

RESEARCH ARTICLE

Open Access

An asparagine residue at the N-terminus affects the maturation process of low molecular weight glutenin subunits of wheat endosperm

Eleonora Egidi¹, Francesco Sestili¹, Michela Janni^{1,4}, Renato D'Ovidio¹, Domenico Lafiandra¹, Aldo Ceriotti², William H Vensel³, Donald D Kasarda³ and Stefania Masci^{1*}

Abstract

Background: Wheat glutenin polymers are made up of two main subunit types, the high- (HMW-GS) and low- (LMW-GS) molecular weight subunits. These latter are represented by heterogeneous proteins. The most common, based on the first amino acid of the mature sequence, are known as LMW-m and LMW-s types. The mature sequences differ as a consequence of three extra amino acids (MET-) at the N-terminus of LMW-m types. The nucleotide sequences of their encoding genes are, however, nearly identical, so that the relationship between gene and protein sequences is difficult to ascertain.

It has been hypothesized that the presence of an asparagine residue in position 23 of the complete coding sequence for the LMW-s type might account for the observed three-residue shortened sequence, as a consequence of cleavage at the asparagine by an asparaginyl endopeptidase.

Results: We performed site-directed mutagenesis of a LMW-s gene to replace asparagine at position 23 with threonine and thus convert it to a candidate LMW-m type gene. Similarly, a candidate LMW-m type gene was mutated at position 23 to replace threonine with asparagine. Next, we produced transgenic durum wheat (cultivar Svevo) lines by introducing the mutated versions of the LMW-m and LMW-s genes, along with the wild type counterpart of the LMW-m gene.

Proteomic comparisons between the transgenic and null segregant plants enabled identification of transgenic proteins by mass spectrometry analyses and Edman N-terminal sequencing.

Conclusions: Our results show that the formation of LMW-s type relies on the presence of an asparagine residue close to the N-terminus generated by signal peptide cleavage, and that LMW-GS can be quantitatively processed most likely by vacuolar asparaginyl endoproteases, suggesting that those accumulated in the vacuole are not sequestered into stable aggregates that would hinder the action of proteolytic enzymes. Rather, whatever is the mechanism of glutenin polymer transport to the vacuole, the proteins remain available for proteolytic processing, and can be converted to the mature form by the removal of a short N-terminal sequence.

Keywords: Asparaginyl endopeptidase, Gluten protein maturation, Low molecular weight glutenin subunits, Proteomic analysis, Genetic transformation, Transgenic plants, Wheat

* Correspondence: masci@unitus.it

¹DAFNE, Tuscia University, Viterbo, Italy

Full list of author information is available at the end of the article

Background

Wheat is the most widely consumed food crop in the world, being processed to give a wide range of products, such as bread, pasta, biscuits, and noodles. This unique versatility is mostly due to gluten proteins, the cohesive mass remaining after washing out the starch granules and water soluble components from a wheat dough. Gluten was first described by Beccari in 1728 (translated by Bailey [1]).

Gluten proteins, belonging to the “prolamin superfamily” [2], are composed of gliadins and glutenins. Gliadins are monomeric proteins, whereas glutenins are polymeric proteins, considered among the largest natural protein molecules [3]. Glutenin polymers are made up of protein subunits of high (HMW-GS) and low (LMW-GS) molecular weight, linked together by intermolecular disulfide bonds. The size and composition of the glutenin polymers play an important role in determining dough rheological properties [4,5].

The HMW-GS are better characterized than LMW-GS, since these latter are much more numerous and heterogeneous. LMW-GS can be classified according to different criteria, based on their primary structure, electrophoretic mobility in SDS-PAGE and N-terminal amino acid sequence. LMW-GS are distinguished into typical types, namely those proteins having a particular primary structure, and gliadin-like types, namely those LMW-GS that are gliadins according to the primary structure, but functionally act like glutenin subunits, due to the presence of an odd number of cysteine residues. The odd number of cysteines enables these gliadin-like proteins to form intermolecular disulfide bonds that incorporate them into glutenin polymer. Based largely on their order of increasing electrophoretic mobility in SDS-PAGE, LMW-GS are classified into B, C and D groups. The B group is composed primarily of typical LMW-GS, whereas the majority of C and D subunits correspond to gliadin-like proteins. The C subunits include mostly α - and γ -gliadin-like LMW-GS, whereas the D group includes ω -gliadin-like proteins, which may be slower moving than B subunits (rev. in [6]).

All these proteins are initially targeted to the endoplasmic reticulum, where signal sequence cleavage occurs, and then transported to the protein storage vacuole. While post-translational modifications of storage proteins have been extensively studied in legumes, they have received less attention in cereals where, with the notable exception of rice, prolamins constitute the majority of the accumulated polypeptides. Identification of distinct processing events would be very useful, not only to understand the mechanisms that lead to accumulation of the mature proteins, but also as a tool to monitor their intracellular transport. Information on post-translational events involved in the biosynthesis of prolamins is indeed rather sparse.

For instance, the C-terminal domain of wheat LMW-GS and γ -gliadins presents homology with 2S storage proteins, but, while these proteins are often proteolytically processed after transport to the storage vacuole [7], the wheat proteins maintain an intact C-terminal domain. Proteolytic processing of vacuolar storage proteins is often due to the action of asparaginyl endopeptidases, a class of proteolytic enzymes (also termed vacuolar processing enzymes, VPEs, or legumains) that preferentially cleave after asparagine residues. The repetitive domain of prolamins is rich in proline and glutamine, but poor in asparagine, and thus lacks sites that can be cleaved by these enzymes. However, indication that enzymes belonging to this class may be involved in the maturation of prolamins is now emerging. Cleavage after an asparagine residue has been hypothesized in the processing of certain ω -gliadins [8]. Further evidence indicating that asparaginyl endopeptidases play a role in wheat storage protein maturation is provided by the comparison of the N-terminal sequences of different LMW-GS. Typical LMW-GS are in fact classified according to the N-terminal amino acid residue of the mature protein: serine in LMW-s, methionine in LMW-m and isoleucine in LMW-i types (rev. in [6]). LMW-s and LMW-m are, on the basis of complete nucleotide sequences, practically identical. The N-terminal amino acid sequences of these two protein types differ in that the first three N-terminal amino acids, MET- (or in one minor case, MEN-, [9]) in the mature sequence of LMW-m types are absent from the amino acid sequence of the s-type. However, the nucleotide sequence encoding these three residues is present in the s-type *lmw-gs* genes with the exception that the codon encoding threonine is replaced by that encoding asparagine, at least in a specific LMW-s type gene (*LMW-42K*) [EMBL:Y17845] [10] (Figure 1).

According to signal sequence cleavage site prediction algorithms, cleavage by signal peptidase would generate a QMET- N-terminal amino acid sequence for LMW-m (or QMEN- in case of LMW-s) type subunits (Figure 1). Removal of the N-terminal glutamine residue would therefore be required to generate m- or s-type LMW-GS, although it has been suggested that the signal cleavage might be degenerate, producing both cleavages [12]. On the other hand, the identification of the gene coding for a specific LMW-s type [EMBL:Y17845] [10] showed the presence of a codon encoding an asparagine in place of a threonine before the start of the mature sequence (MENS- instead of METS-). Accordingly, we postulated that the presence of an asparagine in the unprocessed protein in place of a threonine residue could result in preferential processing to produce an N-terminal end of the LMW-s type [10]. This would result from cleavage of the peptide MEN by an asparaginyl endoprotease, similar to a process that occurs in ω -gliadins [8].

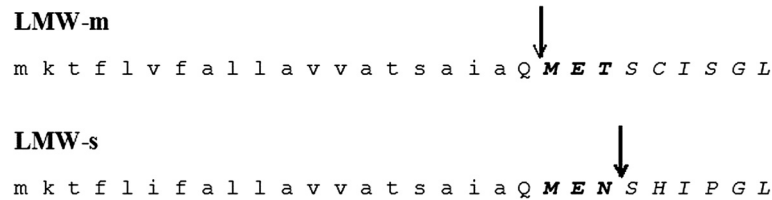


Figure 1 N-terminal amino acid sequences of the immature LMW-m and LMW-s proteins. The arrows indicate the start of the mature sequence. The signal peptide is indicated by lowercase letters. Although there is a slight amino acid variation in the two N-terminal sequences of LMW-GS types [11], here only the two sequences considered in the present work are reported.

In order to test our hypothesis, we expressed native and mutated versions of LMW-m and LMW-s genes in the endosperm tissue of transgenic durum wheats. Proteomic comparisons, MS/MS analyses and Edman degradation, demonstrated the involvement of asparagine in the formation of the LMW-s glutenin subunits.

Our results indicate that LMW-GS can be quantitatively processed most likely by vacuolar asparaginyl endoproteases, and suggest that LMW-GS accumulated in the vacuole are not sequestered into stable aggregates that would hinder the action of proteolytic enzymes. Rather, our results indicate that, whatever is the mechanism of glutenin polymer transport to the vacuole, the proteins remain available for proteolytic processing, and can be converted to the mature form by the removal of a short N-terminal sequence.

Results

Production of wheat transgenic plants

The descriptions of the genes used in this study are reported in the *Plant material and genetic transformation* paragraph (Methods section).

Durum wheat transgenic plants were produced to express native and modified versions of LMW-m (*B1133-WT*, *B1133-T23N*) and LMW-s (*42K-N23T*) type genes in wheat endosperm (Figure 1 and Additional file 1). Eleven, twenty-seven and thirty-two independent regenerated

plants (T_0) were obtained for the *B1133-WT*, *B1133-T23N* and *42K-N23T lmw-gs* genes, respectively, for a transformation efficiency of about 2%, that is within the typical efficiency obtained in stable wheat transformation.

Proteomic analysis

Since the proteomic patterns of the untransformed durum wheat cultivar Svevo and all the null segregant plants obtained from the progenies of the three types of transgenic plants were identical (data not shown), we eventually used only the null segregant plants as a control for the proteomic comparisons, as in the following description.

Proteomic comparison between proteins from the plants transformed with the 42K-N23T construct and proteins from the corresponding null segregant plants

The recipient cultivar Svevo possesses a LMW42K protein with pI 8.28 and molecular weight of 42,419, while the expected pI of the 42K-N23T protein is 7.85 with molecular weight of 37,675. Accordingly, we used IPG strips in the pH range 6-11 in order to maximize separation of the native and transgenic proteins. The proteomic comparison of the B subunits obtained from the null segregant control plants and the transgenic plants permitted us to identify two proteins exclusive to the latter (numbers 1 and 2 in Figure 2B).

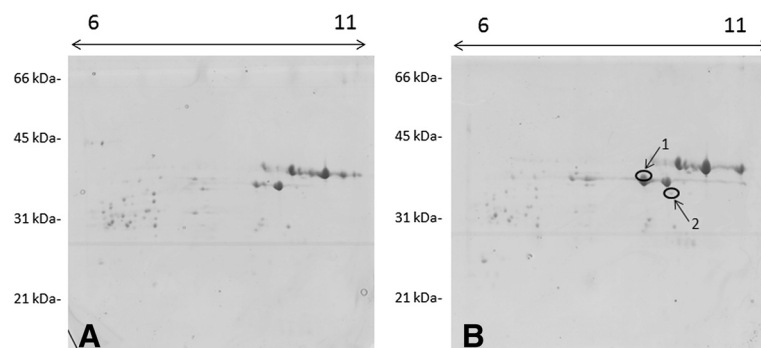


Figure 2 Comparison between the 2D electrophoretic patterns of B subunits of LMW-GS of the 42K-N23T transgenic lines. A) null segregant control line; B) genotype expressing the *42K-N23T* gene. Arrows point at the two proteins (spots nr. 1 and 2) that are absent in the control line. Molecular weight standards (in kDa) are reported on the left side, and the pI range is indicated at the top of the gels.

Analysis of mass spectrometry data from multiple 2D gel separations yielded evidence for the 42K-N23T construct that was identified from 211 total spectra, with 66 exclusive unique spectra, and 54 exclusive unique peptides, resulting in 298 out of 330 amino acids identified, corresponding to 90% sequence coverage, including N-terminal peptides (Figure 3 and Additional file 2). Edman degradation confirmed that the N-terminal amino acid sequence of the transgenic protein present in spot 1 was METSHIP (Table 1). The protein present in spot 2 did not yield sufficient material to provide interpretable N-terminal Edman sequence or tandem mass spectrometry (MS/MS) data, thus the results are available for spot 1 only.

Proteomic comparison between proteins from plants transformed with the B1133-WT construct and proteins from the corresponding null segregant plants

The expected molecular weight of the B1133-WT protein is 34,050 (including tags) and the expected pI is 6.52. Thus, we used IPG strips in the pH range 6-11 on a preparation of total glutenin subunits as a control since this protein is not normally present in the cultivar Svevo. The comparison of gels from transformed and null segregant plants enabled us to identify two main spots present in the transgenic plants, but absent in the control samples. The proteins showed the expected molecular weight, but were visible as what appears to be a charge train (Figure 4B). Immunoblotting performed by using an anti-FLAG antibody to detect proteins of the corresponding gel region gave positive signals matching the same charge train, indicating that the proteins in the spots corresponded to the transgenic proteins (Figure 4C), possibly differing from one another as consequence of glutamine deamidation, which is a common artifact in proteomic analyses [13]. Gel spots were collected and submitted for MS analyses. These results are summarized in Figure 5 and Additional files 3 and 4. The expected sequences are reported and the identified peptides are highlighted or underlined. With Scaffold set to display proteins with a peptide probability of 90%, there were, for example for Spot 3, 38 exclusive unique peptides, 47 exclusive unique spectra, and a total of 239 out of 298 amino acids matched for a coverage of 80%. Both the His- and FLAG tags, as well as the N-terminal peptide, METSCIF-, were identified,

confirming that the protein was processed as a LMW-m type (Figure 5 and Additional file 3). Similar results were obtained for spot 4 (Additional file 4). The presence of a phenylalanine residue in seventh position in Spots 3 and 4 instead of the original serine is due to a point mutation that occurred during the cloning procedure.

Proteomic comparison between proteins of plants transformed with the B1133-T23N construct and proteins from the corresponding null segregant plants

As in the case of the B1133-T23N protein, IPG strips in the pH range 3-10 were used to separate a preparation of the total glutenin subunits, since the expected pI was 6.81. Four spots were identified as a possible charge train at the expected molecular weight (33,629) both in the Coomassie stained gels (Figure 6B) and in the corresponding immunoblots performed with the anti-FLAG antibody (Figure 6C). All identified spots were submitted to MS/MS analysis, and corresponded to the same transgenic protein (Additional file 5). For this reason, these results are summarized in Figure 7, in which the expected sequence is reported and the identified peptides are underlined. In summary, 40 total spectra corresponding to 19 unique peptides identified this protein. In total, 159 amino acids were identified, out of 295 (54% coverage), including both the His- and FLAG tags, as well as the N-terminal peptide, that was, as expected, SCISGLE- (Figure 7). These results confirmed that the protein is processed as a LMW-s type. N-terminal amino acid sequencing on four protein spots corresponding to the transgenic proteins, confirmed that the proteins were of the LMW-s types (Table 1).

Evidence for an N-terminal glutamine

Examination of a number of MS/MS spectra revealed that, in addition to the peptide METSCIF-, there was another peptide that contained an N-terminal glutamine as evidenced by the presence of a singly charged, protonated peptide of mass 998.42. This peptide fragmented to yield an overlapping series of b and y ions that were interpreted by the sequencing software as the peptide sequence -ETSCIF with an unidentified N-terminal mass of 243.1. This would correspond to the sequence QMETSCIF- in which the N-terminal dipeptide (Q-M) had undergone

METSHIPSLEKPLQQPLPLQQILWYQQQPIQQQPQPFPQQPPCSQQQPPLSQQQP
PFSQQQPPFSQQQPVLPQQPPFSQQQQFPQQQPLLPQQPPFSQQQPPFSQQQQP
FSQQQQPLLPQQPPFSQHQQPVLPQQQIPSVQPSILQQLNPCKVFLQQQCSPVAMPQS
LARSQMLWQSSCHVMQQQCCRQLPQIPEQSRYDAIRAIIYSIVLQEQQHGQGLNQPQQQ
QPQQSVQGVSQPQQQKQLGQCSFQQPQQQLGQWPQQQVPQGTLLQPHQIAQLEVMT
SIALRTLPTMCSVNVPVYGTTTIVPFGVGTGVGAY

Figure 3 Complete deduced sequence of the transgenic 42K-N23T protein. The peptides identified by MS/MS in spot nr. 1 (no discernible sequence was identified in spot. Nr. 2) are underlined.

Table 1 Comparison between the expected and observed N-terminal amino acid sequences of the eight transgenic proteins

Genotype	Spot n.	Expected N-terminal sequence	Observed N-terminal sequence
42K-N23T	1	METSHIPLEKPLQQ-	METSHIP-
B1133-WT	3	METSHIFGLERPWQQ-	METSHIFGLERPWQQ-
	4	METSHIFGLERPWQQ-	METSHIFGLERPWQQ-
B1133-T23N	5	SCIPGLERPWQQQPL-	SCIPGLERP-
	6	SCIPGLERPWQQQPL-	SCIPGLERP-
	7	SCIPGLERPWQQQPL-	SCIPGLERP-
	8	SCIPGLERPWQQQPL-	SCIPGLERP-

In the "Genotype" column the three transgenic genotypes here analysed are reported. The "Expected N-terminal sequence" column reports the deduced N-terminal amino acid sequences on the basis of transgene nucleotide sequences, whereas the "Observed N-terminal sequence" corresponds to those obtained after Edman degradation, relatively to each of the eight protein spots indicated in the second column.

rearrangement to pyroglutamic acid (Figure 8). The METSCIF peptide was strongly predominant over the QMETSHIF form. A similar N-terminal peptide was identified by DuPont *et al.* [14] in a presumably *Glu-B3* coded LMW-GS from the bread wheat cultivar Butte 86 (both QMET- and QMEN- type sequences were observed) and also by Mamone *et al.* [15] for a LMW-GS, although in this latter case the sequence was QMDT-. The presence of glutamine as first residue corresponds to the expected N-terminal sequence of LMW-GS, according to signal sequence cleavage site prediction. It is of interest that Edman sequencing of the MET-type proteins does not indicate Q as the first amino acid, but rather M, suggesting that the QMETSHIP form cyclizes rapidly and is thus unavailable to Edman sequencing, whereas the amino acid M is readily recognized for MET-type proteins suggesting that cyclization is partial, that the Q is partly removed by some other enzyme (an aminopeptidase?) or the signal cleavage itself is degenerate, producing cleavages before and after the Q [12]. With or without the Q, they both correspond to the LMW-m type. In any case, the reporting of QMET- (found also in this work), QMEN- and

QMDT- N terminal peptides indicates that this type of processing may occur in LMW-GS.

Finally, Ikeda *et al.* [9] identified LMW-GS with the N-terminal sequence MENSHPGLE, likely as the result of a partial escape from the action of asparaginyl endopeptidase. We did not find any MENSHP in the transgenic polypeptides, although we cannot exclude its presence as a minor component.

Discussion

Wheat gluten proteins are typical secretory proteins in that their synthesis, folding, maturation and deposition take place within the endomembrane system of the plant cell. Among them, LMW-GS are the most heterogeneous group, being represented by multiple proteins, including those with a characteristic LMW-GS sequence and others with a gliadin-like sequence. In regard to the former, the most common proteins are those typically known as LMW-m and LMW-s types, according to the first amino acid of the mature sequence (methionine or serine, respectively). The LMW-m and LMW-s protein types are practically identical, on the basis of the complete

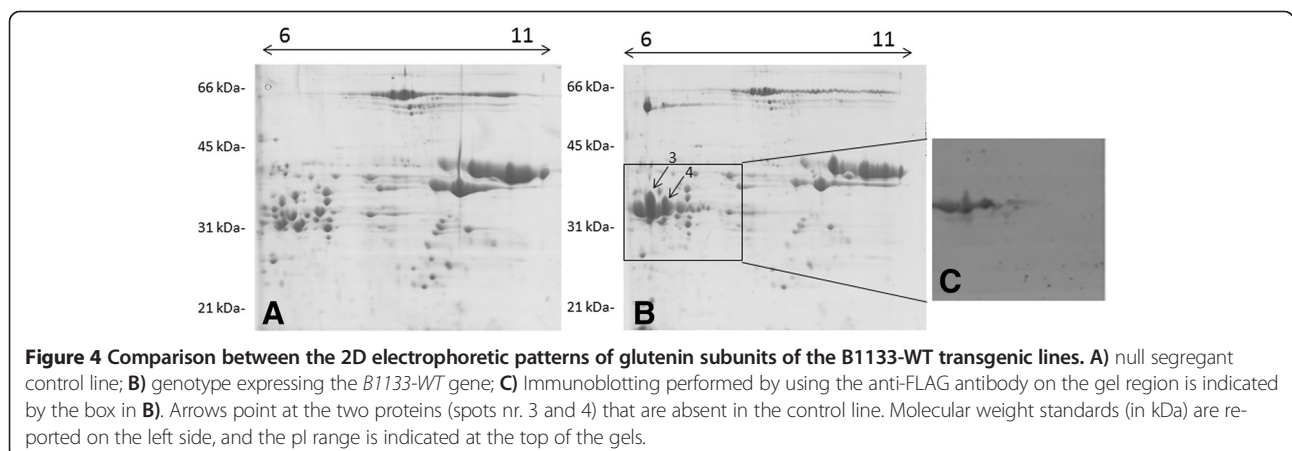


Figure 4 Comparison between the 2D electrophoretic patterns of glutenin subunits of the B1133-WT transgenic lines. A) null segregant control line; **B)** genotype expressing the *B1133-WT* gene; **C)** Immunoblotting performed by using the anti-FLAG antibody on the gel region is indicated by the box in **B)**. Arrows point at the two proteins (spots nr. 3 and 4) that are absent in the control line. Molecular weight standards (in kDa) are reported on the left side, and the pI range is indicated at the top of the gels.

METSCIFGLERPWQQQLPLPPQQSSFSQQPPFSQQQQQPLPQQPSFSQQQPPFSQQQPILS
QQPPFSQQQQPVLPQQSPFSQQQLVLPPQQQQLVQQIPIVQPSVLQLNPCKVFL
QQQCSPVAMPQRLARSQMWQSSCHVMQQCCQLQIPEQSRYEAIRAIIYSIILQEQ
QQGFVQPQQQPQSGVSQSQQLGCSFQPQQLGQPQQQQVLQGTF
LQPHQIAHLEAVTSIALRTLPTMCSVNVPLYSATTSVPFGVGTGVGAYHHHHHHDYKDD
DDK

Figure 5 Complete deduced sequence of the transgenic B1133-WT protein. The peptides identified by MS/MS in spots nr. 3 and 4 are underlined.

nucleotide sequence, since the main difference resides in the absence or the presence of three additional N-terminal amino acids (MET-, or MEN- in some minor cases [9]) in the mature sequence of LMW-m type, that are instead absent in LMW-s types. However, the corresponding nucleotide triplets are present in the sequences encoding both -s and -m type LMW-GS precursors, except that in one case the third amino acid corresponds to threonine and in the other to asparagine [10]. Because of the presence of either a threonine or an asparagine residue in position 23 of the LMW-m and LMW-s precursors (which corresponds to the third position of the mature LMW-m types), we have hypothesized that, in case of LMW-s types, a preferential (and perhaps secondary) processing at the N-terminal end could occur, that would generate the cleavage of the peptide MEN, most likely by an asparaginyl endoprotease. These proteolytic enzymes have been termed also “legumains” or “vacuolar processing enzymes” [16] and consist of a large family of plant and animal Asn-specific cysteine proteinases. In plants, they occur in storage vacuoles or cell wall of seeds and vegetative organs. In seeds, they play a role in both protein maturation and degradation. They in fact are involved in post-translational processing of protein precursors by cleaving asparagine residues in P1 position of peptide bonds [17]. Asparaginyl endoprotease functions depend on the conformational state of the substrate protein. They are involved in maturation processing when they have access to a normal processing site of the precursor protein, but contribute to the

degradation processing by performing an extensive proteolysis, when the protein is misfolded. Thus, seed legumains have a role in protein maturation and also a “structural proof-reading function” [18]. Most of what we know about the characteristics and the biological function of these enzymes is derived from studies on dicot plants. Barley NP1 has been reported to be localized to the cell wall of nucellar cell types [19], while rice REP-2 has been implicated, together with REP-1, in storage protein degradation [20]. These enzymes thus appear to play a role different from storage protein processing. Recently, *OsVPE1*, a rice homolog of the Arabidopsis β VPE gene, has been shown to be responsible for the cleavage of rice storage globulins in the protein storage vacuoles [21].

In order to determine if the hypothesized maturation processing occurs, we produced durum wheat transgenic plants transformed with three gene constructs, two corresponding to the two *lmw-gs* gene types, namely those with triplets giving rise to threonine or asparagine residues in position 23 of the coding sequence, and a third one in which the codon for the threonine residue was mutated to give rise to an asparagine residue. The transgenic proteins were characterized in T₄ kernels by means of a proteomic comparison. MS/MS analyses coupled with Edman sequencing of the identified proteins clearly demonstrated that the processing is dependent on the presence of a threonine or asparagine residue in position 23 of the coding sequences: if an asparagine residue is present, proteins are processed as LMW-s types; conversely, when a

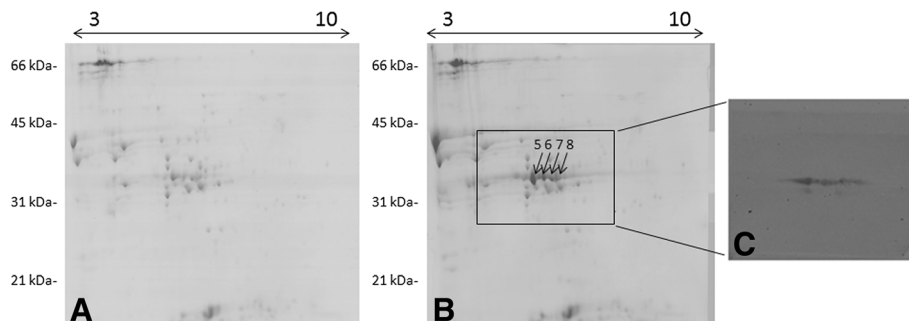


Figure 6 Comparison between the 2D electrophoretic patterns of glutenin subunits of the B1133-T23N transgenic lines. **A)** null segregant control line; **B)** genotype expressing the *B1133-T23N* gene; **C)** Immunoblotting performed by using the anti-FLAG antibody on the gel region indicated by the box in **B)**. Arrows point at the four proteins (spots nr. 5-8) that are absent in the control line. Molecular weight standards (in kDa) are reported on the left side, and the pI range is indicated at the top of the gels.

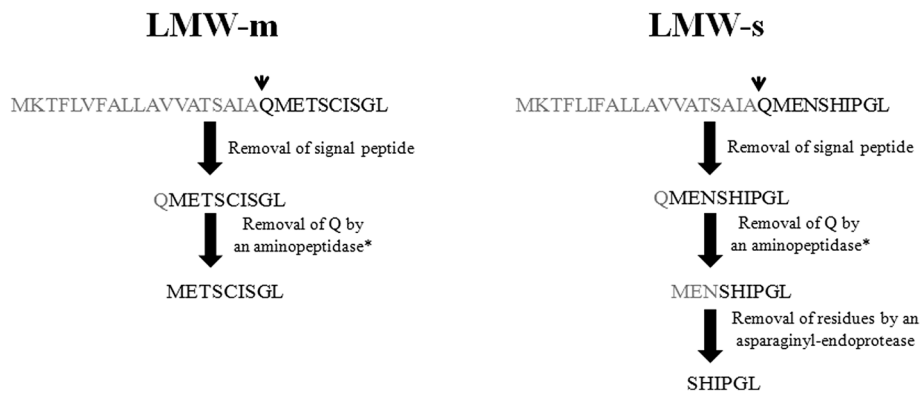


Figure 9 Proposed model for the differential maturation process of LMW-m and LMW-s. Arrow-head indicates signal sequence cleavage site as determined by SignalP (<http://www.cbs.dtu.dk/services/SignalP/>). Asterisk indicates that additional hypotheses on the lack/presence of Q can be proposed (as reported in the sub-paragraph "Evidence for an N-terminal glutamine"). Peptides and amino acids removed during the maturation processes are in grey.

they remain available for proteolytic processing, necessary for finalizing the maturation process.

Methods

Plant material and genetic transformation

The LMW-m and LMW-s type genes and their mutated versions used for wheat transformation were those reported by Ferrante *et al.* [25]. One gene, named *B1133-WT*, corresponds to a native gene that was presumed to code for a LMW-m protein, although it is reported in GenBank as a γ -gliadin [GenBank:M11077] [26]; another one, named *B1133-T23N*, corresponds to the same gene mutated in position 23 to replace a threonine with an arginine. The third gene, named *42K-N23T*, derives from a LMW-s type gene [EMBL:HG529977X], mutated in position 23 to replace an arginine with a threonine. This gene was isolated from genomic DNA of the bread wheat cultivar Yecora Rojo with primers reported in [10]. These three genes were cloned separately into the *SalI-XbaI* restriction sites of pLRPT vector under control of the endosperm-specific HMW-GS Dx5 promoter [27]. The cloning of each gene was achieved by PCR, amplifying the *B1133-WT* or *B1133-T23N* with primers SalHB1133F 5'-acagtcgacatgaagacctctctcttt-3' and XbaHisFlagR 5'-tctagatcactgtcatcgtcatcctgttagtcgtgatggtgatggtgatggt-3', containing the sequences for the His- and FLAG-tags (see also below), whereas the *42K-N23T* gene was amplified with the primers LMW42KSalF 5'-acagtcgacatgaagacctctctcttt-3' and LMW42KXbaIR 5'-tctagatcagtaggca ccaactccggt-3'. Since in the past we experienced multiple problems in *LMW-42K* gene [EMBL:Y17845] isolation, mostly due to rearrangements occurring during the cloning procedure because of the particular organization of the repeated domain [10], we deliberately decided not to add tags to this gene construct, which, although helping in protein identification

and purification, might contribute to cloning difficulties and/or rearrangements.

PCR reactions were prepared in 50 μ l containing 5 μ l of 10X FastStart High Fidelity Reaction Buffer (Roche Diagnostics, Monza, Italy), 100 ng of genomic DNA, dNTP Mix 10 mM, 200 ng of each primer, 2.5 units of Fast Start High Fidelity DNA polymerase (Roche Diagnostics, Monza, Italy). The PCR program was: 95°C 2 min, 1 cycle, 95°C 1 min, 62°C 2 min, 72°C 2 min, 30 cycles; 72°C, 5 min. The amplification products were recovered from a 1.2% agarose gel, digested with *SalI-XbaI* and ligated into pLRPT. Constructs were verified by nucleotide sequencing and the *B1133-WT* showed a single substitution which caused the replacement of a serine with a phenylalanine in seventh position of the deduced mature sequence. Despite this substitution, this gene was used because we reasoned it would not interfere with our hypothesis.

The plasmid pAHC20 [28] carrying the *bar* gene that confers resistance to the bialaphos herbicide, was co-bombarded in immature embryos of the durum wheat cultivar Svevo with each of the above LMW-GS genes in a 1:3 molar ratio by following the procedure reported by Volpi *et al.* [29].

Construct schemes are reported in Additional file 1.

Genomic DNA extraction and PCR analysis of transformed wheat plants

Genomic DNA was extracted from 0.2 g of green tissue as reported in [30].

Transformed T₀ plants were identified by PCR by using primers specific for the HMW-GS DX5 promoter and the terminator (PRDX5F 5'-catgcaggctactctccac, PRDX5R 5'-cggtaggactcatggaattg [31]). The PCR conditions were those reported in the previous paragraph, except for a different extension time that was 1 min and 30 seconds. Positive plants were multiplied up to T₄ generation, in order to

obtain as many homozygous plants as possible to submit to proteomic analyses. Negative plants, corresponding to the null segregant plants, namely those transgenic plants that have lost the transgene by segregation, were also selected, multiplied up to the T₄ generation, and used as control.

Proteomic analyses

Plants, either wild type and transgenic lines, included the null segregant genotypes, were grown together in a growth chamber. T₄ generation plants, previously screened by PCR, were used for proteome analyses, by pooling half seeds of PCR positive plants. Since we were interested only in the presence/absence of the transgenic proteins, we did not use formal replicates, but extracted proteins from each of at least four different positive lines obtained from transformation with each of the three transgenes, and compared to as many biological replicates of the null segregant.

Extraction of glutenin subunits

Seeds from the durum wheat cultivar Svevo as well from its transgenic lines (included null segregants) were crushed and 50 mg of flour were washed three times with 1 mL of 50% (v/v) propanol in order to remove the soluble protein fraction [32]. In case of extraction of total glutenin subunits, the pellet was eventually extracted with a solution (1 mg: 10 µl) of 50% propanol containing 50 mM Tris-HCl pH 8.8, 1 mM EDTA, 10 mM iodoacetamide or 1.4% of 4-vinylpyridine, 1% (w/v) DTT for 1 h at 70°C. After centrifugation at 13,000 rpm for 15 min, four volumes of cold acetone were added to the recovered supernatant and kept overnight at -20°C to precipitate glutenin subunits. After centrifugation at 13,000 rpm for 15 min, the precipitated proteins were collected and dried in a Savant centrifuge.

In the case of the 42K-N23T protein, since tags were not added, in order to facilitate the identification of differences between null segregant and transformed genotypes, we used as a control a protein fraction enriched in B-type low-molecular weight glutenin subunits (similar in structure to the 42K-N23T protein) that was obtained according to [33].

2D electrophoretic analysis (IEF vs SDS-PAGE) of LMW-GS

Quantification of proteins prior to isoelectric focusing (IEF) was performed with the 2-D quant Kit Assay (GE Healthcare).

Total glutenin subunits or B subunits of LMW-GS were suspended in 250 µl of a solution composed of 7 M urea, 2 M thiourea, 4% (w/v) CHAPS, 1.2% (v/v) Destreak Reagent, 0.5% (v/v) IPG buffer pH 3-10 and 6-11 for at least 2 hours. IEF was performed with the *IPGphor™ Isoelectric Focusing System* (Amersham Pharmacia Biotech) and was carried out on immobilized pH gradient (IPG) strips (18 cm, 1 mm) pH 3-10 (for plants transformed with *B1133-WT* and *B1133-T23N* genes)

and pH 6-11 (for plants transformed with the *42 K-N23T* gene). The strips were hydrated with samples overnight (12.30 h) at room temperature. IEF was performed at 90,000 volt-hours at 20°C. After focusing, the strips were equilibrated for 30 min at room temperature in a solution of 6 M urea, 2% (w/v) SDS, 30% (v/v) glycerol, 50 mM Tris-HCl, pH 6.8, and 1% (w/v) DTT. Strips were then loaded on the top of a 1 mm thick by 18 cm SDS polyacrylamide gel, T 11%, C 1.28%, by using the *Protean Plus Dodeca cell* (Bio-Rad). Electrophoretic separation was carried out at 40 mA/gel, with cooling at 10°C. Gels were stained with Coomassie Blue G250 [34] and destained in water before image acquisition.

The gel analyses were performed using the software SameSpots Progenesis (vers. 4.5.4293.47197, Nonlinear Dynamics, UK). This software includes statistical analyses such as ANOVA ($p \leq 0.05$), and determination of False Discovery Rate (FDR, $q \leq 0.05$).

Western blotting for amino acid sequencing

A 9 cm × 7 cm gel piece, corresponding to the region of interest, namely that including proteins in the molecular weight and pI ranges corresponding to the transgenic proteins, was cut out of the unstained 2D gel and electroblotted on Sequi-blot PVDF (polyvinylidene difluoride) membranes (Bio-Rad, Hercules, CA), previously wetted in methanol and rinsed with deionized water for 5 minutes before soaking in electroblot buffer (10 mM CAPS [3-cyclohexylamino-1-propanesulfonic acid], pH 11). Filter paper (Whatman 3 MM) was also soaked in electroblot buffer before use. Gel pieces were soaked in electroblot buffer for 5-10 minutes. Western blottings were performed using a *Mini Trans Blot Cell* module (Bio-Rad). Transfer was performed for 1 hour at 4°C, at a constant voltage of 100 V. The transfer stack was then dismantled and the membrane was rinsed with distilled water for 5 minutes before staining with Coomassie blue (0,025% (w/v) Coomassie R-250, 40% (v/v) methanol) for 5 minutes. The membranes were then de-stained for 5 minutes in 50% (v/v) methanol, briefly rinsed in distilled water and allowed to air dry at room temperature. Spots of interest were excised using a clean razor blade, and amino acid sequencing was performed essentially according to the procedure reported by Tao and Kasarda [35], but using an Applied Biosystems Procise 492 sequencer.

Immunoblotting

For immunoblotting experiments, the gel pieces containing the region of interest were incubated in transfer buffer (25 mM Tris- HCl, pH 8,0, 192 mM glycine and 0,04% SDS) for 15 minutes. The blotting was performed in the *Mini Trans Blot Cell* apparatus (Bio-Rad) using a PVDF membrane (Bio-Rad) according to the manufacturer's

protocol. After transferring, the membrane was saturated in 100 ml of Blocking solution (10 mM Tris-HCl pH 8, 150 mM NaCl, 0.1% Tween20 and 5% non-fat dry milk) at room temperature, on an orbital shaker, for 2 h. The membrane was then washed twice in washing buffer (10 mM Tris-HCl, pH 8, 150 mM NaCl and 0.2% Tween20) and incubated overnight with an anti-Flag-tag polyclonal antibody (Sigma-Aldrich). After removal from the incubation buffer, the membrane was washed extensively and incubated with a horseradish peroxidase-conjugated goat anti-rabbit secondary antibody (Merck Millipore) at room temperature, for 1 hour. The antigen-antibody complex was detected using the “Western blotting Luminol reagent” kit (Santa Cruz Biotechnology, Inc.) following the manufacturer’s procedure.

Mass spectrometry analysis

Coomassie Brilliant Blue stained bands were cut from the polyacrylamide gels and stored in 1.5 mL Eppendorf tubes. Immediately prior to enzymatic digestion, the gel pieces (1-3) were placed into the wells of a 96-well reaction plate that was positioned in an automated xyz robot (DigestPro, Intavis, Langenfeld, Germany) that automatically destained, reduced, alkylated and digested the proteins in the gel plugs with either chymotrypsin, thermolysin or trypsin. Twenty µg of the selected enzyme was used for each 96-well sample plate and digestion was performed at 35°C. At the end of the digestion period, the instrument eluted the samples of enzymatically cleaved peptides into a 96-well receiving plate that was then inserted into the autosampler interfaced with the QSTAR pulsar *i* hybrid quadrupole-TOF mass spectrometer (Applied Biosystems/MDX Sciex, Toronto, Canada) configured with an electrospray ionization (ESI) source (Protana, Odessa Denmark). Mass spectrometric analysis was performed as previously described [36]. When sufficient material was available the samples were reanalyzed using the data obtained from the first mass spec analysis to form an exclusion list to allow previously unidentified spectra to be analyzed.

The resulting data were searched, using Mascot (www.matrixscience.com/) and X!Tandem (<http://www.thegpm.org/>) against a database containing 11,589 wheat protein sequences from NCBI *T. aestivum* plus a list of common laboratory contaminants (<http://www.thegpm.org/>) as well as expected sequences from the mutant and wild type expressed proteins. The results of the searches were combined, analyzed and visualized using Scaffold version 4.075 (<http://www.proteomesoftware.com>).

Additional files

Additional file 1: Plasmids used for biolistic transformation of durum wheat cv. Svevo. pLRPT vector, containing the *Dx5* promoter and 42K-N23T (A), B1133-WT (B) and B1133-T23N (C) genes.

Additional file 2: Data relative to protein spot 1 showing identified sequence and a table showing identified peptides and associated ion statistics. Individual Excel spreadsheets contain the expected sequences of the protein from protein spot 1: the peptides identified by mass spectrometry are highlighted in yellow, modifications are shown in green: C, carbamidomethyl cysteine; M, oxidation; Q, deamidation. Note that not all cysteine residues are colored green although all have been converted to carboxymethyl amino cysteine. Modification of cysteine was defined as both a fixed and a variable modification in the database searching software. The result is that not all cysteine residues were color-coded by the analysis software.

Additional file 3: Data relative to protein spot 3 showing identified sequence and a table showing identified peptides and associated ion statistics. Individual Excel spreadsheets contain the expected sequences of the protein from protein spot 3: the peptides identified by mass spectrometry are highlighted in yellow, modifications are shown in green: C, carbamidomethyl cysteine; M, oxidation; Q, deamidation. Note that not all cysteine residues are colored green although all have been converted to carboxymethyl amino cysteine. Modification of cysteine was defined as both a fixed and a variable modification in the database searching software. The result is that not all cysteine residues were color-coded by the analysis software.

Additional file 4: Data relative to protein spot 4 showing identified sequence and a table showing identified peptides and associated ion statistics. Individual Excel spreadsheets contain the expected sequences of the protein from protein spot 4: the peptides identified by mass spectrometry are highlighted in yellow, modifications are shown in green: C, carbamidomethyl cysteine; M, oxidation; Q, deamidation. Note that not all cysteine residues are colored green although all have been converted to carboxymethyl amino cysteine. Modification of cysteine was defined as both a fixed and a variable modification in the database searching software. The result is that not all cysteine residues were color-coded by the analysis software.

Additional file 5: Data relative to protein spot 5 showing identified sequence and a table showing identified peptides and associated ion statistics. Individual Excel spreadsheets contain the expected sequences of the protein from protein spot 5: the peptides identified by mass spectrometry are highlighted in yellow, modifications are shown in green: C, carbamidomethyl cysteine; M, oxidation; Q, deamidation. Note that not all cysteine residues are colored green although all have been converted to carboxymethyl amino cysteine. Modification of cysteine was defined as both a fixed and a variable modification in the database searching software. The result is that not all cysteine residues were color-coded by the analysis software.

Competing interests

The authors declare that they have no competing interests.

Authors’ contributions

EE carried out the molecular and electrophoretic analyses and made the first draft of the paper; RD and FS contributed to constructs preparation and supervised EE in the molecular analyses; MJ carried out the biolistic transformation that was supervised by RD; AC, DL and DDK critically discussed the results; WHV carried out the MS and Edman analyses and drafted the relevant methods and results; SM conceived the work together with AC, coordinated it and wrote the final draft. All authors read, edited and approved the final manuscript.

Acknowledgements

The authors wish to acknowledge Shireen Javandel for her assistance in data analysis, Edman sequencing and MS/MS. This work was funded by the Italian Ministry of Education, University, and Research (MIUR), project “Laboratorio di GENomica per caratteri di importanza AGRonomica in frumento duro: identificazione di geni utili, analisi funzionale e selezione assistita con marcatori molecolari per lo sviluppo della filiera sementiera nazionale (AGRO-GEN)”. EE PhD fellowship was in part supported by the project Mario Negri Sud “Proteomics/Genomics to analyse flour functionality differences between wheat varieties”. WHV and DDK used funds from USDA Agricultural Research Service CRIS Project 5325-43000-026-00 D.

USDA is an equal opportunity provider and employer. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture.

Author details

¹DAFNE, Tuscia University, Viterbo, Italy. ²IBBA, CNR, Milan, Italy. ³USDA, ARS, WRRRC, Albany, CA, USA. ⁴Present address: Institute of Plant Genetics (IGV), CNR, Via Amendola 165/A, 70126 Bari, Italy.

Received: 13 January 2014 Accepted: 7 March 2014

Published: 14 March 2014

References

- Bailey CH: A translation of Beccari's lecture 'Concerning Grain' 1728. *Cereal Chem* 1941, **18**:555–561.
- Kreis M, Shewry PR, Forde BG, Forde J, Mifflin BJ: Structure and evolution of seed storage proteins and their genes with particular references to those of wheat, barley and rye. *J Mol Biol* 1985, **183**:499–502.
- Wrigley CW: Giant proteins with flour power. *Nature* 1996, **381**:738–739.
- Dachkevitch T, Autran JC: Prediction of baking quality of bread wheats in breeding programs by size-exclusion high-performance liquid chromatography. *Cereal Chem* 1989, **66**:448–456.
- Gupta RB, Popineau Y, Lefebvre J, Cornec M, Lawrence JG, MacRitchie F: Biochemical basis of flour properties in bread wheats. II. Changes in polymeric protein formation and dough/gluten properties associated with the loss of low Mr or high Mr glutenin subunits. *J Cereal Sci* 1995, **21**:103–116.
- D'Ovidio R, Masci S: The low-molecular-weight glutenin subunits of wheat gluten. *J Cereal Sci* 2004, **39**:321–339.
- Shimada T, Yamada K, Kataoka M, Tomoo S, Kenji Y, Miyuki K, Satoru N, Yasuko K, Miwa K, Satoshi T, Tomohiko K, Kazuo S, Motoaki S, Masatomo K, Maki K, Miki N, Ikuko HN: Vacuolar processing enzymes are essential for proper processing of seed storage proteins in *Arabidopsis thaliana*. *J Biol Chem* 2003, **278**:32292–32299.
- DuPont FM, Vensel WH, Encarnacao T, Chan R, Kasarda DD: Similarities of omega gliadins from *Triticum urartu* to those encoded on chromosome 1A of hexaploid wheat and evidence for their post-translational processing. *Theor Appl Genet* 2004, **108**:1299–1308.
- Ikedda TM, Araki E, Fujita Y, Yano H: Characterization of low-molecular-weight glutenin subunit genes and their protein products in common wheats. *Theor Appl Genet* 2006, **112**:327–334.
- Masci S, D'Ovidio R, Lafiandra D, Kasarda DD: Characterization of a low molecular-weight glutenin subunit gene from bread wheat and the corresponding protein that represents a major subunit of the glutenin polymer. *Plant Physiol* 1998, **118**:1147–1158.
- Lew EJL, Kuzmicky DD, Kasarda DD: Characterization of low molecular weight glutenin subunits by Reversed-Phase high-performance liquid chromatography, sodium dodecyl sulfate-polyacrylamide gel electrophoresis, and N-terminal amino acid sequencing. *Cereal Chem* 1992, **69**:508–515.
- Kasarda DD, Adalsteins E, Lew EJL, Lazo GR, Altenbach SB: Farinin: Characterization of a novel wheat endosperm protein belonging to the prolamin superfamily. *J Agric Food Chem* 2013, **61**:2407–2417.
- Sarioglu H, Lottspeich F, Walk T, Jung G, Eckerskorn C: Deamidation as a widespread phenomenon in two-dimensional polyacrylamide gel electrophoresis of human blood plasma proteins. *Electrophoresis* 2000, **21**:2209–2218.
- DuPont FM, Vensel WH, Tanaka CK, Hurkman WJ, Altenbach SB: Deciphering the complexities of the wheat flour proteome using quantitative two-dimensional electrophoresis, three proteases and tandem mass spectrometry. *Proteome Sci* 2011, **9**:10.
- 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**:1709–1723.
- Hara-Nishimura I, Inoue K, Nishimura M: A unique vacuolar processing enzyme responsible for conversion of several proprotein precursors into the mature forms. *FEBS Lett* 1991, **294**:89–93.
- Müntz K, Shutov AD: Legumains and their functions in plants. *Trends Plant Sci* 2002, **7**:340–344.
- Müntz K, Blattner FR, Shutov AD: Legumains - a family of asparagine-specific cysteine endopeptidases involved in propeptide processing and protein breakdown in plants. *J Plant Physiol* 2002, **159**:1281–1293.
- Linnestad C, Doan DN, Brown RC, Lemmon BE, Meyer DJ, Jung R, Olsen OA: Nucellain, a barley homolog of the dicot vacuolar-processing protease, is localized in nucellar cell walls. *Plant Physiol* 1998, **118**:1169–1180.
- Kato H, Sutoh K, Minamikawa T: Identification, cDNA cloning and possible roles of seed-specific rice asparaginyl endopeptidase, REP-2. *Planta* 2003, **217**:676–685.
- Wang YH, Zhu SS, Liu SJ, Jiang L, Chen LM, Ren YL, Han XH, Liu F, Ji SL, Liu X, Wan JM: The vacuolar processing enzyme OsVPE1 is required for efficient glutelin processing in rice. *Plant J* 2009, **58**:606–617.
- Lan Q, Feng B, Xu Z, Zhao G, Wang T: Molecular cloning and characterization of five novel low molecular weight glutenin subunit genes from Tibetan wheat landraces *Triticum aestivum* L. *Genet Resour Crop Ev* 2013, **60**:799–806.
- Naghavi MR, Ahmadi S, Shanejat-Boushehri A-A, Komaei G, Struik PC: Characterization of low-molecular-weight-glutenin subunit genes from the D-genome of *Triticum aestivum*, *Aegilops crassa*, *Ae. cylindrica* and *Ae. tauschii*. *Biochem Syst Ecol* 2013, **50**:23–29.
- Ueda Y, Satoh-Cruz M, Matsusaka H, Takemoto-Kuno Y, Fukuda M, Okita TW, Ogawa M, Satoh H, Kumamaru T: Gene-gene interactions between mutants that accumulate abnormally high amounts of proglutelin in rice seed. *Breeding Sci* 2010, **60**:568–574.
- Ferrante P, D'Ovidio R, Lafiandra D, Ceriotti A, Vensel WH, Kasarda DD, Masci S: Hypothetical differential processing of low molecular weight glutenin subunits Met- and Ser-types at their N-terminal end. In *Gluten Proteins 2006*. Edited by Lookhart GL, Ng PKW. St. Paul, Minnesota: AACCI International; 2007:33–37.
- Okita TW, Cheesbrough V, Reeves CD: Evolution and heterogeneity of the alpha/beta-type and gamma-type gliadin DNA sequences. *J Biol Chem* 1985, **260**:8203–8213.
- Tosi P, D'Ovidio R, Napier J, Bekes F, Shewry PR: Expression of epitope-tagged LMW glutenin subunits in the starchy endosperm of transgenic wheat and their incorporation into glutenin polymers. *Theor Appl Genet* 2004, **108**:468–476.
- Christensen AH, Quail PH: Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. *Transgenic Res* 1996, **5**:213–218.
- Volpi C, Janni M, Lionetti V, Bellincampi D, Favaron F, D'Ovidio R: The ectopic expression of a pectin methyl esterase inhibitor increases pectin methyl esterification and limits fungal diseases in wheat. *Mol Plant Microbe Interact* 2011, **24**:1012–1019.
- Tai T, Tanksley S: A rapid and inexpensive method for isolation of total DNA from dehydrated plant tissue. *Plant Mol Biol Rep* 1991, **8**:297–303.
- Sestili F, Janni M, Doherty A, Botticella E, D'Ovidio R, Masci S, Jones HD, Lafiandra D: Increasing the amylose content of durum wheat through silencing of the SBEIIa genes. *BMC Plant Biol* 2010, **10**:144.
- Singh NK, Shepherd KW, Cornish GB: A simplified SDS-PAGE procedure for separating LMW subunits of glutenin. *J Cereal Sci* 1991, **14**:203–208.
- Masci S, Egorov TA, Ronchi C, Kuzmicky DD, Kasarda DD, Lafiandra D: Evidence for the presence of only one cysteine residue in the D-type low molecular weight subunits of wheat glutenin. *J Cereal Sci* 1999, **29**:17–25.
- Neuhoff V, Arold N, Taube D, Ehrhardt W: Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with a clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. *Electrophoresis* 1988, **9**:255–262.
- Tao HP, Kasarda DD: Two-dimensional gel mapping and N-terminal sequencing of LMW-glutenin subunits. *J Exp Bot* 1989, **40**:1015–1020.
- Vensel WH, DuPont FM, Sloane S, Altenbach S: Effect of cleavage enzyme, search algorithm and decoy database on mass spectrometric identification of wheat gluten proteins. *Phytochemistry* 2011, **72**:154–161.

doi:10.1186/1471-2229-14-64

Cite this article as: Egidi et al.: An asparagine residue at the N-terminus affects the maturation process of low molecular weight glutenin subunits of wheat endosperm. *BMC Plant Biology* 2014 **14**:64.