Gluten aggregation properties as a tool for durum wheat quality assessment: a chemometric approach

Silvia Grassi[1], Mariolina Gullì[2], Giovanna Visioli[2], Alessandra Marti[1,*]

[1]Department of Food, Environmental, and Nutritional Sciences (DeFENS), Università degli Studi di Milano, via G. Celoria 2, 20133 Milan, Italy

[2]Department of Chemistry, Life Sciences and Environmental Sustainability, University of Parma, Parco Area delle Scienze 11/a, 43124 Parma, Italy

* Corresponding author: Alessandra Marti; alessandra.marti@unimi.it

Keywords: semolina; gliadin; glutenin; gluten; gluten aggregation

Abbreviations: AggrEn, aggregation energy; GLI, gliadins; GLU, glutenins; GPE, GlutoPeak equivalent; GPU, GlutoPeak unit; HMW-GS, high molecular weight glutenin sub-units; IGG, Index for glutenin/gliadin; IP, Index for protein content; IQR, Interquartile Range; LMW-GS, low molecular weight glutenin sub-units; MT, maximum torque; PCA, Principal Component Analysis; PMT, peak maximum time; PQI, protein quality index; TGP, total gluten proteins; TotEn, total energy.

Abstract: The study aims at assessing the gluten aggregation properties of wholemeal semolina for the discrimination of durum wheat cultivars with different amount of gluten-quality related fractions. At this purpose, a calibration set of 56 durum wheat cultivars harvested in 2015-2016 were categorized in four quality classes considering both the total gluten protein content and the glutenins to gliadins ratio. Quality classes distribution was explored by Principal Component Analysis (PCA) performed on gluten aggregation parameters alone or considering the gluten fractions. PCA allowed high and medium-high quality varieties discrimination from those with low and medium-low quality. Furthermore, PCA permitted to assess the relationship between gluten fractions and their aggregation kinetics, that was confirmed by a validation step consisting of 22 cultivars harvested in 2016-2017. In conclusion, gluten aggregation properties demonstrated to be good indicators of durum wheat semolina quality, and thus a reliable tool for blend preparation in pasta-making process.

1. Introduction

Durum wheat (*Triticum turgidum* L. spp. *durum* Desf.) is the main ingredient of many products of the Mediterranean diet, from pasta to couscous, bulgur, and bread. The availability of the raw material, the amber yellow color, the high carotenoid content, and its protein functionality are the main aspects accounting for the preference of using durum wheat to other cereals to produce staple foods. The technological performance of the raw material resides in its suitability of creating a continuous matrix composed in larger quantity by starch and gluten proteins, which are the main components influencing dough characteristics and the quality of the end product (Sissons, 2008). As the protein fraction is concerned, strong and extensible gluten is needed for optimal baking properties, whereas strong and tenacious gluten is needed for a firm and less sticky pasta (Sissons, 2008). Given similar grain protein content, gluten strength depends on the types and amount of glutenins (GLU) and gliadins (GLI), as well as their ratios (GLU/GLI), and on the amount of insoluble proteins. Specifically, insoluble proteins and GLU contents are related to gluten strength whereas the GLU/GLI ratio defines the balance between elastic and viscous properties of wheat dough (Gupta, Khan, & Macritchie, 1993; Aussenac, Rhazi, & Branlard 2020; De Santis et al., 2020). These traits have a strong genetic component but environmental factors – including temperature and precipitation during grain filling, agronomic managements as nitrogen availability and soil characteristics - could also play an important role in determining both gluten content and the ratio between the different gluten fractions, thus affecting the final quality of the flour produced by the same cultivar (*cv*) in different cultivation years and/or environments (Visioli et al., 2018; Graziano et al., 2019; Branlard et al., 2020). Differences in grain quality among different growing seasons might impact on both farmers - due to variations in wheat prices based on quality parameters - and millers that may encounter difficulties in the standardization of optimal flour blends for processing purposes.

The different extraction methods for the quantification of quality-related gluten fractions is time-consuming and labor-intense, since it requires the extractions of the gluten fractions - typically by modified Osborne fractionation - followed by their quantification by reversed-phase high-performance liquid chromatography or by colorimetric assay (Thanhaeuser, Wieser, & Koehler, 2015). There are other approaches which allow to estimate gluten strength by empiric rheology that can be used for practical purposes. Among them, the Alveograph test is the most used in the durum wheat chain. Indeed, good correlation between semolina strength index (W) and pasta cooking behavior have been assessed (D'Egidio, Mariani, Nardi, Novaro, & Cubadda, 1990; Sissons, 2008; AbuHammad, Elias, Manthey, Alamri, & Mergoum, 2012). However, some issues arise when time and sample size are limiting factors, such as in breeding programs or at the receiving station of milling industry. In this context, the GlutoPeak test has been recently proposed as a fast and reliable small-scale approach for the evaluation of wheat quality. The device measures the aggregation behavior of gluten upon addition of water and high-speed mixing. Although research on GlutoPeak test is still limited, recent studies pointed out that the indices related to gluten aggregation profile – especially when combined with chemometric approaches – are a powerful tool for screening both refined flour (Lu & Seetharaman, 2014; Marti, Ulrici, Foca, Quaglia, & Pagani, 2015a; Issamy, Cao, Falk, Seetharaman, & Bock., 2017) and wholewheat (Malegori, Grassi, Ohm, Anderson, & Marti, 2018; Wang, Hou, Liu, Wang, & Bock 2018). A relationship between gluten aggregation properties and gluten protein fractions has been highlighted in bread wheat (Marti, Augst, Cox, & Koehler, 2015b; Geisslitz, Wieser, Scherf, & Koehler, 2018), encouraging the measurement of gluten aggregation profile for predicting bread volume (Bouachra, Begemann, Aarab, & Hüsken 2017; Geisslitz et al., 2018; Rakita, Dokić, Dapčević Hadnađev, Hadnađev, & Torbica, 2018; Güçbilmez et al., 2019).

Although less research has been carried out on durum wheat, results agree in suggesting the GlutoPeak test as a fast and reliable approach for semolina characterization. Indeed, gluten aggregation kinetic of semolina provides information about its behavior in a dough system, with the advantages of requiring a few minutes of analysis (less than 5 min) and a small amount of sample (less than 10 g) (Marti, Seetharaman, & Pagani, 2013; Marti, Cecchini, D'Egidio, Dreisoerner, & Pagani, 2014a; Sissons, 2016; Sisson & Smit, 2018). However, to the best of our knowledge, it is largely unknown how the quantitative composition of gluten proteins and the relation between the different gluten protein fractions of durum wheat affects the gluten aggregation properties of semolina assessed by the GlutoPeak test.

Therefore, the aim of this study was to assess the ability of the indices related to gluten aggregation behavior to distinguish durum wheat semolina based on quality traits defined considering the amount and type of gluten fractions. This study would lay the groundwork for the prediction of durum wheat semolina quality classes related to the gluten fractions from the measurement of gluten aggregation properties by means of a rapid, small-scale, shear-based method.

2. Materials and Methods

2.1 Materials

Durum wheat (*Triticum turgidum* subsp. *durum* Desf.) cultivars (*cvs)* utilised in this study were grown in the field at the experimental farm of CREA-CI (Research Centre for Cereal and Industrial Crops, Foggia, Italy; 41°27’44.9’’ N 15°30’03.9’’ E). As calibration set, we utilised fifty-six *cvs* grown in 2015-2016 season and as validation set twenty-two *cvs* grown in 2016-2017. The main characteristics of the different *cvs* are shown in Table S1.

2.2 Seed storage protein extraction and quantification

For each *cv*, 35 g of whole grains were milled with Knifetec 1095 (Foss, Hillerød, Denmark) to obtain a fine powder of wholemeal semolina. Gluten protein fractions (i.e., GLI, HMW-GS, and LMW-GS) were extracted with a sequential extraction protocol described by Visioli et al. (2016). In brief, the alcohol soluble GLI fraction was extracted from 30 mg of flour with 1.5 mL of 550 mL/L propan-2-ol with continuous mixing at 65°C for 20 min, followed by centrifugation at 10 000 × g for 5 min. The supernatant was dried in a vacuum centrifuge to obtain the majority of GLI. The procedure was repeated twice to wash the pellet from residues of GLI. The remaining pellet containing the alcohol insoluble fraction of both HMW-GS and LMW-GS was extracted in a 400 µL solution of 550 mL/L propan-2-ol, 0.08 mol/L tris(hydroxymethyl) aminomethane hydrochloric acid (Tris–HCl, pH 8.3) and 10 g/L 1,4-dithiothreitol (DTT, as reducing agent) by incubation for 30 min at 60 °C with continuous mixing. After centrifugation at 14 000 × *g* for 5 min, the HMW-GS and LMW-GS supernatant was collected. HMW-GS were separated from LMW-GS by acetone precipitation. The fractions of GLI, HMW-GS and LMW-GS were dissolved in 500 mL/L acetonitrile with 1 mL/L trifluoroacetic acid and relative quantification was determined by colorimetric Bradford assay (Bio-Rad, Hercules, CA, USA). GLU was calculated as the sum of HMW-GS and LMW-GS fractions, whereas total gluten proteins (TGP) as the sum of GLI and GLU fractions extracted with this methodology. Three technical replicates were performed for each sample.

2.3 Gluten aggregation properties

Gluten aggregation properties of wholemeal semolina were assessed using the GlutoPeak® (Brabender GmbH and Co KG, Duisburg, Germany), according to the method reported by Cecchini, Antonucci, Costa, Marti, & Menesatti (2020). Briefly, semolina (9 g) was dispersed in distilled water (1:1 ratio; adjusting the quantity of sample on 14 g/100 g moisture basis) and mixed at 2750 rpm, keeping the bowl temperature at 35 °C. The main indices considered and automatically calculated by the software (Brabender GlutoPeak v. 2.1.2) were: i) maximum torque (MT; expressed in GlutoPeak Units, GPU), corresponding to the peak occurring as gluten aggregates; ii) peak maximum time (PMT; expressed in s), corresponding to time of maximum torque; iii) aggregation energy (AggrEn; expressed in GlutoPeak Equivalent, GPE), corresponding to the area under the portion of the curve 15 s before and 15 s after MT; iv) total energy (TotEn; expressed in GlutoPeak Equivalent, GPE), corresponding to the area of the curve from the beginning of the test and 15 s after MT.

2.4 Data processing

A quality class was assigned to each sample considering both the TGP and the GLU/ GLI ratio following the equation:

Eq.1, Protein Quality Index (PQI) = (IP × 2) + ((IGG) × 3)

where IP is the index assigned according to the TGP and IGG is the index related to the GLU/GLI ratio according to the thresholds reported in **Table 1.** A greater weight was attributed to GLU/GLI compared to TGP, since it is well known that the balance between dough strength and extensibility, which is an important quality trait in pasta making process, is directly related to the GLU/GLI ratio (Samaan, El-Khayat, Manthey, Fuller, & Brennan, 2006; Sissons, 2008).

Four quality classes were then identified based on the PQI: low quality (PQI ≤ 8) accounting for seven *cvs*, medium-low quality (9 ≤ PQI ≤ 11) accounting for thirty-one *cvs*, medium-high quality (12 ≤ PQI ≤ 13) including thirty-five *cvs*, high quality (PQI > 13) including five *cvs*.

The gluten aggregation indices obtained for the samples harvested in 2015-2016 season were submitted to Principal Component Analysis (PCA) alone or after merging them with the quality-related gluten fractions to build a unique dataset (56 *cvs* x 12 variables). The same procedure was followed for the creation of a validation set composed by the 22 *cvs* harvested in the following 2016-2017 season to be used to validate the PCA model, thus assessing the consistency of this approach along different growing seasons. By a simple data visualization, the PCA based on gluten aggregation indices will provide an understanding of the relationship between the calculated indices, their load in the sample distribution, and their contribution in durum wheat semolina quality classes grouping (Jolliffe, 2003). The PCA built considering the gluten aggregation indices together with the quality-related gluten fractions will permit the comparison of sample patterns with gluten aggregation indices -PCA and to assess the relationship between protein quality parameters and GlutoPeak indices.

Data analysis was carried out in Matlab environment (v. 2017b, Mathworks, Inc., Natick, MA), using the PLS toolbox (v. 8.5, Eigenvector Research, Inc., Seattle, WA, USA).

3. Results and Discussion

3.1 Quantification of gluten protein fractions

Fig. 1 reports as boxplots the amounts of TGP, GLI, HMW-GS, LMW-GS, GLU, and both the GLU/GLI and HMW-GS/LMW-GS ratios measured in all samples in the two years, while the detailed data are reported in **Table S2**.

The samples harvested in 2015-2016 (**Fig. 1 A-B**) were characterised by a TGP between 26.24 (Platani *cv.*) and 40.56 mg/g (Canizzo *cv.*), with an Interquartile Range (IQR, between Q1 and Q3) of 33.82 - 36.50 mg/g. Since GLI are the most abundant gluten protein fraction, Platani and Canizzo *cvs* also exhibited the lowest (11.24 mg/g) and highest (22.47 mg/g) GLI content, respectively. The IQR for GLI content was 17.38 - 19.39 mg/g (**Table S2**).

The IQR for GLU content was 16.18 - 17.63 mg/g with minimum and maximum of 14.49 and 18.80 mg/g, for Saadi and Neolatino *cvs*, respectively. Considering the LMW-GS, their IQR values ranged from 8.27 to 9.35 mg/g with minimum and maximum of 7.55 (Aureo *cv*) and 10.22 (Rusticano *cv*) mg/g, respectively. Finally, the IQR for HMW-GS was 7.65 - 8.33 mg/g with minimum and maximum of 6.08 and 9.30 mg/g, for Platani and Iride *cvs*, respectively (**Table S2**). The IQRs for GLU/GLI and HMW-GS/LMW-GS ratios were 0.84-0.99, 1.01-1.19, and 0.85-0.95, respectively. Specifically, Colosseo *cv.* exhibited the lowest GLU/GLI ratio and Imhotep *cv.* the highest one; whereas Platani and Cappelli *cvs* showed the lowest and the highest HMW-GS/LMW-GS ratio, respectively (**Table S2**).

Even if in the 2016-2017 growing season a small number of samples was analysed (n=22), the variability recorded was higher for most of the considered parameters (**Fig. 1 C-D**). Indeed, TGP content was between 23.68 (Altar *cv*) and 43.36 mg/g (Kofu *cv*), with an IQR of 29.10 - 39.23 mg/g. The IQR for GLU content was 13.33 - 8.96 mg/g with minimum and maximum of 11.77 (Valerio *cv*) and 21.48 mg/g (Aureo *cv*), respectively (**Table S2**). The IQR for GLI content was 13.13 - 20.37 mg/g with minimum and maximum similar to the one observed for samples harvested in 2016, i.e. 11.72 (Altar84 *cv*) and 22.91 mg/g (Lesina *cv*), respectively. The IQR for LMW-GS content was 5.54 - 10.66 mg/g with minimum and maximum of 4.53 (Valerio *cv*) and 11.80 mg/g (Avispa *cv*), respectively (**Table S2**). HMW-GS content for samples harvested in 2017 was similar to the one observed for 2016 samples, indeed the IQR for HMW-GS content of 2017-samples was 7.31 - 8.68 mg/g with minimum and maximum of 5.40 (Dylan *cv.*) and 10.87 mg/g (Baio *cv*), respectively. The IQRs for total GLU/GLI and HMW-GS/LMW-GS ratios were 0.88 - 1.25, 0.80 - 1.14 and 0.78 - 1.42, respectively (**Table S2**).

All the samples were grown in the same location and under the standardized growing conditions, thus differences in the content and composition of gluten fractions are genotype dependent. The effect of environment was clear when comparing the two growing seasons. Indeed, environmental factors typically account for up to two-thirds of the variation in wheat protein content and composition with the remaining one-third being controlled by genetic factors with additional interaction between the two factors (Shewry, 2009). The effect of the growing season on either the content or the composition of gluten fractions was particularly significant for Aureo *cv.* This *cv.* is widely considered among the high quality wheat for pasta production (Visioli et al., 2018), as also shown by the data related to the 2016-2017 growing season; however, in the previous season (2015-2016) Aureo *cv* exhibited a significant decrease in GLU content, which negatively affected its classification as medium-low quality sample (**Table 2**). The 2016-2017 growing season was favorable also for Cappelli, Iride, Karalis, and Tablur *cvs*, which shifted from the medium-low (in 2015-2016) to the medium-high (in 2016-2017) quality class, while very few *cvs* among those tested in the two years – i.e., Altar.84 (medium-low quality class), Avispa, Biensur, Kofa and PR22D89 (medium-high quality class) - showed the same gluten protein quality traits in the two seasons. As previously reported, the growing season 2015-2016 in Foggia (Italy) had very different weather conditions as compared to the growing season 2016-2017 (Ronga et al., 2020). Specifically, temperatures during the grain filling period in 2016 were below 27 °C, unusual for the climate of Southern Italy, while in 2017 they exceeded 30 °C, and the pre- and post-anthesis periods had higher rainfall in 2016 than in 2017. At this purpose, samples like Aureo *cv* selected to grow exclusively in these areas might have experienced a different balance between the different gluten fractions thus varying the GLU/GLI ratio which was considered in our model to have a greater weight as quality indicators.

The environmental conditions of the second year are those typical of the Mediterranean climate, and can determine a certain level of stress (both drought and heat stress) during grain filling which may cause a higher aggregation level of gluten subunits, and it was suggested that moderately high temperatures have a positive effect on dough properties(Flagella, Giuliani, Giuzio, Volpi, & Masci 2010).

3.2 Gluten aggregation properties

Fig. 2 shows the gluten aggregation profile of the samples with the greatest differences in the amount and type of gluten fractions, and thus belonging to different quality classes (**Table 2**). Specifically, Saadi *cv* – which belongs to the low-quality class and contains the lowest GLU content – is characterized by poor aggregation properties (i.e., low PMT and MT, thus low energy values). On the opposite side for GLU content and overall quality, the gluten aggregation of Neolatino *cv* is characterized by longer PMT and higher MT, resulting in high energy value. Although belonging to the same quality class (i.e., medium-low), Platani and Canizzo *cvs* exhibited the lowest and the highest GLI content, respectively.

Gluten functionality is defined not only by the amount of GLI and GLU *per se*, but also by their ratio: GLU/GLI defines the balance between the elastic and viscous properties of the matrix. Usually, dough exhibiting high elasticity gave a pasta with high firmness (Marti, D'Egidio, Dreisoerner, Seetharaman, & Pagani, 2014b). Indeed, the ideal raw material for pasta production should be strong enough to resist to the mechanical stress occurring during the extrusion step, and elastic to counteract starch swelling during cooking, resulting in a product with high firmness and low stickiness and bulkiness (Marti, D'Egidio, & Pagani, 2016). The GLU/GLI of the sample reported in **Fig. 2** followed the order Saadi *cv* (0.78) < Canizzo *cv* (0.81) < Neolatino *cv* (1.04) < Platani *cv* (1.33).

Besides differences in the indices, samples showed unique gluten aggregation profiles, suggesting differences in the way gluten proteins interact forming the network. During the test, the sample is mixed at high speed (2750 rpm) and as GLI and GLU interact, the gluten is formed, and the device registers an increase in torque. Further mixing destroys the network, and the torque curve declines (**Fig. 2**). Generally, GlutoPeak profiles characterized by low PMT, MT, and/or energy (i.e., area under the curve) suggest poor gluten aggregation capacity, resulting in a structure with a scarce capacity to keep starch granules inside the protein matrix (Marti et al., 2014a). In other words, durum wheat semolina presenting high energy values and/or high PMT gave a product characterized by low stickiness and bulkiness (Marti et al., 2014a; Sissons, 2016).

Fig. 3 reports the boxplots for the indices related to gluten aggregation profile as assessed by the GlutoPeak test. The indexes related to gluten aggregation referred to samples grown in 2015-2016 season (**Fig. 3 A-B-C**) were characterised by IQRs of 92.3 - 129.0 s for PMT, 35.5 - 40.8 GPU for MT, 1630.0 - 2164.5 GPE for TotEn, and 942.5 - 1079.0 GPE for the AggrEn (**Table S2**). The faster aggregation (i.e., lowest PMT) was observed for Saadi *cv* (50.0 s), that also showed the lowest MT (23.2 GPU), AggrEn (590.0 GPE), and TotEn values (792.0 GPE) (**Table S2**). On the other hand, Altar *cv* showed the highest PMT and TotEn, whereas Ariosto *cv* the highest values for both MT and AggrEn (**Table S2**).

Considering the profiles of the samples grown in 2016-2017 (Fig. 3 D-E-F), the gluten aggregation indexes were characterised by IQRs of 44.00 - 69.00 s for PMT, 44.30 - 57.80 GPU for MT, 1971.0 - 2479.0 GPE for TotEn and 1279.0-1500.0 GPE for the AggrEn (Table S2). Finally, comparing the two growing seasons (**Table S2**), PMT showed a lower variability in 2017 samples (IQR = 44.00 - 69.00 s), MT was characterized by values (IQR = 44.30 - 57.80 GPU) higher than the one observed in 2016 (IQR = 32.50 - 40.75 GPU) as well as TotEn (IQR2017 = 1971.0 - 2479.0 GPE) and AggrEn (IQR2017 = 1379.0 - 1500.0 GPE).

3.3 Principal Component Analysis

A PCA was built by considering the indices provided by the gluten aggregation profiles of 56 durum wheat *cvs* of the 2015-2016 growing season (**Fig. 4**), further a validation procedure was performed with the 22 *cvs* harvested in 2016-2017 season. Results seem to be promising for grouping the samples according to the protein quality classes defined using the gluten protein fractions. PC1 (46.61% of the total variance) discriminates between samples belonging to the low quality class to those belonging to the high quality class: the higher the PC1 values, the higher the sample quality. Indeed, most of the samples belonging to the low and medium-low quality classes assume negative PC1 values, whereas those in the high and medium-high quality classes are characterised by positive PC1 scores (**Fig. 4A**). However, Tresor and Latino *cvs* (grown in 2015-2016), even if considered of medium-high quality, assumed negative PC1 scores, thus being in the II quadrant where low and medium-low quality samples are located (**Fig. 4C**). Their similarity is linked to their gluten aggregation properties, in particular both *cvs* are characterised by low AggrEn values, which assumed the highest PC1 values in the loadings plot (**Fig. 4B**). The same for Simeto *cv* (grown in 2016-2017) (**Fig. 4D**). A positive correlation between the energy (i.e., area under the curve) and gluten strength has been shown in durum wheat semolina (Marti et al., 2014a). Also Altar *cv* (grown in 2015-2016) shows a different behaviour from the *cvs* of medium-low quality, as it assumes high positive PC1 and PC2 values (**Fig. 4C**) due to the high value of PMT (482 s) and thus TotEn (4369 GPE), similarly to the neighbour Iride *cv* (2016-2017) (Fig. 4D), characterised by PMT of 442 s and TotEn of 3680 GPE.

Worth to be mentioned, Aureo and Kofu *cvs* harvested in 2016-2017 (**Fig. 4D**) resulted quite diverse from the other samples due to high values of AggrEn (8470 and 9074 s) and MT (76.9 and 77.7 GPU) influencing their displacement in the IV quadrant, where these variables are located (Fig. 4B). Indeed, the loadings plot (**Fig. 4B**) highlights that AggrEn and MT provide correlated information, being close to each other and are responsible for the discrimination of medium-high and high quality *cvs* from the other samples. TotEn plays also a relevant role in the discrimination of medium-high and high quality *cvs* from the rest of quality classes but with lower rate, which is even lower for PMT. The latter is the gluten aggregation index with lower weight in discriminating quality groups with a PC1 loading close to zero (0.28). The role of PMT in defining wheat quality has not been clearly defined yet. Indeed, some authors suggested a significant negative correlation between PMT and dough resistance to mixing (Sissons, 2016; Geisslitz et al., 2018; Sissons & Smit, 2018), others pointed out that among the GlutoPeak indices, the PMT was the less effective in discriminating gluten strength in either common (Marti et al., 2015a) or durum (Marti et al., 2014a) wheat. Although differences in samples, growing environment, and particle size would affect absolute values, this would not explain the difference in correlations among previous studies. Another possibility is that the applied method (i.e., sample: water ratio) may influence the data from the GlutoPeak test.

A second PCA model was calibrated considering all the variables. i.e. quality parameters of TGP and the different gluten fractions combined with the indexes obtained from the GlutoPeak test, to identify the most suitable parameters to predict durum wheat quality and the relation between the gluten protein fractions and gluten aggregation properties (**Fig. 5**). The PCA built with the data corresponding to the varieties harvested in 2015-2016 showed a good separation of quality classes along the PC2, explaining almost the 35% of the total variance of the data (**Fig. 5A**). Indeed, varieties identified as high quality by the combined index PQI obtained high positive PC1 values, by reducing the PQI it was observed a reduction of PC2 scores. This behaviour is consistent in the validation set, i.e. when the data corresponding to samples harvested in 2016-2017 were projected in the new space (PC1 *vs* PC2): the higher the PQI the higher the PC2 score (**Fig. 5A**). Apart from the displacement of samples along PC2, it is possible to notice few samples assuming extreme positive and negative PC1 scores. This is the case of two samples (Imhotep and Platani *cvs*) attributed to medium-low quality class among the samples grown in 2015-2016 and assuming PC1 scores lower than -7 (**Fig. 5C**). Specifically, as shown in Table S2, Imhotep and Platani *cvs* were characterized by high GLU/GLI ratio, i.e. 1.33 and 1.39 which are out from the IQR for 2015-2016, and the lowest HMW-GS/LMW-GS ratio for 2015-2016 season, i.e. 0.68. On the other hand, data referred to Canizzo *cv* grown in 2015-2016 - characterized by the highest TGP and GLI content - resulted in a displacement of this sample to high positive PC1 values (**Fig. 5C**). Also, Colosseo *cv* appears distant from the rest of the medium-low quality *cvs*, reaching low PC2 values, similarly to *cvs* belonging to low quality class (**Fig. 5C**). Indeed, Colosseo *cv.* showed the lowest GLU/GLI ratio (**Table S2**).

As far as the data of the 2016-2017 growing season (**Fig. 5D**), it is possible to notice that those samples are more spread in the PC1-PC2 space than the sample related to the 2015-2016 growing season (**Fig. 5C**). This is linked to the high variability observed for most of the considered parameters, as previously discussed. Most of the medium-high quality *cvs* grown in 2016-2017 are positioned in the II quadrant of the plane, where the GLU/GLI ratio has a higher influence. Indeed, samples in this quality class were characterised by GLU/GLI ratio in the IQR of 1.09-1.34 mg/g. On the opposite side, in the IV quadrant, are isolated two (K26 and Mimmo *cvs*) out of three *cvs* identified as low quality among the 2016-2017 samples (**Fig. 5D**). Their displacement is highly influenced by the GLI/GLU ratio.

A detailed analysis of the loadings plot (**Fig. 5B**), in comparison with the corresponding scores plot, allowed to understand the relationships between variables and samples and among the variables. The variables contributing the most to the displacement of samples to high positive PC2 values are GLU, LMW-GS and HMW-GS (PC2 loadings higher than 0.4), followed by AggrEn, MT, GLU/GLI ratio, TotEn, and TGP. Whereas the variable responsible for samples displacement towards negative PC1 is the ratio GLU/GLI. On the opposite side, GLI, GLI/GLU and TGP are responsible for high positive PC1 scores.

These findings agree with previous studies on common wheat (Marti et al., 2015b). Specifically, total gluten protein and gliadin content in winter wheat varieties was positively correlated to MT, while glutenins was positively correlated to the area under the curve (Marti et al., 2015b). Gluten aggregation indices were also related to glutenins and its fractions when a set of common wheat, spelt, durum wheat, emmer, and einkorn was analysed (Geisslitz et al., 2018). In addition, in a previous study on durum wheat semolina, a positive correlation was measured between the amount of insoluble proteins and the gluten strength indicators with different approaches, including the GlutoPeak test, suggesting that greater quantities of HMW-GS or unextractable glutenin are associated with stronger dough (Marti et al., 2015b; Sissons, 2016).

4. Conclusions

The gluten aggregation properties of wholemeal semolina are related to the amount of the different gluten protein fractions and their ratio, but they are more easily accessible compared to the labor-intensive extraction and quantitation assay. Thus, the fast and easy GlutoPeak test is a useful tool to predict the quality of durum wheat semolina. Indeed, PCA on gluten aggregation indices demonstrated the possibility of separating high and medium-high varieties from low and medium-low quality varieties, which is mainly related to AggrEn and MT information. The possibility of a classification based on the gluten aggregation indices should be further investigated with supervised classification approaches, which require more balanced classes' numerosity. This will need the inclusion of more samples falling in the high and low quality classes. In addition, investigating the gluten aggregation properties of durum wheat varieties with the GlutoPeak test might provide milling industry with a fast and reliable standardization tool for blend preparation to be used in pasta-making process.

Declaration of competing interest

None.

Acknowledgements

The authors thank Alice Bonini and Gaetano Cardone (Università degli Studi di Milano) for excellent technical assistance.

References

AbuHammad, W. A., Elias, E. M., Manthey, F. A., Alamri, M. S., & Mergoum, M. (2012). A comparison of methods for assessing dough and gluten strength of durum wheat and their relationship to pasta cooking quality. *International Journal of Food Science & Technology, 47(12)*, 2561-2573. https://doi.org/10.1111/j.1365-2621.2012.03135.x

Aussenac, T., Rhazi, L., & Branlard, G. (2020). Molecular weight distribution of polymeric proteins in wheat grains: The rheologically active polymers. *Foods*, *9(11)*, 1675. https://doi.org/10.3390/foods9111675

Bouachra, S., Begemann, J., Aarab, L., & Hüsken, A. (2017). Prediction of bread wheat baking quality using an optimized GlutoPeak®-Test method. *Journal of Cereal Science, 76,* 8-16. https://doi.org/10.1016/j.jcs.2017.05.006

Branlard, G., Faye, A., Rhazi, L., Tahir, A., Lesage, V., & Aussenac, T. (2020). Genetic and environmental factors associated to glutenin polymer characteristics of wheat. *Foods*, *9(5)*, 683. https://doi.org/10.3390/foods9050683

Cecchini, C., Antonucci, F., Costa, C., Marti, A., & Menesatti, P. (2020). Application of near-infrared handheld spectrometers to predict semolina quality. *Journal of the Science of Food and Agriculture, 101(1)*, 151-157. https://doi.org/10.1002/jsfa.10625

De Santis, M. A., Cunsolo, V., Giuliani, M. M., Di Francesco, A., Saletti, R., Foti, S., & Flagella, Z. (2020). Gluten proteome comparison among durum wheat genotypes with different release date. *Journal of Cereal Science*, *96*, 103092. https://doi.org/10.1016/j.jcs.2020.103092

D'Egidio, M. G., Mariani, B. M., Nardi, S., Novaro, P., & Cubadda, R. (1990). Chemical and technological variables and their relationships: A predictive equation for pasta cooking quality. *Cereal Chemistry, 67(3)*, 275-281.

Flagella, Z., Giuliani, M. M., Giuzio, L., Volpi, C., & Masci, S. (2010). Influence of water deficit on durum wheat storage protein composition and technological quality. *European Journal of Agronomy, 33(3)*, 197-207. https://doi.org/10.1016/j.eja.2010.05.006

Geisslitz, S., Wieser, H., Scherf, K. A., & Koehler, P. (2018). Gluten protein composition and aggregation properties as predictors for bread volume of common wheat, spelt, durum wheat, emmer and einkorn. *Journal of Cereal Science, 83*, 204-212. https://doi.org/10.1016/j.jcs.2018.08.012

Graziano, S., Marando, S., Prandi, B., Boukid, F., Marmiroli, N., Francia, E., Pecchioni, N., Sforza, S., Visioli, G., & Gullì, M. (2019). Technological quality and nutritional value of two durum wheat varieties depend on both genetic and environmental factors. *Journal of Agriculture and Food Chemistry, 67(8)*, 2384-2395. https://doi.org/10.1021/acs.jafc.8b06621

Güçbilmez, Ç. M., Şahin, M., Akçacık, A. G., Aydoğan, S., Demir, B., Hamzaoğlu, S., Gür, S., & Yakışır, E. (2019). Evaluation of GlutoPeak test for prediction of bread wheat flour quality, rheological properties and baking performance. *Journal of Cereal Science, 90*, 102827. https://doi.org/10.1016/j.jcs.2019.102827

Gupta, R. B., Khan, K., & Macritchie, F. (1993). Biochemical basis of flour properties in bread wheats. I. Effects of variation in the quantity and size distribution of polymeric protein. *Journal of Cereal Science, 18(1)*, 23-41. https://doi.org/10.1006/jcrs.1993.1031

Issarny, C., Cao, W., Falk, D., Seetharaman, K., & Bock, J. E. (2017). Exploring functionality of hard and soft wheat flour blends for improved end-use quality prediction. *Cereal Chemistry, 94(4)*, 723-732. https://doi.org/10.1094/CCHEM-09-16-0248-R

Jolliffe, I. T. (2003). Principal component analysis. *Technometrics*, *45(3),* 276.

Lu, Z., & Seetharaman, K. (2014). Suitability of Ontario-grown hard and soft wheat flour blends for noodle making. *Cereal Chemistry, 91(5)*, 482-488. https://doi.org/10.1094/CCHEM-10-13-0218-R

Malegori, C., Grassi, S., Ohm, J. B., Anderson, J., & Marti, A. (2018). GlutoPeak profile analysis for wheat classification: skipping the refinement process. *Journal of Cereal Science, 79*, 73-79. https://doi.org/10.1016/j.jcs.2017.09.005

Marti, A., Seetharaman, K., & Pagani, M. A. (2013). Rheological approaches suitable for investigating starch and protein properties related to cooking quality of durum wheat pasta. *Journal of Food Quality, 36(2)*, 133-138. https://doi.org/10.1111/jfq.12015

Marti, A., Cecchini, C., D'Egidio, M. G., Dreisoerner, J., & Pagani, M. A. (2014a). Characterization of durum wheat semolina by means of a rapid shear-based method. *Cereal Chemistry, 91(6)*, 542-547. https://doi.org/10.1094/CCHEM-10-13-0224-R

Marti, A., D’Egidio, M. G., Dreisoerner, J., Seetharaman, K., & Pagani, M. A. (2014b). Temperature-induced changes in dough elasticity as a useful tool in defining the firmness of cooked pasta. *European Food Research Technology, 238(2)*, 333-336. https://doi.org/10.1007/s00217-013-2117-4

Marti, A., Ulrici, A., Foca, G., Quaglia, L., & Pagani, M. A. (2015a). Characterization of common wheat flours (*Triticum aestivum* L.) through multivariate analysis of conventional rheological parameters and gluten peak test indices. *LWT- Food Science and Technology*, *64(1)*, 95-103. https://doi.org/10.1016/j.lwt.2015.05.029

Marti, A., Augst, E., Cox, S., & Koehler, P. (2015b). Correlations between gluten aggregation properties defined by the GlutoPeak test and content of quality-related protein fractions of winter wheat flour. *Journal of Cereal Science, 66*, 89-95. https://doi.org/10.1016/j.jcs.2015.10.010

Marti, A., D'Egidio, M. G., & Pagani, M. A. (2016). Pasta: quality testing methods. In C. Wrigley, H. Corke, K. Seetharaman, & J. Faubion, J. (Eds.), *Encyclopedia of food grains* (pp. 161-165) Vol 3. Academic Press.

Rakita, S., Dokić, L., Dapčević Hadnađev, T., Hadnađev, M., & Torbica, A. (2018). Predicting rheological behavior and baking quality of wheat flour using a GlutoPeak test. *Journal of Texture Studies*, *49(3)*, 339-347. https://doi.org/10.1111/jtxs.12308

Ronga, D., Laviano, L., Catellani, M., Milc, J., Prandi, B., Boukid, F., Sforza, S. Dossena, A. Graziano, S. Gulli, M., Visioli, G., Marmiroli, N., De Vita, P., Pecchioni, N., & Francia, E. (2020). Influence of environmental and genetic factors on content of toxic and immunogenic wheat gluten peptides. *European. Journal of Agronomy*, *118*, 126091. https://doi.org/10.1016/j.eja.2020.126091

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

Shewry, P. R. (2009). Wheat. *Journal of Experimental Botany, 60(6)*, 1537-1553. https://doi.org/10.1093/jxb/erp058

Sissons, M. (2008). Role of durum wheat composition on the quality of pasta and bread. *Food, 2(2)*, 75-90.

Sissons, M. (2016). GlutoPeak: A breeding tool for screening dough properties of durum wheat semolina. *Cereal Chemistry, 93(6)*, 550-556. https://doi.org/10.1094/CCHEM-03-16-0063-R

Sissons, M., & Smit, J. (2018). Small-scale methods to assess the gluten properties of durum wheat. *Cereal Chemistry, 95(3)*, 456-468. https://doi.org/10.1002/cche.10048

Thanhaeuser, S. M., Wieser, H., & Koehler, P. (2015). Spectrophotometric and fluorimetric quantitation of quality-related protein fractions of wheat flour. *Journal of Cereal Science*, *62*, 58-65. https://doi.org/10.1016/j.jcs.2014.12.010

Visioli, G., Bonas, U., Dal Cortivo, C., Pasini, G., Marmiroli, N., Mosca, G., & Vamerali, T. (2018). Variations in yield and gluten proteins in durum wheat varieties under late-season foliar versus soil application of nitrogen fertilizer in a northern Mediterranean environment. *Journal of Science and Food Agriculture, 98(6)*, 2360-2369. https://doi.org/10.1002/jsfa.8727

Visioli, G., Galieni, A., Stagnari, F., Bonas, U., Speca, S., Faccini, A., Pisante, M., & Marmiroli, N. (2016). Proteomics of durum wheat grain during transition to conservation agriculture. *PLoS One*, *11(6)*, e0156007. https://doi.org/10.1371/journal.pone.0156007

Wang, J., Hou, G. G., Liu, T., Wang, N., & Bock, J. (2018). GlutoPeak method improvement for gluten aggregation measurement of whole wheat flour. *LWT-Food Science and Technology, 90*, 8-14. https://doi.org/10.1016/j.lwt.2017.11.059

Figure 1. Gluten protein fractions of durum wheat cultivars grown in 2015-2016 (A, B) and in 2016-2017 (C, D). A) and C) Total gluten proteins (TGP), glutenins (GLU), gliadins (GLI), low molecular weight (LMW) and high molecular weight (HMW) glutenin sub-units (GS); B) and D) GLU/GLI and HMW-GS/LMW-GS ratios.

Figure 2. Gluten aggregation properties of selected durum wheat cultivars.

Figure 3. Indices related to gluten aggregation properties (as assessed by the GlutoPeak test) of durum wheat cultivars grown in 2015-2016 (A, B, C) and in 2016-2017 (D, E, F). A) and D) Peak Maximum Time (PMT); B) and E) Maximum Torque (MT); C) and F) Aggregation Energy (AggrEn) and Total Energy (TotEn).

Figure 4. Principal Component Analysis of the indexes obtained from the GlutoPeak test: A) scores plot of the calibration and validation sets projected together; B) loadings plot (PMT, Peak Maximum Time; MT, Maximum Torque; TotEn, Total Energy; AggrEn, Aggregation Energy); C) scores plot of the calibration set alone; D) scores plot of the validation set alone.

Figure 5. Principal Component Analysis of the quality parameters of grain gluten protein content (TGP) and gluten fractions combined with the indexes obtained from the GlutoPeak test: A) scores plot of the calibration and validation sets projected together; B) loadings plot (glutenins, GLU; gliadins, GLI; Low molecular weight glutenin sub-units, LMW-GS and High molecular weight glutenin sub-units, HMW-GS; PMT, Peak Maximum Time; MT, Maximum Torque; TotEn, Total Energy; TPG, Total gluten proteins, AggrEn, Aggregation Energy); C) scores plot of the calibration set alone; D) scores plot of the validation set alone.

Table 1. Definition of IP (Index for the total gluten proteins) and IGG (Index for glutenin/gliadin).

Parameter	Range	Quality class	IP
TGP (mg/g flour)	20 – 27	Low	1
	27.1 – 35	Medium	2
	>35	High	3
			IGG
GLU/GLI	0.7-0.85	Low	1
	0.86-1	Medium	2
	>1	High	3

TGP, total gluten proteins; GLU/GLI, glutenin to gliadin ratio

Table 2. Quality characteristics of total gluten proteins (TGP) and different gluten fractions (glutenins, GLU; gliadins, GLI; low molecular weight (LMW) and high molecular weight (HMW) glutenin sub-units (GS). GlutoPeak indexes obtained from the GlutoPeak test (PMT, Peak Maximum Time; MT, Maximum Torque; TotEn, Total Energy; AggrEn, Aggregation Energy). Data are presented according to the quality class previously defined.

Quality class		TGP (mg/g)	GLU (mg/g)	GLI (mg/g)	LMW-GS (mg/g)	HMW-GS (mg/g)	GLU/GLI	HMW-GS/ LMW-GS	PMT (s)	MT (GPU)	TotEn (GPE)	AggrEn (GPE)
Growing Season 2015 - 2016												
Low (n=4)	Min	32.71	14.49	18.02	7.85	6.40	0.78	0.79	50.0	23.2	792.0	590.0
	Q1	33.03	14.64	18.19	8.00	6.58	0.79	0.82	89.8	30.9	1540.5	753.5
	median	33.24	14.89	18.45	8.07	6.73	0.80	0.82	110.0	34.5	1854.5	843.0
	Q3	33.61	15.12	18.79	8.14	6.95	0.82	0.85	127.5	36.8	1961.3	926.3
	max	34.45	15.23	19.22	8.27	7.38	0.83	0.94	159.0	40.6	2088.0	1071.0
Medium-low (n=26)	min	26.24	15.00	11.24	7.55	6.08	0.72	0.68	51.0	24.7	1074.0	618.0
	Q1	34.16	16.01	17.64	8.24	7.56	0.84	0.87	87.0	35.9	1542.5	939.5
	median	35.27	16.32	19.15	8.44	7.88	0.85	0.92	111.0	39.5	1826.5	1032.0
	Q3	36.32	16.60	20.14	8.71	8.14	0.92	0.98	123.0	40.8	2017.8	1118.0
	max	40.56	18.09	22.47	9.50	9.30	1.33	1.62	482.0	55.6	4369.0	1382.0
Medium-high (n=23)	min	27.07	15.73	11.34	7.97	6.39	0.88	0.68	80.0	23.5	1066.0	620.0
	Q1	33.81	17.04	16.38	8.81	8.18	0.91	0.84	93.0	36.7	1692.0	980.0
	median	35.65	17.45	18.47	9.28	8.24	0.98	0.89	120.0	38.6	1989.0	1048.0
	Q3	36.54	17.89	19.06	9.50	8.45	1.06	0.94	141.0	39.6	2346.0	1073.0
	max	38.20	18.56	20.25	10.22	8.96	1.47	1.04	442.0	47.2	3680.0	1278.0
High (n=3)	min	36.13	18.15	17.98	9.76	8.29	1.01	0.84	89.0	26.8	1279.0	758.0
	Q1	36.50	18.41	18.02	9.81	8.49	1.01	0.85	94.5	32.5	1539.0	893.0
	median	36.87	18.67	18.06	9.86	8.68	1.01	0.87	100.0	38.1	1799.0	1028.0
	Q3	37.00	18.74	18.27	9.92	8.86	1.03	0.90	114.0	41.3	1982.0	1066.5
	max	37.14	18.80	18.47	9.99	9.04	1.04	0.93	128.0	44.4	2165.0	1105.0

Growing Season 2016 - 2017

Low (n=3)	min	27.22	11.77	15.44	4.53	7.25	0.58	1.50	49.0	41.8	2031.0	1216.0
	Q1	29.91	12.00	17.91	4.60	7.37	0.59	1.55	50.5	44.4	2056.0	1298.0
	median	32.60	12.23	20.37	4.67	7.49	0.60	1.60	52.0	46.9	2081.0	1380.0
	Q3	33.21	12.34	20.87	4.82	7.52	0.68	1.61	56.0	51.3	2218.0	1381.0
	max	33.82	12.46	21.36	4.97	7.56	0.76	1.62	60.0	55.6	2355.0	1382.0
Medium-low (n=5)	min	23.68	11.96	11.72	4.65	6.67	0.80	0.70	39.0	44.0	1981.0	1279.0
	Q1	25.38	13.22	12.05	5.51	7.31	0.80	0.84	43.0	44.3	2186.0	1288.0
	median	25.93	13.33	12.71	5.54	7.68	1.02	1.39	54.0	47.8	2259.0	1373.0
	Q3	36.47	16.25	20.22	9.57	7.81	1.04	1.42	64.0	55.1	2318.0	1507.0
	max	41.24	18.33	22.91	9.99	8.34	1.11	1.57	65.0	64.1	2725.0	1635.0
Medium-high (n=12)	min	27.36	15.18	12.18	6.03	5.40	0.88	0.50	33.0	31.0	1569.0	823.0
	Q1	30.77	17.30	13.35	9.74	7.22	1.09	0.69	41.5	43.0	1696.5	1218.5
	median	32.15	18.17	13.93	10.19	8.64	1.25	0.87	55.0	49.7	1971.0	1322.0
	Q3	37.01	19.37	17.73	10.67	9.63	1.34	0.92	97.5	61.7	2515.0	1450.0
	max	43.36	20.63	22.89	11.80	10.87	1.45	1.80	165.0	77.7	16125.0	9074.0
High (n=2)	min	39.23	20.19	19.05	11.29	8.45	1.06	0.72	67.0	50.3	2102.0	1323.0
	Q1	39.66	20.51	19.14	11.40	8.88	1.07	0.77	88.3	57.0	5349.8	3109.8
	median	40.08	20.83	19.24	11.52	9.32	1.08	0.81	109.5	63.6	8597.5	4896.5
	Q3	40.50	21.16	19.34	11.63	9.75	1.09	0.86	130.8	70.3	11845.3	6683.3
	max	40.92	21.48	19.44	11.74	10.19	1.11	0.90	152.0	76.9	15093.0	8470.0

Min, minimum; max, maximum; Q1, lower quartile; Q2, upper quartile. GPU, GlutoPeak Unit; GPE, GlutoPeak Equivalent; n, number of samples belonging to each quality class.

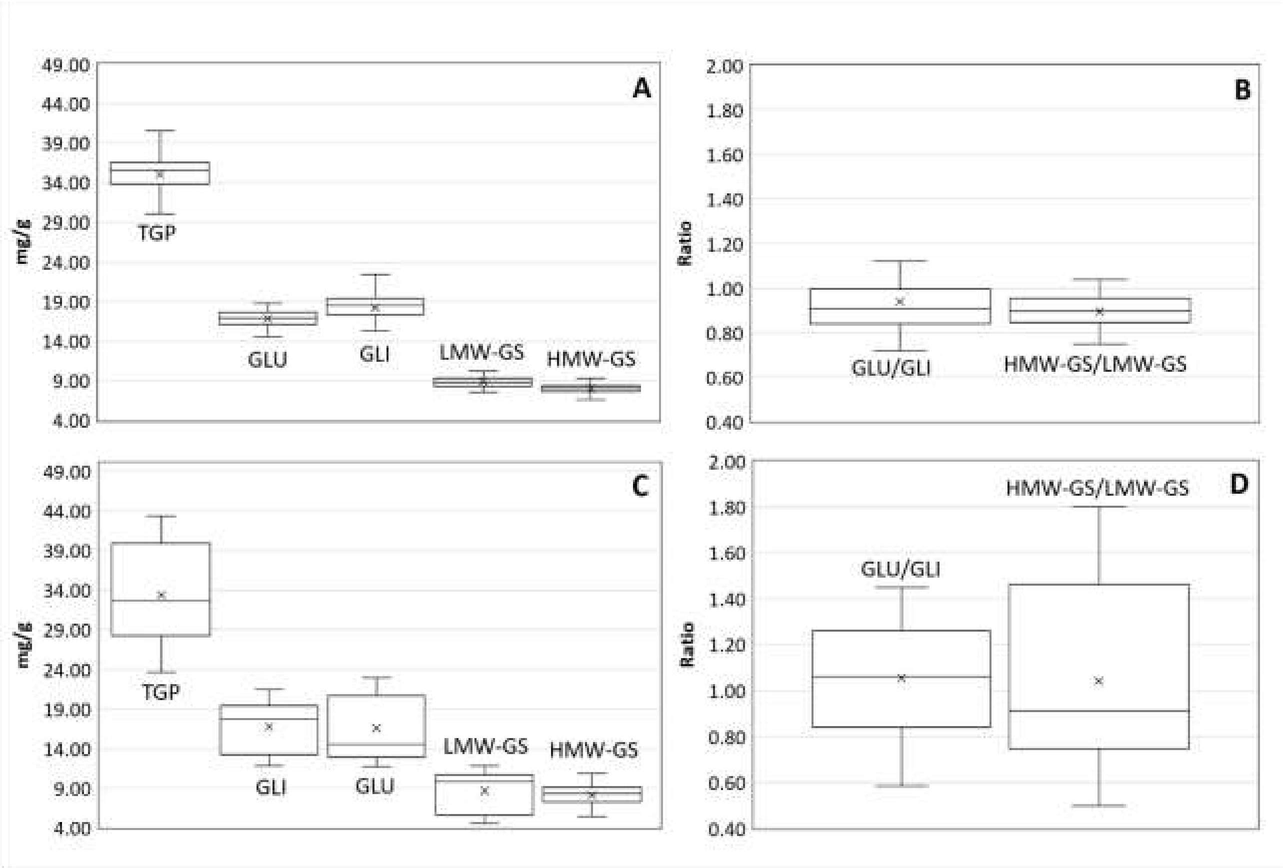

Figure 1. Gluten protein fractions of durum wheat cultivars grown in 2015-2016 (A, B) and in 2016-2017 (C, D). A) and C) Total gluten proteins (TGP), glutenins (GLU), gliadins (GLI), low molecular weight (LMW) and high molecular weight (HMW) glutenin sub-units (GS); B) and D) GLU/GLI and HMW-GS/LMW-GS ratios.

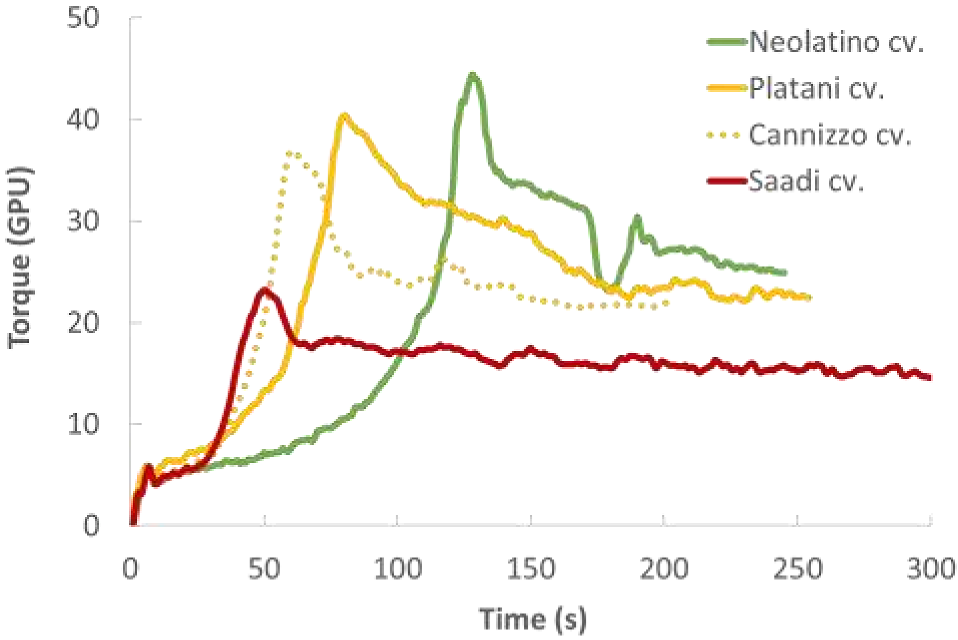

Figure 2. Gluten aggregation properties of selected durum wheat cultivars. GPU, GlutoPeak Unit;

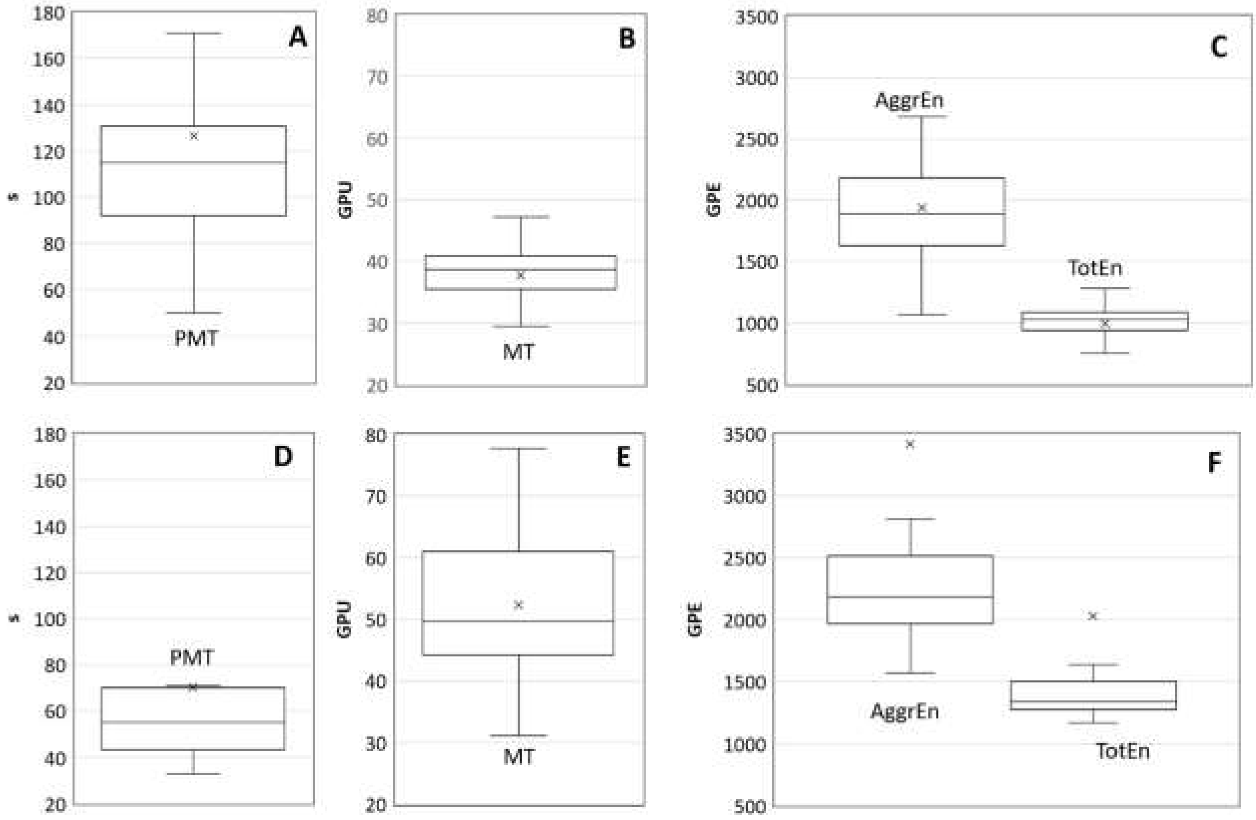

Figure 3. Indices related to gluten aggregation properties (as assessed by the GlutoPeak test) of durum wheat cultivars grown in 2015-2016 (A, B, C) and in 2016-2017 (D, E, F). A) and D) Peak Maximum Time (PMT); B) and E) Maximum Torque (MT); C) and F) Aggregation Energy (AggrEn) and Total Energy (TotEn). GPU, GlutoPeak Unit; GPE, GlutoPeak Equivalent.

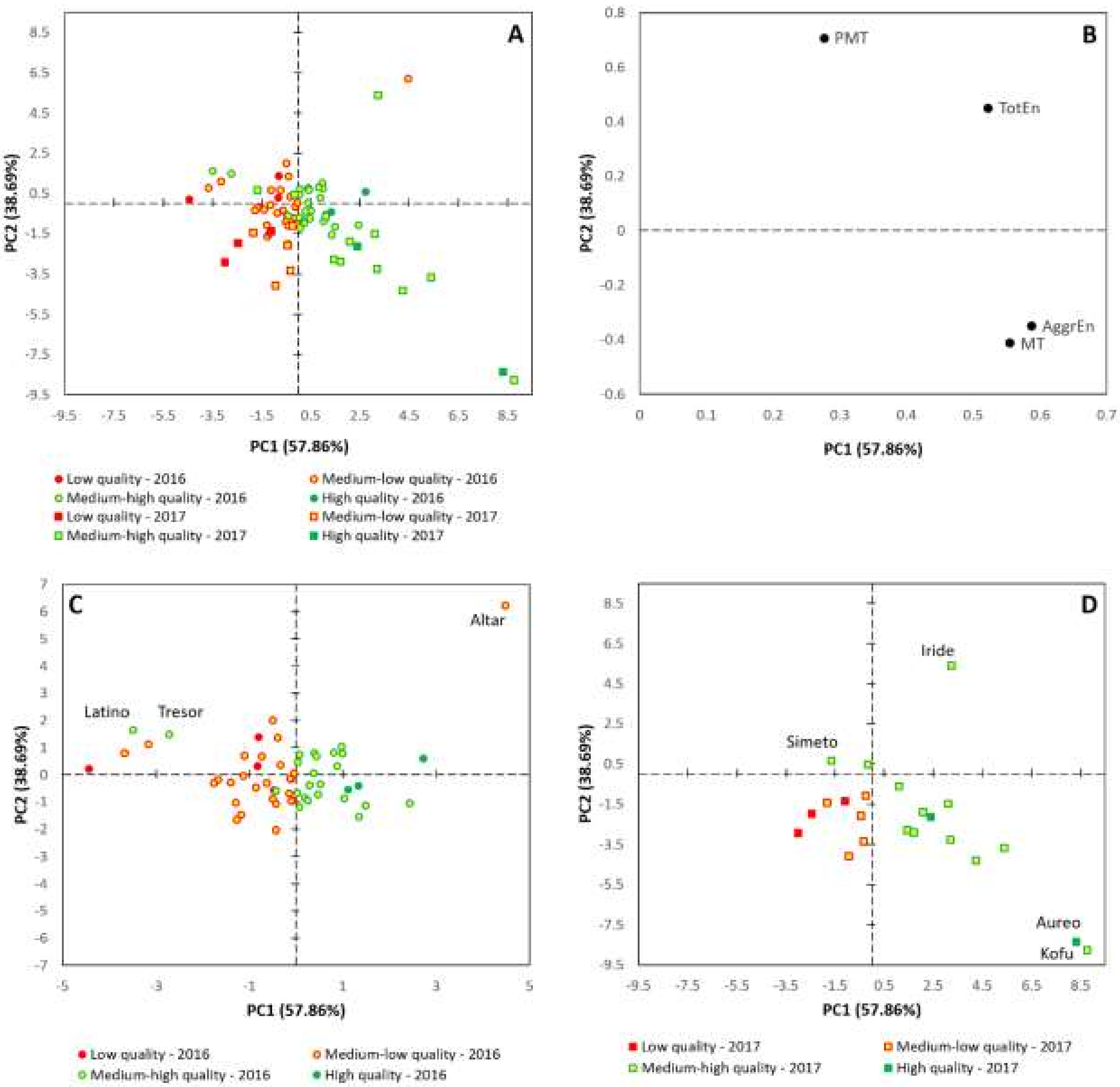

Figure 4. Principal Component Analysis of the indexes obtained from the GlutoPeak test: A) scores plot of the calibration and validation sets projected together; B) loadings plot (PMT, Peak Maximum Time; MT, Maximum Torque; TotEn, Total Energy; AggrEn, Aggregation Energy); C) scores plot of the calibration set alone; D) scores plot of the validation set alone.

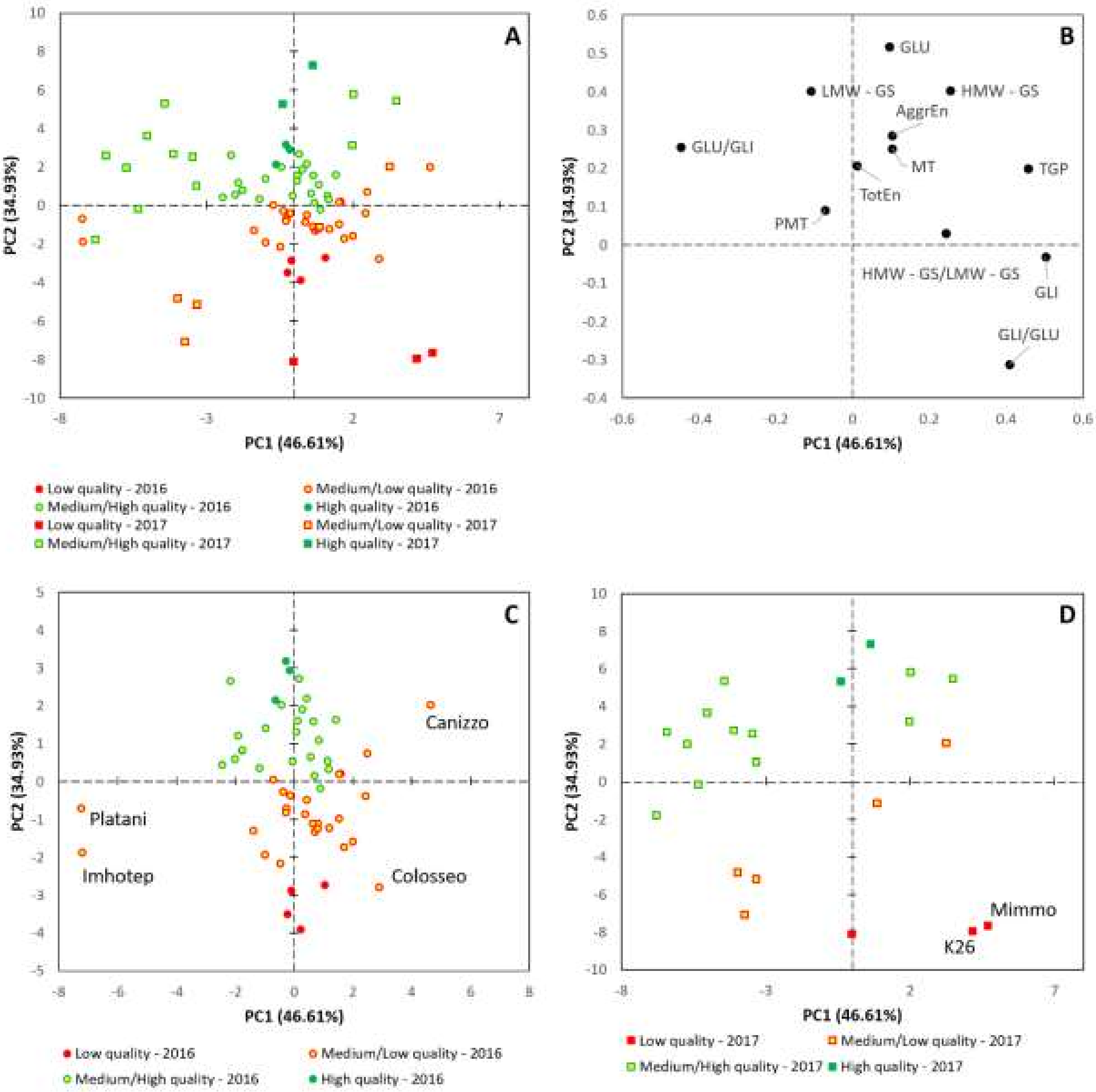

Figure 5. Principal Component Analysis of the quality parameters of grain gluten protein content (TGP) and gluten fractions combined with the indexes obtained from the GlutoPeak test: A) scores plot of the calibration and validation sets projected together; B) loadings plot (glutenins, GLU; gliadins, GLI; Low molecular weight glutenin sub-units, LMW-GS and High molecular weight glutenin sub-units, HMW-GS; PMT, Peak Maximum Time; MT, Maximum Torque; TotEn, Total Energy; TPG, Total gluten proteins, AggrEn, Aggregation Energy); C) scores plot of the calibration set alone; D) scores plot of the validation set alone.