glu-A1a/Glu-B1i/Glu-D1d/Glu-A3c/Glu-B3b/Glu-D3b* (N = 62, alveograph W = 315 J*10^{-4}, P/L = 1, loaf volume = 873 mL). On the other hand, for products such as cookies requiring weak and extensible gluten (in addition to soft texture, very infrequent in CIMMYT germplasm), the best combination found was *Glu-A1b/Glu-B1a/Glu-D1d/Glu-A3c/Glu-B3h/Glu-D3b* (N = 74, alveograph W = 186 J*10^{-4}, P/L = 0,9), but this is far from the ideal quality profile for this type of products (W<100 and P/L<0.6) that can be found in wheat from areas dedicated to this market such as US Pacific Northwest.

Despite the analysis of an unbalanced data set obtained from lines grown across different agronomic cycles, the results obtained in this study are in evident agreement with those previously reported. As also stated by Eagles et al. (2002), these outcomes confirm that using the quality results obtained from a large wheat breeding program is feasible and could be effectively used to successfully estimate the effect of the glutenin alleles on wheat quality variations. Also, these results confirm that the characterization and selection of superior glutenin loci is important to improve the overall gluten and end-use quality of a sample, contributing to the reduction of the samples with an undesired quality profile that are advanced in the breeding pipeline. For this reason, the analysis of the glutenin profile of the bread wheat lines selected for the crossing-block, appear to be an effective strategy.

5. Conclusions

In the present study, the effect of the HMWGS and LMWGS variation was investigated using a set of 2550 bread wheat lines derived from the

CIMMYT spring bread wheat breeding program. Results of this analysis highlight the importance of the glutenins variation on different wheat quality aspects and confirm the usefulness of determining the glutenin profile to improve the selection efficiency for improved wheat quality. Among the different analysed traits, gluten strength was the one more strongly influenced by the glutenin variations, with the *Glu-B1*, *Glu-D1* and *Glu-B3* loci being the most influential. The other traits (SDSS, dough extensibility, bread loaf volume) also appeared to be influenced by changes in the glutenin profile, but only after considering the differences in protein content and environmental conditions. Among the different glutenin alleles, *Glu-A1a* (subunit 1), *Glu-A1b* (subunit 2*), *Glu-B1al* (subunits 7^{OE}+8), *Glu-B1i* (subunits 17 +18), *Glu-B1f* (13 +16), *Glu-D1d* (subunits 5 +10), *Glu-A3b*, *Glu-A3d*, *Glu-A3f*, *Glu-B3c* and *Glu-B3d* were associated in general with greater gluten strength, good extensibility and higher bread loaf volume. On the contrary, alleles *Glu-A1c* (Null), *Glu-B1a* (subunit 7), *Glu-B1d* (subunits 6 +8), *Glu-D1a* (subunits 2 +12), *Glu-A3e* and *Glu-B3j* were associated with an overall poor quality profile. The frequency of these alleles should be minimized in breeding programs focused on the development of varieties with strong gluten.

CRediT authorship contribution statement

Carlos Guzmán: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Resources, Writing – original draft, Writing – review & editing. **Jose Crossa:** Formal analysis, Writing – review & editing. **Suchismita Mondal:** Resources. **Velu Govindan:** Resources. **Julio Huerta:** Resources, Writing – review & editing. **Leonardo Crespo-Herrera:** Resources. **Mateo Vargas:** Formal analysis. **Ravi P. Singh:** Resources, Funding acquisition, Writing – review & editing. **Maria Itria Ibba:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Resources, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

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

Acknowledgments

Funding for wheat quality improvement at CIMMYT during the years in which the data for this study was generated was provided by CGIAR CRP WHEAT, Durable Genetic Gains in Wheat Project (Bill and Melinda Gates Foundation and DFID through a grant to Cornell University), Accelerating Genetic Gain (AGG) in Maize and Wheat Project Grant INV-003439 funded by the Bill and Melinda Gates Foundation (BMGF), the Foreign and Commonwealth Development Office (FCDO) and Foundation for Food and Agriculture Research (FFAR), USAID-Crops to End Hunger (CtEH)-AGG Wheat Supplement grant, and MasAgro Trigo (Sagarpa/Sader, Mexico). Carlos Guzman gratefully acknowledges the European Social Fund and the Spanish State Research Agency (Ministry of Science, Innovation and Universities) for financial funding through the Ramon y Cajal Program (RYC-2017-21891). Funding for open access charge: Universidad de Córdoba / CBUA.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.fcr.2022.108585.

References

AACC, 2010. Approved Methods of the American Association of Cereal Chemists, St. Paul, MN, USA.

Blechl, A., Lin, J., Nguyen, S., Chan, R., Anderson, O.D., Dupont, F.M., 2007. Transgenic wheats with elevated levels of Dx5 and/or Dy10 high-molecular-weight glutenin subunits yield doughs with increased mixing strength and tolerance. J. Cereal Sci. 45, 172–183. https://doi.org/10.1016/j.jcs.2006.07.009.

Bonafede, M.D., Tranquilli, G., Pflüger, L.A., Peña, R.J., Dubcovsky, J., 2015. Effect of allelic variation at the Glu-3/Gli-1 loci on breadmaking quality parameters in hexaploid wheat (Triticum aestivum L.). J. Cereal Sci. 62, 143–150. https://doi.org/10.1016/j.jcs.2015.02.001.

Branlard, G., Dardevet, M., Saccomano, R., Lagoutte, F., Gourdon, J., 2001. Genetic diversity of wheat storage proteins and bread wheat quality. Euphytica 119, 59–67.

Branlard, G., Dardevet, M., Amiour, N., Igrejas, G., 2003. Allelic diversity of HMW and LMW glutenin subunits and omega-gliadins in French bread wheat (Triticum aestivum L.). Genet. Resour. Crop Evol. 50, 669–679. https://doi.org/10.1023/A:1025077005401.

Carrillo, J.M., Rousset, M., Qualset, C.O., Kasarda, D.D., 1990. Use of recombinant inbred lines of wheat for study of associations of high-molecular-weight glutenin subunit alleles to quantitative traits – 1. Grain yield and quality prediction tests. Theor. Appl. Genet. 79, 321–330. https://doi.org/10.1007/BF01186074.

Clarke, F.R., Clarke, J.M., Ames, N.A., Knox, R.E., Ross, R.J., 2010. Gluten index compared with SDS-sedimentation volume for early generation selection for gluten strength in durum wheat. Can. J. Plant Sci. 90, 1–11. https://doi.org/10.4141/CJPS09035.

Dreisigacker, S., Xiao, Y., Sehgal, D., Guzman, C., He, Z., Xia, X., Peña, R.J., 2020. SNP markers for low molecular glutenin subunits (LMW-GSs) at the Glu-A3 and Glu-B3 loci in bread wheat. PLoS One 15, 1–16. https://doi.org/10.1371/journal.pone.0233056.

Eagles, H.A., Hollamby, G.J., Gororo, N.N., Eastwood, R.F., 2002. Estimation and utilisation of glutenin gene effects from the analysis of unbalanced data from wheat breeding programs. Aust. J. Agric. Res. 53, 367. https://doi.org/10.1071/AR01074.

FAOSTAT, 2019. [WWW Document]. URL: ⟨http://www.fao.org/faostat/en/#home⟩.

Feldman, M., 1995. Wheats. In: Smartt, J., Simmonds, N.W. (Eds.), Evolution of Crop Plants. Longman Scientific and Technical, Harlow, Uk, pp. 185–192.

Gupta, R.B., MacRitchie, F., 1994. Allelic variation at glutenin subunit and gliadin loci, Glu-1, Glu-3 and Gli-1 of common wheats. II. Biochemical basis of the allelic effects on dough properties. J. Cereal Sci. https://doi.org/10.1006/jcrs.1994.1004.

Guzmán, C., Posadas-Romano, G., Hernández-Espinosa, N., Morales-Dorantes, A., Peña, R.J., 2015. A new standard water absorption criteria based on solvent retention capacity (SRC) to determine dough mixing properties, viscoelasticity, and bread-making quality. J. Cereal Sci. 66, 59–65. https://doi.org/10.1016/j.jcs.2015.10.009.

Guzmán, C., Ammar, K., Govindan, V., Singh, R., 2019. Genetic improvement of wheat grain quality at CIMMYT. Front. Agric. Sci. Eng. 6, 265–272. https://doi.org/10.15302/J-FASE-2019260.

He, Z.H., Liu, L., Xia, X.C., Liu, J.J., Peña, R.J., 2005. Composition of HMW and LMW glutenin subunits and their effects on dough properties, pan bread, and noodle quality of Chinese bread wheats. Cereal Chem. 82, 345–350. https://doi.org/10.1094/CC-82-0345.

Hernández, N., Posadas, G., Cervantes, F., González, H.I., Santacruz, A., Benítez, I., Peña, R.J., 2013. Contribución de combinaciones de gluteninas a las características del gluten en poblaciones recombinantes de trigo. Rev. Fitotec. Mex. 36, 45–51.

Hernández-Espinosa, N., Payne, T., Huerta-Espino, J., Cervantes, F., Gonzalez-Santoyo, H., Ammar, K., Guzmán, C., 2019. Preliminary characterization for grain quality traits and high and low molecular weight glutenins subunits composition of durum wheat landraces from Iran and Mexico. J. Cereal Sci. 88, 47–56. https://doi.org/10.1016/j.jcs.2019.05.007.

Ibba, M.I., Kiszonas, A.M., Guzmán, C., Morris, C.F., 2017a. Definition of the low molecular weight glutenin subunit gene family members in a set of standard bread wheat (Triticum aestivum L.) varieties. J. Cereal Sci. 74. https://doi.org/10.1016/j.jcs.2017.02.015.

Ibba, Maria I., Kiszonas, A.M., Morris, C.F., 2017b. Influence of low-molecular-weight glutenin subunit haplotypes on dough rheology in elite common wheat varieties. Cereal Chem. 94, 1016–1027. https://doi.org/10.1094/CCHEM-07-17-0137-R.

Jackson, E.A., Morel, M.H., Sontag-Strohm, T., Branlard, G., Metakovsky, E.V., Redaelli, R., 1996. Proposal for combining the classification systems of alleles of Gli-1 and Glu-3 loci in bread wheat (Triticum aestivum L.). J. Genet. Breed. 50, 321–336.

Kiszonas, A.M., Ibba, M.I., Boehm, J.D., Morris, C.F., 2021. Effects of Glu-D1 gene introgressions on soft white spring durum wheat (Triticum turgidum ssp. durum) quality. Cereal Chem. 98, 1112–1122. https://doi.org/10.1002/cche.10459.

Liu, L., He, Z., Yan, J., Zhang, Y., Xia, X., Peña, R.J., 2005. Allelic variation at the Glu-1 and Glu-3 loci, presence of the 1B.1R translocation, and their effects on mixographic properties in Chinese bread wheats. Euphytica 142, 197–204. https://doi.org/10.1007/s10681-005-1682-4.

Liu, L., Ikeda, T.M., Branlard, G., Peña, R.J., Rogers, W.J., Lerner, S.E., Kolman, M.A., Xia, X., Wang, L., Ma, W., Appels, R., Yoshida, H., Wang, A., Yan, Y., He, Z., 2010. Comparison of low molecular weight glutenin subunits identified by SDS-PAGE, 2-DE, MALDI-TOF-MS and PCR in common wheat. BMC Plant Biol. 10 https://doi.org/10.1186/1471-2229-10-124.

Maryami, Z., Huertas-García, A.B., Azimi, M.R., Hernández-Espinosa, N., Payne, T., Cervantes, F., Govindan, V., Ibba, M.I., Guzman, C., 2020. Variability for glutenins, gluten quality, iron, zinc and phytic acid in a set of one hundred and fifty-eight common wheat landraces from Iran. Agronomy 10, 1797. https://doi.org/10.3390/agronomy10111797.

Mcintosh, A., Dubcovsky, J., Rogers, J., Xia, C., Raupp, J., 2020. Catalogue of gene symbols for wheat: 2020 Supplement. Annu. Wheat Newsl. 66, 109–128.

Mohan, D., Gupta, R.K., 2015. Gluten characteristics imparting bread quality in wheats differing for high molecular weight glutenin subunits at Glu D1 locus. Physiol. Mol. Biol. Plants 21, 447–451. https://doi.org/10.1007/s12298-015-0298-y.

Park, C.S., Kang, C.S., Jeung, J.U., Woo, S.H., 2011. Influence of allelic variations in glutenin on the quality of pan bread and white salted noodles made from Korean wheat cultivars. Euphytica 180, 235–250. https://doi.org/10.1007/s10681-011-0385-2.

Payne, P.I., 1987. Genetics of wheat storage proteins and the effect of allelic variation on bread-making quality. Annu. Rev. Plant Physiol. 38, 141–153. https://doi.org/10.1146/annurev.pp.38.060187.001041.

Peltonen, J., Salopelto, J., Rita, H., 1993. The optimal combination of HMW glutenin subunits coded at gene loci Glu-A1 and Glu-B1 for bread-making quality in Scandinavian wheats. Hereditas 118, 71–78. https://doi.org/10.1111/j.1601-5223.1993.t01-3-00071.x.

Peña-Bautista, R.J., Hernandez-Espinosa, N., Jones, J.M., Guzmán, C., Braun, H.J., 2017. CIMMYT series on carbohydrates, wheat, grains, and health: wheat-based foods: their global and regional importance in the food supply, nutrition, and health. Cereal Foods World 62, 231–249. https://doi.org/10.1094/CFW-62-5-0231.

Rathan, N.D., Mahendru-Singh, A., Govindan, V., Ibba, M.I., 2020. Impact of high and low-molecular-weight glutenins on the processing quality of a set of biofortified common wheat (Triticum aestivum L.) lines. Front. Sustain. Food Syst. 4, 1–9. https://doi.org/10.3389/fsufs.2020.583367.

Ribeiro, M., Rodríguez-Quijano, M., Giraldo, P., Pinto, L., Vázquez, J.F., Carrillo, J.M., Igrejas, G., 2017. Effect of allelic variation at glutenin and puroindoline loci on bread-making quality: favorable combinations occur in less toxic varieties of wheat for celiac patients. Eur. Food Res. Technol. 243, 743–752. https://doi.org/10.1007/s00217-016-2788-8.

Roy, N., Islam, S., Yu, Z., Lu, M., Lafiandra, D., Zhao, Y., Anwar, M., Mayer, J.E., Ma, W., 2020. Introgression of an expressed HMW 1Ay glutenin subunit allele into bread wheat cv. Lincoln increases grain protein content and breadmaking quality without yield penalty. Theor. Appl. Genet. 133, 517–528. https://doi.org/10.1007/s00122-019-03483-1.

Singh, N.K., Shepherd, K.W., 1985. The structure and genetic control of a new class of disulphide-linked proteins in wheat endosperm. Theor. Appl. Genet. 71, 79–92. https://doi.org/10.1007/BF00278258.

Singh, N.K., Shepherd, K.W., Cornish, G.B., 1991. A simplified SDS—PAGE procedure for separating LMW subunits of glutenin. J. Cereal Sci. 14, 203–208. https://doi.org/10.1016/S0733-5210(09)80039-8.

Trethowan, R.M., Peña, R.J., Van Ginkel, M., 2001. The effect of indirect tests for grain quality on the grain yield and industrial quality of bread wheat. Plant Breed. 120, 509–512. https://doi.org/10.1046/j.1439-0523.2001.00651.x.

Wrigley, C., Batey, I., Skylas, D., Sharp, P. (Eds.), 2006. Gliadin and Glutenin: The Unique Balance of Wheat Quality. AACCI Press, St. Paul, USA.