Theor Appl Genet (2004) 108:1349–1358
DOI 10.1007/s00122-003-1547-y

ORIGINAL PAPER

Y. Yan · J. Zheng · Y. Xiao · J. Yu · Y. Hu · M. Cai · Y. Li · S. L. K. Hsam · F. J. Zeller

Identification and molecular characterization of a novel y-type *Glu-D^t 1* glutenin gene of *Aegilops tauschii*

Received: 30 September 2003 / Accepted: 12 November 2003 / Published online: 23 January 2004

Abstract A novel y-type high-molecular-weight glutenin subunit possessing a slightly faster mobility than that of subunit 1Dy12 in SDS-PAGE, designated 1Dy12.1^t in *Aegilops tauschi*, was identified by one- and two-dimensional gel and capillary electrophoresis. Its coding gene at the *Glu-D^t 1* locus was amplified with allele-specific-PCR primers, and the amplified products were cloned and sequenced. The complete nucleotide sequence of 2,807 bp containing an open reading frame of 1,950 bp and 857 bp of upstream sequence was obtained. A perfectly conserved enhancer sequence and the −300 element were present at positions of 209–246 bp and 424–447 bp upstream of the ATG start codon, respectively. The deduced mature protein of 1 Dy12.1^t subunit comprised 648 amino acid residues and had a *M*r of 67,518 Da, which is slightly smaller than the 1Dy12 (68,695 Da) but larger than the 1Dy10 (67,495 Da) subunits of bread wheat, respectively, and corresponds well with their relative mobilities when separated by acid-PAGE. The deduced amino acid sequence indicated that the 1Dy12.1^t subunit displayed a greater similarity to the 1Dy10 subunit, with only seven amino acid substitutions, suggesting that this novel gene could have positive effect on bread-making quality. A phenetic tree produced by nucleotide sequences showed that the x- and y-type subunit genes were respectively clustered together and that the *Glu-D^t 1y12.1* gene of *Ae. tauschii* is closely related to other y-type subunit genes from the B and D genomes of hexaploid bread wheat.

Communicated by H.F. Linskens

Y. Yan · J. Zheng · Y. Xiao · J. Yu · Y. Hu · M. Cai · Y. Li
Key Laboratory of Genetics and Biotechnology,
Department of Biology,
Capital Normal University,
100037 Beijing, China

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

Present address:
J. Yu, The Institute of Agricultural Crops (ILK),
Federal Centre for Breeding Research on Cultivated Plants (BAZ),
18190 Gross Lüsewitz, Germany

Introduction

It is well established that the high-molecular-weight glutenin subunits (HMW-GS) are important determinants of wheat technological properties (Payne 1987; Shewry et al. 1992; Gianibelli et al. 2001). They constitute part of the polymeric proteins linked by disulphide bonds and have molecular weights ranging from 80,000 Da to 120,000 Da. Genetic studies have shown that HMW-GS are encoded at the *Glu-1* loci on the long arms of bread wheat homoeologous group 1 chromosomes (Payne 1987). Each locus consists of two genes that are not always expressed, namely the x- and y-type subunits with slow and high electrophoretic mobilities, respectively (Shewry et al. 1992). To date, only some of the HMW-GS genes from bread wheat have been cloned and sequenced (Forde et al. 1985; Sugiyama et al. 1985; Thompson et al. 1985; Halford et al. 1987; Anderson and Greene 1989; Anderson et al. 1989; Halford et al. 1992; Reddy and Appels 1993; Bustos et al. 2000). All HMW glutenin subunits consist of a large central repeat region, N- and C-terminus domains, which are considered to be the most important characteristics relating to the technological properties of dough (Wrigley 1996).

The diploid Asian goatgrass, *Aegilops tauschii* (Coss.) Schmal. (2*n*=2*x*=14, DD; syn. *Ae*. *squarrosa*), is the donor of the bread wheat D genome (Dvorak et al. 1998). Extensive allelic variations in both seed storage protein composition and isozymes have been detected in *Ae. tauschii* (Lagudah et al. 1987; Lagudah and Halloran 1989; Gianibelli et al. 2001; Yan et al. 2003a, 2003b). The diploid wild species, therefore, is expected to be a useful source of genetic material for improving the quality of bread wheat. Some investigations have shown

that the HMW glutenin genes from the *Ae. tauschii* have a significant influence on bread-making properties in synthetic hexaploid wheats (Lagudah et al. 1987; Hsam et al. 2001; Wieser et al. 2003), and some of the resulting lines have exhibited a shorter mixing time and improved milling and baking characteristics when compared to parental lines (Tilley et al. 2000).

To date, only a limited number of HMW-GS genes in *Ae. tauschii* have been cloned and characterized (Mackie et al. 1996; Tilley et al. 2000; Gianibelli and Solomon 2003). Recently, we have detected genetic polymorphisms at the *Glu-D*^{t} and *Gli-D*^{t} loci of 198 *Ae. tauschii* accessions using various electrophoretic methods (Yan et al. 2003a, 2003b). The accession TD159 was found to possess only a y-type HMW glutenin subunit, which was designated 1Dy12.1^t. We report here the biochemical and molecular characterizations of this subunit, which may provide new insights into gene structures and the molecular evolutionary relationship of the glutenin gene family.

Materials and methods

Plant materials

The *Aegilops tauschii* (Coss.) Schmal. accession TD159 analysed was kindly provided by Genebank Braunschweig, Germany. The wheat variety *Chinese Spring* was used as the standard for HMW-GS identification.

One- and two-dimensional gel and capillary electrophoresis

One-dimensional sodium dodecyl sulphate (SDS)- and acid (A)-polyacrylamide gel electrophoresis (PAGE) and two-dimensional A-PAGE×SDS-PAGE identifications of HMW glutenin subunit 1Dy12.1^t were carried out on a Bio-Rad Mini-PROTEAN II electrophoretic apparatus (Bio-Rad, Hercules, Calif.) according to Yan et al. (1999; 2003d). The capillary electrophoretic separation and characterization of the HMW-GS were performed on a BioFocus 3000 instrument with phosphate-based buffers according to Yan et al. (2003c).

Genomic DNA extraction and PCR amplification

Total genomic DNA was isolated from 1–2 g of leaves sampled from single plants about 3–5 weeks old and from dry seeds following the methods of Dellaporta et al. (1983) and Plaschke et al. (1995), respectively.

On the basis of published HMW-GS gene sequences (Mackie et al. 1996; Anderson et al. 1989; Halford et al. 1987), two pairs of allele-specific (AS) PCR primers were designed to amplify different regions of the *Glu-D*^{t} *y12.1* gene as shown below. The nucleotide primers synthesized (Sangong) were as follows: (P_1) 5′-ATGGCTAAGCGGTTGGTCCT-3′; (P_2) 5′-TCACTGGCTAG-CCGACAATG-3′; (P_3) 5′-ACCACAGTTTGCTCATATTGTCT-TG-3′; (P_4) 5′-ACGTCTACACTTCTGCAAACAATACC-3′.

PCR was carried out using a Perkin-Elmer Cetus DNA Thermal Cycler (PE Applied Biosystems, Foster, City, Calif.). Samples of genomic DNA (leaf DNA: 100 ng; seed DNA: 800 ng) were subjected to PCR in a 50-µl reaction volume containing 2.5 U *La Taq* polymerase (TaKaLa), 25 µl 2×GC Buffer II ($MgCl^{2+}$ Plus), 0.4 m*M* dNTP and 0.5 µ*M* of each oligonucleotide primer. The PCR profile consisted of an initial denaturation step of 5 min at 94°C, followed by one cycle of a 45-s denaturation at 94°C, a 60-s annealing at 63°C and a 120-s extension at 72°C and concluded with a final extension step for 10 min at 72°C. PCR products were analysed on 1% agarose gels.

Cloning, sequencing and comparative analyses of HMW-GS genes

The PCR products were purified and ligated into a pGEM-T plasmid vector (Promega, Madison, Wis.). The vector was used to transform *Escherichia coli* competent cells of the DH α strain. The cloned fragment was sequenced on an automatic DNA sequencer (TaKaRa Biotech). Nucleotide sequences were assembled, and comparative analyses among different HMW-GS genes were carried out using BIOEDIT 5.09 software.

Results and discussion

Characterization of 1Dy12.1^t by different gel and capillary electrophoreses

The *Ae. tauschii* accession TD159 possesses only a single y-type subunit, 1Dy12.1^t (Yan et al. 2003b). In order to further characterize this subunit, we used different electrophoresis methods to separate the HMW-GS composition of TD159. The patterns of one- and two-dimensional gel electrophoresis are shown in Fig. 1. It

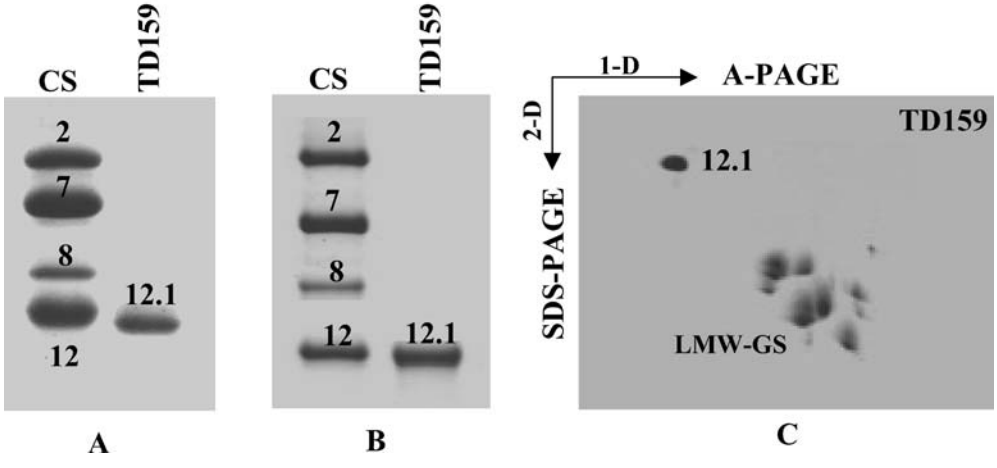

Fig. 1A–C The electrophoretic characterization of 1Dy12.1^t HMW glutenin subunit from *Aegilops tauschii* accession TD159. **A** SDS-PAGE pattern, **B** A-PAGE pattern, **C** two-dimensional pattern by A-PAGE×SDS-PAGE. Chinese Spring (*CS*) was used as control

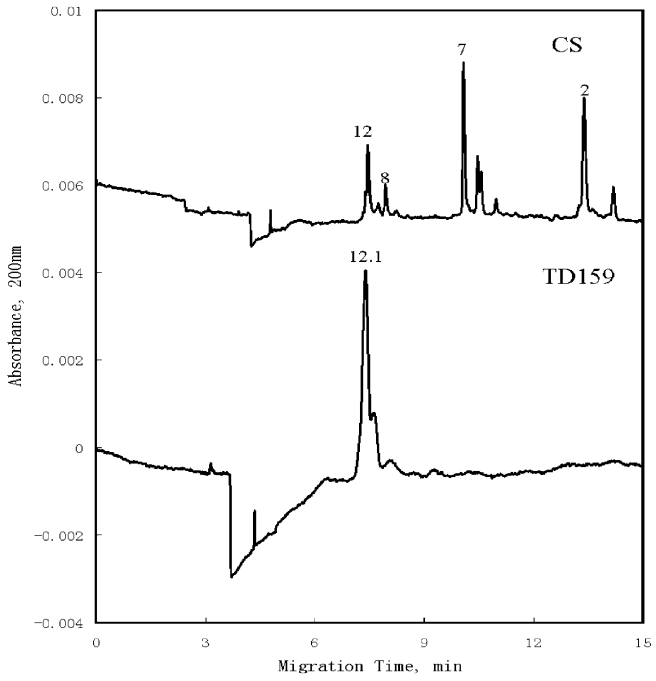

Fig. 2 Capillary electrophoretic characterization of 1Dy12.1^t subunit present in TD159. Capillary electrophoretic characterization was carried out in 0.1 *M* phosphate buffer, pH 2.5, containing 20% acetonitrile and 0.05% hydroxypropylmethyl-cellulose at 12.5 kV and 40°C with 25 μm (ID)×27-cm uncoated fused-silica capillaries. Chinese Spring (*CS*) was used as control

is apparent that TD159 possesses a y-type subunit that is slightly faster than the 1Dy12 subunit of bread wheat when both are run on SDS-PAGE and A-PAGE. This faster subunit was designated 1Dy12.1^t. There is also only one clear protein spot when it is separated by two-dimensional A-PAGE×SDS-PAGE. In addition, the capillary electrophoretic pattern of this subunit also shows a single protein peak at about 7.5 min and its migration time is also slightly shorter than that of the 1Dy12 subunit (Fig. 2). This result confirmed that the 1Dx-type subunit in this accession is not expressed and, therefore, that its coding gene is silent. In general, while gene silencing of 1Dy-type subunits at *Glu-A1* and *Glu-B1* loci exists it is exceptional to observe it at the *Glu-D1* locus in bread wheat.

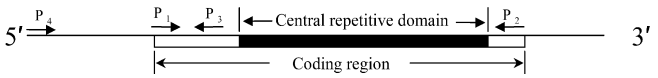

Fig. 3 Location of the oligonucleotide primers P_1+P_2 and P_3+P_4 used to amplify the coding region and the upstream sequence of the 1Dy12.1^t HMW-GS gene, respectively. The coding region of the 1Dy12.1^t subunit gene is *boxed* and the central repetitive domain is *shaded*

PCR analysis of the 1Dy12.1^t HMW-GS subunit gene

Two pairs of PCR primers, P1+P2 and P3+P4, were used to amplify the coding region and upstream sequence of the 1Dy12.1^t HMW-GS subunit gene of *Ae. tauschii* (Fig. 3), respectively, and the results are shown in Fig. 4. It is obvious that only a single PCR product of 1,947 bp is obtained with primer P1+P2, which corresponds to the length of the coding region of the 1Dy-type subunit gene (Fig. 4A). Previous reports on glutenin genes (Thompson et al. 1985; Halford et al. 1987; Anderson et al. 1989) indicated that the coding sequences of the 1Dy-type subunit genes are about 2,000 bp, so the size of the amplified product is in agreement with the size of the coding region of the 1Dy12.1^t subunit gene. Using primer P3+P4, we succeeded in amplifying one apparent single band of about 1,200 bp that corresponded to the upstream sequences of the 12.1^t gene. The DNA extracted from both leaves and dry seeds showed the same results with respect to PCR amplification.

The nucleotide and deduced amino acid sequences of the 1Dy12.1^t subunit gene

The PCR products of both the coding region and upstream sequences were cloned and sequenced by means of primer walking. The complete sequence of 2,807 bp, containing an intron-less open reading frame of 1,950 bp encoding a protein of 648 residues plus 857 bp of upstream sequence, was obtained, which ends at a double stop codon (Fig. 5).

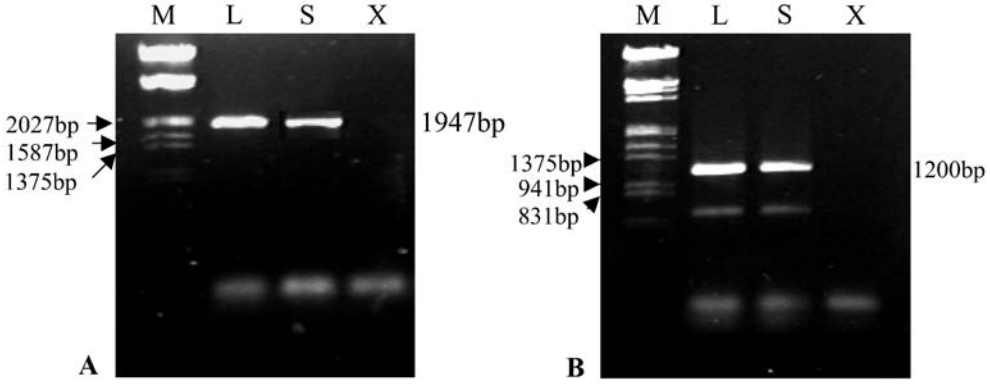

Fig. 4 Agarose-gel separation of amplified products using the primer P_1+P_2 (**A**) and P_3+P_4 (**B**). Both leaf DNA (*L*) and seed DNA (*S*) were used. *M* Lambda DNA/*Hin*dIII markers, *X* negative control

Some typical eukaryotic sequences were found in the upstream region; for instance, a TATA box and a CCACC sequence (Kozak 1987) are present at positions of 85–91 bp and 4–8 bp upstream of the ATG start codon, respectively. In addition, three CCAAT-like sequences—CCAAT (74–78 bp), CCAT (116–120 bp) and CAAAT (190–195)—are also present, as in other plant genes (Messing et al. 1983).

As shown in Fig. 5, at a position of 209–246 bp upstream of the ATG start codon, a 38-bp, perfectly conserved enhancer sequence (5′-GTTTTGCAAAGCTC-CAATTG CTCCTTGCTTATCCAGCT) of HMW-GS genes identified by Thomas and Flavell (1990) is present, which is a major regulatory element for gene expression (Shewry and Halford 2002). In addition, there is a 24-bp sequence from −447 to −424 bp, which is highly homologous with part of the −300 element described by Forde et al (1985). This sequence is also conserved in the 5′ flanking regions of B hordeins and other prolamin genes.

The deduced protein sequence of the 1Dy12.1^t subunit is similar to that of 1Dy12^t (Mackie et al. 1996) and the other 1Dy-type HMW subunits present in bread wheat (Thompson et al. 1985; Halford et al. 1987; Anderson et al. 1989). The first 21 amino acids constitute a signal peptide (Watson 1984), which is cleaved off as the protein is transported through the endoplasmic reticulum. This is homologous with the signal peptides of the other HMW glutenin subunits and other cereal prolamins (Kreis et al. 1985; Halford et al. 1987). As with the other HMW glutenin subunits, the protein structure of the mature 1Dy12.1^t subunit possesses three clear structural characters: a non-repetitive N-terminal domain of 104 residues, a repetitive central domain of 481 residues and a non-repetitive C-terminal domain of 42 residues. The cysteine residues present in the N- and C-terminal regions are considered to form intermolecular disulphide bonds that can create very high molecular weight linear polymers (Wrigley 1996). Therefore, they are important in providing viscoelastic properties to dough. Seven cysteine residues are present in the 1Dy12.1^t subunit: five in the N-terminal domain, one in the C-terminal domain and one in a nonapeptide repeat near the C-terminal end of the array of repeats.

The repetitive domain consists of tandem and interspersed repeats based on hexapeptide (consensus PGQGQQ) and nonapeptide (consensus GYYPTSLQQ) motifs marked in Fig. 5, and no tripeptide (consensus GQQ) is present. There are 49 hexapeptide and 19 nonapeptide repeat units; consequently, hexapeptide units are either adjacent or separated by a nonapeptide unit. Nonapeptides are always separated by a hexapeptide unit and never present in tandem arrays. However, like any of the other y-type subunits present in bread wheat (Forde et al. 1985; Thompson et al. 1985; Halford et al. 1987), two longer repeat units (11 followed by 12 residues) are present adjacent to each other at the beginning of the repetitive central domain. This may correspond to degenerated nonapeptides, probably resulting from the dipeptide and tripeptide insertion. Based on the above, the properties of the 1Dy12.1^t subunit are consistent with the structure of typical 1Dy-type HMW glutenin subunits. The nucleotide and deduced amino acid sequences of the coding region of the1Dy12.1^t subunit gene have been deposited in the EMBL database under accession number AY248704.

Comparative analysis of y-type subunits among diploid and hexaploid species

As indicated above, the nucleotide and deduced amino acid sequences of the 1Dy12.1^t subunit gene are closely homologous with those of the 1Dy12^t subunit (Mackie et al. 1996). A comparison of the amino acid sequence of the 1Dy12.1^t subunit with that of other y-type subunits from diploid and hexaploid wheat species is shown in Fig. 6. All five y-type subunits possess the same number and distribution of cysteine residues. In particular, three subunits (1Dy12.1^t, 1Dy12^t and 1Dy10) possess the same 648 residues and show high homology, with only a few single amino acid substitutions, whereas they exhibit greater differences to subunits 1Dy12 and 1By9 present in bread wheat. The mature 1Dy12.1^t and 1Dy12^t subunits differ in five residues—one in the N-terminal domain and four in the repetitive central domain. Both subunits display a greater similarity to subunit 1Dy10 than to 1Dy12 of bread wheat. The N-terminal domain of 1Dy12.1^t possesses a lysine residue at position 41 compared to a glutamic acid in the other four subunits, while there is a single amino acid difference between 1Dy12.1^t, 1Dy12^t (isoleucine) and the other three subunits (valine) at position 81. The remaining amino acid sequences in the N- and C-terminal domains of the 1Dy12.1^t, 1Dy12^t, 1Dy10 and 1Dy12 subunits are exactly the same.

In the repetitive central domain, 1Dy12.1^t and 1Dy12^t differ in four residues at the positions of 462 (tyrosine/serine), 499 (proline/serine), 503 (glutamine/proline) and 580 (glutamine/histidine), respectively. It is possible to draw the conclusion that these five single amino acid substitutions are responsible for the minor molecular difference between 1Dy12.1^t and 1Dy12^t on the SDS-PAGE pattern.

There are only seven amino acid substitutions between the 1Dy12.1^t and 1Dy10 subunits—at positions 41 (lysine/glutamic acid), 81 (isoleucine/valine), 285 (glutamine/ arginine), 462 (tyrosine/serine), 499 (proline/serine), 503 (glutamine/proline) and 630 (leucine/valine). However, although subunits 1Dy12.1^t and 1Dy12 exhibited very similar mobilities on both SDS- and A-PAGE gels, there were more differences between them than between the 1Dy12.1^t and 1Dy10 subunits with respect to amino acid sequence. As shown in Fig. 6, a dipeptide (GQ at 523–524) and two hexapeptide (QIGQGQ at 217–222 and 244–249) deletions exist in 1Dy12.1^t as well as in 1Dy12^t and 1Dy10, while there is a dipeptide deletion (GQ at 398–399) in the 1Dy12 subunit. In addition, there are 15 other single amino acid changes between subunits

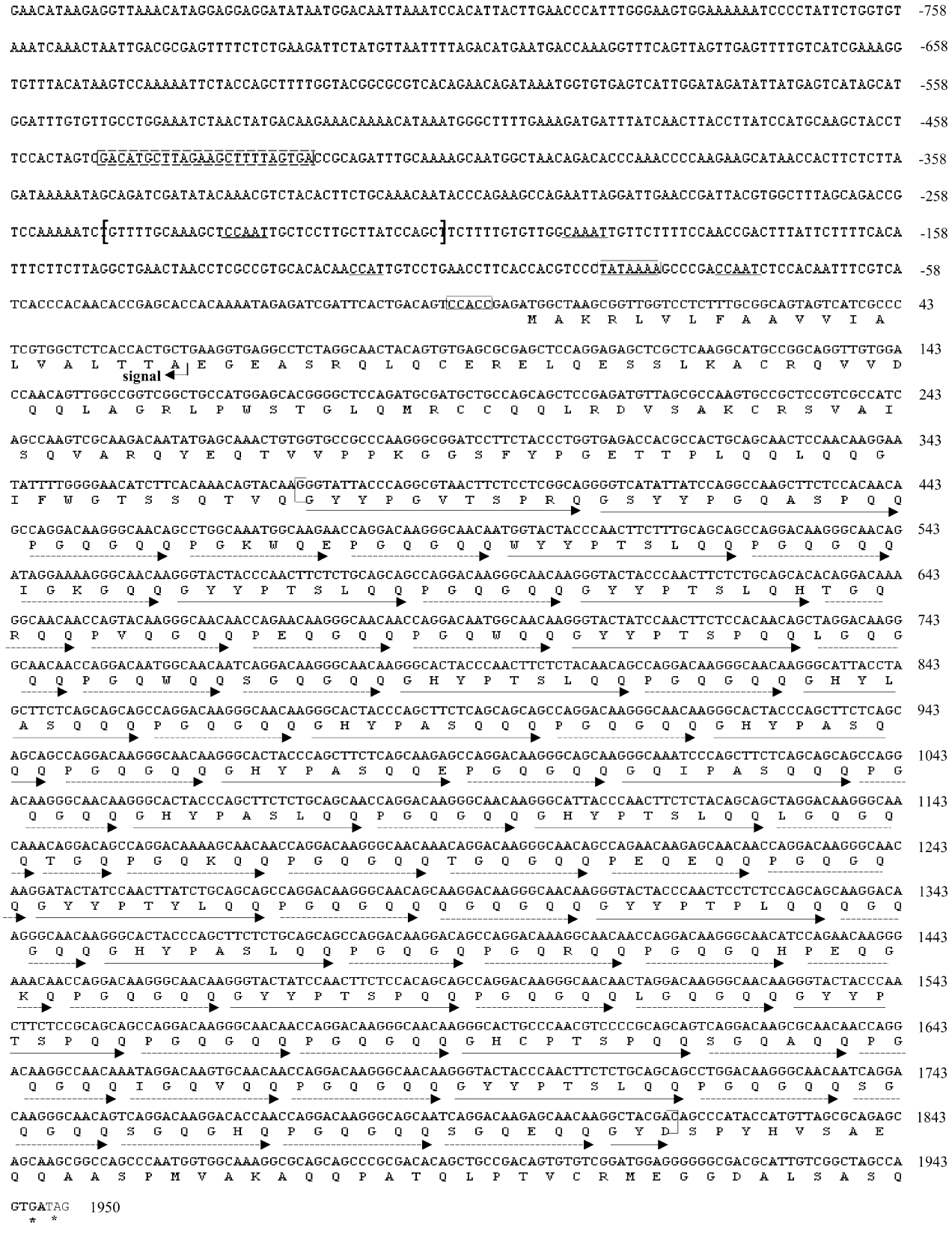

Fig. 5 Nucleotide and deduced amino acid sequences of the 1Dy12.1^{t} HMW-GS genes. The putative TATA box and CCACC sequence are *boxed*. CCAAT-like sequences are *underlined*. An enhancer element with 38 bp identified by Thomas and Flavell (1990) is enclosed in *brackets*. A possible "-300 element" described by Forde et al. (1985) is *boxed* with a *broken line*. The putative signal sequence of the encoded protein (21 residues beginning at the start codon) is indicated. The nonapeptide (consensus GYYPTSLQQ) and hexapeptide (consensus PGQGQQ) repeat motifs are indicated with a *solid* and *broken arrow*, respectively. Two stop codons are marked by *asterisks*

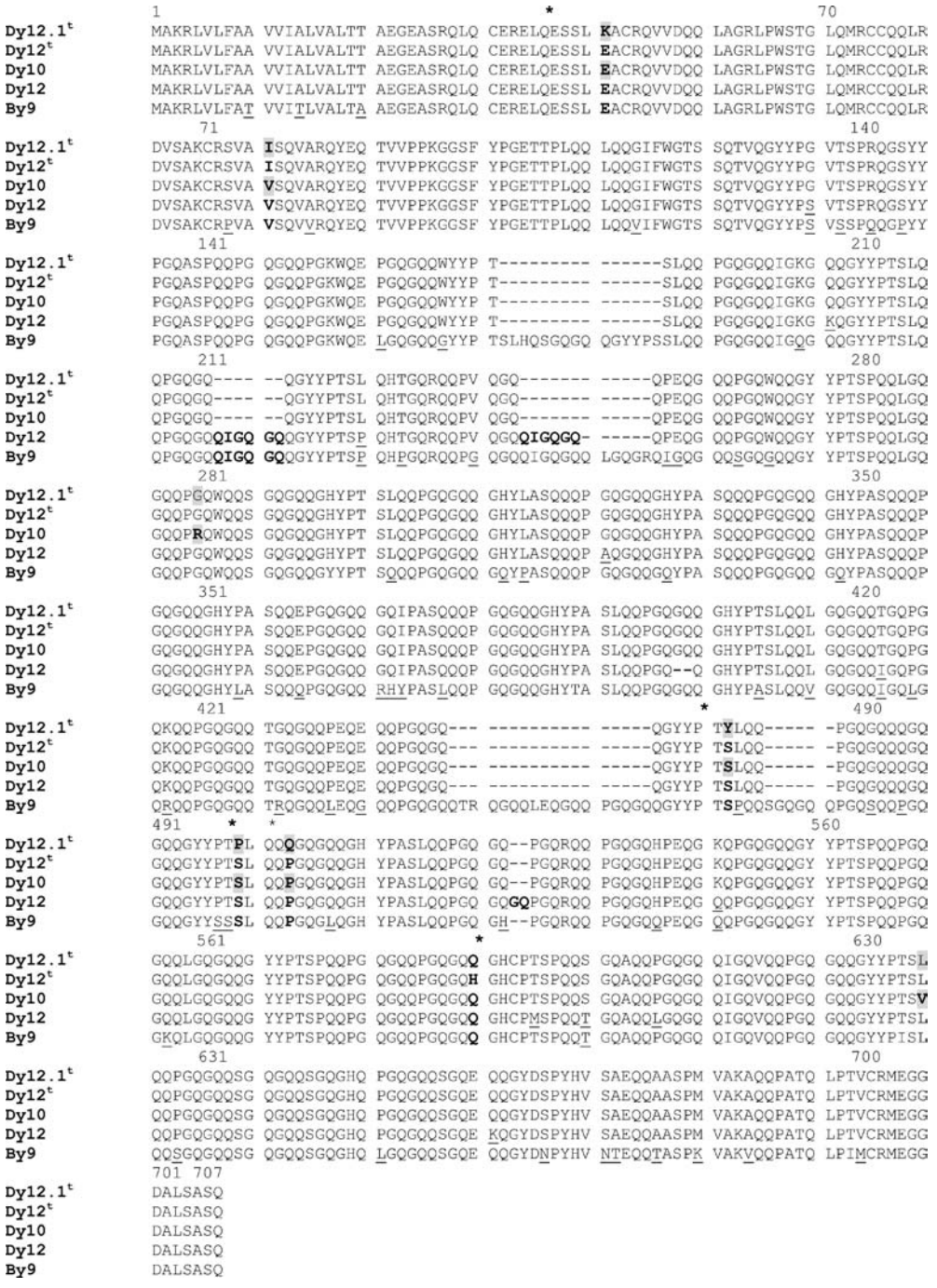

Fig. 6 Comparison of the derived amino acid sequence of the Dy12.1^t subunit with five other y-type HMW glutenin subunits from diploid and hexaploid *Triticum* species. The differences between subunits 1Dy12.1^t and 1Dy12^t are indicated with an *asterisk*. Some other differences among the y-type subunits are *underlined* or marked with *bold letters*. Deletions in all sequences are indicated by a *dash*

1Dy12.1^t and 1Dy12; these occur at positions 41, 81, 130, 201, 230, 321, 416, 462, 499, 503, 541, 585, 590, 596 and 661.

Table 1 lists the frequency of amino acid variations from tripeptide, hexapeptide and nonapeptide consensus at each of the 18 positions in six y-type and six x-type subunits at different loci from diploid and hexaploid species. It is obvious that there are more variations at positions 1 and 4 of the hexapeptide and positions 2, 5 and 7 (with the higher frequency of a proline or glutamine substitution) of the nonapeptide, whereas the remaining positions as well as the tripeptide repeats in x-type subunits are highly conserved. The mean numbers of amino acid differences from the consensus per repeat unit including hexapeptide and nonapeptide units show that the repeat units of y-type subunit 1Dy10 and x-type subunits 1Dx2, 1Dx5, 1Bx17 and 1Bx7 were closer to the consensus than the other subunits.

In order to gain an understanding of the evolutionary relationships among HMW-GS genes from diploid and hexaploid species, we used the nucleotide sequences to produce a dendrogram of the 1Dy12.1^t subunit together

Table 1 Comparison of the repeat unit variations with the consensus among y- and x-type subunits from diploid and hexaploid species (number of amino acids which differ from consensus in tripeptide, hexapeptide and nonapeptide)

HMW Glutenin subunits	Tripeptide				Hexapeptide							Nonapeptide										M[b]
	1 G	2 Q	3 Q	Total units[a]	1 P	2 G	3 Q	4 G	5 Q	6 Q	Total units[a]	1 G	2 Y	3 Y	4 P	5 T	6 S	7 L	8 Q	9 Q	Total units[a]	
1Dy12.1[t]	-	-	-	0	17	5	3	13	6	6	49(1.02)	2	10	2	1	7	2	9	0	2	19(1.84)	1.25
1Dy12[t]	-	-	-	0	16	5	3	13	5	7	49(1.00)	2	10	2	1	7	0	9	0	1	19(1.68)	1.19
1Dy10	-	-	-	0	15	6	3	12	4	5	49(0.88)	1	10	2	1	7	0	10	0	2	19(1.74)	1.15
1Dy12	-	-	-	0	18	6	3	13	6	5	51(1.00)	1	10	2	1	8	0	10	0	2	19(1.79)	1.21
1By9	-	-	-	0	26	6	2	10	6	3	57(0.93)	1	9	1	3	11	0	11	1	1	20(1.90)	1.18
1Ay	-	-	-	0	17	8	4	16	3	5	48(1.10)	1	6	3	2	5	1	4	1	3	15(1.67)	1.25
1Dx5	1	0	0	23(0.04)	18	13	1	14	4	16	73(0.90)	2	0	1	2	0	1	15	3	4	19(1.47)	1.02
1Dx2	1	1	0	20(0.10)	20	13	0	14	3	15	72(0.90)	1	0	1	2	0	0	15	3	3	19(1.32)	0.99
1Ax1	3	1	0	18(0.22)	30	15	1	22	7	14	69(1.29)	3	1	0	3	1	0	20	3	0	21(1.48)	1.33
1Ax2*	3	1	0	16(0.25)	31	16	2	22	6	14	67(1.36)	3	1	0	3	1	0	20	3	0	21(1.48)	1.39
1Bx7	1	0	0	4(0.25)	31	7	1	17	3	11	67(1.05)	3	0	1	0	2	0	20	2	0	23(1.22)	1.09
1Bx17	1	0	0	4(0.25)	30	6	1	16	3	11	64(1.05)	3	0	1	0	2	0	18	2	0	21(1.24)	1.09

[a] The numbers in brackets indicate the number of variants per unit
[b] M, Average number of variants per unit, including hexapeptides and nonapeptides

with other 15 HMW subunits from bread wheat and *Ae. tauschii* (Fig. 7), namely 1Bx7 (Anderson and Greene 1989, EMBL accession number X13927), 1Bx17 (Reddy and Appels 1993, JC2099), 1Ax2* (Anderson and Greene 1989, M22208), 1Ax1 (Halford et al. 1992, X61009), 1Ax-null (Bustos et al. 2000, AF145590), 1Dx5 (Anderson et al. 1989, X12928), 1Dx2[t] (AF480485), 1Dx2.1[t] (AF480485), 1Dx2 (Sugiyama et al. 1985, X03346), 1Ay (Forde et al. 1985, X03042), 1By9 (Halford et al. 1987, X61026), 1Dy10 (Anderson et al., 1989, X12929), 1Dy12[t] (Mackie et al. 1996, U39229), 1Dy13[t] (AY174159), and 1Dy12 (Thompson et al. 1985, X03041).On the resulting dendrogram it is clear that x- and y-type HMW subunit genes cluster together in separate groups and, in particular, that the 1Dy12.1[t] subunit gene is closely related to the 1Dy10 gene with a homology of 98%. This further confirms the 1Dy12.1[t] as a y-type subunit. Within a cluster, HMW glutenin subunits encoded by the *Glu-D1* locus on the D genome of diploid and hexaploid species display a greater similarity. Furthermore, the y-type subunits (except for the 1Ay silent gene) show more similarity than the x-type subunits.

Table 2 lists the molecular weight of 11 HMW glutenin subunits that was calculated on the basis of nucleotide sequences and their relative migration orders under different separation methods. It is known that some HMW glutenin subunits (subunits 1, 2* and 10) differ in their mobilities when compared by SDS-PAGE and A-PAGE methods (Morel 1994; Yan et al. 2003b). In general, A-PAGE and acidic high-performance capillary electrophoresis (A-HPCE) as well as SDS-PAGE and SDS-HPCE have similar migration orders. However, the relative mobilities of HMW glutenin subunits under A-PAGE are well in agreement with their true molecular weights. For example, the relative mobility of subunit 1Dy10 (67,495 Da) is faster than that of the subunit 1Dy12 (68,695 Da) when separated by A-PAGE, while the 1Dy12.1[t] (67,518 Da) subunit is located between the subunits 1Dy12 and 1Dy10.

Allelic variation and evolution of HMW glutenin subunit genes in *Ae. tauschii*

The hexaploid bread wheats (*Triticum aestivum* L., 2*n*=6*x*=42, AABBDD) originated from interspecific hybridization between tetraploid *Triticum turgidum* ssp. *dicoccum* (2*n*=4*x*=28, AABB) and diploid *Ae. tauschii* (Dvorak et al. 1998). Investigations on HMW-GS composition have shown that extensive allelic variation exists at the *Glu-D[t] 1* locus in diploid *Ae. tauschii* species, and several novel HMW subunits have been detected—for example 1Dy12.1[t], 1Dy12.2[t], 1Dy12.3[t], 1Dy12.4[t], 1Dy12.5[t], 1Dy10.1[t], 1Dy10.2[t], 1Dy10.3[t], 1Dy10.4[t], 1Dy10.5[t], 1Dy1.5[t], 1Dy1.5*[t] etc.—which have not been found in hexaploid bread wheat (Lagudah and Halloran, 1989; Gianibelli et al. 2001; Yan et al. 2003b). This suggests that only a small number of *Ae. tauschii* genotypes of restricted geographic origin were involved

Table 2 The molecular weight of HMW glutenin subunits calculated from nucleotide sequences and their relative migration order by different separating methods

HMW-GS	*M*r(Da)	SDS-PAGE[a]	Urea SDS-PAGE[a]	A-PAGE[a]	A-HPCE[b]	SDS-HPCE[c]
1Dx5	88137	4	1	2	2	2
1Ax1	87680	1	3	3	5	3
1Dx2	86987	3	2	1	1	1
1Ax2*	86317	2	4	4	6	4
1Bx7	82873	5	5	5	4	5
1Bx17	80750	6	7	6	3	6
1By9	73526	7	11	7	7	7
1Dy12	68695	9	6	8	9	9
1Dy12.1^t	67518	10	9	9	10	10
1Dy10	67495	8	10	10	8	8
1Dy12^t	67411	9	8	8	9	9

[a] Number indicates the order of HMW glutenin subunits in relative mobilities from low to high obtained by different gel electrophoresis methods based on Yan et al. (2003b) and Mackie et al. (1996b)

[b] Separation order in relative migration time from high to low as shown by acidic high performance capillary electrophoresis (A-HPCE) with 0.1 *M* phosphate-glycine buffer (pH 2.5, containing 20%CAN and 0.05% HPMC according to Yan et al. (2003c)

[c] Separation order in relative migration time from high to low as shown by SDS-HPCE using the ProSort SDS-protein analysis kit modified by the addition of 5% methanol according to Yan et al (2003c)

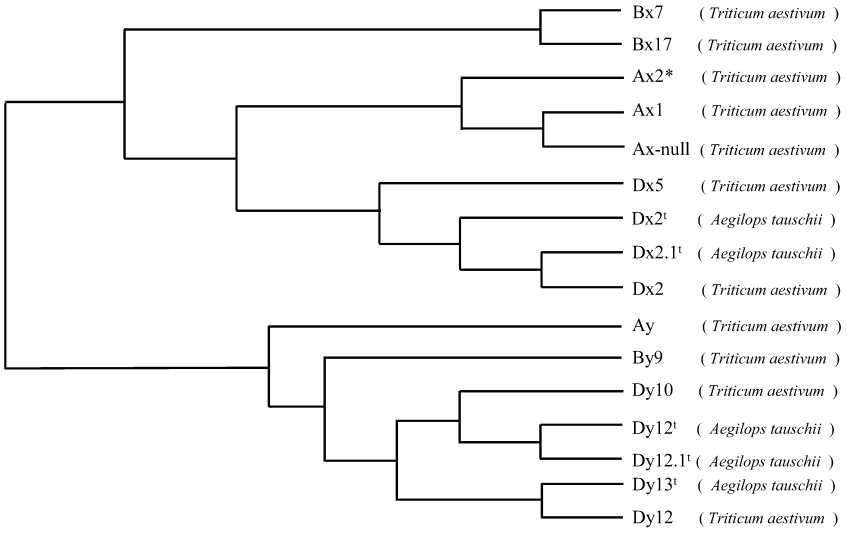

Fig. 7 A phenetic tree derived from the sequence alignment of the HMW-GS genes from different *Triticum* species

in the polyloidization process during the origin of *T. aestivum*. Therefore, a narrow genetic diversity for the D-genome may exist in bread wheat. It is expected that the extensive allelic variations of HMW glutenin subunits present in the wild diploid species could widen the genetic background of bread wheat for the improvement of baking quality.

It is apparent that molecular weight differences among HMW glutenin subunits are mainly due to the central repeat region size. Our comparison of the amino acid sequence of the 1Dy12.1^t and 1Dy12 subunits showed that they differ by 15 amino acid substitutions, two hexapeptide and one dipeptide deletions and one dipeptide insertion (Fig. 6), resulting in subunit 1Dy12.1^t being slightly smaller than subunit 1Dy12. Recent reports have also revealed that subunit 1Dy12.4^t, a small HMW glutenin subunit detected in *Ae. tauschii* (Gianibelli et al. 2001) displays deletions in the central repetitive motifs (Gianibelli and Solomon 2003). Therefore, as proposed by Anderson and Greene (1989), the following modes may result in sequence alteration and evolution: (1) single base and single repeat changes, (2) deletions or additions within a repeat, (3) deletions or duplications of blocks of repeats, with possible unequal crossover and slip-mismatching during the evolutionary process of HMW-GS genes.

The phylogenetic tree that we constructed on the basis of nucleotide sequence showed that the 1Dy12.1^t subunit gene sequence was most closely related to those of the 1Dy12^t and 1Dy10 subunit genes. Sixteen x-type and y-type HMW-GS genes from wild diploid and hexaploid species apparently clustered into two respective groups. Shewry et al. (1989) showed that these two lineages of glutenin genes diverged from an ancestral sequence prior to the separation of the wheat genomes. In general, the y-type genes clustered together more closely than the x-type

genes, suggesting that the two lineages may be evolving at differential rates. Our results are similar to those of Reddy and Appels (1993) in that they show that the y-type subunit genes have accumulated fewer changes and, therefore, have evolved at a slower rate than the x-type subunit genes during the evolution of the storage protein gene family.

Relationship between glutenin structure and dough quality

The strength of gluten is highly correlated with different combinations of the HMW glutenin subunits. For example, the 1Dx5+1Dy10 subunits are associated with good bread-making quality, while the 1Dx2+1Dy12 subunits as well as 1Bx20 are associated with poor bread-making quality (Redaelli et al. 1997; Shewry et al. 2003). However, the molecular basis that results in quality differences between HMW glutenin subunits is not clear. According to the results obtained from the present study and previous reports, the HMW glutenin proteins consist of non-repetitive N- and C-termini amino acid stretches and a large central repeated domain, including tripeptide (x-type subunits), hexa and nonapeptide (y-type subunits) repeat units. The cysteine residues present in N-termini (normally three to five) and C-termini (only one) form intermolecular disulphide bonds to create very large polymers that play an important role in providing viscoelastic properties to dough. The central repeated domain adopts a β-spiral structure that confers elasticity to the protein molecule (Gianibelli et al. 2001). In addition to disulphide cross-links in the glutenin polymers, dityrosine cross-links may play an important role in determining glutenin structure and functionality (Tilley et al. 2001). Flavell et al (1989) found that subunit 1Dy10 has a higher proportion of repeats of the consensus type than subunit 1Dy12 and subsequently postulated that this produces a more regular pattern of repetitive β-turns in the polymers. Recent reports have shown that subunit 1Bx 20 displays two cysteine residue substitutions by tyrosines in the N-terminal domain when compared with subunit 1Bx7; this is considered to be responsible for the negative effect on dough strength by decreasing the number—and affecting the pattern—of disulphide cross-links in the glutenin polymers (Shewry et al. 2003).

It is noteworthy that subunits 1Dy12.1^t and 1Dy10 are highly similar in their sequence and repeated structure, in particular that both subunits possess seven cysteine residues and a similar proportion of repeats of the hexa and nonapeptide consensus types. This suggests that subunit 1Dy12.1^t could have a positive effect on bread-making quality similar to that of the good quality subunit 1Dy10. Further investigations on the expression and functional properties of 1Dy12.1^t gene are currently underway.

Acknowledgements This investigation was financially supported by research grants (nos. 39870481 and 5002003) from the National Natural Science Foundation of China and Beijing Natural Science Foundation, and German Academic Exchange Service (DAAD) Scholarship to Y. Yan, respectively.

References

Anderson OD, Greene FC (1989) The characterization and comparative analysis of HMW glutenin genes from genomes A and B of hexaploid bread 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

Bustos AD, Rubio P, Jouve N (2000) Molecular characterization of the inactive allele of the gene Glu-A1 and the development of a set of AS-PCR markers for HMW glutenins of wheat. Theor Appl Genet 100:1085–1094

Dellaporta SL, Woods J, Hicks JB (1983) A plant DNA mini preparation II. Plant Mol Biol 1:9-21

Dvorak J, Luo MC, Yang ZL, Zhang HB (1998) The structure of the *Aegilops tauschii* gene pool and the evolution of hexaploid wheat. Theor Appl Genet 97:657–670

Flavell RB, Doldsbrough AP, Robert LS, Schnick D, Thompson (1989) Genetic variation in wheat HMW glutenin subunits and the molecular basis of bread-making quality. Biotechnology 7:1281–1285

Forde BG, Heyworth A, Pywell J, Kreis M (1985) Nucleotide sequence of a B1 hordein gene and the identification of possible upstream regulatory elements in endosperm storage protein genes from barley, wheat and maize. Nucleic Acids Res 13:7327–7339

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

Gianibelli MC, Solomon RG (2003) A novel y-type high Mr glutenin subunit (12.4^t) present in *Triticum tauschii*. J Cereal Sci 37:253–256

Gianibelli MC, Gupta RB, Lafiandra D, Margiotta B, MacRitchie F (2001) Polymorphism of high *Mr* glutenin subunits in *Triticum tauschii*: characterization by chromatography and electrophoresis methods. J Cereal Sci 33:39–52

Halford NG, Forde J, Anderson OD, Greene FC, Shewry PR (1987) The nucleotide and deduced amino acid sequence of an HMW glutenin subunit gene from chromosome 1B of bread wheat (*Triticum aestivum* L.) and comparison with those of genes from chromosome 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 AS, Shewry PR (1992) Analysis of HMW glutenin subunits encoded by chromosome 1A of bread wheat (*Triticum aestivum* L.) indicates quantitative effects on grain quality. Theor Appl Genet 83:373–378

Hsam SLK, Kieffer R, Zeller FJ (2001) Significance of *Aegilops tauschii* glutenin genes on breadmaking properties of wheat. Cereal Chem 78:521–525

Kozak M (1987) An analysis of 5′-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res 15:8125–8148

Kreis M, Shewry PR, Forde BG, Miflin BJ (1985) Structure and evolution of seed storage proteins and their genes with particular reference to those of wheat, barley and rye. Oxford Surv Plant Mol Cell Biol 2:253–317

Lagudah ES, Halloran GM (1989) Phylogenetic relationships of *Triticum tauschii*, the D genome donor to hexaploid wheat. 3. Variation in, and the genetics of, seed esterase (Est-5). Theor Appl Genet 77:851–856

Lagudah ES, MacRitchie F, Halloran GM (1987) The influence of high-molecular-weight subunits of glutenin from *Triticum*

tauschii on flour quality of synthetic hexaploid wheat. J Cereal Sci 5:129–138

Mackie AM, Sharp PJ, Lagudah ES (1996) The nucleotide and derived amino acid sequence of a HMW glutenin gene from *Triticum tauschii* and comparison with those from the D genome of bread wheat. J Cereal Sci 24:73–78

Messing J, Geraghty D, Heidecker G, Hu N-T, Kridl J, Rubenstein I (1983) Plant gene structure. In: Kosuge T, Meredith CP, Hollaender A (eds) Genetic engineering of plants: an agricultural perspective. Plenum, New York

Morel MH (1994) Acid-polyacrylamide gel electrophoresis of wheat glutenins: a new tool for the separation of high and low molecular weight subunits. Cereal Chem 7:238–242

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

Plaschke J, Ganal MW, Röder MS (1995) Detection of genetic diversity in closely related bread wheat using microsatellite markers. Theor Appl Genet 91:1001–1007

Redaelli R, Pogna NE, Ng PKW (1997) Effects of prolamins encoded by chromosomes 1B and 1D on the rheological properties of dough in near-isogenic lines of bread wheat. Cereal Chem 74:102–107

Reddy P, Appels R (1993) Analysis of a genomic DNA segment carrying the high-molecular-weight (HMW) glutenin Bx17 subunit and its use as an RFLP marker. Theor Appl Genet 85:616–624

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 (1989) The high molecular weight subunits of wheat, barley and rye: genetics, molecular biology, chemistry and role in wheat gluten structure and functionality. Oxford Surv Plant Mol Cell Biol 6:163–219

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

Shewry PR, Gilbert SM, Savage AWJ, Tatham AS, Wan YF, Belton PS, Wellner N, D'Ovidio R, Bekes F, Halford NG (2003) Sequence and properties of HMW subunit 1Bx20 from pasta wheat (*Triticum durum*) which is associated with poor end use properties. Theor Appl Genet 106:744–750

Sugiyama T, Rafalski A, Peterson D, Soll D (1985) A wheat HMW glutenin subunit gene reveals a highly repeated structure. Nucleic Acids Res 13:8729–8737

Sutton KH, Bietz JA (1997) Variation among high molecular weight subunits of glutenin detected by capillary electrophoresis. J Cereal Sci 25:9-16

Thomas MS, Flavell RB (1990) Identification of an enhancer element for the endosperm-specific expression of high molecular weight glutenin. Plant Cell 2:1171–1180

Thompson RD, Bartels D, Harberd NP (1985) Nucleotide sequence of a gene from chromosome 1D of wheat encoding a HMW-glutenin subunit. Nucleic Acids Res 13:6833–6846

Tilley KA, Benjamin RE, Bagorogoza KE, Okot-Kotber BM, Prakash O, Kwen H (2001) Tyrosine cross-links: molecular basis of gluten structure and function. J Agric Food Chem 49:2627–2632

Tilley M, Bean SR, Seib PA, Sears RG, Lookhart GL (2000) PCR amplification and DNA sequencing of high molecular weight glutenin subunits 43 and 44 from *Triticum tauschii* accession TA2450. In: Shewry PR, Tatham AS (eds) Wheat gluten. The Royal Society of Chemistry, UK, pp 105–108

Watson MEE (1984) Compilation of published signal sequences. Nucleic Acids Res 12:5145–5164

Wieser H, Hsam SLK, Zeller FJ (2003) Relationship between the qualitative and quantitative compositions of gluten protein types and technological properties of synthetic hexaploid wheat derived from *Triticum durum* and *Aegilops tauschii*. Cereal Chem 80:247–251

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

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 Comm 27:251–257

Yan Y, Hsam SLK, Yu JZ, Jiang Y, Zeller FJ (2003a) Genetic polymorphisms at *Gli-D*^{t} gliadin loci in *Aegilops tauschii* as revealed by acid polyacrylamide gel and capillary electrophoresis. Plant Breed 122:120–124

Yan Y, Hsam SLK, Yu JZ, Jiang Y, Zeller FJ (2003b) Allelic variation of the HMW glutenin subunits in *Aegilops tauschii* accessions detected by sodium dodecyl sulphate (SDS-PAGE), acid polyacrylamide gel (A-PAGE) and capillary electrophoresis. Euphytica 130:377–385

Yan Y, Yu JZ, Jiang Y, Hu Y, Cai M, Hsam SLK, Zeller FJ (2003c) Capillary electrophoresis separation of high molecular weight glutenin subunits in bread wheat (*Triticum aestivum*) and related species with phosphate-based buffers. Electrophoresis 24:1429–1436

Yan Y, Hsam SLK, Yu JZ, Jiang Y, Ohtsuka I, Zeller FJ (2003d) HMW and LMW glutenin alleles among putative tetraploid and hexaploid spelt wheat (*Triticum spelta* L.) progenitors. Theor Appl Genet 107:1321–1330