3 Biotech (2018) 8:56
https://doi.org/10.1007/s13205-017-1081-2

ORIGINAL ARTICLE

Proteomic analysis of low-molecular-weight glutenin subunits and relationship with their genes in a common wheat variety

Hye-Rang Beom[1] · Jin Sun Kim[1] · You-Ran Jang[1] · Sun-Hyung Lim[1] · Chang-Kug Kim[1] · Choung Keun Lee[1] · Jong-Yeol Lee[1]

Received: 7 August 2017 / Accepted: 28 December 2017 / Published online: 5 January 2018

Abstract

Although many studies on low-molecular-weight glutenin subunit (LMW-GS) function have been reported, a comprehensive comparison between specific genes and their protein product is still lacking. This study aimed to link the 43 genes isolated from the Korean wheat variety "Jokyoung" in the authors' previous study to their protein products. Proteins were separated using two-dimensional gel electrophoresis (2-DGE) and identified by tandem mass spectrometry (MS/MS) at the gene haplotype level. Using MS/MS analysis of 17 protein spots, two spots were identified in the *Glu-A3* locus and the corresponding haplotype was *GluA3-13*(*Glu-A3c*). Six spots were identified in the *Glu-B3* locus and the corresponding haplotypes were *GluB3-33* and *GluB3-43* (*Glu-B3h*). Eight spots were identified in the *Glu-D3* locus and the corresponding haplotypes were *GluD3-11*, *GluD3-21*, *GluD3-31*, *GluD3-5*, and *GluD3-6* (*Glu-D3a*), and one spot was contaminated with gamma gliadin. Phylogenetic analysis and alignment of nucleotide and amino acid sequences assigned 35 of the 43 genes to seven haplotypes: *GluA3-13*, *GluB3-43*, *GluD3-11*, *GluD3-21*, *GluD3-31*, *GluD3-42*, and *GluD3-5*. Taken together, except for *GluB3-33* and *GluD3-6*, which were not isolated, linking of each gene to the corresponding protein products at the gene haplotype level was accomplished using proteomic tools and phylogenetic analysis.

Keywords Low-molecular-weight glutenin subunit genes · Protein products · Gene haplotype · Wheat quality

Introduction

Gluten proteins play a crucial role in determining the unique baking quality of wheat by conferring water absorption capacity, cohesion, viscosity, and elasticity in dough. The major constituents of gluten are divided into soluble gliadin and insoluble glutenin, depending on the solubility in the aqueous alcohol solution [60–70% (v/v) ethanol or 50% (v/v) propan-1-ol] (Wieser 2007).

Glutenin protein is divided into high-molecular-weight glutenin subunits (HMW-GS) and low-molecular-weight glutenin subunits (LMW-GS), based on molecular weight assessed by SDS-PAGE (Payne 1987). LMW glutenin subunits are encoded by the multigene families at the *Glu-3* loci of the A, B, and D chromosomes of common wheat (D'Ovidio and Masci 2004). LMW-GS was classified as LMW-s, LMW-m, and LMW-i based on the first residue of N-terminal amino acid, and each amino acid is serine, methionine, and isoleucine, respectively. Although the N-terminal amino acid sequence of the LMW-s type subunits is only present as SHIPGL-, the LMW-m type subunits are variously present as METSHIGPL-, METSRIRGL-, and METSCIPGL- (Kasarda et al. 1988; Lew et al. 1992; Masci et al. 1995; Tao and Kasarda 1989). The LMW-i type, which lacks the N-terminal region and immediately starts the ISQQQQ-repeat sequence after signal protein, was identified (Cassidy et al. 1998; Cloutier et al. 2001; Ikeda et al. 2002; Pitts et al. 1988; Zhang et al. 2004).

The LMW-GS gene alleles were divided into 20 band patterns by SDS-PAGE mobility, including six alleles (*a*, *b*, *c*, *d*, *e*, and *f*) on the *Glu-A3* locus, nine alleles (*a*, *b*, *c*, *d*, *e*, *f*, *g*, *h*, and *i*) on the *Glu-B3* locus, and five alleles (*a*,

Hye-Rang Beom and Jin Sun Kim contributed equally.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s13205-017-1081-2) contains supplementary material, which is available to authorized users.

✉ Jong-Yeol Lee
jy0820@korea.kr

[1] National Institute of Agricultural Science, RDA, Jeonju 54874, Republic of Korea

b, *c*, *d*, and *e*) on the *Glu-D3* locus from 222 wheat cultivars (Gupta and Shepherd 1990). Twelve groups of LMW-GS genes have been identified in the Norin 61 cDNA library (Ikeda et al. 2006). Additionally, 12 active genes in Glenlea were identified, with 1, 2, and 9 in 1A, 1B, and 1D, respectively (Huang and Cloutier 2008). Using gene-specific primers, LMW-GS genes were identified from Aroona near-isogenic lines (NILs) or a set of standard cultivars containing specific alleles classified by protein electrophoretic mobility (Liu et al. 2010; Wang et al. 2009, 2010; Zhang et al. 2004; Zhao et al. 2006, 2007). Both the molecular marker system and the gene cloning method were applied to investigate the composition of LMW-GS genes in large populations, including Aroona near-isogenic lines and the micro-core collections (MCC) of Chinese wheat germplasm (Zhang et al. 2012, 2013).

There have been studies that analyzed the relationships of LMW-GS genes and their proteins using individual cultivars. A Chinese wheat variety, Xiaoyan 54, with 14 unique LMW-GS genes, was identified using BAC (bacterial artificial chromosome) library screening and proteomic analysis. Four genes were identified in *Glu-A3*, three genes in *Glu-B3*, and seven genes in *Glu-D3* (Dong et al. 2010). A Korean wheat variety, Keumkang, contained a total of 36 LMW-GS genes and pseudogenes which were amplified, including 11 *Glu-3* gene haplotypes, 2 from the *Glu-A3* locus, 2 from the *Glu-B3* locus, and 7 from the *Glu-D3* locus. To establish relationships between gene haplotypes and their protein products, a glutenin protein fraction was separated by 2-DGE, and 17 protein spots were analyzed by N-terminal amino acid sequencing and tandem mass spectrometry (MS/MS) (Lee et al. 2016). Recently, some reports using proteomics-based methods have characterized the wheat endosperm proteins (Dong et al. 2010; Ikeda et al. 2006; Liu et al. 2010).

Despite these advanced efforts, difficulties in LMW-GS research are caused by (1) bread wheat being a hexaploid with a complex genome size of 16 GB; (2) LMW-GS are encoded by multicopy genes which were estimated to have 30–40 copy numbers (Cassidy et al. 1998; Sabelli and Shewry 1991); (3) LMW-GS is difficult to isolate due to overlapping with the gliadins (Nielsen et al. 1968) and the larger number of repeat sequences; and (4) wheat genome sequencing is not complete, so it is difficult to study the accurate characterization of LMW-GS gene and the function of its proteins. For these reasons, the study of linkage between the specific LMW-GS genes and their protein products is difficult and rarely reported.

The aim of this study was to link the 43 genes (isolated during the authors' previous studies) to their proteins at the gene haplotype level by using spots identified in 2-DGE and then applying MS/MS. According to this study, the more accurate results of LMW-GS genes and their corresponding proteins will be used as a molecular basis to contribute to LMW-GS characteristic studies in the future.

Materials and methods

Plant materials

A Korean common wheat variety, Jokyoung, wheat seeds were grown and harvested from the National Institute of Crop Science, Iksan, Korea, under natural light conditions in 2016. The wheat cultivars, Cheyenne (*Glu-A3c*), Rescue (*Glu-B3h*), and Chinese Spring (*Glu-D3a*), which were used as standards for LMW-GS identification, were provided by the National Plant Germplasm System of USDA-ARS (Albany, CA, USA).

Glutenin extraction, SDS-PAGE, and 2-DGE

Glutenin extraction was performed by Singh et al. (1991) where 25 mL of 50% 1-propanol was added to 0.5 g of ground wheat flour and reacted at 60 °C for 30 min. The supernatant was then removed by centrifugation at 10,000×*g* for 10 min. This step was repeated once more for complete gliadin removal. To extract glutenin, 2.5 mL of 1% propan-1-propanol and 0.08 M Tris–HCl, pH 8.0, with 1% DTT was added to the pellet and incubated at 65 °C for 30 min, and then centrifuged at 10,000×*g* for 5 m. For protein alkylation, 2.5 mL of 50% 1-propanol and 0.08 M Tris–HCl, pH 8.0, with 1.4% 4-vinylpyridine was added to the pellet, reacted for 15 min at 65 °C, and then centrifuged at 10,000×*g* for 2 min. To precipitate protein, the supernatant was stored at 4 °C for 1 day. The supernatant was then centrifuged at 12,000 rpm for 2 min, immersed in acetone containing 15% TCA, and stored at − 20 °C until use. During the SDS-PAGE analysis, the glutenin protein stored at − 20 °C was centrifuged at 14,000 rpm for 10 min, and the supernatant was removed. After drying the remaining pellet, 5 µL was loaded onto 12.5% SDS-PAGE (SE260 Ruby, Hofer) gel by adding 70 µL of sample buffer. After electrophoresis at 70 V for 1 h and at 130 V for 5 h, the cells were stained with Coomassie Blue R-250 (CBB) solution and destained.

To perform 2-DGE, the glutenin protein stored at − 20 °C was centrifuged at 14,000 rpm for 10 min, and the supernatant was removed to dry the pellet. After the addition of 70 µL of rehydration buffer, the pellet was completely dissolved. The Bradford method was used for protein determination and BSA was used as a standard for the standard curve. An 18 cm IPG strip (pH 6–11, GE Healthcare, USA) was used for IEF performance and was confirmed using SDS-PAGE (18 × 18 cm, Bio-Rad, USA).

Comparative gene haplotype analysis of LMW-GS genes

A total of 43 wheat LMW-GS genes, isolated from the authors' previous study (Lee et al. 2010), were aligned with

the reported *Glu-A3*, *Glu-B3*, and *Glu-D3* genes (Wang et al. 2009, 2010; Zhao et al. 2006, 2007). A phylogenetic tree was made using the neighbor-joining method (Saitou and Nei 1987) and MEGA 6 software (Tamura et al. 2013) to confirm the similarity between the Jokyoung genes and the previously reported low-molecular-weight glutenin gene.

Identification of proteins by specific enzyme treatment and LC ESI–MS/MS

Two-dimensional electrophoresis gel spots were excised and treated with chymotrypsin for MS/MS analysis. The chymotryptic peptides were subsequently analyzed using a Thermo Scientific Q Exactive Hybrid Quadrupole-Orbitrap instrument (Thermo Scientific, USA) equipped with a Dionex U 3000 RSLCnano HPLC system. Mass spectrometric analyses were performed using a Thermo Scientific Q Exactive Hybrid Quadrupole-Orbitrap instrument mass spectrometer, with a nano-electrospray ionization source and fitted with a fused silica emitter tip (New Objective, Woburn, MA). Fractions were reconstituted in solvent A [water/acetonitrile (ACN) (98:2 v/v), 0.1% formic acid] and then injected into LC–nano ESI–MS/MS system. Solvent A consisted of water/ACN (98:2 v/v) with 0.1% formic acid for the high aqueous mobile phase. Samples were first trapped on an Acclaim PepMap 100 trap column (100 µm × 2 cm, nanoViper C18, 5 µm, 100 Å, Thermo Scientific, part number 164564) and washed for 6 min with 98% solvent A [water/ACN (98:2 v/v), 0.1% formic acid] at a flow rate of 4 µL/min, and then separated on an Acclaim PepMap 100 capillary column (75 µm × 15 cm, nanoViper C18, 3 µm, 100 Å, Thermo Scientific, part number 164568) at a flow rate of 300 nL/min. The LC gradient was run at 2–35% solvent B over 30 min, then from 35 to 90% over 10 min, followed by 90% solvent B for 5 min and, finally, 5% solvent B for 15 min. The resulting peptides were electrosprayed through a coated silica tip (FS360-20-10-N20-C12, PicoTip emitter, New Objective) at an ion spray voltage of 2000 eV. Mass data were acquired automatically using Proteome Discoverer 1.3 (Thermo Scientific, USA). MS/MS results are summarized in Supplementary File 2.

Results and discussion

Allelic analysis of LMW-GS using SDS-PAGE and 2-DGE

Allelic composition of LMW-GS in the Korean wheat variety Jokyoung is shown using SDS-PAGE (Fig. 1) and 2-DGE (Fig. 2). The glutenin fraction of Jokyoung was extracted and resolved in each experiment with standard wheat cultivars Cheyenne (*Glu-A3c*), Rescue (*Glu-B3h*), and Chinese Spring (*Glu-D3a*) as *Glu-3* allele controls. Figures 1 and 2 show Jokyoung protein bands corresponding to *Glu-A3c*, *Glu-B3h*, and *Glu-D3a*, the same allele as Cheyenne, Rescue, and Chinese Spring, respectively. This result demonstrated that LMW-GS alleles of the Korean wheat Jokyoung variety for *Glu-A3*, *Glu-B3*, and *Glu-D3* were *c*, *h*, and *a*, respectively, consistent with the results of Lee et al. (2017).

Identification of proteins using LC ESI–MS/MS

To identify LMW-GS proteins, 17 spots were excised from the 2-DGE (Fig. 3) and then in gel digested with chymotrypsin and subjected to MS/MS analysis. Chymotrypsin was used because it has a broader specificity and is better suited for ESI–MS/MS analysis than trypsin, and provided an efficient spot digestion with a greater number of medium-sized peptides from LMW-GS (Mamone et al. 2005; Vensel et al. 2014). The results of the MS/MS analysis showed that NCBI DB was limited to *Triticum* and the major amino acid sequence coverage of each spot was 34–65% (Table 1). The 13th spot in the 17 spots was identified as gamma gliadin (Fig. 3).

Spots 1–4 showed the best matches with CAB40553, when spectral data were searched against the NCBI data (Table 1). CAB40553 was best matched with the protein encoded by the *GluB3-33* haplotype, corresponding to the *Glu-B3h* allele [ACA63869 and EU369717 (Wang et al.

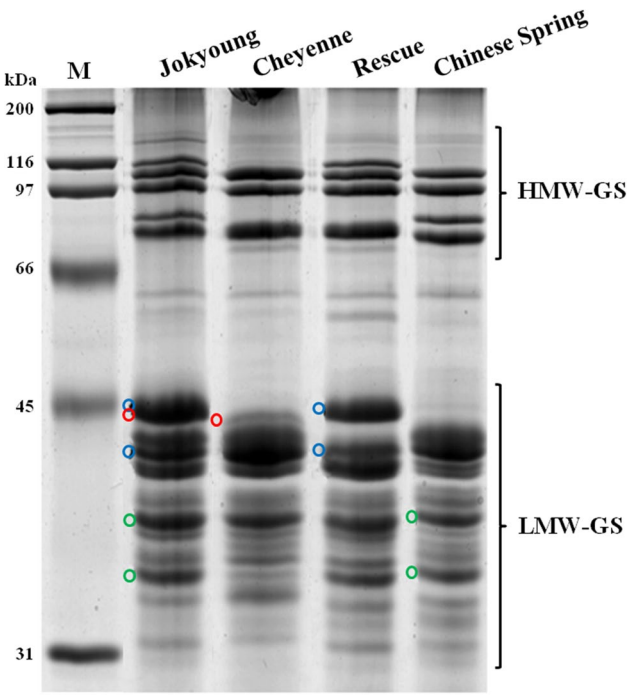

Fig. 1 Identification of LMW-GS allelic composition by SDS-PAGE in Jokyoung and standard cultivars Cheyenne, Rescue, and Chinese Spring. Red, blue, and green circles represent identified alleles of *Glu-A3c*, *Glu-B3h*, and *Glu-D3a* of LMW-GS, respectively

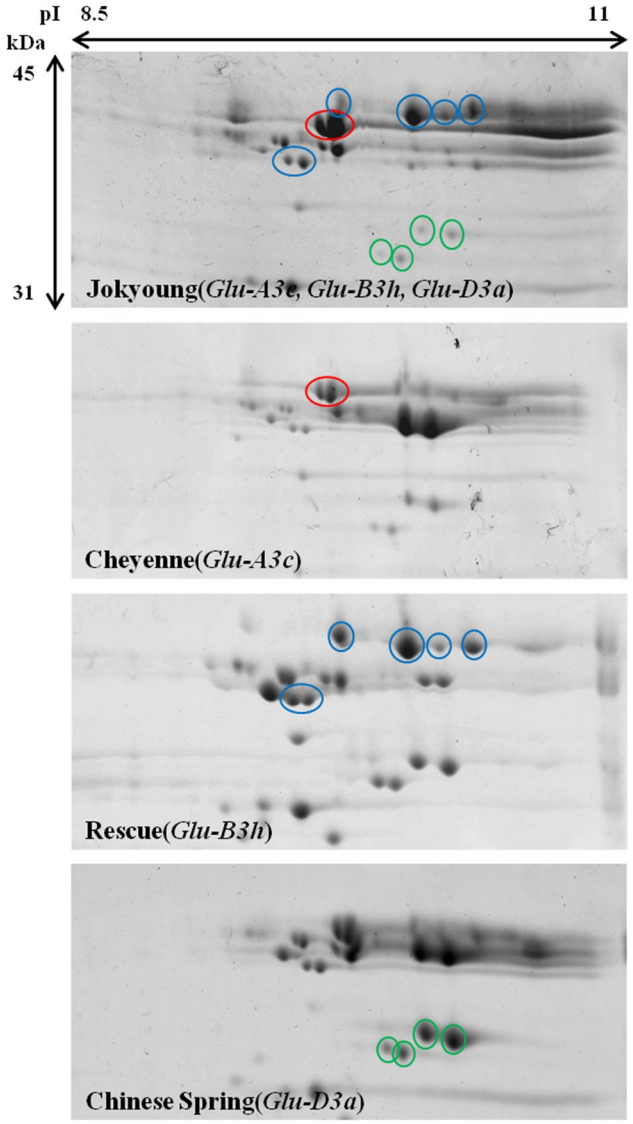

Fig. 2 Comparison of LMW-GS allelic variations in standard cultivars Cheyenne, Rescue, and Chinese Spring with Jokyoung by two-dimensional gel electrophoresis (2-DGE). Red, blue, and green circles represent identified alleles of *Glu-A3c*, *Glu-B3h*, and *Glu-D3a*, respectively

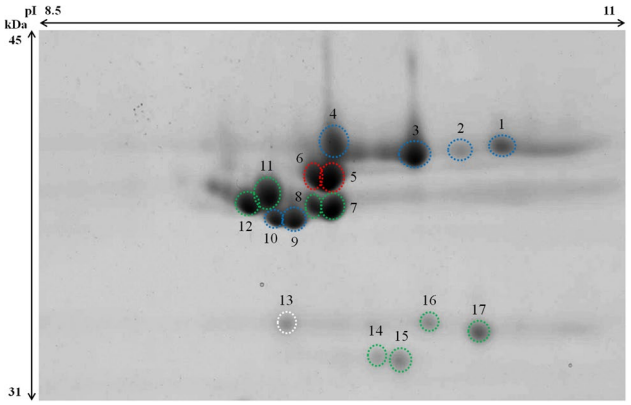

Fig. 3 2-DGE of LMW-GS fraction of Jokyoung. The individual spots analyzed by LC–ESI MS/MS are shown in red, blue, green, and white representing *Glu-A3c*, *Glu-B3h*, *Glu-D3a*, and gamma gliadin, respectively. Detailed MS/MS information is represented in Table 1

2009)] (Table 2). Spots 5 and 6 were best matched with AGK83179 (Table 1). AGK83179 corresponds to the protein encoded by the *GluA3-13* corresponding to the *Glu-A3c* allele [ACT98423 and FJ549930 (Wang et al. 2010)], but was missing a portion of the signal peptide with 16 amino acids (MKTFLVFALLALAAAS). Similarly, spots 5 and 6 were matched with the protein encoded by *LMW73* in Jokyoung with high sequence coverage (Table 2). Spots 7 and 8 matched with AEI00677, which is identical to the proteins encoded at *GluD3-31* [ABC84366 and DQ357057 (Zhao et al. 2006)] (Tables 1, 2). Likewise, spots 7 and 8 were matched to the protein encoded by *LMW48* in Jokyoung with high sequence coverage and a large number of unique peptides (Table 2). Spots 9 and 10 were matched with ACZ51337, showing high sequence similarity with proteins encoded by the *GluB3-43* haplotype (Tables 1, 2). The ACZ51337 had nearly the same amino acid sequence as the protein encoded by the *GluB3-43* haplotype (ACA63879 and EU369727 [Wang et al. 2009)]. These two spots were assigned to *LMW52* with the same sequence as *GluB3-43* (Table 2). Spot 11 was the best match with ACY08820, which was identical to the proteins encoded at *GluD3-5* [ABE77188 and DQ457419 (Zhao et al. 2007)] and spot 11 was similar to *LMW74* in Jokyoung (Table 2). Spot 12 best matched AEI00694, which is identical to the proteins encoded by *GluD3-11* [ABC84361 and DQ357052 (Zhao et al. 2006)], and this spot was assigned to *LMW71* with the same sequence as *GluD3-11* (Table 2). Spots 14 and 15 were matched to AAB48475, which is identical to the proteins encoded by *GluD3-6* [ABE77189 and DQ457420 (Zhao et al. 2007)] (Table 2). Finally, spots 16 and 17 were the best matches to AGU91700, which is identical to the proteins encoded by *GluD3-21* [ABC84363 and DQ357054 (Zhao et al. 2006)], and the two spots were similar to *LMW61* in Jokyoung.

Although 43 LMW-GS genes were identified in Jokyoung, only 17 spots of LMW-GS were resolved by 2-DGE. It is possible that each spot contained many proteins because the LMW-GS gene family shows a high similarity of sequences with their repeat sequence, consistent with the results of the Keumkang variety (Lee et al. 2016). Single proteins separated by 2-DGE often show multiple spots (Deng et al. 2012; Vensel et al. 2014). Deng et al. (2012) reported that individual spots of such charge trains on 2-D gels often represent the same protein and there are at least three theories about these charge trains: (1) PTM (post-translational modification); (2) artifacts of the analytical

Table 1 MS/MS analysis of individual spots separated by 2-DGE of LMW-GS fractions from Jokyoung using the NCBI non-redundant database limited to *Triticum*

Spot no.[a]	Predominant protein	Score	Coverage[b]	No. of proteins[c]	No. of unique peptides[d]	No. of peptides[e]	No. of PSMs[f]	No. of AAs[g]	MW	p*I*
1	CAB40553	118.97	36.31	145	2	11	45	369	42.2	8.35
2	CAB40553	171.36	36.31	136	2	12	60	369	42.2	8.35
3	CAB40553	257.35	36.31	145	1	12	92	369	42.2	8.35
4	CAB40553	175.61	34.15	133	1	10	64	369	42.2	8.35
5	AGK83179	262.23	65.00	113	0	26	90	360	41.6	8.15
6	AGK83179	199.15	65.00	85	1	25	69	360	41.6	8.15
7	AEI00677	251.81	43.79	194	3	19	92	354	40.0	8.15
8	AEI00677	165.68	43.79	228	3	17	60	354	40.0	8.15
9	ACZ51337	167.07	57.46	52	3	16	60	315	36.3	7.93
10	ACZ51337	156.98	57.78	53	4	17	54	315	36.3	7.93
11	ACY08820	148.26	59.13	70	2	17	49	345	39.6	8.32
12	AEI00694	164.26	50.29	107	1	15	55	350	39.8	8.15
13	AAF42989	50.45	35.06	85	1	9	19	308	35.2	7.91
14	AAB48475	58.00	35.57	38	6	10	22	298	33.9	8.68
15	AAB48475	136.55	38.26	36	4	13	52	298	33.9	8.68
16	AGU91700	145.86	41.51	125	5	10	51	265	30.5	8.43
17	AGU91700	235.50	41.51	188	0	11	84	265	30.5	8.43

[a]From Fig. 3

[b]Coverage = percentage of the protein sequence covered by identified peptides

[c]No. of proteins = the number of identified proteins in the protein group of a master protein

[d]No. of unique peptides = the number of peptide sequences unique to a protein group

[e]No. of peptides = the number of distinct peptide sequences in the protein group

[f]No. of PSMs = the total number of identified peptide sequences (peptide spectrum matches) for the protein, including those redundantly identified

[g]No. of AAs = the sequence length of the protein

procedure (e.g., 2-DE), including sample preparation; and (3) conformational changes or complex formation. Such charge trains are explained by isoform differences or by putative post-translational modifications, including phosphorylation, glycosylation, and others (Deng et al. 2012). Several reports have confirmed the high accuracy of spots by MS/MS analysis, as used in this study, but a further study in wheat proteomics will be required.

Reinterpretation of isolated 'Jokyoung' genes and gene haplotype

Forty-three LMW-GS genes were isolated using LMW-GS specific primers from Ikeda et al. (2002) and Lee et al. (2010). They divided 34 for LMW-m types, 8 for LMW-s type, and 1 for LMW-i type, according to the first amino acid sequence of the N-terminus of the 43 isolated genes. Phylogenetic analysis was performed to identify gene haplotypes in Jokyoung using the MEGA 6 software with 9 *Glu-A3*, 17 *Glu-B3*, and 10 *Glu-D3* LMW-GS previously reported gene haplotypes (Wang et al. 2009, 2010; Zhao et al. 2006, 2007). Allelic variants of LMW-GS genes from Jokyoung generally clustered into individual groups. Eight of the 43 genes (*LMW3*, *LMW13*, *LMW14*, *LMW18*, *LMW19*, *LMW25*, *LMW46*, and *LMW66*) were excluded from the analysis because no gene haplotype was identified, and 8 genes were estimated to be sequencing errors. The 35 genes were classified into seven haplotypes, including *GluA3-13*, *GluB3-43*, *GluD3-11*, *GluD3-21*, *GluD3-31*, *GluD3-42*, and *GluD3-5* (Fig. 4).

Accordingly, one gene was identified as an *LMW73* gene from *Glu-A3* locus, and the amino acid sequence started with ISQQQQQ-, i-type LMW-GS. The *LMW73* gene showed a high level of sequence similarity with the *GluA3-13* haplotype [FJ549930 (Wang et al. 2010)]. When compared with the *GluA3-13* gene haplotype, the two amino acid sequences changed due to two single nucleotide polymorphisms (SNPs) (Supplementary Figs. 1 and 2). A previous study of *Glu-A3c* reported one active gene, *GluA3-13*, and two pseudogenes, *GluA3-22* and *GluA3-33*, in Aroona NILs (Wang et al.

Table 2 Relationship between LMW-GS spots and their gene haplotypes in Jokyoung

Spot no.[a]	Allele or loci	MS/MS identification	Gene haplotype	Putative corresponding genes[b] (reference accession number)	Reference
1	*Glu-B3h*	CAB40553	*GluB3-33*	*Ni*[c]	Wang et al. (2009)
2	*Glu-B3h*	CAB40553	*GluB3-33*	*Ni*[c]	Wang et al. (2009)
3	*Glu-B3h*	CAB40553	*GluB3-33*	*Ni*[c]	Wang et al. (2009)
4	*Glu-B3h*	CAB40553	*GluB3-33*	*Ni*[c]	Wang et al. (2009)
5	*Glu-A3c*	AGK83179	*GluA3-13*	*LMW73* (FJ549930)	Wang et al. (2010)
6	*Glu-A3c*	AGK83179	*GluA3-13*	*LMW73* (FJ549930)	Wang et al. (2010)
7	*Glu-D3a*	AEI00677	*GluD3-31*	*LMW48* (DQ357057)	Zhao et al. (2006)
8	*Glu-D3a*	AEI00677	*GluD3-31*	*LMW48* (DQ357057)	Zhao et al. (2006)
9	*Glu-B3h*	ACZ51337	*GluB3-43*	*LMW52* (EU369727)	Wang et al. (2009)
10	*Glu-B3h*	ACZ51337	*GluB3-43*	*LMW52* (EU369727)	Wang et al. (2009)
11	*Glu-D3a*	ACY08820	*GluD3-5*	*LMW74* (DQ457419)	Zhao et al. (2007)
12	*Glu-D3a*	AEI00694	*GluD3-11*	*LMW71* (DQ357052)	Zhao et al. (2006)
14	*Glu-D3a*	AAB48475	*GluD3-6*	*Ni*[c]	Zhao et al. (2007)
15	*Glu-D3a*	AAB48475	*GluD3-6*	*Ni*[c]	Zhao et al. (2007)
16	*Glu-D3a*	AGU91700	*GluD3-21*	*LMW61* (DQ357054)	Zhao et al. (2006)
17	*Glu-D3a*	AGU91700	*GluD3-21*	*LMW61* (DQ357054)	Zhao et al. (2006)

Reference for gene haplotype used Wang et al. (2010) for Glu-A3c, Wang et al. (2009) for *Glu-B3h*, and Zhao et al. (2006, 2007) for *Glu-D3a*. Putative corresponding gene sequences had been submitted to Genbank, viz. *LMW73* (HQ619933), *LMW48* (HQ619912), *LMW52* (HQ619905), *LMW74* (HQ619932), *LMW71* (HQ619926), and *LMW61* (HQ619923) (Lee et al. 2010)

[a]From Fig. 3

[b]Putative corresponding genes having high sequence similarity with each gene haplotype

[c]*Ni* not isolated in this study

2010). Genes corresponding to *GluA3-22* and *GluA3-33* were not isolated in this study.

Nine genes were isolated in the *Glu-B3* locus, but only the *GluB3-43* gene haplotype was identified and the *GluB3-33* gene haplotype was not identified. Nine genes showed a high level of sequence similarity with the *GluB3-43* gene haplotype [EU369727 (Zhao et al. 2007)]. The Jokyoung genes similar to *GluB3-43* were *LMW51*, *LMW52*, *LMW9*, *LMW36*, *LMW8*, *LMW41*, *LMW43*, *LMW54*, and *LMW5*. The *LMW52* was the same as *GluB3-43*, while others had one or two amino acid substitutions in the corresponding proteins due to one or three SNPs (Supplementary Figs. 3 and 4).

Appearing in the *Glu-D3* locus, 25 genes were identified, and five gene haplotypes were distinguished, which were *GluD3-11*, *GluD3-21*, *GluD3-31*, *GluD3-42*, and *GluD3-5*, respectively. The *GluD3-6* haplotype was not identified. The highest number of genes were identified for *GluD3-11* and *GluD3-31*, with nine and eight genes, respectively. Regarding the *GluD3-11* gene haplotype, the Jokyoung genes were *LMW28*, *LMW68*, *LMW 59*, *LMW 60*, *LMW 67*, *LMW 70*, *LMW 69*, *LMW 71*, and *LMW 72*. The *LMW70*, *LMW67*, and *LMW69* showed two amino acid sequences changed due to the change of the three SNPs. The amino acid sequence of *LMW71* was identical to *GluD3-11*, while *LMW70* had one SNP, but there was no change in the amino acid sequence. The *LMW60*, *LMW67*, and *LMW69* showed one amino acid substitution due to one SNP. Additionally, *LMW68* and *LMW59* changed two amino acid sequences due to two SNPs, and LMW28 and *LMW72* changed three amino acid sequences due to four and five SNPs, respectively (Supplementary Figs. 5 and 6). Five genes, *LMW61*, *LMW63*, *LMW64*, *LMW65*, and *LMW62*, were like the *GluD3-21* gene haplotype. These five genes had one or three SNPs, resulting in one or two amino acid substitutions (Supplementary Figs. 7 and 8). The *GluD3-31* gene haplotype was similar to *LMW57*, *LMW48*, *LMW35*, *LMW42*, *LMW34*, *LMW55*, *LMW49*, and *LMW31*. Among the genes, *LMW35* had four amino acid sequences changed due to four SNPs. The *LMW57* showed three amino acid substitutions due to three SNPs (Supplementary Figs. 9 and 10). The remaining four *LMW34*, *LMW55*, *LMW49*, and *LMW31* genes were exactly similar to the *GluD3-31* gene haplotype. The *GluD3-42* gene haplotype was similar to *LMW23* and *LMW24*. *LMW23* had only one SNP without any change in the amino acid sequence, but *LMW24* had three amino acid sequences changed by three SNPs (Supplementary Figs. 11 and 12). Finally, *LMW74* was found to be similar to the *GluD3-5* gene haplotype and the corresponding gene showed seven

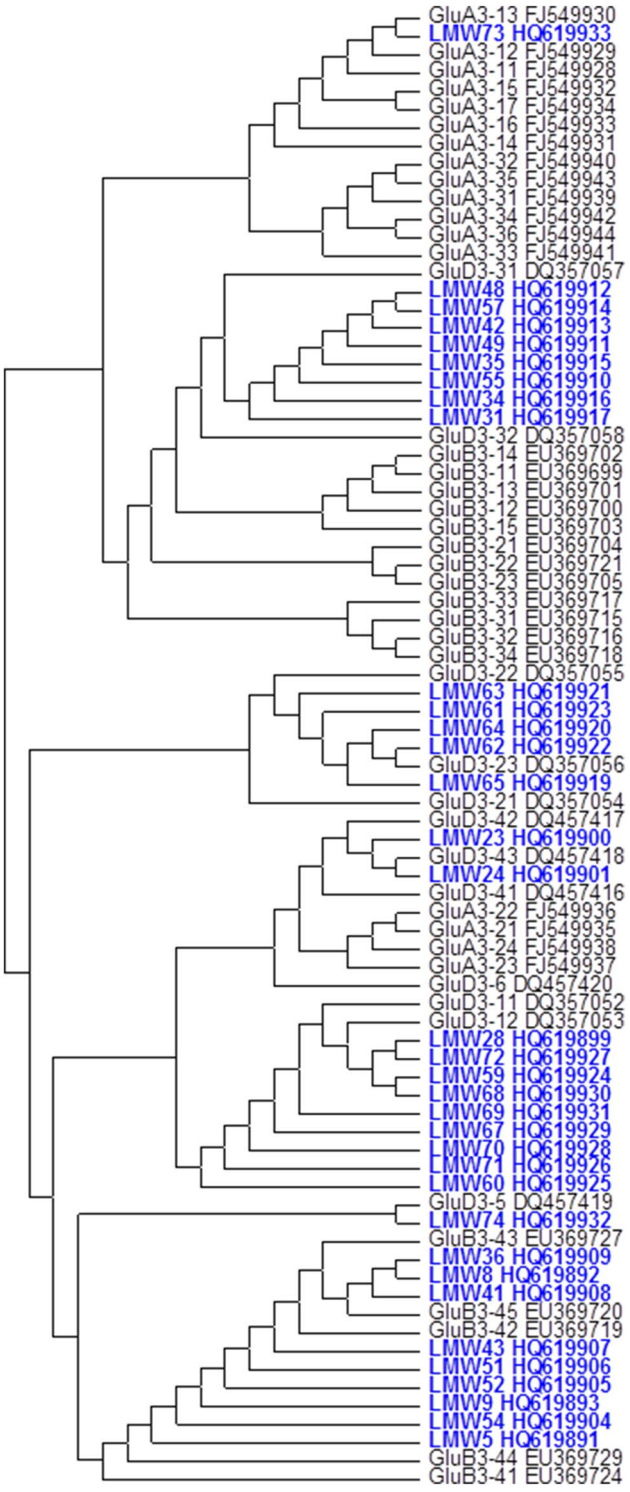

Fig. 4 Phylogenic tree of LMW-GS genes isolated in a previous study by Lee et al. (2010) (blue) and LMW-GS gene haplotypes at *Glu-A3*, *Glu-B3*, and *Glu-D3* loci (black) described previously (Wang et al. 2009, 2010; Zhao et al. 2006, 2007)

SNPs, but they showed four substitutions in the amino acid sequence (Supplementary Figs. 13 and 14).

To summarize, gene haplotypes for each locus were determined in Jokyoung. The *Glu-A3* locus was a *GluA3-13* haplotype, while in the *Glu-B3* locus was *GluB3-43* and in the *Glu-D3* locus were *GluD3-11*, *GluD3-21*, *GluD3-31*, *GluD3-42*, and *GluD3-5*, respectively.

Relationship between the LMW-GS proteins and the genes

Aroona NILs were used to analyze the effect of LMW-GS gene alleles on wheat processing qualities, which showed superior bread-making quality in *Glu-A3d*, *Glu-B3b*, *Glu-B3g*, and *Glu-B3i* loci (Zhang et al. 2012). Identification of all LMW-GS genes, their loci, and their corresponding genes in a single cultivar, however, is difficult to confirm from the copy number (Cassidy et al. 1998; Dong et al. 2010; Huang and Cloutier 2008). Functional analysis of individual LMW-GS genes, therefore, is needed. During this study, genes isolated from the Jokyoung cultivar (Lee et al. 2010) were reinterpreted to understand the correlation between the LMW-GS genes and their protein products by MS/MS analysis, based on 2-DGE.

Previously, 43 LMW-GS genes were identified using the Ikeda classification method (Ikeda et al. 2002) in the Jokyoung cultivar, but in this study 35 genes were divided into seven gene haplotypes. Consequently, only the 35 genes from among 43 genes isolated from the Jokyoung cultivar were analyzed These 35 genes were classified into seven haplotypes including *GluA3-13*, *GluB3-43*, *GluD3-11*, *GluD3-21*, *GluD3-31*, *GluD3-42*, and *GluD3-5* (Fig. 4). These analysis results was then confirmed and the genes linked to the best matched proteins at the haplotype level, by combining the results of both MS/MS data and phylogenetic analysis.

Table 2 summarizes the linkage of both genes and proteins of LMW-GS and the individual 2-DGE spots from a glutenin protein fraction that were well matched with their gene haplotypes. Among the 17 protein spots, 2 (putative corresponding gene; *LMW73*) were associated with the *Glu-A3* locus, 2 (putative corresponding gene; *LMW52*) with the *Glu-B3* locus, and 8 (putative corresponding genes; *LMW48*, *LMW61*, *LMW71*, and *LMW74*) with the *Glu-D3* locus. Previous reports have shown that the *Glu-D3* locus encoded the most abundant LMW-GS (Ikeda et al. 2006; Lee et al. 2016; Zhang et al. 2012), consistent with the results in this study. Regarding Aroona NILs, Wang et al. (Wang et al. 2010) reported one active gene, *GluA3-13*. Jokyoung also showed only one *GluA3-13* haplotype at the *Glu-A3c* locus in this study (Fig. 4 and Table 2). Regarding the *Glu-B3* locus, only one *GluB3-43* haplotype was identified in Jokyoung, designated as *LMW52*. These results suggest that the haplotype

genes of *GluB3-33* were not isolated in the Jokyoung cultivar, unlike the results of previous reports (Lee et al. 2016). Zhao et al. (2006, 2007) also reported *Glu-D3a* allele haplotypes, among them this study showed similar results for *GluD3-11*, *GluD3-21*, *GluD3-31*, *GluD3-42*, and *GluD3-5* (Fig. 4 and Table 2). To conclude, these results suggest that the links between individual LMW-GS proteins and their genes are relatively conserved between different wheat varieties. Additionally, reinterpretation of LMW-GS genes and their protein products using gene haplotypes can be more precisely correlated. These approaches will be useful tools to distinguish the individual LMW-GS genes with their protein products, and give insight into the research of the function of specific LMW-GS proteins and the allergenic potential of specific gluten proteins.

Acknowledgements This work was supported by grants from the National Institute of Agricultural Science (RDA PJ012458) and the Next-Generation BioGreen 21 Program (RDA PJ011021 and RDA PJ011057), Republic of Korea.

Author contributions JYL designed and conducted the experiments and wrote the manuscript; HRB and JSK conducted the experiments and wrote the manuscript; YRJ conducted the experiments and contributed to the writing of the manuscript; and SHL, CKK, and CKL contributed to the experimental design and writing of the manuscript.

Compliance with ethical standards

Conflict of interest The authors declare no conflict of interest.

References

Cassidy BG, Dvorak J, Anderson OD (1998) The wheat low-molecular-weight glutenin genes: characterization of six new genes and progress in understanding gene family structure. Theor Appl Genet 96:743–750. https://doi.org/10.1007/s001220050797

Cloutier S, Rampitsch C, Penner GA, Lukow OM (2001) Cloning and expression of a LMW-i glutenin gene. J Cereal Sci 33:143–154. https://doi.org/10.1006/jcrs.2000.0359

D'Ovidio R, Masci S (2004) The low-molecular-weight glutenin subunits of wheat gluten. J Cereal Sci 39:321–339. https://doi.org/10.1016/j.jcs.2003.12.002

Deng X, Hahne T, Schröder S, Redweik S, Nebija D, Schmidt H, Janssen O, Lachmann B, Wätzig H (2012) The challenge to quantify proteins with charge trains due to isoforms or conformers. Electrophoresis 33:263–269. https://doi.org/10.1002/elps.201100321

Dong LL, Zhang XF, Liu DC, 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. https://doi.org/10.1371/journal.pone.0013548

Gupta RB, Shepherd KW (1990) Two-step one-dimensional SDSPAGE analysis of LMW subunits of glutenin. I. Variation and genetic control of the subunits in hexaploid wheats. Theor Appl Genet 80:65–74. https://doi.org/10.1007/BF00224017

Huang XQ, Cloutier S (2008) Molecular characterization and genomic organization of low molecular weight glutenin subunit genes at the *Glu-3* loci in hexaploid wheat (*Triticum aestivum* L.). Theor Appl Genet 116:953–966. https://doi.org/10.1007/s00122-008-0727-1

Ikeda TM, Nagamine T, Fukuoka H, Yano H (2002) Identification of new low-molecular-weight glutenin subunit genes in wheat. Theor Appl Genet 104:680–687. https://doi.org/10.1007/s001220100756

Ikeda TM, Araki E, Fujita Y, Yano H (2006) Characterization of low molecular-weight glutenin subunit genes and their protein products in common wheats. Theor Appl Genet 112:327–334. https://doi.org/10.1007/s00122-005-0131-z

Kasarda DD, Tao HP, Evans PK, Adalsteins AE, Yuen SW (1988) Sequencing of protein from a single spot of a 2-D gel pattern: N-terminal sequence of a major wheat LMW-glutenin subunit. J Exp Bot 39:899–906

Lee JY, Kim YT, Kim BM, Lee JH, Lim SH, Ha SH, Ahn SN, Nam MH, Kim YM (2010) Cloning of low-molecular-weight glutenin subunit genes and identification of their protein products in common wheat (*Triticum aestivum* L.). Korean J Breed Sci 42:547–554

Lee JY, Beom HR, Altenbach SB, Lim SH, Kim YT, Kang CS, Yoon UH, Gupta R, Kim ST, Ahn SN, Kim YM (2016) Comprehensive identification of LMW-GS genes and their protein products in a common wheat variety. Funct Integr Genom 16:269–279. https://doi.org/10.1007/s10142-016-0482-3

Lee JY, Jang YR, Beom HR, Altenbach SB, Lim SH, Lee CK (2017) Allelic analysis of low molecular weight glutenin subunits using 2-DGE in Korean wheat cultivars. Breed Sci 67:398–407. https://doi.org/10.1270/jsbbs.16106

Lew EJL, Kuzmicky DD, Kasarda DD (1992) 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 69:508–515

Liu L, Ikeda TM, Branlard G, Peña RJ, Rogers WJ, Lerner SE, Kolman MA, Xia XC, Wang LH, Ma WJ, 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. https://doi.org/10.1186/1471-2229-10-124

Mamone G, Addeo F, Chianese L, Di Luccia A, De Martino A, Nappo A, Formisano A, De Vivo P, Ferranti P (2005) Characterization of wheat gliadin proteins by combined two-dimensional gel electrophoresis and tandem mass spectrometry. Proteomics 5:2859–2865. https://doi.org/10.1002/pmic.200401168

Masci S, Lew EJL, Lafiandra D, Porceddu E, Kasarda DD (1995) Characterization of low molecular weight glutenin subunits in durum wheat by reversed-phase high-performance liquid chromatography and N-terminal sequencing. Cereal Chem 72:100–104

Nielsen HC, Beckwith AC, Wall JS (1968) Effect of disulfide-bond cleavage on wheat gliadin fractions obtained by gel filtration. Cereal Chem 45:37–47

Payne PI (1987) Genetics of wheat storage proteins and the effect of allelic variation on bread-making quality. Ann Rev Plant Physiol 38:141–153. https://doi.org/10.1146/annurev.pp.38.060187.001041

Pitts EG, Rafalski JA, Hedgcoth C (1988) Nucleotide sequence and encoded amino acid sequence of a genomic gene region for a low molecular weight glutenin. Nucleic Acids Res 16:11376

Sabelli PA, Shewry PR (1991) Characterization and organization of gene families at the Gli-1 loci of bread and durum wheats by restriction fragment analysis. Theor Appl Genet 83:209–216. https://doi.org/10.1007/BF00226253

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425. https://doi.org/10.1093/oxfordjournals.molbev.a040454

Singh NK, Shepherd KW, Cornish GB (1991) A simplified SDS-PAGE procedure for separating LMW subunits of glutenin. J Cereal Sci 14:203–208. https://doi.org/10.1016/S0733-5210(09)80039-8

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2724–2729. https://doi.org/10.1093/molbev/mst197

Tao HP, Kasarda DD (1989) Two-dimensional gel mapping and N-terminal sequencing of LMW-glutenin subunits. J Exp Bot 40:1015–1020. https://doi.org/10.1093/jxb/40.9.1015

Vensel WH, Tanaka CK, Altenbach SB (2014) Protein composition of wheat gluten polymer fractions determined by quantitative two-dimensional gel electrophoresis and tandem mass spectrometry. Proteome Sci 12:8. https://doi.org/10.1186/1477-5956-12-8

Wang LH, Zhao XL, He ZH, Ma W, Appels R, Peña 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. https://doi.org/10.1007/s00122-008-0918-9

Wang LH, Li G, Pena RJ, Xia X, He Z (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. https://doi.org/10.1016/j.jcs.2010.01.005

Wieser H (2007) Chemistry of gluten proteins. Food Microbiol 24:115–119. https://doi.org/10.1016/j.fm.2006.07.004

Zhang W, Gianibelli MC, Rampling L, Gale KR (2004) Characterization 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. https://doi.org/10.1007/s00122-003-1558-8

Zhang XF, Jin H, Zhang Y, Liu DC, Li GY, Xia XC, 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. https://doi.org/10.1186/1471-2229-12-243

Zhang XF, Liu DC, Zhang JH, Jiang W, Luo GB, Yang WL, Sun J, Tong Y, 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. https://doi.org/10.1093/jxb/ert070

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. https://doi.org/10.1007/s00122-006-0379-y

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. https://doi.org/10.1007/s00122-006-0445-5