Journal of Cereal Science 65 (2015) 9–18

Contents lists available at ScienceDirect

Journal of Cereal Science

journal homepage: www.elsevier.com/locate/jcs

Characterization of novel high-molecular-weight glutenin subunits and their coding sequences in *Aegilops markgrafii*

Xin Li [a], Dong-cheng Liu [a], Jia-zhu Sun [a], Wen-long Yang [a], Xiao-li Guo [c], Dao-wen Wang [a, b], Ai-min Zhang [a, b, *]

[a] State Key Laboratory of Plant Cell and Chromosome Engineering, National Center for Plant Gene Research, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, No. 1 West Beichen Road, Chaoyang District, Beijing 100101, PR China
[b] The Collaborative Innovation Center for Grain Crops in Henan, Henan Agricultural University, 63 Nongye Road, Zhengzhou 450002, Henan Province, PR China
[c] College of Biological Sciences, China Agricultural University, No. 2 West Yuanmingyuan Road, Haidian District, Beijing 100193, PR China

ARTICLE INFO

Article history:
Received 4 January 2015
Received in revised form
27 April 2015
Accepted 4 May 2015
Available online 12 June 2015

Keywords:
Aegilops markgrafii
HMW-GS
Glutenin
Wheat flour quality

ABSTRACT

Wheat's wild relatives are potential resources for the genetic improvement of common wheat quality. Novel high-molecular-weight glutenin subunits (HMW-GS) from *Aegilops markgrafii* were identified using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS). Three complete coding sequences of novel HMW-GS were cloned and sequenced. The lengths of these sequences, encoding 881, 847, and 644 amino acid residues, were 2649, 2571, and 1938 bp (designated as Cx1, Cx3, and Cy3), respectively. Sequence comparison showed that added SNPs and InDels were responsible for major variations in the new genes. Analysis of the deduced amino acid sequences revealed that the primary structures of the x- and y-type HMW-GS of *Ae. markgrafii* were similar to those of previously published HMW-GS. In particular, all new Cx subunits had higher molecular weights than the previously reported HMW-GS in *Ae. markgrafii*, which have potential value in improving bread quality. Phylogenetic analysis supported the information on the origin of the B genome in wheat. In conclusion, our study not only broadens our understanding of the genetics and phylogenetic knowledge of the HMW-GS protein family but also provides new insight into wheat breeding for end-use quality improvement.

1. Introduction

High-molecular-weight glutenin subunits (HMW-GS), low-molecular-weight glutenin subunits, and gliadins are the three categories of polypeptide that comprise the seed storage proteins of common wheat (*Triticum aestivum* L.). Dough elasticity and bread-making quality of wheat are determined mainly by the composition and quantity of gluten-forming storage proteins, and, in particular, the HMW-GS (Payne et al., 1987; Shewry et al., 1995).

HMW-GS are encoded by the *Glu-1* loci on the long arms of homoeologous group 1 chromosomes. There are two closely linked genes (*Glu-1-1* and *Glu-1-2*) at three loci (*Glu-A1*, *Glu-B1*, and *Glu-D1*) in common wheat (AABBDD, 2n = 6x = 42), which encode one larger x-type (80–88 kDa) and one smaller y-type (67–73 kDa) subunits, respectively (Mackie et al., 1996). Allelic variation at the *Glu-1* loci in modern cultivars of common wheat is limited. A broader gene pool of modern wheat cultivars and improved end-use quality of wheat flour are highly desirable in wheat-breeding programs.

In wheat and its wild relative species, HMW-GS share a primary structure that includes the conserved signal peptide, N- and C-terminal domains, and the central repetitive region composed of tri-, hexa- and nona-peptide motifs. Generally, the N- and C-terminal domains and the number of cysteine residues are highly conserved in HMW-GS, with four cysteine residues present in most x-type and seven in y-type subunits (Shewry et al., 1992). The central repeated domain adopts a β-spiral structure that confers elasticity to the protein molecule. The β-spirals are assembled into gluten polymers by inter-molecular disulfide bonds between the

Abbreviations: HMW-GS, high-molecular-weight glutenin subunits; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; MALDI-TOF-MS, matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry.

* Corresponding author.
E-mail address: amzhang@genetics.ac.cn (A.-m. Zhang).

http://dx.doi.org/10.1016/j.jcs.2015.05.014
0733-5210/

cysteine residues in the α-helical domains near the N- and C- terminals (Chen et al., 2012; Shewry and Tatham, 1997). Thus, the number and position of cysteine residues are important features of HMW-GS, as they are involved in the formation of inter- or intramolecular disulfide bonds. Compared with wheat, several HMW-GS from related species possess an unusual number and/or location of cysteine residues, and accordingly have an important effect on the strength and elastic properties of the gluten complex (Ma et al., 2013). Additionally, Belton (1999) proposed that inter-chain hydrogen bonds, formed, in particular, between glutamine residues present in the repetitive domains, are also important in conferring elasticity.

As a genus related to wheat, *Aegilops* possesses an extensive range of HMW-GS variation that is expected to provide potential genetic resources for improving wheat quality. *Aegilops* contains more than 20 diploid, tetraploid, and hexaploid species with unique diploid genomes, classified as S, C, N, U, M, D, and T types (Lilienfeld and Kihara, 1951). Numerous HMW-GS have been identified and characterized from *Ae. tauschii* (D^tD^t) (Wang et al., 2012), *Ae. umbellulata* (UU) (Rodríguez-Quijano et al., 2001), *Ae. comosa* (MM) (Rodríguez-Quijano et al., 2001), *Ae. markgrafii* (CC) (Liu et al., 2003; Rodríguez-Quijano et al., 2001; Wan et al., 2000), *Ae. searsii* (S^sS^s) (Sun et al., 2006), *Ae. cylindrica* (CCDD) (Wan et al., 2002), *Ae. speltoides* (SS), and *Ae. kotschyi* (UUSS) (Ma et al., 2013). These HMW-GS could serve as valuable new resources for wheat quality improvement. Recently, a wheat cultivar "Chinese Spring" chromosome substitution line CS-$1S^1$ (1B), in which the 1B chromosome was substituted by $1S^l$ from *Ae. longissima*, was developed and found to possess superior dough and bread-making quality (Wang et al., 2013).

Ae. markgrafii (2n = 2x = 14, CC) is one of the diploid species whose genome is involved in forming many polyploid *Aegilops* species (Friebe et al., 1992). In this study, we examined 12 *Ae. markgrafii* accessions collected from Greece, Turkey, and Iran. A series of novel HMW-GS and their coding sequences were identified and characterized.

The objectives of this study were: 1) to isolate novel HMW-GS genes from *Ae. markgrafii* and compare their primary structures with other known HMW-GS in wheat, 2) to investigate their evolutionary relationship among available HMW-GS, and 3) to explore the genetic potential of transformation of HMW-GS for improving the bread-making qualities of wheat. The results would not only broaden the understanding of the genetics and phylogenetic knowledge of this HMW-GS protein family but also provide opportunities to improve the end-use quality in wheat-breeding programs.

2. Materials and methods

2.1. Plant materials

Aegilops markgrafii (Greuter) Hammer (syn. *Ae. caudata* L., 2n = 2x = 14, CC) accessions PI203431, PI254863, PI263554, PI298889, PI542198, PI542200, PI542206, PI542209, PI551121, PI551124, PI551129, and PI551139 were kindly provided by Dr. Yiwen Li, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, China. The wheat cultivars 'Cheyenne' (1, 7 + 9, 5 + 10) and 'Jimai20' (1, 13 + 16, 5 + 10) were used as standards for HMW-GS identification.

2.2. Glutenin preparation and SDS-PAGE analysis

The glutenin proteins were extracted from mature seeds of *Ae. markgrafii* and common wheat according to the method described by Singh et al. (1991) with slight modifications. Three seeds of each *Ae. markgrafii* accession were crushed into a fine powder. Glutenins were extracted with 70% isopropanol twice at 65 °C for 40 min. After centrifuging at 12,000 g for 5 min, the pellets were used to extract the HMW-GS using a buffer containing 50% isopropanol, 0.08 M Tris–HCl (pH 8.0) and 1% DTT. After a 5 min centrifugation, the proteins in the supernatant were alkylated with 1.4% 4-vinylpyridine at 65 °C for 30 min. After centrifuging at 12,000 g for 5 min, the supernatant was transferred into a sample buffer containing 4% SDS, 25% glycerol, 0.04% bromophenol blue and 1 M Tris–HCl (pH 6.8). Glutenins were separated in a discontinuous 1-mm SDS-PAGE gel. The concentration of the separating gel was 12% and that of the stacking gel was 5%. The pH for the separating gel was 8.8. The gel was run for 13 h at 12 mA. The SDS-PAGE gel was stained with Coomassie Brilliant Blue R-250 and then scanned to detect protein bands.

2.3. MALDI-TOF-MS analysis

All the Cx and Cy subunits bands of *Ae. markgrafii* were cut from the SDS-PAGE gel, followed by in-gel digestion with trypsin. The samples of the digested protein were sent to Public Technology Service Centre (IMCAS, Beijing), for MALDI-TOF-MS analysis following the standard protocol.

2.4. DNA extraction and PCR amplification

Genomic DNA was extracted from young leaves following a standard procedure with the CTAB method. The PCR primers used to amplify the full coding sequences of HMW-GS genes were published by Liu et al. (2008) and synthesized by Life Technologies (Beijing). The sequences of two primers were HMW-F: 5′-ATGGCTAAGCGGYTRGTCCTCTTTG-3′ and HMW-R: 5′-CTATCACTGGCTRGCCGACAATGCG-3′. PCR amplification was performed in a total volume of 20 μL containing 1.5-μL DNA (50–100 ng/μL), 10-μL 2 × GC buffer I ($MgCl_2$ plus, TaKaRa), 1.5 μL of each primer (5 μM), 0.8 μL dNTPs (10 mM), 0.2-μL *LA Taq* polymerase (5 U/μL, TaKaRa), and 4.5-μL ddH_2O. The reaction was carried out using the following PCR conditions: pre-denaturation at 94 °C for 5 min, followed by 35 cycles of 94 °C for 1 min, 68 °C for 4 min and a final extension step at 68 °C for 7 min.

2.5. Isolation and confirmation of complete coding sequences

PCR products were separated on a 1% agarose gel. DNA fragments of expected sizes were cut from agarose gels and then ligated into the pGEM-T vector (Promega). The ligation mixtures were transformed into competent cells of *Escherichia coli* DH10B following standard procedures (Promega). Positive clones were selected using blue/white color screening and were sequenced by SinaGenoMax Co., Ltd., Beijing. To obtain the complete coding sequences, specific primers were designed (Table 1). The identification and sequencing of the sub-clones were carried out using the procedure described above. The complete coding sequences were determined using the DNAMAN 8.0 software (Lynnonon Biosoft,

Table 1
Primers for sub-sequencing in complete coding sequences assembly.

Primer pair	Sequence (5′-3′)	Coding sequence assembled
P1	GGATACTACCCAACTTCTCTGC TTGCCCTTGTCCTGACTC	Cx1
P2	GCAGGTTTCATACTATCCAGG TGCCCTTGTCCTGACTCC	Cx3
P3	CCAGAACAAGGGCAACAA TGCCCTTGTCCTGACTCC	Cx3