Mol Breeding (2018) 38: 68
https://doi.org/10.1007/s11032-018-0827-9

Development of haplotype-specific molecular markers for the low-molecular-weight glutenin subunits

Maria I. Ibba · Alecia M. Kiszonas · Craig F. Morris

Received: 14 November 2017 / Accepted: 19 April 2018 / Published online: 12 May 2018

Abstract Low-molecular-weight glutenin subunits (LMW-GSs) are one of the major components of gluten, and their allelic variation has been widely associated with different wheat end-use quality parameters. These proteins are encoded by multigene families located at the orthologous *Glu-3* loci (*Glu-A3*, *Glu-B3*, and *Glu-D3*); the genes at each locus are divided by large intergenic and highly recombinogenic regions. Among the methods used for the LMW-GS allele identification, polymerase chain reaction (PCR)-based molecular markers have the advantages of being simple, accurate, and independent from the plant stage of development. However, the available LMW-GS molecular markers are either incapable of capturing the complexity of the LMW-GS gene family or difficult to interpret. In the present study, we report the development of a set of PCR-based molecular markers specific for the LMW-GS haplotypes present at each *Glu-3* locus. Based on the LMW-GS gene sequences available in GenBank, single nucleotide polymorphisms (SNPs) specific for each *Glu-3* haplotype were identified and the relevant PCR primers were designed. In total, we developed three molecular markers for the *Glu-A3* and *Glu-B3* loci, respectively, and five molecular markers for the *Glu-D3* locus. The markers were tested on 44 bread wheat varieties previously characterized for their LMW-GS genic profile and found to be equally or more efficient than previously developed LMW-GS PCR-based markers. This set of markers allows an easier and less ambiguous identification of specific LMW-GS haplotypes associated with gluten strength and can facilitate marker-assisted breeding for wheat quality.

Keywords Low-molecular-weight glutenin subunits · Gluten · *Glu-3* · Alleles · Molecular markers · PCR

Mention of trademark or proprietary products does not constitute a guarantee or warranty of a product by the US Department of Agriculture and does not imply its approval to the exclusion of other products that may also be suitable. This article is in the public domain and not copyrightable.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11032-018-0827-9) contains supplementary material, which is available to authorized users.

M. I. Ibba
Department of Crop and Soil Sciences, Washington State University, Pullman, WA, USA

A. M. Kiszonas · C. F. Morris (✉)
USDA-ARS Western Wheat Quality Laboratory, Pullman, WA, USA
e-mail: craig.morris@ars.usda.gov

Introduction

Gluten is a macro polymer that confers to wheat dough its unique viscoelastic properties. The main components of gluten are gliadins and glutenins, two classes of wheat seed storage proteins that, depending on their quantity and quality, greatly influence gluten technological properties (Shewry et al. 2002). Gliadins are monomeric proteins divided into α/β, γ, and ω gliadins depending on their mobility in acid-polyacrylamide gel electrophoresis (Barak et al. 2015). Glutenins are

polymeric proteins crosslinked with each other through disulfide bonds. Under reducing conditions, glutenins can be separated based on their molecular weight into high-molecular-weight glutenin subunits (HMW-GSs, 70–90 KDa) and low-molecular-weight glutenin subunits (LMW-GSs, 20–45 KDa) (D'Ovidio and Masci 2004).

Compared to gliadins and LMW-GSs, the HMW-GSs are less abundant, representing only 10% of the total wheat seed storage proteins (Wang et al. 2010). However, their fundamental role in wheat end-use quality has been extensively reported in the literature (He et al. 2005; Jin et al. 2013; Liu et al. 2005; Lukow et al. 1989; Payne et al. 1987) and the HMW-GS allelic profile is often used as a predictor of gluten quality. The LMW-GSs, in contrast, represent around one third of the total wheat seed storage proteins and around 60% of the glutenins (Bietz and Wall 1973). Their importance in different wheat end-use quality parameters has been well-established (Bonafede et al. 2015; Liu et al. 2005; Nieto-Taladriz et al. 1994; Payne 1987; Zhang et al. 2012); however, they have received much less attention compared to the HMW-GSs mainly because of their greater complexity at both the gene and protein level.

The typical LMW-GSs are usually divided into m-type (methionine), s-type (serine), or i-type (isoleucine) based on the first amino acid present at the N-terminal region of the mature protein. They all possess a highly similar structure consisting of a signal peptide, an N-terminal region (absent in the i-type LMW-GSs), a repetitive region, and a C-terminal region. Differences in the repetitive domain are primarily responsible for LMW-GS gene and protein size variation (D'Ovidio and Masci 2004).

The LMW-GS proteins are encoded by multigene families located in the short arms of the group 1 chromosomes at the *Glu-A3*, *Glu-B3*, and *Glu-D3* loci. The number of LMW-GS genes varies depending on the locus and the variety, but generally more than 15 LMW-GS genes are present in individual bread wheat varieties (Zhang et al. 2013). The genes at each *Glu-3* locus are separated by large intergenic regions of up to 748 Kb between two neighboring genes at the *Glu-B3* locus (Ibba et al. 2017a). Also, recombination was detected at the *Glu-A3*, *Glu-B3*, and *Glu-D3* loci (Dong et al. 2010; Ibba et al. 2017a) and, based on this evidence, each *Glu-3* locus was divided into multiple haplotypes (Ibba et al. 2017b).

The most common methods for the analysis of the LMW-GS profile are sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and reversed-phase high-performance liquid chromatography (RP-HPLC). However, the complexity of the protein bands/peaks obtained using these methods has limited their routine use for the selection of specific LMW-GS alleles (Zhang et al. 2011). Other protein-based methods with greater resolution have been adopted for the LMW-GS analysis, like two-dimensional gel electrophoresis (2-DE, IEF × SDS-PAGE) and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS), but also these methods possess several limitations that have restricted their use: time consuming (2-DE), requirement of skilled personnel, and high cost (Liu et al. 2010).

PCR-based markers have several advantages over biochemical markers, namely, easy to perform, low cost, fast, require a minimum amount of DNA, and can be applied at any stage of the plant life cycle (Kiszonas and Morris 2017; Kumar 1999). For these reasons, PCR-based markers are usually preferred for the identification of favorable alleles in breeding programs. Currently, several PCR-based molecular markers have been developed for the identification of different *Glu-3* allelic variants (Appelbee et al. 2009; Wang et al. 2009, 2010; Zhang et al. 2004; Zhao et al. 2006, 2007). However, even though these markers have been efficiently used for the discrimination of *Glu-A3* and *Glu-B3* alleles (Liu et al. 2010), they were not able to capture the complete LMW-GS genic profile, but only the polymorphism of a few of the LMW-GS genes.

In contrast, the PCR-based molecular marker system developed by Zhang et al. (2011) took advantage of the genetic structure of the LMW-GS genes allowing their differentiation based on the length polymorphism related to the repetitive region of each LMW-GS gene. With this marker system, most of the LMW-GS genes of a given bread wheat variety could be amplified and differentiated by using as few as two PCR reactions (Ibba et al. 2017b). This marker system has been effectively used in numerous studies (Espí et al. 2014; Ibba et al. 2017a, b, c; Zhang et al. 2012, 2013) allowing the sequencing of a greater number of LMW-GS genes, a better understanding of the *Glu-3* loci organization, and the connection between specific LMW-GS genic profiles and the observed *Glu-3* allelic variants. However, this method has some disadvantages, for example, the complexity of the PCR profiles, presence of overlapping

PCR amplicons, impossibility to selectively amplify LMW-GS genes known to be involved in wheat end-use quality, and impossibility to easily differentiate genes coming from the three *Glu-3* loci.

Recently, in the study of Ibba et al. (2017b), the molecular marker system developed by Zhang et al. (2011) was adopted to investigate the role that each haplotype has in the variation of gluten strength and wheat end-use quality parameters and specific haplotypes with greater impact on these phenotypes were identified. Results of that study highlighted the need for developing new molecular markers that specifically target single LMW-GS haplotypes in order to be able to restrict the selection to only the LMW-GS genes mainly involved in wheat end-use quality.

For these reasons, the objective of the present study was to develop a set of haplotype-specific PCR-based molecular markers in order to effectively differentiate specific *Glu-3* haplotype polymorphisms and resolve the ambiguities generated from the molecular marker system developed by Zhang et al. (2011). These markers could greatly improve marker-assisted breeding for wheat quality.

Materials and methods

Plant material

A total of 44 bread wheat (*Triticum aestivum* L.) varieties were selected for the development and validation of the LMW-GS haplotype-specific molecular markers. The LMW-GS genic profile of these varieties was previously determined by Ibba et al. (2017b, c) using the LMW-GS molecular marker system developed by Zhang et al. (2011). The location of the DNA markers was confirmed using the variety Chinese Spring and its chromosome 1 long arm ditelosomic lines (Dt1AL, TA#3102; Dt1BL, TA#3113; Dt1DL, TA#3131) obtained from the Kansas State University Wheat Genetics Resource Center.

Primer design

For primer design, a total of 90 LMW-GS gene sequences (AY453158, EU369701, EU369704-EU369710, EU369715-EU369721, EU369724, EU369729, FJ549929, FJ549932, FJ549933, FJ549939, JF271919, JF339156, JF339184, JF339186, JF339201, JX877780-JX878230, KJ666154, and KX879102) were selected and used for nucleotide alignment through Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) using the default parameters.

From the alignment, specific SNPs unique for the LMW-GS genes at each haplotype located before and/or after the repetitive region of each sequence were identified. Based on the identified SNPs, primer pairs specific for each LMW-GS haplotype were manually designed, and when possible, by including the mismatched nucleotides at the 3′ end of each primer. Their sequences were successively uploaded in the software Primer3 v0.4 (Untergasser et al. 2012), and the relative sequence lengths were adjusted to meet the default primer standards.

The primer sets developed for the *Glu-A3*-3 haplotype were refined by including the forward primer GluA3eF (5′-CAATGAAAACCTTCCTCGTCTG-3′) (Zhang et al. 2004). This additional primer identified the differences between the two variants of the gene A3-646, associated with alleles *Glu-A3e* and *Glu-A3g*. Similarly, the primer set developed for the haplotype *Glu-B3*-3 was modified by including the primer GluB3hF (5′-CATTACCACCACAACAAACATTAA-3′) (adapted from Wang et al. 2009). This primer discriminated the gene B3-688-2 associated with allele *Glu-B3h*.

DNA extraction and PCR amplification

Genomic DNA was extracted from the leaves of 10–14-day-old glasshouse-grown seedlings using the DNeasy 96 Plant Kit (Qiagen, Redwood City, CA, USA). All PCR reactions were performed using the M13 tail fluorescent labeling method. The sequence of the labeled universal primer was 5′-CACGACGTTGTAAAACGAC-3′. The same sequence was added at the 5′ end of each forward primer.

PCR reactions for the Glu-A3-1, Glu-A3-2, Glu-B3-1, Glu-D3-1/5, Glu-D3-4, Glu-D3-5, and Glu-D3-7 markers were performed each in 25-μl reaction volumes containing 100 ng of genomic DNA, 1× Standard *Taq* Buffer (10 mM Tris-HCl, 50 mM KCl, 2.5 mM $MgCl_2$, pH 8.3), 200 mM of dNTPs, 0.2 mM of reverse primer, 0.04 mM of unlabeled forward primer, 0.16 mM of labeled universal primer, and 1 unit of *Taq* DNA Polymerase. PCR reactions for the Glu-A3-3, Glu-B3-2, Glu-B3-3, Glu-D3-2, and Glu-D3-6 molecular markers were performed using the same abovementioned

conditions but with different primer concentrations: 0.28 mM of reverse primer, 0.07 mM of unlabeled forward primer(s), and 0.21 mM of labeled universal primer.

Amplifications were performed in a Peltier Thermal Cycler PTC-200 using the same conditions for all the markers: 5 min at 94 °C, 40 cycles of 30 s at 94 °C, 50 s at 63 °C, and 1 min at 72 °C, and a final extension step of 10 min at 72 °C.

PCR product analysis

PCR products labeled with either FAM or VIC fluorescent dyes were diluted 1:8 in water. Three microliters of the diluted product was added to 9 μl of HiDi-formamide together with either 1 μl of 685-LIZ internal size standard or 0.1 μl of 1200-LIZ internal size standard (Applied Biosystems, Foster City, CA, USA). The mixtures were denatured in a Peltier Thermal Cycler PTC-200 at 95 °C for 5 min and successively analyzed on a 3130xl DNA analyzer (Applied Biosystems, Foster City, CA, USA). The size of each PCR amplicon was determined using the software GeneMarker v3.7 (SoftGenetics, LLC, USA).

Results

Marker development

From the analysis of the LMW-GS gene alignment, SNPs specific for each *Glu-3* haplotype were identified and used to develop a primer set specific for each of these linkage blocks (Supplementary File 1). The only exceptions were the haplotypes *Glu-D3*-1 and *Glu-D3*-5 that were amplified by the same marker. No markers were developed for the haplotype *Glu-D3*-6 due to the absence of polymorphism at this group as reported by Zhang et al. (2013) (Table 1). All the markers were designed in order to amplify the repetitive region of the LMW-GS genes and, consequently, to distinguish the allelic variants of the genes at each haplotype based on length polymorphism.

In addition to the haplotype-specific primers, the GluA3eF and the GluB3hF primers were included in the marker sets specific for the haplotypes *Glu-A3*-3 and *Glu-B3*-3, respectively. The addition of these markers facilitated the discrimination of the genes A3-646 and B3-688 based on their sequence polymorphism.

To ensure the correct annealing location of each primer set, the markers were applied to the common wheat variety Chinese Spring and its ditelosomic lines Dt1AL, Dt1BL, and Dt1DL. As shown in Supplementary Fig. 1, the molecular markers specific for the haplotypes at the *Glu-A3*, *Glu-B3*, and *Glu-D3* loci did not produce any PCR amplicon in the Dt1AL, Dt1BL, and Dt1DL lines, respectively. The exception was the marker Glu-D3-1/5 that produced two tightly close unspecific bands in the Dt1DL line. These bands, however, were absent in Chinese Spring and the Dt1AL and Dt1BL lines.

Marker validation

To validate marker efficiency, each haplotype-specific LMW-GS molecular marker was applied to a set of 44 bread wheat varieties of known LMW-GS genic profile. In order to effectively distinguish each allelic variant at all of the *Glu-3* haplotypes, the PCR results were separated by capillary electrophoresis, producing electropherograms as the ones shown in Supplementary Fig. 2.

According to previous studies, only one LMW-GS gene is located at the haplotype *Glu-A3*-1 (Ibba et al. 2017b). Five different allelic variants of this gene were previously identified among the bread wheat variety validation set (Ibba et al. 2017b,c), namely, A3-353, A3-370, A3-373, A3-391, and a Null allele. These alleles were effectively identified by the Glu-A3-1 molecular marker and corresponded to amplicons with a length of 246, 264, 267, and 286 nt, and no amplicon, respectively (Tables 2 and 3).

The molecular marker Glu-A3-2 was also able to correctly identify the allelic variation at the *Glu-A3*-2 haplotype. Among the 44 bread wheat varieties analyzed, six different alleles were identified. These alleles were named A3-374, A3-394, A3-400, A3-402, A3-408, A3-411, and Null allele (Zhang et al. 2013) and produced, respectively, PCR products with a length of 196, 215, 224, 227, 233, and 236 nt, and no amplicon (Table 2). For all the analyzed varieties, the results previously reported in the literature matched the results obtained with the Glu-A3-2 molecular marker (Table 3).

In contrast to the previous two *Glu-A3* haplotypes, at *Glu-A3*-3 there are typically present two to four i-type LMW-GS genes. Among the varieties analyzed were identified seven different variants of this haplotype. By using the molecular marker Glu-A3-3, each allelic variant of the genes constituting the *Glu-A3*-3 haplotype

Table 1 Haplotype-specific LMW-GS markers

Marker	Forward primer (5′-3′)[a]	Reverse primer (5′-3′)	Reference
Glu-A3-1	GAGAGACMATGGCAGSAG	TTGTTGCAGAASCTGTTGGTA	–
Glu-A3-2	GAGAGACMATGGCAGSAG	TTGCGGTAKAACAGGTTGTGT	–
Glu-A3-3	CGTTGCGCAAATTTCACAG	GYGACCTAGCAAGACATYGCT	–
	CAATGAAAACCTTCCTCGTCTG		Zhang et al. (2004)
Glu-B3-1	CTTATGGTACCACCAACAACAAC	GCAAGATAGATGGATGAACATATGG	–
Glu-B3-2	TCCCTAGCTTGGAGAAACCATT	TGCAAGATAGATGGCTGAACAG	–
Glu-B3-3	TTTGGAGAGACCATCGCAG	GATTTGCGGCAACTGCT	–
	CATTACCACCACAACAAACATTAA		Wang et al. (2009)
Glu-D3-1/5	CAACAACGACCACCATTTTCT	ATGGCAACTGCTCTGCCA	Zhao et al. (2007)
Glu-D3-2	AATTGCGCAGATGGAGACTAGA	GAGATTGGATGGAACCCTGAA	–
Glu-D3-4	TGGCAGCARCAACCATTACA	GATGAACAATAGGGATTTGTTGG	–
Glu-D3-6	ATTACCACCACAACAACAACCA	ACCTAGCAAGACGTTGCGA	–
Glu-D3-7	CAACCATTACCACCACAACAGT	GGCTGAACAATAGGGATTTGC	–

[a] At the 5′ end of each forward primer has been added the M13 universal primer sequence 5′-CACGACGTTGTAAAACGAC-3′

Table 2 List of the LMW-GS genes and their relative fragment size as identified by the haplotype-specific LMW-GS molecular markers

Haplotype	Genes[a]	Fragment size (nt)	Haplotype	Genes	Fragment size (nt)	Haplotype	Genes	Fragment size (nt)
Glu-A3-1	A3-353	246	*Glu-B3*-1	B3-548	440	*Glu-D3*-1	D3-575	303
	A3-370	264		B3-557	452	*Glu-D3*-2	D3-432	637
	A3-373	267	*Glu-B3*-2	B3-510	456		D3-441	646
	A3-391	286		B3-530	477		D3-444	648
Glu-A3-2	A3-374	196	*Glu-B3*-3	B3-544	606	*Glu-D3*-4	D3-385′	292
	A3-394	215		B3-578	640		D3-393	299
	A3-400	224		B3-587	650	*Glu-D3*-5	D3-583	335
	A3-402	227		B3-593	656		D3-586	339
	A3-408	233		B3-601	665		D3-589	342
	A3-411	236		B3-604	667		D3-591	345
Glu-A3-3	A3-480	508		B3-607	670		D3-594	348
	A3-484	514		B3-613	677		D3-597	-
	A3-502	531		B3-621	684	*Glu-D3*-6	D3-525	513
	A3-565	596		B3-624	687		D3-528	516
	A3-568	598		B3-642	705	*Glu-D3*-7	D3-394	292
	A3-573	605		B3-688	751			
	A3-620	651		B3-691	754			
	A3-626	656		B3-813	-[b]			
	A3-640	671		*Glu-B3h*	725			
	A3-643	674						
	A3-646	678						
	A3-662	694						
	Glu-A3e	729						

[a] The LMW-GS gene name refers to the nomenclature adopted by Zhang et al. (2013)

[b] A hyphen indicates the absence of the amplicon

Table 3 *Glu-A3* genic profile of the 44 bread wheat varieties used to validate the *Glu-A3* haplotype-specific molecular markers

Variety	*Glu-A3* haplotype								
	Glu-A3-1	*Glu-A3*-2		*Glu-A3*-3					*Glu-A3e*
Norin-61, Keldin	246[a]	-[b]	224	514	596	598	-	694	-
Hira-1	264	-	-	514	596	598	-	694	-
Nanbu-Komugi, Hira-2	264	-	218	514	596	598	-	694	-
Darius, Evina, HE9817_1_2, Orca	264	-	227	514	596	598	-	694	-
Amadina, India 115-2, Radja, Rescue, Thatcher	264	-	233	531	-	-	-	678	729
Glenlea	264	-	233	531	-	-	-	678	-
ClearstoneCL2, ID1101	264	-	236	531	-	-	-	656	-
Fengmai 27	267	196	218	508	-	-	-	651	-
Gabo, Bolac, Hartog, Opata 85, Pavon 76	267	196	218	531	-	-	-	674	-
Cappelle-Desprez, Insignia, JD, Lincoln, NSA10_2196, Pepital, WA8236, WA8237, WA8238	286	-	224	531	-	-	605	671	-
Seri 82, Gawain, Jufy-1, Theese	286	-	224	531	-	-	-	651	-
Bungulla	286	-	224	531	-	-	-	678	729
India 115-1[c]	286	-	224	531	-	-	-	651/678	729
Chinese Spring	286	-	227	531	-	-	-	651	-
Cheyenne, Courtot, Soissons	-	-	-	531	-	-	-	651	-
Farnum, WA8207	-	-	218	531	-	-	605	671	-

[a] The numeric values represent the amplicon length (nt) obtained with the *Glu-A3* haplotype-specific molecular markers

[b] A hyphen indicates the absence of the amplicon

[c] Underlined varieties exhibit a different *Glu-A3* genic profile compared to previously published data

was identified by a specific amplicon (Table 2). As expected, the results obtained with the molecular marker Glu-A3-3 matched with previously reported results, with the exception of variety India 115-2 (Table 3). This variety in fact exhibited both alleles A3-620 (651 nt) and A3-646 (678 nt) whereas in previous studies, it exhibited only allele A3-620. Also, by including in the Glu-A3-3 marker set the primer GluA3eF, it was possible to distinguish the two different variants of the gene A3-646 previously reported (Zhang et al. 2013), and associated with the *Glu-A3* alleles *Glu-A3e* and *Glu-A3g*. When the A3-646 variant associated with allele *Glu-A3e* was present, two amplicons were obtained (678 and 729 nt), whereas when only a single amplicon was present (the smaller with 678 nt), the A3-646 variant allele *Glu-A3g* was identified (Supplementary Fig. 3).

The marker Glu-B3-1 was able to discriminate the three allelic variants detected at the haplotype *Glu-B3*-1: B3-548 (440 nt), B3-557 (452 nt), and a Null allele (absence of an amplicon). For all the varieties, the Glu-B3-1 profile identified here was in accordance with previous results (Tables 2 and 4).

The marker Glu-B3-2 was developed for the haplotype *Glu-B3*-2. The analyzed varieties exhibited three different allelic variants at this group: B3-510, B3-530, and a Null allele. All these alleles were correctly identified across the varieties by the Glu-B3-2 molecular marker and produced amplicons with lengths of 456 and 477 nt, and no amplicon, respectively (Tables 2 and 4).

Similar to the haplotype *Glu-A3*-3, at the haplotype *Glu-B3*-3 there are generally multiple s-type LMW-GS genes (1 to 3) tightly linked together. According to previous studies (Ibba et al. 2017b, c), the set of bread wheat varieties analyzed here exhibited 12 different variants of the *Glu-B3*-3 haplotype. The molecular marker Glu-B3-3, developed to amplify the s-type LMW-GS genes from the *Glu-B3* locus, was able to successfully detect the gene polymorphisms associated

Table 4 *Glu-B3* genic profile of the 44 bread wheat varieties used to validate the *Glu-B3* haplotype-specific molecular markers

Variety	*Glu-B3* haplotype						
	Glu-B3-1	*Glu-B3*-2	*Glu-B3*-3				*Glu-B3h*
Theese	440[a]	456	640	606	684	-[b]	-
Opata 85, Jufy-1	440	456	-	-	751	754	-
Glenlea, HE9817_1_2, NSA10_2196	440	477	640	606	684	-	-
Darius	440	477	640	606	687	-	-
Chinese Spring	440	477	640	656	687	-	-
Fengmai 27, Gawain, Nanbu-Komugi	440	477	640	665	684	-	-
ClearstoneCL2, Evina, Keldin	440	477	640	667	684	-	-
Gabo, Bolac, India 115-2, Radja, Soissons	440	477	640	670	684	-	-
Courtot[c]	440	477	640	677	684	-	-
Norin-61	440	477	-	-	751	754	-
Hartog, Hira-1, Hira-2, India 115-1, Lincoln, Pavon 76, Rescue, Thatcher	440	477	-	-	751	m[d]	725
Insignia, Orca, Pepital	440	477	-	-	751	-	-
JD, WA8236, WA8237, WA8238	452	477	640	650	684	-	-
Cappelle-Desprez	-[d]	477	640	606	684	-	-
Cheyenne, ID1101	-	477	640	667	684	-	-
Bungulla	-	477	640	705	684	-	-
Farnum, WA8207	-	477	-	-	751	m	725
Amadina, Seri 82	-	-	-	-	-	-	-

[a] The numeric values represent the amplicon length (nt) obtained with the *Glu-B3* haplotype-specific molecular markers

[b] A hyphen indicates the absence of the amplicon

[c] Underlined varieties exhibit a different *Glu-B3* genic profile compared to previously published data

[d] The "m" indicates that an amplicon is missing compared to previously published data

with the *Glu-B3*-3 variants previously identified. The only exception was the allele B3-813, which did not produce any PCR product (Table 2). From the results of this marker, there were also detected some alleles that were not previously characterized like B3-613 (676 nt) and B3-642 (704 nt). Few discrepancies between the present and previous results were identified. Specifically, according to the results obtained from the marker Glu-B3-3, the variety Courtot exhibited allele B3-613 (676 nt), whereas according to previous results, this variety did not exhibit any allelic variants related to that gene. Similarly, varieties JD, WA8236, WA8237, and WA8238 all exhibited allele B3-587 (650 nt) according to this study. In contrast, according to previous studies, these varieties did not possess any of the allelic variants of that gene (Table 4). In order to better characterize the genes at this haplotype, one additional primer was included in the marker set specific for *Glu-B3*-3. This marker allowed the discrimination of the B3-688-3 variant (opposed to B3-688-1, B3-688-2, and B3-688-4) (Supplementary Fig. 4). When the allele B3-688-3 was present, two amplicons were produced with lengths of 729 and 751 nt, whereas the other B3-688 allelic variants produced only one amplicon of 751 nt (Supplementary Fig. 4).

For the haplotypes *Glu-D3*-1 and *Glu-D3*-5, one molecular marker (Glu-D3-1/5) was developed due to the high sequence similarity of the genes at these haplotypes. This marker produced a 303-nt amplicon, corresponding to the *Glu-D3*-1 gene D3-575, and an amplicon with a length of 335, 339, 342, 345, or 348 nt, or no amplicon, corresponding to the allelic variants of the gene of the haplotype *Glu-D3*-5 (Table 2). The genic profile of these haplotypes matched the results previously reported. The exception was the variety Cappelle-Desprez that, according to Ibba et al. (2017c), exhibited allele D3-597 whereas here it exhibited allele D3-594 (348 nt) (Table 5).

Table 5 *Glu-D3* genic profile of the 44 bread wheat varieties used to validate the *Glu-D3* haplotype-specific molecular markers

Variety	*Glu-D3* haplotype					
	Glu-D3-1	*Glu-D3*-2	*Glu-D3*-4	*Glu-D3*-5	*Glu-D3*-6	*Glu-D3*-7
Farnum	303[a]	637	292	342	513	292
Pepital, Courtot, Gawain, HE9817_1_2[b], Theese, WA8207	303	637	292	-[c]	-	-
Bolac	303	637	299	339	516	292
Norin-61, ClearstoneCL2	303	637	299	342	513	292
Glenlea, ID1101, Orca, Thatcher	303	637	299	342	516	292
Soissons, Keldin	303	637	299	345	513	292
Insignia	303	637	299	345	516	292
Cappelle-Desprez, Evina	303	637	299	348	513	292
NSA10_2196	303	637	299	-	-	-
Nanbu-Komugi, Radja	303	646	299	335	513	292
Hira-1, Hira-2	303	646	299	335	516	292
Opata 85, Hartog, India 115-2, Pavon 76, Rescue	303	646	299	339	513	292
Gabo, Jufy-1	303	646	299	339	516	292
Chinese Spring, Bungulla, Cheyenne, JD, WA8236, WA8237, WA238	303	646	299	342	513	292
Lincoln	303	646	299	348	513	292
Seri 82	303	646	299	-	513	-
Fengmai 27	303	648	299	339	513	292
Darius	303	-	299	342	513	292
India 115-1	303	-	-	-	513	292
Amadina	303	637	292	-	-	-

[a] The numeric values represent the amplicon length (nt) obtained with the *Glu-D3* haplotype-specific molecular markers

[b] Underlined varieties exhibit a different *Glu-D3* genic profile compared to previously published data

[c] A hyphen indicates the absence of the amplicon

The marker Glu-D3-2 was developed to capture the variation at the haplotype *Glu-D3*-2. This marker generated an amplicon of 637, 646, or 648 nt, or no amplicon, corresponding to alleles D3-432, D3-441, D3-444, and a Null allele, respectively (Table 2). The present results obtained with this molecular marker were consistent with previous results (Table 5).

To detect polymorphism at the haplotype *Glu-D3*-4, the molecular marker Glu-D3-4 was developed. This marker produced PCR amplicons of 292 or 299 nt, or no amplicon corresponding, respectively, to alleles D3-385′, D3-393, and a Null allele (Table 2). The results obtained here (Table 5) were consistent with results previously reported with the exception of variety NSA10_2196. According to Ibba et al. (2017b), this variety possessed allele D3-385′ whereas from the present analysis, the variety had allele D3-393 (Table 5).

The marker Glu-D3-6 was developed to discriminate the alleles at the haplotype *Glu-D3*-6. Among the varieties analyzed, three different alleles were detected, namely D3-525, D3-528, and a Null allele. They respectively amplified a 513- or 516-nt amplicon, or no amplicon (Table 2). The *Glu-D3*-6 genic profile previously reported was consistent with the results obtained here apart from varieties HE9817_1_2 and WA8207. These varieties had previously been shown to carry allele D3-375 (Ibba et al. 2017b), but here no amplicons were observed using Glu-D3-6 (Table 5).

Polymorphism at the haplotype *Glu-D3*-7 was detected by the molecular marker Glu-D3-7. The set of bread wheat varieties chosen for the analysis exhibited either allele D3-394 or a Null allele. Both of the alleles were correctly detected by the marker and were associated with a PCR product of 292 nt or with no amplicon

(Table 2). Exceptions were varieties Farnum and WA8207 that in this study exhibited a Null allele at the haplotype *Glu-D3*-7 whereas, according to previous studies (Ibba et al. 2017b), possessed a previously uncharacterized allele (D3-382) (Table 5).

Discussion

The LMW-GSs are one of the major components of gluten and both their quantity and quality strongly influence wheat dough functionality (Shewry et al. 2002). For these reasons, numerous methods have been developed for the identification and selection of favorable LMW-GS alleles. These methods can be divided into two main groups: protein-based methods (SDS-PAGE, RP-HPLC, 2-DE, MALDI/TOF) and DNA-based methods (PCR-based molecular markers) (Kiszonas and Morris 2017; Liu et al. 2010).

Typically, PCR-based molecular markers are the preferred method in breeding programs for the selection of desired alleles because they are easy to use, fast, accurate, and independent of the plant stage of development (Kumar 1999). Based on the available LMW-GS genes, allele specific markers were developed for the *Glu-A3* and *Glu-B3* loci (Wang et al. 2009, 2010). However, the relatively high number of primer pairs required for the *Glu-A3* and *Glu-B3* allelic characterization (7 and 10, respectively) and the impossibility of these markers to capture the complexity of the *Glu-3* loci limited their use. In contrast, the PCR-based molecular marker system developed by Zhang et al. (2011) was able to characterize most of the LMW-GS genes present in a given bread wheat variety through their length polymorphism. The development of this marker system facilitated a deeper understanding of the *Glu-3* loci structure, the LMW-GS gene sequences, and the connection between different LMW-GS genic profiles and their relative LMW-GS alleles (Ibba et al. 2017a, b; Zhang et al. 2013).

In the present study, we developed a set of PCR-based haplotype-specific LMW-GS molecular markers. This set, different from the molecular markers developed by Wang et al. (2009, 2010), is able to identify and differentiate most of the LMW-GS genes present in each bread wheat variety. Also, different from the marker system developed by Zhang et al. (2011), these markers are easier to score, less ambiguous, take into account the presence of recombination at each *Glu-3* locus and can be used to selectively amplify only the LMW-GS haplotypes more highly involved in wheat end-use quality. The molecular markers developed by Zhang et al. (2011), in fact, produced complex electropherograms displaying multiple DNA fragments with lengths ranging from 353 to 813 nt (for the LMWGS1 marker). The interpretation of this output could be challenging since (1) the molecular marker amplifies indiscriminately LMW-GS genes from the *Glu-A3*, *Glu-B3*, and *Glu-D3* loci and (2) DNA fragments with the same length but representing different alleles could be present.

The primer sets used in this study were developed based on numerous LMW-GS genes and were found to be specific for each *Glu-3* locus and for the haplotypes at each *Glu-3* locus. In contrast to the molecular markers developed by Zhang et al. (2011), the main purpose here was to selectively detect polymorphism in specific *Glu-3* haplotypes known to have an effect on wheat end-use quality rather than indistinctly amplify all the LMW-GS genes of a given bread wheat variety. According to previous studies (Dong et al. 2010; Ibba et al. 2017a), there is recombination at the *Glu-A3*, *Glu-B3*, and *Glu-D3* loci. These findings required a re-evaluation of the *Glu-3* loci as sets of multiple haplotypes rather than as unique linkage blocks with "allele" status. Consequently, the methods of analysis to assess LMW-GS polymorphism also required reconsideration. On the one hand, the selection of specific LMW-GS alleles based on the polymorphism of a unique gene (Wang et al. 2009, 2010) could no longer be considered as representative of the *Glu-3* loci polymorphism but, on the other hand, the analysis of the entire LMW-GS genic profile of a given bread wheat variety could be excessively complex, especially because many of the LMW-GS gene are pseudogenes and/or do not strongly affect wheat end-use quality (Ibba et al. 2017b).

In general, the molecular markers developed in this study were equally or more efficient than the molecular marker system of Zhang et al. (2011), and the results obtained in this study almost always corresponded with previously reported results (Ibba et al. 2017b, c). This result is notable given the large genetic diversity of the included varieties, specifically chosen to represent the genetic variation at these three LMW-GS loci.

Among the three sets of markers developed for the haplotypes at the *Glu-A3* locus (Glu-A3-1, Glu-A3-2, and Glu-A3-3), discrepancies were identified only for the *Glu-A3*-3 profile of the variety India 115-1. Here, India 115-1 exhibited both genes A3-620 (651 nt) and

A3-646 (678 nt), whereas according to Ibba et al. (2017c), it possessed only the gene A3-646. Zhang et al. (2013) reported that these two genes are allelic variants of the same gene so it is unlikely that they are both present in a single variety. It is possible that these discrepancies are the result of seed/DNA mislabeling or seed contamination. According to Ibba et al. (2017c) the LMW-GS genic profile of the haplotype *Glu-A3*-3 was representative of the *Glu-A3* alleles that were established by SDS-PAGE. However, by using the molecular markers of Zhang et al. (2011), it was not possible to distinguish alleles *Glu-A3e* and *Glu-A3g*. The varieties exhibiting these two alleles, in fact, possessed the same *Glu-A3*-3 genic profile represented by alleles A3-502 and A3-646, but differed in the sequence of allele A3-646. By including in the Glu-A3-3 marker set the primer GluA3eF developed by Zhang et al. (2004), it was possible to differentiate the two A3-646 gene sequences indicating the possibility to use this marker set to directly identify the *Glu-A3* alleles present in a single bread wheat variety.

Among the three sets of markers developed for the *Glu-B3* haplotypes, differences between the results obtained here and previous results were identified only for the haplotype *Glu-B3*-3. Interestingly, all of the varieties that according to the studies of Ibba et al. (2017b, c) exhibited only alleles B3-578 and B3-621, in this study, they produced a third amplicon. Specifically, varieties JD, WA8236, WA8237, and WA8238 all possessed allele B3-587 (650 nt), whereas varieties Courtot and Bungulla exhibited two novel alleles B3-613 (677 nt) and B3-642 (705 nt), respectively. In the case of allele B3-587, it is possible that this allele was not identified through the marker system of Zhang et al. (2011) because its amplicon overlapped with the *Glu-D3* gene D3-589, present in all the four abovementioned varieties. The two other alleles (B3-613 and B3-642), in contrast, were not previously detected, likely because of their relatively low amplification efficiency. Similar to the haplotype *Glu-A3*-3, the genic profile of the haplotype *Glu-B3*-3 was associated with different *Glu-B3* allelic variants (Ibba et al. 2017c). By including in the Glu-B3-3 marker set the primer GluB3hF modified from Wang et al. (2009), it was possible to discriminate the B3-688-3 variant from the other B3-688 alleles. According to Zhang et al. (2012), allele B3-688-3 is associated with allele *Glu-B3h*. Results obtained here suggest that the Glu-B3-3 marker set could be effectively used to discriminate most of the *Glu-B3* alleles with the exceptions of the alleles *Glu-B3c* and *Glu-B3d*, which exhibit the same *Glu-B3*-3 profile, and the alleles *Glu-B3ac* and *Glu-B3ad*, which appear to be determined by both the *Glu-B3*-2 and the *Glu-B3*-3 genic profile (Ibba et al. 2017c).

As found in the *Glu-A3* and *Glu-B3* genic profiles, the *Glu-D3* genic profile obtained here was consistent with the one previously reported (Ibba et al. 2017b, c) with few exceptions. Specifically, from the analysis of the haplotype *Glu-D3*-4, variety NSA10_2196 exhibited allele D3-393 (299 nt) whereas according to the previous study (Ibba et al. 2017b), it possessed allele D3-385′. In this case, it is possible that the allele D3-385′ was incorrectly identified. This allele in fact produces an amplicon with the same length as another *Glu-D3* gene (D3-385), present in the variety NSA10_2196, making its identification through the molecular marker system of Zhang et al. (2011) problematic. In contrast, at the haplotype *Glu-D3*-5, varieties Cappelle-Desprez and Evina both exhibited allele D3-594 (348 nt) whereas according to Ibba et al. (2017b, c), they possessed allele D3-597. From these results, it appears that either the molecular marker developed here is unable to distinguish the allele D3-597 or that this allele was previously incorrectly identified. More varieties exhibiting either of the two alleles should be analyzed in order to reach a conclusion. Finally, according to the study of Ibba et al. (2017b), varieties HE9817_1_2 and WA8207, and varieties Farnum and HE9817_1_2 hypothetically possessed the previously uncharacterized alleles D3-375 (*Glu-D3*-6) and D3-382 (*Glu-D3*-7), respectively. The molecular markers Glu-D3-6 and Glu-D3-7 did not detect these novel alleles. These differences could be determined by both an incorrect gene assignment by Ibba et al. (2017b), but also by a possible limitation of the Glu-D3-6 and Glu-D3-7 molecular markers. More studies will be needed in order to characterize the alleles D3-375 and D3-382 previously identified and to determine if they are actually allelic variants of the genes at the haplotypes *Glu-D3*-6 and *Glu-D3*-7 or represent unique genes.

Conclusion

In the present study, three, three, and five PCR-based molecular marker sets specific for the haplotypes at the *Glu-A3*, *Glu-B3*, *Glu-D3* locus, respectively, were developed. Each of these molecular markers was found to

be specific for the different *Glu-3* loci and for each haplotype. The LMW-GS genic profile detected using these molecular markers corresponded to the results obtained in previous studies confirming the reliability of these haplotype-specific molecular markers. Also, in many cases, the markers developed here were more efficient than those previously reported. Specifically, compared to the molecular markers developed for the *Glu-A3* and *Glu-B3* allele discrimination (Wang et al. 2009, 2010; Zhang et al. 2004), the markers developed here were more accurate since they were able to detect most of the LMW-GS genes rather than the polymorphism of single genes. Also, they could discriminate different *Glu-A3* alleles, and most of the *Glu-B3* alleles by performing a single PCR reaction as opposed to seven (Zhang et al. 2004), ten (Wang et al. 2009), or four (Wang et al. 2010). In contrast, compared to the molecular marker system developed by Zhang et al. (2011), the haplotype-specific molecular markers developed here are easier to interpret, less ambiguous, and selectively amplify only specific LMW-GS haplotypes rather than a relatively large number of amplicons. Even if more studies are needed to confirm the efficiency of this set of markers, the primers developed here could be effectively used for the discrimination of the known *Glu-A3* and *Glu-B3* alleles, to solve the ambiguities from the molecular marker of Zhang et al. (2011) and, in the future, could be employed for the selection of specific LMW-GS haplotypes more highly involved in wheat end-use quality.

Acknowledgments The authors thank Dr. Derick Jiwan for constructive suggestions on this work. Stacey Sykes and Shawna Vogl assisted in the preparation and submission of the manuscript.

Funding information This research was supported in part by NIFA AFRI Grant No. 2013-67013-21226 and CRIS Project 2090-43440-007-00-D.

References

Appelbee MJ, Mekuria GT, Nagasandra V, Bonneau JP, Eagles HA, Eastwood RF, Mather DE (2009) Novel allelic variants encoded at the *Glu-D3* locus in bread wheat. J Cereal Sci 49: 254–261

Barak S, Mudgil D, Khatkar BS (2015) Biochemical and functional properties of wheat gliadins: a review. Crit Rev Food Sci 55:357–368

Bietz JA, Wall JS (1973) Isolation and characterization of gliadin-like subunits from glutelin. Cereal Chem 50(5):537–547

Bonafede MD, Tranquilli G, Pflüger LA, Peña RJ, Dubkovsky J (2015) Effect of allelic variation at the *Glu-3*/*Gli-1* loci on breadmaking quality parameters in hexaploid wheat (*Triticum aestivum* L.). J Cereal Sci 62:143–150

D'Ovidio R, Masci S (2004) The low-molecular-weight glutenin subunits of wheat gluten. J Cereal Sci 39:321–339

Dong L, Zhang X, Liu D, Fan H, Sun J, Zhang Z, Qin H, Li B, Hao S, Li Z, Wang D, Zhang A, Ling HQ (2010) New insights into the organization, recombination, expression and functional mechanism of low molecular weight glutenin subunit genes in bread wheat. PLoS One 5:e13548

Espí A, Rodríguez-Quijano M, Vazquez JF, Carrillo JM, Giraldo P (2014) Molecular characterization of *Glu-B3* locus in wheat cultivars and segregating populations. J Cereal Sci 60:374–381

He ZH, Liu L, Xia XC, Liu JJ, Peña RJ (2005) Composition of HMW and LMW glutenin subunits and their effects on dough properties, pan bread, and noodle quality of Chinese bread wheats. Cereal Chem 82:345–350

Ibba MI, Kiszonas AM, Morris CF (2017a) Evidence of intralocus recombination at the *Glu-3* loci in bread wheat (*Triticum aestivum* L.). Theor Appl Genet 130(5):891–902

Ibba MI, Kiszonas AM, Morris CF (2017b) Influence of low-molecular-weight glutenin subunit haplotypes on dough rheology in elite common wheat varieties. Cereal Chem 94: 1016–1027

Ibba MI, Kiszonas AM, Guzmán C, Morris CF (2017c) Definition of the low molecular weight glutenin subunit gene family members in a set of standard bread wheat (*Triticum aestivum* L.) varieties. J Cereal Sci 74:263–271

Jin H, Zhang Y, Li G, Mu P, Fan Z, Xia X, He Z (2013) Effects of allelic variation of HMW-GS and LMW-GS on mixograph properties and Chinese noodle and steamed bread qualities in a set of Aroona near-isogenic wheat lines. J Cereal Sci 57: 146–152

Kiszonas AM, Morris CF (2017) Wheat breeding for quality: a historical review. Cereal Chem 95:17–34

Kumar LS (1999) DNA markers in plant improvement: an overview. Biotechnol Adv 17:143–182

Liu L, He Z, Yan J, Zhang Y, Xia X, Peña RJ (2005) Allelic variation at the *Glu-1* and *Glu-3* loci, presence of the 1B.1R translocation, and their effects on mixographic properties in Chinese bread wheats. Euphytica 142:197–204

Liu TM, Ikeda TM, Branlard G, Peña RJ, Rogers WJ, Lerner SE, Kolman MA, Xia X, Wang L, Ma W, Appels R, Yoshida H, Wang A, Yan Y, He Z (2010) Comparison of low molecular weight glutenin subunits identified by SDS-PAGE, 2-DE, MALDI-TOF-MS and PCR in common wheat. BMC Plant Biol 10:124

Lukow OM, Payne PI, Tkachuk R (1989) The HMW glutenin subunit composition of Canadian wheat cultivars and their association with bread-making quality. J Sci Food Agric 46: 451–460

Nieto-Taladriz MT, Perretant MR, Rousset M (1994) Effect of gliadins and HMW and LMW subunits of glutenin on dough properties in the F6 recombinant inbred lines from a bread wheat cross. Theor Appl Genet 88:81–88

Payne PI (1987) Genetics of wheat seed storage proteins and the effect of allelic variation on bread-making quality. Annu Rev Plant Physiol 38:141–153

Payne PI, Nightingale MA, Krattiger AF, Holt LM (1987) The relationship between HMW glutenin subunit composition and the bread making quality of British grown wheat varieties. J Sci Food Agric 40:51–65

Shewry PG, Halford NG, Belton PS, Tatham AS (2002) The structure and properties of gluten: an elastic protein from wheat grain. Philos Trans R Soc Lond 357:133–142

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3—new capabilities and interfaces. Nucleic Acids Res 40(15):115

Wang LH, Li GY, Pena RJ, Xia XC, He ZH (2010) Development of STS markers and establishment of multiplex PCR for *Glu-A3* alleles in common wheat (*Triticum aestivum* L.). J Cereal Sci 51:305–312

Wang LH, Zhao XL, He ZH, Ma W, Appels R, Pena RJ, Xia XC (2009) Characterization of low-molecular-weight glutenin subunit *Glu-B3* genes and development of STS markers in common wheat (*Triticum aestivum* L.). Theor Appl Genet 118:525–539

Zhang W, Gianibelli MC, Rampling LR, Gale KR (2004) Characterisation and marker development for low molecular weight glutenin genes from *Glu-A3* alleles of bread wheat (*Triticum aestivum* L). Theor Appl Genet 108:1409–1419

Zhang X, Jin H, Zhang Y, Liu D, Li G, Xia X, He Z, Zhang A (2012) Composition and functional analysis of low-molecular-weight glutenin alleles with Aroona near-isogenic lines of bread wheat. BMC Plant Biol 12:243

Zhang X, Liu D, Yang W, Liu K, Sun J, Guo X, Li Y, Wang D, Ling H, Zhang A (2011) Development of a new marker system for identifying the complex members of the low-molecular-weight glutenin subunit gene family in bread wheat (*Triticum aestivum* L.). Theor Appl Genet 122:1503–1516

Zhang X, Liu D, Zhang J, Jiang W, Luo G, Yang W, Sun J, Tong Y, Cui D, Zhang A (2013) Novel insights into the composition, variation, organization, and expression of the low-molecular-weight glutenin subunit gene family in common wheat. J Exp Bot 64:2027–2040

Zhao XL, Xia XC, He ZH, Gale KR, Lei ZS, Appels R, Ma W (2006) Characterization of three low-molecular-weight *Glu-D3* subunit genes in common wheat. Theor Appl Genet 113: 1247–1259

Zhao XL, Xia XC, He ZH, Lei ZS, Appels R, Yang Y, Sun QX, Ma W (2007) Novel DNA variations to characterize low molecular weight glutenin *Glu-D3* genes and develop STS markers in common wheat. Theor Appl Genet 114:451–460