IDENTIFICATION OF EARLY-REPRESENTED GLUTEN PROTEINS DURING DURUM WHEAT GRAIN DEVELOPMENT

Maria Fiorella Mazzeo[1], Luigia Di Stasio[1,2], Chiara D'Ambrosio[3], Simona Arena[3], Andrea Scaloni[3], Simona Corneti[4], Aldo Ceriotti[5], Roberto Tuberosa[4], Rosa Anna Siciliano[1], Gianluca Picariello[1] and Gianfranco Mamone[1]*

[1] Institute of Food Sciences National Research Council (CNR), 83100 Avellino, Italy

[2] Department of Agriculture, University of Naples "Federico II", 80100 Portici, Italy

[3] Proteomics & Mass Spectrometry Laboratory, ISPAAM, National Research Council (CNR), 80147 Naples, Italy

[4] Department of Agricultural Sciences, University of Bologna, 40127 Bologna, Italy

[5] Institute of Agricultural Biology and Biotechnology, National Research Council, 20133 Milan, Italy

*Corresponding author: Dr. Gianfranco Mamone

Institute of Food Sciences – CNR, via Roma 64, I-83100 Avellino, Italy, Phone +39 0825 299521;

Fax: +39 0825 781585, e-mail: mamone@isa.cnr.it

ABSTRACT

The time course of biosynthesis and accumulation of storage proteins in developing grain of durum wheat (*Triticum turgidum* ssp. *durum* (Desf.) Husn.) pasta-quality reference cv. Svevo was investigated at the protein level for the first time. Seeds were harvested at key kernel developmental stages, namely 3 (seed increase threefold in size), 5 (kernel development, water-ripe stage), 11 (kernel development, water-ripe stage), 16 (kernel full development, water-ripe stage), 21 (milk-ripe stage) and 30 (dough stage) days post-anthesis (dpa). Gliadins and glutenins were fractionated according to their different solubility and individually analyzed after fractionation by reversed-phase high performance liquid chromatography and sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Proteins were identified by liquid chromatography-tandem mass spectrometry of proteolytic peptides. The α- and γ-gliadin were already detected at 3 dpa. The biosynthesis of high molecular mass glutenin Bx7 was slightly delayed (11 dpa). Most of gluten proteins accumulated rapidly between 11 and 21 dpa, with minor further increase up to 30 dpa. The expression pattern of gluten proteins in *Triticum durum* at the early stages of synthesis provides reference datasets for future applications in crop breeding and growth monitoring.

Keywords: durum wheat; grain development; proteomics; gliadin; glutenin

INTRODUCTION

Wheat cereals are among the most important food sources for humans. Wheat kernels include starch and non-starch polysaccharides along with a heterogeneous fraction of proteins, which are typically classified in water saline soluble albumins/globulins and insoluble gluten proteins. Albumins/globulins constitute approximately 20% of the endosperm protein and rule the metabolic pathways during grain development and germination. Gluten constitutes the major storage proteins (about 80%) and contributes to confer the unique elasticity and extensibility rheological properties to wheat doughs.[1] Gluten proteins embrace monomeric gliadins and polymeric glutenin networks. Gliadins are alcohol soluble proteins traditionally sub-classified into α/β-, γ- and ω-fractions, according to their electrophoretic mobility.[2] Glutenins are constituted by low-molecular weight (LMW) and high-molecular weight (HMW) subunits, organized in complex hetero-polymers, which are stabilized by both inter- and intra-molecular disulfide bonds.[3,4] Therefore, glutenin subunits can be extracted in alcohol solution after reduction of disulfide bonds.[1]

The identification of wheat proteins is essential to understand the gene-function relationships and the effects of growth and storage conditions on wheat quality,[5] providing useful indicators of the technological quality of dough [4] Over the last decades, -omic approaches have been exploited as powerful tools to assess the diversity and the functional role of proteins expressed during wheat grain development, particularly in *T. aestivum*. Using a transcriptomic approach, Shewry and coworkers characterized the time course of synthesis and accumulation of the storage proteins in the endosperm of hexaploid wheat.[6] Transcripts of over 30,000 genes were found in the developing soft wheat grain.[7] Proteomics was applied on mature endosperm for assessing the main mechanisms involved in accumulation of storage proteins, during grain filling of soft wheat and other cereals.[9-13] Characterizations at protein level, carried out on both whole protein extracts and purified protein fractions, have mainly highlighted the expression changes of albumin/globulin (metabolic) proteins during kernel development. The assessment of albumins/globulins has been typically addressed by canonical proteomic tools, such as mono- and/or two-dimensional electrophoresis coupled with

tandem mass spectrometry (MS/MS).[14] However, due to peculiar structural features, the comprehensive characterization of gluten proteins suffers from a series of drawbacks: i) the extreme heterogeneity of protein subunits with closely related primary sequences;[15-16] ii) low occurrence of arginine and lysine residues, which dramatically complicates the MS/MS peptide sequencing of tryptic peptides;[17] iii) Coomassie staining suppression effects in gel electrophoresis of gliadins;[18] iv) still incomplete database annotation of protein sequences and divergence among the wheat accessions.[19] Altogether, these drawbacks have hampered a detailed definition of the expression pattern of gluten proteins over various stages of grain formation.

In the present study, for the first time the expression of storage proteins in durum wheat (*Triticum turgidum* ssp. *durum* (Desf.) Husn.) cv. Svevo was monitored at six stages of development during grain filling, between 3 and 30 day post anthesis (dpa), using specifically devised proteomic strategies.

MATERIAL AND METHODS

Chemicals

Trypsin and chymotrypsin, Tris-HCl, ammonium bicarbonate (AMBIC), KCl, guanidine-HCl, EDTA, dithiothreitol, modified Lowry assay kit, methanol, polyclonal antigliadin antibody, monoclonal peroxidase conjugated anti-rabbit IgG antibody, Tween 20, glacial acetic acid and HPLC-MS grade solvents were all provided by Sigma-Aldrich (Milan, Italy). Electrophoresis reagents were all from Bio-Rad (Milan, Italy).

Plant material and grain preparation

Triticum turgidum ssp. *durum* (Desf.) Husn.) cv. Svevo was selected for this study, because it is a well-known pasta-quality reference cultivar, widely used as a crossing parent world-wide.[20-22] In early spring 2015, Svevo was grown in a growth chamber under completely controlled conditions of photoperiod, air temperature and humidity, as well as horizontal laminar air flux. The growing cycle lasted for three months. Initially the plants were grown at 14/10 h day/night photoperiod and 18/14°C temperature up to the first node elongation stage. The plants were then grown under long days conditions (16/8 h day/night photoperiod) at 20/16 °C. Temperatures were increased to 24/18 °C from flowering to maturity. The experimental unit consisted of 4 L cylindrical pots, filled with a mix of 6 L peat moss, 3 L silt-loam soil mix, 3 L perlite, 4 L vermiculite, 4 L sand, 300 mg of complex mineral fertilizer. Eight petri-dish germinated Svevo's seedlings were transplanted per pot. Subsequently, at the two-leaf stage the growing seedlings were thinned to allow four uniform adult plants to grow to maturity per pot. A complex fertilizer with N:P:K at ratio 15:9:15 and microelements (NPK® Original Gold®, Compo Expert, Italy) was applied at the following consecutive stages: sowing (incorporated into the soil), three leaf stage, tillering, first and second elongated nodes and booting (50 g m^{-2} for the whole growing cycle). Soil moisture was kept at

optimal water availability conditions throughout the whole plant cycle by manually irrigating the pots to full field capacity every two days.

Each replicate consisted of six pots in total (24 adult plants in total) that were used for subsequent destructive grain sampling at the chosen grain developmental phases. The experiment consisted of three replicates in total. Wheat kernels were harvested at key developmental stages, namely at 3 d (embrio increase three-fold in size), 5 d (kernel development at 1/3 of full developmental size, water-ripening stage), 11 d (kernel development at half-full size, water-ripening stage), 16 d (kernel full development, water-ripening stage), 21 d (milk-ripening stage) and 30 d (dough stage) days post-anthesis (dpa), according to the reference kernel developmental stages reported in the CerealDB (http://bio-gromit.bio.bris.ac.uk/cerealgenomics/cgi-bin/grain3.pl). Kernel's development in growth chamber was daily inspected to ensure that samplings were conducted at the right developmental stage. Triplicate samples were harvested according to the above-mentioned six developed stages: 3 days post-anthesis (dpa) (T0), 5 dpa (T1), 11 dpa (T2), 16 dpa (T3), 21 dpa (T4) and 30 dpa (T5). Immediately after harvesting, kernels were stored at -80 °C, until their use for proteomic analysis.

Sample preparation

Kernels were dried and ground to a powder using a mortar and pestle. Proteins were fractionated according to the Osborne procedure, based on different solubility in aqueous and organic solvents. Pre-extraction of albumins and globulins from seed powder (100 mg) was achieved by stepwise addition (three times) of 1 mL of 50 mM Tris-HCl, 100 mM KCl, 5 mM EDTA, pH 7.8, containing a protease inhibitor cocktail (Sigma Aldrich). The supernatant was further investigated for proteomic analysis of albumins and globulins.[23] Resulting pellet was rinsed with 0.5 mL of 60% v/v ethanol (three times) for gliadin extraction.[16] Finally, glutenins were extracted with 50% 1-propanol, 80 mM Tris–HCl, pH 8.5, 1% w/v dithiothreitol at 60 °C, for 45 min, and alkylated with 4-vinylpirydine for 15 min, at 60 °C, as previously reported;[24] resulting products were precipitated

at -20 °C with 1-propanol, overnight. Glutenin pellet was solubilized in 6 M guanidine-HCl, 0.3 M Tris-HCl, 1 mM EDTA, pH 8.0. Each extraction step included magnetic stirring for 20 min, followed by centrifugation at 14,500 x *g*, for 15 min, at 4 °C. Protein concentration wad determined by a modified version of the Lowry assay.[25] Samples were aliquoted for further electrophoretic and chromatographic analysis.

SDS-PAGE analysis

Proteins were analysed by mono-dimensional SDS-PAGE (Bio-rad, Mini-Protean) on precast linear gels (TGX Gel 12% acrylamide). Glutenin samples were precipitated with 1-propanol for 2 h, at 4 °C; resulting protein pellets were solubilized in 20 µL of 6 M urea. Gliadin samples were dried using a rotating evaporator and protein pellets were then solubilized in 1 mL of 5% v/v formic acid. Aliquots of gliadin extracts (125 µL) were dried and protein pellets were solubilized in 10 µL of 6 M urea. Laemli buffer was added to 1 µL-samples aliquots, and samples were denatured in a boiling water bath for 5 min. Electrophoresis was performed at 100 V mA for about 1 h, at room temperature. After migration, gel was stained for 2 h with 0.5% w/v Coomassie Brilliant Blue R-250 in 50% methanol/10% glacial acetic acid (v/v) and destained with 30% methanol/10% glacial acetic acid (v/v), overnight. Gels were imaged with an Image-Scanner (GE Helthcare) and analysed by ImageQuant TL software (GE Helthcare).

Western blot analysis

Gliadin protein extracts were analysed by Western blotting. Following the SDS-PAGE analysis as above, gels were electroblotted onto a nitrocellulose membrane (Transblot nitrocellulose, Bio-Rad) at 120 V for 60 min. The membrane was blocked with 5% w/v non-fat milk (Bio-Rad) in Tris-buffered saline, 0.05% w/v Tween 20 (TBS-T) for 1 h, at room temperature. Subsequently, the membrane was incubated overnight at 4 °C with polyclonal anti-gliadin antibody, developed by immunizing rabbit with a whole gliadin extract, which was previously diluted 1/1,500 in TBS-T.

After washing with TBS-T, monoclonal peroxidase-conjugated anti-rabbit IgG antibody (Sigma_Aldrich) diluted in TBS-T (1/15,000) was applied to the membrane for 2 h, at room temperature. The membrane was extensively rinsed with TBS-T (3 × 10 min) and finally with TBS (1 × 10 min) before development. Chemiluminescence reagents (ECL Plus WB reagent, GE Healthcare) and X-ray film (Kodak, Chalons/Saône, France) were used to visualize the immunoreactive protein bands at various exposure times (0.5-5 min range).

HPLC analysis

Alcohol soluble gliadin and glutenin aliquots were analyzed immediately after extraction, using a modular HPLC Agilent 1100 series (Palo Alto, CA, USA). Chromatography was carried out using a narrow-bore reversed phase column C8 (Phenomenex 250 cm, 2 mm i.d, 3.6 ∝m particle diameter), at a flow-rate of 0.2 mL/min. The column temperature was 60 °C, and the eluent was monitored at 214 nm. Solvent A was 0.1% v/v trifluoroacetic acid (TFA) in water and solvent B was 0.1% v/v TFA in acetonitrile. Proteins were eluted using a linear gradient (25-55%) of solvent B in 100 min, following 10 min of isocratic elution at 25% B. Fractions were manually collected, dried down in a speed-vac and stored at -20 °C until further analysis. Protein amount in chromatographic fractions was grossly estimated through peak area integration.

Protein digestion

HPLC fractions were suspended in 100 mM AMBIC, pH 7.8, and enzymatically digested by trypsin or chymotrypsin (1:50 enzyme to protein, w/w ratio) overnight at 37 °C.
Coomassie stained protein band were manually excised from gels and destained by repeated washing with 25 mM AMBIC/acetonitrile (1/1, v/v). Proteins were in-gel hydrolyzed at 37 °C with trypsin or chymotrypsin (12.5 ng/µL) in 25 mM AMBIC. The peptides were extracted three times in 5% formic acid/acetonitrile (1/1, v/v). Peptide digests were were dried in a speed-vac, re-dissolved in 0.1% formic acid (35 µL) and analysed by LC-MS/MS.

LC-MS/MS analysis

Mass spectrometry analysis was performed using a Q Exactive Orbitrap mass spectrometer (Thermo Scientific, San Jose, CA, USA), online coupled with an Ultimate 3000 ultra-high performance liquid chromatography instrument (Thermo Scientific). Lyophilized digests (about 1 µg/analysis) were re-suspended in 0.1% v/v formic acid solution, loaded through a 5 mm long, 300 µm i.d. pre-column (LC Packings, USA) and separated by an EASY-Spray™ PepMap C18 column (2 µm, 15 cm x 75 µm) 3 µm particles, 100 Å pore size (Thermo Scientific™). Eluent A was 0.1% v/v formic acid in Milli-Q water; eluent B was 0.08% v/v formic acid in 80% acetonitrile. The column was equilibrated at 4% B. Peptides were separated applying a 4–40% gradient of eluent B over 40 min; the flow rate was 300 nL/min. The mass spectrometer operated in data-dependent mode and all MS1 spectra were acquired in the positive ionization mode with an *m/z* scan range of 350 to 1600. Up to 10 most intense ions in MS1 were selected for fragmentation in MS/MS mode. A resolving power of 70,000 full width at half maximum (FWHM), an automatic gain control (AGC) target of 3 x 106 ions and a maximum ion injection time (IT) of 100 ms were set to generate precursor spectra. MS/MS fragmentation spectra were obtained at a resolving power of 17,500 FWHM an automatic gain control (AGC) target of 1 x 105 ions and a maximum ion injection time (IT) of 80 ms. In order to prevent repeated fragmentation of the most abundant ions, a dynamic exclusion of 30 s was applied. Ions with one or more than six charges were excluded from fragmentation. Mass spectra were elaborated using the Proteome Discoverer 2.1 software (Thermo Fisher), restricting the research to a specific gliadin sequence database and a specific glutenin sequence database extracted from the NCBI Triticum protein database. Parameters for gliadin database searching were the following: Met oxidation and pyroglutamic for N-terminus Gln as variable protein modifications, a mass tolerance value of 10 ppm for precursor ion and 0.02 Da for MS/MS fragments, chymotrypsin as proteolytic enzyme, and a missed cleavage maximum value of

5. Parameters for glutenin database searching were the same, except for pyridylethyl Cys as a constant modification and trypsin as proteolytic enzyme.

The false discovery rate (FDR) and protein probabilities were calculated by Target Decoy PSM Validator working between 0.01 and 0.05 for strict and relaxed searches, respectively. Proteins were considered confidently identified relying on at least two sequenced peptides. Identifications based on only two peptides were validated by manual inspection of MS/MS spectra.

RESULTS

The Italian pasta-quality reference cv. Svevo was selected for investigating the time-course of gluten protein synthesis and accumulation in durum wheat. Growing grains were harvested at six stages of grain development (from 3 to 30 dpa) during the post-anthesis period, corresponding to early development (lag phase), accumulation of storage compounds (filling phase) and maturation-desiccation. Morphological details at each developmental stage are reported in supplementary material (Table S1). After depletion of albumin/globulin fractions, whose evolving sub-proteomes were compared separately,[23] gliadins and glutenins were selectively extracted based on their different solubility. The fractionation of the gluten proteins was necessary to roughly reduce the complexity of the crude protein system, thereby facilitating the identification of the components. The protein sub-families were individually analyzed by HPLC and SDS-PAGE, also to compare the time course of protein synthesis and accumulation of the two gluten protein fractions during grain development.

Analysis of gliadin fractions

The accumulation of individual gliadin proteins was monitored by SDS-PAGE analysis (Figure 1). No gliadin was detected at T0 stage, although a band at 14 kDa clearly appeared. This protein, also detected over the entire grain filling course, was identified as the non-gluten protein histone H4 (accession number AAA34292 1) (Table 1). An additional protein band close to 14 kDa, most likely corresponding to a different histone, appeared at T4 and T5. Gliadin content increased rapidly between T2 and T4, whereas only a faint further increase was appreciated at T5. Thus, at T4 the gliadin synthesis should be considered complete, showing the typical protein pattern, including α-, γ- and ω-gliadin subfamilies in order of increasing mobility (Figure 1).[19]

The comprehensive characterization of the gliadin proteome was obtained by LC-MS/MS analysis of the chymotryptic digests of HPLC-isolated subunits (Figure 2; Table S1). Although very reliable because of the high accuracy of the MS and MS/MS measurements, the identification of

gliadins often grounded on a few peptide matches, due to the poor quality of the fragmentation spectra of non-tryptic peptides, as well as to the incomplete database protein inventory of *T. durum*. Unlike SDS-PAGE, the HPLC analysis demonstrated that few gliadin protein components were expressed, though at a low level, even at T0 (Figure 2). These proteins eluted at retention times of 59.8 and 61.8 min, broadly in the region of α-gliadins,[24] as inferred by comparison with the retention times of ω-, γ- and α-gliadins extracted at T4 and T5. MS/MS-based sequencing of chymotryptic peptides revealed that the early expressed proteins in both fractions were α - and γ-gliadin gene products (Table 1; Table S3). Interestingly, during synthesis, gliadin products other than these latter became quantitatively predominating (Figure 2).

Western blotting of gliadin fractions

To overcome the limit of Coomassie-based detection of gluten proteins, the evolution of gliadin fraction over the early stages of maturation was monitored by Western blotting analysis using a commercial anti-gliadin (wheat) antibody (Figure 3). The immunochemical approach confirmed the expression of a protein band migrating at 50 kDa at a very early stage (T0), with intensity increasing over the time. No other immune reactive proteins in the T0-T2 timeframe were observed.

Analysis of glutenin fractions

Glutenins were extracted from maturing caryopses with a denaturing/reducing buffer and their accumulation was monitored by both SDS-PAGE (Figure 4) and HPLC (Figure 5). Unlike gliadins, the SDS-PAGE analysis of glutenins exhibited a comparable or even higher sensitivity than HPLC. The first glutenin subunit appearing was HMW Bx7 of about 90 kDa (Table 2; Table S4), which was detected as a faint band at T2 (11 dpa), while it escaped the HPLC detection. In contrast, no glutenin was detected at T0 and T1 stages. The major expression of glutenins started at T3, when HWM-glutenin Bx7 (band 1) and By8 (band 2) were clearly visible. The identity of HMW Bx7 and

By8 was also confirmed by exploring large sized tryptic peptides by MALDI–TOF analysis (not showed), according to a previously developed strategy.[26] In concomitance with synthesis of HMW-glutenins, a set of LWM glutenin subunits (bands 3-6, Figure 4) appeared at T3 (Table 3; Table S3). Accumulation of glutenins progressed at T4 and was complete at T5 stage. The glutenin patterns of *T. durum cv*. Svevo at T5 matched the corresponding subproteome characterized in a previous work.[26]

DISCUSSION

Several authors have investigated the changes of gluten proteins during the development of wheat kernel. However, due to the different analytical methods used, data about the temporal appearance and accumulation of individual gluten proteins at the subunit level remain fragmentary. Early studies based on acid-PAGE indicated that gliadins appear in the *T. aestivum* kernel at 10 dpa, while both HMW and LMW glutenins began to be detectable at 14 dpa.[27] Interestingly, at all stages of the plant development, the appearance of gluten protein subsets as well as the time-dependent trend of protein accumulation did not depend on the wheat cultivars. Two ω-gliadin subunits were found as firstly expressed in *T. aestivum* at 10 dpa.[6]

By using Western-blotting, *Iametti et al.* found that the accumulation of immune-reactive gluten proteins in *T. durum* started after 15 or 17 dpa, depending on the wheat accession.[28] By monitoring the expression of selected gluten subunit genes at the RNA level, *Katagiri et al.* found that gliadin and glutenin trascripts in soft wheat appeared even at 7 dpa, with changes in the expression pattern related to the cultivar.[29] These authors pointed out the presence of low molecular weight proteins (~15 kDa) at very early stages of grain-filling, which later disappeared probably due to the increased expression of other gene products. Most likely, these low molecular weight components correspond to the histones identified in this study. As proteomics, most of comparative studies aimed to investigate grain development at protein level were carried out by analyzing whole protein extracts from wheat seeds, and did not consider the first stages of grain development (3 and 5 dpa).[30-31]

In the present investigation, the HPLC-based monitoring allowed us to anticipate the detection of specific gliadin and glutenin components up to 3 (T0) and 11 (T2) dpa, respectively. The early-synthesized subunits were here identified as α- and γ-gliadins (Table 2), clearly detected even at 3 dpa (T0). Whilst SDS-PAGE outcomes were in part in line with previous results,[6] HPLC analysis provided deeper insights. This apparent incongruence is probably the result of the higher sensitivity

of the HPLC detection of gliadins compared to that of SDS-PAGE (with Coomassie detection), due to poor staining of these proteins.[18]

Gliadins and glutenins steadily increase in number and amount during grain development. In agreement with previous results, wheat grains at the milk phase (T4, 21 dpa), coinciding with the end of the cellular division phase, exhibited a content of gluten proteins very similar to that at full maturation.[27-28]

Immature wheat, harvested before the milk stage, has been proposed as a potential source of functional nutrients, because of the significant content of fructo-oligosaccharides and the incomplete pattern of possible immunogenic proteins.[28] In the light of the results herein, a dedicated determination of gluten proteins is required to propose immature kernel as a possible foodstuff suited for gluten-sensitive individuals.[29]

In line with the current outcomes, most of the previous investigations highlighted a delayed synthesis and a broadly slower accumulation of glutenin subunits during the early- and mid-stage of development, compared to gliadins.[6,32] HMW glutenins appear before LMW ones, suggesting that a backbone of HMW glutenins could create a polymeric scaffold in the developing endosperm onto which LMW subunits might branch.[33] In the mid phase, LMW subunits prevail over the HMW ones, whereas at later stages HPLC profiles showed a relative increased amount of HMW glutenin subunits, which could further progress during the desiccation stage of the kernel starting at 42 dpa.[6]

In conclusion, the accumulation of gluten proteins is a complex process undergoing spatial and temporal regulation and affected by environmental biotic and abiotic signaling.[34] The use of HPLC-based separation of gluten proteins allowed us to detect specific gluten proteins expressed at earlier stages than previously described. The evolution of gluten proteins observed in this study contribute to a better understanding of the mechanisms that regulate the expression and the accumulation of gluten proteins during durum grain filling and to the elucidation of physiological aspects of the gene-function correlation. The identification of the early-expressed subunits can help to clarify

some still unaddressed issues related to the formation of the gluten network, which is a key information from a technological standpoint.

In view of the forthcoming release of the genome sequence of cv. Svevo (Cattivelli et al, unpublished data), these findings provide a timely reference dataset for future studies, especially for those aimed to evaluate the genetic control of the development of durum wheat storage proteins under a range of stress conditions or environmental changes. Though preliminary, this approach can further support trascriptomic analysis with supplementary and complementary data, since it comprises the study of functional rather than informational molecules.

Abbreviations used

AMBIC, ammonium bicarbonate, dpa, days post-anthesis; HMW, high molecular weight; LMW, low molecular weight; MS/MS, tandem mass spectrometry; T0, 3 days post-anthesis; T1, 5 days post-anthesis; T2, 11 days post-anthesis ;T3, 16 days post-anthesis; T4, 21 days post-anthesis; T5, 30 days post-anthesis; TCA, trifluoroacetic acid.

Acknowledgements

This work was supported by a grant from MIUR-CNR: Progetto Bandiera "*InterOmics" Sviluppo di una piattaforma integrata per l'applicazione delle scienze "omiche" alla definizione dei biomarcatori e profili diagnostici, predittivi, e teranostici*".

Supporting information description

Table S1, Morphological details at each developmental stage of Svevo cv.

Table S2, Gliadin proteins identified by LC-MS/MS of derived peptides: Gliadin at T5 stage were fractionated by HPLC (Figure 1) and digested by chymotrypsin. The identity of gliadin proteins were obtained by LC-MS/MS analysis.

Table S3, Identification of gliadin proteins at T0 stage: Gliadin T0 stage were fractionated by HPLC (Figure 1) and digested by chymotrypsin. The identity of gliadin proteins were obtained by LC-MS/MS analysis.

Table S4, Identification of glutenin proteins at T3 stage: Glutenin at T3 stage (Figure 5) were identified by MS-based proteomic analysis.

REFERENCES

(1) Bonomi, F.; Ferranti, P.; Mamone, G. Wheat Flour: Chemistry and Biochemistry. In *Bakery Products Science and Technology*. Weibiao, Z., Eds.; Wiley-Blackwel. **2014**, 55-74.

(2) Woychik, J. H.; Boundy, J. A.; Dimler, R. J. Starch gel electrophoresis of wheat gluten proteins with concentrated urea. *Arch. Biochem. Biophys.* **1961**, 94, 477–482.

(3) Wieser, H. Chemistry of gluten proteins. *Food Microbiol.* **2007**, 24, 115-119.

(4) Wrigley, C.W. Giant proteins with flour power. *Nature.* **1996**, 381, 738–739.

(5) Lafiandra, D.; Masci, S.; Blumenthal, C; Wrigley, C.W. The formation of glutenin polymer in practice. *Cereal Foods World.* **1999**, 44, 572–578.

(6) Shewry, P.R.; Underwood, C.; Wan, Y., Lovegrove, A.; Bhandari, D.; Toole, G.; E.N. Mills, C.; Denyer, K.; Mitchell, R.A.C. Storage product synthesis and accumulation in developing grains of wheat. *J. Cereal Sci*. **2009**, 50, 106-112.

(7) Wan Y.; Poole R.L.; Huttly A.K.; Toscano-Underwood C.; Feeney K.; Welham S.; Gooding M.J.; Mills C.; Edwards K.J.; Shewry P.R.; Mitchell R.A. Transcriptome analysis of grain development in hexaploid wheat. *BMC Genomics*. **2008**, 6, 9-121.

(8) Vensel, W.; Tanaka, C. K.; Cai, N.; Wrong, J.H.; Buchanan, B.B.; Hurkman, W.J. Developmental changes in the metabolic protein profiles of wheat endosperm. *Proteomics*. **2005**, 5, 1594–1611.

(9) Hurkman, W. J.; Tanaka, C. K.; Extraction of wheat endosperm proteins for proteome analysis. *J. Chromatogr*. **2007**, 849, B344–B350.

(10) Nadaud, I.; Girousse, C.; Debiton, C.; Chambon, C.; Bouzidi, M.F.; Martre, P.; Branlard, G. Proteomic and morphological analysis of early stages of wheat grain development. *Proteomics.* **2010** 10, 2901-10.

(11) Yang, M., Dong, J., Zhao, W., Gao, X. Characterization of proteins involved in early stage of wheat grain development by iTRAQ. *J Proteomics*. **2016**, 136:157-66.

(12) Finnie, C.; Melchior, S.; Roepstorff, P.; Svensson, B. (2000). Proteome analysis of grain filling

393 and seed maturation in barley. *Plant Physiol*. **2002**, 129, 1308-19.
394 (13) Xu, H.; Zhang, W.; Gao, Y.; Zhao, Y.; Guo, L. Wang J. Proteomic analysis of embryo
395 development in rice (Oryza sativa). *Planta.* **2012**, 235, 687-701.
396 (14) Mamone, G.; Picariello, G.; Caira, S.; Addeo, F., Ferranti, P. Analysis of food peptides by
397 mass spectrometry – based techniques. *J. Chromatography A*. **2009**, 216, 7130-42.
398 (15) Shewry PR, Halford NG, Lafiandra D. Genetics of wheat gluten proteins. *Adv Genet*. **2003**, 49,
399 111-84.
400 (16) Mamone, G.; Ferranti, P.; Chianese, L.; Scafuri, L.; Addeo, F. Qualitative and quantitative
401 analysis of wheat gluten proteins by liquid chromatography and electrospray mass
402 spectrometry. *Rapid Commun Mass Spectrom.* **2000**, 14, 897-904.
403 (17) Mamone G, Addeo F, Chianese L, Di Luccia A, De Martino A, Nappo A, Formisano A, De
404 Vivo P, Ferranti P. Characterization of wheat gliadin proteins by combined two-dimensional
405 gel electrophoresis and tandem mass spectrometry. *Proteomics*. **2005**,11, 2859-65.
406 (18) Clements, R.L. Alternative methods for fixing and staining gliadins in polyacrylamide gels.
407 *Electrophoresis.* **1990**, 11, 123-8.
408 (19) Ferranti, P.; Mamone, G.; Picariello, G.; Addeo, F. Mass spectrometry analysis of gliadins in
409 celiac disease. *J Mass Spectrom.* **2007**, 2, 1531-48.
410 (20) Arduini, I.; Masoni, A.; Ercoli, L.; Mariotti, M. Grain yield, and dry matter and nitrogen
411 accumulation and remobilization in durum wheat as affected by variety and seeding rate.
412 *Eur J Agron*, **2006**, 25, 309-318.
413 (21) Maccaferri, M.; Sanguineti, M.C.; Corneti, S.; Araus Ortega, J.L.; Ben Salem, M.; Bort J.;
414 DeAmbrogio, E.; Garcia del Moral, L.F.; Demontis, A.; El-Ahmed, A.; Maalouf, F.;
415 Machlab, H.; Martos, V.; Moragues, M.; Motawaj, J.; Nachit, M.; Nserallah, N.; Ouabbou,
416 H.; Royo, C.; Slama, A.; Tuberosa, R. Quantitative trait loci for grain yield and adaptation
417 of durum wheat (*Triticum durum* Desf.) across a wide range of water availability. *Genetics.*
418 **2008**, 178, 489-511.

(22) Blanco, A.; Mangini, G.; Giancaspro, A.; Giove, S.; Colasuonno, P.; Simeone, R.; Signorile, P.; De Vita, P.; Mastrangelo, A.M.; Cattivelli, L.; Gadaleta, A. Relationships between grain protein content and grain yield components through quantitative trait locus analyses in a recombinant inbred line population derived from two elite durum wheat cultivars. *Molecular Breeding.* **2012**, 30, 79-92.

(23) Arena, S.; Vitale, M.; Mazzeo, F; Mamone, G., Di Stasio, L.; Maccaferri, M.; Curci PL, Corneti S, Tuberosa, R.; Zambrano, N.; Scaloni, A.;D'Ambrosio, C. Differential representation of albumins and globulins during grain development in durum wheat. *Submitted* ***2*****016**.

(24) Marchylo, B.A.; Hatcher, D.W; Kruger, J.E. Identification of wheat cultivars by reversed-phase high-performance liquid chromatography of storage proteins. *Cereal Chem.* **1998**, 65, 28-40.

(25) Peterson, G.L. A simplification of the protein assay method of Lowry et al. which is more generally applicable. *Anal Biochem*. **1977**, 83, 346-56.

(26) Mamone, G., De Caro, S.; Di Luccia, A.; Addeo, F.; Ferranti P. Proteomic-based analytical approach for the characterization of glutenin subunits in durum wheat. *J Mass Spectrom*. *2009* 44, 1709-23.

(27) Johansson, E.; Oscarson, P.; Heneen, W.K.; Lundborg, T. Differences in Accumulation of Storage Proteins between Wheat Cultivars during Development. *J Sci Food* Agric. **1994**, 64, 305-313.

(28) Iametti, S.; Bonomi, F.; Pagani, M.A.; Zardi, M.,; Cecchini, C.; D'Egidio, M.G. Properties of the protein and carbohydrate fractions in immature wheat kernels. *J Agric Food Chem.* **2006**, 54, 10239-44.

(29) Katagiri, M.; Masuda, T.; Tani, F.; Kitabatake, N. Expression and Development of Wheat Proteins during Maturation of Wheat Kernel and the Rheological Properties of Dough Prepared from the Flour of Mature and Immature Wheat. *Food Sci. Technol. Res.*, **2011**, 17, 111–120.

(30) Ma C, Zhou J, Chen G, Bian Y, Lv D, Li X, Wang Z, Yan Y. iTRAQ-based quantitative proteome and phosphoprotein characterization reveals the central metabolism changes involved in wheat grain development. *BMC Genomics*. **2014**, 27, 15:1029.

(31) Cao H, He M, Zhu C, Yuan L, Dong L, Bian Y, Zhang W, Yan Y. Distinct metabolic changes between wheat embryo and endosperm during grain development revealed by 2D-DIGE-based integrative proteome analysis. *Proteomics*. **2016**, 10, 1515-36.

(32) Hurkman, W. J.; Tanaka, C. K.; Vensel, W.H.; Thilmony, R.; Altenbach, S.B. Comparative proteomic analysis of the effect of temperature and fertilizer on gliadin and glutenin accumulation in the developing endosperm and flour from Triticum aestivum L. cv. Butte 86. *Proteome Sci*, **2013**, 11, 8.

(33) Rathmell, W.G.; Skylas, D.J.; Bekes, F.; Wrigley, C.W. Wheat-grain proteomics; the full complement of proteins in developing and mature grain. In *Wheat Gluten*; Shewry, P.R., Tatham, A.S.; Eds.; Royal Society of Chemistry, Cambridge, UK, 2000; pp. 117–121.

(34) Giuliani, M.M.; Palermo, C.; De Santis, M.A.; Mentana, A.; Pompa, M.; Giuzio, L.; Masci, S.; Centonze, D.; Flagella, Z. Differential Expression of Durum Wheat Gluten Proteome under Water Stress during Grain Filling. *J Agric Food Chem*. **2015**, 63, 6501-12.

FIGURE CAPTIONS

FIGURE 1

SDS-PAGE (12%) of gliadin extracts from developing caryopses (T0-T5). Gel was stained with Coomassie Blue R-250. Gliadin-type proteins were distinguished into three groups based on their electrophoretic mobility, *i.e.* ω-, α/β-, and γ-gliadins [*Ferranti et al 2007*]. Molecular mass standards are indicated on the left side. Band 1 at T0 stage was identified as the histone H4 (accession number AAA34292.1) (Table 1).

FIGURE 2

Comparison among HPLC chromatograms of gliadin proteins from developing caryopses (T0-T5). Gliadins were assigned to α-, γ-, and ω-gliadin subfamilies according to retention times [Marchylo 1996]. Insets at T0 shows an expanded vied of the two closely eluting α- and γ- gliadin gene products (Table 1 and Table S3). Fractions collected at T5 were individually analysed by LC-MS/MS for a comprehensive characterization of the gliadin proteome (Table S2).

FIGURE 3

Western blotting analysis of gliadin protein extracts from developing caryopses (T0-T5) with commercial anti-gliadin (wheat) antibody. Arrows indicate the immunoreactive protein bands at T0, T1 and T2, which were not detectable by Coomassie R250 staining.

FIGURE 4

SDS-PAGE (12%) of glutenin extracts from developing caryopses (T0-T5a). Gel was stained with Coomassie Blue R250. Gluten-type proteins were distinguished into HMW and LMW subunits

according to their migration. Molecular mass standards are shown on the left side. Bands from T2 and T3 were excised, and analysed by LC-MS/MS after *in-gel* trypsin digestion. Protein contained in band 1 at T2 was HMW glutenin Bx7 (Table 2 and Table S4).

FIGURE 5

Comparison among HPLC chromatograms of glutenin proteins from developing caryopses (T0-T5). Glutenin subunits were classified in HMW and LMW according to their retention times.[24] Proteins were clearly visible by HPLC only after T3 stage.

505

506 **Table 1:** Gliadin proteins identified by LC-MS/MS of derived peptides. Detailed information on

507 protein identification are shown in supplementary Table SI.

dpa stage	Sample	Protein name	Accession
T0	Band 1 (SDS-PAGE)	Istone H4	AAA34292.1
	LC1 (HPLC, peak 1)	α-gliadin α-gliadin α-gliadin γ-gliadin γ-gliadin	SCW25728.1 BAA12318.1 CAB76962.1 ACJ03505.1 ACI04110.1
	LC2 (HPLC, peak 2)	α-gliadin α-gliadin α-gliadin γ-gliadin γ-gliadin	BAA12318.1 SCW25728.1 AGI15858.1 CAB76962.1 ACI04110.1

508

509 **Table 2:** Glutenin proteins identified by LC-MS/MS of derived peptides. A detailed information on

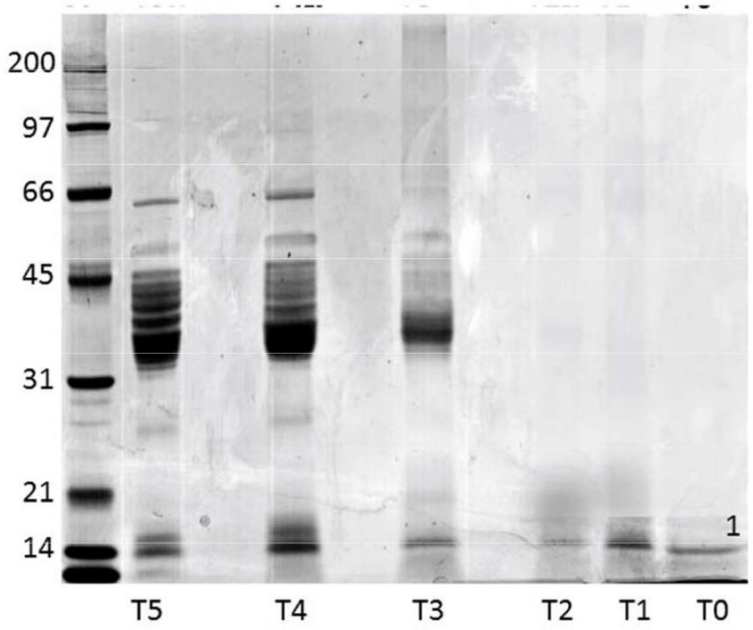
510 protein identification are shown in supplementary Table SII.

511

dpa stage	Sample	Protein name	Accession
T2	Band 1 (SDS-PAGE)	HMW Bx7 [a]	ABY59654.1
T3	Band 1 (SDS-PAGE)	HMW Bx7 [a]	ABY59654.1
	Band 2 (SDS-PAGE)	HMW By8	AE045111.1
	Band 3 (SDS-PAGE)	LMW B-type LMW B-type LMW B-type group 3 type II LMW B-type LMW B-type type I LMW-glutenin LMW-i1 LMW glutenin subunit protein 1-50 LMW glutenin LMW glutenin LMW glutenin LMW-s1 LMW glutenin LMW-m glutenin subunit 53 LMW glutenin subunit Glu-D3 LMW-s KS2 LMW-glutenin P3-5 LMW-i glutenin subunit 1 LMW glutenin	ACA63864.1 BAJ09461.1 BAB78740.1 AAS10193.1 AAV75997.1 ABQ50899.1 ABX84154.1 AFB35206.1 ALN96402.1 ABG76009.1 ABM73527.1 AAS10192.1 AGU91705.1 AHN55176.1 BAD42431.1 ABE77188.1 AGU91655.1 AFB35204.1
	Band 4 (SDS-PAGE)	LMW-s1 LMW B-type group 3 type II	ABM73527.1 BAB78740.1
	Band 5 (SDS-PAGE)	LMW glutenin subunit B3-1, LMW glutenin LMW glutenin LMW glutenin glutenin subunit LMW glutenin subunit D3-3	ACZ51337.1 CAA59339.1 AEI00683.1 ACA63864.1 ABM73527.1 ACY08822.1
	Band 6 (SDS-PAGE)	γ-gliadin B LMW glutenin	P06659.1 ACA63864.1

512

513 **Figure 1**

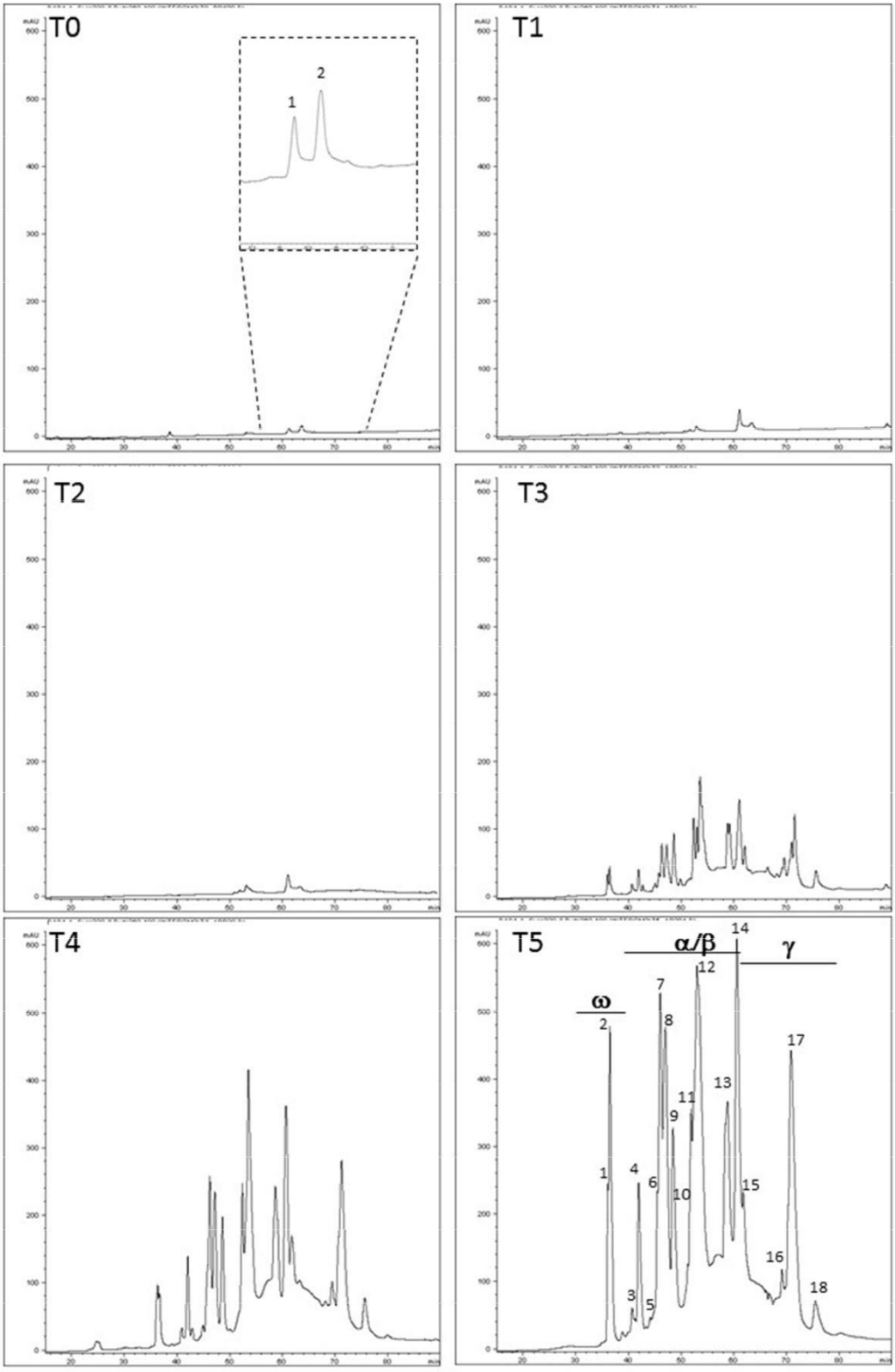


514

515

516

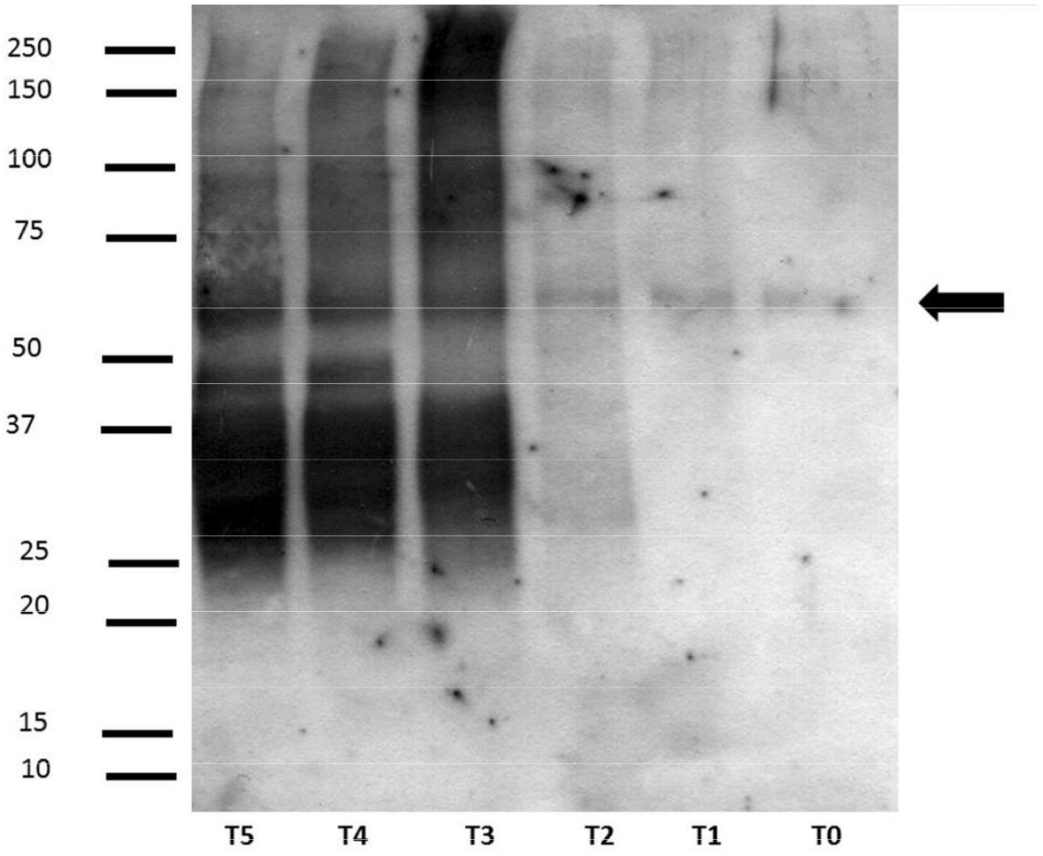
517 **Figure 2**



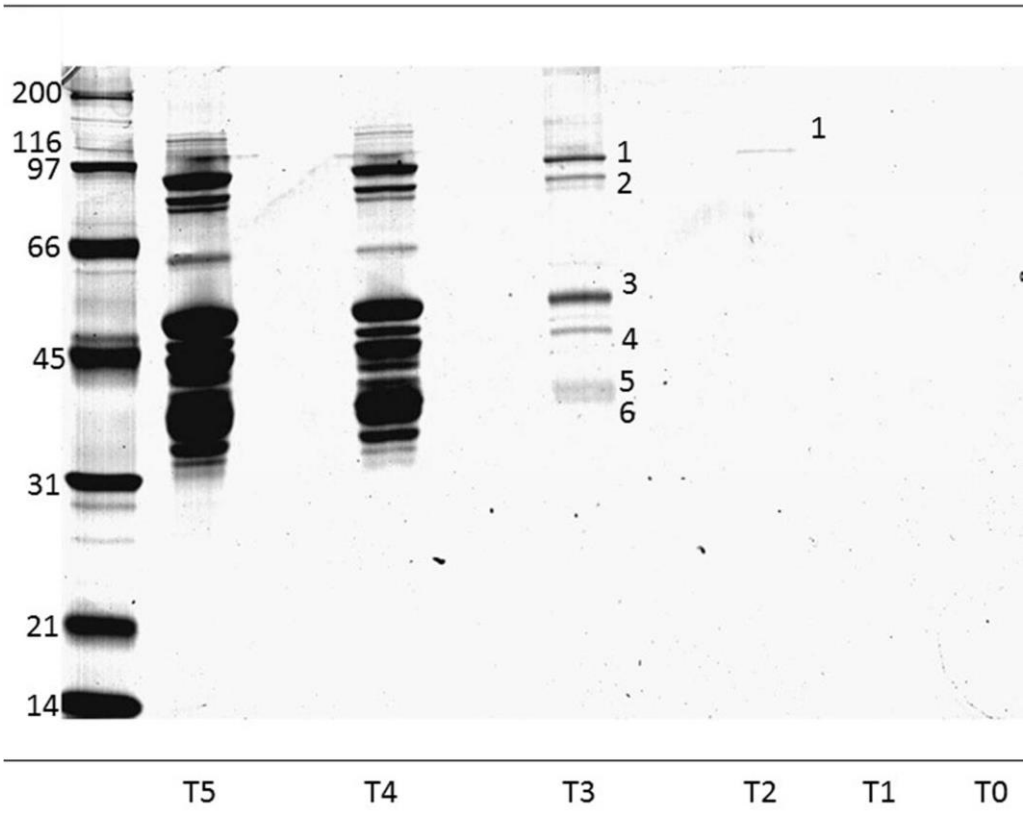
518

519

Figure 3



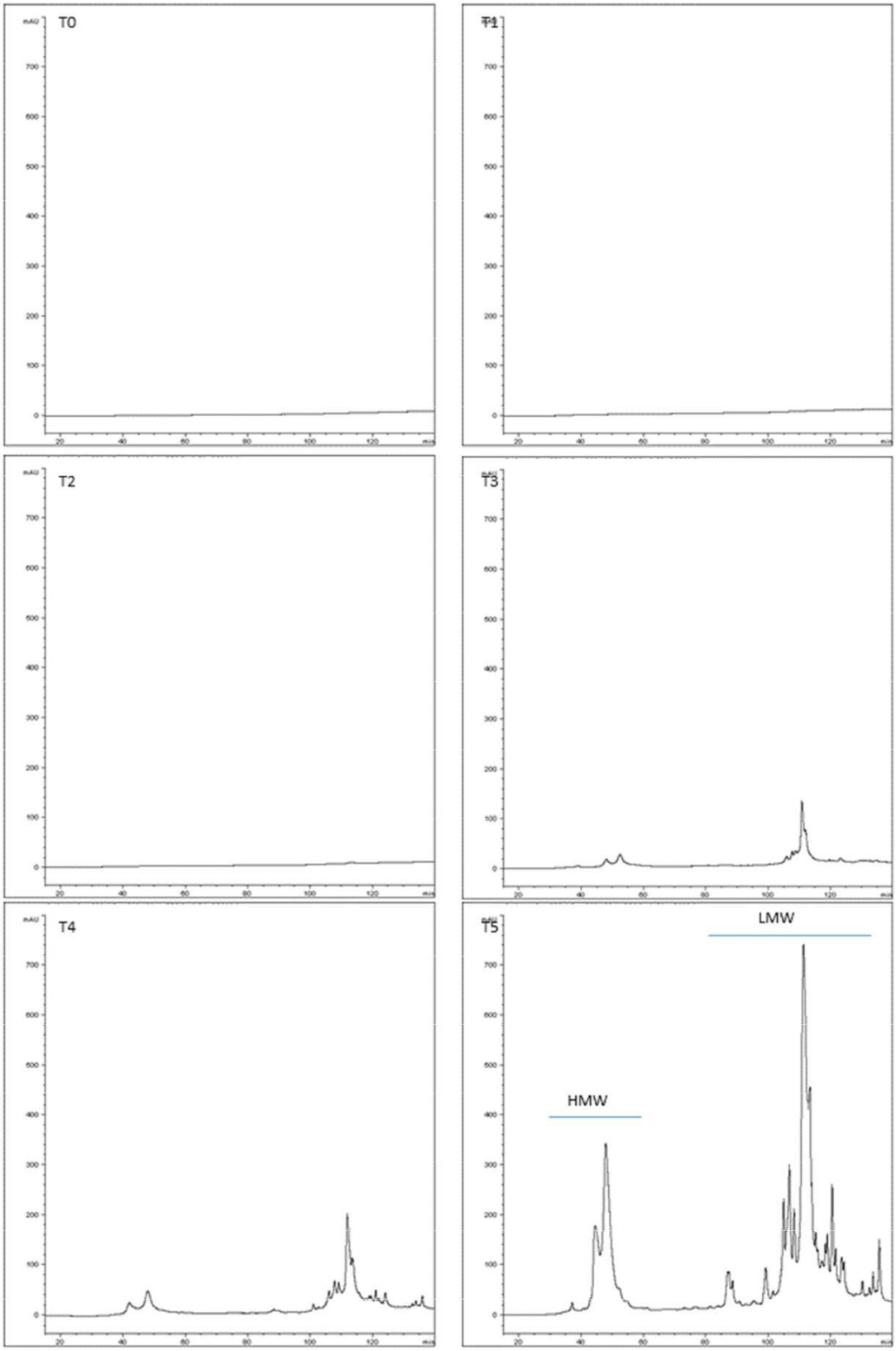
526 **Figure 4**



527

528

Figure 5



Graphic for table of contents