Molecular characterization of HMW-GS 1Dx3′ and 1Dx4′ genes from Aegilops tauschii and their potential value for wheat quality improvement

K. WANG1, X. L. AN1, L. P. PAN1, K. DONG1, L. Y. GAO1, S. L. WANG1, Z. Z. XIE3, Z. ZHANG, R. APPELS2, W. MA2 and Y. M. YAN1

1Key Laboratory of Genetics and Biotechnology, College of Life Science, Capital Normal University, Beijing, PR China
2Western Australian Dept of Agriculture and Food, State Agriculture Biotechnology Centre, Centre for Comparative Genomics, Murdoch University, Perth, WA, Australia


Two x-type high molecular weight glutenin subunits (HMW-GS) in Aegilops tauschii, 1Dx3′ and 1Dx4′ were identified by SDS-PAGE and MALDI-TOF-MS. Their complete coding sequences were isolated by AS-PCR. 1Dx3′ and 1Dx4′ genes consist of 2535 bp and 2508 bp and encode 845 and 836 amino acid residues, respectively. The deduced molecular masses of 1Dx3′ and 1Dx4′ gene products are 87655.26 Da and 86664.24 Da, respectively, well corresponding to the molecular masses measured by MALDI-TOF-MS. A total of 18 SNPs were identified between 1Dx3′ and 1Dx4′. Comparing with 1Dx5 subunit, 1Dx3′ had a six amino acid insertion at 146-151 while the 1Dx4′ had a nine amino acid deletion when compared with 1Dx3′ subunit. The authenticity of the cloned 1Dx3′ and 1Dx4′ genes were confirmed by successful expression of their ORFs in E. coli. Comparison and phylogenetic tree based on the amino acid and nucleotide sequences confirmed that 1Dx3′ was most closely related to 1Dx5 subunit that is widely accepted as a superior subunit for bread-making property. The secondary structure prediction demonstrated that 1Dx3′ subunit has significantly high α-helix and β-strand contents, suggesting it might have positive effects on dough quality.

Yueming Yan, Key Laboratory of Genetics and Biotechnology, College of Life Science, Capital Normal University, Beijing, PR China. E-mail: yanyym@hotmail.com

Bread-making quality is largely determined by the high molecular weight glutenin subunits (HMW-GS), the major class of seed storage proteins in wheat (Payne 1987; Shewry et al. 1992). HMW-GS are encoded by the Glu-1-1 (x type) and Glu-1-2 (y type) genes at the Glu-1 loci on 1A, 1B and 1D chromosomes (Lawrence and Shepherd 1981). Extensive investigations showed that there existed considerable diversity at the Glu-1 loci, especially in the related species of bread wheat. Payne and Lawrence (1983) presented a catalogue of HMW-GS alleles in hexaploid wheat, including 20 alleles and 25 subunits. Most of these subunit genes have been cloned and well characterized. Only few of these genes, such as Dx3, Dx4, remain uncloned.

The molecular structures of HMW-GS comprise three distinct domains, a central large repetitive domain flanked by short N- and C-terminal non-repetitive domains (Shewry et al. 1995; Gianibelli et al. 2001a). The size differences of HMW-GS are mainly resulted from repetitive domain variations (Shewry et al. 1992). Many studies have demonstrated that some HMW subunits, such as 1Dx5 + 1Dy10, 1Ax1, are highly related to good bread-making quality, and have been successfully used for wheat quality improvement by a range of genetic approaches especially by gene transformation (Altpeter et al. 1996; Barro et al. 1997; Shewry et al. 2006). However, the number of such superior glutenin genes is still limited. Therefore, searching, cloning and characterizing new candidate genes are an eminent approach for further wheat quality improvement through genetic transformation.

Aegilops tauschii (2n = 2x = 14, DD) is the progenitor of the D genome in cultivated bread wheat (Kihara 1944, McFadden and Sears 1946, Dvorka et al. 1998). The D genome of Ae. tauschii was considered to be the main contributor to bread-making properties of bread wheats (Dong et al. 1991). Since hexaploid wheat is related to only a small number of Ae. tauschii genotypes of certain geographic origin (Lagudah et al. 1991), this resulted in a narrow genetic diversity for the D genome in Triticum aestivum. Extensive HMW-GS allelic variations are present in Ae. tauschii compared to common wheat (Pfluger et al. 2001; Gianibelli et al. 2001b; Yan et al. 2003, 2004). Thus, Ae. tauschii is expected to be utilized as a potentially useful gene source for bread wheat improvement.

© 2012 The Authors. This is an Open Access article. DOI: 10.1111/j.1601-5223.2011.02215.x
To date, only limited numbers of x-type HMW glutenin genes have been isolated from Ae. tauschii accessions (Lu et al. 2005; Wan et al. 2005; Zhang et al. 2008b; An et al. 2009). In the current, two novel x-type HMW glutenins 1Dx3t and 1Dx4t were cloned and characterized from Ae. tauschii; their potential effects on dough strength and consequently bread-making quality are discussed.

MATERIAL AND METHODS

Plant material

Aegilops tauschii (Coss.) Schmal. accessions T67 and T132 were used in this study; kindly provided by Genebank Gatersleben (TD67: Triticum tauschii 423 Gatersleben; TD 132: Aegilops tauschii AE 432/80 Gatersleben). Cultivars Chinese Spring (CS) and Contrast (which contains the 1Dx5 subunit) were used as the control for HMW-GS identification.

SDS-PAGE, DNA extraction and PCR amplification

HMW-GS were extracted from a half kernel (about 20 mg) and analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) based on Yan et al. (2003). Genomic DNA of single dry seeds was extracted following the procedure of An et al. (2006). About 20 mg of crushed seed were placed in a 1.5 ml tube followed by maceration in 1 ml of extraction buffer (200 mM Tris-Hcl (pH 7.5), 288 mM NaCl, 25 mM EDTA and 0.5% SDS). The supernatant was retained and 700 μl phenol: chloroform: iso-amyl alcohol (25:24:1) was added and 0.5% SDS. The supernatant was washed two times by 70% ethanol and dissolved in H2O. Primers were designed based on the published HMW-GS gene sequences to amplify the coding sequence of 1Dx3t and 1Dx4t subunit genes from T67 and T132, respectively. Primers for 1Dx3t were: P1F: 5′-ATGGCTAAGCGGC/TTA/GGTCCTCTTTG-3′, P2 R: 5′-CTATCACTGGCTG/AGCCGACAATGC-3′, while primers for 1Dx4t gene were: Dx1-: 5′-ATGGCTAAGCGGTAGTCTG-3′, Dx2-: 5′-GCTCGAGAGTTCTAC-3′.

PCR amplifications were carried out in a total volume of 50 μl containing 2.5 U LA Taq polymerase (TaKaRa), 60 ng of templet DNA, 25 μl 2× GC buffer I (MgCl2 plus), 0.4 mM dNTP, 0.5 μM of each primer, and double distilled H2O added to a final volume of 50 μl. The reaction was performed in a PTC-100 (MJ Research) according to the following protocol: heat lid turned on, initial denaturation at 94°C for 4 min, cycled 30 times at 94°C for 45 s, 64/56°C for 1 min and 72°C for 130 s, and a final extension at 72°C for 10 min. PCR products were analyzed by agarose electrophoresis with 1% gel in Tris-acetic acid-EDTA buffer.

Molecular cloning and sequencing

The PCR fragments of expected size were purified from the gel using the Gel Extraction Kit (Omega). Purified products were ligated into pMD18-T vector (TaKaRa, Otsu, Shiga, Japan) and transformed into Escherichia coli TOP 10 cells. DNA sequencing from three clones of each PCR fragment was carried out by Sunbiotech.

Expression of the cloned HMW-GS genes in E. coli

The 1Dx3t and 1Dx4t genes were re-amplified to remove the signal peptides by designing a new pair of primers pETyF (5′-AGTCATATGGAAGGTGGCCTTCT GAG-CAACTAC-3′) and pETyR (5′-ATGAATTCTACTG-GCTGGCCGACAA TGC-3′). Ncol and EcoRI sites (underlined) were incorporated into the 5′ ends of the pETyF and pETyR, respectively. After purification, the PCR products of 1Dx3t and 1Dx4t genes were ligated into the expression vector pET-28a and pET-30a (Novagen), and transformed into E. coli BL21 (DE3) phylS cells. The expressed protein extraction and separation were carried out by SDS-PAGE according to Li et al. (2007).

MALDI-TOF-MS

The heterologous expressing proteins and their corresponding subunits in the seed endosperm were detected by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) according to previous reported methods (Pei et al. 2007; Zhang et al. 2008a; Yan et al. 2009). Shimadzu corporation AXIMA-CFRTM Plus MS apparatus (Japan) and the matrix of sinapinic acid (SA, α-cyano-4-hydroxycinnamic acid) were used. International two-dot calibration with standard sample albumin–aldrase with respective masses of 39 212.88Da and 66 431.08Da was used.

SNPs and InDels identification and secondary structure prediction

Identification of single-nucleotide polymorphisms (SNPs) and insertions/deletions (InDels) in HMW glutenin genes was performed with software Bioedit 7.0. Prediction of secondary structure of deduced amino acid sequences was carried out by using PSIPRED server (http://bioinf.cs.ucl.ac.uk/psipred/psiform.html) (McGuffin et al. 2000; Bryson et al. 2005; Li et al. 2007).

Phylogenetic analysis

Multiple alignment of homologous nucleotide sequences was created by using ClustalW program. This alignment
file was used to constructed phylogenetic tree based on the complete coding regions with software MEGA4.0.

RESULTS

Identification of 1Dx3t and 1Dx4t subunits

SDS-PAGE (Fig. 1a) analysis showed that the electrophoretic mobility of both 1Dx3t and 1Dx4t was between subunits 1Dx2 and 1Dx5, being slightly faster than that of 1Dx2 and slower than that of 1Dx5. The 1Dx3t subunit was slightly slower than 1Dx4t subunit. MALDI-TOF-MS analysis revealed that 1Dx3t and 1Dx4t subunits had the molecular mass of 87653.3 Da and 86934.5 Da, respectively (Fig. 1b–c). These two figures are consistent with the mobility order on SDS-PAGE gel.

Molecular characterization of 1Dx3t and 1Dx4t genes

PCR amplification and sequencing showed that both 1Dx3t and 1Dx4t genes were successfully isolated from Aegilops tauschii accessions T67 and T132, respectively. The 1Dx3t gene consisted of 2535 bp encoding 845 amino acid residues and the deduced molecular mass was 87655.26 Da, which well corresponded to the result from MALDI-TOF-MS (87653.3 Da). The 1Dx4t gene comprised 2508 bp encoding 836 amino acid residues. The deduced molecular mass was 86664.24 Da, 270.26 Da lower than that of MALDI-TOF-MS (86934.5 Da), but within the range of experiment error for HMW-GS molecular mass measurement (Hickman et al. 1995). The nucleotide and deduced amino acid sequences of the coding region of the 1Dx3t and 1Dx4t subunit genes were deposited in the GenBank with accession number HM347447 and DQ307383, respectively.

The general deduced protein structures of 1Dx3t, 1Dx4t, 1Dx2 and 1Dx5 subunits and their comparison were shown in Fig. 2. Four x-type subunits had similar structural characteristics, including a signal peptide of 21 amino acids, a non-repetitive N-terminal domain of 89 amino acid sequences, a long repetitive domain, and a non-repetitive domain of 42 amino acids in the C-terminal. Comparing with 1Dx5 subunit, 1Dx3t had a six amino acid insertion at 146–151 (SGQGQQ). The 1Dx4t subunit had a nine amino acid deletion compared with the 1Dx3t subunit. Therefore, both subunits were more similar to 1Dx5 than 1Dx2 subunit.

The coding sequences of 1Dx3t and 1Dx4t genes were compared with 9 other x-type subunit genes that are previously characterized or deposited in GenBank, and the detected SNPs and InDels were showed in Table 1. A total of 18 SNPs were identified over different positions, and 8 and 12 SNPs were present in 1Dx3t and 1Dx4t genes, respectively. The majority of the changes in nucleotides at the SNP sites were due to A-G or C-T transitions (accounting for 72%). One deletion at the position of 2132–2159 was detected in the 1Dx4t. Of the 18 SNPs, 9 were non-synonymous SNP, viz. position 232 (valine → methionine), 256 (phenylalanine → leucine), 443 (glutamine → arginine), 971 (praline → leucine), 1729 (threonine → alanine), 1810 (glycine → arginine), 2042 (praline → leucine), 2150 (valine → glycine) and 2171 (alanine → glutamic acid).

Heterologous expression analysis

The cloned ORF without signal peptide of 1Dx3t and 1Dx4t genes were ligated into the expression vector pET-28a and pET-30a (Novagen), transformed into E. coli BL21 (DE3) plysS cells. The positive hybrid plasmid expressed a new protein by the induction of IPTG. As shown in Fig. 3a, the expressed protein had a slightly faster mobility than 1Dx3t in T67 (indicated by the arrow). Due to the sequences of his-tag were incorporated into the end of the expressed genes, the hybrid protein was about 2186 Da larger than the quondam protein. Figure 3b clearly showed that the band indicated by
arrow co-migrated with the authentic 1Dx4’ subunit from seed endosperm of T132.

Secondary structure analysis of 1Dx3’ and 1Dx4’ subunits

Protein secondary structure is the base of high complex spatial conformation. In this study, PSIPRED method was used to predict secondary structures of the 1Dx3’ and 1Dx4’ subunits and the results were shown in Table 2. Previous reports indicated that β-strand was generally considered to endow the protein with high elasticity and to improve the capability to resist distortion, and therefore the content of β-strands in HMW-GS and LMW-GS might have positive effects on dough quality (Tatham et al. 1987, 1990; Jiang et al. 2008). HMW-GS contain β-turns as a structural motif within the elastomeric domains, which enable the proteins to withstand significant deformations without rupture before returning to their original state after the stress is removed (Tatham et al. 1985). Comparing to the 1Dx5 subunit, which was considered to have a positive effect on bread-making quality, the percent of β-strands in both 1Dx3’ and 1Dx4’ subunits were higher than those in 1Dx5 subunit, and 1Dx3’ had much higher α-helix content (37.55%). These results suggested that both subunits were likely to confer good dough properties.

Fig. 2. Comparison of the deduced amino acid sequences of four Glu-1Dx subunits. The four subunits have an identical primary structure: signal peptide, N-terminal domain, repetitive domain and C-terminal domain were indicated.
To investigate the phylogenetic relationships among HMW-GS gene family from different genomes, the complete coding sequences of 26 genes were used to construct a phylogenetic tree, which included the 1Dx3t and 1Dx4t genes obtained in this study and other 24 genes from GenBank, namely 1Dx5" (DQ681076), 1Dx5.1" (DQ681077), 1Dx5 (X12928), 1Dx2 (X03346), 1Dx2.1" (AF480486), 1Dx1.5" (AY594355), 1Dx1.6" (DQ857243), 1Dx2.1 (AY517724), 1Dx2.2 (AY159367), 1Dx2.2" (AJ893508), 1Dx5.2" (DQ307384), 1Ax1 (X61009), 1Ax2" (M22208), 1Bx7 (X13927), 1Bx7OE (DQ119142), 1Bx23 (AY553933), 1Bx20 (AJ437000), 1Bx13 (EF540764), 1Bx14 (AY367771), 1Bx17 (JC2099), 1By8 (AY245797), 1By9 (X61026), 1Dy10 (X12929) and 1Dy12 (X03041). The results were showed in Fig. 4.

The phylogenetic tree was apparently clustered into two clear branches, corresponding to x-type and y-type subunit genes, respectively. In the x-type group, three subgroups were clearly separated by the A, B and D genomes. The x-type genes encoded by A and D genomes were clustered into a major branch. In particular, the 1Dx3t and 1Dx5 genes were clustered to a subgroup, suggesting their similar structural features and close phylogenetic relationship.

**DISCUSSION**

Aegilops tauschii contains many useful genes for bread wheat cultivar improvement, and therefore in a long time it serves as an important genetic resource. Until now, different resistance and tolerance genes as well as glutenin genes related to quality have been found in Ae. tauschii, and have been introgressed and utilized in wheat cultivar improvement programs via hybridization between bread wheats or tetraploid species and Ae. tauschii (Eastwood et al. 1991; Schachtman et al. 1991; Cox et al. 1995; Hsam et al. 2001; Huang and Gill 2001; Zhang et al. 2009). It has been confirmed that the HMW glutenin genes from the Ae. tauschii species have significant influence on bread-making properties in synthetic hexaploid wheats (Lagudah et al. 1987; Hsam et al. 2001), and some of the resulting lines exhibited shorter mixing time and improved milling and baking characteristics when compared to parental hexaploid lines (Tilley et al. 2000).

In recent years, wheat genetic transformation has made a great progress, especially for quality improvement (Cheng et al. 1997; Alvarez et al. 2001; Sahrawat et al. 2003; Shewry et al. 2006). Several HMW-GS genes, such as 1Dx5 + 1Dy10, 1Ax1, have been successfully transformed through shoot gun, and the transgenic lines performed much improved quality properties (Barro et al. 2003; Altpeter et al. 2004). However,
Both genes displayed high similarity with good quality subunit gene 1Dx5, implying their potential values for wheat quality improvement. Studies have shown that the strength of gluten is highly correlated with different combinations of the HMW-GS. For instance, 1Dx5 + 1Dy10 subunits associated with good bread-making quality while 1Dx2 + 1Dy12 subunits as well as 1Bx20 associated with poor bread-making quality (Redaelli et al. 1997; Shehry et al. 2003).

It was believed that β-turns in HMW-GS play a positive role in dough quality (Tatham et al. 1985). The central repeated domain of HMW-GS, mainly resulting in variations in size difference of HMW-GS and adopting a beta-spiral structure confers elasticity to the protein molecule (Gianibelli et al. 2001a). It is clear that the cysteine residues presented in N-termini (normally 3) and C-termini (only one) form intermolecular disulphide bonds to create very large polymers, which are important to provide the viscoelastic properties to dough. An extra cysteine residue present in 1Dx5 subunit may be responsible for its good quality property (Anderson et al. 1989). Subunit 1Bx20 displays two cysteine residues substitutions by tyrosine in the N-terminal domain when compared with subunit 1Bx7, which 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 (Shehry et al. 2003). Flavel et al. (1989) found that subunit 1Dy10 has a higher proportion of repeats of the consensus type than subunit 1Dy12 and postulated that this produces a more regular pattern of repetitive beta-turns in the polymers. Gupta et al. (1996)

Table 2. The secondary structure prediction of the eight deduced HMW-GSs.

<table>
<thead>
<tr>
<th>HMW-GS</th>
<th>Structure motifs</th>
<th>Content %</th>
<th>Total</th>
<th>N-terminal domain</th>
<th>Repetitive domain</th>
<th>C-terminal domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>1Dy10</td>
<td>α-helix</td>
<td>13.40</td>
<td>6</td>
<td>0.5</td>
<td>4.5</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>β-strand</td>
<td>1.12</td>
<td>2</td>
<td>–</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1Dy12</td>
<td>α-helix</td>
<td>16.80</td>
<td>9</td>
<td>0.5</td>
<td>6.5</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>β-strand</td>
<td>0.31</td>
<td>1</td>
<td>–</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>1Dx2</td>
<td>α-helix</td>
<td>4.65</td>
<td>4</td>
<td>2</td>
<td>–</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>β-strand</td>
<td>1.47</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1Dx2.2</td>
<td>α-helix</td>
<td>32.84</td>
<td>16</td>
<td>2</td>
<td>14</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>β-strand</td>
<td>0.74</td>
<td>2</td>
<td>2</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>1Dx3'</td>
<td>α-helix</td>
<td>37.55</td>
<td>13</td>
<td>2.5</td>
<td>11.5</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>β-strand</td>
<td>1.60</td>
<td>3</td>
<td>2</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>1Dx4'</td>
<td>α-helix</td>
<td>5.70</td>
<td>7</td>
<td>4</td>
<td>–</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>β-strand</td>
<td>1.82</td>
<td>4</td>
<td>3</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>1Dx5</td>
<td>α-helix</td>
<td>6.60</td>
<td>8</td>
<td>4</td>
<td>–</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>β-strand</td>
<td>1.47</td>
<td>3</td>
<td>2</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>1Dx5.2'</td>
<td>α-helix</td>
<td>36.10</td>
<td>14</td>
<td>3</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>β-strand</td>
<td>1.84</td>
<td>4</td>
<td>2</td>
<td>–</td>
<td>2</td>
</tr>
</tbody>
</table>
found that variation in accumulation rate could lead to their quality differences because 1Dx5 + 1Dy10 accumulate larger polymers more quickly than 1Dx2 + 1Dy12. It could be concluded that allelic variations resulting in quality differences at Glu-1 loci are mainly attributed to structural characteristics of HMW glutenin subunits.

Interestingly, 1Dx3\textsuperscript{t} subunit had a six amino acid insertion at 146–151 (SGQQGQQ) and the deduced mature protein possessed 845 amino acid residues, being nine more than the 1Dx5 subunit. In addition, the number of β-strands in both 1Dx3\textsuperscript{t} and 1Dx4\textsuperscript{t} subunits was higher than that in 1Dx5 subunit. Particularly, 1Dx3\textsuperscript{t} had much higher α-helix content (37.55%) than 1Dx5 (6.60%). These characteristics suggested that both subunits, especially 1Dx3\textsuperscript{t}, may have positive effects on dough properties and consequently possess high potential for wheat quality improvement. Compared to other cloned genes, 1Dx3\textsuperscript{t} represents the largest x-type HMW-GS gene with high glutamine content among Glu-DI-1 alleles. There is evident from mixing studies that larger HMW-GS as well as LMW-GS rich in glutamines have a greater positive effect on dough strength than smaller subunits (Békes et al. 1995; Li et al. 2008). Theoretically, the high glutamine content through forming more hydrogen bonds can stabilize the polymeric structure of glutenin (Gilbert et al. 2000), and the length of the repetitive region in the HMW-GS may play a role in dough visco-elastic properties (Belton 1999). Therefore, the longer repetitive domain and higher α-helix and β-strand contents of 1Dx3\textsuperscript{t} subunit make it potentially a valuable candidate gene for wheat quality improvement.

Acknowledgements – This research was financially supported by grants from the National Natural Science Foundation of China (30830072, 30771334), the Chinese Ministry of Science and Technology (2009CB118303, 12006AA10Z186) and the Key Developmental Project of Science and Technology, Beijing Municipal Commission of Education (KZZ200910028003).

REFERENCES

Altpeter, F., Popelka, J. C. and Wieser, H. 2004. Stable expression of 1Dx5 and 1Dy10 high-molecular-weight glutenin subunit genes in transgenic rye drastically increases the polymeric glutelin fraction in rye flour. – Plant Mol. Biol. 54: 783–792.


Kihara, H. 1944. Discovery of the DD analyser, one of the ancestors of Triticum vulgare. – Agric Horticult. 19: 13–14.


Shewry, P. R., Gilbert, S. M., Savage, A. W. et al. 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.

