Theor Appl Genet (2005) 111: 272–280
DOI 10.1007/s00122-005-2021-9

ORIGINAL PAPER

Zhen Luo · Fanguo Chen · Deshun Feng · Guangmin Xia

LMW-GS genes in *Agropyron elongatum* and their potential value in wheat breeding

Received: 26 January 2005 / Accepted: 29 March 2005 / Published online: 31 May 2005

Abstract To study the usefulness of low-molecular-weight glutenin subunits (LMW-GS) of *Agropyron elongatum* (Host) Nevski to wheat (*Triticum aestivum* L.) quality improvement, we characterized LMW-GS genes of *A. elongatum*. Nine LMW-GS genes of *A. elongatum*, which were named *AeL1* to *AeL9*, were cloned by genomic PCR. After sequencing, we obtained complete open reading frames from *AeL2* to *AeL8* and partial genes of *AeL1* and *AeL9*. All nine sequences are homoeologous to those of wheat and related grasses. Comparison of the deduced amino acid sequences with those of published LMW-GS suggests that the basic structures of all the subunits are very similar. However, except for AeL4 and AeL5, which contain the identical N-terminal sequence with LMW-m, other LMW-GS sequences separated from *A. elongatum* cannot be classified according to previous criteria for the three types: LMW-m (methionine), LMW-s (serine), and LMW-i (isoleucine), and then 12 groups. In addition, there are some characters in the LMW-GS sequences of *A. elongatum*: AeL2, AeL3, and AeL6 involve a Cys residue in the signal peptide respectively, which is absent in most of LMW-GS; AeL3, AeL6, AeL8, and AeL9 start their first Cys residues in the N-terminal repetitive domains, respectively; both AeL2 and AeL5 have nine Cys residues, with an extra Cys residue in the N-terminal repetitive domain and the repetitive and glutamine-rich domain; AeL2, AeL3, AeL6, and AeL9 comprise long repetitive domains. Phylogenetic analysis indicates that there is a relatively weak sequence identity between the LMW-GS genes from *A. elongatum* cloned in this study and those reported from other plants. Three LMW-GS sequences, AeL2, AeL3, and AeL6, are clustered to *Glu-A3* from wheat than to those from other plants. The possible use of these genes in relation to the high quality of hybrid wheat is discussed.

Keywords *Agropyron elongatum* · Somatic hybrid line · LMW-GS · Coding sequence · Evolution · Wheat quality

Communicated by P. Langridge

Z. Luo · F. Chen · D. Feng · G. Xia (✉)
School of Life Sciences, Shandong University,
Jinan, 250100, PR China
E-mail: xiagm@sdu.edu.cn
Tel.: +86-531-8364525
Fax: +86-531-8565610

Introduction

It is well known that seed-storage proteins determine the quality of common wheat flour. These storage proteins are classified into two major groups, glutenin and gliadin; the latter are monomeric proteins, whereas glutenins are aggregates of high-molecular-weight subunits (HMW-GS) and low-molecular-weight subunits (LMW-GS) held together by disulfide bonds (Payne 1987). Of total storage proteins of wheat, about 10% are HMW-GS and 40% are LMW-GS. LMW-GS are classically divided into B, C, and D groups according to the molecular weight and isoelectric point (Jackson et al. 1983). The genes encoding LMW-GS occur on the short arms of group 1 chromosomes at the *Glu-A3*, *Glu-B3*, and *Glu-D3* loci that are tightly linked to the *Gli-1* locus (Singh and Shepherd 1988; Pogna et al. 1990). The LMW-GS gene family, due to its complexity and heterogeneity, has so far been characterized in only a very limited number of common wheat and durum wheat cultivars (Ikeda et al. 2002), as well as the related grasses (Rodríguez-Quijano et al. 1997; Xu et al. 2004). D'Ovidio et al. (1992) first cloned an LMW-GS gene from *Triticum durum* by PCR. This direct cloning of genomic sequences provides an easy approach to obtain the genes from different cereals and grasses. It is known that LMW-GS sequences have been classified into LMW-s (serine), LMW-m (methionine), and LMW-i (isoleucine) types based on the first amino acid of mature subunits (Lew et al. 1992; Cloutier et al. 2001). Recently, Ikeda et al. (2002) divided LMW-GS sequences into 12 groups

and then classified the 12 groups into six types, based on N- and C-terminal domains and positions of functional Cys residues, respectively.

It is generally accepted that HMW-GS is mainly responsible for bread-making quality. LMW-GS, containing more Cys residues than HMW-GS, also plays a significant role in the formation of large polymers. Griffin (1989) found that HMW-GS played only a minor role in regulating environmental variability for bread making, the whole-gluten protein fractions appearing to be important, not just the HMW-GS. Maruyama-Funatsuki et al. (2004) reported that LMW-GS enrichment of durum dough resulted in stronger, more-elastic dough than did HMW-GS enrichment, confirming that LMW-GS is the major contributor to durum dough strength. The result of stepwise regression analysis also indicated that the quality prediction by using both two types of glutenins was more reliable than that of each component (Ge et al. 2002). Other studies have also shown that allelic variation of HMW-GS and LMW-GS are both associated with difference in the technological qualities of wheat flour (Autran et al. 1987; Payne 1987; Gupta et al. 1989; Nieto-Taladriz et al. 1994).

Agropyron elongatum (syn. *Lophopyrum elongatum*; *Thinopyrum ponticum*; $StStE^{e}$ E^{b} E^{x}, $2n = 70$) is an important wild source for wheat breeding because it possesses a high content of seed proteins and resistance to stress and disease (Xia et al. 2003). There are a series of high-quality hybrid cultivars derived from sexual hybridization between common wheat and *A. elongatum*, e.g., 'Xiaoyan 6' (Zhou et al. 1995) and 'Xiaoyan 54' (Liu et al. 2001). In addition, we obtained some somatic hybrid strains between wheat and *A. elongatum* with higher qualities than their parental wheat (Xia et al. 2003; Feng et al. 2004). The high-quality hybrids from both sexual and somatic hybridizations between wheat and *A. elongatum* contain different HMW-GS, which are absent in the parental wheat (Fan and Guo 2000; Zhao et al. 2003). It was shown that the coding sequences of excellent HMW-GS, 1Bx14 and 1By15 in hybrid lines of 'Xianyan 6' and 'Xiaoyan 54', as well as homologues of 1Bx13 and 1By16 in the somatic hybrid lines, did not originate from *A. elongatum* (Fan and Guo 2000; Feng et al. 2004). So the quality improvements of the hybrids may be related with the LMW-GS, gliadin of *A. elongatum*. It is necessary to investigate the origin of these LMW-GS and other storage proteins in the hybrids through comparing their coding sequences among *A. elongatum*, parent wheat and hybrids. In this paper, we characterize the nuclear acid sequences of LMW-GS in *A. elongatum*.

Materials and methods

Plant materials

Seeds of *A. elongatum* and *Triticum aestivum* 'Jinan 177' were stored in our laboratory. They were planted in greenhouse separately in order to avoid cross-pollination from other grasses and cultivars.

Isolating and cloning of LMW-GS genomic DNA

Total DNA was prepared from leaves of *A. elongatum* by the CTAB method according to Doyle and Doyle (1990). LMW-GS gene-specific degenerate primers were designed based on published LMW-GS gene sequence data (Y14104, D'Ovidio et al. 1997; AJ007746, D'Ovidio et al. 1999; U86026–U86030, Anderson et al. 2001; AB062860–AB062878, Ikeda et al. 2002): P1 (5′-GCA TCAA/GACCAAGCAAA/CAC-3′) and P2 (5′-TTAT CAGTAGG/CCACCAACTC-3′). PCR was performed in a total volume of 20 µl containing 1.5 mmol/l of $MgCl_2$, 0.1 mmol/l of each dNTP, 10 pmol/l of each primer, and 1 U of LA GC *Taq* DNA polymerase (a highly fidelity polymerase; TaKaRa, Japan), 1X PCR buffer (TaKaRa), and 300 ng of total DNA. The reaction was performed according to the following protocol: denaturing at 95°C for 5 min, 10 cycles of 94°C for 30 s, 68°C for 1 min (each cycle decrease 1°C), 72°C for 90 s; 35 cycles of 94°C for 30 s, 58°C for 30 s, 72°C for 90 s; and a final extension at 72°C for 7 min.

The purified PCR products were separated in 4% gel polyacrylamide electrophoresis and retrieved by DNA purification kit (Banner, Beijing, China).

Sequencing and analyzing of the LMW-GS genes

All reclaimed PCR products were cloned into pUCm-T vectors (Bioasia, Shanghai, China) and then transformed into *Escherichia coli* DH10B competent cells. Cloning and transforming methods are followed Sambrook et al. (1989). DNA sequencing was performed by a commercial company (Bioasia). Over two repeat sequences of each clone were used for analysis. The DNAMAN program and relative programs found on the National Center for Biotechnology Information and European Bioinformatics Information networks were used for the sequence analysis, as well as the alignments of deduced amino acid sequences of LMW-GS genes.

Results

Sequencing and characterization of the nine LMW-GS genes

Primary structure of the sequence

DNA profile indicated that the compositions of LMW-GS genes in *A. elongatum* were different from 'Jinan 177' (Fig. 1). Nine of the LMW-GS gene segments, which were divided on the polyacrylamide gel, were cloned and sequenced. All the nine sequences, as a whole, have a similar structure of the sequence with other LMW-GS

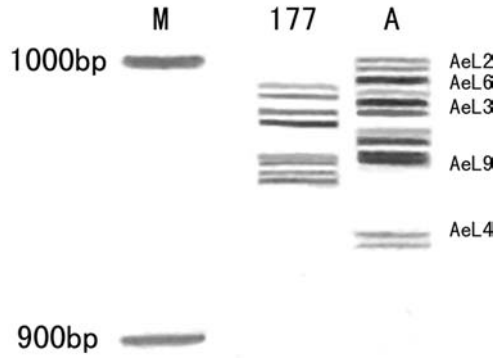

Fig. 1 PCR amplification of low-molecular-weight (LMW)-glutenin gene open reading frames (ORFs) from *A. elongatum* and common wheat. *A Agropyron elongatum*, *M* DNA marker (50-bp ladder), *JN177* wheat cultivar 'Jinan 177'. *AeL2*, *AeL3*, *AeL6*, and *AeL9*: LMW-glutenin gene fragments cloned

reported previously. However, there are distinctions at the N- and C-terminal domains and repeating sequences, including some sequence site mutations, base substitutions, and deletions/insertions in different subunits, and so on. Seven of them contain complete open reading frames (ORFs) of the genes; two were partial genes absent of 5′ terminals. The general information of the eight complete mature proteins deduced from *A. elongatum* sequences are summarized in Table 1.

Characterization of the amino acid sequences

The deduced amino acid sequences of the nine genes are listed in Fig. 2. Comparison of the amino acid sequences between *AeL1* and *AeL9* and those published from wheat and other relative grasses shows that they have similar structures: signal peptide domain, N-terminal repetitive domain, and C-terminal domain. Their C-terminal domains were more conservative than N-terminal domains as in most known LMW-GS sequences. The different-sized polypeptides in the nine genes are due to the presence of various deletions/insertions within repetitive and glutamine-rich domains (Fig. 2). In the nine sequences of *A. elongatum*, AeL1, lacking signal peptide and partial N-terminal sequences, has similar sequence with wheat BAB78753 (Ikeda et al. 2002) and *Aegilops tauschii* AAW28853 (Huang et al., submitted). As for the other eight N-terminal sequences (Table 1), AeL4 and AeL5, having a uniform N-terminal conserved domain, METSCIPGL, share the same N-terminal sequence with the LMW-m (Lew et al. 1992). AeL4 shows a relative strong sequence identity (95.30%) with BAB78760 (Ikeda et al. 2002); and AeL5 reveals a weak sequence identity (85.55%) to wheat AAV92002 (Ozdemir and Cloutier, submitted). Both are different from published LMW-m type genes at the fifth position, where a Cys residue replaces a His residue.

The remaining six sequences, AeL2, AeL3, AeL6, AeL7, AeL8, and AeL9, cannot be divided into any of the normal LMW-GS types by previous criteria (Lew et al. 1992; Cloutier et al. 2001; Ikeda et al. 2002), because they lack typical N-terminal conserved sequences. AeL2 to AeL3 and AeL6 possess the same N-terminal sequence that was not reported previously, and have a same Cys residue at position 7 in the signal peptide (Fig. 3), as well as contain a homoeologous sequence encoded by *Glu-A3* (Zhang et al. 2004; AAS10189, AAS10190). The similarity between AeL2 and AAS10189 is 56.99%, whereas AeL3 and AeL6 have weak identity to AAS101906 with frequencies of 62.30% and 61.19%, respectively.

The frequency of homologous sequence between *AeL2* and *AeL3* is 88.99%. Both have a TTTTTT sequence in N-terminal repetitive domain, which likely influences the hydrophobicity of the subunits (Fig. 2). The similar hydrophobic structure was reported by Ikeda et al. (2002), with a different sequence. In contrast to the typical sequences of LMW-GS genes, *AeL2* has a partially extra repeated signal peptide TFLICALLAIAATSAIAQ in the N- terminal repetitive domain, which makes the number of Cys residues in it become two (Fig. 2). In order to avoid the feint from second structure of DNA by wrong PCR, dimethylsulfoxide was added to the PCR system. It was showed that the gene fragment of *AeL2* is longer than *AeL3* (Fig. 4). Thus, the sequence present actually, which maybe caused by unequal crossing over in the LMW-GS evolution (Alkan et al. 2002).

Table 1 Classification of the deduced amino acid sequences of the eight LMW-GS genes from *Agropyron elongatum*

Genes	GenBank accession	Amino acid number	N-terminal sequence	C-terminal sequence	Similar published sequence (locus, sequence identity)
AeL2	AY646285	345	LPISQQQQ	VGTGVGAY	AAS10189 (Glu-A3, 56.99%)
AeL3	AY606257	315	LPISQQQQ	VGTGVGAY	AAS10190 (Glu-A3, 62.30%)
AeL4	AY639024	295	METSCIPG	VGSGVGAY	BAB78760 (95.30%)
AeL5	AY724441	248	METSCIPG	VGTGVSAY	AAV92002 (85.55%)
AeL6	AY724436	323	LPISQQQQ	VLAPEWCY	AAS10190 (Glu-A3, 61.19%)
AeL7	AY724438	239	QQQLPQQP	VGTGVSAY	AAO53264 (75.45%)
AeL8	AY724439	277	QQQLPQQP	VGTGVGAY	AAO53264 (86.64%)
AeL9	AY724440	315	LPISQQQQ	VGTGVGAY	AAV92080 (55.44%), AAO17157(55.44%)

Fig. 2 Comparison of the deduced amino acid sequences of the LMW-glutenin subunits (GS) genes. Nine sequences were aligned. Cys are *boxed*. N-terminal conserved domain of AeL4 and AeL5 are *underlined*. AeL2- and AeL3-specific hydrophobic clusters are *double underlined*

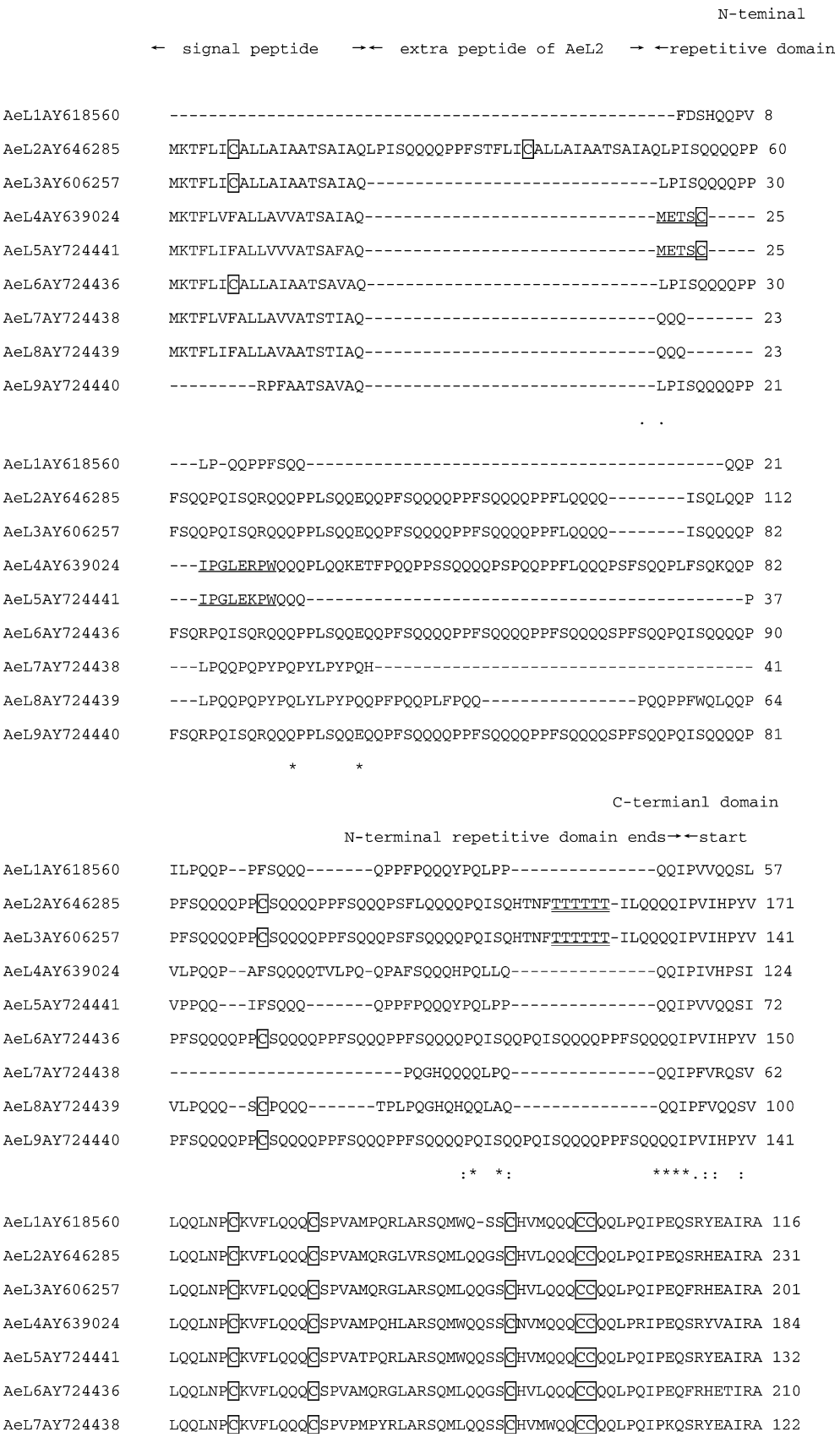

Fig. 2 (Contd.)

```
AeL8AY724439     LQHLNPCKVFLQQQCSPVAMPYRLARSQMLQQSSCQVMQQQCCQQLPQIPEQSRYEAIHA 160
AeL9AY724440     LQQLNPCKVFLQQQCSPVAMQRGLARSQMLQQGSCHVLQQQCCQQLPQIPEQFRHETIRA 201
                  **:***************.    *.**** * .**:*: ********:**:* *: :*:*

AeL1AY618560     IIYSIILQEQQQ--GFVQPQQQQPQQAGQGVSQPQQQSQQQQQLGQCFFQLPQQQQQLGQ 174
AeL2AY646285     IVYSIILQEQQQGQGFIQPQQQQPQQSAQCVSQPQQQSQQ--QLGR----QPQQQQ-LGQ 284
AeL3AY606257     IVYSIIPQEQQQGQGFIQPQQQQPQQSAQCVSQPQQQSQQ--QLGQ----QPQQQQ-LGQ 254
AeL4AY639024     IILSIILQEQQ---GFVQPQQQQPQQSVQGVYQPQQQSQQ--QLGQCSFQQPQQQ--LGQ 237
AeL5AY724441     IIYSIILQEQQQ--GFVQPQQQQPQQSGQGVSQHQQQSQQQQQLGQCSFQQPQQQQ-LGQ 189
AeL6AY724436     IVYSIIPQEQQQGQGFIQPQQQQPQQSAQRVSQPQQQSQQ--QLGQ----QPQQQQ-LGQ 263
AeL7AY724438     IVYSIILQEQQQGQGFVQPQQQ-PIQSVQGVSQPQQQSWQQ-QFVQCSFQQPQPQQ-LGQ 179
AeL8AY724439     IVYSIILQEQQQGQGFVQPQQQ-PIQSVQGVSQPQQQSGQQ-QLVQCSFQQPQPQQ-LGQ 217
AeL9AY724440     IVYSIIPQEQQQGQGFIQPQQQQPQQSAQRVSQPQQQSQQ--QLGQ----QPQQQQ-LGQ 254
                    *: *** ****   **:***** * *: * * * **** *  *: :    ** *  ***

AeL1AY618560     QP-QQQQVPQGT-LQPHQIAQLDVMTSIALRTLPTMCSVNVPLYGTTTIVPFGVGTGVGA 232
AeL2AY646285     QPQQQQQVLQGTFLQPHQIAQLEAMTSIALRTLPRMCSVNVPLYGTASSVSFGVGTGVGA 344
AeL3AY606257     QPQQQQQVLQGTFLQPHQIAQLEAMTSIALRTLPRMCSVNVPLYGTASSVSFGVGTGVGA 314
AeL4AY639024     QP-QQQ-HVQGT-LQPHQIARLEVMTSIALSTLPTMCSVNVPLYSSITSVPFGVGSGVGA 294
AeL5AY724441     QP-QQQ-VPQGIFLQPHQICQLEVMTSIALRTLPTMCSVNVPLYSSTTIVPFSVGTGVSA 247
AeL6AY724436     QPQQQQQVLQGTFLQPHQIAQLEAMTSIALRTLPRMCSVNVPLYGTASSVSFVLAP-EWC 322
AeL7AY724438     QP-QQQQVPQGAFLQPHQIAQLEVRASIALHTLPRMCNVNVPLYGASTSVPFGVGTGVSA 238
AeL8AY724439     QP-QQQQVPQGALLQPHQIAQLEVRTSIALHTLPRMCNVNVPLYGASTSVPFGVGTGVGA 276
AeL9AY724440     QPQQQQQVLQGTFLQPHQIAQLEAMTSIALRTLPRMCSVNVPLYGTASSVSFGVGTGVGA 314
                    ** ***   **  ******.:*:. :**** *** **.******.: : *.* :..   .

AeL1AY618560     Y 233
AeL2AY646285     Y 345
AeL3AY606257     Y 315
AeL4AY639024     Y 295
AeL5AY724441     Y 248
AeL6AY724436     Y 323
AeL7AY724438     Y 239
AeL8AY724439     Y 277
AeL9AY724440     Y 315
                 *
```

AeL7 and *AeL8* share the same N-terminal sequence QQQLPQQP that was not found before (Table 1). Both exhibit a weak similarity to one of the LMW-GS in the *A. intermedium* (AAO53264, Xu et al. 2004), with frequencies of 75.45% and 86.64%, respectively. AeL9 contains only a partial signal peptide, RPFAATSAVAQ (Fig. 2). It is also a novel type, having a very weak sequence identity (55.4%) to two LMW-GS genes from wheat, AAV92080 (Ozdemir and Cloutier, submitted) and AAO17157 (Wicker et al. 2003).

Cys residues in deduced amino acid sequences of the eight genes

To the mature peptides, AeL2, AeL3, AeL6, AeL8 and AeL9 start their first Cys residues in the N-terminal repetitive domains (Fig. 3), the same as a few in wheat (Masci et al. 1998; Ikeda et al. 2002), but different from most other LMW-GS. AeL2 possesses an additional Cys residue in the N-terminal repetitive domains, in agreement with LAi1 in an addition line of wheat with *A.*

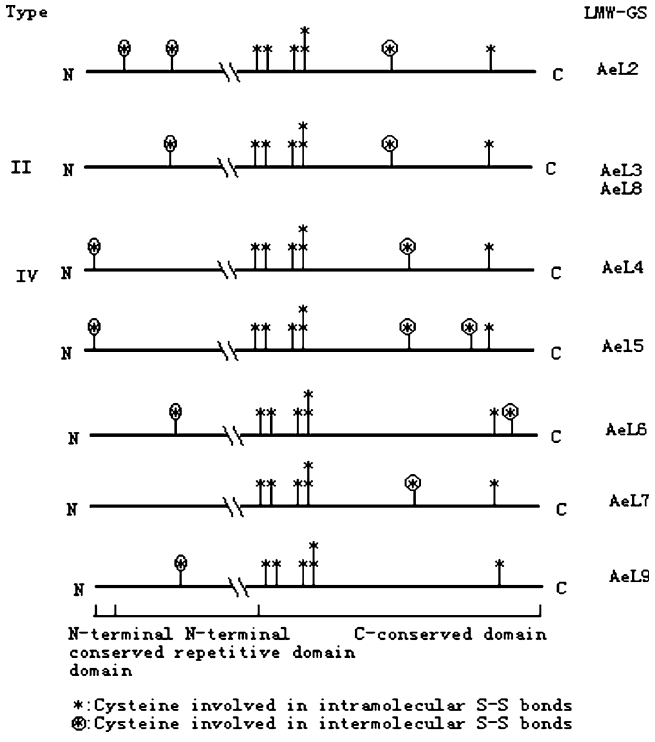

Fig. 3 Classification of the deduced amino acid sequences of the LMW-GS genes based on the distribution of Cys. The positions of Cys are shown as *asterisks*

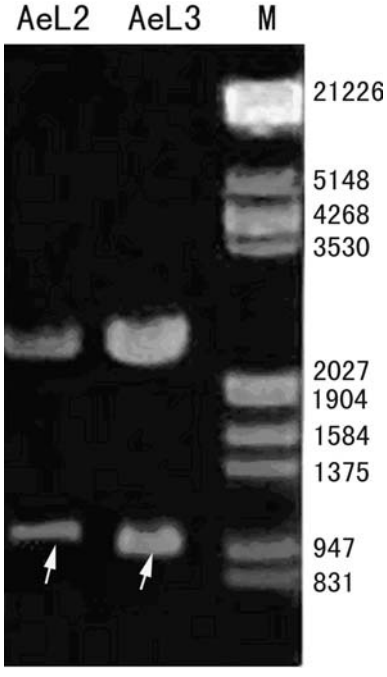

Fig. 4 The result of *AeL2* and *AeL3* plasmids digested with *Eco*RI and *Hin*dIII. *M* DNA marker (λDNA digested with *Eco*RI and *Hin*dIII)

intermedium (Xu et al. 2004); AeL5 has an extra Cys residue in the repetitive and glutamine-rich domain, which has not been reported before; both AeL6 and AeL9 lack Cys residues in this domain, but AeL6 has an extra Cys residue in the C-terminal conserved domain. Statistic analysis of eight LMW-GS sequences reveals that, except Cys in the signal peptide, most contain eight Cys residues, which are conservative in most of the previously published LMW-GS sequences (Ikeda et al. 2002), and three possess a new locus of Cys, respectively (Fig. 3). On the other hand, six of these loci are relatively conservative, devoted to intra-molecular disulfide bonds, as described by Andrea et al. (2001). Whereas others, which are related to inter-molecular disulfide bonds, alter with a high frequency (Fig. 3).

Evolution relationship in the LMW-GS coding sequences

The ORFs of the seven completely sequenced genes were used to produce a dendrogram, together with some other previously characterized LMW-GS genes from *T. aestivum* (Van et al. 1995, X84959-X84960; Zhang et al. 2004, AY453156), *T. durum* (AJ293097, AJ293099, D'Ovidio et al. 1992), *A. tauschii* (AY585350, AY585356, Johal et al. 2004), and *A. intermedium* (AY214452, AY214454, and AY214458, Xu et al. 2004) (Fig. 5). *AeL2* clustered with *AeL3* and *AeL7* with *AeL8*, respectively. Some of LMW-GS genes of *A. elongatum*, e.g., *AeL2*, *AeL3*, and *AeL6*, are related with one of *Glu-A3*. Cloutier et al. (2001) have reported positive effects for the *Glu-A3* locus on dough strength. The result implies that *A. elongatum* is important as a wild resource in the wheat quality improvement.

Pseudogenes

The *AeL1* from *A. elongatum* contains an in-frame stop codon resulting from the mutation of CAG (GLn) to TAA (stop) on residue 88 in the N-terminal repetitive domain. The *AeL4* comprises two in-frame stop codons, which originate from mutation-replacing of CAG (Gln) by TAG (stop) and AAG (Lys) by TAG (stop), located on the residues 196 and 247 in the repetitive and glutamine-rich domains, respectively. The mutations in *AeL1* and the former of *AeL4* come from base transition, whereas the latter of *AeL4* derives from base transversion. These mutations indicate that the two genes are unlikely to express full-length proteins. Pseudogenes are common in the cereal prolamins (Anderson and Hsia 2001), including LMW-GS (Benmoussa et al. 2000; Ikeda et al. 2002; Johal et al. 2004).

Discussion

Although both HMW-GS and LMW-GS are responsible for the visco-elasticity of wheat flour, scientists have paid major attention to HMW-GS and their encoding genes in the past 30 years. There have been many fewer reports on LMW-GS than on HMW-GS. We selected a

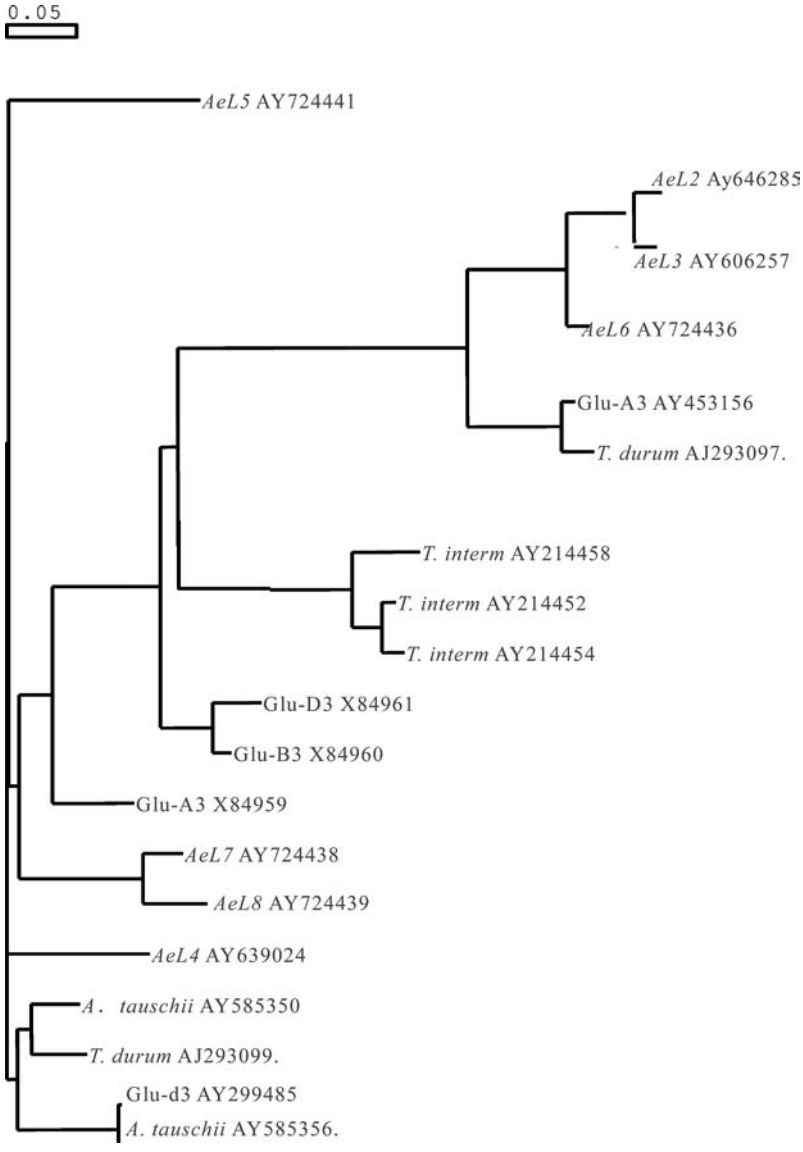

Fig. 5 Phylogenetic tree based on the ORFs of *A. elongatum* GS and representatives of published LWM-GS alignments. The GenBank accession numbers are to the *right of the names* of subunit genes

closely related grass (*A. elongatum*), which has been successfully used for wheat quality breeding (Zhong et al. 2002; Feng et al. 2004), to investigate the possible effect of its LMW-GS to wheat quality.

According to the first amino acid residue of the predicted mature protein, LMW-GS have been classified into three classes: type s, type m, and type i (Lew et al. 1992; Cloutier et al. 2001). Ikeda et al. (2002) further divided wheat LMW-GS genes into 12 groups based on alignment of the conserved N- and C-terminal domains of the deduced amino acid sequences. In our research on the LMW-GS genes from *A. elongatum*, only two genes, *AeL4* and *AeL5*, are in the range of groups 8 or 9, based on the N-terminal domain sequences (without considering the C-terminal domain sequences, which are different from any of those in the 12 groups). There are only weak sequence identity between *A. elongatum* and other published LMW-GS genes (Table 1). This means that most LMW-GS genes of *A. elongatum* cloned are unique and do not complete suitable the classes in wheat, although these classes in favor of LMW-GS genes from bread wheat, durum wheat, and some relatives, such as *A. tauschii*, and so on (Johal et al. 2004).

HMW-GS and LMW-GS, having both intra- and inter-molecular disulfide bonds, result in the formation of the "glutenin polymer" (Masci et al. 1998). The size distribution and composition of the polymers in the glutenin fraction are strongly correlated to flour technological characteristics (Wrigley 1996; Masci et al. 2002). Thus, it is important to define all Cys residues in the primary structures of the LMW-GS. Most of LMW-GS groups contain eight Cys residues, which are conserved among all of the previously published LMW-GS sequences (Ikeda et al. 2002), besides LAi1 from *A. intermedium*, which has nine Cys residues and was suggested relative with the high quality of several addition lines of wheat/*A. intermedium* (Xu et al. 2004). It is found that two LMW-GS sequences, AeL2 and AeL5, in

the *A. elongatum* containing nine Cys residues (Fig. 3). Both can also form three extra inter-molecular disulfide bonds and are also very important to the visco-elasticity of flour (Xu et al. 2004).

According to previous reports (Shewry and Tatham 1997; Gianibelli et al. 2001), six of the seven Cys residues of C-terminal domain could form three intra-molecular disulfide bonds, which are quite stable in these (Fig. 3) and other known LMW-GS (Ikeda et al. 2002; Xu et al. 2004). The other Cys residue in the C-terminal domain was also stable in that of *A. intermedium* (Xu et al. 2004) and *T. aestivum* (Ikeda et al. 2002), but variable from zero to two in *A. elongatum* (Fig. 3). In the N-terminal domain, a significant alteration appeared in *A. intermedium* (Xu et al. 2004), *T. aestivum* (Ikeda et al. 2002) and *A. elongatum*, with the number ranging from zero to two (Fig. 3). The variation in both loci and numbers of the Cys residues in *A. elongatum* are more than in other published plants, which provide abundant samples to investigate the relationship between the structure and function.

Based on the position of functional Cys, Ikeda et al. (2002) also divided wheat LMW-GS groups into six types. Of our eight ORFs, both AeL3 and AeL8 share the same functional Cys residues with type II; AeL4 has also the first and seventh Cys residues uniform with type IV (Fig. 3). Thus, all three genes can be included into the types of Ikeda et al. (2002). AeL2 contain two Cys residues in agreement with type II, except for an additional functional Cys on the N-repetitive domain. AeL9 start its first Cys on the repeat domain, in consistent with those in types I and II (Ikeda et al. 2002), but its seventh Cys is absent, so that we cannot divide it into both types. AeL5 has both Cys residues identity to type IV, but containing an extra functional Cys on C-terminal. AeL6 has the primary structure similar to type II, except lacking the seventh Cys and giving one after the eighth. AeL7, lacking the first Cys as described in type IV and revealing the same Cys as type I, also cannot be classed into any type. Thus, Cys locations of *A. elongatum* reveal the diversity of functional structure on LMW-GS sequences. Johal et al. (2004) reported that all *A. tauschii* cloned could be included in the four of six types. Our result provides new information for further functional Cys class in different LMW-GS sequences.

Besides the Cys residues, a long N- terminal repetitive domain was also speculated to have positive influence on quality of wheat flour (Masci et al. 1998, 2000). Significant variation occurs in the number of repeating units and their amino acid compositions of AeL1 to AeL9. Of the nine genes, four (AeL2, AeL3, AeL6 and AeL9) have a long N- terminal repetitive domain, containing over 16 repeating unit (Fig. 2), e.g., AeL2, with nine Cys residues, has 20 repeating units (Fig. 2), which is similar to LAi1 (Xu et al. 2004) and the 42-kDa LMW-GS from wheat (Masci et al. 1998), both were deduced responsible for a high quality of wheat, containing 25 and 26 repeating units, respectively.

From above analysis, we can infer that the high quality in the somatic hybrid lines (Zhao et al. 2003) is likely relative with LMW-GS of *A. elongatum*. It is worth investigating further.

Acknowledgements National Natural Science Foundation of China, no. 30370857, Major Project of Ministry of Education in China and National 863 High Technology Research and Development project no. 2001AA241032 supported this study.

References

Alkan C, Bailey JA, Eichler EE, Sahinalp SC, Tuzun E (2002) An algorithmic analysis of the role of unequal crossover in alpha-satellite DNA evolution. Genome Inform Ser Workshop Genome Inform 13:93–102

Anderson OD, Hsia CC (2001) The wheat γ-gliadin genes: characterization of ten new sequences and further understanding of γ-gliadin gene family structure. Theor Appl Genet 103:323–330

Anderson OD, Hsia CC, Adalsteins AE, JL Lew E, K-asarda DD (2001) Identification of several new classes of low-molecular-weight wheat gliadin-related proteins and genes. Theor Appl Genet 103:307–315

Andrea O, Francesca S, Aldo C (2001) Role of individual disulfide bonds in the structural maturation of a low molecular weight glutenin subunit. J Biol Chem 276(34):32322–32329

Autran JC, Laignelet B, Morel MH (1987) Characterization and quantification of low-molecular-weight glutenins in durum wheat. Biochimie 69:699–711

Benmoussa M, Vezina LP, Page M, Yelle S, Laberge S (2000) Genetic polymorphism in low molecular weight glutenin genes from *Triticum aestivum* variety Chinese Spring. Theor Appl Genet 100:789–793

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

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 CL, Porceddu E (1999) Sequence similarity between allelic Glu-B3 genes related to quality properties of durum wheat. Theor Appl Genet 98:455–461

Doyle JJ, Doyle JI (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15

Fan SH, Guo AG (2000) A study on the origin of HMW glutenin subunit 14 and 15 in Xiao Yan 6. Acta Univ Agric Boreali-Occidentalia (in Chinese with English abstract) 28(6):1–5

Feng DS, Xia GM, Zhao SY, Chen FG (2004) Two quality-associated HMW glutenin subunits in a somatic hybrid line between *Triticum aestivum* and *Agropyron elongatum*. Theor Appl Genet 110: 136–144. DOI 10.1007/s00122-004-1810-x.

Ge SJ, Xie LQ, Wang JH, Lu SY, Zhang RZ, Zhou SX (2002) Studies on relations between the low molecular weight glutenin subunits and agronomic and quality traits in common wheat. J Agric Univ Hebei (in Chinese with English abstract) 225(3):6–9

Gianibelli MC, Larroque OR, Macritchie F, Wrigley CW (2001) Biochemical, genetic, and molecular characterization of wheat glutenin and its component subunits. Cereal Chem 78:635–646

Griffin WB (1989) Influence of high-molecular-weight gluten in subunits on environmentally affected variations of wheat quality. In: Proceedings of the Cereal Science Conference, pp 46–50

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

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

Lew EJL, Kuzmicky DD, Kasarda DD (1992) Characterization of low-molecular-weight glutenin subunits by reversed-phase high-performance liquid chromatography, sodium dodecyl sulfate-polyacrylamide gel electrophoresis, and N-terminal amino-acid sequencing. Cereal Chem 69:508–515

Liu QY, Tong YP, Sun JH, Li JY (2001) Study on optimal sowing density and fertilizer application rate for cultivating high quality wheat Xiaoyan 54 in Beijing Area. Acta Bot Boreal–Occident Sin (in Chinese with English abstract) 21(3):92–496

Maruyama-Funatsuki W, Takata K, Nishio Z, Tabiki T, Yahata E, Kato A, Saito K, Funatsuki H, Saruyama H, Yamauchi H (2004) Identification of low-molecular weight glutenin subunits of wheat associated with bread-making quality. Plant Breed 123:355

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 glutenin polymer. Plant Physiol 118:1147–1158

Masci S, D'Ovidio R, Lafiandra D, Kasarda DD (2000) A 1B coded low-molecular-weight glutenin subunit associated with quality in durum wheats show strong similarity to subunit present in some bread wheat cultivars. Theor Appl Genet 100:396–400

Masci S, Rovelli L, Kasada DD, Vensel WH, Afiandra DL (2002) Characterization and chromosomal localization of C-type LMW-GS in the bread wheat cultivars Chinese Spring. Theor Appl Genet 104:422–428

Nieto-Taladriz MT, Perretant MR, Rousset M (1994) Effect of gliadin and HMW and LMW subunits of glutenin on dough properties in the F6 recombinant inbred lines from a bread wheat cross. Theor Appl Genet 88:81–88

Payne PI (1987) Genetics of wheat storage proteins and the effect of allelic variation on bread-making quality. Annu Rev Plant Physiol 38:141–153

Pogna NE, Autraun JC, Mellini F, Lafiandra D, Feillet P (1990) Chromosome 1B-encoded gliadins and glutenin subunits in drum wheat: genetics and relationship to gluten strength. J Cereal Sci 11:15–34

Rodríguez-Quijano M, Nieto-Taladriz MT, Carrillo JM (1997) Variation in B-LMW glutenin subunits in Einkorn wheats. Genet Resour Crop Evol 44:539–543

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor

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

Singh NK, Shepherd KW (1988) Linkage mapping of the genes controlling endosperm proteins in wheat. 1. Genes on the short arms of group-1 chromosomes. Theor Appl Genet 75:628–641

Van CS, Vander SJ, Sagi L, Volckaert G (1995) Locus-specific primers for LMW glutenin genes on each of the group 1 chromosomes of hexaploid wheat. Theor Appl Genet 91:313–319

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 the A and A^m genomes of wheat. Plant Cell 15(5):1186–1197

Wrigley CW (1996) Giant proteins with flour power. Nature 381:738–739

Xia GM, Xiang FN, Zhou AF, Wang H, He SX, Chen HM (2003) Asymmetric somatic hybridization between wheat (*Triticum aestivum* L.) and *Agropyron elongatum* (Host) Nevski. Theor Appl Genet 107:299–305

Xu H, Zhang XQ, Wang XP, Guo AG (2004) Molecular cloning of *Agropyron intermedium* low-molecular-weight glutenin subunit genes from a *Triticum aestivum–A. intermedium* addition line TAI-13. Acta Bot Sin 46(5):595–602

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 TJ, Quan TY, Xia GM, Chen HM (2003) Glutenin and SDS sedimentation analysis of the F_5 somatic hybrids between *Triticum aestivum* and *Agropyron elongatum*. J Shandong Univ (Nat Sci) (in Chinese with English abstract) 38(3):112–116

Zhong GZ, Mu SM, Zhang ZB (2002) Study on remote hybridisation of wheat. Chinese Science, Beijing, pp 92–97

Zhou HP, Li M, Li ZS (1995) The study of breeding blu-grain gene translocation of wheat. Acta Bot Agric Boreali-Occidentalia Sin (in Chinese with English abstract) 15:125–128