

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.)

Xiaofei Zhang · Dongcheng Liu · Wenlong Yang · Kunfan Liu · Jiazhu Sun · Xiaoli Guo · Yiwen Li · Daowen Wang · Hongqing Ling · Aimin Zhang

Received: 4 August 2010 / Accepted: 5 February 2011 / Published online: 23 February 2011
© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract Low-molecular-weight glutenin subunits (LMW-GSs) play an important role in determining the bread-making quality of bread wheat. However, LMW-GSs display high polymorphic protein complexes encoded by multiple genes, and elucidating the complex LMW-GS gene family in bread wheat remains challenging. In the present study, using conventional polymerase chain reaction (PCR) with conserved primers and high-resolution capillary electrophoresis, we developed a new molecular marker system for identifying LMW-GS gene family members. Based on sequence alignment of 13 LMW-GS genes previously identified in the Chinese bread wheat variety Xiaoyan 54 and other genes available in GenBank, PCR primers were developed and assigned to conserved sequences spanning the length polymorphism regions of LMW-GS genes. After PCR amplification, 17 DNA fragments in Xiaoyan 54 were detected using capillary electrophoresis. In total, 13

fragments were identical to previously identified LMW-GS genes, and the other 4 were derived from unique LMW-GS genes by sequencing. This marker system was also used to identify LMW-GS genes in Chinese Spring and its group 1 nulli-tetrasomic lines. Among the 17 detected DNA fragments, 4 were located on chromosome 1A, 5 on 1B, and 8 on 1D. The results suggest that this marker system is useful for large-scale identification of LMW-GS genes in bread wheat varieties, and for the selection of desirable LMW-GS genes to improve the bread-making quality in wheat molecular breeding programmes.

Introduction

Bread wheat (*Triticum aestivum* L., AABBDD, $2n = 6x = 42$) can be processed into bread, pasta, noodles and other food products due to its unique dough viscoelastic properties. These biochemical properties are mainly conferred by seed storage proteins (Shewry et al. 1995; Weegels et al. 1996) called prolamins, comprising two major groups—gliadins and glutenins—based on their different solubilities in alcohol–water mixtures (Osborne 1924). Wheat glutenins are classified as high-molecular-weight glutenin subunits (HMW-GSs) and low-molecular-weight glutenin subunits (LMW-GSs) according to their mobility during sodium dodecyl sulphate polyacrylamide-gel electrophoresis (SDS-PAGE; Payne 1987; Gupta and Shepherd 1990). These glutenin subunits are polymerised by intermolecular disulphide bonds, which contribute to the viscoelastic and extensible properties of dough and consequently influence the quality of wheat food products (Shewry et al. 1995). Thus, the characterisation of specific HMW-GS and LMW-GS genes encoding these glutenin subunits was highlighted in the improvement of wheat

Communicated by J. Snape.

X. Zhang and D. Liu contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-011-1550-7) contains supplementary material, which is available to authorized users.

X. Zhang · D. Liu · W. Yang · K. Liu · J. Sun · Y. Li · D. Wang · H. Ling · A. Zhang (✉)
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, 1 West Beichen Road, Chaoyang District, Beijing 100101, China
e-mail: amzhang@genetics.ac.cn

X. Guo
College of Biology, China Agricultural University,
Yuanmingyuan Xi Lu 1, Haidian District, Beijing 100193, China

quality (Payne 1987; Gupta et al. 1989, 1991; Metakovsky et al. 1990). HMW-GSs have been extensively investigated because of their importance in dough strength and their relative simplicity at the DNA and protein levels (Shewry et al. 1995). Unlike HMW-GSs comprising three to five members in a bread wheat variety, LMW-GSs display high polymorphic protein complexes encoded by a multigene family. Generally, these genes are located at the *Glu-A3*, *Glu-B3* and *Glu-D3* loci on the short arms of chromosomes 1A, 1B and 1D, respectively (Singh and Shepherd 1988; D'Ovidio and Masci 2004). The total gene copy number was estimated to vary from 10–15 (Harberd et al. 1985) to 30–40 (Sabelli and Shewry 1991; Cassidy et al. 1998) in hexaploid wheat based on Southern hybridisation analysis. However, the exact copy number of LMW-GS genes is still unknown, mostly due to lack of efficient methods to distinguish members of this multigene family. Even so, several hundred LMW-GS genes have been isolated from bread wheat and its relatives because of their significant effect on dough strength and extensibility (D'Ovidio and Masci 2004). Based on LMW-GS sequences deposited in GenBank (<http://www.ncbi.nlm.nih.gov/>), the primary structure of a typical LMW-GS has been proposed: a signal peptide of 20 amino acids, a short N-terminal domain (13 amino acids), a central repetitive domain and a C-terminal domain, which is further subdivided into three distinctive regions—namely, C-terminal I, II and III (D'Ovidio and Masci 2004). Sequence alignment has indicated that the signal peptides and the C-terminal I and III regions are well conserved, whereas the repetitive domains are highly polymorphic in length and the number of repeat units in the repetitive domain is mainly responsible for the length variation of different LMW-GSs (D'Ovidio et al. 1999; D'Ovidio and Masci 2004).

To investigate the polymorphic LMW-GS complex in bread wheat, one-dimensional SDS-PAGE (Singh et al. 1991; Liu et al. 2010), reversed-phase high-performance liquid chromatography (RP-HPLC; Margiotta et al. 1993) and matrix-assisted laser desorption/ionisation time-of-flight (MALDI-TOF; An et al. 2006; Liu et al. 2010) were used. However, their efficiency was hindered by the complex band/peak patterns of LMW-GSs and the overlapping mobility between LMW-GSs and gliadins, particularly when testing allelic variations of LMW-GSs in various wheat varieties (Mamone et al. 2009). Recently, several proteomics studies based on two-dimensional gel electrophoresis (2-DE) and mass spectrometry (MS) were performed to characterise proteins in wheat grain endosperm (Amiour et al. 2002; Andon et al. 2002; Balmer et al. 2006; Ikeda et al. 2006; Dong et al. 2010; Liu et al. 2010). This proteomics-based method was imperfect in identifying the complex members of the LMW-GS gene family because of the high-sequence similarity among different LMW-GSs, the high cost and the

strenuous work required. Much attention has been given to allelic variations of LMW-GS proteins, and the genes encoding these proteins have also been well investigated via cDNA and bacterial artificial chromosome (BAC) library screening (Ikeda et al. 2002, 2006; Huang and Cloutier 2008; Dong et al. 2010), providing general elucidation of complex members of the LMW-GS gene family in a particular wheat variety. However, the strategies and technologies (e.g. construction of a DNA or cDNA library) are complex and cannot be widely used to identify LMW-GS genes in a large number of wheat varieties. Meanwhile, molecular markers for LMW-GS genes were investigated to characterise the composition and discriminate the members of the LMW-GS gene family in bread wheat (Long et al. 2005; Ikeda et al. 2006; Zhao et al. 2006, 2007; Huang and Cloutier 2008; Wang et al. 2009, 2010; Dong et al. 2010). These markers might be useful in marker-assisted breeding programmes for the improvement of wheat quality, although using several pairs of primers specific to >10 LMW-GS protein alleles was challenging for wheat breeders. Additionally, these molecular markers were all designed based on sequence polymorphisms among the identified LMW-GS haplotypes, and the number of the cloned genes and identified haplotypes are limited. Thus, distinguishing complex members of the LMW-GS gene family in bread wheat using these molecular markers might be difficult.

From the information presented above, elucidating the complex LMW-GS gene family in several bread wheat varieties is demanding due to the lack of a simple and efficient approach to identify the whole gene family. This difficulty has limited the establishment of a clear relationship between LMW-GSs and bread-making quality, and has hindered the use of LMW-GS in wheat quality improvement. In the present study, we developed a marker system comprising conventional polymerase chain reaction (PCR) technology with conserved primers and a high-resolution capillary electrophoresis system. Using the marker system, almost all LMW-GS gene family members could be amplified by conventional PCR and visualised accurately by capillary electrophoresis. Further investigation of LMW-GS genes in the wheat varieties Xiaoyan 54, Chinese Spring, Norin 61 and Glenlea indicated that this marker system could be used on a large scale to accurately discriminate and elucidate the complex members of the LMW-GS gene family in bread wheat.

Materials and methods

Plant materials

The wheat varieties Xiaoyan 54, Glenlea and Norin 61 were used to validate the conserved primers specific to

LMW-GS genes in bread wheat. Xiaoyan 54 possessing the LMW-GS alleles *Glu-A3d*, *Glu-B3d* and *Glu-D3c* (Liu 2008) is an elite winter Chinese wheat variety with superior bread-making quality. Fourteen unique LMW-GS genes (accession numbers FJ755302–FJ755307 and FJ755309–FJ755316) in Xiaoyan 54 have been characterised with complementary approaches (gene identification through sequencing selected BAC clones, transcript profiling and proteomic analysis; Dong et al. 2010). LMW-GS genes and alleles in Glenlea and Norin 61 had been well understood, i.e. *Glu-A3g*, *Glu-B3g* and *Glu-D3c* in Glenlea, *Glu-A3d*, *Glu-B3i* and *Glu-D3c* in Norin 61 (Ikeda et al. 2002; Huang and Cloutier 2008; Liu et al. 2010). Chinese Spring (*Glu-A3a*, *Glu-B3a* and *Glu-D3a*) and its nulli–tetrasomic lines N1AT1D (nullisomic 1A–tetrasomic 1D), N1BT1D and N1DT1B were used to evaluate the efficiency of the conserved primers and to determine the chromosomal locations of the identified LMW-GS genes.

Sequence analysis of LMW-GS genes from Xiaoyan 54 and GenBank

LMW-GS protein sequences (ACY08809–ACY08819) deduced from 11 active genes in Xiaoyan 54 were collected and clustered with ClustalW software (<http://www.ebi.ac.uk/Tools/clustalw/index.html>) using the default parameters. The phylogenetic tree of 11 LMW-GS in Xiaoyan 54 was constructed using MEGA4.1 (Tamura et al. 2007) with the neighbour-joining method. Bootstrap tests were performed using 1,000 replications.

Thirteen LMW-GS genes (FJ755302–FJ755306 and FJ755309–FJ755316) from Xiaoyan 54, excluding the gene FJ755307 whose coding region was interrupted by a transposon insertion, were clustered using ClustalW. The conserved sequences flanking the repetitive domain were identified and used to design conserved primers specific to LMW-GS genes (Table 1). To optimise the conserved primers, all the other LMW-GS genes with complete coding sequences available in GenBank (<http://www.ncbi.nlm.nih.gov/>) were retrieved using the keywords “(low glutenin) OR (lmw glutenin) OR (LMW-GS) AND triticum AND complete”. The coding sequences of these retrieved LMW-GS genes were clustered using ClustalW. The conserved primers designed above were matched with the clustered sequences. Additionally, to determine whether the designed primers are universal to all known LMW-GS genes, we performed BLAST searches from the EST database using the conserved primers. Based on the results of the analysis, the conserved primers were modified and their loyalty to LMW-GS genes was improved (Table 1). The primers were synthesised using PAGE purification (Invitrogen Biotechnology Co., Ltd.,

Table 1 Three sets of conserved primers designed based on the 13 LMW-GS gene sequences from Xiaoyan 54 BAC clones and those available in GenBank

Name	Sequences (5′–3′)
LMWGS1	
LMWGS1,2F	
LMWGS1,2F ^a	ATGAAGACCTTCCTCRTCCTTTG
LMWGS1R	
LMWGS1R-1	CATGGGTTTAGCTGCTGCA
LMWGS1R-2	CACGGGTTTAGCTGCTGCA
LMWGS1R-3	CATGGGTTTAGTTGCTGCA
LMWGS1R-4	CATGGGTT <u>C</u> AGCTGCTGCA
LMWGS2	
LMWGS1,2F	
LMWGS1,2F ^a	ATGAAGACCTTCCTCRTCCTTTG
LMWGS2R	
LMWGS2R-1	CAACATTGTTGYTGCATCACAT
LMWGS2R-2	CAACATTGT <u>C</u> GCTGCATCACAT
LMWGS2R-3	CAACATTGTTGCCGCATCACAT
LMWGS3a	
LMWGS3aF	
LMWGS3aF-1	CCTGGTTTGGAGAAACCATG
LMWGS3aF-2	CCTGGTTTGGAGAG <u>A</u> CCATG
LMWGS3aR	
LMWGS3aR-1 ^a	CATYTYGACCTAGCAAGACG
LMWGS3aR-2 ^a	CATCTGCGACCTAGCAAGAT <u>T</u>
LMWGS3b	
LMWGS3bF	
LMWGS3bF-1	CCTAGCTTGGAGAAACCATT <u>T</u>
LMWGS3bF-2	CCTGGCTTGGAGAAACCATC
LMWGS3bF-3	CCTGGTTTGGAGAG <u>A</u> CCATC
LMWGS3bF-4	CCTGGTTTGGAGAAACCATC
LMWGS3bR	
LMWGS3bR ^a	CATTTGYGACCTAGCAAGACT
LMWGS3c	
LMWGS3cF	
LMWGS3cF	GCCGTTGCGCAAATTTAC
LMWGS3cR	
LMWGS3cR ^a	CATTTGTGACCTAGCAAGACATC

LMWGS1 and LMWGS2 each contains only one pair of primers, and both share the same forward primer, LMWGS1,2F, whereas LMWGS3 comprises three pairs of primers (LMWGS3aF and LMWGS3aR, LMWGS3bF and LMWGS3bR and LMWGS3cF and LMWGS3cR). The primers with unique nucleotides (underlined) were all synthesised and then mixed to work as the forward or reverse primers

^a 5′ terminal of the primer was labelled with 6FAM

Shanghai, China). Additionally, one primer of each pair was labelled with 6-carboxyfluorescein (6-FAM) at the 5′ terminal (Takara Biotechnology Co., Ltd., Dalian, China).

DNA and RNA isolation and PCR amplification

Genomic DNA was extracted from fresh leaf material collected from 10-day-old wheat seedlings grown in a glasshouse using the cetyl trimethyl ammonium bromide (CTAB) procedure (Saghai-Marooft et al. 1984). Total RNA samples were prepared from developing seeds of Xiaoyan 54 at 14 days post-anthesis using TRIzol[®] reagent according to the manufacturer's protocol (Invitrogen, Carlsbad, CA) and were converted into cDNAs using Moloney Murine Leukemia Virus Reverse Transcriptase (M-MLV RT; Promega, Madison, WI). PCR was performed in 20- μ L reaction volumes containing 1.0 U *LA Taq* DNA polymerase (Takara Bio, Otsu, Japan), 100 ng of genomic DNA, 1 \times GC buffer I (Mg²⁺ plus), 6 pmol of each PCR primer and 8 nmol of each dNTP. PCR conditions were 95°C for 3 min, followed by 35 cycles of 30 s at 94°C, 30 s at 59–61°C, 60 s at 72°C and a final extension step of 10 min at 72°C. PCR reactions were performed in an ABI 9700 thermal cycler (Applied Biosystems, Foster City, CA).

Analysis of PCR products using the Applied Biosystems 3730 DNA Analyzer

PCR products were diluted 1:30 in water, and 1 μ L of the diluted products was added to 7 μ L of HiDi-formamide and 0.1 μ L of GeneScan 1200 LIZ size standard (Applied Biosystems Foster City, CA). After denaturation at 95°C for 10 min, the mixtures were analysed in the Applied Biosystems 3730 DNA Analyzer using the default genotyping module and the G5 dye set. The patterns of the DNA fragments in PCR products were analysed with GeneMapper Software v3.7 according to the manufacturer's instructions.

Cloning and sequencing of DNA fragments

PCR products were separated on 1.2% agarose gels and the expected fragments were recovered from the gels with the TIANgel Midi Purification Kit (Tiangen Biotech Co., Ltd., Beijing, China). Purified DNA fragments were ligated into the pGEM-T vector (Promega) and transformed into competent cells of *Escherichia coli* TOP10 (Tiangen Biotech Co., Ltd.). The positive clones were tested using PCR with the 6FAM-labelled conserved primers, and the size of the PCR fragments was determined using the Applied Biosystems 3730 DNA Analyzer. Clones containing the DNA fragments with the expected sizes were selected and sequenced by SinoGenoMax Co., Ltd. (Beijing, China). Each PCR and sequencing analysis was repeated three to five times to avoid technical errors. Sequence analysis and characterisation were performed using Lasergene software (DNASar; <http://www.dnastar.com/>).

Accession numbers

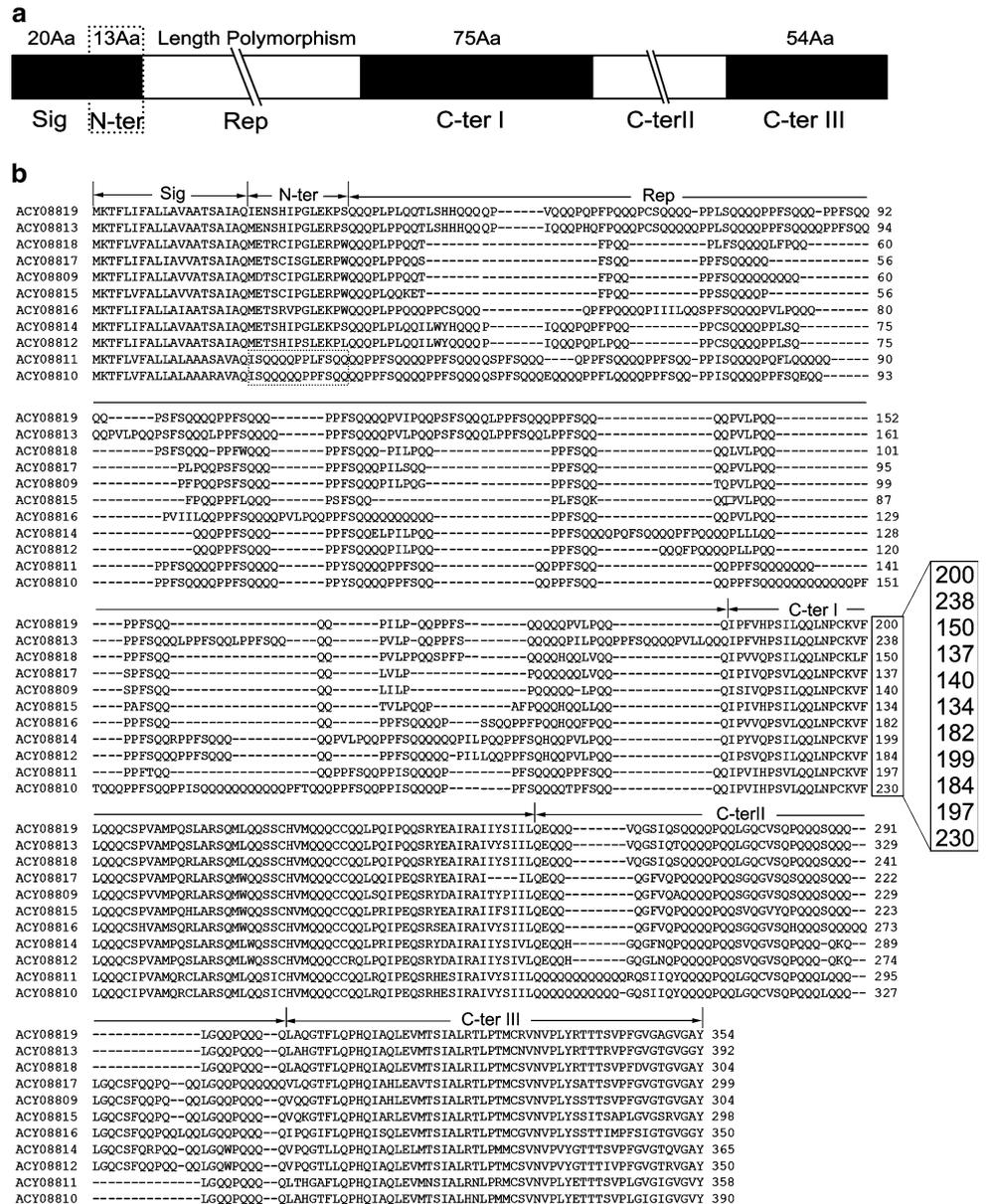
The LMW-GS gene sequences reported in the present study were deposited in GenBank under accession numbers HM640980–HM640985. Specifically, HM640980–HM640983 correspond to the four new LMW-GS genes identified in Xiaoyan 54, i.e. *LGS2-477*, *LGS2-590*, *LGS2-655* and *LGS2-690*, respectively. HM640984 and HM640985 are the two genes cloned from Chinese Spring mentioned below.

Results

Conservation and length polymorphisms among LMW-GSs in Xiaoyan 54

Xiaoyan 54 is an elite winter wheat variety in China with superior bread-making quality. Fourteen unique LMW-GS genes have been identified through BAC clone screening and sequencing (Dong et al. 2010). The deduced peptides of LMW-GS genes in Xiaoyan 54 have the typical structure of LMW-GS, including a signal peptide, a short N-terminal region, a repetitive domain and C-terminal I, II and III regions (Fig. 1a). Among the 14 identified LMW-GS genes, 11 with an intact open reading frame (ORF) were shown to be expressed by reverse transcription-polymerase chain reaction (RT-PCR), 2-DE and MS (Dong et al. 2010). Multiple alignments of the deduced amino acid sequences were performed using ClustalW software (<http://www.ebi.ac.uk/Tools/clustalw/index.html>), which revealed that these subunits shared high-sequence identity ranging from 62 to 94% (Fig. 1b). In particular, the signal peptides and the C-terminal I region were highly conserved. The signal peptide sequences could be summarised as MKTFLI(V)FALL(I)A(I/L)AV(V)AT(A)S(R)AI(V)AQ, and several peptides in the C-terminal I region—LQQLNPCK, FLQQQC, LARSQM and VMQQQC—were completely conserved in these LMW-GSs (Fig. 1b). In contrast, the repetitive domains were highly polymorphic in length among these LMW-GSs, and each LMW-GS has its unique length of repetitive domain, varying from 83 to 192 amino acid residues. The length of the sequences covering the signal peptide, the N-terminal region, the repetitive domain and the partial C-terminal I region (200, 238, 150, 137, 140, 134, 182, 199, 184, 197 and 230 amino acids; Fig. 1b) displayed sufficient polymorphisms among the repetitive domains of these LMW-GSs in Xiaoyan 54. Accordingly, the sequence conservation and length polymorphisms among LMW-GSs in Xiaoyan 54 provided us with the feasibility to distinguish members of the LMW-GS gene family using conventional PCR with conserved primers

Fig. 1 Sequence alignment of the deduced proteins of 11 active LMW-GS genes in Xiaoyan 54. **a** Schematic diagram showing the structure of a typical LMW-GS as deduced from their encoding genes. **b** Multiple alignments of the amino acid sequences deduced from 11 active LMW-GS genes in Xiaoyan 54 (Dong et al. 2010). The s- and m-type subunits, with serine and methionine as the first amino acids of the mature LMW-GS (ACY08809, ACY08812–ACY08819), contain a signal peptide (Sig), an N-terminal domain (N-ter), a repetitive domain (Rep) and a C-terminal domain (C-ter), which is further divided into three regions: C-terminal I (C-ter I), C-terminal II (C-ter II) and C-terminal III (C-ter III). The two i-type subunits (ACY08810 and ACY08811) lack the N-terminal domain (enclosed with broken lines), starting directly with the repetitive domain after the signal peptide. The group of numbers (200, 238, 150, 137, 140, 134, 182, 199, 184, 197 and 230) represents the length of the sequences covering the signal peptide, the N-terminal domain, the repetitive domain and part of the C-terminal I region



that were matched with the conserved regions flanking the repetitive domains.

Development of conserved primers for identifying LMW-GS genes in bread wheat

To design conserved primers, we aligned the coding sequences of the 13 unique LMW-GS genes with complete coding sequences identified in Xiaoyan 54, excluding the LMW-GS gene (FJ755307) interrupted with a transposon insertion (Dong et al. 2010). The sequence alignment revealed that the sequences encoding the signal peptide and C-terminal I region shared high identity, up to 90%. Thus, the conserved nucleotide sequences were used for the primer design (Fig. 2). To ensure the reproducibility and

reliability of the analysis, three sets of conserved primers that matched with different conserved regions—LMWGS1, LMWGS2 and LMWGS3—were developed (Fig. 2; Table 1). LMWGS1 and LMWGS2 each contained only one pair of primers, whereas LMWGS3 comprised three pairs of primers (LMWGS3aF and LMWGS3aR, LMWGS3bF and LMWGS3bR and LMWGS3cF and LMWGS3cR; Fig. 2). LMWGS1 and LMWGS2 were designed based on the high-sequence identity of the coding sequences for the signal peptide and C-terminal I region. They shared the same forward primer, LMWGS1,2F, located in the coding region for the signal peptides and starting from the initiation codon of the LMW-GS genes, while their reverse primers, LMWGS1R and LMWGS2R, were located in different regions of C-terminal I (Fig. 2).

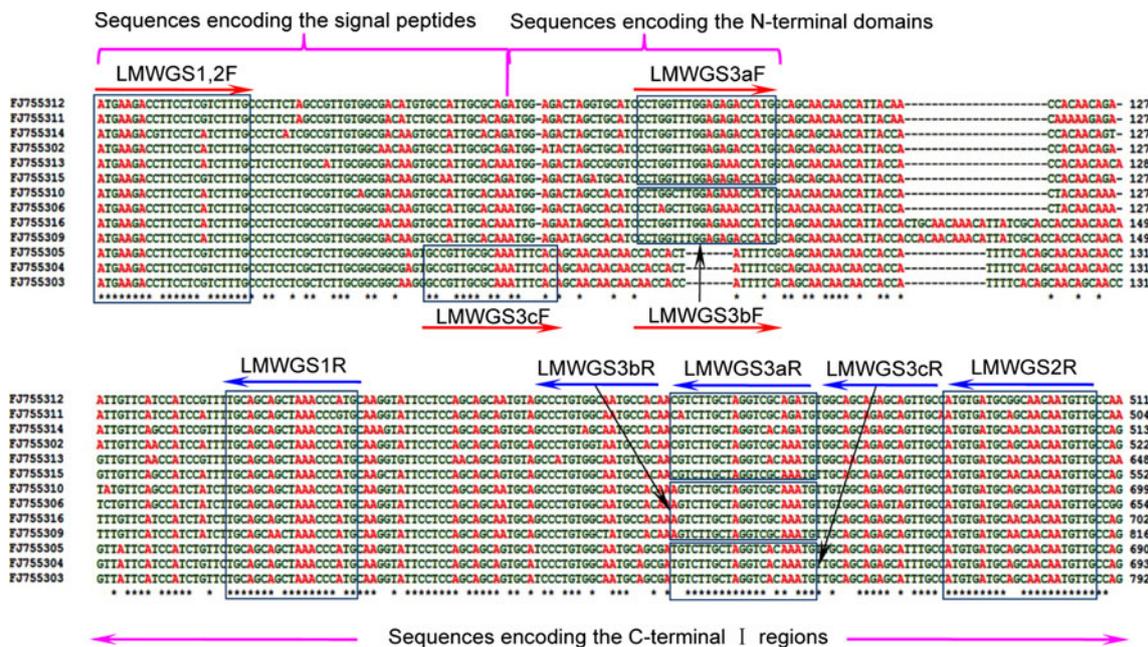


Fig. 2 Design of the conserved primers based on nucleotide sequence conservation of the LMW-GS genes in Xiaoyan 54. Based on the conserved sequences of LMW-GS genes, three sets of conserved primers were designed: LMWGS1, LMWGS2 and LMWGS3. LMWGS1 comprised LMWGS1,2F and LMWGS1R. LMWGS2 contained the pair of primers LMWGS1,2F and LMWGS2R, whereas LMWGS3 comprised the three pairs of conserved primers LMWGS3a

(LMWGS3aF and LMWGS3aR), LMWGS3b (LMWGS3bF and LMWGS3bR) and LMWGS3c (LMWGS3cF and LMWGS3cR). All of the conserved primers were matched to the sequences flanking the repetitive sequences. Specifically, all the forward primers were located in the sequences encoding the signal peptide or the N-terminal domain, while the reverse primers were located in the sequences encoding the C-terminal I region

The different locations of the primers resulted in a 108-bp longer amplicon of LMW-GS2 than that of LMW-GS1 for each LMW-GS gene. Theoretically, all the LMW-GS genes could be amplified simultaneously by conventional PCR using each set of the conserved primers: i.e. >10 amplicons might be obtained with LMWGS1 or LMWGS2. To simplify the PCR amplification, a third set of primers, LMWGS3, was designed that contained three pairs of primers specific to a subset of LMW-GS genes. The sequence alignment of the subunits deduced from 11 active LMW-GS genes in Xiaoyan 54 revealed that these LMW-GSs can be categorised into four groups (I, II, III and IV; Supplementary Fig. 1). Accordingly, three pairs of conserved primers were designed. LMWGS3a was specific to groups I and II, LMWGS3b to group III and LMWGS3c to group IV. Primers LMWGS3aF and LMWGS3bF were derived from the N-terminal coding region, while LMWGS3cF was designed based on the unique coding region for the signal peptide and the repetitive region of LMW-i types. The reverse primers with unique nucleotides—LMWGS3aR, LMWGS3bR and LMWGS3cR—were positioned at the same coding region for the C-terminal I region (Fig. 2). Thus, each of these group-specific primers could amplify a subset of LMW-GS genes, and all three subsets of amplified genes would comprise the entire LMW-GS gene family in a bread wheat variety.

To optimise the conserved primers, the LMW-GS gene sequences deposited in GenBank were used to explore the conservation of these primers designed from LMW-GS genes in Xiaoyan 54. The database was searched with the keywords “(low glutenin) OR (lmw glutenin) OR (LMW-GS) AND triticum AND complete”, and 384 LMW-GS genes with complete coding sequences were retrieved, including eight LMW-GS genes in seven BAC sequences. The coding sequences of these LMW-GS genes were used for cluster analysis using ClustalW. Three sets of conserved primers, designed according to LMW-GS genes in Xiaoyan 54, were matched with counterpart clustered nucleotide sequences, which were slightly modified, if necessary. Additionally, dbEST of GenBank was searched by BLAST with the improved primers. Based on the analysis of BLAST hits, the conserved primers were modified and made specific to all known LMW-GS genes. These modifications improved the representativeness of the conserved primers, and the improved primers matched >85% of the LMW-GS gene sequences deposited in GenBank.

Validation of the conserved primers in Xiaoyan 54

Based on the conserved sequences and length polymorphisms among LMW-GS genes in Xiaoyan 54 and those available in GenBank, three sets of conserved primers

Fig. 3 Analysis of polymerase chain reaction (PCR) products amplified from Xiaoyan 54 with the conserved primers.

a Electrophoresis of the PCR products amplified from Xiaoyan 54 with the conserved primers in an agarose gel.

M DNA ladder Marker III (200, 500, 800, 1,200, 2,000, 3,000 and 4,500 bp; Tiangen Biotech Co., Ltd.).

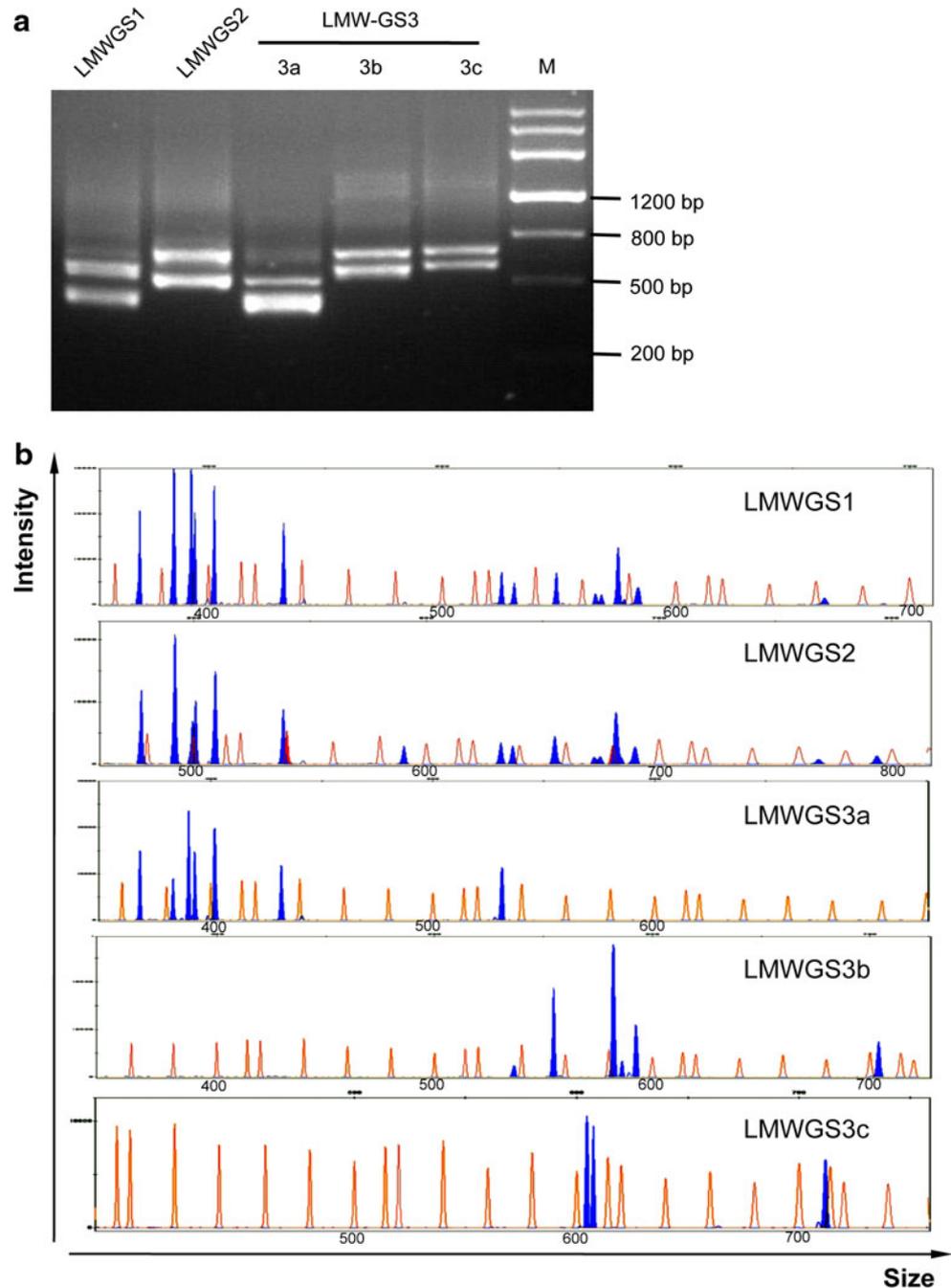
b Electropherograms showing capillary electrophoresis separation of the DNA fragments amplified from Xiaoyan 54 with the conserved primers.

The *horizontal axis* shows the size of the detected DNA fragments, while the *vertical axis* displays the intensity of the signal (i.e. the concentration of DNA fragments in the PCR products).

The *orange peaks* match the standard DNA fragments in the GeneScan 1200 LIZ size standard, while the *blue ones* representing the DNA fragments in the PCR products, should match theoretically with the LMW-GS genes in Xiaoyan 54.

The *numbers on the horizontal axis* represent the size of the corresponding peak in the GeneScan 1200 LIZ size standard (*orange*). In Xiaoyan 54, 14 *blue peaks* (DNA fragments) could be detected with the conserved primer LMWGS1, 16 with LMWGS2 and 16 with LMWGS3 (7 with LMWGS3a, 6 with LMWGS3b and 3 with LMWGS3c).

The data shown are representative of four independent sets of PCR amplifications and DNA fragment analyses



flanking the repetitive domain were designed. To ensure the efficiency of these primers in characterising LMW-GS genes in bread wheat, Xiaoyan 54 was analysed. Only two main bands for each pair of primers were detected in the PCR products using 1.5% agarose gel electrophoresis (Fig. 3a). To effectively and clearly display all the DNA fragments in the PCR products, capillary electrophoresis was applied to DNA fragment analysis, resulting in a high resolution (up to 1 bp; Butler et al. 2004). The electropherograms revealed that 15 DNA fragments (blue peaks) were detected in the PCR products with the primer LMWGS1, 16 DNA

fragments with LMWGS2, 7 DNA fragments with LMWGS3a, 6 DNA fragments with LMWGS3b and 3 DNA fragments with LMWGS3c (Fig. 3b). These data suggested that 15 or 16 DNA fragments could be amplified from Xiaoyan 54 using conventional PCR with one set of the conserved primers, consistent with the previous observation that 14 unique LMW-GS genes were identified in Xiaoyan 54. Thus, the conserved primers and high-resolution capillary electrophoresis comprised a new molecular marker system for simple and efficient characterisation of the composition of LMW-GS genes in a bread wheat variety.

Table 2 Size of the DNA fragments amplified from Xiaoyan 54 with three sets of conserved primers and subsequently detected with the Applied Biosystems 3730 DNA Analyzer

Gene	LMWGS1		LMWGS2		LMWGS3a		LMWGS3b		LMWGS3c		Accession no. ^c	Identity ^d (%)
	Theo ^a	Exp ^b	Theo	Exp	Theo	Exp	Theo	Exp	Theo	Exp		
<i>LGS2-477</i> ^{e,f}	377	370	485	477	366	368	–	–	–	–	AJ519836	98
<i>D3-4</i>	392	385	500	492	381	383	–	–	–	–	FJ755311	100
<i>D3-5</i> ^f	399	393	507	499	388	390	–	–	–	–	FJ755312	100
<i>D3-7</i>	401	394	509	501	390	393	–	–	–	–	FJ755314	100
<i>A3-1</i>	410	402	518	509	399	402	–	–	–	–	FJ755302	100
<i>D3-2</i>	440	432	548	538	429	432	–	–	–	–	FJ755315	100
<i>LGS2-590</i> ^{e,f}	493	N ^g	601	590	–	–	–	–	512	N	FJ549941	99
<i>D3-6</i>	536	525	644	632	525	531	–	–	–	–	FJ755313	100
<i>B3-1</i>	542	530	650	637	–	–	531	537	–	–	FJ755306	100
<i>LGS2-655</i> ^{e,f}	561	549	669	655	–	–	550	556	–	–	EU369731 ^h 6	100
<i>A3-3</i> ^f	578	566	686	672	–	–	–	–	597	605	FJ755305	100
<i>A3-2</i>	581	568	689	675	–	–	–	–	600	608	FJ755304	100
<i>D3-1</i>	587	575	695	682	–	–	576	583	–	–	FJ755310	100
<i>D3-3</i>	590	N	698	N	–	–	579	586	–	–	FJ755316	100
<i>LGS2-690</i> ^{e,f}	596	584	704	690	–	–	585	593	–	–	AY585349	97
<i>A3-4</i>	680	664	788	768	–	–	–	–	699	715	FJ755303	100
<i>B3-2</i>	704	688	812	792	–	–	693	707	–	–	FJ755309	100
<i>B3-3</i> ^f	– ^h	–	–	–	–	–	–	–	–	–	FJ755307	–

^a Theoretical size of DNA fragments

^b Experimental size of DNA fragments

^c FJ755302–FJ755307 and FJ755309–FJ755316 were identified in Xiaoyan 54 BAC clones reported by Dong et al. (2010)

^d Nucleotide sequence identity of the DNA fragments to previously reported LMW-GS genes

^e The new genes from Xiaoyan 54 identified in the present study

^f The genes were not expressed in the developing seeds of Xiaoyan 54

^g Not detected

^h Not applicable

To check the consistency of the DNA fragments revealed by the new marker system with the LMW-GS genes previously identified in Xiaoyan 54 BAC clones, the experimental size of the DNA fragments was calculated with the GeneScan 1200 LIZ size standard using GeneMapper Software v3.7. Next, the experimental size of the DNA fragments and the theoretical size of 13 genes with complete sequences in Xiaoyan 54 [excluding the LMW-GS gene (FJ755307) interrupted with a transposon insertion] were compared (Table 2). The experimental size of the amplicons of LMW-GS1 ranged from 370 to 688 bp, and that of LMW-GS2 from 477 to 792 bp (Table 2). The experimental size differences between the amplicons of both primers were 104–107 bp, which was consistent with the theoretical difference of 108 bp. The third conserved primer set, LMWGS3, comprised three pairs of primers (LMWGS3a, LMWGS3b and LMWGS3c), and each could amplify an independent but complementary subset of LMW-GS genes in a given bread wheat variety (Table 2). Thus, 16 unique DNA fragments

in Xiaoyan 54 were identified with the third primer set, which was consistent with the data from LMWGS1 and LMWGS2 (Table 2). The integrated data from the three sets of conserved primers demonstrated that 17 DNA fragments can be identified from Xiaoyan 54 (Table 2). Except for four new DNA fragments, the other 13 fragments were well matched with the LMW-GS genes isolated in Xiaoyan 54, although small deviations existed between the experimental size and theoretical sizes (Table 2). These kinds of deviations have been reported to be universal and were probably caused by the capillary electrophoresis and the internal standard system (GeneScan 1200 LIZ size standard; Rosenblum et al. 1997; Akbari et al. 2008). Additionally, to investigate whether all the DNA fragments were derived from LMW-GS genes, the 17 DNA fragments were cloned, sequenced and aligned with LMW-GS genes. In total, 13 DNA fragments that were matched in size with the LMW-GS genes in Xiaoyan 54 were identical to the corresponding LMW-GS genes (Dong et al. 2010), and 4 new DNA fragments were

identified with >97% identity to the LMW-GS genes in GenBank by BLAST searches (Table 2). The four LMW-GS newly identified genes in Xiaoyan 54, corresponding to DNA fragments 477, 590, 655 and 690 (size of amplicons from primer LMWGS2), were named *LGS2-477*, *LGS2-590*, *LGS2-655* and *LGS2-690*, respectively (Table 2). These genes may all be pseudogenes because stop codons were present in their coding regions. Collectively, using this new marker system, 17 unique LMW-GS genes could be detected in Xiaoyan 54, and only one, *B3-3* (FJ755307) reported by Dong et al. (2010), could not be detected because it was interrupted by a large transposon insertion. These findings suggest that >18 unique LMW-GS genes exist in the bread wheat variety Xiaoyan 54 (Table 2).

To validate our new marker system, the transcription profiles of LMW-GS genes in developing seeds of Xiaoyan 54 were also investigated. RT-PCR analysis using total RNA extracted from developing seeds of Xiaoyan 54 and the electropherogram displaying 11 blue peaks indicated that 11 unique genes were expressed (Supplementary Fig. 2; Table 2). These expressed genes matched exactly with the 11 active genes (*A3-1*, *A3-2*, *A3-4*, *B3-1*, *B3-2*, *D3-1*, *D3-2*, *D3-3*, *D3-4*, *D3-6* and *D3-7*) observed by Dong et al. (2010) previously.

These results of the analysis of LMW-GS genes in Xiaoyan 54 at the DNA and cDNA levels demonstrated that this new marker system, comprising conventional PCR with conserved primers and high-resolution capillary electrophoresis, could be used to effectively discriminate the LMW-GS genes of the complex gene family in Xiaoyan 54, and might be universally applied to identify LMW-GS genes in bread wheat.

Characterisation of the LMW-GS genes in Chinese Spring with the marker system

The efficiency of the marker system in characterising LMW-GS genes was evaluated in another bread wheat variety, Chinese Spring. Conventional PCR was performed with the genomic DNA of Chinese Spring using these conserved primers, and DNA fragments in the PCR products were identified using the Applied Biosystems 3730 DNA Analyzer. The data showed that >10 DNA fragments could be displayed in the PCR products amplified with each set of conserved primers. Specifically, 15 DNA fragments were identified with the primer LMWGS1, 14 with LMWGS2, 7 with LMWGS3a, 7 with LMWGS3b and 2 with LMWGS3c (Fig. 4a). With the help of the GeneScan 1200 LIZ size standard, the experimental size of the DNA fragments was calculated and compared (Supplementary Table 1), revealing that most of the DNA fragments in each PCR product matched well with each other.

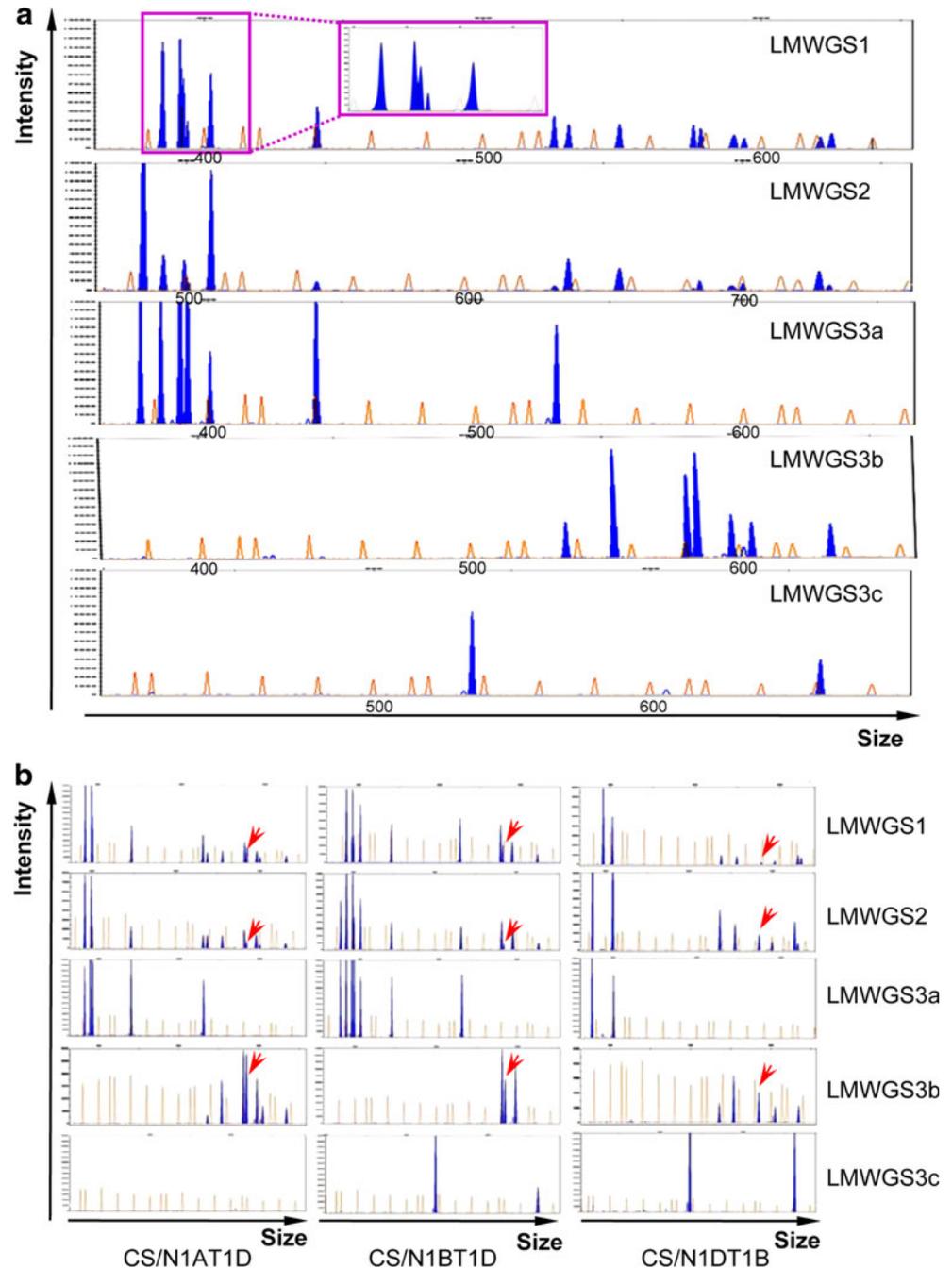
The only exception was DNA fragment 537, which could only be detected in the PCR products with the primer LMWGS3c (Supplementary Table 1; Fig. 4a). To confirm whether it belonged to the LMW-GS gene family, fragment 537 was cloned and sequenced. Using the sequence of fragment 537 (HM640984), BLAST searches revealed that its corresponding gene might encode a LMW-i-type subunit. With the new marker system, we identified ≥ 16 DNA fragments in Chinese Spring representing 16 LMW-GS genes.

Typical LMW-GS genes are located at the *Glu-A3*, *Glu-B3* and *Glu-D3* loci on the short arms of chromosome 1A, 1B and 1D, respectively (Singh and Shepherd 1988; D'Ovidio and Masci 2004). To determine the chromosomal locations of individual LMW-GS genes, three Chinese Spring nulli-tetrasomic lines—N1AT1D, N1BT1D and N1DT1B—were used. Conventional PCR was performed with three sets of conserved primers in the marker system, and the PCR products were displayed using the Applied Biosystems 3730 DNA Analyzer, followed by analysis using GeneMapper Software v3.7. Compared to the fragments detected in Chinese Spring, four, four and seven distinct DNA fragments were missing in the nulli-tetrasomic lines N1AT1D, N1BT1D and N1DT1B, respectively (Fig. 4b; Table 3), suggesting that four LMW-GS genes were located on chromosome 1A, four on 1B and seven on 1D. Unexpectedly, DNA fragment 684 could be detected in all three Chinese Spring nulli-tetrasomic lines (Fig. 4b; Table 3), although this fragment corresponding to the *D3-3* gene in Xiaoyan 54 was mapped to the *Glu-D3* locus (Dong et al. 2010; Table 2). This inconsistency suggested that some other sequences might be amplified or another LMW-GS gene with the same DNA fragment size might be located at loci other than *Glu-D3*. To investigate its origins, DNA fragment 684 was cloned and sequenced, and two sequences were obtained. One was identical to *D3-3* (Dong et al. 2010) and the other was homologous to a LMW-GS pseudogene (EU057598) with 99% identity by online BLAST. Further analysis revealed that this new LMW-GS gene (HM640985) was located on chromosome 1B for its absence in N1BT1D. Collectively, using the developed marker system, 17 LMW-GS genes in Chinese Spring, represented by the DNA fragments, were shown to be located on the group 1 chromosomes, with 4 on chromosome 1A, 5 on 1B and 8 on 1D.

Discussion

Previously, several molecular markers were developed based on sequence polymorphisms among LMW-GS genes in bread wheat (Zhang et al. 2004; Long et al. 2005; Ikeda et al. 2006; Zhao et al. 2006, 2007; Wang et al. 2009, 2010).

Fig. 4 Electropherograms displaying the patterns of the DNA fragments detected in Chinese Spring and its nulli-tetrasomic lines with the marker system. **a** Identification of the LMW-GS genes in Chinese Spring with the developed marker system. **b** Determination of the location of the DNA fragments with the Chinese Spring nulli-tetrasomic lines N1AT1D, N1BT1D and N1DT1B. The horizontal and vertical axes are the same as those described in Fig. 3b. In Chinese Spring, 15 blue peaks (DNA fragments) could be detected with the conserved primer LMWGS1, 14 with LMWGS2 and 16 with LMWGS3 (7 with LMWGS3a, 7 with LMWGS3b and 2 with LMWGS3c). The data collected from three sets of conserved primers indicated that 12 blue peaks (DNA fragments) were detected in N1AT1D, 12 in N1BT1D and 9 in N1DT1B. However, DNA fragment 684 (marked with red arrows) could be detected in all three nulli-tetrasomic lines. All data provided are representative of four independent sets of PCR amplifications and analyses using the Applied Biosystems 3730 DNA Analyzer



Here, we focussed on the conserved structure and length polymorphisms among LMW-GS genes. Comprising conserved primers and capillary electrophoresis, a new marker system was developed to characterise multiple LMW-GS genes in bread wheat varieties. This system was validated in Xiaoyan 54 and Chinese Spring, confirming that it could be efficiently used for large-scale characterisation of the complex LMW-GS gene family in bread wheat.

The new marker system, comprising conserved primers and capillary electrophoresis, is useful to elucidate the complex members of the LMW-GS gene family

Although several hundred LMW-GS genes have been cloned and deposited in GenBank, accurate characterisation of the complex members of the LMW-GS gene family in a bread wheat variety is still difficult regarding wheat quality improvement (Ikeda et al. 2002; Huang and Cloutier 2008; Dong et al. 2010). In the present study, we focussed on sequence conservation and length polymorphisms of LMW-GS genes and established a new marker system with which all LMW-GS genes could be characterised using conventional PCR. Previously, LMW-GS genes in three bread wheat varieties—Norin 61, Glenlea

Table 3 DNA fragments amplified from Chinese Spring and its nulli–tetrasomic lines N1AT1D, N1BT1D and N1DT1B

Chromosome	Chinese Spring	N1AT1D	N1BT1D	N1DT1B	Accession no. ^a
1A	485 ^b	N ^c	485	485	– ^d
1D	492	492	492	N	AK330800 ^e
1D	499	499	499	N	–
1D	501	501	501	N	DQ457417 ^f
1A	509	N	509	509	–
1D	547	547	547	N	FH615311
1A	596	N	596	596	–
1D	632	632	632	N	DQ357052/AY214450
1B	637	637	N	637	X84960
1B	655	655	N	655	–
1D	682	682	682	N	X84961
1B/1D	684	684	684	684	DQ357057(1D)
1D	695	695	695	N	–
1B	700	700	N	700	–
1A	725	N	725	725	AT453154/FJ876821
1B	729	729	N	729	–

^a The LMW-GS genes/alleles, identified from Chinese Spring reported in GenBank, share the same size and/or high identity with the DNA fragments detected in the present study

^b Sizes of the DNA fragments were represented with the data from the primer LMWGS2

^c Not detected

^d Not applicable

^e From the cDNA clone in Chinese Spring, which was not annotated to be a LMW-GS gene

^f Having high-sequence identity with the sequences of the corresponding DNA fragment, but containing some deletions in the repetitive region

and Xiaoyan 54—were extensively investigated using complementary approaches, and >10 genes had been identified in each variety (Ikeda et al. 2002; Huang and Cloutier 2008; Dong et al. 2010). Thus, the data from these three varieties would be useful to validate the efficiency of our marker system. In Xiaoyan 54, using three sets of conserved primers in our marker system, 17 DNA fragments were detected, sequenced and confirmed to be LMW-GS genes, of which 13 sequences were identical to the LMW-GS genes identified from BAC clones (Dong et al. 2010; Table 2; Fig. 3b). The other four new genes were mapped to the *Glu-3* locus by this marker system and by microsatellite markers with a population of recombinant inbred lines derived from a cross between Xiaoyan 54 and Jing 411 (data not shown). Of the 18 unique LMW-GS genes identified in Xiaoyan 54, 6 were located at the *Glu-A3* locus, 4 at the *Glu-B3* locus and 8 at the *Glu-D3* locus (Table 2). In Norin 61, 18 unique DNA fragments were identified, with most being identical in size to those from Xiaoyan 54 (Supplementary Table 2). Compared to the sequences reported by Ikeda et al. (2002), 12 DNA fragments could be matched with 12 genes with high similarities in the coding sequences for N-terminal and C-terminal regions, whereas 4 of the genes contained deletions in the repetitive domain, resulting in the different sizes of the

corresponding sequences between the report by Ikeda et al. (2002) and the present study (Supplementary Table 2). The other six DNA fragments—460, 499, 590, 655, 675 and 695 (PCR products with LMWGS2)—might have been derived from new LMW-GS genes that were detected and reported for the first time to our knowledge in Norin 61 (Supplementary Table 2). Using the marker system, 16 DNA fragments were isolated from the genomic DNA of Glenlea, while 12 active genes and seven pseudogenes were isolated by Huang and Cloutier (2008). Nine DNA fragments were identical in length with the 12 active genes (Supplementary Table 2) and the other 7 corresponded as expected to the 7 pseudogenes described by Huang and Cloutier (2008). Similar to the situation of fragment 684 in Chinese Spring detected by the marker system, two fragments—492 and 637—have three and two corresponding active genes (EU189090, EU189091 and EU189092; EU189089 and EU189097), respectively, reported by Huang and Cloutier (2008). Sequence alignment revealed that the three active genes—EU189090, EU189092 and EU189097—were unique in GenBank since they contained special single-nucleotide polymorphisms (SNPs) only present in Glenlea. Future experiments should be performed to confirm whether these three genes are universal in bread wheat. In Chinese Spring, of 12 LMW-GS genes

retrieved from GenBank, 8 were identical in size to the DNA fragments identified by our marker system (Supplementary Table 2), and the other 4 genes were unique, with deletions in the repetitive region that might have been introduced during the PCR process (Masci et al. 1998; Ikeda et al. 2002). These data suggested that 9 LMW-GS genes corresponding to 9 of the 17 DNA fragments might be reported for the first time in Chinese Spring. In summary, the comparison demonstrated that the marker system could help us detect not only most of the LMW-GS genes reported previously, but also some new LMW-GS genes in Xiaoyan 54, Norin 61, Glenlea and Chinese Spring that were not found in other studies. The marker system could potentially be used to identify LMW-GS genes and elucidate the complex of this gene family in bread wheat.

The sizes of the DNA fragments detected in Xiaoyan 54, Chinese Spring, Norin 61 and Glenlea were compared (Supplementary Table 2). The data showed that these varieties shared several genes with the same size of DNA fragments: *D3-4* (492), *D3-5* (499), *D3-7* (501), *B3-1* (637), *LGS2-655* (655), *D3-1* (682) and *D3-3* (684). The alignment analysis demonstrated that the sequences of these genes were extremely conserved, with only a few SNPs among different varieties. Allelic variations in the other DNA fragments and their corresponding sequences in GenBank were aligned as well, indicating that allelic variations of a certain gene were also conserved in nucleotide sequences, although certain kinds of deletions were present in each allelic variation. On the other hand, allelic variations of the genes (i.e. *LGS2-477*, *A3-1*, *LGS2-590*, *A3-2*, *A3-3* and *A3-4* at the *Glu-A3* locus, *B3-2* and *B3-3* at the *Glu-B3* locus, *D3-2*, *D3-6* and *LGS2-690* at the *Glu-D3* locus) displayed high polymorphism in the length of the repetitive domains (Supplementary Table 2). The allelic variation of *D3-2* gene, e.g., corresponded to the DNA fragment 547 in *Glu-D3a* allele (Chinese Spring), and 538 in *Glu-D3c* allele (Xiaoyan 54, Norin 61 and Glenlea). These data suggested that the differences in the molecular weight among LMW-GS alleles, which were characterised with SDS-PAGE and MALDI-TOF-MS (Singh et al. 1991; Dworschak et al. 1998; Liu et al. 2010), might also be detected by the developed marker system. Further work is necessary to characterise LMW-GS genes in hundreds of wheat varieties with this developed molecular marker system. Comparison of the composition of the DNA fragments and alignment of the sequences of LMW-GS genes among wheat varieties might deeply display the conservation and polymorphism among allelic variations and facilitate the construction of the relationship between each unique DNA fragment and a specific LMW-GS gene. Thus, this new marker system is useful in characterising the LMW-GS genes and illuminating the composition of this gene family in bread wheat.

Advantages of the new marker system in identifying LMW-GS genes

Combining the conserved primers and capillary electrophoresis, we have provided an efficient marker system to identify LMW-GS genes in bread wheat. Genomic DNA or cDNA from several bread wheat varieties was amplified with the conserved primers, and subsequently, the PCR products were analysed using the Applied Biosystems 3730 DNA Analyzer. This system has advantages over traditional protein-based techniques and other marker systems for identifying LMW-GS genes.

First, the marker system is efficient and applicable to all wheat varieties in identifying LMW-GS genes. Compared to protein-based methods, DNA-based molecular markers are simpler to operate and have been widely used as diagnostic markers for the identification of LMW-GS in bread wheat (D'Ovidio 1993; D'Ovidio and Porceddu 1996; D'Ovidio et al. 1999; Long et al. 2005; Ikeda et al. 2006; Zhao et al. 2006, 2007; Huang and Cloutier 2008; Wang et al. 2009, 2010; Dong et al. 2010). Group-specific markers have been designed based on polymorphic sites of the LMW-GS genes identified in Norin 61 and Glenlea (Ikeda et al. 2006; Huang and Cloutier 2008). These markers might be useful in characterising LMW-GS genes in a special variety, but their efficiency in identifying LMW-GS genes in other varieties has not been validated. Recently, following the analysis of LMW-GS gene sequences in different haplotypes, several sequence-tagged-site (STS) markers for *Glu-A3*, *Glu-B3* and *Glu-D3* subunits were developed and validated in >100 wheat varieties (Zhao et al. 2006, 2007; Wang et al. 2009, 2010). Additionally, some other gene-specific markers were developed based on polymorphic sites of the cloned LMW-GS genes (D'Ovidio 1993; D'Ovidio and Porceddu 1996; D'Ovidio et al. 1999; Dong et al. 2010). However, the LMW-GS gene family in bread wheat was complex and their exact composition was still not well understood (D'Ovidio and Masci 2004; Liu et al. 2010). Thus, the above molecular markers were useful in detecting the known LMW-GS genes, but were limited and difficult in identifying and characterising some other new genes in the LMW-GS gene family in bread wheat. In the present study, we developed a LMW-GS gene molecular marker system based on the conservation and polymorphism among LMW-GS genes. In the marker system, conserved primers ensured the amplification of all the LMW-GS genes in bread wheat, which made the marker system applicable to all wheat varieties, and the high polymorphism in length of the repetitive domain facilitated the separation of the amplified LMW-GS genes by capillary electrophoresis. After conventional PCR using conserved primers, >10 DNA fragments could be displayed using the Applied

Biosystems 3730 DNA Analyzer, and each fragment was determined to be derived from a LMW-GS gene (Figs. 3b, 4a). The marker system has been evaluated using several varieties, in which LMW-GS genes have been investigated (Ikeda et al. 2006; Huang and Cloutier 2008; Dong et al. 2010). Except for the known LMW-GS genes, several new LMW-GS genes and their allelic variations (e.g. *LGS2-477*, *LGS2-590* and *LGS2-690*) have been identified and reported for the first time.

Second, the marker system is accurate with the use of high-resolution capillary electrophoresis. PCR products using the marker system were separated by high-resolution capillary electrophoresis, which has higher resolution (up to 1 bp) than traditional agarose or denatured/native polyacrylamide gel electrophoresis. Capillary electrophoresis instruments have been widely used for DNA fragment analysis: e.g. in simple sequence repeat, amplified fragment length polymorphism and DNA footprinting analyses (Maguire et al. 2002; Broza et al. 2007; Zianni et al. 2006). Additionally, with the help of GeneScan 1200 LIZ size standards, the size of DNA fragments can be measured digitally using GeneMapper Software v3.7, which could facilitate further comparison of LMW-GS genes among various varieties.

Third, the marker system is time-saving and high throughput. Conventional PCR and capillary electrophoresis in the marker system are highly automatic, and the run time and subsequent analysis can be completed in 4–5 h as compared to 2–4 days for protein isolation and subsequent identification. Protein-based methods entail protein isolation and separation, gel spot excision, in-gel digestion, MS analysis and database searching (Mamone et al. 2009; Dong et al. 2010). All these steps add considerable labour and expense to the experiment for LMW-GS identification in one sample, whereas our marker system contains two highly automatic steps: conventional PCR and capillary electrophoresis. Thus, with this marker system, several hundred samples could be easily analysed by one person in 1 day. This marker system does, however, have the disadvantage of being difficult to distinguish active LMW-GS genes in the bread wheat genome, but it is still efficient in identifying LMW-GS genes expressed in developing kernels, such as those shown in Xiaoyan 54 (Supplementary Fig. 2).

In summary, we developed a simple and reliable marker system to elucidate the complex LMW-GS gene family in bread wheat. The molecular marker system comprises the use of conserved primers for LMW-GS genes and high-resolution capillary electrophoresis and was confirmed to be efficient in characterising LMW-GS genes in Xiaoyan 54, Norin 61, Glenlea and Chinese Spring. Thus, the marker system, involving conserved primers and capillary electrophoresis, is valid and powerful to elucidate the

members of the LMW-GS gene family in bread wheat. In near future, the marker system may help to improve the characterisation of LMW-GS genes in numerous bread wheat varieties and its relatives and to facilitate the establishment of a clear relationship between the allelic variations in LMW-GS genes and bread-making quality in wheat.

Acknowledgments We thank Drs Xianchun Xia (Chinese Academy of Agricultural Sciences, Beijing, China) and Xiuqiang Huang (University of Manitoba, Winnipeg, Canada) for constructive suggestions on this work. This work was supported by the Ministry of Science and Technology of China (2009CB118300 and 2006AA10Z1B2) and the Ministry of Agriculture of China for transgenic research (2008ZX08009-003 and 2008ZX08002-004).

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Akbari A, Marthinsen G, Lifjeld JT, Albrechtsen F, Wennerberg L, Stenseth NC, Jakobsen KS (2008) Improved DNA fragment length estimation in capillary electrophoresis. *Electrophoresis* 29:1273–1285
- Amiour N, Merlini M, Leroy P, Branlard G (2002) Proteomic analysis of amphiphilic proteins of hexaploid wheat kernels. *Proteomics* 2:632–641
- An X, Zhang Q, Yan Y, Li Q, Zhang Y, Wang A, Pei Y, Tian J, Wang H, Hsam SLK, Zeller FJ (2006) Cloning and molecular characterization of three novel LMW-i glutenin subunit genes from cultivated einkorn (*Triticum monococcum* L.). *Theor Appl Genet* 113:383–395
- Andon NL, Hollingworth S, Koller A, Greenland AJ, Yates JR III, Haynes PA (2002) Proteomic characterization of wheat amyloplasts using identification of proteins by tandem mass spectrometry. *Proteomics* 2:1156–1168
- Balmer Y, Vensel WH, DuPont FM, Buchanan BB, Hurkman WJ (2006) Proteome of amyloplasts isolated from developing wheat endosperm presents evidence of broad metabolic capability. *J Exp Bot* 57:1591–1602
- Broza YY, Danin-Poleg Y, Lerner L, Broza M, Kashi Y (2007) *Vibrio vulnificus* typing based on simple sequence repeats: insights into the biotype 3 group. *J Clin Microbiol* 45:2951–2959
- Butler JM, Buel E, Crivellente F, McCord BR (2004) Forensic DNA typing by capillary electrophoresis using the ABI Prism 310 and 3100 genetic analyzers for STR analysis. *Electrophoresis* 25:1397–1412
- 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
- D'Ovidio R (1993) Single-seed PCR of LMW glutenin genes to distinguish between durum wheat cultivars with good and poor technological properties. *Plant Mol Biol* 22:1173–1176
- D'Ovidio R, Masci S (2004) The low-molecular-weight glutenin subunits of wheat gluten. *J Cereal Sci* 39:321–339
- D'Ovidio R, Porceddu E (1996) PCR-based assay for detecting 1B-genes for low molecular weight glutenin subunits related to

- gluten quality properties in durum wheat. *Plant Breed* 115:413–415
- D’Ovidio R, Marchitelli C, Ercoli Cardelli L, Porceddu E (1999) Sequence similarity between allelic *Glu-B3* genes related to quality properties of durum wheat. *Theor Appl Genet* 98:455–461
- Dong LL, Zhang XF, Liu DC, Fan HJ, Sun JZ, Zhang ZJ, Qin HJ, Li B, Hao ST, Li ZS, Wang DW, Zhang AM, 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(10):e13548. doi: [10.1371/journal.pone.0013548](https://doi.org/10.1371/journal.pone.0013548)
- Dworschak RG, Ens W, Standing KG, Preston KR, Marchylo BA, Nightingale MJ, Stevenson SG, Hatcher DW (1998) Analysis of wheat gluten proteins by matrix-assisted laser desorption/ionization mass spectrometry. *J Mass Spectrom* 33:429–435
- Gupta RB, Shepherd KW (1990) Two-step one-dimensional SDS-PAGE analysis of LMW subunits of glutelin. 1. Variation and genetic control of the subunits in hexaploid wheats. *Theor Appl Genet* 80:65–74
- Gupta RB, Singh NK, Shepherd KW (1989) The cumulative effect of allelic variation in LMW and HMW glutenin subunits on dough properties in the progeny of two bread wheats. *Theor Appl Genet* 77:57–64
- Gupta RB, Békés F, Wrigley CW (1991) Prediction of physical dough properties from glutenin subunit composition in bread wheats: correlation studies. *Cereal Chem* 68:328–333
- Harberd NP, Bartels D, Thompson RD (1985) Analysis of the gliadin multigene loci in bread wheat using nullisomic–tetrasomic lines. *Mol Gen Genet* 198:234–242
- 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
- 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
- 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
- Liu L (2008) Identification of glutenin composition in common wheat by proteomics technology. Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing
- 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 AL, Yan YM, He ZH (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*. doi: [10.1186/1471-2229-10-124](https://doi.org/10.1186/1471-2229-10-124)
- Long H, Wei YM, Yan ZH, Baum B, Nevo E, Zheng YL (2005) Classification of wheat low-molecular-weight glutenin subunit genes and its chromosome assignment by developing LMW-GS group-specific primers. *Theor Appl Genet* 111:1251–1259
- Maguire TL, Peakall R, Saenger P (2002) Comparative analysis of genetic diversity in the mangrove species *Avicennia marina* (Forsk.) Vierh. (Avicenniaceae) detected by AFLPs and SSRs. *Theor Appl Genet* 104:388–398
- Mamone G, De Caro S, Di Luccia A, Addeo F, Ferranti P (2009) Proteomic-based analytical approach for the characterization of glutenin subunits in durum wheat. *J Mass Spectrom* 44:1709–1723
- Margiotta B, Colaprico G, D’Ovidio R, Lafiandra D (1993) Characterization of high *Mr* subunits of glutenin by combined chromatographic (RP-HPLC) and electrophoretic separations and restriction fragment length polymorphism (RFLP) analyses of their encoding genes. *J Cereal Sci* 17:221–236
- Masci S, D’Ovidio R, Lafiandra D, Kasarda DD (1998) Characterization of a low-molecular-weight glutenin subunit gene from bread wheat and the corresponding protein that represents a major subunit of the glutenin polymer. *Plant Physiol* 118:1147–1158
- Metakovsky EV, Wrigley CW, Békés F, Gupta RB (1990) Gluten polypeptides as useful genetic markers of dough quality in Australian wheats. *Aust J Agric Res* 41:289–306
- Osborne TB (1924) *The vegetable proteins*. Longmans Green and Co., London
- 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
- Rosenblum BB, Oaks F, Menchen S, Johnson B (1997) Improved single-strand DNA sizing accuracy in capillary electrophoresis. *Nucleic Acids Res* 25:3925–3929
- 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
- Saghai-Maroo MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. *Proc Natl Acad Sci USA* 81:8014–8018
- Shewry PR, Tatham AS, Barro F, Barcelo P, Lazzari P (1995) Biotechnology of breadmaking: unraveling and manipulating the multi-protein gluten complex. *Nat Biotechnol* 13:1185–1190
- Singh NK, Shepherd KW (1988) Linkage mapping of genes controlling endosperm storage proteins in wheat. 1. Genes on the short arms of group 1 chromosomes. *Theor Appl Genet* 75:628–641
- Singh NK, Shepherd KW, Cornish GB (1991) A simplified SDS-PAGE procedure for separating LMW subunits of glutenin. *J Cereal Sci* 14:203–208
- Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596–1599
- 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
- Wang LH, Li GY, Peña 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
- Weegels PL, Hamer RJ, Schofield JD (1996) Functional properties of wheat glutenin. *J Cereal Sci* 23:1–17
- Zhang W, Gianibelli MC, Rampling LR, 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
- 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
- Zianni M, Tessanne K, Merighi M, Laguna R, Tabita FR (2006) Identification of the DNA bases of a DNase I footprint by the use of dye primer sequencing on an automated capillary DNA analysis instrument. *J Biomol Tech* 17:103–113