Theor Appl Genet (2005) 111: 1251–1259
DOI 10.1007/s00122-005-0024-1

ORIGINAL PAPER

Hai Long · Yu-Ming Wei · Ze-Hong Yan
Bernard Baum · Eviatar Nevo · You-Liang Zheng

Classification of wheat low-molecular-weight glutenin subunit genes and its chromosome assignment by developing LMW-GS group-specific primers

Received: 5 May 2005 / Accepted: 20 June 2005 / Published online: 27 September 2005

Abstract On the basis of sequence analysis, 69 known low-molecular-weight glutenin subunit (LMW-GS) genes were experimentally classified into nine groups by the deduced amino acid sequence of the highly conserved N-terminal domain. To clarify the chromosomal locations of these groups, 11 specific primer sets were designed to carry out polymerase chain reactions (PCR) with the genomic DNA of group 1 ditelosomic lines of *Chinese Spring*, among which nine primer sets proved to be LMW-GS group-specific. Each group of LMW-GS genes was specifically assigned on a single chromosome arm and hence to a specific locus. Therefore, these results provided the possibility to predict the chromosome location of a new LMW-GS gene based on its deduced N-terminal sequence. The validity of the classification was confirmed by the amplifications in 27 diploid wheat and *Aegilops* accessions. The length polymorphisms of LMW-GS genes of groups 1 and 2, and groups 3 and 4.1 were detected in diploid A-genome and S-genome accessions, respectively. The diploid wheat and *Aegilops* species could be used as valuable resources of novel allele variations of LMW-GS gene in the improvement of wheat quality. The nine LMW-GS group-specific primer sets could be utilized to select specific allele variations of LMW-GS genes in the marker-assisted breeding.

Keywords *Aegilops* · Alignment analysis · Length polymorphism · N-terminal domain · *Triticum aestivum*

Electronic Supplementary Material Supplementary material is available for this article at http://dx.doi.org/10.1007/s00122-005-0024-1

Communicated by B. Friebe

Hai Long and Yu-Ming Wei are the two authors who have contributed equally to this paper

H. Long · Y.-M. Wei · Z.-H. Yan · Y.-L. Zheng (✉)
Triticeae Research Institute, Sichuan Agricultural University,
Dujiangyan, Sichuan, 611830, China
E-mail: ylzheng@sicau.edu.cn
Tel.: +86-835-2882007
Fax: +86-835-2883153

B. Baum
Agriculture and Agri-Food Canada,
Eastern Cereal and Oilseed Research Centre,
Ottawa, K1A 0C6, ON, Canada

E. Nevo
Institute of Evolution, University of Haifa,
Mt. Carmel, Haifa, 31905, Israel

Introduction

Wheat (*Triticum aestivum* L.) quality is largely determined by the seed storage proteins present in the endosperm of the grain (Shewry and Halford 2002). Glutenins mainly consist of two types of subunits: high-molecular-weight glutenin subunits (HMW-GS) and low-molecular-weight glutenin subunits (LMW-GS), which are held together by intermolecular disulphide bonds to form the glutenin polymer. The HMW-GS represent approximately 10% of the total seed storage proteins. Their important role has been well established (Shewry et al.1992, 1995). LMW-GS represent about one-third of total seed storage proteins and 60% of glutenins (Bietz and Wall 1973). They are classically subdivided into B, C, and D groups according to their electrophoretic mobility in SDS-PAGE and their isoelectric points (Jackson et al. 1983).

Coding genes of the LMW-GS are at *Glu-A3*,*Glu-B3*, and *Glu-D3* loci, and located on the short arms of chromosome 1A, 1B, and 1D, respectively (Singh and Shepherd 1988). The estimates of gene copy number varied from 10–15 (Harberd et al. 1985) to 35–40 (Cassidy et al. 1998; Sabelli and Shewry 1991). Over the past few years, more efforts have been focused on analysis at the molecular level. To date, more than 90 cDNA and genomic DNA clones of LMW-GS gene have been reported, and the gene structure has been well established. On the basis of the sequence data, Van Campenhout et al. (1995) designed several primer sets specific for each of the *Glu-3* loci and successfully determined the

chromosomal locations of five LMW-GS genes. D'Ovidio et al. (1997) and Zhang et al. (2004) developed primer sets specific for *Glu-B3* locus and *Glu-A3* alleles, respectively. Ikeda et al. (2002) isolated several LMW-GS gene clones from a soft wheat cultivar and classified them into 12 groups based on the N- and C-terminal sequences. Some of them were assigned to specific chromosomes. However, the comprehensive gene classification and chromosomal assignment focusing on this complex multigene family is not well established.

The objective of the present study was to experimentally classify the LMW-GS genes comprehensively based on the known sequences and to clarify their exact chromosomal locations by performing PCR in Group 1 ditelosomic lines of *Chinese Spring*, henceforth CS, with LMW-GS group-specific primer sets. A total of 27 diploid wheat and *Aegilops* accessions, consisting of four closely related species to cultivated wheat, were used to perform further verification of the validity of the classification and the specificity of the primers.

Materials and methods

Plant material

Group 1 ditelosomic lines of *CS* were kindly provided by Dr. D.C. Liu (Triticeae Research Institute of Sichuan Agricultural University, China), which were originally obtained from Prof. E.R. Sears. A total of 27 diploid wheat and *Aegilops* accessions were used to verify the group-specific primer sets (Table 1). There were ten *T. monococcum* ($2n=2$ $x=14$, A^m A^m), eight *T. urartu* ($2n=2x=14$, A^u A^u), five *A. tauschii* ($2n=2$ $x=14$, DD), and four *Ae. speltoides* ($2n=2$ $x=14$, SS) accessions, whose genomes are closely related to A-, B-, and D-genome of the cultivated wheat (*T. aestivum*, $2n=6$ $x=42$, AABBDD), respectively. These diploid wheat and *Aegilops* accessions were kindly supplied by USDA-ARS (Table 1).

LMW-GS gene sequences and analysis

The LMW-GS gene sequences in this study were retrieved from the GenBank database (http://www.ncbi.nlm. nih.gov/Genbank). The coding regions of LMW-GS genes were used for nucleotide and amino acids alignment using CLUSTAL X program (ver.1.83, Thompson et al. 1997). MEGA program (ver.2, Kumar et al. 2001) was used to setup phylogenetic trees based on clustering of deduced amino acid sequences of N-terminal domain using Neighbor-joining method (Saitou and Nei 1987).

DNA isolation and PCR amplification

Genomic DNA was isolated from 3- to 5-day-old seedlings with cetytrimethylammonium bomide (CTAB) procedure as reported by Murray and Thompson (1980). Polymerase chain reaction (PCR) amplifications were performed in 25µl reaction volume, consisting of 1U *Taq* DNA polymerase (*TaKaRa*), 2.5µl PCR buffer (supplied with *Taq* DNA polymerase), 50 ng genomic DNA, 1.5 mM $MgCl_2$ and 100 mM of each dNTP. PCR amplifications were conducted according to the following program: 95°C for 5 min denaturation followed by 35 cycles of 45 s at 95°C, 45 s at 50–70°C (depending on the primer sets, see Table 3), and 45 s at 72°C. PCR products were separated in 2% agarose gels.

Results

Sequence analysis and gene classification

More than 90 complete and partial LMW-GS gene sequences from *T. aestivum* (AABBDD), *T. durum* (AABB), *T. dicoccoides* (AABB), *T. monococcum* (AA), and *Ae. tauschii* (DD) were retrieved from GenBank. The sequences of 5′ and 3′ non-translated regions were eliminated to concentrate on the coding region alone, and several ultra-short sequences and all pseudogenes were removed. The deduced amino acids of 69 LMW-GS gene sequences that were left were used for alignment using CLUSTAL X (not shown, alignment available on request or as supplement to this paper as "Supplement Fig."). Phylogenetic analysis based on the highly conserved N-terminal domain indicated that the 69 sequences could be divided into seven groups by the first eight amino acid residues in the N-terminal domain (Table 2 and Fig.1a).

Table 1 Diploid wheat species analyzed in this study

Species	Accessions
T. monococcum	PI428149, PI428150, PI428152, PI428154, PI428156, PI428171, PI428173, PI428174, PI428175, PI428176
T. urartu	PI428183, PI487270, PI428221, PI428269, PI428208, PI428197, PI487268, PI428328
Ae. speltoides	AS928*, PI560751, PI542269, PI542256
Ae. tauschii	AS63*, PI499262, PI486266, PI48272, PI508362

Accessions marked by "*" were supplied by Triticeae Research Institute of Sichuan Agricultural University, China

Table 2 Classification of LMW-GS genes based on deduced amino acid sequence of N-terminal domain

Groups		N-terminal sequence	Accession	Reference
1		ISQQQ-	AB062876-78, AJ293097-98*, AY146588, AY453154-60	Ikeda et al. (2002), Wicker et al. (2003), Zhang et al. (2004)
			AY214456*, X07747, U86030	Pitts et al. (1988), Cassidy et al. (1998)
2		MDTSCIP-	AB062868-71, AJ293099*	Ikeda et al. (2002)
			X62588	D'Ovidio et al. (1992)
3		MENSHIP-	Y18159, Y17845, AB062853-62	Masci et al. (1998), D'Ovidio et al. (1999), Ikeda et al. (2002)
			AB119006, AB119007	Maruyama-Funatsuki et al. (2005)
4	4.1	METSHIPG-	Y14104, AB062852 AJ537508*	D'Ovidio et al. (1997), Ikeda et al. (2002)
	4.2	METSHIPS-	AB062851, AY585354	Ikeda et al. (2002), Johal et al. (2004)
5		METSRV-	AB062865-67, AY585355, AY841013-14, AY214450*	Ikeda et al. (2002), Johal et al. (2004), Huang et al. (2005)
			AY299457	Wang et al. (2005)
6		METRCIP-	X13306, AB062875, AY263369, AY299485, AY585356	Colot et al. (1989), Ikeda et al. (2002), Zhao et al. (2004), Wang et al. (2005), Johal et al. (2004)
			U86027, U86029, AJ519835	Cassidy et al. (1998), Chardot et al. (2002)
7	7.1	METSCIP-	X51759, X84959, U86026, AB062873-74,, AY299458, AY695380*, AY748826*	Cassidy and Dvorak (1991), Van Campenhout et al. (1995), Cassidy et al. (1998), Ikeda et al. (2002), Wang et al. (2005)
	7.2	METSCIS-	AB062872, AY585350	Ikeda et al. (2002), Johal et al. (2004)
			U86028, AY296753	Cassidy et al. (1998), Long et al. (2004)

"*" directly submitted to database

Chromosome assignment of each group of LMW-GS genes by group-specific primer sets

A specific-primers-based PCR assay was carried out to identify the chromosome locations of each group of LMW-GS genes. Seven specific primer sets were designed. The forward primers were at the beginning of N-terminal domain, and the reverse primers were in the C-terminal domain. Therefore, the PCR products should contain the whole N-terminal and repetitive domain. Owing to the high sequence identity in conserved domains of LMW-GS genes, an additional primer-template mismatched base was added at position 3 of 3′end of a few primer sequences (See detail in Table 3) following the strategy described by Kwok et al. (1990). Genomic DNA of six group 1 ditelosomic lines of *CS* was used to carry out the specific PCR amplifications. Therefore, each primer set should yield uniformly sized products in five of six ditelosomic lines. Furthermore, the chromosomal location would be consequently recognized by the absence of PCR products from one particular ditelosomic line which lost specific chromosome arms (Fig. 2).

Five out of the seven groups of LMW-GS gene were successfully assigned to specific chromosome arms (Table 3). The genes of LMW-GS groups 1 (N-terminal sequence: ISQQQ-) and 2 (MDTSCIP-) were assigned on the short arm of chromosome 1A. The genes of LMW-GS group 3 (MENSHIP-) were located on the chromosome arm 1BS, and the genes of LMW-GS groups 5 (METSRV-) and 6 (METRCIP-) were assigned to the short arm of chromosome 1D (Fig. 2). However, the primers specific to LMW-GS group 4 (METSHIP-) resulted in two close bands, about 620 and 580 bp in ditelosomic lines Dt1AS, Dt1AL, Dt1BS, and Dt1DS, whereas only one different band resulted in ditelosomic lines Dt1BL and Dt1DL (Fig. 2). Similarly, LMW-GS group 7–specific primers yielded two PCR fragments in all ditelosomic lines except Dt1AL and Dt1DL, in which only a single band was present.

These results suggest that both LMW-GS groups 4 and 7 could consist of two subgroups with high sequence similarity and located on different chromosome arms. A separated phylogenetic analysis of the N-terminal sequences revealed that LMW-GS groups 4 and 7 could each be subdivided into two subgroups: groups 4.1 and 4.2, and groups 7.1 and 7.2 (Fig 1b, c), with two and one amino acid residue difference in the N-terminal domain, respectively. On the basis of these results, four additional primer sets were designed for each subgroup, and the chromosomal locations of the four subgroups were consequently determined. The genes of subgroups 7.1 and 4.1 were assigned to the short arms of chromosome 1A and 1B, respectively, and the genes of subgroups 7.2 and 4.2 were both located on the chromosome arm of 1DS.

Variations in diploid wheat and *Aegilops* species

To obtain further verification of the classification of LMW-GS genes to groups and their chromosome assignment, the nine group-specific primer sets were tested in 27 diploid wheat and *Aegilops* accessions, whose genomes are closely related to the A-, B-, and D-genomes of the cultivated wheat.

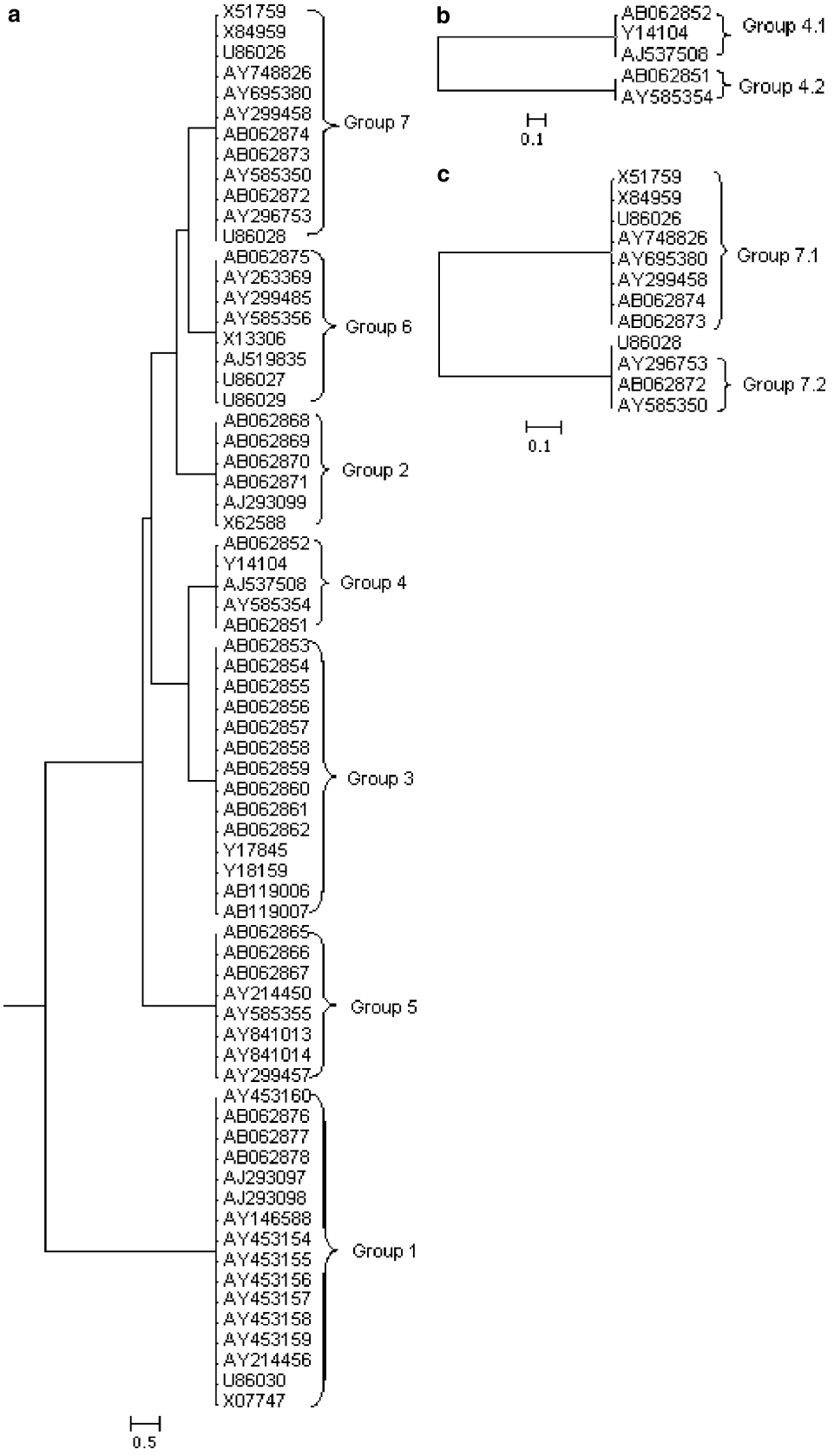

Fig. 1 Phylogenetic trees of LMW-GS genes. Total of 69 sequences were originally classified into seven groups as showed in (**a**). Four subgroups were additionally identified by the separated analysis of the genes of groups 4 and 7 (**b**) and (**C**)

Table 3 Group-specific primer sets and chromosome locations of each group of LMW-GS genes

Group	Marker	Primer	Sequence (5′-3′)	Product size (bp)	Chromosome location	Annealing temperature (°C)
1	Glu3A.1	P1F	GCCGTTGCGCAAATTTCACAG	450 and 600	1AS	55
		P1R	AACAGATGGATGAATAACTGGTAT			
2	Glu3A.2	P2F	AGTGCCATTGCGCAGATGAAT	350	1AS	60
		P2R	AACGGATGGTTGAACAATAGA			
3	Glu3B.1	P3F	GCACAAATGGAGAATAGCCAC	500	1BS	59
		P3R	AACAAATGGTATTTGTTGTTG			
4	Glu3.1	P4F	ATGGAGACTAGCCACATCCCT	620 and 580	1BS, 1DS	61
		P4R	CACATGGCAACTACTCTGCCA			
4.1	Glu3B.2	P5F	CCTAGCTTGGAGAAACCATT	450	1BS	50
		P5R	CAAGATAGATGGCTGAATAG			
4.2	Glu3D.1	P6F	CCTGGCTTGGAGAAACCATC	500	1DS	50
		P6F	CAAGATAGATGGCTGAATAT			
5	Glu3D.2	P7F	ATGGAGACTAGCCGCGTCCCT	540	1DS	69.5
		P7R	TGACCTAGCAAGACGTTGCGA			
6	Glu3D.3	P8F	ATGGAGACTAGATGCATCCCT	600	1DS	60
		P8R	AGATTGGATGGAACCCTGAAC			
7	Glu3.2	P9F	TGCCATTGCACAGATGCAG	680 and 700	1AS, 1DS	54
		P8R	CTGCAAAAAGGTACCCTT			
7.1	Glu3A.3	P10F	ATGGAGACTAGCTGCATCC	680	1AS	60
		P10R	CTGCAAAAAGGTACCCTTTT			
7.2	Glu3D.4	P11F	ATGGAGACTAGCTGCATCT	700	1DS	52
		P11R	CTGCAAAAAGGTACCCTGTA			

Underlined represents additional mismatched bases. *N-ter* N-terminal domain

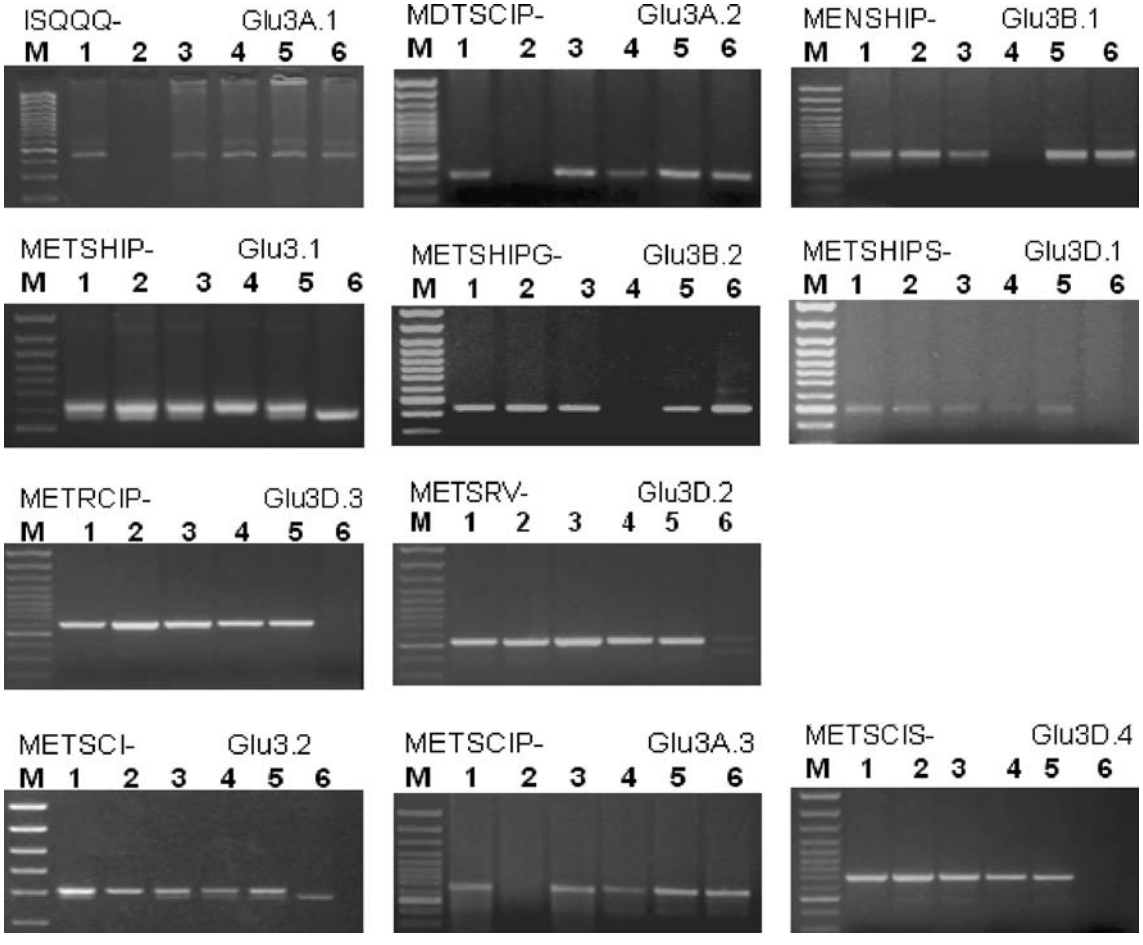

Fig. 2 The amplification products from group 1 ditelosomic lines of *Chinese Spring* (*1* Dt1AS; *2* Dt1AL; *3* Dt1BS; *4* Dt1BL; *5* Dt1DS and *6* Dt1DL) using LMW-GS group-specific primer sets. The fragment size was listed in Table 3. *M* = 3-kb DNA Ladder (*MBI*)

LMW-GS genes of groups 1, 2, and 7.1 were located on chromosome 1AS, and the corresponding primer sets Glu3A.1, Glu3A.2, and Glu3A.3 failed to yield any amplification products in the accessions of *Ae. speltoides* and *Ae. tauschii*. The length polymorphism of PCR products were detected in both *T. monococcum* and *T. urartu* accessions (Fig. 3). There were three to four amplified bands in each *T. monococcum* accession with fragment lengths ranging from 480 and 600 bp, resulting in five different patterns in ten *T. monococcum* accessions (Fig.3a). In *T. urartu*, two to three bands with fragment length between 430 and 600 bp were detected in each accession, resulting in four patterns in eight accessions by primer set Glu3A.1. A 340 bp fragment of the same size as that found in the ditelosomic lines was detected in seven out of ten *T. monococcum* accessions by primer set Glu3A.2. Two new alleles with approximate lengths of 350 and 320 bp were found in three other accessions: PI428174, PI428176 and PI428171 (Fig. 3b). It is noteworthy that PCR products were absent from all the eight accessions of *T. urartu*, which has been widely accepted as the A-genome donor of cultivated wheat by primer Glu3A.2. Therefore, the LMW-GS genes of group 2 might have originated from *T. monococcum*, which is closely related to *T. urartu*. However, more evidence needs be obtained to support this hypothesis. LMW-GS group 7.1–specific PCR product was present in both *T. monococcum* and *T. urartu* accessions, and its size was identical to that of cultivated wheat.

Primer sets Glu3B.1 and Glu3B.2 were specific to the LMW-GS genes of groups 3 and 4.1, respectively, which were both assigned to chromosome 1BS. No product was obtained from any accession of the A-genome and D-genome diploid species with the two primer sets. However, two bands were obtained from four *Ae. speltoides* accessions with the Glu3B.1 primer set, resulting in three allele variants. A 530 bp fragment was specifically amplified from accession AS928 (Fig. 3c). Primer set Glu3B.2 amplified a uniform 450 bp band from three *Ae. speltoides* accessions, which was similar to that in *CS* ditelosomic lines, whereas a 360 bp fragment was obtained from accession AS928 (Fig. 3d). Genes of LMW-GS group 4.2, 5, 6, and 7.2 were all located on 1DS, and the corresponding primer sets only amplified the products with no length polymorphism from *Ae. tauschii* accessions.

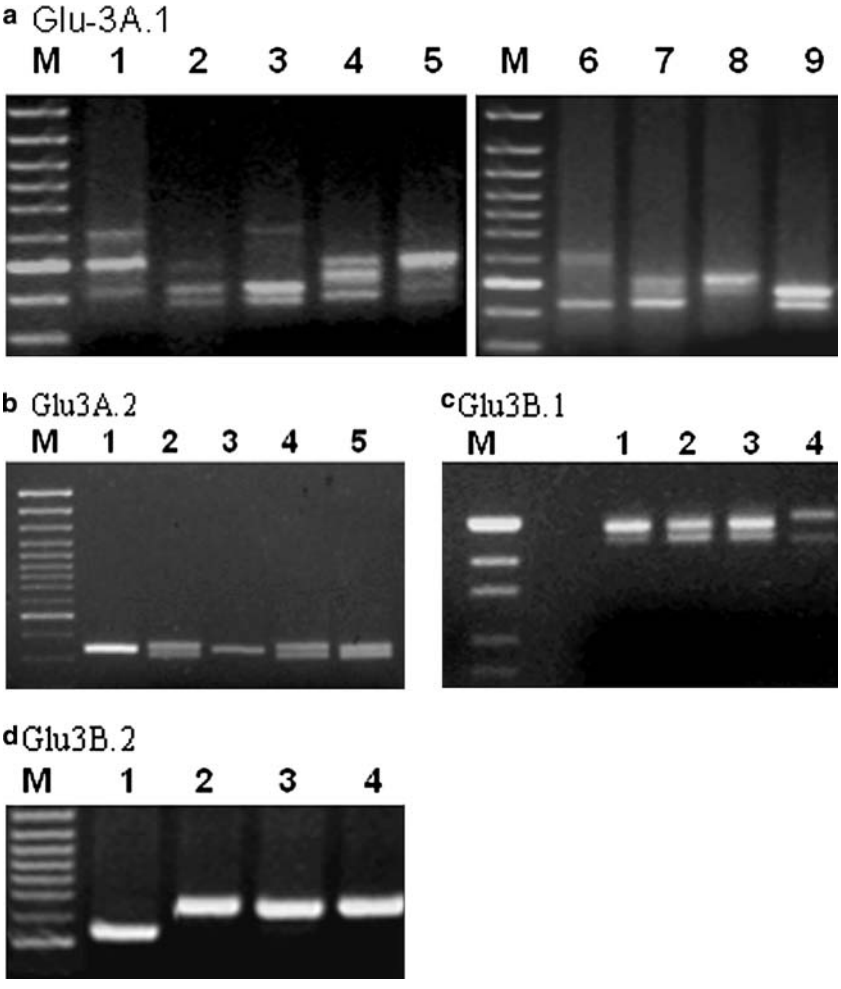

Fig. 3 Length polymorphism of LMW-GS genes in diploid wheat and *Aegilops* accessions detected by primer sets Glu3A.1, Glu3A.1.2, Glu3B.1 and Glu3B.2. **a** *T. monococcum* accessions *1* PI428175, *2* PI428156, *3* PI428152, *4* PI428149 and *5* PI428149; *T. urartu* accessions *6* PI428183, *7* PI487270, *8* Pi428221 and *9* PI428197. **b**: *1* PI428173; *2* PI428174; *3* PI428175; *4* PI428176; *5* PI428171. **c** *Ae. speltoides* accessions *1* PI560751, *2* PI542269, *3* PI542256, *4* AS928. **d** *Ae.speltoides* accessions *1* AS928; *2* PI542269; *3* PI542256; *4* PI56075. *M* = DNA marker

Discussion

Wheat LMW-GS play an important role in determining dough quality. However, the contribution of individual LMW-GS to wheat quality is not fully understood. This is in part due to the complexity of LMW-GS gene family. In recent years, more attention has been paid to the isolation of LMW-GS genes from cultivated wheat and its wild relatives. More than 90 LMW-GS gene sequences have been reported. However, a classification of LMW-GS gene into groups, in such a complex multigene family, has yet to be established. Ikeda et al. (2002) isolated several LMW-GS genes and classified them into 12 groups based on the N- and C-terminal sequences. This was the first relatively comprehensive research on classifying the composition of the LMW-GS gene. However, their work was limited by the shortage of sequence information.

In the present study, sequence alignment analysis of 69 genes, derived from hexaploid, tetraploid, and their closely related diploid wheat species (*Triticum* and *Aegilops*), was performed and it revealed a relatively high degree of homology between the LMW-GS genes (data not shown). Most of LMW-GS genes contained a short (13 amino acid residues) and highly conserved N-terminal domain. In contrast, the C-terminal domain comprised over two-thirds of the polypeptide length (Lee et al. 1999) and was more variable than the N-terminal domain. We classified the LMW-GS genes into seven groups (Table 2 and Fig. 1) based on the deduced amino acid sequences of N-terminal domain, and came to different results from Ikeda et al. (2002). A marker-based PCR was employed to assign the chromosome location(s) of each LMW-GS group. Our results indicated that five out of the seven groups of LMW-GS genes were successfully assigned to specific chromosome arms, whereas the LMW-GS genes of groups 4 and 7 were located on two different chromosome arms. Moreover, two subgroups with relatively high-sequence identity were identified from LMW-GS groups 4 and 7, respectively. In all, nine groups and subgroups were identified from the 69 LMW-GS genes. Three, two and four groups or subgroups were assigned to the short arms of chromosome 1A, 1B, and 1D at *Glu-A3, Glu-B3*and *Glu-D3* loci, respectively. The amplification results from diploid wheat and *Aegilops* accessions were in agreement with those in the CS ditelosomic lines, although length variations were additionally detected. These results gave further support to the validity of the classification results of the LMW-GS genes. Therefore, the results in this study provided the possibility to predict the chromosome location of a new LMW-GS gene based on its deduced N-terminal sequence. The classification based on the sequence similarities of a short and conserved domain could simplify the characterization of LMW-GS sequences.

In this study, a total of 11 sets of specific primers were utilized for chromosome assignment analysis, and nine of them proved to be LWM-GS group-specific. These primers anneal to the beginning of the conserved N-terminal domain and the C-terminal domain, so that the amplified products contained the whole repetitive domain of LMW-GS gene. All nine primer sets resulted in uniform amplification products in the CS group 1 ditelosomic lines, whereas length polymorphisms were detected in the A- and S-genome accessions. This was probably due to the limited number of genotypes of diploid species from which cultivated wheat was derived. Therefore, the wild and cultivated relatives of bread wheat containing novel alleles of LMW-GS could be utilized as a valuable source of new gene variation with the potential to improve wheat grain properties. The primers reported in the present study could be used for selecting specific LMW-GS genes in the breeding programs.

A schematic model of structure comparison of nine LMW-GS groups was constructed (Fig. 4) based on the alignment of the deduced amino acid sequences of LMW-GS genes. The general structure of the nine LMW-GS groups shares a very high degree of similarity, as reported by Cassidy et al. (1998). However, we now

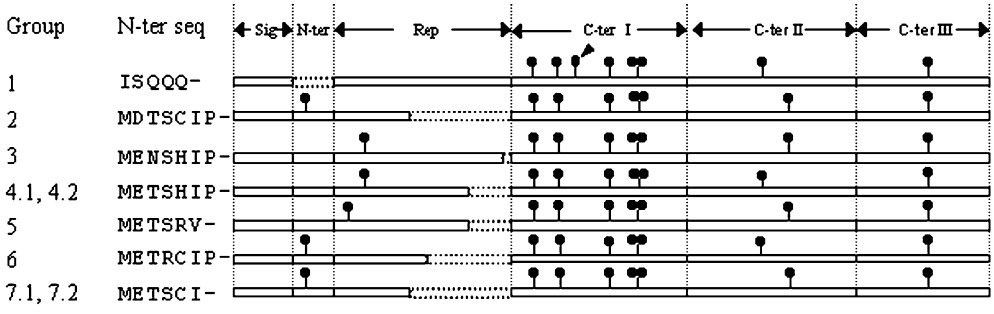

Fig. 4 Schematic model of the structure comparison of each group of LMW-GS genes. Cysteine residues and their locations were shown as *bullethead*. A unique cysteine (indicated by *arrowhead*) appears in the LMW-GS genes of group 1, and differs from other groups by its location. *Sig* Signal peptide; *N-ter* N-terminal domain; *Seq* sequence; *Rep* repetitive domain; *C-ter* C-terminal domain. *Solid bar* represents nucleotide sequences of LMW-GS genes while *empty bar* indicates deletions in respect of group 1 genes

present additional features specific to each LMW-GS group, as shown in Fig. 4. A short N-terminal domain following an extremely conserved signal peptide is a common feature in nearly all the typical LMW-GS genes. Six main types of deduced amino acid sequence of N-terminal domain were established. Each type is composed of 13 amino acid residues with a methionine at its first position. However, LMW-GS group 1, which was denominated as LMW-i type (Cloutier et al. 2001), is distinguished from others by its lack of typical N-terminal domain and starts directly with the repetitive domain, beginning with ISQQQ.

The repetitive domain of LMW-GS has been known for its high glutamine content and variability. The number of repeats present in this domain is mainly responsible for length variation of LMW-GS and for the general hydrophilic character as well (D'Ovidio et al. 1999, and also review in D'Ovidio and Masci 2004). The length polymorphism between nine LMW-GS groups resulted from the alignment of the 69 sequences. The longest repetitive domain was observed in LMW-GS groups 1 and 3, containing about 25 repeats and nearly 190 amino acid residues. LMW-GS groups 4.1, 4.2, and 5 have shorter repetitive domains compared to those of LMW-GS groups 1 and 3. They are composed of about 150 and 135 amino acid residues, respectively, and could be divided into about 20 repeats. LMW-GS group 5 contains about 100 amino acid residues in its repetitive domain, including approximately 15 repeats. In LMW-GS group 2 which features only 13 repeats (less than 90 amino acid residues), the length of repetitive domain is similar to groups 7.1 and 7.2, was the shortest among the nine LMW-GS groups.

The length variation of the repetitive domain does also exist within LWM-GS groups (data not shown). This was probably caused by unequal crossing-over and/or slippage during replication as suggested for the evolution of other prolamins (Shewry et al. 1989; D'Ovidio and Masci 2004). But another cause that does not need to be ignored is that some cloned genes may contain deletions of the repetitive domain that might have occurred during the cloning process (Masci et al. 1998).

Disulphide bonds play a key role in determining the structure and properties of wheat gluten proteins (reviewed in Shewry and Tatham 1997). Kasarda (1989) proposed that the LMW-GS genes have eight cysteines: six are involved in intramolecular disulphide bonds and two are free to form intermolecular disulphide bonds. The number and positions of cysteines in LMW-GS were summarized here and were found to be similar to the results of Ikeda et al. (2002). All nine LMW-GS groups of LMW-GS genes contain eight cysteines, among which six (five in C-terminal domain I and one in the C-terminal domain III) are extremely conserved in all LMW-GS genes. They may form intramolecular disulphide bonds. The remaining two cysteines are distributed in six patterns (illustrated in Fig. 4), in which LMW-GS groups 2, 7.1, and 7.2, and groups 4.1 and 4.2 share the same positions, respectively. One of the two free cysteines is present in C-terminal domain II and is distributed in two varying positions. The another one is distributed in three different domains among nine LMW-GS groups: groups 2, 6, 7.1, and 7.2 have a free cysteine at position 5 in N-terminal domain; in group 5, the free cysteine is present in the repetitive domain and closer to N-terminal domain than that in groups 3, 4.1, and 4.2; in group 1, it is present in C-terminal domain I and is therefore different from all other groups. Whether or not different distributions of cysteines have a particular effect on the structure of the glutenin polymer, the end-using quality of the wheat grain might merit further investigation.

Acknowledgements The seed accessions with PI numbers were kindly provided by Dr. Bockelman of the American Germplasm Resources Information Network (GRIN). We wish to thank the two anonymous reviewers for helpful comments that improved the manuscript. This work was supported by the National High Technology Research and Development Program of China (863 program 2003AA207100), National Natural Science Foundation of China (No 30300219), and the FANEDD project (200357 and 200458) from Ministry of Education, China. Y.-L. Zheng was supported by the Program for Changjiang Scholars and Innovative Research Teams in University of China (IRT0453).

References

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

Cassidy BG, Dvorak J (1991) Molecular characterization of a low-molecular-weight glutenin cDNA clone from *Triticum durum*. Theor Appl Genet 81:653–660

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 Gene 96:743–750

Chardot T, Do T, Perret L, Lauriere M (2002) Heterogeneity of genes encoding the low molecular weight glutenin subunits of bread wheat. In: Renard D, Delle Valle G, Popineau Y (eds) Plant biopolymer science: food and non-food applications. Royal Society of Chemistry, Cambridge, pp 24–30

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

Colot V, Bartels D, Thompson R, Flavell R (1989) Molecular characterization of an active wheat LMW glutenin gene and its relation to other wheat and barley prolamin genes. Mol Gen Genet 216:81–90

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

D'Ovidio R, Tanzarella OA, Porceddu E (1992) Nucleotide sequence of a low- molecular-weight glutenin from *Triticum durum*. Plant Mol Biol 18:781–784

D'Ovidio R, Simeone M, Masci S, Porceddu E (1997) Molecular characterization of a LMW-GS gene located on chromosome 1B and the development of primers specific for the *Glu-B3* complex locus in durum wheat. Theor Appl Genet 95:1119–1126

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

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 Z, Long H, Yan ZH, Wei YM, Zheng YL (2005) Sequence analysis of LMW-GS genes from *Aegilops tauschii* containing HMW glutenin subunits 5 + 12 with good quality. High Technol Lett 15(7):67–72

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

Jackson EA, Holt LM, Payne PI (1983) Characterization of high-molecular-weight gliadin and low-molecular-weight glutenin subunits of wheat endosperm by two-dimensional electrophoresis and the chromosomal localization of their controlling genes. Theor Appl Genet 66:29–37

Johal J, Gianibelli MC, Rahman S, Morell MK, Gale KR (2004) Characterization of low-molecular-weight glutenin genes in *Aegilops tauschii*. Theor Appl Genet 109:1028–1040

Kasarda DD (1989) Glutenin structure in relation to wheat quality. In: Pomeranz Y (ed) Wheat is unique. Am Assoc Cereal Chem, St Paul, pp 277–302

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

Kumar S, Tamura K, Jakobsen IB, Nei M (2001) *MEGA2*: molecular evolutionary genetics analysis software. Arizona State University, Tempe

Kwok S, Kellogg DE, McKinney N, Spasic D, Goda L, Levenson C, Sninsky JJ (1990) Effects of primer-template mismatches on the polymerase chain reaction: human immunodeficiency virus type 1 model studies. Nucleic Acids Res 18:999–1005

Lee Y-K, Ciaffi M, Appels R, Morell MK (1999) The low-molecular-weight glutenin subunit protein of primitive wheat. II. The genes from A-genome species. Theor Appl Genet 98:126–134

Long H, Wei YM, Yan ZH, Zheng YL (2004) Molecular cloning of a novel low- molecular- weight glutenin subunit gene from wheat (*Triticum aestivum* L.), variety "chuannong 16". Acta Agronomica Sinica 30:1179–1184

Maruyama-Funatsuki W, Takata K, Funatsuki H, Tabiki T, Ito M, Nishio Z, Kato A, Saito K, Yahata E, Saruyama H, Yamauchi H (2005) Identification and characterization of a novel LMW-s glutenin gene of a Canadian Western extra-strong wheat. J Cereal Sci 41:47–57

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 Physio 118:1147–1158

Murray M, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res 8:4321–4325

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 P, Shewry PR (1991) Characterization and organization of gene families at the *Gli-1* loci of bread and durum wheat. Theor Appl Genet 83:428–434

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Ecol 4:406–425

Shewry PR, Halford NG (2002) Cereal seed storage proteins: structures, properties and role in grain utilization. J Exp Bot 53:947–958

Shewry PR, Tatham AS (1997) Disulfide bonds in wheat gluten proteins. J Cereal Sci 25:207–227

Shewry PR, Halford NG, Tatham AS (1989) The high molecular weight subunits of wheat, barley and rye. In: Miflin BJ (ed) Genetics, molecular biology, chemistry and role in wheat gluten structure and functionality, Oxford survey plant molecular and cellular biology, vol 6. University Press, New York, pp 163–219

Shewry PR, Halford NG, Tatham AS (1992) High molecular weight subunits of wheat glutenin. J Cereal Sci 15:105–120

Shewry PR, Tatham AS, Barro F, Barcelo P, Lazzeri P (1995) Biotechnology of bread-making: unraveling and manipulating the multi-protein gluten complex. Biotechnology 13:1185–1190

Singh NK, Shepherd KW (1988) The structure and genetic control of a new class of disulphide-linked proteins in wheat endosperm. Theor Appl Genet 71:79–92

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882

Van Campenhout S, Vander Stappen J, Sagi L, Volckaert G (1995) Locus-specific primers for LMW glutenin genes on each of the group1 chromosomes of hexaploid wheat. Theor Appl Genet 91:313–319

Wang ZQ, Long H, Zheng YL, Yan ZH, Wei YM, Lan XJ (2005) Cloning and analysis of LMW-GS Genes from *Triticum aestevum* ssp. *Tibetanum Shao*. Acta Genetica Sinica 32:86–93

Wicker T, Yahiaoui N, Guyot R, Schalgenhauf E, Liu Z-D, Dubcovsky J, Keller B (2003) Rapid genome divergence at orthologous low molecular weight glutenin loci of the A and A (m) Genomes of Wheat. Plant Cell 15:1186–1197

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

Zhao HX, Guo AG, Hu SW, Fan SH, Zhang DP, Ren SL, Wang RJ (2004) Development of primers specific for LMW-GS Genes at *Glu-D3* and *Glu-B3* loci and PCR Amplification. Acta Agronomica Sinic 30:126–130