Molecular cloning and analysis of a novel HMW-GS gene Glu1-St1.5 from Elymus sibiricus in Qinghai-Tibetan Plateau

LIU Jin1,2,3#, ZHANG Bo1#, LIU Bao-long1, LI Hong-qin1,2, WANG Xin1,2, LIU Qi1,2, WEI Le1,2
LIU Deng-cai1,4 and ZHANG Huai-gang1

1Key Laboratory of Adaptation and Evolution of Plateau Biota, Northwest Institute of Plateau Biology Chinese Academy of Sciences, Xining 810008, China.
2Graduate University, Chinese Academy of Sciences, Beijing 100049, China.
3Yunnan Institute of Tropical Crops, Xishuangbanna 666100, China.
4Triticeae Research Institute, Sichuan Agricultural University, Wenjiang, Chengdu 611130, Sichuan, China.

Accepted 30 May, 2012

In the present study, a HMW-GS allele Glu1-St1.5 was isolated and characterized from Elymus sibiricus. Nucleotide sequence analysis indicated that the St1.5 shared the highest similarity with St1.3 from Elymus canadensis, but the “TA” insertion mutation at 15 to 16 bp in signal peptide sequence made it an inactive gene, because three premature stop codons were generated by “TA” insertion in signal peptide region, N-terminal region and C-terminal region, respectively. To identify the type of St1.5, the inserted “TA” was eliminated factitiously and the deduced amino acid sequence was analyzed, the result showed that St1.5 was a y-type HMW-GS pseudo gene. Phylogenetic tree demonstrated that St1.5 had the closest relationship to the St1 allele of Pseudoroegneria stipifolia, which indicated that St1.5 was a novel gene of St Genome. These results provide insights into the evolutionary biology of Glu1-St1.5 and other HMW-GS genes.

Key words: Glu1-1St1.5, HMW-GS, Elymus sibiricus.

INTRODUCTION

The end-use quality of wheat (Triticum aestivum L.) flour is largely determined by the composition of storage proteins in seed endosperm (Payne, 1987; Gianibelli, 2001a, b). The high molecular weight glutenin subunit (HMW-GS) is one of most important storage proteins in seed endosperm of wheat and related species and accounting for approximately 10% of the total proteins in wheat seed (Shewry et al., 1995). In hexaploid wheat, HMW-GS is encoded by the Glu-A1, Glu-B1 and Glu-D1 locus on the long arm of homologous chromosomes 1A, 1B and 1D, respectively (Shewry et al., 1992). There are two tightly linked genes encoding a larger x-type (80-88 kDa) and a smaller y-type (67 to 73 kDa) subunit at each locus. But only three to five HMW-GSs could be detected in hexaploid wheat seed by conventional SDS-PAGE analysis, because the y-type gene of Glu-A1 locus is often silenced and the x-type gene of Glu-A1 locus and y-type gene of Glu-B1 locus are silence occasionally (Harberd et al., 1986; Payne, 1987; Shewry et al., 1992, 2001, 2002).

Elymus L. is the largest genus in the tribe Triticeae with about 150 species distributed in most places of the world, especially in some high altitude, cold and arid areas. (Lu, 1993; Ma et al., 2009). As the type species of the genus Elymus, Elymus sibiricus L. is a perennial, self-pollinating and allotetraploid grass indigenous to Northern Asia, possessing the St and H genome (Dewey 1974, 1984). Its geographic distribution extends from Sweden to Japan and even to parts of Alaska and Canada (Bowden and
Figure 1. The SDS-PAGE characters of HMW-GS in *Elymus sibiricus* (lane 1) and the common wheat line Chinese Spring (lane M) with HMW-GS composition 1Dx2, 1Bx7, 1By8 and 1Dy12 was used as control. The *Elymus sibiricus* possesses two glutenin subunits, both of them moved slower than 1Dy12 of the control.

Cody, 1961) and then extends southerly to Qinghai-Tibet Plateau, which is the highest plateau in the world. *E. sibiricus* usually grows on wet meadows, riverside sands and among open forest or shrubs. In the subalpine meadows with less than 4000 meter altitude in Qinghai-Tibet Plateau, *E. sibiricus* usually serves as an important forage species.

Orthologous HMW glutenin subunits have been found in many Triticeae grasses including various *Aegilops* species and rye (De Bustos et al., 2001; De Bustos and Jouve, 2003; Liu et al., 2003; Wan et al., 2002; William et al., 1993). Up to now, no research work has been reported on HMW-GS gene isolation and characterization in *Elymus sibiricus*.

In the present work, a novel HMW-GS gene *Glu1-St1.5* was isolated and characterized from *Elymus sibiricus*. The gene structure was analyzed by comparison with HMW-GS genes in other wheat related wild species. The results will help us understanding the evolutionary relationship between *Glu1-St1.5* and other HMW-GS genes.

### MATERIALS AND METHODS

#### Plant material

SDS-PAGE analysis of 105 individual plants of *Elymus* L. distributed in 56 different longitude and dimensionality using seed protein extracts, revealed the presence of two putative HMW glutenin subunits in one of the *Elymus sibiricus* collected from Qinghai-Tibetan Plateau (Figure 1). The longitude and dimensionality of collected site was E 100.913663 and N 36.372140, respectively and the altitudes was 3535.16 m. Therefore the *Elymus sibiricus* was chosen as the material for the follow-up experiment. SDS-PAGE analysis was conducted according to Wan et al. (2002).

#### Genomic DNA extraction and PCR amplification

Seeds of *Elymus sibiricus* were germinated under the dark at room temperature for two weeks. Genome DNA was extracted from young leaves by conventional CTAB method with some modifications, ethanol precipitated, re-suspended in 50 ul of sterile distilled water and then stored at -20°C.

In order to amplify HMW-GS genes by genomic PCR, a pair of degenerated primers P1 (5’-ATGGCTAAGCGGC/TTA/GGTCCTCTTTG-3’) and P2 (5’-CTATCACTGGCTA/GGCCGACAATGCG-3’), which could amplify the complete ORFs of HMW-GS genes of wheat related species were used (Liu et al., 2003; Li et al., 2008). Primer P1 contains the start codon of the HMW-GS ORF whereas primer P2 contains the two tandem stop codons that are conserved in the HMW-GS ORFs characterized so far (Liu et al., 2003). PCR were carried out using ABI Gene Amp PCR System 9700, the high fidelity polymerase HiFi Taq with GC buffer (TransGen Biotech, Beijing, China) were used. Different annealing temperatures were tested and 65°C appeared to be optimal for PCR amplification of the complete ORFs of HMW-GS genes. The PCR reaction programmed at 94°C for 5 min to denature the DNA, followed by 30 cycles each with a 50 s denaturing step at 94°C, a 1 min annealing step at 65°C and a 1.5 min extension step at 72°C. The final extension step was for 10 min, followed by a 4°C soak step. PCR products were separated by 1% agarose gel electro-phoresis in TAE buffer.

#### DNA cloning, sequencing and phylogenetic analysis

The PCR product of expected size was extracted from agarose gel; the fragment was ligated into the pEASY-T1 vector (TransGen Biotech, Beijing, China) and transformed into competent cells of *Escherichia coli* DH-10B strain. The positive clones were identified using blue/white screening and colony PCR. The complete sequence was acquired by sequencing three different positive clones; the sequencing was performed by AUGCT Biotechnology Company (Beijing, China). Nucleotide sequences assemble and phylogenetic analysis were carried out by Invitrogen Advance 10 program.

### RESULTS

#### Isolation and sequencing of HMW-GS gene *Glu1-St1.5*

PCR amplification was carried out using the two degenerated primer P1 and P2 which could amplify the complete ORFs of HMW-GS genes in all of the wheat
related species. The agarose electro-phoresis analysis of the PCR products showed that only one fragment have been amplified, the size of the fragment was just approximate 1300bp (Figure 2). Previous studies showed that the DNA sequence length of 1Dy12 in the control Chinese Spring was 1977 bp (Cunsolo et al., 2003), so the PCR fragment was not correspondent with the two subunit in SDS-PAGE, the PCR fragment may be another HMW-GS gene which have not been detected by SDS-PAGE.

To elucidate whether the PCR fragment in Figure 2 was a HMW-GS gene and whether it could translate into glutenin subunit, the fragment was cloned into the pEASY-T1 vector and sequenced. The sequence was deposited in GenBank under the accession number HM804857.

Characterization of St1.5 gene

BLAST analysis in NCBI revealed that the DNA sequence of St1.5 showed very high homology to that of previously cloned HMW-GS genes in all kinds of wheat related species, Glu1-St1.3 in Elymus canadensis was the highest one, coefficient of similarity between these two genes was 94.50%, so the St1.5 was a family member of high molecular weight glutenin subunit genes.

Compared the DNA sequences of St1.5 and St1.3, 29 different sites were observed, two sites were nucleotide insertion, others were nucleotide substitution or deletion mutation (boxed in Figure 3). The most crucial difference of these two sequences was the “TA” insertion mutation at 15 to 16 bp downstream the start codons of St1.5, which caused three stop codons at 34 to 36 bp, 397 to 399 bp, 1237 to 1239 bp downstream the start codons, these stop codons made St1.5 could not be translated into protein normally. The other nucleotide insertion occurred at the central repetitive domain, but it did not caused any mutation of the reading frame of St1.5.

Regardless of its inactivation, elucidating whether St1.5 had the typical HMW-GS structure before the “TA” insertion in its evolutionary processing was much more valuable to us. So the inserted “TA” of St1.5 was eliminated factitiously, the deduced amino acid sequences of St1.5 (TA) and St1.3 were compared.

The deduced amino acid sequences of both St1.5 (TA) and St1.3 shared the typical HMW-GS structure, including the signal peptides, N-terminal region, central repetitive domain and C-terminal region (Figure 4). The signal peptides, N-terminal region and C-terminal region were conservative, but the central repetitive domain which consisted of nonapeptide (GYYPTSP/LQQ) and hexapeptide (consensus PGQGQQ) was semi-conservative. There were eleven different sites between the deduced amino acid sequences of St1.5 (TA) and St1.3, seven of them were amino acid substitution, and the other four were amino acid insertion or deletion.

Both of St1.5 (TA) and St1.3 had six cysteine residues, five in N-terminal region and the other one in C-terminal region. Previous studies showed that x-type HMW-GS has three conservative cysteine residues in its N-terminal region and one in its C-terminal region, but the y-type HMW-GS has five conservative cysteine residues in its N-terminal region and one in its C-terminal region, both of them have a semi-conservative cysteine residue the central repetitive domain (Shewry and Tatham, 1997; Belton, 1999; Veraverbek and Delcour, 2002; Wieser, 2007). So it was obvious that St1.5 was a y-type HMW-GS gene.

Phylogenetic analysis of HMW-GS genes

Sequence comparisons among different HMW-GS alleles could provide useful information relevant to the evolutionary relationship among them. There are nearly two hundred HMW-GS genes from more than thirty different wheat related species which belong to 12 different genera have been cloned till now (www.ncbi.nlm.nih.gov), it was difficult to analyze the phylogenetic relationship of all the cloned genes, so 23 x-
Figure 3. Nucleotide sequences of the novel inactive HMW-GS gene St1.5 compared with the ORFs of another active gene St1.3 which came from *Elymus canadensis*, the GenBank accession number for St1.5 and St 1.3 were HM804857 and GU998882.1, respectively. Solid boxes point out the sites of difference among the two nucleotide sequences, the start codons and stop codons are marked by bold line, the “TA” insertion of St1.5 was marked by a bold box. Broken bold line point out the stop codons induced by the “TA” insertion mutation in St1.5.

Figure 4. The deduced amino acid sequences of St1.5 while eliminated the inserted “TA” factitiously, compared with the deduced amino acid sequences of the ORFs of St1.3. Solid boxes point out the sites of difference among the two sequences, the signal peptides are underlined with solid line, while the N-terminal region and the C-terminal region are underlined with double lines and triplicate lines, respectively. The central repetitive domain which consisted of nonapeptide (GYYPTSP/LQQ) and hexapeptide (consensus PGQGQQ) are indicated by bold line and cysteine residues are indicated by black trigones.

and y-type genes from 12 different genera of *Triticeae* were chosen to analyze the phylogenetic position of the novel gene St1.5 in this study.

As shown in Figure 5, the phylogenetic tree was obviously divided into two halves. The y-type subunit genes were clustered into one subclass and the x-type ones were clustered into another. The newly cloned Glu-St1.5 gene from *Elymus sibiricus* was a y-type gene.
St1.5 of *Elymus sibiricus* and St1 of *Pseudoroegneria stipifolia* were clustered into one group inferring their closest relationship. It is known that *Pseudoroegneria stipifolia* was one of wild wheat related species with the genome of St (2n = 2x = 14), so we speculated St1.5 was a novel HMW-GS allele in St Genome.

**DISCUSSION**

In the present study, a novel HMW-GS gene *Glu1-St1.5* was isolated and characterized, the insertion mutation of “TA” at signal region resulted three stop codons in N-terminal region, central repetitive domain and C-terminal region, respectively, which made St1.5 gene not being translated into protein. Previous studies showed that not all HMW-GS genes could be translated, the representative case is *Glu1-Ay* in common wheat (Harberd et al., 1987; Halford et al., 1989; Sun et al., 2004) and the main reason of HMW-GS gene silence was the insertion transposon element or the presence of stop codon (Xiang et al., 2010). So the St1.5 represented a new silence mechanism compared with the previous published inactive HMW-GS genes.

The molecular size of x-type HMW-GS genes was bigger than y-type genes, but the cysteine number of y-type was more than x-type (Shewry and Tatham, 1997; Belton, 1999; Veraverbek and Delcour, 2002; Wieser, 2007) and they were useful significant characteristics to distinguish the x- and y-type HMW-GS alleles. Wheat flour quality was affected by the different composition of HMW-GS alleles because cysteine number affected the volume and quantity of glutenin macro polymer, which positively correlated with the visco-elastic property of wheat dough, so the y-type genes were more important than x-type for wheat flour quality. Although the St1.5 was an inactive y-type gene, but more research should been carried out in other y-type alleles of *Elymus sibiricus* and if there existed active y-type allele, it could be used as a target gene to modified the quality of wheat flour by transgenic approach.

The evolutionary origin of different HMW-GS alleles from different Triticeae grasses and genome was an interesting issue for scientists. The St Genome of *Elymus L.* was probably derived from *Pseudoroegneria* (Nevski) A. Love, which only has the St Genome. The phylogenetic tree of 24 HMW-GS genes from different genome showed that the alleles from St (*Elymus sibiricus*, *Elymus canadensis*, *Pseudoroegneria stipifolia*, *Thinopyrum intermedium*) have closer relationship with those from V genome (*Dasypyrum breviaristatum*), R genome (*Secale cereale*) and D genome (*Triticum aestivum 1Dy*) than those from P genome (*Agropyron cristatum*), K genome (*Crithopsis delileana*), B genome (*Triticum aestivum 1By*), E* genome (*Lophopyrum elongatum*) and Ns genome (*Leymus mollis*). Interestingly, the “TA” insertion resulted gene silence was observed in the present study. The silence mechanism of St1.5 was different from the previous studies. Further investigations are underway to examine if the above feature is common to the HMW glutenin subunits encoded by other species in *Elymus L.* from Qinghai-Tibetan Plateau.

**ACKNOWLEDGEMENTS**

The work was supported by grants from Project of...
Knowledge Innovation Engineering, CAS (KSCX3-EW-N02), the Ministry of Agriculture of China (2008ZX08002-004) and National Natural Science Foundation of China (31101140).

REFERENCES


