Theor Appl Genet (2004) 109: 1028–1040
DOI 10.1007/s00122-004-1711-z

ORIGINAL PAPER

J. Johal · M. C. Gianibelli · S. Rahman · M. K. Morell · K. R. Gale

Characterization of low-molecular-weight glutenin genes in *Aegilops tauschii*

Received: 20 January 2004 / Accepted: 22 April 2004 / Published online: 26 May 2004

Abstract This paper reports the characterization of the low-molecular-weight (LMW) glutenin gene family of *Aegilops tauschii* (syn. *Triticum tauschii*), the D-genome donor of hexaploid wheat. By analysis of bacterial artificial chromosome (BAC) clones positive for hybridization with an LMW glutenin probe, seven unique LMW glutenin genes were identified. These genes were sequenced, including their untranslated 3′ and 5′ flanking regions. The deduced amino acid sequences of the genes revealed four putative active genes and three pseudogenes. All these genes had a very high level of similarity to LMW glutenins characterized in hexaploid wheat. The predicted molecular weights of the mature proteins were between 32.2 kDa and 39.6 kDa, and the predicted isoelectric points of the proteins were between 7.53 and 8.06. All the deduced proteins were of the LMW-m type. The organization of the seven LMW glutenin genes appears to be interspersed over at least several hundred kilo base pairs, as indicated by the presence of only one gene or pseudogene per BAC clone. Southern blot analysis of genomic DNA of *Ae. tauschii* and the BAC clones containing the seven LMW glutenin genes indicated that the BAC clones contained all LMW glutenin-hybridizing bands present in the genome. Two-dimensional gel electrophoresis of an LMW glutenin extract from *Ae. tauschii* was conducted and showed the presence of at least 11 distinct proteins. Further analysis indicated that some of the observed proteins were modified gliadins. These results suggest that the actual number of typical LMW glutenins may in fact be much lower than previously thought, with a number of modified gliadins also being present in the polymeric fraction.

Communicated by P. Langridge

J. Johal · M. C. Gianibelli · S. Rahman · M. K. Morell · K. R. Gale (✉)
Commonwealth Scientific and Industrial Research Organisation, Plant Industry,
GPO Box 1600 Canberra, ACT, 2601, Australia
e-mail: kevin.gale@csiro.au
Tel.: +61-62-465317
Fax: +61-62-465345

Introduction

Wheat quality is largely determined by the seed storage proteins present in the endosperm of the grain (Shewry and Halford 2002). These proteins consist of two major components: the polymeric glutenins and the monomeric gliadins. The glutenins consist of high-molecular-weight (HMW) and low-molecular-weight (LMW) subunits held together by intermolecular disulfide bonds to form the glutenin polymer. The glutenin polymer interacts with gliadins largely through noncovalent interactions in the presence of water to form gluten, which confers unique visco-elastic properties to wheat flour dough and determines the quality of end-use products (Lindsay and Skerritt 1999; Shewry et al. 2001). The role of HMW glutenins in wheat quality, particularly dough strength, is well established (Payne et al. 1979; Payne 1987; Shewry et al. 1992). LMW glutenins also have significant effects on wheat quality through both additive and epistatic interactions, particularly with HMW glutenins (Damidaux et al. 1978; Gupta et al. 1989, 1994; Gupta and MacRitchie 1994; Nieto-Taladriz et al. 1994; Sontog-Strohm et al. 1996; Sissons et al. 1998; Luo et al. 2001; Eagles et al. 2002; Tranquilli et al. 2002). However, the exact role or contribution of individual LMW glutenin subunits or their alleles to wheat quality is not fully understood. This is at least in part because of the complexity of LMW glutenin gene family.

LMW glutenins are usually subdivided into B, C, and D groups according to their electrophoretic mobility in SDS-PAGE and their isoelectric points (pIs) (Jackson et al. 1983, Masci et al. 1993). The LMW glutenins in the B group (40–50 kDa) are the most basic subunits, while the C-group subunits (30–40 kDa) have a wide range of pIs. It has been found by N-terminal amino sequencing that the C group of LMW glutenins contain up to 95% modified or mutant gliadins that have become incorporated into the polymer via disulfide bonding (Tao and Kasarda 1989; Lew et al. 1992; Masci et al. 2002). The D-group subunits are modified ω-gliadins (44–74 kDa) with at least one cysteine residue in their sequences, and they are the most

acidic of all endosperm proteins (Masci et al. 1991, 1993). The monomeric gliadins (30–75 kDa) are usually classified as α/β-, γ-, and ω-type based on their electrophoretic mobility in PAGE under acidic conditions and by their N-terminal sequences (Kasarda et al. 1983).

Structurally, LMW glutenin subunits are composed of unique N- and C-terminal domains with a central repetitive domain. On the basis of N-terminal amino acid sequences, they can be divided into LMW-m, LMW-s, and LMW-i types (Kasarda et al. 1988; Tao and Kasarda 1989; Lew et al. 1992; Cloutier et al. 2001). The LMW-s type have been found to be the predominant type by N-terminal sequencing of proteins and belongs to the B group of LMW glutenins (Lew et al. 1992; Masci et al. 2002). The LMW-m type is the most predominant type found by gene sequencing and belong to both the B and C groups of LMW glutenins. Recently, LMW-i type gene sequences have been characterized, and these belong to the B group of LMW glutenins (Cloutier et al. 2001; Ikeda et al. 2002; Zhang et al. 2004).

LMW glutenins are encoded by genes present at the *Glu-3* loci on the short arms of chromosomes 1A, 1B, and 1D of hexaploid wheat (Singh and Shepherd 1988). They are present as a multigene family in close association with the *Gli-1* loci, encoding the γ- and ω-gliadins (Metakovsky et al. 1997). The α/β-gliadins are encoded by the *Gli-2* loci on the short arm of group 6 chromosomes. An estimate of 30–40 LMW glutenin genes has been made for the hexaploid wheat cvs. Cheyenne and Chinese Spring based on Southern blot analysis (Cassidy et al. 1998). By N-terminal sequencing of purified LMW glutenin fractions, 39 different sequences were identified in cv. Yecora Rojo, out of which 13 were either α- or γ-type gliadins and the remainder LMW glutenins (Lew et al. 1992). A number of partial and complete LMW glutenin gene sequences have been reported from different wheat cultivars (Bartels and Thompson 1983; Okita et al. 1985; Pitts et al. 1988; Colot et al. 1989; Cassidy and Dvorak 1991; D'Ovidio et al. 1992, 1995; Van Campenhout et al. 1995; Cassidy et al. 1998; Masci et al. 1998; Lee et al. 1999; Ciaffi et al. 1999; Benmoussa et al. 2000; Cloutier et al. 2001; Ikeda et al. 2002; Zhang et al. 2004). The complete characterization and organization of this complex multigene gene family is however still lacking.

Here, we report the characterization of the LMW glutenin gene family in *Aegilops tauschii*, the D-genome progenitor species of hexaploid wheat, using large insert clones from a bacterial artificial chromosome (BAC) library (Moullet et al. 1999). We report the isolation and sequencing of seven different LMW glutenin type sequences, of which three appeared to represent pseudogenes.

Materials and methods

BAC clones

Twenty-four LMW glutenin-positive BAC clones used in this study were isolated by standard screening procedures described in Turnbull et al. (2003) from the BAC library of the *Ae. tauschii*, accession AUS18913 (Moullet et al. 1999). This library contains 144,000 clones, with an average insert size of 120 kb. The probe used for BAC screening was specific for LMW glutenin genes and was generated by PCR amplification from genomic DNA of *Ae. tauschii* (accession CPI110799) using the following primers: forward, 5′-AGATGCATCCCTGGTTTGGAG-3′ and reverse, 5′-AATGGAAGTCATCACCTCAAG-3′.

These PCR primers were designed in the N-terminal domain and near the end of the C-terminal domain of an LMW glutenin gene (*LMWG-1D1*) from the D genome of *Triticum aestivum* (Colot et al. 1989) and amplify a product of approximately 750 bp.

BAC DNA extraction

Rapid preparation of BAC DNA was done using the alkaline lysis method (Sambrook et al. 1989) from 5 ml LB cultures containing 15 μg/ml tetracycline. For BAC sequencing, high-quality DNA free of contaminating genomic DNA and nicked or damaged BAC DNA was extracted using the Qiagen Large-Construct kit from 500 ml LB cultures containing 15 μg/ml tetracycline.

Restriction enzyme analysis of BAC clones

BAC DNA was digested with *Hin*dIII or *Eco*R1 and analyzed on 0.8% agarose gels, followed by Southern blot hybridization. Southern blot analyses were carried out following standard procedures (Sambrook et al. 1989). The probe was generated by PCR amplification of genomic DNA of *Ae. tauschii*, accession AUS 24230 using the following primers: lmwF5, 5′-CTTGCTAGGTCGCAAATGTG-3′ and lmwR, 5′-TGCCAACGCCGAATGGCAC-3′.

These primers were designed in the C-terminal domain of the LMW gene (*LMWG-1D1*) from the D genome of hexaploid wheat (Colot et al. 1989) and amplify a 403 bp fragment. The PCR fragment was purified using a QIAquick PCR purification kit (Qiagen) and horseradish peroxide (HRP) labeled, using the North^2 South Direct HRP labeling and detection kit (Pierce). For Southern blot analysis of *Ae. tauschii* genomic DNA, the probe was labeled with ^{32}P, using the Megaprime DNA labeling system (Amersham).

Direct sequencing of BAC clones

High-sensitivity cycle sequencing reactions were conducted using 0.5–1.0 μg high-quality BAC DNA and 5–10 pmol of primer in 40 μl reactions with Big Dye chemistry and run on an ABI 377 DNA sequencer (ABI Systems, Columbia, Mo., USA).

Primers used for DNA sequencing

The initial four primers used for DNA sequencing were based on the sequence of an LMW glutenin gene from the D genome of hexaploid wheat (Colot et al. 1989). Further primers were designed based on the sequence information obtained to complete the gene sequences, along with the flanking 3′ and 5′ untranslated regions. The list of primers used and their sequences are shown in Table 1.

Nucleotide and amino acid sequence analysis

DNA sequence alignment, translation, identification of open reading frames, and analysis of protein sequence were conducted using the Australian National Genomic Information Service computing facility at the University of Sydney (http://www.angis.org.au). Expressed sequence tag (EST) database searches were conducted using The Institute of Genome Research (TIGR) site (http://www.tigr.org/tdp/tgi/). Global nucleic acid and protein comparisons were done by BLAST (http://www.ncbi.nlm.nih/gov).

Table 1 List of primers used for sequencing low-molecular-weight (LMW) glutenin genes in *Aegilops tauschii*

Name	Sequence (5′- 3′)
lmw.colot (forward)	AGA TGC ATC CCT GGT TTG GAG
lmw.colot (reverse)	AAT GGA AGT CAT CAC CTC AAG
ZLLMW5′	ATG AAG ACC TTC CTC GTC TTT GC
ZLLMW3′	CCT TAT CAG TAG GCA CCA ACT CC
J.LMF1	GCT AAA CCC ATG CAA GGT ATT CC
J. LMF2	CGC TAT GAG GCA ATC CGT GC
J.LMF3	TTG CAG CCA CAC CAG ATA GC
J.LMF4	CTT GAG GTG ATG ACT TCC ATT GC
J.LMF5	AGT TGG TGC CTA CTG ATA AG
J.LMF6	ACA CTA GTT AAC ACC AAT CCA CCA
J.LMR1	TGC CAT GGT CTC TCC AAA CCA G
J.LMR1b	ATG GTT TCT CCA AGC CAG GGA
J.LMR2	GAA TAC CTT GCA TGG GTT TAG C
J.LMR3	GCA AAG ACG AGG GTC TTC AT
J.LMR4	AAC ATT TGC GAC CTA GCA AGA CG
J.LMR5	GAT CTT TTC TTA TCA GTA GGC ACC
J.LMR6	GGG TGG GTC ACA CAT GAC ATT G

Sample preparation and two-dimensional gel electrophoresis

The total or polymeric protein extracts from *Ae. tauschii*, accession no. AUS18913, were used to rehydrate 11 cm immobilized pH gradient (IPG) strips (pH range of 3–10, 4–7, or 6–11). For total protein preparation, 10 mg wholemeal flour was suspended in 800 μl rehydration buffer (8 M urea, 2% CHAPS, 0.01% bromophenol blue, and 50 mM DTE). For polymeric protein extraction (Gupta and MacRitchie 1991), the monomeric gliadins were first removed by dimethyl sulphoxide extraction at room temperature for 1 h. The polymeric proteins were then extracted, using 70% ethanol containing 3% (v/v) β-mercaptoethanol at 65°C for 1 h. The extracted polymeric proteins were precipitated by addition of acetone to 80% (v/v). The resulting pellet was resuspended in rehydration buffer. Isoelectric focusing was performed, using a Multiphore II unit (Amersham Pharmacia Biotech, Sweden) using the step-wise protocol for 40–50 kVh. For the second dimension, the strips were reduced with DTE, alkylated with iodoacetamide, and run on 10% SDS-PAGE gels. Proteins were visualized by staining with Coomassie Brilliant Blue (G-250) and destained with distilled water (Neuhoff et al. 1988). The stained gels were sent to a commercial service provider for N-terminal sequencing.

Results

Southern blot analysis of BAC clones and *Ae. tauschii* genomic DNA

Twenty-four BAC clones positive by hybridization with the LMW glutenin probe were subjected to fingerprint analysis based on the pattern of restriction digests obtained with *Hin*dIII or *Eco*R1. Gels were blotted and hybridizations performed, using a 403 bp PCR probe derived from the 3′ end of an LMW glutenin gene from the *Glu-D3* locus (Colot et al. 1989). All clones contained one or two LMW-hybridizing bands and on the basis of the pattern of the bands observed on gels and Southern blots, the 24 BAC clones were divided into seven groups. One representative clone (S1–S4, S6, S7, S13) from each of the seven groups was selected for further analysis. Southern blot analysis of the *Hin*dIII-digested genomic DNA of *Ae. tauschii* in parallel with the seven BAC clones is shown in Fig. 1. All the LMW-hybridizing bands present in the genomic DNA were represented in the collection of BAC clones, except for one HMW band, which is probably due to uncut genomic DNA. Similar results were obtained using *Eco*R1 as the restriction enzyme for Southern blot analysis.

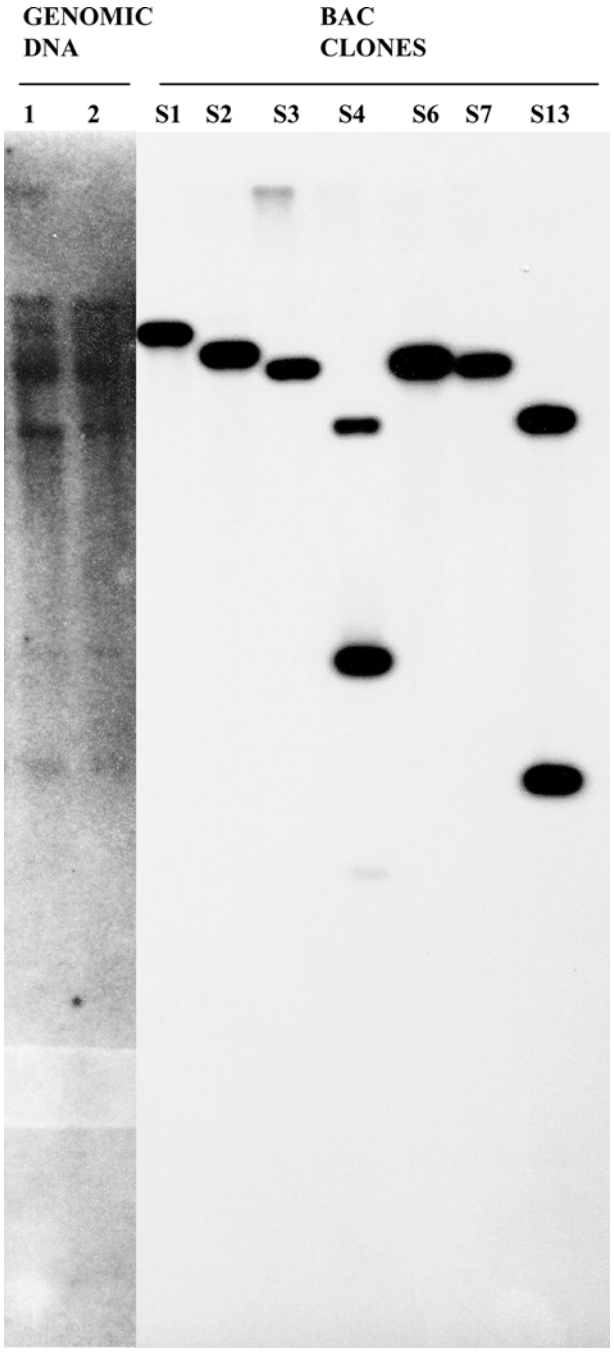

Fig. 1 Southern blot analysis of low-molecular-weight (LMW) glutenin genes in genomic DNA of *Aegilops tauschii* and seven LMW glutenin positive bacterial artificial chromosome (*BAC*) clones, all from the same accession of *Ae. tauschii*. The *first two lanes* represent two different preparations of genomic DNA. All DNA was digested with *Hin*dIII and analyzed on 0.8% agarose gel. The 403-bp probe used was based on the sequence of a D-genome LMW glutenin gene (Colot et al. 1989)

LMW glutenin gene sequence analysis

High-quality DNA extracted from the seven representative BAC clones was used for direct sequencing, using primers specific for the LMW genes based on the available LMW gene sequence from the D genome (Colot et al. 1989). These primers were designed in the conserved region of the genes. Using these primers, seven LMW glutenin genes were identified (one per BAC) and designated *LMW.S1–LMW.S4*, *LMW.S6*, *LMW.S7*, and *LMW.S13*, respectively. Based on the initial sequence information obtained for the LMW glutenin genes, internal primers were designed to complete the sequencing of the seven LMW glutenin genes. The GenBank accession numbers for these sequences are as follows: *LMW.S1*, AY585350; *LMW.S2*, AY585349; *LMW.S3*, AY585351; *LMW.S4* 5′ end, AY585353; *LMW.S4* 3′ end, AY585352; *LMW.S6*, AY585354; *LMW.S7*, AY585355; and *LMW.S13*, AY585356.

The pairwise percentage sequence similarity scores between the genes are shown in Table 2. The similarity scores were greater than 85% between the paralogous genes. Based on sequence similarity scores, *LMW.S2* and *LMW.S6* genes seemed to be most closely related and were also most distinct from the other LMW glutenin genes.

DNA sequence alignment of the seven genes and their untranslated or noncoding 5′ and 3′ flanking regions is shown in Fig. 2. The total length of the genes sequenced with their flanking regions was variable: 1,781 bp for *LMW.S1*, 1,883 bp for *LMW.S2*, 1,819 bp for *LMW.S3*, 1,189 bp at the 5′ end and 1,118 bp at the 3′ end for *LMW.S4* (this gene was interrupted by an insertion, see below), 2,115 bp for *LMW.S6*, 1,562 bp for *LMW.S7*, and 1,785 bp for *LMW.S13*. In the 5′ noncoding regions of the genes, all the previously characterized promoter sequence motifs were identified. The TATA box with sequence motif TATAAATA is present about 70 bp upstream of the ATG start codon in all the genes except for *LMW.S4*. About 300 bp upstream of the translation start site is the "endosperm box" or "−300 element," consisting of two motifs: the endosperm or E motif at the 5′ end and the GCN4-like or N motif at the 3′ end of the box. The E motif (TGTAAAGT) is absolutely conserved and is present in all the genes, except for *LMW.S4*, for which no sequence information was available in this region. In contrast, the N motif is variable at three positions with the consensus sequence being (T/C)(G/A) A(G/T)TCAT. In the coding regions of the genes, many single nucleotide polymorphisms (SNPs) and insertions/deletions (INDELs) are evident between the paralagous genes. In the 3′ noncoding regions of the genes, three polyadenylation signals (AATAAA) were identified. In *LMW.S3*, the first polyadenylation signal was present as AATATA.

BLAST (National Center for Biotechnology Information) and EST database searches (TIGR) were conducted to establish the relationships of these genes to the LMW glutenin genes present in the databases. The best matches obtained for each of the genes are listed in Table 3. *LMW.S4* seemed to be interrupted by a sequence that is 98% identical to the 3′ untranslated flanking region of the γ-gliadins in hexaploid wheat (Anderson et al. 2001). The sequence also had 95% identity to two different noncoding regions on a BAC clone containing genes related to disease resistance (Brooks et al. 2002). Based on these comparisons the insertion was calculated to be 2,170–

Table 2 The pairwise percentage identity of the coding region of the seven LMW glutenin genes characterized from *Ae. tauschii*

	LMW.S1	*LMW.S2*	*LMW.S3*	*LMW.S4*	*LMW.S6*	*LMW.S7*	*LMW.S13*
LMW.S1	100						
LMW.S2	86.60	100					
LMW.S3	88.59	86.64	100				
LMW.S4	87.67	86.73	94.74	100			
LMW.S6	86.16	98.69	86.76	87.07	100		
LMW.S7	86.90	85.74	88.14	86.44	84.66	100	
LMW.S13	91.23	85.90	88.93	89.43	86.34	86.69	100

2,172 bp in length and is 750 bp from the 5′ end of the gene (Fig. 2). The *LMW.S4* gene sequence without the insertion was used for subsequent BLAST searches. Overall, these *Ae. tauschii* LMW glutenin genes have a high level of sequence identity (up to 99%) to LMW glutenin genes present in different cultivars of hexaploid wheat. In particular, these genes had identities ranging from 94% to 99% to the genes characterized in *T. aestivum* cv. Norin 61, including those present at the *Glu-D3* locus (Ikeda et al. 2002). The percentage identities and the accession numbers of these genes are listed in Table 3. The genes *LMW.S1*, *LMW.S3*, and *LMW.S13* had up to 99% identity to LMW glutenin genes characterized in *T. aestivum* cv. Cheyenne (Okita et al. 1985; Anderson et al. 1997). The genes *LMW.S6*, *LMW.S7*, and *LMW.S13* had >98% identity to genes in *T. aestivum* cv. Chinese Spring (Colot et al. 1989; Van Campenhout et al. 1995; Xu et al. 2003, unpublished). *LMW.S13* also had 99% identity with an LMW glutenin gene present in *T. aestivum* cv. Neepawa (Chardot et al. 2002). The genes *LMW.S1*, *LMW.S2*, *LMW.S4*, and *LMW.S6* showed >90% identity to LMW glutenin genes from tetraploid wheat, particularly *Triticum turgidum* subspecies durum cv. Langdon (D'Ovidio et al. 1997).

Deduced amino acid sequence analysis

Deduced amino acid analysis of the seven genes showed that four genes (*LMW.S1*, *LMW.S6*, *LMW.S7*, and *LMW.S13*) have a single open reading frame (ORF), and the other three genes (*LMW.S2*, *LMW.S3*, and *LMW.S4*) have internal stop codons, indicating them to be putative pseudogenes.

Sequence alignment of the predicted amino acid sequences of the seven genes/pseudogenes was conducted (Fig. 3). All genes encode a predicted, highly conserved signal peptide of 20 amino acids and a short N-terminal region of 13 amino acids. This is followed by a repetitive domain rich in proline and glutamine residues, with PPFSQ being the most common repeat. The size of this domain is variable between the genes, depending upon the number of the repeats. A polyglutamine stretch of 20 residues was observed in this domain for *LMW.S2*. Two hydrophobic repeats (PIIIIL and PVIIIL) similar to those identified in LMW-GS group 5 by Ikeda et al. (2002) were also observed in the *LMW.S7*-predicted protein. The repetitive domain is followed by the C-terminal domain, comprising more than half of the protein. This region is glutamine rich with interspersed proline residues and is less variable in size than the repetitive domain.

The number and position of the stop codons were found to be variable in the pseudogenes. *LMW.S2* had one stop codon in the repetitive domain. *LMW.S4* had two stop codons, one in the repetitive domain and the other in the C-terminal domain. *LMW.S3* had three stop codons in the C-terminal domain. Double stop codons were present at the end of the coding regions in all the genes, including the pseudogenes.

All the deduced proteins belong to the LMW m-type (methionine at position 1 of the putative mature peptide). Based on their N- and C-terminal sequences, they could be classified as group 1 (*LMW.S6* and *LMW.S2*), group 5 (*LMW.S7*), group 7 (*LMW.S1*), group 8 (*LMW.S3*), or group 10 (*LMW.S13* and *LMW.S4*), according to the nomenclature of Ikeda et al. (2002). In position 2 in the N-terminal sequence of *LMW.S3*, glutamine is present instead of glutamate, the result of single point mutation.

All the deduced LMW glutenin proteins (including those of the three pseudogenes) showed the presence of eight conserved cysteine residues. An extra cysteine residue is present in the deduced protein sequence of the pseudogene *LMW.S3*. The positions of the first and the seventh cysteine residues were found to be variable as observed previously for other LMW glutenins (Ikeda et al. 2002). In *LMW.S1*, *LMW.S3*, *LMW.S4* and *LMW.S13*, the first cysteine residue is present at position 5 in the N-terminal domain. In LMW.S7, the first cysteine residue is present at position 26 (central repetitive domain). In *LMW.S2* and *LMW.S6*, the first cysteine residue is present at position 45 (central repetitive domain). The positions of the other six cysteine residues are conserved, and all are located in the C-terminal domain of the proteins. In *LMW.S13*, the seventh cysteine residue is present nine residues closer to the N-terminus relative to the other deduced proteins. Based on the position of the cysteine residues, all the genes could be classified as type I (*LMW.S2* and *LMW.S6*), type III (*LMW.S7*), type IV (*LMW.S1*, *LMW.S4*) and type V (*LMW.S13*), according to Ikeda et al (2002).

The polypeptide lengths, predicted molecular weights, and pIs of the four deduced proteins from the genes that contain a single ORF are as follows: *LMW.S1*, 284 residues, 32,215 Da, pI 7.53; *LMW.S6*, 345 residues, 39,660 Da, pI 8.06; *LMW.S7*, 346 residues, 39,596 Da, pI 7.84; and *LMW.S13*, 285 residues, 32,509 Da, pI 7.82.

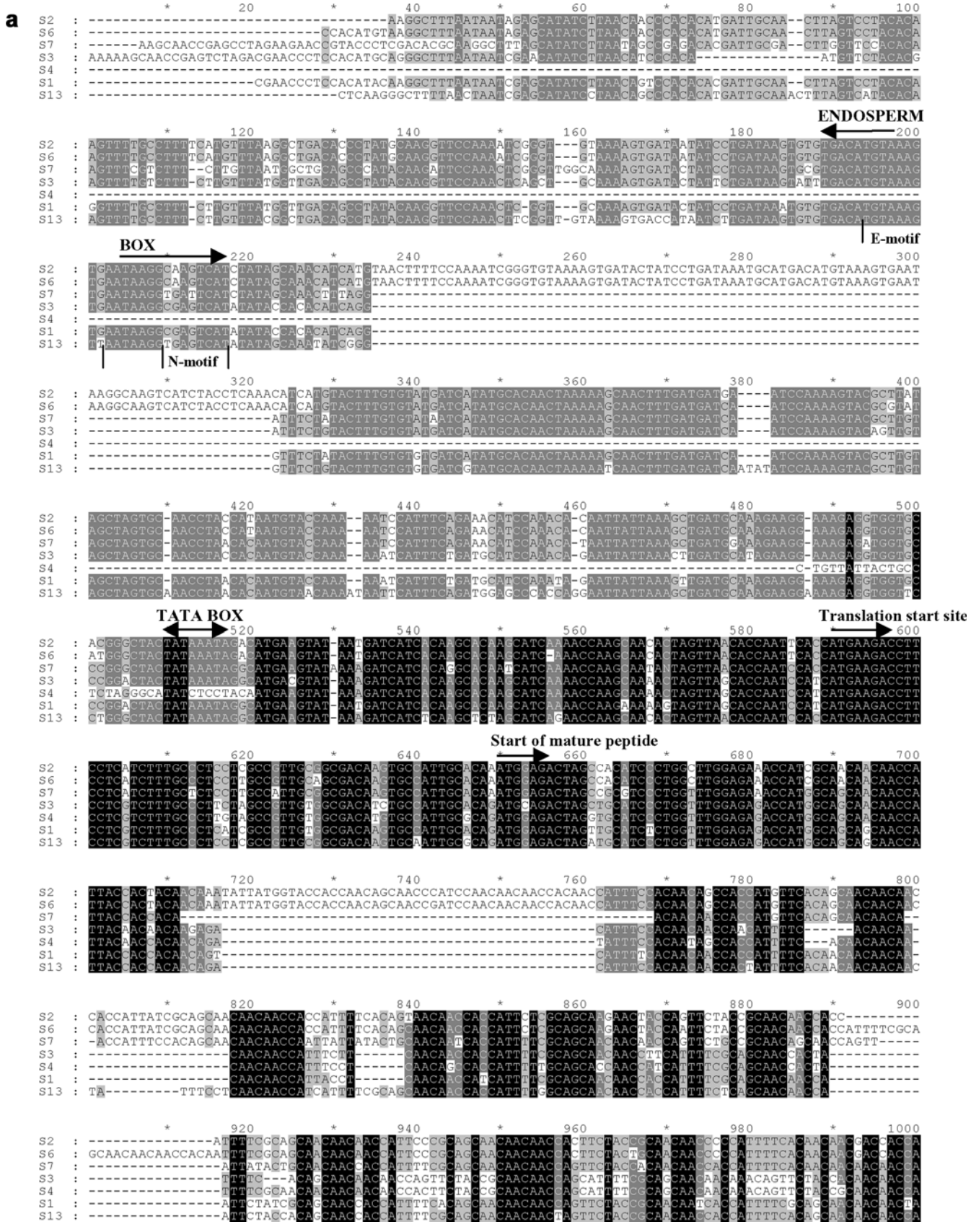

Fig. 2a, b Annotated DNA sequence alignment of seven *Ae. tauschii* LMW glutenin genes/pseudogenes and their untranslated 5′ and 3′ flanking regions

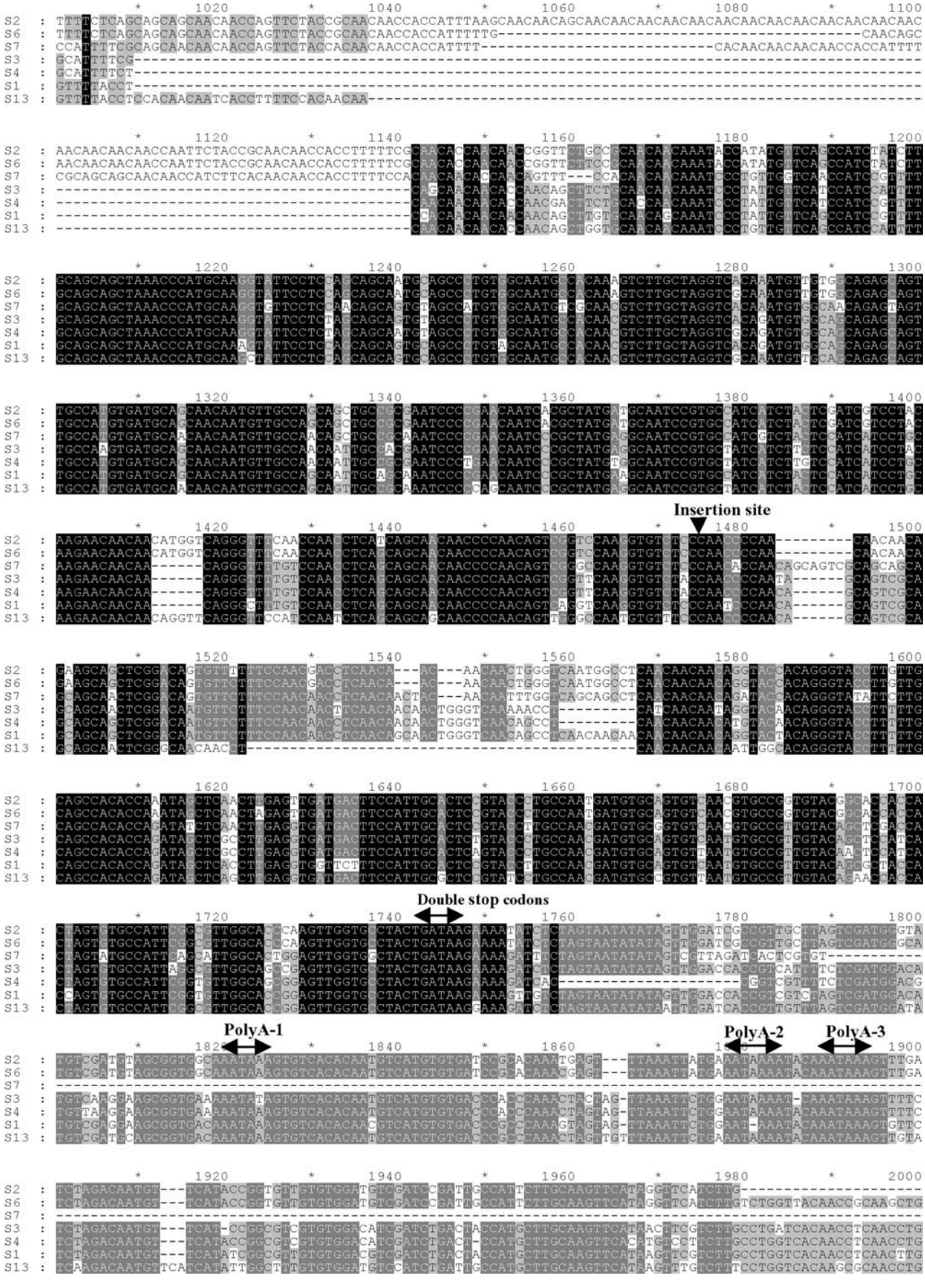

Fig. 2a, b (continued)

Table 3 Database similarity scores of the seven LMW glutenin genes characterized from *Ae. tauschii*

LMW glutenin	Accession numbers[a]	Percentage identity	Cultivar	Reference
LMW.S1	M11077	98	*Triticum aestivum* cv. Cheyenne	Okita et al.1985
	ABO62872	97	*T. aestivum* cv Norin 61	Ikeda et al. 2002
	U86028	97	*T. aestivum* cv. Cheyenne	Anderson et al. 1997
	AJ293099	93 (partial)	*Triticum turgidum* sub species durum cv. Langdon	D'Ovidio et al. 2000 (unpublished)
LMW.S2	AB062851	97–99	*T. aestivum* cv. Norin 61	Ikeda et al. 2002
	AF419256	99	*T. aestivum* cv. PH 82–2-2	Li et al. 2001 (unpublished)
	X84961	98	*T. aestivum* cv. Chinese spring	Van Campenhout et al. 1995
	AB062852	95–96	*T. aestivum* cv. Norin 61	Ikeda et al. 2002
	Y14104	95–96	*T. turgidum* subspecies durum cv. Langdon	D'Ovidio et al. 1997
LMW.S3	AB062873	97	*T. aestivum* cv. Norin 61	Ikeda et al. 2002
	U86026	97	*T. aestivum* cv. Cheyenne	Anderson et al. 1997
	AB062874	97	*T. aestivum* cv. Norin 61	Ikeda et al. 2002
	AJ519836	92 (partial)	*T. aestivum* cv. Neepawa	Chardot et al. 2002
	AB007763	92 (partial)	*T. aestivum* cv. hard red spring ICW	Maruyama et al. 1998
LMW.S4 insertion	AF446141	98	*Ae. tauschii* AUS18913	Brooks et al. 2002
	AF234649	98	*T. aestivum* cv. Cheyenne	Anderson et al. 2001
	AF446141	96	*Ae. tauschii* AUS18913	Brooks et al. 2002
LMW.S4 (without insertion)	AB062873	94	*T. aestivum* cv. Norin 61	Ikeda et al. 2002
	AB062874	94	*T. aestivum* Norin 61	Ikeda et al. 2002
	U86026	94	*T. aestivum* cv. Cheyenne	Anderson et al. 1997
	AB007763	93	*T. aestivum* cv. hard red spring ICW	Maruyama et al. 1998
	X51759	92 (partial)	*T. turgidum* subspecies durum cv. Langdon	Cassidy et al. 1998
	AJ519838	92 (partial)	*T. aestivum* cv. Neepawa	Chardot et al. 2002
	AB062871	92	*T. aestivum* cv. Norin 61	Ikeda et al. 2002
LMW.S6	AB062851	99	*T. aestivum* cv. Norin 61	Ikeda et al. 2002
	X84961	99	*T. aestivum* cv. Chinese spring	Van Campenhout et al. 1995
	AF419256	99 (partial)	*T. aestivum* cv. PH 82-2-2	Li et al. 2001 (unpublished)
	X84960	96	*T. aestivum* cv. Chinese spring	Van Campenhout et al. 1995
	Y14104	95 (partial)	*T. turgidum* subspecies durum cv. Langdon	D'Ovidio et al. 1997
	AB062852	96	*T. aestivum* cv. Norin 61	Ikeda et al. 2002
LMW.S7	AB062867	96–100 (partial)	*T. aestivum* cv. Norin 61	Ikeda et al. 2002
	AB062865	99–100 (partial)	*T. aestivum* cv. Norin 61	Ikeda et al. 2002
	AY214450	98 (partial)	*T. aestivum* cv. Chinese spring	Xu et al. 2003 (unpublished)
	AB062866	98 (partial)	*T. aestivum* cv. Norin 61	Ikeda et al. 2002
	AY214457	99 (partial)	*T. aestivum*/Thinopyrum intermedium alien addition line cv. TAI-13	Xu et al. 2003 (unpublished)
LMW.S13	AB062875	99	*T. aestivum* cv. Norin 61	Ikeda et al. 2002
	X13306	99	*T. aestivum* cv. Chinese Spring	Colot et al. 1989
	AJ519835	99	*T. aestivum* cv. Neepawa	Chardot et al. 2002
	U86027	99	*T. aestivum* cv. Cheyenne	Anderson et al. 1997
	U86028	98	*T. aestivum* cv. Cheyenne	Anderson et al. 1997

[a]GenBank

Analysis of LMW glutenin proteins expressed in *Ae. tauschii* by two-dimensional gel electrophoresis

An initial analysis of total and polymeric proteins extracted from wholemeal flour of *Ae. tauschii*, accession AUS18913 (the same accession as used for BAC library construction) was done using 11-cm, wide-range, linear pH 3–10 IPG strips. For further resolution of spots, 11-cm pH 4–7 and pH 6–11 IPG strips were used. At least 11 distinct spots were observed for LMW glutenins in the extracted polymeric proteins, and these clustered into two regions: one in the range of pH 6.5–7.5 and the other in the range of pH 9–10 (Fig. 4). One spot, marked as 6, was observed between these two clusters. Of particular interest

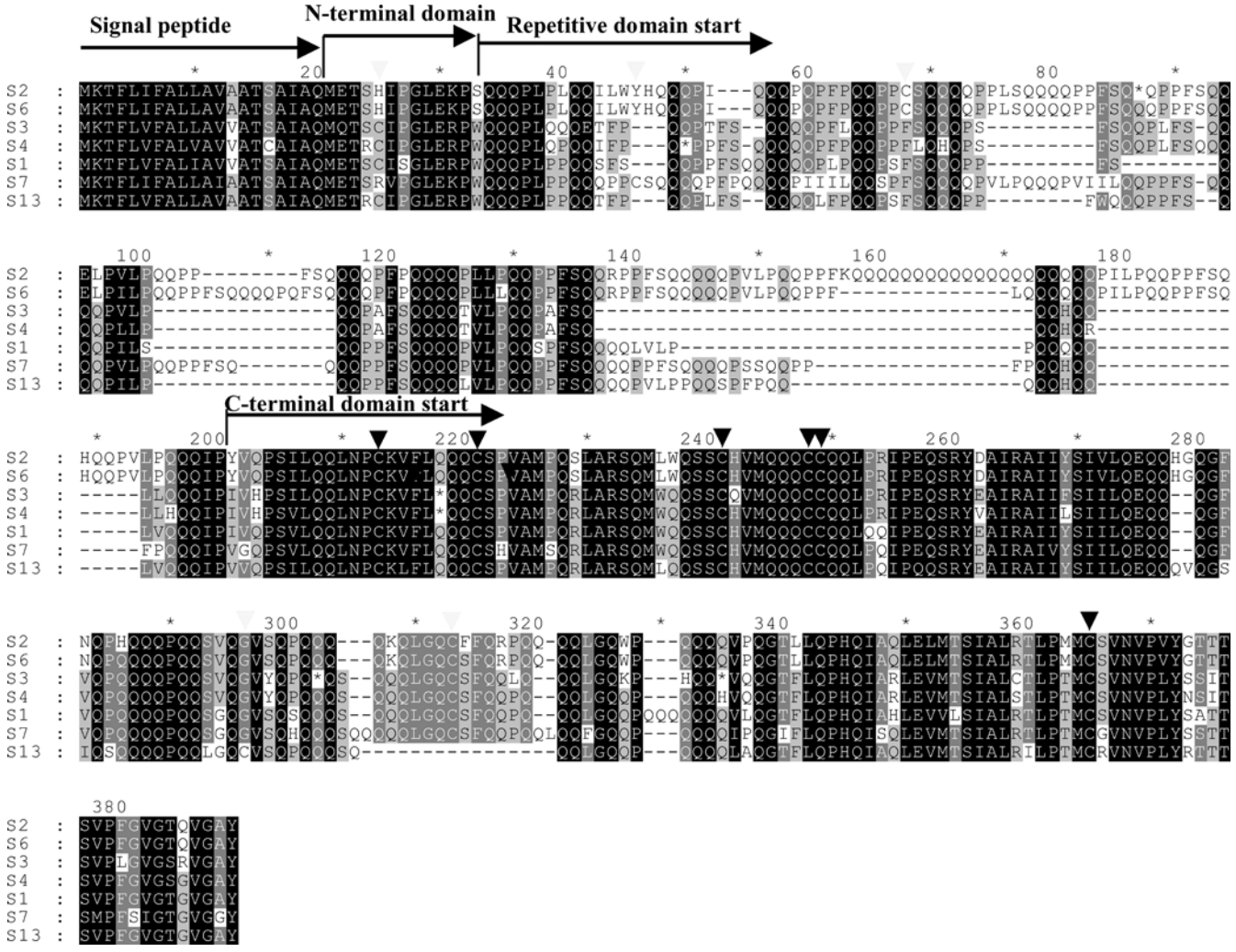

Fig. 3 Sequence alignment of deduced precursor proteins of seven *Ae. tauschii* LMW glutenin genes and pseudogenes. The positions of cysteine residues at variable locations are marked as *grey arrows*, while the positions of the conserved cysteine residues are marked as *black arrows*

are the four spots in the pH range 9–10, with two of these (A and B) having mobilities in the range of 42–44 kDa (B subunits) and the other two (C and D), with mobilities in the range of 35–38 kDa (C subunits). Analysis of some spots by N-terminal sequencing has shown that six spots (spots 1–6, Fig. 4) have typical α/β-gliadin N-terminal sequences, and one spot (spot D, Fig. 4) was found to have a typical γ-gliadin N-terminal sequence (results not shown).

Discussion

Ae. tauschii is the D-genome (diploid) progenitor of hexaploid wheat, *T. aestivum* (Kihara 1944; McFadden and Sears 1946; Friebe and Gill 1996). The colinearity between the genome of *Ae. tauschii* and the D genome of *T. aestivum* has also been established (Gill et al. 1991; Dvorak et al. 1998; von Buren 2001). This study reports the characterization of the LMW glutenin gene family in *Ae. tauschii* and the comparison of these genes with genes characterized previously in *T. aestivum*.

By the analysis of BAC clones positive for LMW glutenin-hybridizing sequences, seven LMW glutenin genes were identified (one per BAC clone) and sequenced. These appear to represent all of the genes detected by Southern blotting in genomic DNA of the same *Ae. tauschii* accession as used for BAC library construction. The presence of two LMW glutenin-hybridizing bands observed for BAC clones S4 and S13 by Southern blot analysis using *Hin*dIII are due to the presence of an internal *Hin*dIII restriction site in these LMW glutenin genes.

The similarity between the genes characterized in this study was greater than 85%. The level of identity of these genes to those in *T. aestivum* is very high (up to 100%, for instance, over a 328-bp region in *LMW.S7*), confirming the close relationship of the *Ae. tauschii* accession used (AUS18913) to the D-genome of hexaploid wheat. DNA-sequence alignment of these genes showed the presence of INDELs and SNPs that may now be targeted for the

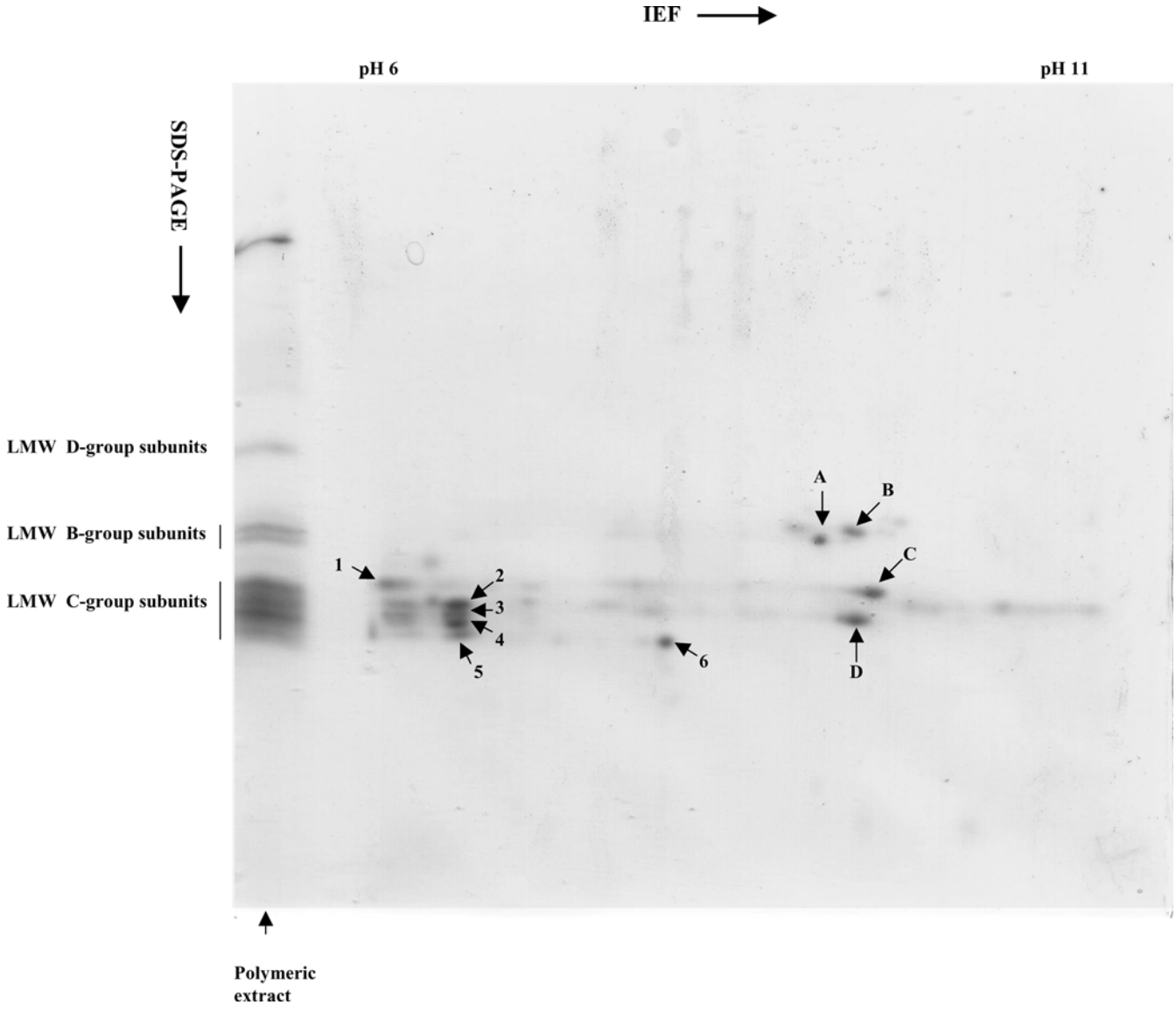

Fig. 4 Two-dimensional gel analysis [isoelectric focusing (IEF) × SDS-PAGE] of polymeric gluten proteins extracted from *Ae. tauschi*. The proteins visualized in one dimension are represented in *lane one*

development of gene-specific markers. These gene-specific markers will be useful for identification of the corresponding genes in hexaploid wheat and the analysis of allelic variation at the *Glu-D3* locus.

The LMW glutenin genes described here, like all other prolamin genes analyzed to date, do not contain introns. *LMW.S4*, however, contains a 2,170–2,172-bp insertion with 96–98% sequence identity to the intergenic region sequenced in the *Gli-1* locus in both *Ae. tauschii* and *T. aestivum* (Brooks et al. 2002; Anderson et al. 2001). Further characterization of this sequence is required to establish the mechanism of the observed insertion event.

Three of the seven *Ae. tauschii* LMW glutenin genes identified contain internal stop codons, indicating them to be putative pseudogenes. Interestingly, these pseudogenes also have a high level of sequence identity (up to 99%) to genes present in *T. aestivum*, including matches to cDNA/EST sequences. This indicates that functional orthologous genes corresponding to these pseudogenes may be present in hexaploid wheat. The presence of LMW glutenin pseudogenes has been observed previously (Benmoussa et al. 2000; Ikeda et al. 2002). Some pseudogenes have also been reported in other gene families, including the γ-gliadin (Rafalski 1986), HMW glutenin (Forde et al. 1985) and α-gliadin (Anderson et al. 1997) gene families. It is interesting to note that the 5′ noncoding regions of the pseudogenes and putative active genes are highly conserved, with the presence of TATA and endosperm boxes. The N-motif in the endosperm box was, however, variable between the genes. The 5′ noncoding region of the pseudogene *LMW.S4*, was different to the other genes characterized in this study, including the TATA box. The sequence 5′ of the TATA box was not obtained for *LMW.S4*.

The *Ae. tauschii* LMW glutenin genes characterized in this study encode a group of predicted proteins with a

typical LMW glutenin structure: a conserved signal peptide of 20 amino acids, a short N-terminal "unique" domain of 13 amino acids, a central repetitive domain comprising a variable number of amino acids, and a conserved C-terminal domain, which comprises more than half of the protein. All the proteins were LMW-m type, having a methionine residue at the N-terminus. Sequence comparison of the predicted proteins from the genes and pseudogenes showed the presence of eight conserved cysteine residues, which is characteristic of LMW glutenins (Ikeda et al. 2002). The number and position of cysteine residues in glutenins is critical in determining the structure of the proteins, their intermolecular disulfide bonding, and ultimately, the size and functionality of the glutenin polymer (Shewry and Tatham 1997). It has been proposed that the first and seventh cysteine residues present in LMW glutenins are likely to participate in intermolecular disulfide bond formation, while the remaining six cysteine residues are involved in intramolecular disulphide bond formation (Lew et al. 1992; Shewry and Tatham 1997; Müller et al. 1998; D'Ovidio et al. 1999; Masci et al. 1998).

Regarding organization of the *Glu-D3* locus, it appears that in *Ae. tauschii* the LMW glutenin genes are interspersed over a distance of at least several hundred kilobase pairs, as all seven genes are located on different BAC clones. Similar results have been reported recently in *Triticum monococcum* (Wicker et al. 2003), where two LMW glutenin genes were separated by more than 150 kb, with a third gene less than 100 kb from these genes.

On the distal end of the chromosome 1DS of *Ae. tauschii*, 17 restriction fragment length polymorphism (RFLP) markers have been mapped (Spielmeyer et al. 2000). An analysis of 106 kb of contiguous DNA sequence in this region has revealed a high gene density (one gene per 8.9 kb on average, Brooks et al. 2002). The genes in cereal genomes are mostly found in gene-rich regions or "gene islands" (Gill et al. 1993; 1996; DeScenzo and Wise 1996; Büschges et al. 1997; Keller and Feuillet 2000). Two major gene-rich regions have been identified on the short arm of homeologous group 1 chromosomes (Gill et al. 1996), and by comparative mapping, 75 genes and 93 RFLP markers have been identified in this region (Sandhu et al. 2001). Complete DNA sequence analysis of the *Ae. tauschii* LMW glutenin locus is required to determine the actual gene density in this region.

Southern blot analysis of genomic DNA of *Ae. tauschii* in comparison with the seven LMW glutenin positive BAC clones demonstrated that these appear to contain all hybridizing bands observed in the corresponding genomic DNA. To establish the corelation of the complexity of the gene family characterized in this study with the expressed LMW glutenin subunits present in *Ae. tauschii*, an analysis of polymeric proteins extracted from wholemeal flour of the same accession of *Ae. tauschii* as used for BAC library construction was performed. Two genes (*LMW.S6* and *LMW.S7*), encode predicted LMW glutenins with molecular weights of 39,660 and 39,595 Da, and are thus classified as B-group subunits, and two (*LMW.S1* and *LMW.S13*) with molecular weights of 32,215 and 32,509 Da, corresponding to C-group subunits. In one-dimensional SDS-PAGE of the polymeric protein extract, two bands of B-group subunits and at least six bands of C-group subunits were observed. The LMW glutenins in the B group are normally considered as "typical" LMW glutenins. The calculated molecular weight and pI of the deduced proteins was lower than that of the spots A–D that were observed by electrophoresis. Higher estimated molecular weight by SDS-PAGE analysis and higher observed pIs compared with the calculated values from the amino acid sequence have been observed previously for LMW glutenins (Zhang et al. 2004; M.C. Gianibelli, personal communication). By N-terminal amino acid analyses, most of the C-group LMW glutenins have been found to be modified or mutant gliadins (Masci et al. 2002). It has been proposed that these gliadins become incorporated into the polymeric glutenin fraction due to an altered number of cysteine residues and the subsequent formation of intermolecular disulfide bonds (Masci et al. 2002). The number of protein spots observed by two-dimensional (2-D) gel electrophoresis analysis of polymeric proteins extracted from *Ae. tauschii* in the current study was approximately 11. Six of these spots were subsequently found by N-terminal sequence analysis to be α/β-gliadins and one spot to be a γ-gliadin. The presence of these modified gliadins in the polymeric protein extract of *Ae. tauschii* is consistent with the unexpected complexity of the observed protein species in the 2-D gels relative to gene family complexity. It is not known at this time if any of the observed complexity of protein species is due to posttranslational modification of the gluten proteins.

In summary, this study provides a model for understanding the sequence variation, complexity, and organization of the LMW glutenin gene family in wheat and has provided information that is currently being used to develop markers for specific *Glu-D3* alleles in hexaploid wheat.

References

Anderson OD, Litts JC, Greene FC (1997) The alpha-gliadin gene family. I. Characterization of ten new alpha-gliadin genomic clones, evidence for limited sequences conservation of flanking DNA, and Southern analysis of the gene family. Theor Appl Genet 95:50–58

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

Bartels D, Thompson RD (1983) The characterization of cDNA clones coding for wheat storage proteins. Nucleic Acids Res 11:2961–2977

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

Brooks SA, Huang L, Gill BS, Fellers JP (2002) Analysis of 106 kb of contiguous DNA sequence from the D genome of wheat reveals high gene density and a complex arrangement of genes related to disease resistance. Genome 45:963–972

Buren M von (2001) Polymorphisms in two homologous γ-gliadin genes and the evolution of cultivated wheat. Genet Res Crop Evol 48:205–220

Büschges R, Hollritcher K, Panstruga R, Simons G, Wolters M, Frijters A, van Daelen R, van der Lee T, Diergaarde P, Groenendijk J, Topsch S, Vos P, Salamini F, Schulze-Lefert P (1997) The barley *Mlo* gene: a novel control element of plant pathogen resistance. Cell 88:695–705

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

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

Ciaffi M, Lee YK, Tamas L, Gupta R, Sherritt J, Appels R (1999) The low-molecular-weight glutenin subunit proteins of primitive wheats. III. The genes from D-genome species. Theor Appl Genet 98:135–148

Cloutier S, Rampitsch C, Penner GA, Lukow OM (2001) Cloning and expression of an 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

Damidaux R, Grignac P, Feillet P (1978) Mise en évidence de relations applicables en sélection entre l'électrophorègramme des gliadines et les propriétés viscoélastiques du gluten de *Triticum durum* Desf. C R Acad Sci (Paris) 287:701–704

DeScenzo RA, Wise RP (1996) Variation in the ratio of physical to genetic distance in intervals adjacent to the *Mla* locus on barley chromosome 1H. Mol Gen Genet 251:472–482

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, Kasarda DD (1995) Nucleotide sequence of a γ-type glutenin gene from a durum wheat: correlation with a γ-type glutenin subunit from the same biotype. Cereal Chem 72:443–449

D'Ovidio R, Simeone M, Masci S, Porceddu E (1997) Molecular characterization of an 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, Cardelli LE, Porceddu E (1999) Sequence similarity between allelic *Glu-B3* genes related to quality properties of durum wheat. Theor Appl Genet 93:455–461

Dvorak J, Luo M-C, Yang Z-L, Zhang H-B (1998) The structure of the *Aegilops tauschii* genepool and the evolution of hexaploid wheat. Theor Appl Genet 97:657–670

Eagles HA, Hollamby GJ, Gororo NN, Eastwood RF (2002) Estimation and utilization of glutenin gene effects from the analysis of the unbalanced data from the wheat breeding programs. Aust J Agric Res 53:367–377

Forde J, Malpica J-M, Halford NG, Shewry PR, Anderson OD, Green FC, Miflin BJ (1985) The nucleotide sequence of a HMW glutenin subunits gene located on chromosome 1A of wheat (*Triticum aestivum* L.). Nucleic Acids Res 13:6817–6832

Friebe B, Gill BS (1996) Chromosome banding and genome analysis in diploid and cultivated polyploid wheats. In: Jauher PP (ed) Methods of genome analysis in plants. CRC, New York, pp 39–60

Gill KS, Lubbers EL, Gill BS, Raupp WJ, Cox TS (1991) A genetic linkage map of *Triticum tauschii* (DD) and its relationship to the D genome of bread wheat (AABBDD). Genome 34:362–374

Gill KS, Gill BS, Endo TR (1993) A chromosome region-specific mapping strategy reveals gene-rich telomeric ends in wheat. Chromosome 102:374–381

Gill KS, Gill BS, Endo TR, Boyko EV (1996) Identification and high-density mapping of gene-rich regions in chromosome group 5 of wheat. Genetics 143:1001–1012

Gupta RB, MacRitchie F (1991) A rapid one-step one-dimensional SDS-PAGE procedure for analysis of subunit composition of glutenin in wheat. J Cereal Sci 14:105–109

Gupta RB, MacRitchie F (1994) Allelic variation at glutenin subunit and gliadin loci, *Glu-1-Glu-3* and *Gli-1*, of common wheats. II. Biochemical basis of the allelic effects on dough properties. J Cereal Sci 19:19–29

Gupta RB, Singh NK, Shepherd KW (1989) The cumulative effect of allelic variation in LMW and HMW glutenin subunits on dough properties in the progeny of two bread wheats. Theor Appl Genet 77:57–64

Gupta RB, Paul JG, Cornish GB, Palmer GA, Bekes F, Rathjen AJ (1994) Allelic variation at glutenin subunit and gliadin loci, *Glu-1-Glu-3* and *Gli-1*, of common wheats I. Its additive and interaction effects on dough properties. J Cereal Sci 19:9–17

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

Kasarda DD, Autran JC, Lew EJ-L, Nimmo CC, Shewry RP (1983) N-terminal amino acid sequences of ω-gliadins and ω-secalins: implications for the evolution of prolamin genes. Biochim Biophys Acta 747:138–150

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

Keller B, Feuillet C (2000) Colinearity and gene density in grass genomes. Trends Plant Sci 5:246–251

Kihara H (1944) Discovery of the DD-analyzer, one of the ancestors of *Triticum vulgare* (in Japanese). Agric Hortic 19:13–14

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

Lew EJ-L, 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

Lindsay MP, Skerritt JH (1999) The glutenin macropolymer of wheat flour doughs: structure-function perspectives. Trends Food Sci Technol 10:247–253

Luo C, Griffin GB, Branlard G, McNeil DL (2001) Comparison of low- and high molecular weight wheat glutenin allele effects on flour quality. Theor Appl Genet 102:1088–1098

Maruyama N, Ichise K, Katsube T, Kishimota T, Kawase S, Matsumura Y, Takeuchi Y, Sawada T, Utsumi S (1998) Identification of major wheat allergens by means of the *Escherichia coli* expression system. Eur J Biochem 255:739–745

Masci S, Rovelli L, Kasarda DD, Vensel WH, Lafiandra D (2002) Characterisation and chromosomal localization of C-type low-molecular-weight glutenin subunits in the bread wheat cultivar Chinese Spring. Theor Appl Genet 104:422–428

Masci S, Porceddu E, Colaprico G, Lafiandra D (1991) Comparison of the B and D subunits of glutenin encoded at the *Glu-D3* locus in two biotypes of the common wheat cultivar Newton with different technological characteristics. J Cereal Sci 14:35–46

Masci S, Lafiandra D, Porceddu E, Lew EJ-L, Tao HP, Kasarda DD (1993) D-glutenin subunits: N-terminal sequences and evidence for the presence of cysteine. Cereal Chem 70:581–585

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

McFadden ES, Sears ER (1946) The origin of *Triticum spelta* and its free-threshing hexaploid relatives. J Hered 37:81–89, 107–116

Metakovsky EV, Branlard G, Chernakov VM, Upelniek VP, Redaelli R, Pogna NE (1997) Recombination mapping of some chromosome 1A-, 1B-, 1D- and 6B-controlled gliadins and low molecular-weight glutenin subunits in common wheat. Theor Appl Genet 94:788–795

Moullet O, Zhang HB, Lagudah ES (1999) Construction and characterization of a large DNA insert library from the D genome of wheat. Theor Appl Genet 99:303–313

Müller S, Vensel WH, Kasarda DD, Köhler P, Wieser H (1998) Disulphide bonds of adjacent cysteine residues in low molecular weight subunits of wheat glutenin. J Cereal Sci 27:109–116

Neuhoff V, Arold N, Taube D, Ehrhardt W (1988) Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. Electrophoresis 9:255–262

Nieto-Taladriz MT, Perretant MR, Rousset M (1994) Effects of gliadins and HMW and LMW properties in the F_6 recombinant inbred lines from a bread wheat cross. Theor Appl Genet 88:81–88

Okita TW, Cheesbrough V, Reeves CD (1985) Evolution and heterogeneity of the α/β-type and γ-type gliadin DNA sequences. J Biol Chem 260:8203–8213

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

Payne PI, Corfield KG, Blackman JA (1979) Identification of a high-molecular-weight subunit of glutenin whose presence correlates with bread-making quality in wheats of related pedigree. Theor Appl Genet 55:153–159

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

Rafalski JA (1986) Structure of wheat gamma-gliadin genes. Gene 43:221–229

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

Sandhu D, Champoux JA, Bondareva SN, Gill KS (2001) Identification and physical localization of useful genes and markers to a major gene-rich region on wheat group 1S chromosomes. Genetics 157:1735–1747

Santag-Strohm T, Payne PI, Salovaara H (1996) Effect of allelic variation of glutenin subunits and gliadins on baking quality in the progeny of two biotypes of bread wheat cv. Ulla. J Cereal Sci 24:115–124

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 (1992) High-molecular-weight subunits of wheat glutenin. J Cereal Sci 15:105–120

Shewry RP, Popineau Y, Lafiandra D, Belton P (2001) Wheat glutenin subunits and dough elasticity: findings of the EUROWHEAT project. Trends Food Sci Technol 11:433–441

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

Sissons MJ, Bekes F, Skerritt JH (1998) Isolation and functionality testing of low molecular weight glutenin subunits. Cereal Chem 75:30–36

Spielmeyer W, Moullet O, Laroche A, Lagudah ES (2000) Highly recombinogenic region at seed storage protein loci on chromosome 1DS of *Aegilops tauschii*, the D-genome donor of wheat. Genetics 155:361–367

Tao HP, Kasarda DD (1989) Two-dimentional gel mapping and N-terminal sequencing of LMW-glutenin subunits. J Exp Bot 40:1015–1020

Tranquilli G, Cuniberti M, Gianibelli MC, Bullrich, L, Larroque, OR, MacRitchie F, Dubcovsky J (2002) Effect of *Triticum monococcum* glutenin loci on bread-making and cookie quality. J Cereal Sci 36:9–18

Turnbull KM, Turner M, Mukai Y, Yamamoto M, Morell MK, Appels R, Rahman S (2003) The organization of genes tightly linked to the *Ha* locus in *Aegilops tauschii*, the D-genome donor to wheat. Genome 46:330–338

Van Campenhout S, Stappen JV, 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 Z-D, Dubcovsky J, Keller B (2003) Rapid genome divergence at orthologous low molecular weight glutenin loci of the A and Am genome of wheat. Plant Cell 15:1186–1197

Zhang W, Gianibelli MC, Rampling L, 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