Check for updates

RESEARCH ARTICLE

A Complete Mass Spectrometry (MS)-Based Peptidomic Description of Gluten Peptides Generated During In Vitro Gastrointestinal Digestion of Durum Wheat: Implication for Celiac Disease

Fatma Boukid,[1,2] Barbara Prandi,[1,3] Andrea Faccini,[4] Stefano Sforza[1,2]

[1]Food and Drug Department, University of Parma, Parco Area delle Scienze 27/A, 43124, Parma, Italy
[2]Interdepartmental Center SITEIA.PARMA, University of Parma, Parco Area delle Scienze 23/A, Parma, 43124, Italy
[3]Department of Human Sciences and Promotion of Quality of Life, Telematic University San Raffaele Roma, via Val Cannuta, 247, Rome, Italy
[4]Interdepartmental Centre for Measurements, University of Parma, Parco Area Delle Scienze 23/A, 43124, Parma, Italy
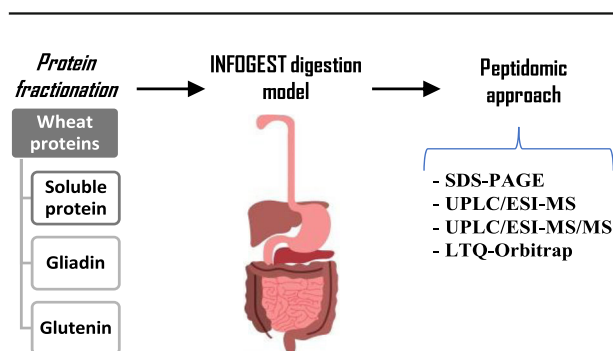
Abstract. Resistance of gluten to gastrointestinal digestion is involved in immune-mediated adverse reactions to wheat, since several peptides produced by the incomplete digestion are able to trigger, in predisposed individuals, the immune response responsible, for instance, of celiac disease (CD) and other adverse reactions. Even if several peptides have been identified, an exhaustive description of the peptidome generated by wheat digestion is lacking. To this end, in the present work, durum wheat proteins were fractionated, digested, and then subjected to various proteomic techniques, including single stage and multiple stage mass spectrometry (MS) (SDS-PAGE, UPLC/ESI-MS, UPLC/ESI-MS/MS, and LTQ-Orbitrap). Based on SDS-PAGE, although proteins were severely degraded after in vitro gastrointestinal digestion, some differences were observed among protein profile of the different digests. Through untargeted UPLC techniques, 227 peptide sequences were identified, with only few sequences shared by the different digests. In particular, 9 gluten peptides involved in CD were identified. Based on target proteomic, the quantification of these peptides revealed significant ($p \leq 0.05$) differences among the different extracts. Taken together, all the proteomic tools confirmed that gluten digestion is closely related to the matrix regardless of wheat genotype.
Keywords: Gluten, Digestion, SDS-PAGE, UPLC-MS, UPLC-MS/MS, Celiac disease

Introduction

W heat grain proteins are divided into two major groups: non-gluten proteins (20%) and gluten proteins (80%). Non-gluten proteins are soluble proteins (albumin and globulin), playing mainly structural and metabolic functions and a minor role in wheat quality [1]. These proteins might be involved in IgE-mediated food allergies such as Baker's asthma and intestinal inflammation [2, 3]. Gluten proteins are a mixture of storage proteins (gliadins and glutenins), conferring the rheological properties to wheat dough. Gluten proteins are rich in glutamine (30 to 35%) and proline (10 to 15%) residues, which make them resistant to complete proteolytic digestion, a property that contributes to the immunogenic nature of gluten for patients with celiac disease (CD) [4–6]. Besides CD, gluten-related disorders also include wheat-dependent exercise-

induced anaphylaxis, atopic dermatitis, urticaria, and gluten sensitivity [7, 8].

Accurate, reliable, and sensitive detection methods for gluten are, indeed, mandatory to support current EU regulations [9]. Enzyme-linked immunosorbent assays (ELISA) remain the method of choice [10]. However, the accuracy of commercially available ELISA kits is controversial due to the lack of certified reference material; cross-reactivity of antibodies; and large variety of kit with different antibody specificity, extraction conditions, and matrix effects [11, 12]. Over the last decades, the field of proteomics has seen a huge expansion and has been driven by the development of new, mostly MS-based, technologies for protein identification, separation, and quantification [13, 14]. Mass spectrometry–based approaches have then emerged as a core tool for large-scale gluten analysis owing to their specificity, sensitivity, and ability to identify hydrolyzed gluten [15]. Anyway, for a complete proteomic analysis, a full description of the compounds generated by gluten digestion is still lacking.

Considering that gluten is entrapped in a complex matrix of starch, protein, and lipids, the surface area accessible for digestive enzymes is very much dependent by the food matrices [16, 17]. This makes particularly problematic a reproducible identification and quantification of peptides deriving from digestion, which is particularly important in the case of gluten epitopes associated with CD. Ideally, digestion should be studied *in vivo* but this is not always possible for ethical and economic reasons. *In vitro* digestion models mimicking the gastrointestinal tract have been widely used to study the digestive fate of gluten, from single static systems to multicompartmental and dynamic systems. Nevertheless, the diversity of the systems used makes the comparison between the different studies almost impossible [12, 18–21]. For quantitative proteomics applications, the reproducible generation of peptides is closely related to several parameters including the type and amount of enzymes, salts, and pH values used in these methods, which resulted in substantial variability [12]. The type of cleavage enzyme was extensively studied because it greatly influenced both the sequence and the amounts of wheat peptides identified after digestion. For instance, 341 peptides were identified with chymotrypsin, 407 with thermolysin, and 105 with trypsin [18], while a multi-enzymatic digestion (LysC, trypsin, and chymotrypsin) enabled the identification of 434 peptide sequences from gluten [20].

Moving toward a standardized model, COST INFOGEST network has developed an internationally harmonized static model that simulates digestive processes by defining key parameters and conditions [22]. In order to test this model for assessing the digestive fate of gluten (pure extracts, or within wheat flour), several samples were analyzed in this work, mostly to test two hypotheses: (i) if gluten digestion within wheat matrix or as an extract modifies the resulting peptides/protein profile; and (ii) if changes in protein accessibility lead to changes in the amounts of peptides associated with CD. To this end, the sequential extraction of wheat proteins was carried out to obtain soluble fraction (SP), gliadins (GLIA), and

glutenins (GLUT) as well as the pellet was recovered. These extracts were dried and then subjected to digestion. A mixture of the extracts and the pellets (MIX) was remixed, dried, and digested. These samples (SP, GLIA, GLUT, PELLET, and MIX) were compared to the digesta of whole wheat flour (WWF). To track to the fate of gluten, an extensive characterization (qualitative and quantitative) of the peptides generated after digestion using both targeted and untargeted proteomics methods was used. Untargeted proteomics was used to obtain the global peptide profile of the digested wheat protein samples, while targeted proteomics was used to specifically quantify peptides which contain sequences able to trigger celiac disease.

Material and Methods

Plant Material

Grains of durum wheat (*Triticum durum* Desf.) of different pure varieties (Saragolla, Svevo, Maestrale, Creso, Simeto, and Cappelli) were milled using a laboratory mill (Ika Werke, Staufen, Germany). Whole wheat fine flour was sieved (160 μm sieve), and then stored in plastic bags at 4 °C until analysis. Table 1 summarizes some characteristics of the samples, in which they were ordered according to the release date.

Sequential Extraction of Wheat Proteins

The sequential extraction of wheat proteins was performed as previously described by [23] with minor modifications. Flours (100 mg) were extracted with a buffered salt solution (2 times × 1 mL 0.067 mol/L K_2HPO_4/KH_2PO_4-buffer, 0.4 mol/L NaCl, pH = 7.6) at 22 °C (room temperature) to obtain soluble fraction (albumin and globulin, SP). The residues were extracted with ethanol (60%, *v/v*; 3 times × 0.5 mL) at 22 °C (gliadins, GLIA) followed by the glutenin extraction solvent (2 times × 1 mL; 50% (*v/v*) 1-propanol, 0.1 mol/L TRIS-HCl, pH 7.5, 0.06 mol/L (*w/v*) dithiothreitol (glutenins, GLUT). After addition of the respective solvent, each flour suspension was vortexed for 2 min, stirred for 30 min, and centrifuged for 20 min at 3550*g* and 22 °C. The pellet was also recovered. Protein fractions (SP, GLIA, and GLUT) and the pellets were dried under nitrogen and stored until analysis.

For each variety, mixtures (MIX) were prepared by combining the total volume of the protein extracts (SP, GLIA, and GLU) and their corresponding pellets, and then dried under nitrogen flux and stored until analysis.

In Vitro Gastrointestinal Digestion

Dried samples (protein fractions (SP, GLIA, and GLUT), pellets, mixtures) and whole wheat flours (WWF) were subjected to simulated gastrointestinal digestion [22]. In brief, 1 g of WWF was incubated 2 min with 1 mL simulated saliva containing porcine amylase (Sigma-Aldrich, St. Louis, MI, USA; 75 U/mL of digesta); then, 2 mL of simulated gastric juice containing porcine pepsin (Sigma-Aldrich, St. Louis, MI,

Table 1. Pedigree and Year of Release of the Studied Wheat Genotypes

Genotypes	Pedigree	Year of release
Cappelli	Strampelli selection from Tunisian landrace "Jean khotifah"	1915
Creso	Yaktana-54/Norin 10-B//2*Cappelli-63/3/3*Tehuacan-60/4/Capelli-B144	1974
Simeto	Capeiti 8 × Valnova	1988
Svevo	Cimmyt line × Zenit	1996
Maestrale	Iride × Svevo	2004
Saragolla	Iride × PSB 014 line	2004

USA; 2000 U/mL of digesta) was added and the sample was incubated for 2 h after adjusting the pH to 3. Subsequently, 4 mL of duodenal juice containing porcine pancreatin (Sigma-Aldrich, St. Louis, MI, USA; 100 U trypsin activity/mL of digesta) and porcine bile (Sigma-Aldrich, St. Louis, MI, USA; 10 mmol/L in the total volume) was added and the sample was incubated for 2 h after adjusting the pH to 7. All the digestion steps were carried out at 37 °C under constant gentle mixing. Then, to inactivate the enzymes, the sample was boiled for 10 min at 95 °C. After centrifugation (3220g, 4 °C, 45 min), 295 μL of each sample supernatant was added to 5 μL of internal standard solution (TQQPQQPF(d_5)PQQPQQPF(d_5)-PQ; 1.6 mmol L−1). The standard is the labeled form (on the two phenylalanine residues) of the most abundant immunogenic peptide, which was synthesized following the method of Prandi and others [24].

Protein Profile

Protein contents of samples were quantified using a Qubit fluorometer (Invitrogen, Grand Island, NY, USA). The volume of gel-ready protein was computed considering that the quantity required in each well was 20 μg. If the required volume was > 10.5 μL, it was dried under nitrogen flux and then resolved in 15 μL of reducing buffer solution (protein concentration in the sample buffer was 1.3 μg/μL). Protein fractions (SP, GLIA, and GLUT) before and after digestion as well as the digests of pellets, mixtures, and whole wheat flours were subjected to SDS-PAGE [25].

Peptidomic Profile: UHPLC/ESI-MS/MS Analysis

The digests were chromatographically separated by reverse-phase (RP) column (Aeris Peptide 1.7 μm XB-C18, 150 × 2.10 mm, Phenomenex, Torrance, CA, USA) in an UHPLC/ESI-MS/MS (Dionex Ultimate 3000, Thermo Scientific, Waltham, MA, USA, equipped with a triple quadrupole TSQ Vantage (Thermo Scientific, Waltham, MA, USA). Eluent A was water with 0.1% formic acid and 0.2% acetonitrile, eluent B was acetonitrile with 0.1% formic acid and 0.2% water; gradient 0–7 min, 100% A; 7–50 min, from 100% A to 50% A; 50–52.6 min, 50% A; 52.6–53 min, from 50% A to 0% A; 53–58.2 min, 0% A; 58.2–59 min, from 0% A to 100% A; 59–72 min, 100% A. Product ion scan modality was carried using a variable collision energy based on the mass and charge of the ion to be fragmented. The setting of the UHPLC/ESI-MS/MS analysis was flow 0.2 mL/min; analysis time 72 min; column temperature 35 °C; sample temperature 18 °C; injection volume 3 μL; acquisition time 7–58.2 min; ionization type positive ions; scan range 100–1500 m/z, micro scans 1, scan time 0.50, Q1 PW 0.70, spray voltage 3200 V, capillary temperature 250 °C, vaporizer temperature 250 °C, sheath gas flow 22 units. The peptide sequences were assigned based on the mass spectra obtained. Briefly, the software FindPept (http://web.expasy.org/findpept/) was used to find the peptide sequences whose molecular weight matched with the experimental data. Then, the software Proteomics Toolkit (http://db.systemsbiology.net/proteomicsToolkit/FragIonServlet.html) was used to verify the correspondence between the theoretical MS/MS fragmentation and the obtained spectra.

Untarget Peptidomic: LTQ-OrbiTrap Analyses

Digested samples were analyzed by μHPLC-LTQ-ORBITRAP using a C18 column (Phenomenex Jupiter 4 μm Proteo 90 Å 150 mm × 0.3 mm) equipped with an enrichment cartridge (μ-Precolumn Cartridge, Acclaim PepMap100 C18 5 μm, 100 Å, 300 μm × 5 mm; loading flow 30 μL/min, 50% eluent A and 50% eluent B). Eluent A was water with 0.2% formic acid and eluent B was acetonitrile with 0.2% formic acid (gradient 0–4 min 10% B, 4–60 min linear from 10% B to 50% B, 60–62 min from 50 to 95% B, 62–72 min 95% B, 72–73 min from 95% B to 10% B, 73–82 min 10% B). The analysis parameters were flow 5 μL/min; analysis time 82 min; column temperature 35 °C; sample temperature 10 °C; injection volume 5 μL; acquisition time 0–72 min; ionization type positive ions. High-resolution mass spectrometry (HRMS) acquisition was performed through 5 subsequent events: event 1: full scan acquisition from 250 to 2000 m/z in high-resolution mode (resolution at 400 m/z = 30,000); events from 2 to 5: data-dependent scan, at each cycle, the four most intense ions (with charge z > 1 and with a minimum signal of 500 counts) identified in event 1 are fragmented. The same ion (tolerance 10 ppm and isolation window 2 m/z) can be observed for a maximum of 2 cycles, and then it is automatically inserted in the exclusion list for a maximum time of 20 s. Fragmentation is performed in the linear trap of the instrument in CID mode with collision energy of 35. Proteins were identified either with Peaks Studio (Bioinformatics Solutions, Waterloo, ON, Canada). Parameters were precursor ion tolerance 5 ppm, fragment ion tolerance 0.8 Da, decoy database search: strict 0.01, relaxed 0.05, fixed modifications: cysteine carbamidomethylation, variable modifications: methionine oxidation, hydroxyproline, and hydroxylysine. The spectral

datasets were searched against the *Poaceae*. Results filtration parameters were peptide $-10\lg P \geq 15$, protein $-10\lg P \geq 20$, protein-unique peptides ≥ 0, de novo ALC score ≥ 50. False discovery rates were as follows: soluble proteins (FDR peptide-spectrum matches 23.0%, FDR peptide sequences 38.9%, FDR protein 5.5%), gliadins (FDR peptide-spectrum matches 6.7%, FDR peptide sequences 12.0%, FDR protein 3.0%), glutenins (FDR peptide-spectrum matches 12.2%, FDR peptide sequences 13.3%, FDR protein 0.0%), pellet (FDR peptide-spectrum matches 0.0%, FDR peptide sequences 0.0%, FDR protein 0.0%), mix (FDR peptide-spectrum matches 2.1%, FDR peptide sequences 3.3%, FDR protein 0.0%), and whole wheat flour (FDR peptide-spectrum matches 2.7%, FDR peptide sequences 6.5%, FDR protein 0.0%).

Target Peptidomic: UPLC/ESI-MS Analysis

Digested samples were separated using UPLC/ESI-MS system (ACQUITY Ultra-Performance UPLC with a single quadrupole mass spectrometer SQD, Waters Milford, MA, USA) a RP column (ACQUITY UPLC BEH 300, C18, 1.7 mm, 2.1 × 150 mm; Waters, Milford, MA, USA) equipped with a Van-Guard Pre-column (ACQUITY UPLC BEH C18, 130 Å, 1.7 μm, 2.1 × 5 mm, Waters, Milford, MA, USA). The gradient elution is made with eluent A (bi-distilled water solution with 0.1% formic acid and 0.2% acetonitrile 0.2%) and eluent B (acetonitrile solution with 0.1% formic acid). Gradient elution was carried out as follows: 0–7 min 100% eluent A; 7–50 min from 100 to 50% eluent A; 50–52.6 min 50% eluent A; 52.6–53 min from 50 to 0% eluent A; 53–58.2 min 0% eluent A; 58.2–59 min from 0 to 100% eluent A; 59–72 min 100% eluent A. The digests were analyzed with UPLC/ ESI-MS in the Full Scan mode. Flow is 0.2 mL/min; analysis time 72 min; column temperature 35 °C; sample temperature 18 °C; injection volume 5 μL; acquisition time 7–58.2 min; ionization type is positive ions; scan range 100–2000 m/z; capillary voltage 3.2 kV; cone voltage 30 V; source temperature 150 °C; desolvation temperature 300 °C; cone gas flow 100 L/h; desolvation gas flow 650 L/h. Two determinations were performed for each sample. For CD-related peptide quantification, the areas of the identified peptides and internal standard (TQQPQQPF(d_5)PQQPQQPF(d_5)PQ; 1.6 mmol L^{-1}) were integrated with the MassLynx software. Due to the high number of peptides ($n = 10$) to be quantified, the main immunogenic peptide found (IP4) was synthesized and used as labeled standard.

Statistical Analysis

To study significant differences among the variables, analysis of variance (ANOVA) was performed, with a confidence interval of 95% ($p \leq 0.05$). Principal components analysis was performed based on the correlation matrix. All the statistical analyses were determined using the program SPSS for Windows (Version 24.0, SPSS Inc., Chicago, IL, USA).

Results and Discussion

Efficiency of Gluten In Vitro Digestion

Wheat proteins were extracted using a modified Osborne protocol for the recovery of soluble fraction, gliadin, and glutenin. To assess the efficiency of gluten digestion, protein fractions were subjected to SDS-PAGE before and after in vitro digestion (Figure 1).

The characteristic bands for undigested protein fractions were observed at the corresponding MW ranges of 7,000–97,000 for soluble proteins, 31,000–66,000 for gliadins, and 14,000–120,000 for glutenin. After in vitro digestion, only minor traces were observed in the digested pure protein fractions due to their severe hydrolysis by digestive enzymes.

Also, the recovered pellet from the sequential extraction of proteins was digested and analyzed using SDS-PAGE (Figure 1), and in this case, the results showed that after digestion, no bands were anymore detectable. On the other side, in the case of the mixture obtained by blending all protein extracts and pellet, some bands were observed at MW ranges of 25,000–66,000 Da, indicating an incomplete digestion. The same was also observed for the whole flour digest, with the observed bands ranging from 14,000–97,000 Da. Therefore, it can be concluded that protein profile after digestion is closely related to the matrix surrounding the proteins.

The Fate of Gluten

A combined peptidomic approach was used to track the fate of gluten after digestion. To this end, the digests were chromatographically separated by UPLC/ESI-MS to select all the peptide ions that yielded detectable peaks. For each of these peaks, the molecular weight was determined.

By comparing the peak profile among the different digests, it was immediately apparent that, qualitatively speaking, the detected profiles did not vary among the different wheat varieties used, whereas differences were observed among the different extracts and total wheat flour within each variety. As an example, total ion chromatograms obtained from UPLC/ESI-MS analysis of the digests relative to the variety Cappelli are shown in Figure 2. For the acquisition time, from 0 to 7 min of
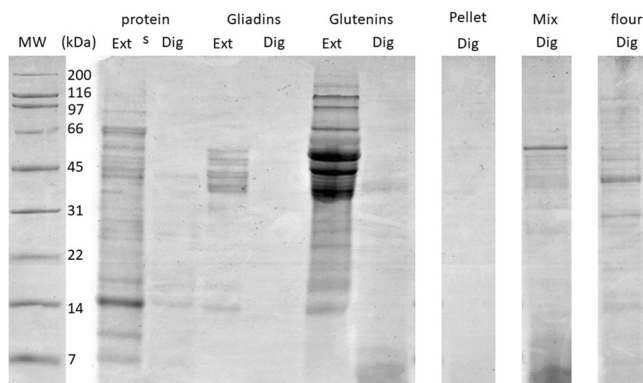
Figure 1. Changes in protein profile before and after digestion (Cappelli is shown as example) (Ext, extract; Dig, digest)

the chromatographic run was diverted to waste due to the high salt and sugar content; from 12 to 30 min, peptides ranging from MW 200 to 3,600 were eluted; and from 35 to 72 min, bile salts deriving from the digestive solutions were present. To compare the peak profile of the six digests (SP, GLI, GLU, PEL, MIX, and WWF), exclusively, the segment ranging from 12 in to 30 min was analyzed. It is very easy to see that the relative response of peptides was different between the different digests in this range, which can be associated with the compositional differences among the digested samples (different samples with varying amount of each component to start with) (Figure 2).

In this range, peaks were significantly higher in WWF and in the pellet. The digest of the mix was different, showing less intensity in the peptide area. As for protein fractions, the chromatograms of gliadin and soluble proteins were apparently very similar, whereas that of glutenin showed less peaks from 7 to 15 min and no peaks at all beyond 15 min.

Then, the sequences of the most relevant peptides were determined through a further UHPLC/ESI-MS/MS and LTQ-OrbiTrap analysis, by determining the single specific sequences by analysis of the independent fragment peaks matching the theoretical peptide fragments (Supp. Material 1). This combined MS-based peptidomic approach enabled the identification of 227 sequences (Supp. Material 2) belonging to several wheat proteins, thereby giving a good representation of the wheat gluten peptidome after digestion. The number of peptides and the type of proteins differed among the different digests (Table 2). In soluble protein digests, the 27 identified peptides were mainly fragments from alpha-amylase/trypsin inhibitor CM3 and beta-amylase (85%), γ-gliadin (11%), and LMW-

glutenin (4%). CM3 and beta-amylase, in fact, belong to the soluble fraction of wheat proteins. Small amount of γ-gliadin and LMW-glutenin were also probably co-extracted during the fractionation procedure. The peptides identified in the digest of gliadin were almost equally resulting from α- and γ-gliadin (49% and 42%, respectively) followed by β-gliadin (5%) and LMW-glutenin (5%), consistently with their solubility features. The sequences found in the glutenin digests were mainly from LMW-glutenin (77%), followed by α-gliadin, (14%) and γ-gliadin (14%). Most peptides identified in the digested pellet were from α-gliadin (80%), followed by γ-gliadin (13%) and HMW-glutenin (7%). As for the digest of the mixture, peptides were mainly from γ-gliadin (45%), α-gliadin (30%), and HMW-glutenin (24%). Digestion of the whole wheat flour generated 80 identifiable peptides, comparable to previous findings ([12] (84 peptides); [26](77 peptides)). These peptides were mostly deriving from α-gliadin (78%), γ-gliadin (13%), and HMW-glutenin (8%). Few peptides deriving from non-gluten proteins were also identified, arising from α-amylase/trypsin inhibitor CM3. No peptides from the ω-gliadin were identified in any of the digested samples.

The fractionation of wheat proteins allowed to identify more peptides (27 (SP) + 43 (GLI) + 14(GLU) + 30 (PELLET) = 114 peptides), when compared to that of digested whole wheat flour (80 peptides), maintaining the same acquiring time for the data.

Results showed different types and relative amounts of peptides in the fractions, pellet, and mix. This finding is suggestive for the fact that gluten protein digestion is hampered by the matrix effect in whole flour. This can be due to a limitation in enzyme accessibility to proteins, or to an increased competition among them for accessing the enzyme-active sites.
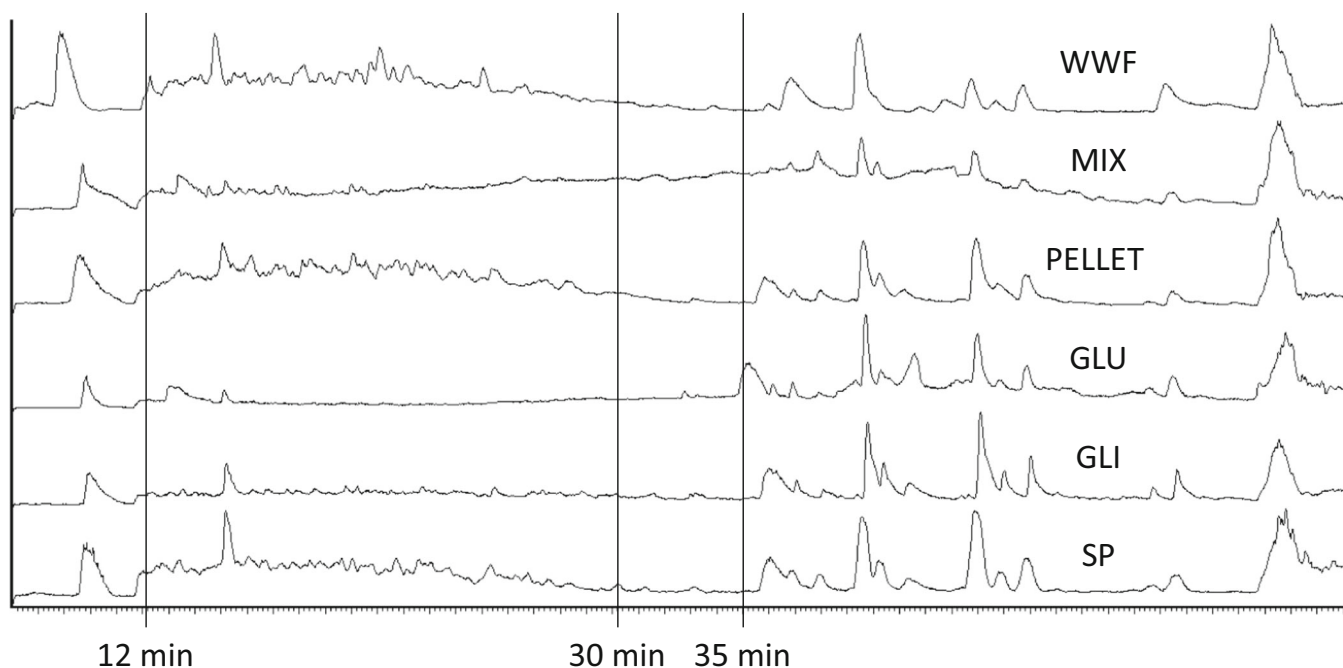

Figure 2. Total ion chromatograms obtained from UPLC/ESI-MS of the different digests (Cappelli is shown as example). Peptides elute from 12 to 30 min

Table 2. Identified Peptides (%) in the Digests of Protein Fractions, Pellet, Mix and Whole Wheat Flour

Code	Digest	n	Gluten proteins						Soluble proteins
			α	β	γ	ω	LMW	HMW	
SP	Soluble protein	27	0%	0%	11%	0%	4%	0%	85%
GLI	Gliadin	43	49%	5%	42%	0%	5%	0%	0%
GLU	Glutenin	14	14%	0%	14%	0%	71%	0%	0%
PEL	Pellet	30	80%	0%	13%	0%	7%	0%	0%
MIX	Mixture	33	30%	0%	45%	0%	24%	0%	0%
WWF	Whole wheat flour	80	78%	0%	13%	0%	8%	1%	1%

n: number of peptides identified in each digest

Peptides Eliciting Adverse Reactions in CD Patients

Identification of Peptides Associated with CD In the context of CD, some gluten peptides were identified in the digests already known in literature to be involved in the adaptive or the innate immune response. Table 3 summarizes the list of gluten peptides associated with CD, which were identified in the digested samples, which was consistent with previous works [24, 26, 27]. These peptides may be subdivided into two groups: sequences triggering the adaptive or the immune response. Seven immunogenic peptides deriving from γ-gliadin were identified as involved in the adaptive immune response ("immunogenic peptide," TI), containing within their sequences the epitope DQ2.5-glia-γ4c (QQPQQPFPQ) [28]. Three peptides deriving from α-gliadin were identified as involved in the innate immune response ("toxic peptides," TT), containing within their sequences toxic epitopes (PSQQ, QQQP, QQPY, or QPYP) [29].

The 33-mer peptide (LQLQPFPQPQLPYPQPQLPYPQPQLPYPQPQPF) from α2-gliadin is generally referred as the most immunodominant gluten peptide (Arentz-Hansen et al., 2000; Shan et al., 2002). In the case of CD patients, the ingestion of gluten can result in the initiation of a strong immune response because it contains three T cell epitopes (PFPQPQLPY (DQ2.5-glia-α1a), PYPQPQLPY (DQ2.5-glia-α1b), and PQPQLPYPQ (DQ2.5-glia-α2)) [30, 31]. However, the complete fragment of 33-mer was not identified in all the studied digests deriving from durum wheat. Indeed, several studies [23, 32] found the entire sequence of this peptide exclusively in common wheat (*Triticum aestivum*) flour because it is encoded by D-genome.

Quantification of Peptides Associated with CD Through target proteomic (UPLC/ESI-MS), performing a quantification based on a labeled internal standard, the changes in the contents of peptides associated with CD were monitored among the different digests (Table 4 (peptides triggering adaptive immune response); Table 5 (peptides triggering innate immune response)). Regarding peptides triggering the adaptive immune response (Table 4), digests of whole wheat flours recorded the highest amounts than those deriving from the extracts. Such a result might be attributed to the fact that the more complete digestion in the extracts rather than WWF allows digestive enzymes to reduce immunogenic peptides into smaller sequences. In the case of WWF, immunogenic peptides are encased within a compact network with starch particles thereby likely reducing their accessibility toward digestive enzymes. As well, the mixture (MIX = SP + GLU + GLU + PELLET) contained less amounts of peptides compared to the whole wheat flour (WWF), which can be attributed to the fact that extraction possibly increased the availability of proteins to be cleaved by the digestive enzyme. The same trend was observed in all the six studied varieties. The amount of peptides in soluble fraction was found to be fluctuating with no clear trend, consistently with the gluten peptides being present in this fraction only as contaminants, or/and the fact that globulins and gliadins shared some common sequences [4, 33]. Such results can suggest that the

Table 3. Identified Peptides Associated with CD

Identified peptides related to CD	Protein
Peptides eliciting the adaptive immune response	
TQQPQQPFPQ	γ-gliadin
SQQPQQPFPQPQ	γ-gliadin
QAFPQQPQQPFPQ	γ-gliadin
TQQPQQPFPQQPQQPFPQ	γ-gliadin
PQTQQPQQPFPQFQQPQQPFPQPQQP	γ-gliadin
FPQQPQLPFPQQPQQPFPQPQQPQ	γ-gliadin
QQPQQPFPQPQQTFPQQPQLPFPQQPQQPFP	γ-gliadin
Peptides eliciting the innate immune response	
LQPQNPSQQQPQ	α-gliadin
RPQQPYPQPQPQ	α-gliadin
LQPQNPSQQQPQEQVPL	α-gliadin

Table 4. Monitoring Peptides Triggering Adaptive Immune Response (ppm) Associated with CD in the Digests

Digest	TQQQQPFPQ	SQQPQQPFPQPQ	QAFPQQQQQPFPQ	TQQPQQPFP QQPQQPFPQ	PQTQQPQQPFPQ FQQPQQPFPQPQQP	FPQQPQLPFPQQ PQQPFPQQQQQPQ	QQPQQPFPQPQQ TFPQQPQLPFPQ QPQQPFP	TI
Simeto								
SI-SP	14±1 b	19±1 b	7±1 b	22±2 b	3±0 a	2±0 a	1±0 a	68±3b
SI-GLIAD	20±2 c	35±2c	8±0 b	49±5 c	3±0 a	4±1 a	5±0 a	124±5c
SI-GLU	0±0 a	1±0 a	1±0 a	1±0 a	1±0 a	0±0 a	0±0 a	4±0 a
SI-PELLET	42±2 d	96±2 e	72±4 cd	106±0 d	85±11 b	7±0 b	32±1 b	440±16 d
SI-MIX	1±0 a	64±2 d	86±4 d	397±34 e	351±8 c	176±5 d	119±15 c	1193±30 f
SI-WWF	83±3 e	107±4 f	38±3 c	413±18 ef	316±17 c	71±4 c	106±8 c	1135±8e
Cappelli								
CA-SP	21±5c	23±4c	11±0bc	31±4c	13±1b	3±0a	3±0a	105±6c
CA-GLIAD	8±1 b	10±1b	7±0b	10±0b	1±0e	1±0a	1±0a	36±1b
CA-GLU	1±0 a	0±0a	0±0a	0±0a	1±1a	2±0a	2±0a	6±0a
CA-PELLET	33±10 d	178±8f	90±2f	530±31e	417±35c	119±1b	206±16c	1573±43e
CA-MIX	42±2d	80±9d	36±5d	220±5d	228±11b	176±8c	128±2bc	908±36d
CA-WWF	110±1e	113±12ef	75±1e	682±24 f	467±15c	143±10bc	241±14c	1830±49f
Svevo								
SV-SP	8±0b	16±1bc	12±2b	42±14c	2±1a	2±0a	1±0a	83±18bc
SV-GLIAD	16±0bc	14±1b	10±0b	18±0b	2±0a	2±1a	1±1a	63±1b
SV-GLU	1±0a	1±0a	0±0a	0±0a	1±0a	2±1a	2±0a	7±1a
SV-PELLET	53±10c	108±8d	51±4c	310±49de	206±33b	24±2b	83±13b	835±29d
SV-MIX	1±0a	104±6d	65±6 cd	278±55d	269±16bc	121±22d	130±23c	967±27de
SV-WWF	105±1d	91±4 cd	51±1c	432±10ef	386±6 cd	89±13c	157±5 cd	1308±19f
Maestrale								
MA-SP	10±1b	17±4b	0±0a	18±0b	2±1a	0±0a	1±0a	47±4b
MA-GLIAD	24±1c	20±4b	1±0a	7±1ab	2±0a	2±0a	1±0a	56±6b
MA-GLUT	0±0a	0±0a	0±0a	1±0a	1±0a	1±0a	1±1a	4±1a
MA-PELLET	47±8d	83±28d	44±14b	284±15c	186±24b	21±11b	61±26c	724±26c
MA-MIX	0±0a	54±1c	1±0a	302±9 cd	285±19c	106±3d	247±21d	996±45 cd
MA-WWF	131±8e	147±17e	120±10c	771±16e	653±15d	150±27d	19±10b	1989±40e
Saragolla								
MA-SP	0±0a	13±3c	8±1b	18±1b	1±0a	0±0a	0±0a	41±6b
MA-GLIAD	13±2b	9±0b	9±1b	15±1b	1±1a	1±1a	1±0a	46±4b
MA-GLU	0±0a	0±0a	0±0a	1±0a	1±0a	0±0a	3±1a	6±0a
MA-PELLET	34±2c	40±6d	55±10d	201±6d	148±11b	13±1b	65±6c	553±31c
MA-MIX	8±0ab	42±1d	34±0c	146±18c	265±29c	3±1a	4±0a	500±10c
SA-WWF	80±1d	71±4e	74±3de	472±10e	432±9d	74±3c	191±11d	1391±27d
Creso								
CR-SP	2±1a	17±3b	1±0a	12±1b	1±0a	0±0a	1±1a	34±4b
CR-GLIAD	2±0a	6±0ab	0±0a	1±0a	2±0a	0±0a	1±0a	13±1ab
CR-GLU	1±0a	0±0a	0±0a	4±1a	1±0a	1±0a	1±0a	8±1a
CR-PELLET	178±21c	143±14e	79±21c	423±8d	302±1c	77±8bc	150±8b	1351±21d
CR-MIX	1±0a	40±8c	34±0b	212±14c	198±1b	52±12b	177±19c	712±16c
CR-WWF	118±3b	113±14d	114±5d	675±52e	584±77d	221±14d	359±45d	2183±204e

Data were listed as means ± standard deviations of duplicate measurements; for each variety, different letters in the same column mean significant difference (Duncan's method, $p \leq 0.05$)

Table 5. Monitoring Peptides Triggering Innate Immune Response (ppm) Associated with CD in the Digests

| Digest	LQPQNPSQQQPQ	RPQQPYPQPQPQ	LQPQNPSQQQPQEQVPL	TT
Simeto				
SI-SP	9 ± 0 bc	0 ± 0a	1 ± 0a	10 ± 0b
SI-GLIAD	13 ± 0 c	0 ± 0a	0 ± 0a	14 ± 0b
SI-GLU	0 ± 0 a	0 ± 0a	0 ± 0a	1 ± 0a
SI-PELLET	93 ± 8 d	8 ± 1b	41 ± 5b	142 ± 13c
SI-MIX	0 ± 0 a	1 ± 0a	132 ± 0c	133 ± 0c
SI-WWF	129 ± 5 e	95 ± 5c	39 ± 4b	263 ± 14d
Cappelli				
CA-SP	13 ± 2c	1 ± 0a	1 ± 0a	15 ± 1c
CA-GLIAD	3 ± 1b	2 ± 1a	0 ± 0a	5 ± 1b
CA-GLU	0 ± 0a	0 ± 0a	0 ± 0a	1 ± 1a
CA-PELLET	159 ± 9c	14 ± 1b	76 ± 2bc	249 ± 12e
CA-MIX	0 ± 0a	2 ± 0a	69 ± 1b	71 ± 1c
CA-WWF	200 ± 4e	245 ± 16c	98 ± 8d	543 ± 11f
Svevo				
SV-SP	15 ± 0	2 ± 0a	0 ± 0a	17 ± 0bc
SV-GLIAD	7 ± 0b	1 ± 0a	0 ± 0a	9 ± 0b
SV-GLU	1 ± 1a	3 ± 0a	0 ± 0a	3 ± 0a
SV-PELLET	91 ± 5d	10 ± 1b	64 ± 4c	165 ± 11d
SV-MIX	1 ± 0a	1 ± 0a	85 ± 0d	87 ± 0c
SV-WWF	168 ± 11e	196 ± 2c	120 ± 1e	484 ± 11e
Maestrale				
MA-SP	13 ± 1b	8 ± 1b	0 ± 0a	20 ± 0b
MA-GLIAD	12 ± 0b	9 ± 1b	0 ± 0a	21 ± 1b
MA-GLU	0 ± 0a	0 ± 0a	0 ± 0a	1 ± 0a
MA-PELLET	92 ± 12c	43 ± 1c	18 ± 1b	151 ± 11c
MA-MIX	0 ± 0a	440 ± 15de	84 ± 4c	525 ± 62d
MA-WWF	243 ± 6d	412 ± 16d	110 ± 13d	764 ± 48e
Saragolla				
MA-SP	11 ± 1c	0 ± 0a	0 ± 0a	12 ± 1ab
MA-GLIAD	5 ± 0b	1 ± 0a	0 ± 0a	7 ± 1b
MA-GLU	1 ± 0a	0 ± 0a	0 ± 0a	1 ± 0a
MA-PELLET	66 ± 6d	60 ± 6b	37 ± 4b	162 ± 8d
MA-MIX	0 ± 0a	89 ± 1c	44 ± 3bc	133 ± 2c
SA-WWF	176 ± 11e	252 ± 9d	139 ± 17d	566 ± 16e
Creso				
CR-SP	12 ± 1b	4 ± 1ab	1 ± 1a	16 ± 1b
CR-GLIAD	0 ± 0a	8 ± 1b	0 ± 0a	8 ± 0ab
CR-GLU	0 ± 0a	0 ± 0a	0 ± 0a	1 ± 0a
CR-PELLET	163 ± 1c	61 ± 0c	63 ± 8c	287 ± 6de
CR-MIX	1 ± 0a	184 ± 20d	49 ± 5b	234 ± 15d
CR-WWF	196 ± 22d	300 ± 40e	296 ± 76e	791 ± 37f

Data were listed as means ± standard deviations of duplicate measurements; for each variety, different letters in the same column mean significant difference (Duncan's method, $p \leq 0.05$)

extraction method greatly influences not only the identified peptides but also their amounts. Peptides triggering the innate immune response which were identified, listed in Table 3, derive from α-gliadin, which are the smallest ones among the gluten proteins. Thus, the salt solution probably allows their extraction, while for higher MW proteins (γ-gliadin and glutenins), the use of aqueous alcohol or denaturing agents is needed.

A better picture was obtained by putting together all the results through principal component analysis (PCA). The first principal components (PC1 and PC2) explained 86% of the total variability. PC1 (44%) was explained by three peptides triggering adaptive immune response (PQTQQPQQPFPQFQQPQQPFPQPQQP, FPQQPQLPFPQQPQQPFPQPQQPQ, and QQPQQPFPQPQQTFPQQPQLPFPQQPQQPFP) and two peptides triggering innate immune response (RPQQPYPQPQPQ and LQPQNPSQQQPQEQVPL), while PC2 (42%) was a function of four peptides triggering the adaptive immune response (TQQPQQPFPQ, SQQPQQPFPQPQ, QAFPQQPQQPFPQ, and TQQPQQPFPQQPQQPFPQ) and the remaining peptide triggering innate immune response (LQPQNPSQQQPQ). The biplots of the first two components (Figure 3a) showed that all the peptides were located in the same quadrant. The overlapping of the digests' scores on the consensus space generated by PC1 and PC2 is presented in Figure 3b. A clear clustering was obtained, where four groups were separated: Group 1 gathered digested WWF, which are characterized by the highest content of peptides associated with CD; group 2 was made by the digested mixtures; group 3 included the digested fractions (SP, GLI, and GLU); and group 4 was made with digested pellet, which was located in the opposite side of WWF. It is interesting to note that the digestion of the extracts (SP, GLI, and GLU) leads at having of all the six varieties
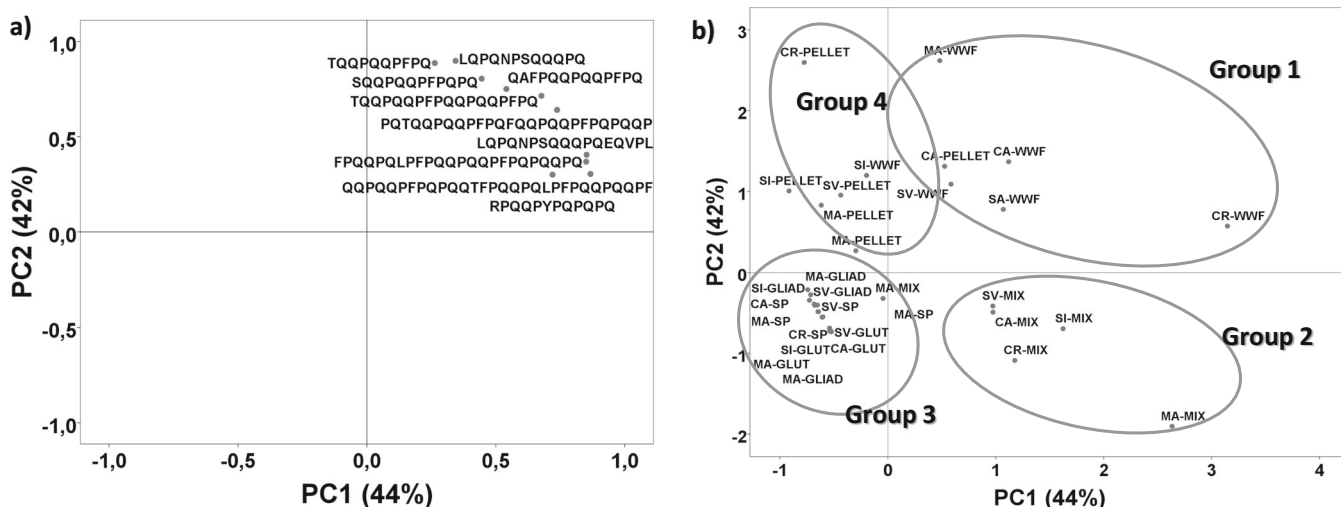
Figure 3. Principal component analysis. (**a**) Load plot of principal component analysis explaining 86% of total variability. (**b**) Score plot with the wheat samples projected onto the first two principal components

tested much closed together, overwhelming possible inter-genotype variations. The digestion of WWF, indeed, allows to better separate the different genotypes on the score plot, enhancing the peculiarities of each variety, as already demonstrated in previous papers [24, 26, 32, 34].

Conclusion

Combining multiple enzymatic digestion and MS is a powerful tool that enabled the full peptidomic description of the digest mixtures, and the detection and the quantification of gluten peptides associated with CD. Particularly, 227 peptides deriving from gluten were identified, in which 9 were known to be involved in offsetting CD. The quantification of these peptides in the digested extracts and whole wheat flour showed significant difference. For both groups, the peptides eliciting the innate and adaptive immune responses, the superiority of the amounts of these peptides was maintained in the digested whole wheat flour compared to digested mixture, independently of wheat genotype. Such findings indicate a significant effect of the matrix on protein availability and thereby their digestibility. Therefore, gluten deriving material (extracts or within matrix) greatly influenced the efficacy of in vitro digestion, with profound effects on the composition of the mixture and the outcome for CD patients and other subjects sensitized to gluten. Overall, by demonstrating that the full wheat matrix leads to generation of more peptides triggering CD, it also outlines how, properly working on the matrix, it might be possible to enhance the digestion of toxic and immunogenic peptides, thereby reducing the adverse potential of wheat to sensitized subjects.

Acknowledgements

The plant set was kindly provided by E. Francia, Department of Life Sciences, University of Modena and Reggio Emilia, Italy.

Compliance with Ethical Standards

The research did not involve any human participants or animals.

Conflict of Interest The authors declare that they have no conflict of interest.

References

1. Guerrieri N., Cavaletto M. Cereals proteins. In: Proteins in food processing. pp. 223 - 244. Rickey Y. Yada (ed.). (Woodhead Publishing Series in Food Science, Technology and Nutrition). United Kingdom: Woodhead Publishing, an imprint of Elsevier (2018)
2. Cianferoni, A.: Wheat allergy: diagnosis and management. J. Asthma Allergy. **9**, 13–25 (2016)
3. Platzer, B., Baker, K., Vera, M.P., Singer, K., Panduro, M., Lexmond, W.S., Turner, D., Vargas, S.O., Kinet, J.-P., Maurer, D., Baron, R.M., Blumberg, R.S., Fiebiger, E.: Dendritic cell-bound IgE functions to restrain allergic inflammation at mucosal sites. Mucosal Immunol. **8**, 516–532 (2015)
4. Ferranti, P., Mamone, G., Picariello, G., Addeo, F.: Mass spectrometry analysis of gliadins in celiac disease. J. Mass Spectrom. **42**, 1531–1548 (2007)
5. Valerii, M.C., Ricci, C., Spisni, E., Di Silvestro, R., De Fazio, L., Cavazza, E., Lanzini, A., Campieri, M., Dalpiaz, A., Pavan, B., Volta, U., Dinelli, G.: Responses of peripheral blood mononucleated cells from non-celiac gluten sensitive patients to various cereal sources. Food Chem. **176**, 167–174 (2015)
6. Stamnaes, J., Sollid, L.M.: Celiac disease: autoimmunity in response to food antigen. Semin. Immunol. **27**, 343–352 (2015)
7. Scherf, K.A., Koehler, P., Wieser, H.: Gluten and wheat sensitivities – an overview. J. Cereal Sci. **67**, 2–11 (2016)

8. Catassi, C., Bai, J., Bonaz, B., Bouma, G., Calabrò, A., Carroccio, A., Castillejo, G., Ciacci, C., Cristofori, F., Dolinsek, J., Francavilla, R., Elli, L., Green, P., Holtmeier, W., Koehler, P., Koletzko, S., Meinhold, C., Sanders, D., Schumann, M., Schuppan, D., Ullrich, R., Vécsei, A., Volta, U., Zevallos, V., Sapone, A., Fasano, A.: Non-celiac gluten sensitivity: the new frontier of gluten related disorders. Nutrients. **5**, 3839–3853 (2013)

9. FAO-WHO, S| C.: Standards | CODEXALIMENTARIUS FAO-WHO. http://www.fao.org/fao-who-codexalimentarius/codex-texts/list-standards/en. Accessed 9/11/2018

10. Koerner, T.B., Abbott, M., Godefroy, S.B., Popping, B., Yeung, J.M., Diaz-Amigo, C., Roberts, J., Taylor, S.L., Baumert, J.L., Ulberth, F., Wehling, P., Koehler, P.: Validation procedures for quantitative gluten ELISA methods: AOAC allergen community guidance and best practices. J. AOAC Int. **96**, 1033–1040 (2013)

11. Mena, M.C.; Sousa, C.:25Analytical tools for gluten detection: Policies and regulation. OmniaSci. Monogra. 527–564 (2015). https://doi.org/10.3926/oms.264

12. Martínez-Esteso, M.J., Nørgaard, J., Brohée, M., Haraszi, R., Maquet, A., O'Connor, G.: Defining the wheat gluten peptide fingerprint via a discovery and targeted proteomics approach. J. Proteome. **147**, 156–168 (2016)

13. Zhang, W., Zhao, X.: Method for rapid protein identification in a large database. Biomed. Res. Int. **2013**, 414069 (2013)

14. Cunningham, R., Ma, D., Li, L.: Mass spectrometry-based proteomics and Peptidomics for systems biology and biomarker discovery. Front. Biol. (Beijing). **7**, 313–335 (2012)

15. Fallahbaghery, A., Zou, W., Byrne, K., Howitt, C.A., Colgrave, M.L.: Comparison of gluten extraction protocols assessed by LC-MS/MS analysis. J. Agric. Food Chem. **65**, 2857–2866 (2017)

16. Zou, Y., Yang, M., Zhang, G., He, H., Yang, T.: Antioxidant activities and phenolic compositions of wheat germ as affected by the roasting process. J. Am. Oil Chem. Soc. **92**, 1303–1312 (2015)

17. Bhattarai, R.R., Dhital, S., Wu, P., Chen, X.D., Gidley, M.J.: Digestion of isolated legume cells in a stomach-duodenum model: three mechanisms limit starch and protein hydrolysis. Food Funct. **8**, 2573–2582 (2017)

18. Vensel, W.H., Dupont, F.M., Sloane, S., Altenbach, S.B.: Effect of cleavage enzyme, search algorithm and decoy database on mass spectrometric identification of wheat gluten proteins. Phytochemistry. **72**, 1154–1161 (2011)

19. Dupont, D., Mackie, A.R.: Static and dynamic in vitro digestion models to study protein stability in the gastrointestinal tract. Drug Discov. Today Dis. Model. **17–18**, 23–27 (2015)

20. Colgrave, M.L., Byrne, K., Howitt, C.A.: Food for thought: selecting the right enzyme for the digestion of gluten. Food Chem. **234**, 389–397 (2017)

21. Zou, W., Sissons, M., Gidley, M.J., Gilbert, R.G., Warren, F.J.: Combined techniques for characterising pasta structure reveals how the gluten network slows enzymic digestion rate. Food Chem. **188**, 559–568 (2015)

22. Minekus, M., Alminger, M., Alvito, P., Ballance, S., Bohn, T., Bourlieu, C., Carrière, F., Boutrou, R., Corredig, M., Dupont, D., Dufour, C., Egger, L., Golding, M., Karakaya, S., Kirkhus, B., Le Feunteun, S.,

Lesmes, U., Macierzanka, A., Mackie, A., Marze, S., McClements, D.J., Ménard, O., Recio, I., Santos, C.N., Singh, R.P., Vegarud, G.E., Wickham, M.S.J., Weitschies, W., Brodkorb, A.: A standardised static *in vitro* digestion method suitable for food – an international consensus. Food Funct. **5**, 1113–1124 (2014)

23. Schalk, K., Koehler, P., Scherf, K.A.: Quantitation of specific barley, rye, and oat marker peptides by targeted liquid chromatography–mass spectrometry to determine gluten concentrations. J. Agric. Food Chem. **66**, 3581–3592 (2018)

24. Prandi, B., Faccini, A., Tedeschi, T., Cammerata, A., Sgrulletta, D., D'Egidio, M.G., Galaverna, G., Sforza, S.: Qualitative and quantitative determination of peptides related to celiac disease in mixtures derived from different methods of simulated gastrointestinal digestion of wheat products. Anal. Bioanal. Chem. **406**, 4765–4775 (2014)

25. Boukid, F., Prandi, B., Buhler, S., Sforza, S.: Effectiveness of germination on protein hydrolysis as a way to reduce adverse reactions to wheat. J. Agric. Food Chem. **65**(28), 5831-5836 (2017). https://doi.org/10.1021/acs.jafc.7b03175

26. Prandi, B., Tedeschi, T., Folloni, S., Galaverna, G., Sforza, S.: Peptides from gluten digestion: a comparison between old and modern wheat varieties. Food Res. Int. **91**, 92–102 (2017)

27. Boukid, F., Prandi, B., Sforza, S., Sayar, R., Seo, Y.W., Mejri, M., Yacoubi, I.: Understanding the effects of genotype, growing year, and breeding on Tunisian durum wheat allergenicity. 1. The Baker's Asthma Case. J. Agric. Food Chem. **65**, (2017). https://doi.org/10.1021/acs.jafc.7b02040

28. Sollid, L.M., Qiao, S.-W., Anderson, R.P., Gianfrani, C., Koning, F.: Nomenclature and listing of celiac disease relevant gluten T-cell epitopes restricted by HLA-DQ molecules. Immunogenetics. **64**, 455–460 (2012)

29. Cornell, H.J., Wills-Johnson, G.: Structure-activity relationships in coeliac-toxic gliadin peptides. Amino Acids. **21**, 243–253 (2001)

30. Almeida, L.M., Gandolfi, L., Pratesi, R., Uenishi, R.H., de Almeida, F.C., Selleski, N., Nóbrega, Y.K.d.M.: Presence of DQ2.2 associated with DQ2.5 increases the risk for celiac disease. Autoimmune Dis. **2016**, 5409653 (2016)

31. Pisapia, L., Camarca, A., Picascia, S., Bassi, V., Barba, P., Del Pozzo, G., Gianfrani, C.: HLA-DQ2.5 genes associated with celiac disease risk are preferentially expressed with respect to non-predisposing HLA genes: implication for anti-gluten T cell response. J. Autoimmun. **70**, 63–72 (2016)

32. Prandi, B., Bencivenni, M., Tedeschi, T., Marchelli, R., Dossena, A., Galaverna, G., Sforza, S.: Common wheat determination in durum wheat samples through LC/MS analysis of gluten peptides. Anal. Bioanal. Chem. **403**, 2909–2914 (2012)

33. Juhász, A., Gell, G., Békés, F., Balázs, E.: The epitopes in wheat proteins for defining toxic units relevant to human health. Funct. Integr. Genomics. **12**, 585–598 (2012)

34. Boukid, F., Prandi, B., Sforza, S., Sayar, R., Seo, Y.W., Mejri, M., Yacoubi, I.: Understanding the effects of genotype, growing year, and breeding on Tunisian durum wheat allergenicity. 2. The Celiac Disease Case. J. Agric. Food Chem. **65**, 5837–5846 (2017)