SHORT COMMUNICATION

Characterization of y-type high-molecular-weight glutenins in tetraploid species of *Leymus*

Yanqi Sun · Zongjun Pu · Shoufen Dai · Xiaoxue Pu · Dengcai Liu · Bihua Wu · Xiujin Lan · Yuming Wei · Youliang Zheng · Zehong Yan

Abstract Three y-type high-molecular-weight (HMW) glutenin gene open reading frames (ORFs), *Chiy1*, *Chiy2*, and *Racy*, were isolated and characterized from *Leymus chinensis* PI499516 and *Leymus racemosus* ssp. *racemosus* W623305. They shared an extra glutamine in the N-terminal and LAAQLPAMCRL peptides in the C-terminal with x-type HMW glutenins but had different N-terminal lengths. Like other y-type HMW glutenins, *Chiy2* and *Racy* had 104 (or 105) amino acid (aa) residues at the N-terminal and started with EGEASR, whereas *Chiy1* had 99 aa in this domain and started with QLQCER because of the deletion of EGEASR. Five other y-type glutenins, including those from *Elymus ciliaris*, *Pseudoroegneria libanotica*, and *Leymus mollis*, were similar to *Chiy1*. The ORF of *Chiy2* was probably not expressed. The ORFs of both *Chiy1* and *Racy* were expressed in bacteria. The maximum likelihood phylogenic tree based on the signal peptide and N-terminal and C-terminal aa residues revealed two clades of y-type HMW glutenins in Triticeae; the first contained *Ay*, *By*, *Cy*, *Dy*, *Eey*, *Gy*, *Ky*, *Ry*, *Tay*, and *Uy*, while the second clade contained the remaining y types, including those from *Leymus*. Within the second clade, HMW glutenins lacking the EGEASR peptide formed a subclade. These y-type HMW glutenins in *Leymus* could not be targeted to the Xm or Ns genome.

Keywords Gene sequencing · *Glu-1* · *Leymus* · NsXm genome

Introduction

High-molecular-weight (HMW) glutenins are the main storage proteins in the endosperm of Triticeae grasses. Their biochemical function is to provide carbon energy sources for seed germination. They also contribute to the rheological and dough-baking properties of wheat flours. In bread wheat, HMW glutenins are encoded by the genes at the *Glu-1* complex loci on the long arms of homologous group 1 chromosomes. At each locus, two tightly linked paralogous *x* and *y* genes encode an x subunit with larger mass and a smaller y-type subunit (Payne et al. 1980).

Wheat HMW glutenins share four similar structural domains, including a signal peptide that is removed in the mature protein, highly conserved N-terminal and C-terminal, and an intermediate repetitive domain (Shewry et al. 2003). In their N-terminal, the x types share 81–86 amino acid (aa) residues and 3 cysteines, while the y types share 104 aa and 5 cysteines. Both the x and y types share 21 aa residues in the signal peptide and 42 aa in the C-terminal. In the larger central domains, there are three repeating units of hexapeptides (PGQGQQ), nonapeptides (GYYPTSPQQ), and tripeptides (GQQ). The hexapeptide and nonapeptide units exist in both x and y types, whereas the tripeptides only occur in x types (Shewry et al. 2003). Isolation and characterization of new HMW glutenin variants from wild relatives of wheat revealed novel modifications in all domains (Guo et al. 2005; Yan et al. 2006; Wang et al. 2006, 2012; Pistón et al. 2007; Liu et al. 2010, 2012a). For

Y. Sun · S. Dai · X. Pu · D. Liu · B. Wu · X. Lan · Y. Wei · Y. Zheng · Z. Yan (✉)
Triticeae Research Institute, Sichuan Agricultural University, No. 211, Huimin Road, Wenjiang Region, Chengdu City 611130, Sichuan, People's Republic of China
e-mail: zhyan104@163.com

Z. Pu
Crop Research Institute, Sichuan Academy of Agricultural Sciences, Chengdu 610066, Sichuan, People's Republic of China

D. Liu
Key Laboratory of Adaptation and Evolution of Plateau Biota, Northwest Institute of Plateau Biology, Chinese Academy of Sciences, Xining 810001, Qinghai, People's Republic of China
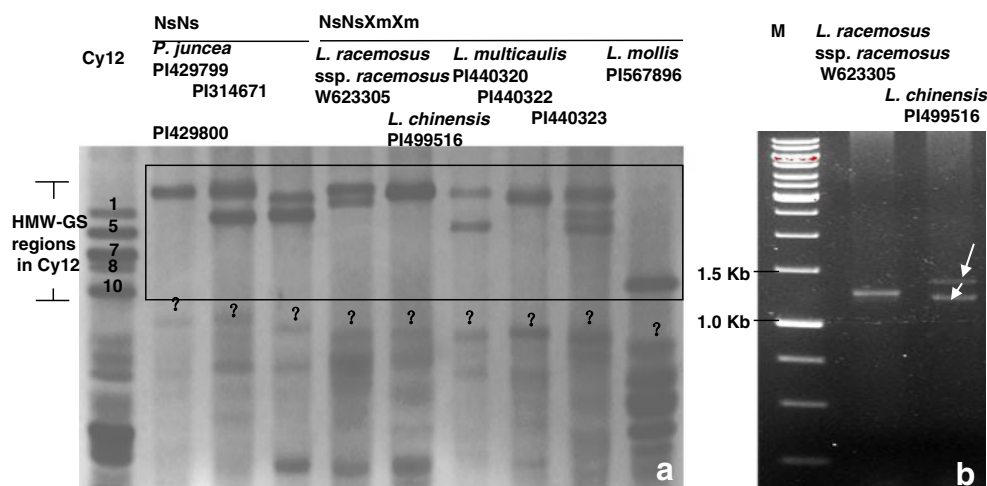
Fig. 1 SDS-PAGE (**a**) and PCR (**b**) patterns of HMW-GSs in *Leymus* species. **a** Using the HMW-GSs in wheat cv. Cy12 as reference, the HMW-GSs encoded by *P. juncea* and *Leymus* species with similar electrophoretic mobilities with those of wheat were *boxed*. Some of the protein bands (*next to the box region with a question mark*), with faster electrophoretic mobilities than those of wheat were putative HMW-GSs.

b One and two DNA fragments containing the ORFs of HMW-GSs from *L. racemosus* ssp. *racemosus* W623305 and *L. chinensis* PI499516 were amplified using degenerated PCR primers. The DNA bands in W623305 and PI499516 corresponded to the ORFs of *Racy*, *Chiy1* (*long arrow*), and *Chiy2* (*short arrow*)

example, some wild wheat species contained an extra glutamine in the N-terminal of y-type HMW glutenin, resulting in 105 aa rather than the 104 in wheat (Yan et al. 2006; Wang et al. 2006, 2012). Modifications of the extreme N-terminal aa were also observed (Wang et al. 2012). Thus, new structural characters were found in the HMW glutenins of Triticeae.

Leymus Hochst. is a genus of ~30 species of polyploid perennial grasses distributed from Eurasia to North America (Dewey 1984). All of the species share two basic genomes: Ns, from *Psathyrostachys*, and Xm, of unknown origin (Wang et al. 1994; Zhang and Dvořák 1991). Chromosome numbers in *Leymus* range from $2n=4x=28$ to $2n=14x=98$. Thus far, the HMW glutenin genes in the Ns and Xm genomes have not been documented. Here, we describe the isolation and characterization of HMW glutenins from tetraploid species of *Leymus*. The results extend our understanding of sequence variation and evolutionary relationships among HMW glutenins in *Leymus* and other Triticeae.

Materials and methods

Plant materials and DNA isolation

Nine accessions belonging to four tetraploid *Leymus* species ($2n=4x=28$, NsNsXmXm), including *Leymus racemosus* ssp. *racemosus* (accession number W623305), *Leymus chinensis* (PI499516), *Leymus multicaulis* (PI440320, PI440322, and PI440323), and *Leymus mollis* (PI567896) as well as *Psathyrostachys juncea* ($2n=2x=$NsNs; PI429799, PI314671, and PI429800), were used for sodium dodecyl
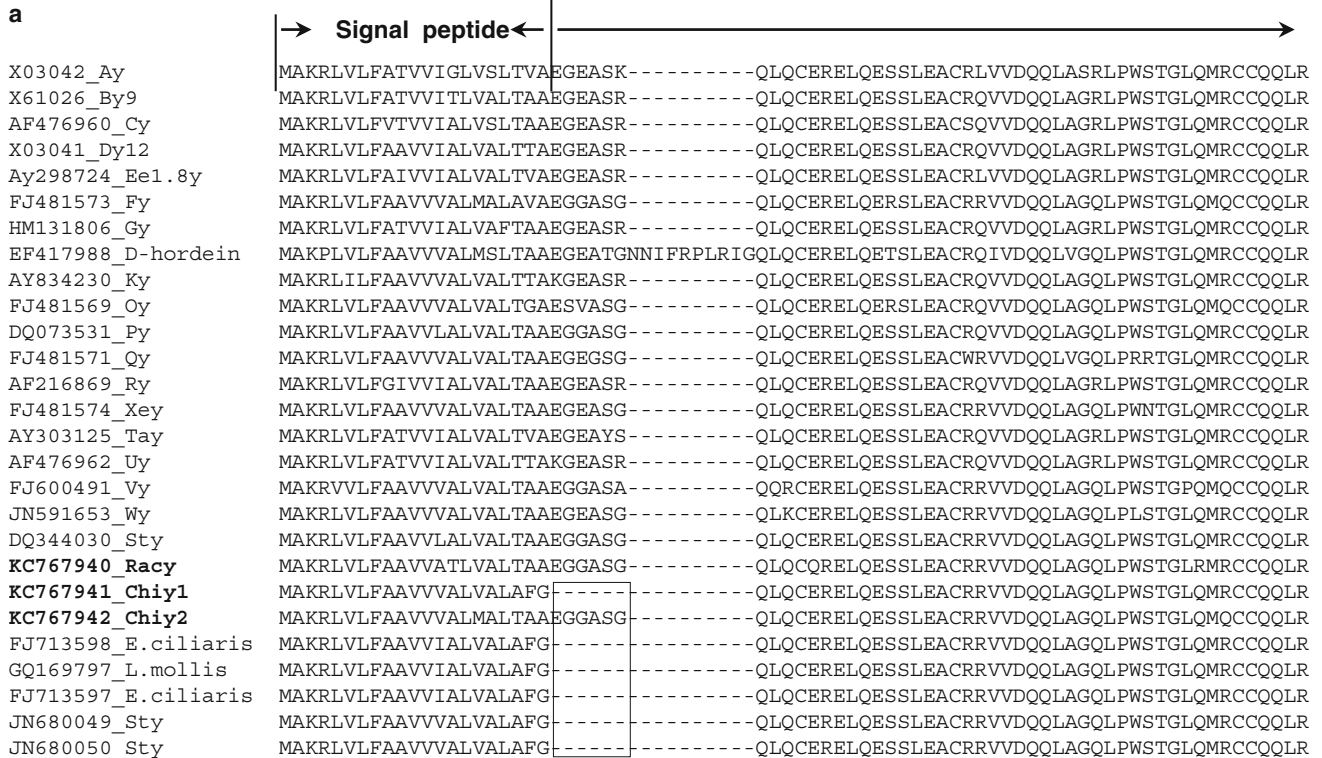
sulfate polyacrylamide gel electrophoresis (SDS-PAGE) separation of HMW glutenin subunits (HMW-GS). Seeds were provided by the USDA-ARS germplasm bank (http://www.ars-grin.gov). Common wheat cultivars (cvs.) Chuanyu12 (Cy12) and Chinese spring (CS), with HMW-GS 1, 7+8 and 5+10 and HMW-GS 7+8 and 2+12, respectively, were used as references to estimate the electrophoretic mobility of the HMW-GS in *Leymus*. Genomic DNA was extracted from mixed young seedlings of PI499516 and W623305 using the CTAB method (Doyle and Doyle 1990).

Fig. 2 Sequence comparison (**a**) and diagram (**b**) of y-type HMW-GSs ▶ from *Leymus* species with orthologous genes from Triticeae species. **a** Sequence alignment shows the typical structure of y types in two *Leymus* species with those of homologous Triticeae species. The deleted EGEASER peptide at the start of the N-terminal, the extra glutamine at its middle, and the LAAQLPAMCRL peptide in the C-terminals were *boxed*. The lacking aa were shown by *short dashes*. The representative y types from Triticeae species were *Ay* (from *Triticum aestivum*, GenBank accession number X03042), *By9* (*T. aestivum*, X61026), *Dy12* (*T. aestivum*, X03041), *Cy* (*Aegilops markgrafii*, AF476960), *Uy* (*Aegilops umbellulata*, AF476962), *Py* (*Agropyron cristatum*, DQ073531), *Wy* (*Australopyrum retrofractum*, JN591653), *Ky* (*C. delileana*, AY834230), *Vy* (*Dasypyrum villosum*, FJ600491), *Fy* (*Eremopyrum distans*, FJ481573), *Xey* (*Eremopyrum triticeum*, FJ481574), *Oy* (*H. persica*, FJ481569), *Qy* (*Heteranthelium piliferum*, FJ481571), D-hordein (*Hordeum chilense*, EF417988), *Ee1.8y* (*Lophopyrum elongatum*, Ay298724), *Sty* (*P. stipifolia*, DQ344030 and *Pseudoroegneria tauri* ssp. *libanotica*, JN680049 and JN680050), *Tay* (*T. crinitum* AY303125), *Ry* (*T. aestivum* cv. 7841 1R/1D substitution line, AF216869), and *Gy* (*Triticum timopheevii* ssp. *araraticum*, HM131806), respectively. Three y types in this study were *Chiy1* (*L. chinensis*, KC767941), *Chiy2* (*L. chinensis*, KC767942), and *Racy* (*L. racemosus* ssp. *racemosus*, KC767940). **b** Diagram of y-type HMW glutenins with those homologous in Triticeae showing typical structure of y types from *Leymus* species in all domains

Extraction of HMW-GS and SDS-PAGE

HMW-GSs were extracted from 4 mg of seed endosperms using 100 μL extraction buffer containing 0.0625 mM/L Tris–HCl (pH 6.8), 2 % w/v SDS, 10 % glycerol, 5 % v/v β-mercaptoethanol, and 0.002 % w/v bromophenol blue. Mixtures were incubated at room temperature for 3 h with occasional vortexing. Extractions were denatured in boiling water for 5 min and centrifuged at 10,000 rpm for 5 min. Five microliters of supernatant was used to
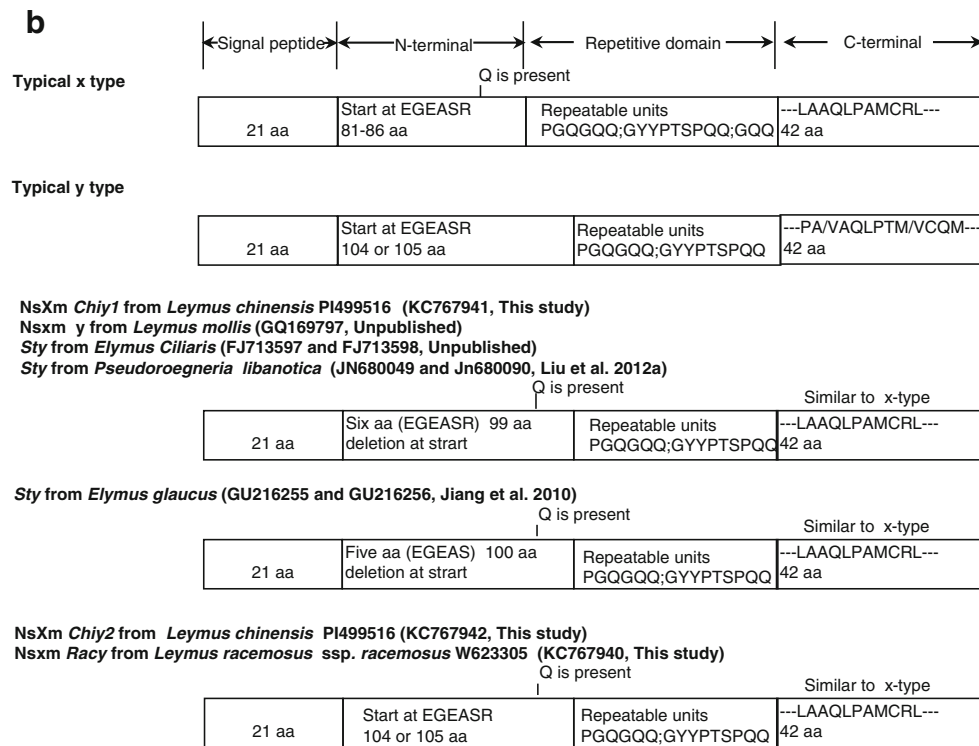
a

```
                          →  Signal peptide ←  →
X03042_Ay          MAKRLVLFATVVIGLVSLTVAEGEASK----------QLQCERELQESSLEACRLVVDQQLASRLPWSTGLQMRCCQQLR
X61026_By9         MAKRLVLFATVVITLVALTAAEGEASR----------QLQCERELQESSLEACRQVVDQQLAGRLPWSTGLQMRCCQQLR
AF476960_Cy        MAKRLVLFVTVVIALVSLTAAEGEASR----------QLQCERELQESSLEACSQVVDQQLAGRLPWSTGLQMRCCQQLR
X03041_Dy12        MAKRLVFAAVVIALVALTTAEGEASR----------QLQCERELQESSLEACRQVVDQQLAGRLPWSTGLQMRCCQQLR
Ay298724_Ee1.8y    MAKRLVLFAIVVIALVALTVAEGEASR----------QLQCERELQESSLEACRLVVDQQLAGRLPWSTGLQMRCCQQLR
FJ481573_Fy        MAKRLVLFAAVVVALMALAVAEGGASG----------QLQCERELQERSLEACRRVVDQQLAGQLPWSTGLQMQCCQQLR
HM131806_Gy        MAKRLVLFATVVIALVAFTAAEGEASR----------QLQCERELQESSLEACRQVVDQQLAGRLPWSTGLQMRCCQQLR
EF417988_D-hordein MAKPLVLFAAVVVALMSLTAAEGEATGNNIFRPLRIGQLQCERELQETSLEACRQIVDQQLVGQLPWSTGLQMRCCQQLR
AY834230_Ky        MAKRLILFAAVVVALVALTTAKGEASR----------QLQCERELQESSLEACRQVVDQQLAGQLPWSTGLQMRCCQQLR
FJ481569_Oy        MAKRLVLFAAVVVALVALTGAESVASG----------QLQCERELQERSLEACRQVVDQQLAGQLPWSTGLQMQCCQQLR
DQ073531_Py        MAKRLVLFAAVVLALVALTAAEGGASG----------QLQCERELQESSLEACRQVVDQQLAGQLPWSTGLQMRCCQQLR
FJ481571_Qy        MAKRLVLFAAVVVALVALTAAEGEGSG----------QLQCERELQESSLEACWRVVDQQLVGQLPRRTGLQMRCCQQLR
AF216869_Ry        MAKRLVLFGIVVIALVALTAAEGEASR----------QLQCERELQESSLEACRQVVDQQLAGRLPWSTGLQMRCCQQLR
FJ481574_Xey       MAKRLVLFAAVVVALVALTAAEGEASR----------QLQCERELQESSLEACRRVVDQQLAGQLPWNTGLQMRCCQQLR
AY303125_Tay       MAKRLVLFATVVIALVALTVAEGEAYS----------QLQCERELQESSLEACRQVVDQQLAGRLPWSTGLQMRCCQQLR
AF476962_Uy        MAKRLVLFATVVIALVALTTAKGEASR----------QLQCERELQESSLEACRQVVDQQLAGRLPWSTGLQMRCCQQLR
FJ600491_Vy        MAKRVVLFAAVVVALVALTAAEGGASA----------QQRCERELQESSLEACRRVVDQQLAGQLPWSTGPQMQCCQQLR
JN591653_Wy        MAKRLVLFAAVVVALVALTAAEGEASG----------QLKCERELQESSLEACRRVVDQQLAGQLPLSTGLQMRCCQQLR
DQ344030_Sty       MAKRLVLFAAVVLALVALTAAEGGASG----------QLQCERELQESSLEACRRVVDQQLAGQLPWSTGLQMRCCQQLR
KC767940_Racy      MAKRLVLFAAVVATLVALTAAEGGASG----------QLQCQRELQESSLEACRRVVDQQLAGQLPWSTGLRMRCCQQLR
KC767941_Chiy1     MAKRLVLFAAVVVALVALAFG-----\----------QLQCERELQESSLEACRRVVDQQLAGQLPWSTGLQMRCCQQLR
KC767942_Chiy2     MAKRLVLFAAVVVALMALTAAEGGASG----------QLQCERELQESSLEACRRVVDQQLAGQLPWSTGLQMQCCQQLR
FJ713598_E.ciliaris MAKRLVLFAAVVIALVALAFG-----\----------QLQCERELQESSLEACRRVVDQQLAGQLPWSTGLQMRCCQQLR
GQ169797_L.mollis  MAKRLVLFAAVVIALVALAFG-----\----------QLQCERELQESSLEACRRVVDQQLAGQLPWSTGLQMRCCQQLR
FJ713597_E.ciliaris MAKRLVLFAAVVIALVALAFG-----\----------QLQCERELQESSLEACRRVVDQQLAGQLPWSTGLQMRCCQQLR
JN680049_Sty       MAKRLVLFAAVVVALVALAFG-----\----------QLQCERELQESSLEACRRVVDQQLAGQLPWSTGLQMRCCQQLR
JN680050_Sty       MAKRLVLFAAVVVALVALAFG-----\----------QLQCERELQESSLEACRRVVDQQLAGQLPWSTGLQMRCCQQLR
```

```
     N-terminal  ←                                                          →  C-terminal  ←
DISAKCRPVALSQVARQYGQ-TAVPPKGGPFYHRETTPLQQLQQGIFGGTSSQTVQSPYHVSVEQQAASPKVAKAHHPVAQLPTMC---QMEGGDALSASQ
DVSAKCRPVAVSQVVRQYEQ-TVVPPKGGSFYPGETTPLQQLQQVIFWGTSSQTVQNPYHVNTEQQTASPKVAKVQQPATQLPIMC---RMEGGDALSASQ
DVSAKCRPVAVSQVARQYEQ-TVLPPTGGSFYPRETTPLQQLQQGIFWGTSSQTVQSPYHVSAEQQATNPKVAKAQHPAAQLPTMC---QMEGGDALSASQ
DVSAKCRSVAVSQVARQYEQ-TVVPPKGGSFYPGETTPLQQLQQGIFWGTSSQTVQSPYHVSAEQQAASPMVAKAQQPATQLPTVC---RMEGGDALSASQ
DVSAKCRPVAVSQVARQYGQ-TAVPPKGGSFYPSETTPLQQLQQGIFWGTSSQIVQSPYHVSVEQQATSPKVAKAQHPVAQLPTMC---QMVGGDALSASQ
DVSPECRRVAVSQVARQYEQQTAVPPKGGSFYPGETTPLQQLQQII-FSGIPTLLRSPYHVSAEHQAASLKVAKAQQLAAQLPAMC---RLEGGDTLSASQ
DVSAKCRPVAVSQVARQYEQ-TAVLPKGGSFYPSETTPLQQLQQVIFWGTSSQTVQSPYHVSAEQQMASPKVAKAQQPATQLPIMC---RMEGGEPLSASQ
DVSPECRPAAVSQVARQYEQQTMVPSKGGSFYPGGTA--RPLQQGR-W--WETYVKSPCHVSTEQPVASLKVAKAQQLAAQLPAMC---RLEGGGALSASQ
DVSAKCRPVAVSQVARQYEQ-TAMPPKGGSFYPGDTTPTQQFQQRIFWGRSSQTVQNPYHVSVGQQAAGLKVAKAQQPVAQLPAMC---RLEGGDALSVRQ
DVSPECRRVAVSQVARQYEQQTAVPPKGGSFYPGETTPLQQLQRIF-W-GIPTLLSSPYHVRAEHQAASLKVAKAQQLAAQLPAMC---RLEGGDALSASQ
DVSPECRPVAVSQVARQYEQQTAAPPKGGAFYPGEATPPQQLQQRIFWGRSSQTVQSPYHVSAEQQVASLKVAKAQQLAAQLPAMC---RLEGGDALSASQ
DVSPECRPVAVSQVARQYEQQTTVPPKGRSFYPGETTPPQQLLQRIFWGRSSQTVQSPYHVSAEQQATSLKVAKAQQLAAQLPAMC---RLEGGDALSASQ
DVSAKCRRVAVSQVARQYEQ-TAVPPKGGSIYPGETTPLQQLQQGIFWGTSSQTVQSPCHVSAEQKATSPKVAKAQQPVAQLPTMC---QMEGGDTLSASQ
DVSPKCRPVAVSQVARQYEQQTVVPPKGVSFYPGETKPPQRLQ-------QSQIVQSSYHVSAEQQAAILKVAKAQQLAAQLPAMC---MEGRDALSASQ
DVSAKCRPIAISQVARQYEQ-TTVPAKGGSFYPVETMPPQQLQQGIFWRTSSQTVQSPYHVSAEQQAASPKVAKAQQPAAQLPTMC---MEGRDALSASQ
DVSAKCRSVAVSQVARQYEQ-TVVPPKGGSFYPGETTPLQQLQQGIFWGTSSQTVQSPYYVSAEQQAASPMVAKAQQPATQLPTVC---RMEGGDALSASQ
DVSPECRPVAVSQVTRQYEQQIVVPPKGGSFYPGKTTPPQQLQQRILWGRSSQTVQSPYHVSPEQQEASLMVAKAQQLAAQLPAMC---RLEGSAALSASQ
DVSPECRPVAVSQVARQYEQQTAVPT-----------------------IPYHVSAEQQAASMKVAKAQQLVAQLPAMCRLCRLEGGDALSASQ
DVSPECRPVAVSQVARQYEQQTAVPPKAGSFYPGEATPTQQLQQRIFWGRSSQTVQSPYHVSAEQQVASLKVAKAQQLAAQLPAMC---RLEGGDALSASQ
DISPECRPVAVSQVARQYEQQTAVPPKGRSFYPGEATPPQQLQQRIFWGRSSQTVQSPYHVSAEKQVASLKVAKAQQLTAQLPAMC---RLEGGDALSASQ
DVSPECRPVAVSQVARQYEQQTAVPPKGGSFYPDETTQPQQLQQRIFWGRSSQPVQSPYHVSAEQQAASLKVAKAQQLAAQLPAMC---RLEEGGDALSASQ
DVSPECRPVAVSQVARQYEQQTAVPPRGGSFYPGEARPPQQLQQRIF-GRSSQTVQSPYHVSAKQQVASLKVAKAQQLAAQLPAMC---RLEEGGDALSASQ
DISPECRPVAVSQVARQYEQQTPVPPKGGSFYPGETTQPQQLQQRIFWGRSLQPVQSPYHVSVEQQAVSLKVAKAQQLAAQLPAMC---RLEGGDALSTNQ
DISPECRPVAVSQVARQYEQQTPVPPKGGSFYPGETTQPQQLQQRIFWGRSLQPVQSPYHVSVEQQAVSLKVAKAQQLAAQLPAMC---RLEGGDALSTNQ
DVSPECRPVAVSQVARQYEQQTPVPPKGGSFYPGETTQPQQLQQRIFWGRSLQPVQSPYHVSVEQQAVSLKVAKAQQLAAQLPAMC---RLEGGDALSTNQ
DVSPECRPVAVSQVARQYEQQTPVPPKGGSFYPGETTPPQQLQQRIFWGRSSQPVQSPYHVSAEQQAASLKVAKAQQLAAQLPAMC---RLEGGDALSASQ
DVSPECRPVAVSQVARQYEQQTPVPPKGGSFYPGETTPPQQLQQRIFWGRSSQPVQSPYHVSAEQQAASLKVAKAQQLAAQLPAMC---RLEGGDALSASQ
```

b

Fig. 2 (continued)

separate HMW glutenins on vertical 10 % SDS-PAGE gels (Yan et al. 2002).

PCR amplification and cloning of HMW glutenin ORFs

Degenerate polymerase chain reaction (PCR) primers (F1: 5′-atggctaagcgg(c/t)t(a/g)gtcctctttg-3′; R1: 5′-ctatcactggct(a/g)gccgacaatgcg-3′) were used to amplify the open reading frames (ORFs) of HMW glutenins. PCRs were performed in a GeneAmp 9700 thermocycler (Applied Biosystems, Foster City, CA, USA) in total volumes of 50 μL. The PCR ingredients were 200–300 ng template DNA, 1.25 U *ExTaq* high-fidelity polymerase (Takara, Dalian, China), 1× *ExTaq* PCR buffer, 0.2 mM of each dNTP, and 1 μM of each primer. PCR conditions were at 94 °C for 5 min to denature the template DNA, followed by 24 cycles of 94 °C for 40 s and 68 °C for 8 min for annealing and extension, then a final incubation at 68 °C for 15 min.

PCR products were separated on 0.8 % agarose gels and recovered with a DNA recovery kit DP1602 (Bioteke, Beijing, China). Target DNA fragments were ligated into *pMD18-T* vectors (Takara, Dalian, China) and the ligated products were transformed into chemically competent cells of *Escherichia coli* DH10B to acquire positive clones. Full-length DNA sequences were obtained by sequencing at least three individual subclones, which were made by nested deletion methods (Yan et al. 2002).

Bacterial expression of ORFs of *Racy* and *Chiy1*

The candidate expressed y-type HMW glutenins in PI499516 and W623305 were *Racy* and *Chiy1*, respectively. To confirm the expression of *Racy* and *Chiy1* ORFs, two plasmid constructs, *PET-30a–Racy* and *PET-30a–Chiy1*, were developed using cloned DNA fragments without signal peptides. Recombinant plasmids were constructed using PCR mutagenesis to add restriction sites for *Nde*I and *Eco*RI. To express the ORF of *Racy* in *L. racemosus* ssp. *racemosus* W623305, the primers were F2: 5′-acccatatg gaaggtggggcctctgggca-3′ (*Nde*I restriction site underlined) and R2: 5′-tcggaattc ctatcac tggctggccgaca-3′ (*Eco*RI restriction site underlined). To express the ORF of *Chiy1* in *L. chinensis* PI499516, the primers were F3: 5′-acccatatg caactacagtgtgagcgcga-3′ (*Nde*I restriction site underlined) and R3: 5′-tcagaattc ctatcactggctggcc gaca-3′ (*Eco*RI restriction site underlined).

The mutant ORFs were introduced into bacterial expression vector *PET-30a* (Invitrogen, Carlsbad, CA, USA) to obtain positive plasmid constructs *PET-30a–Racy* and *PET-30a–Chiy1* using the PCR strategies described previously. Bacterial expression of HMW-GSs was conducted in plasmid constructs *PET-30a–Racy* and *PET-30a–Chiy1* using the bacterial host strain *BL21(DE3)plySs*. Bacterial expression was induced with 1 mM isopropyl-β-D-1-galactoside (IPTG) for 3 h, when the cell concentration reached $OD_{600}=0.6$. Bacterial expression of proteins was gauged using SDS-PAGE with uninduced plasmid constructs as a control.

Phylogenetic relationships among HMW glutenins of *Leymus* and other Triticeae

DNA sequences were translated into aa using the universal genetic code and aligned with HMW glutenin orthologs from representative Triticeae species using the multiple sequence alignment program ClustalW (Thompson et al. 1994). A phylogenetic tree of the deduced aa of the signal peptide and conserved N-terminal and C-terminal was constructed using MEGA5.0 with the complete deletion method (Tamura et al. 2011). We tested for the best model of aa evolution. The Jones–Taylor–Thornton model with a gamma distribution (JTT+G, γ=1.29) had the lowest Bayesian information criterion score and was considered optimal. The corrected Akaike information criterion and maximum likelihood (ML) value for the best model were 3,143.9 and −1,504.9, respectively. The ML tree was constructed using the complete deletion option with respect to gaps in the aligned sequences. Bootstrap values were estimated based on 1,000 replications. Evolutionary distances were calculated using the best model for each pair of aligned sequences.

Results

Analysis of HMW glutenin compositions in *Leymus* species

Theoretically, tetraploid *Leymus* species should have at most four HMW-GSs. Using wheat cv. Cy12 as reference, there were one to three HMW-GSs in each of six accessions from four *Leymus* species with similar electrophoretic mobilities with those of wheat (Fig. 1a, boxed region). It was not confirmed that some of the protein bands with faster electrophoretic mobilities than subunit 10 of wheat were also HMW-GSs (Fig. 1a, shown with a question mark). The electrophoretic mobilities of HMW-GSs in *Leymus* in the boxed region were similar to those of subunits in *P. juncea*, suggesting that these subunits were encoded by the Ns genome. Because the origin of the Xm genome has not been identified, the electrophoretic mobility of HMW-GSs encoded by the Xm genome could not be confirmed.

PCR amplification and characterization of HMW glutenins from *L. racemosus* ssp. *racemosus* and *L. chinensis*

Using the degenerate PCR primers, one and two DNA fragments ranging from 1.0 to 1.5 kb were amplified and sequenced from *L. racemosus* ssp. *racemosus* W623305 and *L. chinensis* PI499516, respectively (Fig. 1b). These results verified the presence of y-type HMW glutenin ORFs in these plants. Their DNA lengths were very short: 1,272 bp (from *L. racemosus* ssp. *racemosus* W623305, abbreviated as *Racy*), 1,389 bp (*L. chinensis* PI499516, *Chiy1*), and 1,234 bp

(*L. chinensis* PI499516, *Chiy2*). The nucleotide sequences were deposited in GenBank with accession numbers KC767940 (*Racy*), KC767941 (*Chiy1*), and KC767942 (*Chiy2*).

The ORFs of *Chiy1* and *Racy* were predicted to form mature proteins of 440 and 401 aa, respectively, after the 21-aa signal peptide was removed. The HMW-GS lengths were largely determined by the repetitive domain length. Compared with typical y-type glutenins, the ORFs of *Chiy1* and *Racy* were very short, with only 29 hexapeptides and 11 nonapeptides and 26 hexapeptides and 6 nonapeptides in their repetitive domains, respectively. The ORF of *Chiy2* was interrupted by frameshift mutations in the N-terminal via two base deletions (CA) at nucleotide positions 354 and 355.

For further comparison, the nucleotide sequences were translated into aa sequences after ignoring the base deletion. Three y-type HMW glutenins shared four structural domains, including a signal peptide, a central repetitive domain, and the flanking N-terminal and C-terminal, as homologous genes from wheat and related wild species (Fig. 2a). The aa compositions in the N-terminal and C-terminal were highly conserved among HMW glutenins. Normally, y-type HMW glutenins have 104 or 105 aa at the N-terminal. *Racy* and *Chiy2* shared 104 aa, but *Chiy1* had only 99 aa in its N-terminal. Compared with homologous y-type HMW-GSs, an 18-bp DNA fragment (ACCGCCGCTGAAGGTGGG encoding EGEASR) at the start of the N-terminal was absent in *Chiy1*. Therefore, the N-terminal in *Chiy1* began with QLQCER. An extra glutamine at the N-terminal and short LAAQLPAMCRL peptides at the C-terminal that were conserved in x types were also present in these y types (Fig. 2b).

Expression of *Racy* and *Chiy1* ORFs in bacterial cells

SDS-PAGE analysis of IPTG-induced cells of plasmid constructs *PET-30a–Racy* and *PET-30a–Chiy1* confirmed bacterial expression of the ORFs of *Racy* (Fig. 3a, lane I) and *Chiy1* (Fig. 3b, lane I). In contrast, no expressed protein was detected in any controls (Fig. 3, lane UI). Compared with CS, the electrophoretic mobility of the expression proteins was very small (Fig. 3). However, the expressed proteins were not verified in seeds because cross-pollination resulted in HMW-GS that differed among seeds.

Phylogenetic trees of y genes from *Leymus* and other Triticeae

An ML tree was constructed using the aa residues of the signal peptide and N-terminal and C-terminal (Fig. 4). The y-type HMW glutenins formed two clades. The first contained *Uy*, *Cy*, *Ee1.8y*, *Ky*, *Gy*, *Tay*, *Ry*, *Ay*, *By9*, and *Dy12* from *Aegilops* ssp., *Elytrigia elongata*, *Crithopsis delileana*, *Taeniatherum crinitum*, *Secale cereale*, and *Triticum* ssp. The y-type HMW-GS from O, Q,W, V, P, St, F, Xe, and H,

Fig. 3 Bacterial expression of the ORFs of **a** *Racy* from *L. racemosus* ssp. *racemosus* (accession number W623305) and **b** *Chiy1* from *L. chinensis* (PI499516). *Lanes UI* were controls showing no expression of HMW glutenins and *lanes I* were bacterial expression of HMW glutenins induced with 1 mM isopropyl β-D-1-thiogalactopyranoside showing the expression of target proteins (*arrowheads*)
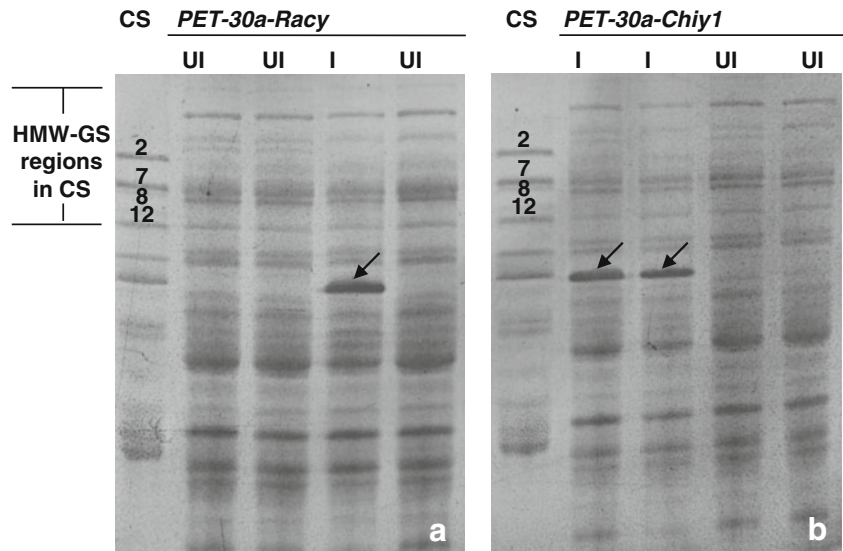
and from tetraploid *Leymus* and *Elymus* formed the second clade. The y-type HMW-GS lacking the EGEASR or EGEAS residues at the start of the N-terminal formed a subclade within the second clade.

Fig. 4 ML tree based on aa sequences of the signal peptide and N-terminal and C-terminal domains of y-type HMW-GS in two *Leymus* species and orthologous genes from Triticeae species. Glutenins lacking the EGEASR or EGEAS peptides at the start of the N-terminal are indicated with *gray backgrounds*
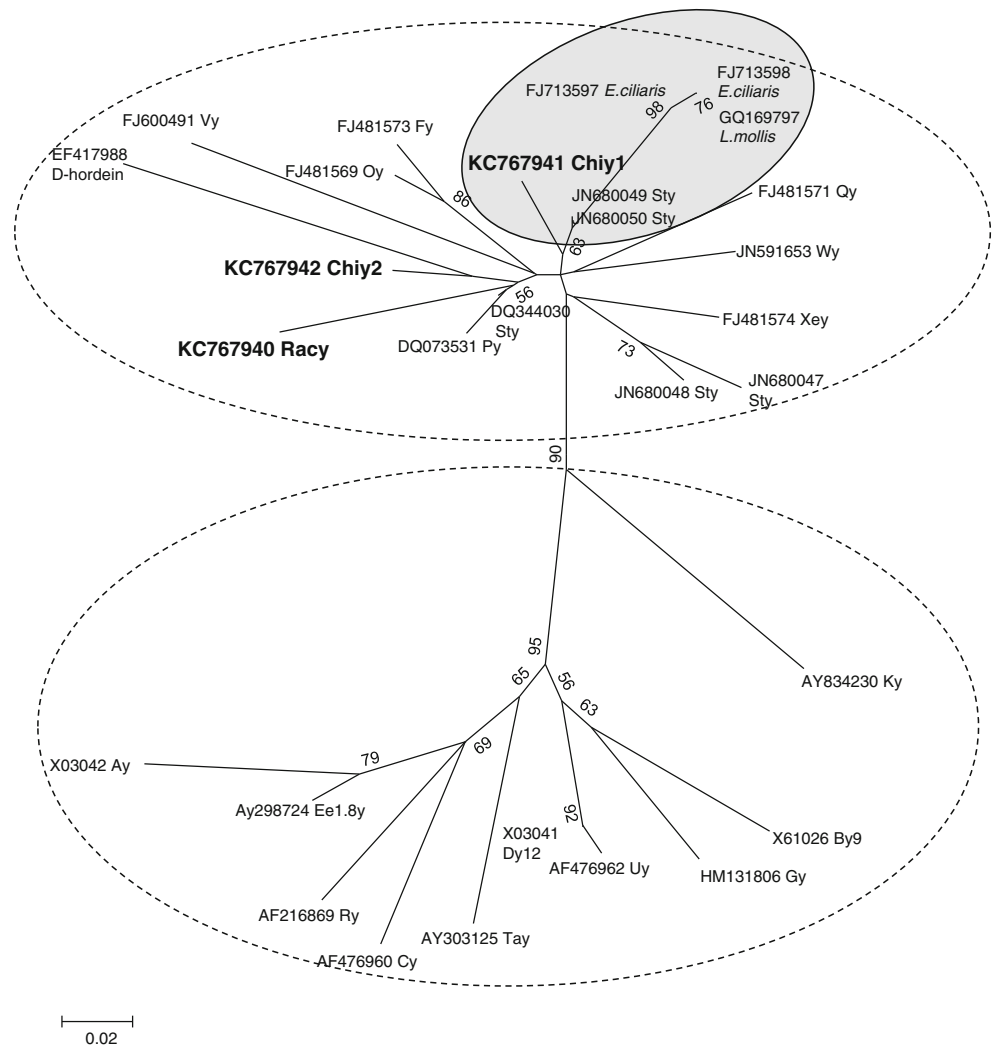
Discussion

Two and one HMW glutenin fragments were amplified from *L. chinensis* PI499516 and *L. racemosus* ssp. *racemosus* W623305, respectively. Two y types existed in *L. chinensis* PI499516, one (*Chiy2*), with 104 aa in its N-terminal by lacking of an amino acid W (tryptophan), started with EGERAS like most y types, and the other (*Chiy1*), with 99 aa in its N-terminal, began with QLQCER. In contrast, only one y-type glutenin (*Racy*) was found in *L. racemosus* ssp. *racemosus* W623305; it was very similar to *Chiy2* in PI499516, which had 104 aa in the N-terminal and started with EGERAS.

Although two y-type glutenin genes existed in PI499516, the ORF of *Chiy2* was probably not expressed because it had a two-base deletion in its N-terminal, resulting in a frameshift mutation and indirectly producing downstream in-frame stop codons (TAG). Previously characterized silence HMW-GS genes were deactivated by the insertion of a transposon or by an in-frame stop codon(s) due to base substitution. Similar to *Chiy2*, the HMW-GS gene *1By* in two common wheat landraces was silenced by a base deletion in a codon that indirectly produced premature stop codons (Yuan et al. 2009). Insertion of two extra bases in signal peptide or N-terminal was also responsible for silencing the HMW glutenin genes *Qy* in *Henrardia persica* PI577712 and *Sty1.5* in *Elymus sibiricus* (Liu et al. 2012b; Wang et al. 2012).

Absence of the short EGERAS peptide at the start of the N-terminal shortened this subunit of *Chiy1*. Five other y-type glutenin genes, including those from *Elymus ciliaris*, *Pseudoroegneria libanotica* (Jiang et al. 2010; Liu et al. 2012a), and *L. mollis* (GenBank accession number GQ169797, unpublished), also lacked the N-terminal EGERAS or EGERA peptides. Interestingly, these genes formed a monophyletic cluster within the second y-type clade (Fig. 4). Sequencing suggested that *Chiy1* and *Racy* had intact ORFs and could be predicted to form mature HMW-GS. However, the molecular sizes of HMW-GSs predicated by *Chiy1* and *Racy* were smaller than those of HMW-GSs in wheat. Bacterial expression of the modified ORFs (without signal peptides) of *Chiy1* and *Racy* confirmed the expression of HMW-GSs with small molecular sizes (Figs. 1 and 3). Therefore, both y-type glutenin genes in *Leymus* were expressed, regardless of whether the N-terminal started with EGEASR or QLQCER. However, the electrophoretic mobilities of the proteins expressed by *Chiy1* and *Racy* were similar with some of the protein bands, with a question mark in Fig. 1, and were faster than Dy12 of wheat cv. CS, suggesting that some of the protein bands were also HMW-GSs.

However, the two y-type HMW glutenin genes could not be targeted to the Xm or Ns genomes because of the reference sequences from the diploid parents and the undefined Xm genome. In previous investigations, other y-type glutenins with small sizes were also reported in wheat wild relatives. For example, the y-type subunits in *Taeniatherum* ssp. (Yan et al. 2006) and *Pseudoroegneria stipifolia* (Liu et al. 2012a) behaved like the y-type glutenins depicted in this study.

Bacterial expression of y-type HMW-GS with small molecular masses also indicated that the x-type HMW glutenins, which have larger molecular mass, were not acquired in this study. Unexpectedly, the PCR primers used here amplified both x-type and y-type glutenins from wild wheat relatives, such as *Aegilops* spp., *Taeniatherum*, and *E. elongata* (Yan et al. 2002, 2006; Wang et al. 2006), but failed to amplify the x-type glutenins from *Leymus*. Perhaps too many base mismatches occurred between the primers and the coding regions. In an attempt to isolate the x-type and y-type HMW-GS genes from *P. juncea* (Ns genome) by sequencing the promoter sequences, at least two base differences were detected in the signal peptide regions of two accessions (Yang et al. 2010). Therefore, better PCR primers are needed to obtain the x-type glutenins of *Leymus* species; this work is now underway.

Because the new y-type glutenins with shortened N-terminal were observed in tetraploid *Leymus* species with genomes NsNsXmXm, as well as in diploid *P. libanotica* (StSt) and tetraploid *E. ciliaris* (StStYY) with St genomes, we speculated that the St genome was one of two or more donors for this type of glutenin.

References

Dewey DR (1984) The genome system of classification as a guide to intergeneric hybridization with the perennial Triticeae. In: Gustafson JP (ed) Gene manipulation in plant improvement. Plenum, New York, pp 209–279

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

Guo ZF, Yan ZH, Wang JR, Wei YM, Zheng YL (2005) Characterization of HMW prolamines and their coding sequences from *Crithopsis delileana*. Hereditas 142:56–64

Jiang QT, Wei YM, Lu ZX, Pu ZE, Lan XJ, Zheng YL (2010) Structural variation and evolutionary relationship of novel HMW glutenin subunits from *Elymus glaucus*. Hereditas 147:136–141

Liu SW, Zhao F, Gao X, Chen FG, Xia GM (2010) A novel high molecular weight glutenin subunit from *Australopyrum retrofractum*. Amino Acids 39:385–392

Liu ST, Zhu XL, Liu SW, Tan Y (2012a) Isolation and characterization of *Glu-1* genes from the St genome of *Pseudoroegneria libanotica*. Gene 499:154–159

Liu J, Zhang B, Liu BL, Li HQ, Wang Z, Liu Q, Wei L, Liu DC, Zhang HG (2012b) Molecular cloning and analysis of a novel HMW-GS gene Glu1-St1.5 from *Elymus sibiricus* in Qinghai-Tibetan Plateau. Sci Res Essays 7:2189–2194

Payne PI, Law CN, Mudd EE (1980) Control by homoeologous group 1 chromosomes of the high-molecular-weight subunits of glutenin, a major protein of wheat endosperm. Theor Appl Genet 58:113–120

Pistón F, Shewry PR, Barro F (2007) D hordeins of *Hordeum chilense*: a novel source of variation for improvement of wheat. Theor Appl Genet 115:77–86

Shewry PR, Halford NG, Tatham AS, Popineau Y, Lafiandra D, Belton PS (2003) The high molecular weight subunits of wheat glutenin and their role in determining wheat processing properties. Adv Food Nutr Res 45:219–302

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28:2731–2739

Thompson JD, Higgins DG, Gibson TJ (1994) Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Wang RR, Bothmer RV, Dvořák J, Fedak G, Linde-Laursen I, Muramatsu M (1994) Genome symbols in the Triticeae (Poaceae), In: Wang RR, Jensen KB, Jaussi C (eds) Proceedings of the 2nd International Triticeae Symposium, Logan, Utah, USA, pp 29–34

Wang JR, Yan ZH, Wei YM, Zheng YL (2006) Characterization of high molecular weight glutenin subunit genes from *Elytrigia elongata*. Plant Breed 125:89–95

Wang XQ, Yan ZH, Dai SF, Liu DC, Wei YM, Zheng YL (2012) Molecular characterization of four HMW glutenin genes from *Heteranthelium piliferum* C. E. Hubbard and *Henrardia persica* (Banks et Solander) Hochstetter. Genet Resour Crop Evol 59: 1309–1318

Yan ZH, Wan YF, Liu KF, Zheng YL, Wang DW (2002) Identification of a novel HMW glutenin subunit and comparison of its amino acid sequence with those of homologous subunits. Chin Sci Bull 47:220–225

Yan ZH, Wei YM, Wang JR, Liu DC, Dai SF, Zheng YL (2006) Characterization of two HMW glutenin subunit genes from *Taenitherum* Nevski. Genetica 127:267–276

Yang F, Yan ZH, Dai SF, Han C, Hu XG (2010) Isolation and characterization of upstream sequences of two high molecular weight glutenin subunit genes from *Psathyrostachys juncea*. J Sichuan Agric Univ 28:397–402

Yuan ZW, Chen QJ, Zhang LQ, Yan ZH, Zheng YL, Liu DC (2009) Molecular characterization of two silenced y-type genes for *Glu-B1* in *Triticum aestivum* ssp. *yunnanese* and ssp. *tibetanum*. J Integr Plant Biol 51:93–99

Zhang HB, Dvořák J (1991) The genome origin of tetraploid species of *Leymus* (Poaceae: Triticeae) inferred from variation in repeated nucleotide sequences. Am J Bot 78:871–884