

Cloning and molecular characterization of three novel LMW-i glutenin subunit genes from cultivated einkorn (*Triticum monococcum* L.)

X. An · Q. Zhang · Y. Yan · Q. Li · Y. Zhang ·
A. Wang · Y. Pei · J. Tian · H. Wang · S. L. K. Hsam ·
F. J. Zeller

Received: 21 October 2005 / Accepted: 19 April 2006 / Published online: 15 June 2006
© Springer-Verlag 2006

Abstract Three novel low molecular weight (LMW) glutenin subunits from cultivated einkorn (*Triticum monococcum* L., $A^m A^m$, $2n = 2x = 14$) were characterized by SDS-PAGE and molecular weights determined by MALDI-TOF-MS. Their coding genes were amplified and cloned with designed AS-PCR primers, revealing three complete gene sequences. All comprised upstream, open reading frame (ORF), downstream and no introns were present. The deduced amino acid sequences showed that all three genes, named as LMW-M1, LMW-M3 and LMW-M5, respectively, belonged to the LMW-i type subunits with the predicted molecular weight between 38.5206 and 38.7028 kDa. They showed high similarity with other LMW-i type genes from hexaploid bread wheats, but also displayed unique features. Particularly, LMW-M5 subunit contained an extra cysteine residue in the C-terminus except for eight conserved cysteines, which

resulted from a single-nucleotide polymorphism (SNP) of the T–C transition, namely arginine → cysteine substitution at position 242 from the N-terminal end. This is the first report that the LMW-i subunit contained nine cysteines residues that could result in a more highly cross-linked and more elastic glutenin suggesting that LMW-M5 gene may associates with good quality properties. In addition, a total of 25 SNPs and one insertions/deletions (InDels) were detected among three LMW-i genes, which could result in significant functional changes in polymer formation of gluten. It is anticipated that these SNPs could be used as reliable genetic markers during wheat quality improvement. The phylogenetic analysis indicated that LMW-i type genes apparently differed from LMW-m and LMW-s type genes and diverged early from the primitive LMW-GS gene family, at about 12.92 million years ago (MYA) while the differentiation of A^m and A genomes was estimated at 3.98 MYA.

Communicated by M. Sorrells

Q. Zhang had the same contribution to this work as the first author.

X. An · Q. Zhang · Y. Yan (✉) · Q. Li · Y. Zhang · A. Wang ·
Y. Pei · J. Tian · H. Wang
Key Laboratory of Genetics and Biotechnology,
College of Life Science, Capital Normal University,
100037 Beijing, China
e-mail: yanym@hotmail.com

S. L. K. Hsam (✉) · F. J. Zeller (✉)
Division of Plant Breeding, Technical University of Munich,
85350 Freising-Weihenstephan, Germany
e-mail: Hsam@wzw.tum.de

F. J. Zeller
e-mail: zeller@wzw.tum.de

Introduction

In bread wheat and related species, the seed storage proteins mainly consist of polymeric glutenins and monomeric gliadins. Under reducing conditions, polymeric glutenins are subdivided into high molecular weight glutenins (HMW-GS) and low molecular weight glutenins (LMW-GS) according to their mobility by SDS-PAGE. Many studies showed that glutenins determine the nutritional and processing properties of bread wheat, especially in dough viscoelasticity (Shewry et al. 1992). So far, HMW glutenins have been extensively studied in gene cloning, functional analyses, genetic evolution and transformation, whereas

relatively less work has been carried out on LMW glutenins because of their complexity in compositions.

Low molecular weight glutenins as well as α -, γ -gliadins and B-hordein of barley are S-rich prolamins, which represent about 40% of the total endosperm proteins and 60–70% of glutenins (Shewry and Halford 2002). Genetic analysis showed that LMW-GS were encoded by the *Glu-3* loci on the short arms of chromosomes 1A, 1B and 1D while some components were found to be encoded by genes on the short arms of the group 6 and 7D chromosomes (D'Ovidio and Masci 2004). Until now, three subgroups of typical LMW-GS have been found on the basis of the first amino acid residue of N-terminal sequences, namely LMW-m, LMW-s and LMW-i types possessing methionine, serine and isoleucine, respectively (D'Ovidio and Masci 2004). LMW-s type is the typical subunit in all genotypes and possesses the N-terminal sequence of SHIPGL-while the LMW-m subunit is METSH(R/C)I-. Pitts et al. (1988) isolated the LMW-i type subunit for the first time, which is a variant form lacking the N-terminal region and starting directly with the repetitive domain after the signal sequence, with the N-terminal sequence of ISQQQQ-. Recent studies have indicated that LMW-i type genes can express normally in the wheat endosperm (Cloutier et al. 2001; Ikeda et al. 2002). Generally, LMW-GS possess eight cysteine residues that are important for forming intra- and inter-molecular disulphide bonds in the gluten macropolymers (GMP).

Low molecular weight glutenins are closely related to dough resistance and extensibility, and also play an important role in determining wheat flour properties. In particular, LMW-2 type subunit in durum wheat is associated with good pasta-making quality. Thus, it is very essential to clone and identify novel LMW-GS candidate genes in wheat quality improvement. More recently, different LMW glutenin genes not only from bread wheat but also from related *Triticum* species and other cereals have been cloned and characterized (D'Ovidio and Masci 2004).

Cultivated einkorn (*Triticum monococcum* L., $2n = 2x = 14$, $A^m A^m$) is closely related to *Triticum urartu* ($2n = 2x = 14$, $A^u A^u$), one of the progenitors of hexaploid bread wheat (Dvorak et al. 1988). Some investigations have shown that there exist extensive glutenin variations in cultivated einkorn, and therefore it is expected to be a potential important source of novel LMW-GS genes for improving bread wheat quality (Borghini et al. 1996). However, less work has been carried out for the LMW glutenin gene cloning and molecular structures of A^m genome in *T. monococcum* (Corbellini et al. 1999; Tranquilli et al. 2002; Wicker et al. 2003). In the present study, three novel LMW-i

type genes from different accessions of *T. monococcum* were isolated, cloned and characterized. In addition the evolutionary relationships with other cereals based on LMW-GS and prolamin genes were described.

Materials and methods

Plant materials

Three accessions of *T. monococcum*, namely Mo-M1 (ATRI 585/74), Mo-M3 (ATRI 644/90) and Mo-M5 (ATRI 896/97), originally obtained from Genebank Gatersleben, Germany, were analyzed.

Protein extraction, SDS-PAGE and N-terminal protein microsequencing

Half of a seed (about 20 mg) was used for LMW-GS extraction according to Yan et al. (1999). Sodium dodecyl sulphate (SDS) polyacrylamide gel electrophoresis (PAGE) was performed on a Bio-Rad Mini-PROTEAN III cell. Running time was prolonged half of hour after the indicator solution line passed out of the gel-edge based on the method of Yan et al. (2003). To confirm the type of LMW subunits identified, N-terminal protein microsequencing was performed on the instrument of PROCISE[®] cLC 491 (Applied Biosystems).

Matrix assisted laser desorption ionization time of flight mass spectrometry (MALDI-TOF-MS)

Extraction of LMW glutenin subunits for MALDI-TOF-MS analysis was similar to the method described above. Pre-cooling acetone was added to the final sample supernatant (the end concentration of acetone was 80% v/v), and then the LMW-GS were allowed to precipitate for 2 h at -20°C and centrifuged for 10 min at 13,000g. After vacuum centrifugation drying, the pellet was dissolved in acetonitrile + 1.0% Trifluoroacetic acid in water (1:1) solution, and then 1 μl sample concentrated by ZIPTip was used for MALDI-TOF-MS analysis on a Shimadzu corporation AXIMA-CFRTM-Plus MS apparatus. The matrix used was sinapinic acid (SA, 3,5-Dimethoxy-4-Hydroxycinnamic acid). Spectra were acquired in a linearity mode using a mass range of 10,000–100,000 Da and about 80–120 laser shots power were averaged to improve the signal-to-noise level. International calibration was performed by using the method named two-dot calibration with standard sample Albumin-Aldrase at masses of 39,212.88 and 66,431.08 Da.

DNA extraction and PCR amplification

Genomic DNA of one dry seed was prepared following the procedure of McDonald et al. (1994). Some of processes were improved, for instance, ground powder of one seed was dipped firstly in chloroform before extraction buffer was added, which would make proteins and enzymes especially DNA enzyme denatured in order to protect DNA.

A pair of AS-PCR primers (LMW-3 and LMW-4) was designed to amplify the upstream coding regions and downstream based on previously cloned LMW glutennin gene sequences (Colot et al. 1989; Cloutier et al. 2001). The sequences of primers were LMW-3: 5'-GCCTTTCTTGTTTACGGCTG-3'. LMW-4: 5'-TCAGATTGACATCCACACAAT-3' (synthesized by Sangong). PCR amplifications were performed in 50 μ l reaction volume containing 2.5 U La Taq polymerase (TaKaRa) 100 ng of template DNA, 25 μ l of 2 \times GC buffer II (MgCl $_2^{2+}$ plus), 0.4 mM dNTP, 0.5 μ M of each primer, and some of ddH $_2$ O. The reactions were carried out in a PTC-100 (MJ Research) using the following protocol: heat lid turned on, 94°C for 2 min to denature the DNA, cycled 35 times at 94°C for 45 s, 58°C for 1 min and 72°C for 2 min, finally extended at 72°C for 10 min.

Molecular cloning and DNA sequencing

PCR products were separated on 1.2% agarose gels and expected fragments were purified from the gels using Quick DNA extraction kit (TaKaRa). Subsequently purified products were ligated into pGEM-T Easy vector (Promega) and transformed into cells of *Escherichia coli* DH-5 α strain. DNA sequences were obtained with three clones using primer walking and performed by TaKaRa Biotech.

Identification of SNPs and InDels

The identification of single-nucleotide polymorphisms (SNPs) and insertions/deletions (InDels) among LMW glutenin genes encoded by *Glu-A3* locus from different *Triticum* species was based on the multiple alignments and performed with software Bioedit 7.0.

Phylogenetic analysis

Multiple alignment of LMW glutenin nucleotide and protein sequences were completed by Bioedit 7.0. DNAMAN 5.2.2 was used to construct a phylogenetic tree. The divergence times were calculated by MEGA3 (Gaut et al. 1996; Kumar et al. 2004).

Results

Characterization of LMW glutenin subunits from *T. monococcum*

The compositions of LMW as well as HMW glutenin subunits of three cultivated einkorn accessions Mo-M1, Mo-M3 and Mo-M5 from *T. monococcum* were separated and characterized by SDS-PAGE, and the results are shown in Fig. 1. The LMW-GS of Mo-M1, Mo-M3 and Mo-M5 had two clear regions, which correspond to B-subunits and C-subunits, respectively. In general, the B-subunits comprised fewer components because of only one coding locus *Glu-A3* is present in *T. monococcum*. Mo-M1 contained only a B-subunit and two B-subunits were present in Mo-M3 and Mo-M5 accessions. In the present study, the genes coding for the B-subunit with approximately 40 kDa (Fig. 1), designated LMW-M1, LMW-M3 and LMW-M5, respectively, were isolated and sequenced (see below).

The N-terminal microsequencing demonstrated that the first 15 amino acid residues of all three subunits were identical (ISQQQQQPPFSQQQQ), and hence they belonged to typical LMW-i type subunits because of the presence of the first amino acid residue isoleucine of the mature protein. To obtain accurate molecu-

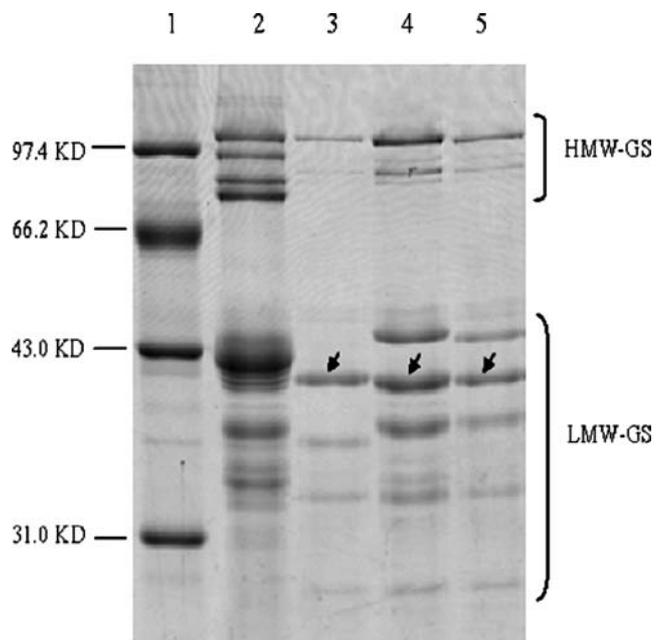


Fig. 1 SDS-PAGE analysis of total glutenins of single seed for *T. aestivum* Chinese Spring, and *T. monococcum* Mo-M1, Mo-M3 and Mo-M5. Lane 1 Protein marker, lane 2 control sample (Chinese Spring), lane 3–5 three accessions of Mo-M1, Mo-M3 and Mo-M5, respectively. The LMW-GS for coding cloned genes are arrowed

lar weights, LMW-GS were further analyzed by MALDI-TOF-MS. As shown in Fig. 2, three accessions possessed an apparent protein peak with the molecular weights of 38.7896, 39.0269 and 39.0080 kDa that are generally consistent with the M_r of LMW-M1, LMW-M3 and LMW-M5, respectively.

PCR amplification and clone of LMW glutenin genes

It is known that LMW-GS genes contain no introns, so the entire gene sequences with no intervention can be amplified by using genomic DNA as a template. A pair of AS-PCR primer LMW-3 and LMW-4 was designed to amplify the complete coding sequences of LMW-M1, LMW-M3 and LMW-M5 subunits, which started at about 380 bp in upstream and ended at 190 bp in downstream. Figure 3 showed the PCR amplification

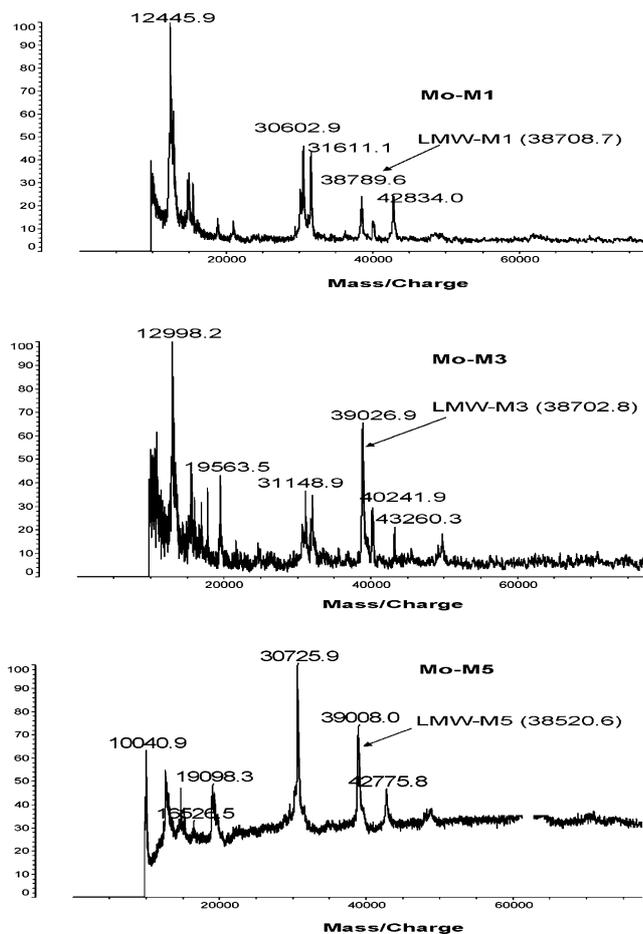


Fig. 2 The molecular weight (M_r) determination of LMW glutenin subunits from three accessions (Mo-M1, Mo-M3 and Mo-M5) of *Triticum monococcum* by MALDI-TOF-MS. The subunits of LMW-M1, LMW-M3 and LMW-M5 and their M_r s were indicated. The values in bracket were the M_r s calculated from deduced amino acid sequences of their coding DNA sequences

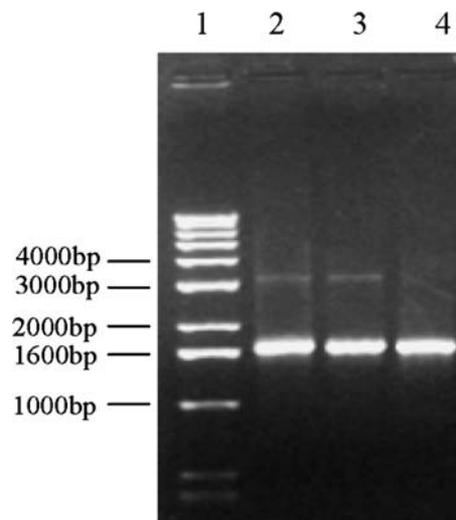


Fig. 3 PCR amplification products from genomic DNA of three *T. monococcum* accessions. Lane 1 DNA mark, lanes 2–4 PCR products of accessions Mo-M1, Mo-M3 and Mo-M5

products of three accessions Mo-M1, Mo-M3 and Mo-M5, with all accessions containing a similar single band with a size of about 1,600 bp. Since most of the complete coding sequences of LMW-GS genes vary between 909 and 1,167 bp (Cassidy et al. 1998; Johal et al. 2004; D'Ovidio and Masci 2004), these special bands were well corresponding to the complete sequences of LMW-GS genes after subtracting upstream and downstream sequences. The amplified products were cloned and named as LMW-M1, LMW-M3 and LMW-M5, respectively.

Molecular characterization of three novel LMW-i glutenin genes

Both strands of LMW-M1, LMW-M3 and LMW-M5 were sequenced by primer walking. LMW-M1 consisted of 1,652 bp containing 385 bp upstream, 1,065 bp open reading frame (ORF), double terminal codons and 196 bp downstream. LMW-M3 (1,638 bp) and LMW-M5 (1,646 bp) consisted of 377 and 386 bp of upstream, same 1,059 bp of ORF and two terminal codons, 196 and 195 bp downstream, respectively.

The upstream nucleotide sequences of the three cloned genes and ten other LMW-GS genes reported previously were aligned by Bioedit 7.0 in order to analyze the characters of promoter sequences and the results are shown in Fig. 4. All typical promoter motifs were found in the upstream regions of each gene. One CAAT box (6–10 bp) and two TATA boxes (57–62, 74–79 bp) were present in the upstream from start codon among most genes. In the position of first

TATA box, single nucleotide A was substituted by G in LMW-M1 or by T in AY585350, AY585354 and Y14104. Endosperm box also named as the –300 element or prolamin box (Forde et al. 1985; Colot et al. 1987) was present between –326 and –160 bp, which is the sequences to bind a trans-acting factor controlling the transcriptional level (Bartels and Thompson 1986). It has been reported that gene containing an endosperm box could be conferred to be a typical LMW-GS gene (Colot et al. 1987). As shown in Fig. 4, an endosperm box including E-motif and N-motif existed in all alignment genes, which located at about 300 bp upstream from start codon.

The analysis of the deduced amino acid of the three novel genes showed that all possessed a single ORF and do not contain internal stop codons. The ORF of LMW-M1 gene encoded 355 amino acids residues while both LMW-M3 and LMW-M5 genes encoded a protein of 353 residues. The coding regions of three genes were all terminated by double stop codons like other LMW-GS and prolamin genes. Three polyadenylation signals (AATAAA) were identified, which are located in 73–78, 131–136 and 142–147 bp in the downstream regions. As the results of protein N-terminal microsequencing, the first amino acid residue of their mature proteins was isoleucine, and hence three

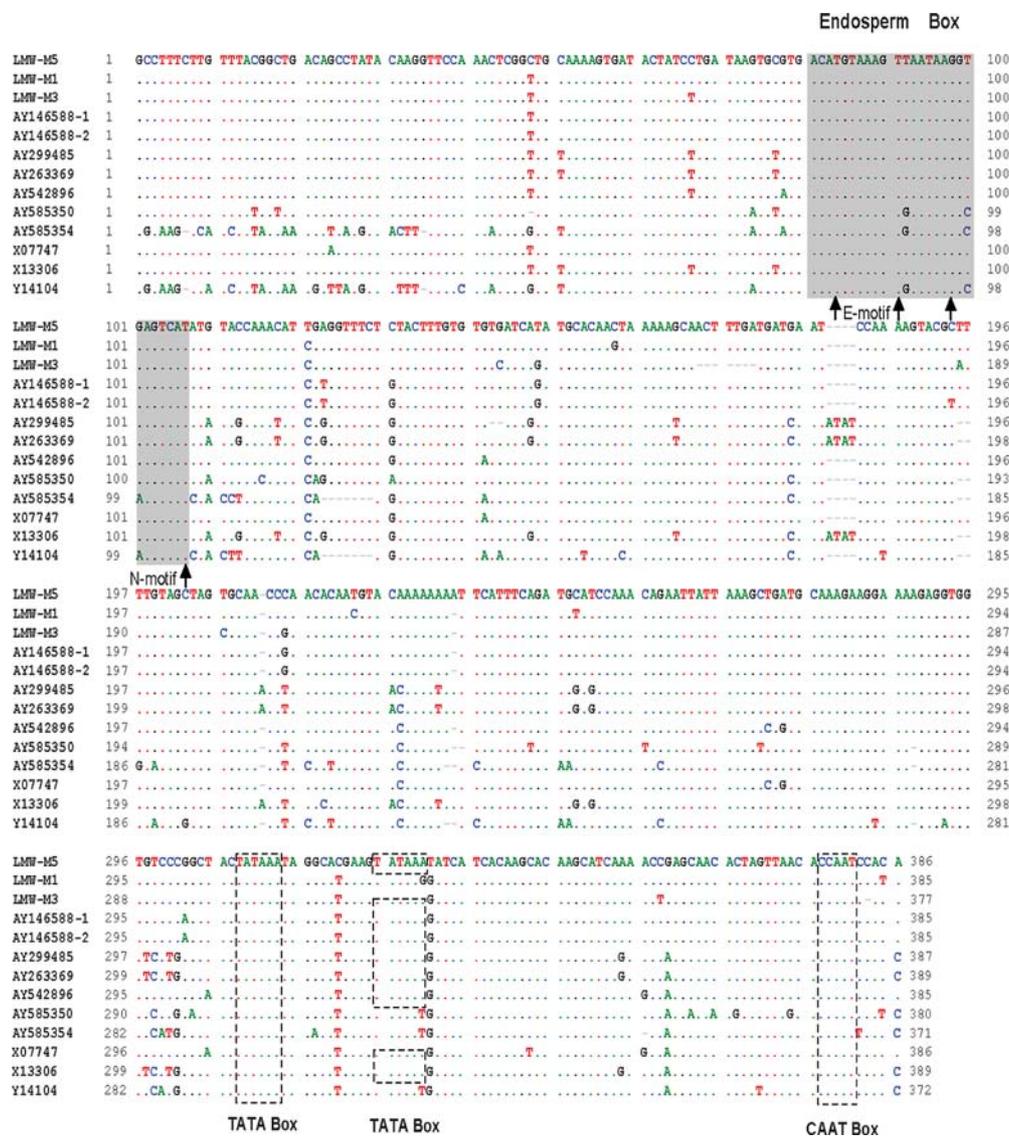


Fig. 4 Comparison of the 5' flanking DNA sequences of LMW-M1, LMW-M3 and LMW-M5 genes with those of other nine LMW glutenin genes. TATA box and CAAT box are boxed by broken lines and the endosperm box is indicated by grey box. In

all genes, dots and dashes indicate the same sequences with LMW-M5 genes and deletion. AY146588-1 and AY146588-2 indicated the two separate genes with the same GeneBank accession AY146588

LMW-GS could belong to LMW-i type subunits. Furthermore, their deduced protein sequences were similar to those of the other LMW-i subunits presented in bread wheat (Ikeda et al. 2002; Cloutier et al. 2001).

The predicted molecular weights from the coding proteins of LMW-M1, LMW-M3 and LMW-M5 genes were 38.7067, 38.7028 and 38.5206 kDa, respectively. It can be seen from Table 1, that the molecular weight of LMW-M1 is well in consistent with that from MS method. For LMW-M3 and LMW-M5 subunits, however, the molecular weights from MALDI-TOF-MS were 324.1 and 478.4 Da higher than those from gene coding sequences, respectively. These results suggested that LMW-M3 and LMW-M5 subunits may have some kinds of protein post-translational modifications (PTMs), such as phosphorylation and glycosylation and so on. Lauriere et al. (1996), on the first time, reported the PTMs of LMW-GS and showed that LMW-GS were N-glycosylated with xylose.

As shown in Fig. 5, the deduced amino acid sequences of the three genes were aligned with other 13 LMW-i, LMW-m and LMW-s type subunits from *T. monococcum*, *Aegilops tauschii*, *T. durum* and *T. aestivum* species. In general, the encoded proteins of three novel genes had similar structures to previously characterized LMW-GS. Each comprised four main structural regions, including a 20 amino acid signal peptide (Sig), a short N-terminal region of 13 amino acids, a repetitive domain rich in glutamine and proline residues and a C-terminal domain. As suggested by Cassidy et al. (1998), the N-terminal and repetitive domains were also named as I and II domains, and the C-terminal domain was further subdivided into three regions: a conserved domain III, a glutamine-rich region IV and a highly conserved sequences region V terminating the LMW-GS.

In the domain I, LMW-i type subunits (LMW-M5, LMW-M1, LMW-M3, EMBL accession AY146588, AY542896, X07747, U86030) with ISQQQQ being the deduced N-terminal sequences actually lack the N-terminus, starting directly with repetitive region. All

LMW-m type subunits (EMBL accession AB585350, X51759, X13306 and U86027) contained a cysteine residue at the position 5.

The repetitive region (domain II) consisted of a typical repeat motif: P1-2FP/S Q2-6 (Cassidy et al. 1998), the number of which result in variable length of protein. The comparison of the three different types of LMW-GS demonstrated that repetitive domain variation resulted from deletions and insertions of some repeat units. The repeat motif of seven LMW-i subunits was mainly P1-2FSQ2-6. In particular, the repetitive numbers of three novel genes LMW-M5, LMW-M1 and LMW-M3 were 9, 13 and 9, respectively. The repeat P1-2 F SQ2-6 also existed in LMW-m and LMW-s subunits, but their repeat motif numbers were fewer than that of LMW-i subunits. The repeat motifs resulted from various single amino acid substitution were also found in repetitive regions, for instance PLFSQ4, PSFSQ4, PPYSQ4 and PPFHQ4 and so on. Four LMW-m subunits (EMBL accession Y17845, Y18159, AB062853, AB119006) had a cysteine residue at the same positions in the repetitive region.

The domain III is a cysteine-rich region usually containing five cysteine residues. Figure 4 showed that all of aligned LMW-m and LMW-s subunits were founded to have five cysteines at five relatively conserved positions while the LMW-i subunits contained six cysteines with an exception of LMW-M5 subunit that had seven cysteines, with an extra cysteine located at the 67th amino acid of the domain III. The domain IV contained the seventh cysteine at two different positions, which is as same as the result of Cassidy et al. (1998). The last domain V contained the eighth or the ninth cysteine, which is highly conserved region and usually used as specific LMW-GS sequences for designing PCR primers.

Although the structures of the five LMW-i type subunits (LMW-M5, LMW-M1, LMW-M3, AY146588-1 and AY146588-2) from *T. monococcum* showed higher identity, striking differences among them could be found. For instance, LMW-M5 displayed 10 amino

Table 1 Composition of nucleotides, deduced amino-acid sequences and molecular weights of three novel LMW-i genes from *Triticum monococcum*

| LMW-GS genes | Length of nucleotide (bp) ^a | | | Residue number of mature protein | Number of cysteine residues | | | | Molecular weights (kDa) | | |
|--------------|--|-------|-----|----------------------------------|-----------------------------|---|---|-------|-------------------------|-----------------|------------|
| | X | Y | Z | | X | Y | Z | Total | By protein sequence | By MALDI-TOF-MS | Difference |
| LMW-M1 | 385 | 1,065 | 196 | 335 | 6 | 1 | 1 | 8 | 38.7087 | 38.7896 | +0.0809 |
| LMW-M3 | 377 | 1,059 | 196 | 333 | 6 | 1 | 1 | 8 | 38.7028 | 39.0269 | +0.3241 |
| LMW-M5 | 386 | 1,059 | 195 | 333 | 7 | 1 | 1 | 9 | 38.5206 | 39.0080 | +0.4874 |

^a X, Y and Z represent upstream, ORF and downstream region, respectively

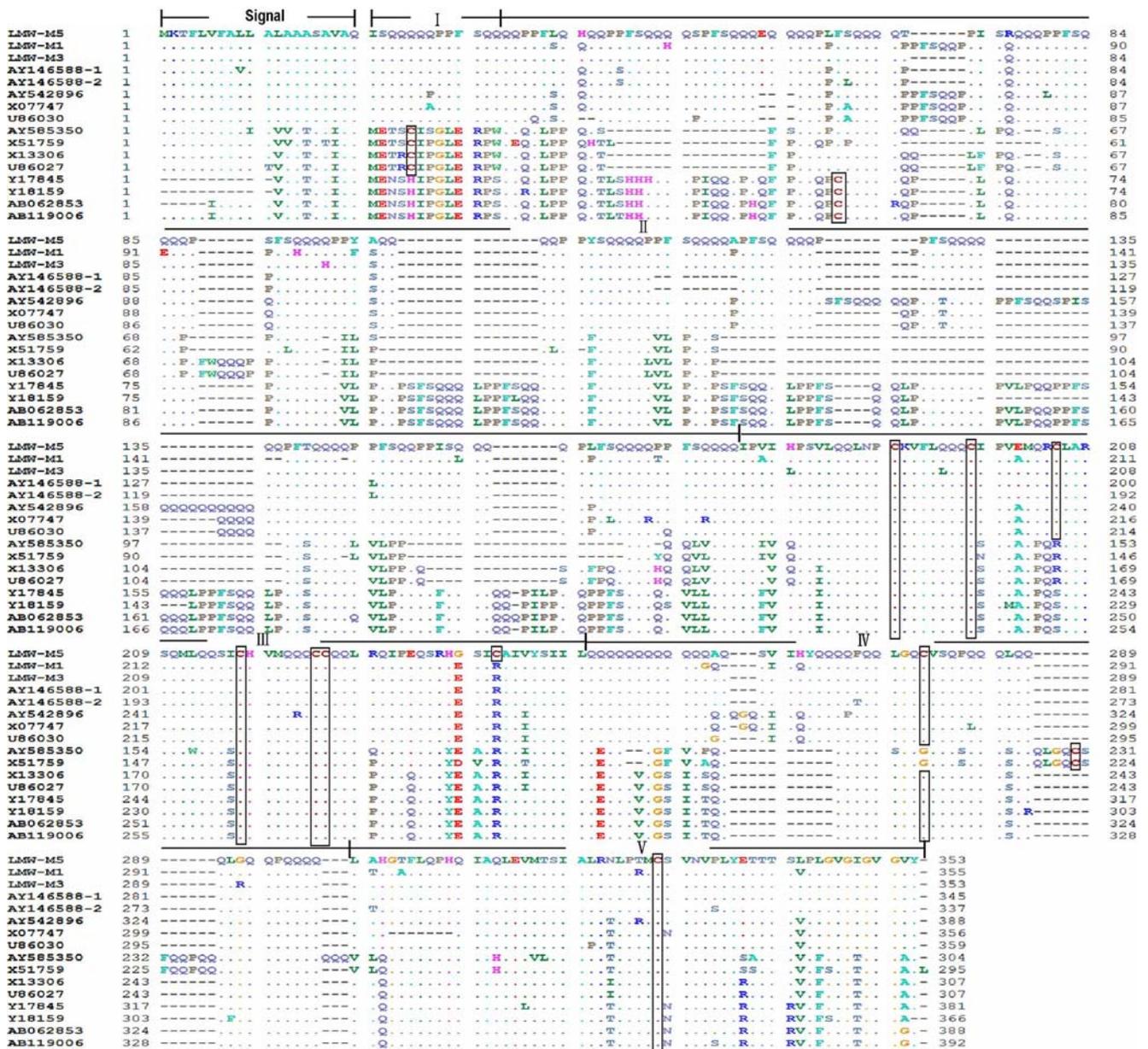


Fig. 5 Multiple alignment of the deduced amino acid sequences of 16 LMW glutenin genes including LMW-i LMW-m and LMW-s three types. Signal, signal peptide, I, II, III, IV and V represent N-terminal domain, repetitive domain and three sub-regions of

C-terminal domain, respectively. Cysteines are boxed by real lines. The same sequences with LMW-M5 genes and deletion are indicated by a dot and a dash

acid substitutions, 8 polypeptide (PFSQQQQA) insertion in domain II and an additional cysteine comparing with AY146588-1, and 14 amino acid substitutions, two 8 polypeptide (PFSQQQQA and PPFSSQQQ) insertions in domain II and an additional cysteine in domain III comparing with AY146588-2. More differences could be observed between LMW-M1 and AY146588-1, AY146588-2: 20 and 21 amino acid substitutions, the same six polypeptide (PFSQQP), eight polypeptide (PFSQQQP) and six polypeptide (PPFSQQ) insertions, a dipeptide (QQ) (only with AY146588-1) and

two single amino acid (Q) deletions, respectively. When compared to AY146588-1 and AY146588-2, LMW-M3 displayed 11 and 14 amino acid substitutions and the same insertions as in LMW-M5.

Identification of SNPs and InDels

In this study, the complete coding sequences were aligned to detect SNPs and InDels in LMW-M1, LMW-M3 and LMW-M5 by comparing with other 12 LMW-GS genes that are all assigned to LMW-i subunits

encoded by 1A and 1A^m chromosomes. These LMW-i genes originated from different *Triticum* species, including *T. monococcum*, *T. durum* and *T. aestivum*. The SNPs and InDels detected in three novel LMW-i genes were listed in Table 2. A total of 25 SNPs was identified at different positions and the numbers of SNPs in LMW-M1, LMW-M3 and LMW-M5 were 5, 13 and 14, respectively. The majority (15 SNPs accounting for 60%) of the changes in nucleotides at the SNP sites resulted from transition (A–G or C–T) and transversion (A–T, A–C, C–G or G–T) that account for slightly more than one-third among the detected SNPs. One C deletion SNP at the position of 529 was detected in LMW-M1. Of the 25 SNPs, 14 were predicted to produce amino acid substitutions, namely nonsynonymous SNP. Particularly, the transition of T–C resulted in arginine → cysteine substitution at position 242 from the N-terminal end in LMW-M5. Four nonsynonymous substitutions were present in both LMW-M5 and LMW-M3, three were unique to the LMW-M5 and LMW-M3, respectively, and four were only present in LMW-M1. The frequency of SNPs varied from one SNP per 76–216 bases. Only one AC InDels was found in LMW-M1, which located at 531–532 bp.

Phylogenetic analysis among LMW-GS and other prolamins genes

Through alignment of 34 nucleotide sequences (the coding region), a homology tree was constructed to analyze evolutionary relationships among different LMW-GS and other prolamins genes using DNAMAN software and the result is shown in Fig. 6. These sequences comprised of 19 LMW-GS genes from differ-

ent genomes of diploid, tetraploid and hexaploid species, six HMW-GS genes from *Triticum aestivum*, six gliadin genes from *T. aestivum* and *T. durum* and three hordein genes from barley (*Hordeum vulgare* L.). The LMW-GS genes included three novel LMW-i genes (this study) and other 16 genes from Genbank, namely AY146588 (Wicker et al. 2003), AB062853 (Ikeda et al. 2002), AB119006 and AB119007 (Maruyama-Funatsuki et al. 2005), AY453158 (Zhang et al. 2004), AY542896 (Cloutier et al. 2001), AY585350 (Johal et al. 2004), AY299457, U86027, U86030, X13306 (Colot et al. 1989), X51759, X07747 (Pitts et al. 1988), Y17845 (Masci et al. 1998) and Y18159 (D'Ovidio et al. 1999). The accession numbers of other prolamins genes are X61009 (Halford et al. 1992), X12928, X13927 (Anderson and Greene 1989), X12929 (Anderson et al. 1989), X61026 (Halford et al. 1987), X03042 (Forde et al. 1985), U51307, AJ870965, AJ937838, AY338392, AY591334, AF280605 (Hsia and Anderson 2001), X53690 (Vicente-Carbajosa et al. 1992), S66938 (Sainova et al. 1993) and AY268139 (Gu et al. 2003).

As shown in Fig. 6, it is obvious that the homology tree is clustered into two clear branches that generally represent HMW and LMW cereal prolamins, respectively. Among LMW prolamins genes, five subgroups were apparently separated, namely LMW-GS, B-Hordeins, γ -gliadins, α -gliadins, and ω -gliadins and C-Hordeins. The LMW-GS genes had 78% identity and contained three clear groups that correspond to LMW-i, LMW-s and LMW-m type subunit genes, respectively. This is in good agreement with their similarity of amino acid sequences shown in Fig. 5. In particular, LMW-i genes showed greater differences to LMW-s and LMW-m genes with the homology of 78% while

Table 2 The positions of SNPs and InDels identified in the three novel LMW-i genes

| LMW-GS gene | 111 | 123 | 162 | 194 | 214 | 237 | 245 | 283 | 294 | 309 | 313 | 529 | 531 | 532 |
|---------------------------------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|-----|
| LMW-M5 | G | T | C | T | A | A | G | T | G | G | G | C | A | C |
| LMW-M3 | G | T | C | T | A | C | A | T | G | A | T | C | A | C |
| LMW-M1 | A | A | T | C | C | A | A | C | C | A | T | — | — | — |
| Twelve other LMW-i Genes ^a | A | A | T | C | C | A | A | C | G | A | T | C | A | C |
| LMW-GS gene | 584 | 650 | 656 | 701 | 747 | 810 | 833 | 841 | 1003 | 1027 | 1102 | 1116 | 1179 | |
| LMW-M5 | A | T | A | A | G | A | G | T | G | C | T | C | T | |
| LMW-M3 | A | T | T | T | A | G | A | C | A | T | T | C | G | |
| LMW-M1 | T | C | A | A | A | A | A | C | G | T | C | T | T | |
| Twelve other LMW-i Genes ^a | A | T | A | A | A | A | A | C | G | T | C | C | T | |

SNPs are indicated in boldface while horizontal dash indicates the deletions of nucleotide

^a Twelve other LMW-i genes encoded by 1A^m or 1A chromosome include AY146588-1 and AY146588-2 (Wicker et al. 2003), X07747 (Pitts et al. 1988), AY542896 (Cloutier et al. 2001), AB062877 (Ikeda et al. 2002), AY453154, AY453155, AY453156, AY453157, AY453158, AY453159 and AY453160 (Zhang et al. 2004)

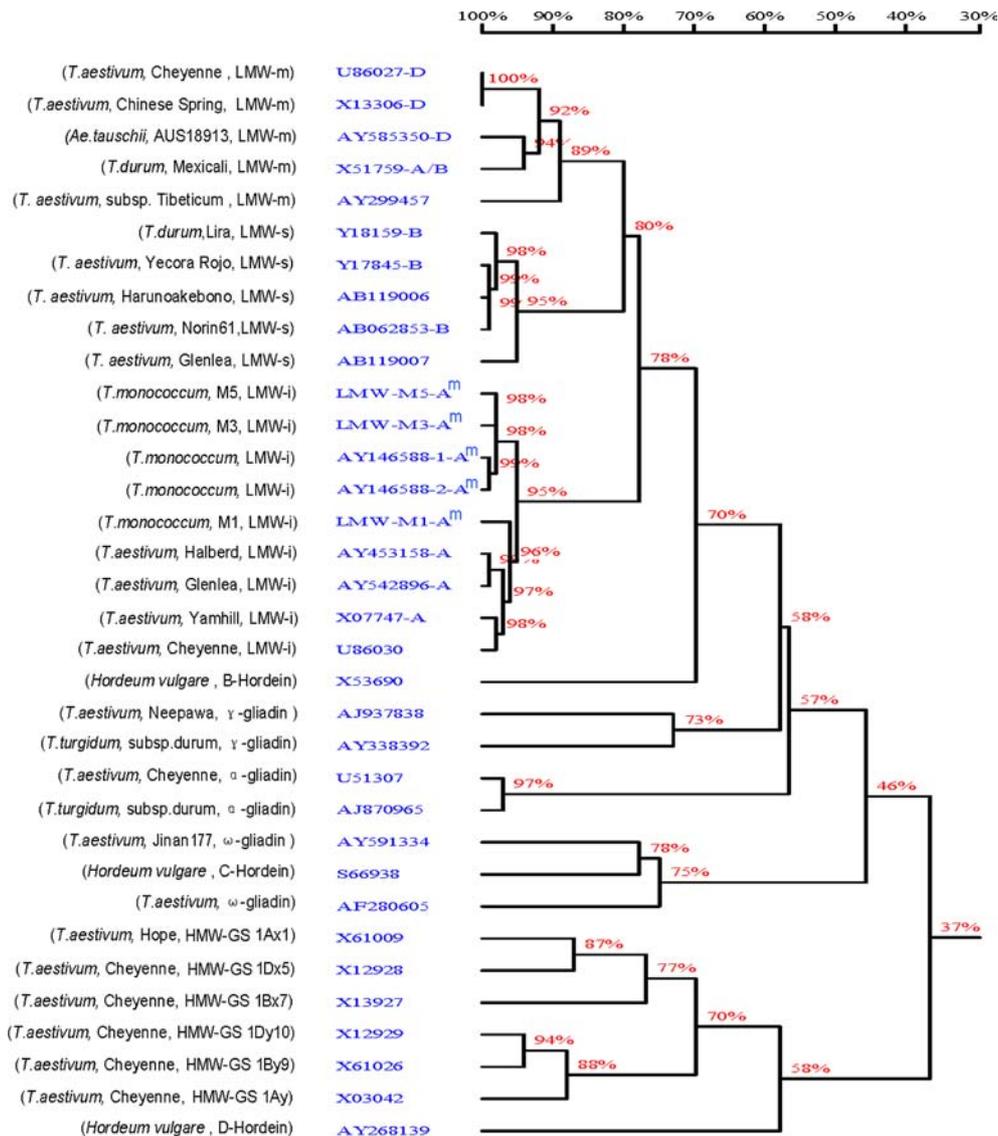


Fig. 6 Homology tree showing the relationship among LMW-GS, HMW-GS and gliadins from different *Triticum* species and hordein genes from barley (*Hordeum vulgare* L.)

the identity between LMW-s and LMW-m genes was 80%. Among three groups of LMW-GS genes, the identity was 95, 95 and 89% in LMW-i, LMW-s and LMW-m genes, respectively.

Among the gene family of cereal seed storage proteins, LMW-GS genes were most related to B-Hordeins of barley, γ -gliadins and α -gliadins with the sequence homology of 70, 58 and 57%, respectively. The identity between LMW-GS and ω -gliadins and C-Hordeins were 46%. The lower sequence homology (37%) was found between LMW-GS and HMW-GS and barley D-hordeins.

In order to further understand the evolutionary relationships between LMW-GS and other prolamin

genes, the nucleotide sequences (coding region) of all genes analyzed in Fig. 6 were aligned, and then their divergent times were calculated with software MEGA3. According to the previous reports (Gaut et al. 1996; Wicker et al. 2003; Li et al. 2004), the evolutionary rate of 6.5×10^{-9} was used. In the evolutionary history of wheat prolamin gene family, the loci of HMW-GS and LMW-GS diverged relatively earlier, at about 45.34 MYA (million years ago). About 40.10 MYA, the loci of LMW-GS genes firstly separated from ω -gliadin loci and relatively later differentiated from α - and γ -gliadin loci (32.18 MYA). Among three types of LMW-GS genes, the estimated divergent time between LMW-i and LMW-m, LMW-s was 12.92

MYA while LMW-m and LMW-s were divergent at 11.76 MYA. In addition, the divergent time of LMW-GS with B-Hordein and D-Hordein loci was 19.33 MYA and 45.34 MYA, respectively.

Discussion

Cultivated einkorn (*T. monococcum*) is the only cultivated diploid species and carries an A^m genome closely related to the A genome of hexaploid wheat (Tranquilli et al. 2002). It has been proved that *T. monococcum* is a potential source of genetic variation for proteins and new quality characteristics according to the variations observed in the bread making properties among several accessions (Degidio et al. 1993; Degidio and Vallega 1994; Borghi et al. 1996). In this study, three novel LMW-i glutenin genes, named as LMW-M1, LMW-M3 and LMW-M5 from cultivated einkorn accessions of Mo-M1, Mo-M3 and Mo-M5 were cloned and sequenced by AS-PCR primers, which are expected to be useful gene resources for quality improvement of bread wheat.

Structures of LMW-GS and implication for their functional properties

The deduced N-terminal amino acid sequences of the three novel genes demonstrated that all were the i-type subunit of LMW glutenins as isoleucine is present as the first amino acid residue (Cloutier et al. 2001). They are unlike m- and s-type LMW glutenin genes and usually possess the truncated N-terminus and start directly with repetitive domain after signal peptide. The predicted molecular weights from the deduced amino acid sequences of the three genes cloned ranged from 38.5206 to 38.7028 kDa, which are in good agreement with the M_r s determined by MALDI-TOF-MS as indicated in Fig. 2. These results demonstrated that LMW-GS lacked extensive glycosylation or other post-translational modifications as some HMW glutenin subunits (Cozzolino et al. 2001; Cunsolo et al. 2004). However, Lauriere et al. (1996) reported N-glycosylation with xylose in LMW-GS, which may be related to the functional properties of wheat. Furthermore, some glycosylation and phosphotyrosine in HMW glutenin subunits were also detected (Tilley et al. 1993; Tilley and Schofield 1995). Until now, little has been known about post-translational modifications of LMW-GS as well as HMW-GS. Thus, further studies are needed.

The repetitive domain and repeat motif numbers showed clear variations in LMW-i, LMW-m and LMW-s genes (Fig. 5) which result in variable length of

proteins. The length of LMW-m type subunits in repetitive domain was shorter than those of LMW-i and LMW-s type subunits, mainly due to the more deletions which occurred in this domain. As shown in Fig. 5, three long deletions were present in the four LMW-m subunit sequences, namely 15 amino-acid deletion at position 45–49, 16 amino-acid deletion at position 112–127 and 38 amino-acid deletion at position 145–182, respectively. Seven LMW-i subunits also displayed clear deletions of repeat unit. For instance, FWQQQP, PSFSQQQLPPFS/LQQ and QQQPIP/LP deletions occurred in the positions of 95–100, 124–137 and 212–219, respectively. Otherwise, some repeat unit insertions, such as QQPSFSQQQLPPFS/LQQ (122–137), SFSQQQLPPFS (144–155) and PFISQQQQQ (207–304), were present in all four LMW-s subunits compared in Fig. 5. It is possible that during replication, the repetitive region diverges rapidly by allowing slippage leading to duplication or deletion of sequences (Cassidy and Dvorak 1991). As suggested for other prolamin evolution (Anderson and Greene 1989), single base, single repeat changes and unequal crossover and so on could be responsible for the variations of the repetitive domain.

As shown in Fig. 5, all LMW glutenin subunits contained eight conserved cysteine residues except for LMW-M5 that possessed nine cysteine residues. In general, LMW-i glutenin genes contain all cysteine residues in C-terminus at precisely conserved position, which differ from LMW-m and LMW-s subunits with a cysteine residue in N-terminus or in repetitive domain. The different distribution and the number of cysteine residues could lead to functional differences, especially to the LMW-i subunits (D'Ovidio and Masci 2004). It is well known that the first and the seventh cysteine form the inter-molecular disulfide band while the remains form three intra-molecular disulfide bands (Lew et al. 1992; Masci et al. 1998; D'Ovidio et al. 1999). Thus, the secondary and three-dimensional structures of LMW-i subunit would be quite different from those of LMW-m and LMW-s subunits. Cloutier et al. (2001) suggested that the cysteine in LMW-i subunit differing to that in other two types of subunits would be located in the middle of a tight loop formed by the disulfide bond between cys2 and cys3, which may have an impact on the viscoelastic properties of protein. In the Canadian bread wheat cultivar Glenlea, the LMW-i glutenin gene (AY542896) assigned to 1A chromosome has been confirmed to co-migrate with LMW-50 subunit that plays an important role in determining the good quality characteristics of Glenlea. Therefore, AY542896 gene was considered to have positive effect on quality properties (Cloutier et al. 2001).

Interestingly, an additional cysteine residue in LMW-M5 was found, which resulted from a SNP variation (T–C transition) resulting in arginine → cysteine substitution. LMW-M5 is the first LMW-i gene containing nine cysteines. Recently, the same number of cysteines has also been reported in the LMW-m gene AY263369 that is inferred to associate with the good properties of Chinese bread wheat cultivar Xiaoyan 6 (Zhao et al. 2004). According to the structural features of LMW-GS, it may be speculated that the additional cysteine residue probably results in a more highly cross-linked with other glutenin subunits and consequently more elastic glutenin structure may form as the situation in HMW 1Dx5 subunit (Shewry et al. 1992). The additional cysteine residue present in LMW-M5 subunit would promote a differential cross-linking and endow dough with increased strength and superior performance. Therefore, it could be considered that the LMW-M5 gene as a new candidate LMW glutenin gene that may have important effect on wheat quality.

Single nucleotide polymorphisms (SNPs) and insertions/deletions (InDels) are the most frequent variations in the genome of any organism, which play an important role in many aspects of genetics and breeding researches. SNPs were first found in the human genome and since then more than 1.4 million SNPs have been identified (Sachidanandam et al. 2001). More recently in wheat, HMW-GS (Lu et al. 2005) and gliadin (Zhang et al. 2003) genes have been found to contain SNPs, which can be used as desirable molecular marker and AS-PCR marker. In this work, we used the complete coding DNA sequences of 10 LMW-i type genes in Genbank to align with the corresponding sequences of LMW-M1, LMW-M3 and LMW-M5, then to identify the SNPs and InDels in the three novel genes cloned from *T. monococcum*. The numbers of SNPs in LMW-M1, LMW-M3 and LMW-M5 are five in 1,065 bp, 13 in 1,059 bp and 14 in 1,059 bp, respectively. The frequency of SNPs is higher than that in HMW-GS, probably due to the fact that LMW-GS loci are more complex and the total gene copy number is highly variable from 10–15 (Harberd et al. 1985) to 35–40 (Cassidy et al. 1998). It could be concluded that SNPs and InDels variations have striking effect on glutenin polymer structure and hence result in different functional properties. It may be anticipated that the SNPs and InDels detected in LMW-GS may provide new tools for further insights into the mechanisms of quality variations. Furthermore, it could be used as reliable genetic markers for the identification of different LMW-GS genes when the desirable subunits were introgressed from cultivated einkorn into bread wheat.

Evolutionary relationships of prolamin genes in different cereal species

As described above, three novel LMW-i genes isolated from *T. monococcum* showed similar structures with other LMW-GS genes, but unique features exist when compared to other LMW-i genes of cultivated einkorn and hexaploid species, especially in extra cysteine residue, SNPs and InDels variations. These novel genes have provided new information for understanding the structural differentiation and evolutionary relationships among prolamin genes in different cereal species. As shown in phylogenetic analysis, the LMW-i type subunits encoded by *Glu-A3* were clearly different from the other two groups of LMW-GS and early diverged from primitive LMW-GS at about 12.92 MYA, which is earlier than the divergent time (7.2–10.0 MYA) of x- and y-type HMW-GS genes and the origin of A, B and D genomes (5.0–6.9 MYA) estimated by Allaby et al. (1999). The LMW-i genes from A^m genome had 95% identity with those from A genome of common wheat and the differentiation of A^m and A genomes occurred 3.98 MYA, close to the estimated divergence time (3.13 MYA) between *T. monococcum* LMW glutenin *Glu-A3-3* and *T. durum* *Glu-A3-1* locus (Wicker et al. 2003).

In the evolutionary history of prolamin genes, the loci of LMW-GS and HMW-GS as well as D-hordeins diverge earlier, at about 45.34 MYA. In relatively recent years, LMW-GS diverged with gliadins about 40.10–32.18 MYA. In fact, many reports have shown that some of the B group, most of the C and D group are the modified gliadins that have acquired the ability to participate in the formation of the glutenin polymers (Lew et al. 1992; D'Ovidio and Masci 2004). Cereal prolamins constitute a protein superfamily with limited sequence homology, of which the S-rich prolamins are the most abundant and structurally diverse. Comparative analysis demonstrated that the C-terminal domains of some LMW-GS as well as other cereal S-rich prolamins possess three conserved regions, called A, B and D (Kreis and Shewry 1989). This suggests that they have evolved by the triplication of a single short ancestral sequence. Various events, including SNPs and InDels variations have resulted in considerable diversity in prolamin structures. Past studies demonstrated that most deletions and insertions were probably caused by unequal crossing-over and slippage during replications. So far, LMW-GS as well as other prolamin genes have shown to lack introns that are generally considered as a primitive feature (Kreis and Shewry 1989), which may provide useful information for understanding the evolutionary relationships among

plant genes. Particularly, the absence of typical N-terminal sequences of the LMW-i type subunits, probably resulted from deletions of DNA fragments and hence resulted in this variation form, may reflect its recent origin from a more ancient LMW-m or LMW-s type gene.

Recent studies have demonstrated that two LMW-GS genes (AY146588-1 and AY146588-2) in *T. monococcum* may be separated from each other by more than 150 kbp. Duplications and deletions of large fragments, probably due to illegitimate recombination have resulted in the differentiation of gene loci and the striking differences in the intergenic landscape between the A^m and A genomes have revealed rapid genome divergence at orthologous LMW-GS loci (Wicker et al. 2003). In addition, their works also show that the wheat genome has undergone dramatic changes even in “recent” evolutionary times because of the various InDels present in *T. monococcum*. Therefore, the extensive variations in LMW-GS of cultivated einkorn and other wheat relatives may provide a rich gene source for improving hexaploid wheat quality properties.

Acknowledgement This research was financially supported by the grants from National Natural Science Fundation of China (30571154), the Ministry of Science and Technology of China (2002CB111300) and the Developmental Project of Science and Technology, Beijing Municipal Commission of Education (KM200410028009), respectively.

References

- Allaby RG, Banerjee M, Brown TA (1999) Evolution of the high molecular weight glutenin loci of the A, B, D, and R genomes of wheat. *Genome* 42:296–307
- Anderson OD, Greene FC (1989) The characterization and comparative analysis of high MW glutenin genes from genomes A & B of hexaploid wheat. *Theor Appl Genet* 77:689–700
- Anderson OD, Greene FC, Yip RE, Halford NG, Shewry PR, Malpica-Romero JM (1989) Nucleotide sequences of the two high-molecular-weight glutenin genes from the D-genome of a hexaploid bread wheat, *Triticum aestivum* L. cv Cheyenne. *Nucleic Acids Res* 17:461–462
- Bartels D, Thompson RD (1986) Synthesis of mRNAs coding for abundant endosperm proteins during wheat grain development. *Plant Sci* 46:117–125
- Borghgi B, Castagna R, Corbellini M, Heun M, Salamini F (1996) Breadmaking quality of einkorn wheat (*Triticum monococcum* ssp. *Monococcum*). *Cereal Chem* 73:208–214
- 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 Genet* 96:743–750
- 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, Robert LS, Kavanagh TA, Bevan MW, Thompson RD (1987) Localization of sequences in wheat endosperm protein genes which confer tissue-specific expression in tobacco. *EMBO J* 6:3559–3564
- 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 prolamins genes. *Mol Gen Genet* 216:81–90
- Corbellini M, Empilli S, Vaccino P, Brandolini A, Borghi B, Heun M, Salamini F (1999) Einkorn characterization for bread and cookie production in relation to protein subunit composition. *Cereal Chem* 76:727–733
- Cozzolino R, Giorgi SD, Fisichella S, Garozzo D, Lafiandra D, Palermo A (2001) Matrix-assisted laser desorption/ionization mass spectrometry peptide mapping of high molecular weight glutenin subunits 1Bx7 and 1Dy10 in Cheyenne cultivar. *Rapid Commun Mass Spectrom* 15:778–787
- Cunsolo V, Foti S, Saletti R, Gilbert S, Tatham S, Shewry PR (2004) Structural studies of the allelic wheat glutenin subunits 1Bx7 and 1Bx20 by matrix-assisted laser desorption/ionization mass spectrometry and high-performance liquid chromatography/electrospray ionization mass spectrometry. *J Mass Spectrom* 39:66–78
- Degidio MG, Vallega V (1994) Bread baking and dough mixing quality of diploid wheat *Triticum monococcum* L. *Ind Aliment S4:6*
- Degidio MG, Nardi S, Vallega V (1993) Grain, flour, and dough characteristics of selected strains of diploid wheat, *Triticum monococcum* L. *Cereal Chem* 70:298–303
- D’Ovidio R, Masci (2004) The low-molecular-weight glutenin subunits of wheat gluten. *J Cereal Sci* 39:321–339
- D’Ovidio R, Marchitelli C, Ercoli CL, Porceddu E (1999) Sequence similarity between allelic *Glu-B3* genes related to quality properties of durum wheat. *Theor Appl Genet* 98:455–461
- Dvorak J, McGuire PE, Cassidy B (1988) Apparent sources of the A genome of wheats inferred from polymorphism in abundance and restriction fragment length of repeated nucleotide sequences. *Genome* 30:680–689
- Forde J, Malpica JM, Halford NG, Shewry PR, Anderson OD, Greene FC, Mifflin BJ (1985) The nucleotide sequence of a HMW glutenin subunit gene located on chromosome 1A of wheat (*Triticum aestivum* L.). *Nucleic Acids Res* 13:6817–6832
- Gaut BS, Morton BR, Mccaig BC, Clegg MT (1996) Substitution rate comparisons between grasses and palms: synonymous rate differences at the nuclear gene *Adh* Parallel rate differences at the plastid gene *rbcL*. *Proc Natl Acad Sci USA* 93:10274–10279
- Gu YQ, Anderson OD, Londeore CF, Kong X, Chibbar RN, Lazo GR (2003) Structural organization of the barley D-hordein locus in comparison with its orthologous regions of wheat genomes. *Genome* 46:1084–1097
- 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
- Halford NG, Forde J, Anderson OD, Greene FC, Shewry PR (1987) The nucleotide and deduced amino acid sequences of an HMW glutenin subunit gene from chromosome 1B of bread wheat (*Triticum aestivum* L.) and comparison with those of genes from chromosomes 1A and 1D. *Theor Appl Genet* 75:117–126
- Halford NG, Field JM, Blair H, Urwin P, Moore K, Robert L, Thompson R, Flavell RB, Tatham AR, Shewry PR (1992)

- Analysis of HMW glutenin subunit encoded by chromosome 1A of bread wheat (*Triticum aestivum* L.) indicates quantitative effects on grain quality. *Theor Appl Genet* 83:373–378
- Hsia CC, Anderson OD (2001) Isolation and characterization of wheat omega-gliadin genes. *Theor Appl Genet* 103:37–44
- Ikeda TM, Nagamine T, Fukupka H, Yano H (2002) Identification of new low-molecular-weight glutenin subunit genes in wheat. *Theor Appl Genet* 104:680–687
- Johal J, Gianibelli MC, Rahman S, Morell MK (2004) Characterization of low-molecular-weight glutenin genes in *Aegilops tauschii*. *Theor Appl Genet* 109:1028–1040
- Kreis M, Shewry PR (1989) Unusual features of cereal seed protein structure and evolution. *BioEssays* 10:201–207
- Kumar S, Tamura K, Nei M (2004) *MEGA3*: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform* 5(2):150–163
- Lauriere M, Bouchez I, Doyen C, Eynard L (1996) Identification of glycosylated forms of wheat storage proteins using two-dimensional electrophoresis and blotting. *Electrophoresis* 17:497–501
- Lew EJJ, Kuzmicky DD, Kasarda DD (1992) Characterization of low molecular weight glutenin subunits by reversed-phase high-performance liquid chromatography, sodium dodecyl sulfate-polyacrylamide gel electrophoresis, and N-terminal amino acid sequencing. *Cereal Chem* 69:508–515
- Li W, Wan Y, Liu Z, Liu K, Li B, Li Z, Zhang X, Dong Y, Wang D (2004) Molecular characterization of HMW glutenin subunit allele 1Bx14: further insights into the evolution of Glu-B1-1 alleles in wheat and related species. *Theor Appl Genet* 109:1093–1104
- Lu C, Yang W, Zhang W, Lu B (2005) Identification of SNPs and development of allelic specific PCR marker for high molecular weight glutenin subunit D^x1.5 from *Aegilops tauschii* through sequence characterization. *J Cereal Sci* 41:13–18
- 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 Physiol* 118:1147–1158
- McDonald MB, Elliot LJ, Sweeney PM (1994) DNA extraction from dry seeds for RAPD analyses in varietal identification studies. *Seed Sci Technol* 22:171–176
- 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. *Nucl Acids Res* 16:11376
- Sachidanandam R et al (The international SNP Map Working Group) (2001) A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* 409:928–933
- Sainova OV, Mekhedov SL, Zhelnin LG, Khokhlova TA, Anan'ev EV (1993) Nucleotide sequence of the barley C-hordein gene. *Genetika* 29:1070–1079
- Shewry PR, Halford NG (2002) Cereal seed storage proteins: structures, properties and role in grain utilization. *J Exp Bot* 53:947–958
- Shewry PR, Halford NG, Tatham AS (1992) High molecular weight subunits of wheat glutenin. *J Cereal Sci* 15:105–120
- Tilley KA, Schofieldt JD (1995) Detection of phosphotyrosine in the high Mr subunits of wheat glutenin. *J Cereal Sci* 22:17–19
- Tilley KA, Lookhart GL, Hosnoney RC, Mawhinney TP (1993) Evidence for glycosylation of the high molecular weight glutenin subunits 2, 7, 8, 12 from Chinese Spring and TAM105 wheat. *Cereal Chem* 70:602–606
- Tranquilli G, Cuniberti M, Gianibelli MC, Bullrich L, Larroque OR, MacRitchie F, Dubcovsky J (2002) Effect of *Triticum monococcum* glutenin loci on cookie making quality and on predictive tests for bread making quality. *J Cereal Sci* 36:9–18
- Vicente-Carbajosa J, Beritashvili DR, Kraev AS, Skryabin KG (1992) Conserved structure and organization of B hordein genes in the Hor2 locus of barley. *Plant Mol Biol* 18:453–458
- Wicker T, Yahiaoui N, Guyot R, Schlagenhauf E, Liu ZD, Dubcovsky J, Keller B (2003) Rapid genome divergence at orthologous low molecular weight glutenin loci of A and A^m genomes of wheat. *Plant Cell* 15:1186–1197
- Yan Y, Prodanovic S, Mladenov N, Milovanovic M (1999) Genetic control of low molecular weight glutenin subunits in wheat by A-PAGE analysis. *Cereal Res Commun* 27:251–257
- Yan Y, Hsam SLK, Yu JZ, Jiang Y, Ohtsuka I, Zeller FJ (2003) HMW and LMW glutenin alleles among putative tetraploid and hexaploid European spelt wheat (*Triticum spelta* L.) progenitors. *Theor Appl Genet* 107:1321–1330
- Zhang W, Gianibelli MC, Ma W, Rampling L (2003) Identification of SNPs and development of allele-specific PCR markers for γ -gliadin alleles in *Triticum aestivum*. *Theor Appl Genet* 107:130–138
- 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 H, Guo A, Hu S, Fan S, Zhang D, Ren S, Wang R (2004) Development of primers specific for LMW-GS genes at *Glu-D3* and *Glu-B3* Loci and PCR amplification. *Acta Agronom Sinica* 30:126–130