Cereal Research Communications 46(3), pp. 499–509 (2018)
DOI: 10.1556/0806.46.2018.018
Published Online: July 19, 2018

Characterization of a Novel 4.0-kb Y-type HMW-GS from *Eremopyrum distans*

S.F. Dai[1], D.Y. Xu[1], Z.J. Wen[1], Z.P. Song[1], H.X. Chen[1], H.Y. Li[1], J.R. Li[1], L.Z. Kang[1] and Z.H. Yan[1,2]*

[1]Triticeae Research Institute, Sichuan Agricultural University, 611130, Wenjiang District, Chengdu, Sichuan, 611130, China

(Received 6 August 2017; Accepted 27 November 2017;
Communicated by F. Békés)

A novel 4.0-kb Fy was sequenced and bacterially expressed. This gene, the largest y-type HMW-GS currently reported, is 4,032-bp long and encodes a mature protein with 1,321 amino acid (AA) residues. The 4.0-kb Fy shows novel modifications in all domains. In the N-terminal, it contains only 67 AA residues, as three short peptides are absent. In the repetitive domain, the undecapeptide RYYPSVTSPQQ is completely lost and the dodecapeptide GSYYPGQTSPQQ is partially absent. A novel motif unit, PGQQ, is present in addition to the two standard motif units PGQGQQ and GYYPTSPQQ. Besides, an extra cysteine residue also occurs in the middle of this domain. The large molecular mass of the 4.0-kb Fy is mainly due to the presence of an extra-long repetitive domain with 1,279 AA residues. The novel 4.0-kb Fy gene is of interest in HMW-GS gene evolution as well as to wheat quality improvement with regard to its longest repetitive domain length and extra cysteines residues.

Keywords: high-molecular-weight glutenin subunits (HMW-GSs), y-type genes, F genome, *Glu-1,* gene sequencing

Introduction

High-molecular-weight glutenin subunits (HMW-GSs) are important storage proteins in the endosperms of Triticeae. HMW-GSs comprise ~10% of all storage proteins but have a profound influence on the rheological and dough-baking properties of wheat flours (Payne 1987). The genes encoding HMW-GSs are located on the long arms of homoeologous group-1 chromosomes. Hexaploid wheat (*Triticum aestivum* L.; $2n = 6x = 42$, AABBDD) has three sets of gene loci, namely *Glu-A1*, *Glu-B1*, and *Glu-D1*, encoding HMW-GSs. At each locus, two tightly linked paralogous x- and y-genes produce a larger x-type and a smaller y-type protein, respectively (Lawrence and Shepherd 1981).

The x- and y-type HMW-GSs among Triticeae species share four similar structural domains, Viz. a signal peptide (removed in the mature protein) and an intermediate repetitive domain flanked by conserved N and C terminal domains (Lawrence and Shepherd 1981). X- and y-type HMW-GSs share 21 amino acid (aa) residues in the signal peptide and 42 aa

*Corresponding Author; E-mail: zhyan104@163.com; Phone/Fax: +86 288265 0350

residues in the C terminal, with the number of aa residues in the large central repetitive domain and the N-terminal are more variable. In the N-terminal, x-types share 81–86 aa residues and y-types share 104 or 105 aa residues. Differences in the central repetitive domain length are largely responsible for size variations among HMW-GSs. Three types of basic repetitive motif units, hex- (PGQGQQ), nona- (GYYPTSPQQ) and tripeptide (GQQ), occur in the repetitive domain. Hex- and nonapeptide units are common to x- and y-types, whereas tripeptide units are specific to x-types. Central repetitive domain lengths are largely determined by the numbers of repeat motif units (Shewry et al. 2003; Rasheed et al. 2014). In addition to differences in the structural domains, cysteines in x- and y-types differ from each other. X-types usually have three cysteines in their N-terminals and y-types have five; both have one cysteine in C terminals.

HMW-GSs with large molecular size have attracted considerable research interest. Some of the largest reported HMW-GS genes include *Ux* from *Aegilops umbellulata* (Liu et al. 2003), *Sx* from the Sitopsis section of *Aegilops* (Sun et al. 2006; Jiang et al. 2012), and the *Dx* subunits Dx 2.2 and Dx 2.2* from hexaploid wheat (Wan et al. 2005). The open reading frames (ORFs) of these genes typically consist of about 3.0 kb in nucleotides and approximately 100–106 kDa in mature protein.

Eremopyrum is a small Triticeae genus that contains two basic Xe and F genomes (Frederiksen 1991). Both the x- and y-type HMW-GS genes in Xe genome and the y-type in F genome have been reported in diploid *Er. triticeum* and *Er. distans* species (Dai et al. 2013), but the putative 4.0-kb x-type gene of *Er. distans* was not reported. The 4.0-kb HMW-GS gene was sequenced and expressed. Not anticipated, it is a novel y-type rather than x-type gene with an extra long ORF and revealed it to be the largest HMW-GS genes. Our results shed new light on sequence variations of HMW-GS genes and provide novel genes that can be used for wheat quality improvement.

Materials and Methods

Cloning of the 4.0-kb HMW-GS ORF from Er. distans PI 193264

The 4.0-kb HMW-GS ORF was amplified from *Er. distans* PI 193264 as described previously (Dai et al. 2013). The 4.0-kb PCR fragment was separated on a 0.8% agarose gel and recovered using a DP1602 DNA recovery kit (Bioteke, Beijing, China). The target DNA fragment was ligated into a *pMD18-T* vector (Takara, Dalian, China) and then transformed into chemically competent cells of *Escherichia coli* DH10B to acquire the resulting plasmid. The full ORF sequence was obtained via transposing sequencing provided by Takara.

Sequence comparison of the 4.0-kb Fy with homologous proteins

The nucleotide of the 4.0-kb Fy ORF was translated into amino acids using universal triplet genetic codes. Multiple sequence alignment of four HMW-GSs from *Eremopyrum* was carried out using Clustal W (Thompson et al. 1994). To determine the reason for the

Table 1. A comparison of the 4.0-kb Fy with the largest x- and y-type HMW-GSs previously reported from wheat and wild relative species

Source	Subunit	GenBank accession	References	ORFs size[a]	Signal peptide size[b]	Molecular mass, KD	N-terminus		Repetitive domain			C-terminus		Total	
							Size[a]	Cysteine residues	Size[a]	Hex-, Nona-, Triploid peptide	Cysteine residues	Size[a]	Cysteine residues	Size[a]	Cysteine residues
Er. triticeum	Fy	KR493381	This study	4032	21	139,945	67	5	1279	134/25/9[c]	2	42	1	1321	8
T. aestivum	Dx2.2*	AJ893508	Wan et al. 2005	3073	21	106,598	89	3	872	93/23/7	0	42	1	1003	4
Ae. umbellulata	Ux	AF476961	Liu et al. 2003	2983	21	103,421	86	3	843	88/22/30	0	42	1	971	4
Ae. bicornis	Sx	AY611727	Sun et al. 2006	2970	21	102,799	86	3	836	86/26/20	0	42	1	967	4
Ae. sharonensis	Sx	JN001485	Jiang et al. 2012	2928	21	101,203	83	3	825	84/24/26	0	42	1	953	4
T. aestivum	Dx2.2	AY159367	Wan et al. 2005	2919	21	100,886	89	3	819	86/23/23	0	42	1	950	4
Ae. bicornis	Sx	JN001481	Jiang et al. 2012	2901	21	100,388	86	3	816	83/24/25	0	42	1	944	4
Ae. longissima	Sx	JN001483	Jiang et al. 2012	2892	21	100,129	86	3	816	83/24/25	0	42	1	944	4
H. chilense	D-hordein	EF417988	Piston et al. 2007	2691	21	92,806	110	5	720	28/31[d]/4[e]	2	42	2	875	9
H. chilense	D-hordein	EF417989	Piston et al. 2007	2613	21	89,794	110	5	697	26/30[d]/4[e]	2	42	2	849	9
Er. triticeum	Xey	FJ481574	Dai et al. 2013	2655	21	91,259	98	5	722	83/23/0	2	42	1	862	8
A. retrofractum	Wy	JN591653	Li et al. 2012	2445	21	84,151	75	5	672	65/31/0	0	45	2	792	7
Ae. bicornis	Sy2.5	AY611728	Sun et al. 2006	2436	21	84,767	104	5	643	60/29/0	1	42	1	789	7

The superscripts a and b indicate the number of nucleotide base pairs and amino acid residues, respectively. The superscripts c, d, and e represent the peptides PGQGQQPGQQGYYPTSSQQ, PGQGQQGYYPSATSPQQ, and PHQGQQTTVS, respectively.

large molecular size of the 4.0-kb Fy, we compared it in detail with the 12 largest known x- and y-type HMW-GSs (Table 1).

Phylogenetic relationships of y-type HMW-GSs from Triticeae

A phylogenetic tree was constructed using the aa residues of conserved N- and C- terminals of y-type HMW-GSs from 19 Triticeae genomes (i.e., A, B, C, D, E, F, G, H, K, O, P, Q, R, St, Ta, U, V, W and Xe) with MEGA 5.0 (Tamura et al. 2011). The GenBank accession numbers for these proteins were shown in the phylogenetic tree. The best amino acid substitution model was selected for use in the analysis according to the Bayesian information criterion. The maximum likelihood 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.

Bacterial expression of the 4.0-kb Fy ORFs

To verify the functionality of 4.0-kb Fy ORFs, a plasmid construct *pET-30a–4.0-kb Fy* was developed using the cloned DNA fragment. The recombinant plasmid was obtained by PCR mutagenesis in order to remove the signal peptide. Because a rare *Nde*I site was present in the gene repetitive domain, two restriction sites, *Nco*I and *EcoR*I were added near the 5′ primer ends. The primer sequences were as follows. FyDF: 5′-accccatggaaggtgaagcctctgggcaa-3′ (*Nco*I site underlined) and FyDR: 5′-ctagaattcctatcactggctggccgacaa-3′ (*EcoR*I site underlined).

The altered ORF was inserted into a *pET-30a* bacterial expression vector (Invitrogen, Carlsbad, CA, USA) to obtain the plasmid construct *pET-30a–4.0-kb Fy.* Bacterial expression of the target plasmid was performed using *E. coil BL21 (DE3) plySs*. Bacterial expression was induced with 1 mM IPTG once the cell concentration OD_{600} reached 0.6 (about 3 h). Bacterial expression of the protein was detected by SDS-PAGE using uninduced plasmids as a control.

Results

Sequencing analysis of the 4.0-kb Fy

Sequencing revealed that the 4.0-kb fragment in *Er. distans* PI 193264 was 4,032 bp in nucleotides and encoded a mature protein with 1,321 aa residues. Sequence comparison indicated a higher similarity to y-types than to x-types. We designated the gene as 4.0-kb Fy (GenBank accession number KR493381). In our previous study, we detected a small y-type gene consisting of 1,593 bp and containing 529 aa residues in the mature protein (herein referred to as the 1.5-kb Fy) in PI 193264 (Dai et al. 2013).

The 4.0-kb Fy shared four similar primary structural domains with the known x- and y-type HMW-GSs from diploid *Eremopyrum* species and wheat but shown novel modifications with them (Fig. 1ab). The 67-residue N-terminal of the 4.0-kb Fy was the shortest

Figure 1. Comparison of 4.0-Kb Fy with those of the HMW-GS genes from two diploid species of *Eremopyrum* (a) and their orthologous from wheat (b). Deletions unique to the 4.0-kb Fy are indicated by dashed boxes in standard x- and y-type HMW-GS. The actual positions of cysteines are indicated by arrowheads

of any HMW-GSs (Fig. 1a). The repetitive domain of HMW-GSs began with two short peptides, RYYPSVTSPQQ and GSYYPGQTSPQQ, but the 4.0-kb Fy lacked the first undecapeptide and the first four residues (GSYY) of the second short dodecapeptide (Fig. 1b). In addition to three standard motif units, PGQGQQ, GQQ, and GYYPTSPQQ in the repetitive domain of x- and y-types, a novel repetitive motif unit, PGQQ, was specified by the 4.0-kb Fy. The 4.0-kb Fy shared the same LAAQLPAMCRL peptide as x-genes at the C-terminals (Fig. 1b) but the cysteine was changed from position 32 to 3 (Fig. 1b). Besides, the 4.0-kb Fy has an extra cysteine in the middle of repetitive domain and totally eight cysteines in comparison with seven cysteines for standard y types (Table 1).

Comparison of the 4.0-kb Fy with other HMW-GSs with large molecular masses

Twelve x- and y-type HMW-GSs with large molecular sizes have been currently reported: seven x-types with 2,892- to 3,073-bp ORFs and 100,129- to 106,598-KDa mature proteins and five y-types with 2,436- to 2, 691-bp ORFs and 84,767- to 92,806-KDa mature proteins (Table 1). A detailed comparison of these genes indicated that their protein molecular sizes or nucleotide lengths were all smaller than those of the 4.0-kb Fy (4,032-bp ORF and 139,945-KDa mature protein). The 4.0-kb Fy is thus the largest currently reported x- or y-type HMW-GSs.

Multiple sequence alignment of the 4.0-kb Fy with homologous proteins – including Dx2, the 1.5-kb Fy, Dy12, and the 12 largest x- and y-type HMW-GSs – revealed that the 4.0-kb Fy was not derived from any of them by direct repeat or larger fragment insertion (not shown).

Expression of the 4.0-kb Fy ORF in bacterial cells

PCR mutagenesis of the 4.0-kb Fy ORF was used to remove the signal peptide (Fig. S1*) to yield the recombinant *pET-30a–4.0-kb Fy*. Because the His-Tag protein was located upstream of the 4.0-kb Fy, the fusion protein was slightly larger than the HMW-GS obtained from seeds. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis of isopropyl-β-D-1-galactoside (IPTG)-induced cells of the plasmid construct *pET-30a–4.0-kb Fy* confirmed bacterial expression of the 4.0-kb Fy ORF (Fig. 2). In contrast, no expressed protein was detected in the control (Fig. 2). Compared with HMW-GSs from seeds of *Er. distans* PI 193264, the expressed protein with the larger subunit had a slower electrophoretic mobility (Fig. 2).

Phylogenetic analysis of Triticeae y-types

A phylogenetic tree of y-type HMW-GS proteins was then constructed for Triticeae species by maximum likelihood analysis of N- and C-terminal residues (Fig. 3). In the tree shown in Fig. 3, the 20 y-types are evenly divided into two clades. Clade I contains Ay,

*Further details about the Electronic Supplementary Material (ESM) can be found at the end of the article.

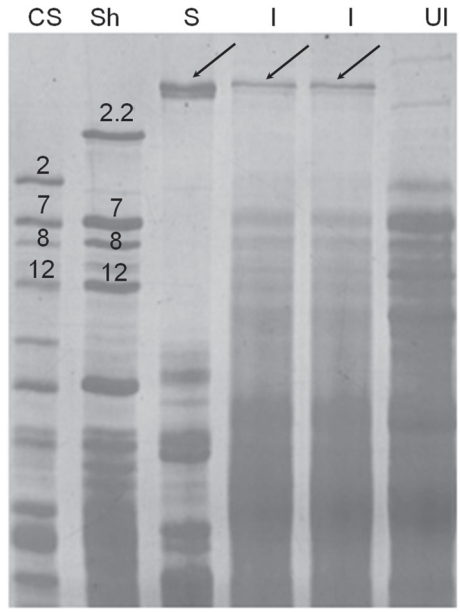

Figure 2. Bacterial expression of the 4.0-kb Fy ORF showing that the electrophoretic mobility of the infusion protein was slightly larger than those from seeds. Lanes CS (Chinese Spring) and Sh (Shinchunaga) are wheat references with known HMW-GSs. Lanes S, I, and UI correspond to HMW-GSs from seeds of PI 193264, from cells induced by 1 mM IPTG, and from control cells not induced by IPTG, respectively

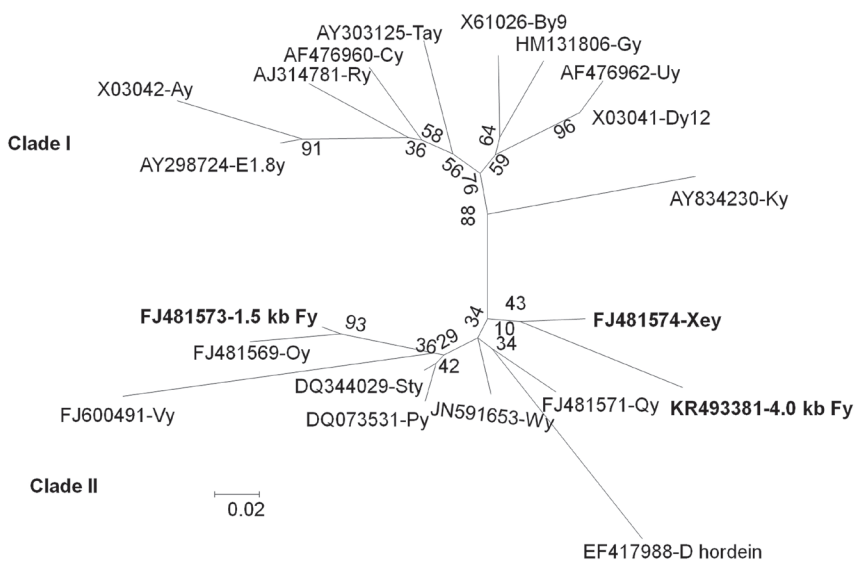

Figure 3. Phylogenetic tree of y-type HMW-GSs of Triticeae based on N- and C-terminal residues

By 9, Dy 12, and Gy proteins from *T. aestivum* and *T. araraticum*, Uy and Cy from *Ae. umbellulata* and *Ae. markgrafii*, Ry from *Secale cereale*, Ky from *Crithopsis delileana*, E1.8y from *Elytrigia elongata* and Tay from *Taeniatherum crinitum*. Clade II comprises the remaining 10 y-types: three Fy and Xey genes from *Er. distans* and *Er. triticeum,* Py from *Agropyron cristatum*, Oy from *Henrardia persica*, Qy from *Heteranthelium piliferum*, Sty from *Pseudoroegneria stipifolia*, Vy from *Dasypyrum villosum*, Wy from *Australopyrum retrofractum* and D-hordein from *Hordeum chilense*.

Discussion

A novel 4.0-kb HMW-GS gene was sequenced from *Er. distans* PI 193264 and its functionality was verified by bacterial expression. Sequencing result confirm that this gene is 4,032-bp in DNA length and encodes a protein with 1,321 AA residues. Unexpectedly, sequence comparison suggest that the 4.0-kb gene is a y-type rather than x-type. However, novel modifications were found in each domain. The N-terminal of the 4.0-kb Fy is shortest among all the x- and y-type HMW-GSs, with only 67 aa residues. A *Tax* from *Ta. crinitum* and a *Wy* from *Au. retrofractum* have also been found to have reduced N-terminal lengths (Yan et al. 2006; Liu et al. 2010, Li et al. 2012). Deletion of short peptides KGGSFYP and KGGSFYPGETTPLQQLQQGIFWGTSSQTVQ, consisting of 7 and 30 residues, respectively, was responsible for the 74 and 75 aa residues present in *Tax* and *Wy*. In contrast, the shortened N-terminal of the 4.0-kb Fy is a consequence of lacking three short peptides – QLPW/NSTG, SQVARQY, and PGETTPLQQLQQIIFSGIPTLLR – involving a total of 37 aa residues. The first deletion was located at the second and third cysteines, whereas the last two deletions occurred between the fifth cysteine and the N-terminal end (Fig. 1b). Three aspects of the repetitive domain of the 4.0-kb Fy are unique. First, standard x- and y-types often share a RYYPSVTSPQQ undecapeptide and a GSYYPGQTSPQQ dodecapeptide at the beginning of the repetitive domain. In the 4.0-kb Fy, however, the undecapeptide has been completely lost and the dodecapeptide is partially absent. A comparison with y-types from *Au.retrofractum* and *Ht. piliferum* reveals that the undecapeptide and dodecapeptide is missing in the former and the dodecapeptide has been lost in the latter (Liu et al. 2010; Wang et al. 2012). Second, a novel repetitive motif, PGQQ, occurs in the 4.0-kb Fy that has not been observed in any other HMW-GS. Other unique repetitive motifs, namely P/SGQGQPQGQGQQ, LGQGQQQQ, and PGQQGQQ, have been found in y-types from *Au. retrofractum*, *Ht. piliferum*, and *H. persica*, respectively (Liu et al. 2010; Wang et al. 2012). Third, the repetitive domain of the 4.0-kb Fy is the largest of all currently characterized x- and y-types. The 4.0 Kb Fy had a much larger repetitive domain length (1,279 residues) than that of the seven x-types and five y-types, in which it varied from 816 to 872 residues and 643 to 720 residues, respectively. The 4.0-kb Fy had more hexapeptides than the 12 largest HMW-GS proteins, which had 40–50 or 51–74 hexapeptide motif units in x- and y-types, respectively (Table 1). Compared with the largest x-protein, Dx 2.2*, the 4.0-kb Fy has more PGQGQQ hexapeptides (41 motif units) and fewer GQQ tripeptides (27 motif units). D-hordein of barley is highly similar to the y-types but has different repetitive motifs,

namely, PGQGQQGYYPSATSPQQ and PHQGQQTTVS. Compared with the largest D-hordein (EF417988), the 4.0-kb Fy has more PGQGQQ hexapeptide (106 motif units) and GYYPTSPQQ nonapeptide (25 motif units) but is lacking 31 PGQGQQGYYPSAT-SPQQ motif units and 4 PHQGQQTTVS motif units. Xey (FJ481574), the second largest y-protein after D-hordein, is smaller than the 4.0-kb because it lacks 51 PGQGQQ hexapeptide motif units (Dai et al. 2013). Some y-genes, including the 4.0-kb Fy and 1.5-kb *Fy*, *Xey*, *Oy*, *Qy*, and *Sty*, share the short C-terminal peptide LAAQLPAMCRL with x-subunits (Shewry et al. 2003). Based on these unique features, we hypothesize that the 4.0-kb Fy is an important HMW-GS evolutionary intermediate.

Novel HMW-GS alleles have been formed by single or multiple aa residue alterations, including point mutations as well as insertions and deletions of large fragments with repeatable motif units. Insertion and duplication of blocks with repeatable motif units are considered responsible for expansion and shrinkage of HMW-GS sizes. Unequal crossover events and slip-mismatching have been suggested as the most likely mechanisms for size variations in HMW-GSs. A detailed comparison of the large HMW-GSs Dx 2.2 and Dx 2.2* as well as Sx 2.9 with Dx 2 have indicated that they have different numbers and lengths of repeatable blocks (Jiang et al. 2012; Wan et al. 2005). Such duplications most likely arose by unequal crossing over during meiosis (D'Ovidio et al. 1996). Another probable explanation for size variations in HMW-GSs is illegitimate recombination (Zhang et al. 2008). In both cases, it is notable that the direct repeat sequence is located immediately adjacent to the region from which it was duplicated or deleted (D'Ovidio et al. 1996; Wan et al. 2005; Jiang et al. 2012). Because we have not found any direct repeat or repeatable blocks in the 4.0-kb Fy, we have no detailed explanation for its formation.

The HMW-GSs have a profound influence on the baking and processing properties of wheat flours. The factors influencing HMW-GSs to form larger, stable polymers are suggested to contribute to their role in determining dough strength. An extra cysteine located at the start of the 1Dx5 repetitive domain is thought to confer the capability to form a higher proportion of macropolymers on lines with 1Dx5+1Dy10 compared with lines having allelic pairs 1Dx2+1Dy12 (Popineau et al. 1994). A point mutation of C-G occur at 1,181 bp in the middle of repetitive region of the gene *Glu-A1x2**B in an old Hungarian wheat variety Bánkfúti 1201, resulting to the Ax2*^{B} subunit has an extra cysteine in its protein sequence, similar to subunit Dx5 (Juhász et al. 2003).The 4.0-kb Fy contains an additional cysteine residue in the central part of repetitive domain, thus possesses a total of eight cysteine residues (Fig. 1b). This suggested that the 4.0-kb Fy would favorable of forming a higher proportion of macropolymers than other y-type HMW-GSs. A positive relationship between the size of HMW-GSs and their effect on dough strength has been established through dough incorporation with 2 g Mixograph, with subunits 1Dx2.2 and 1Dx2.2* showing greater strength than subunit 1Dx2 (Békés et al. 1994). The 4.0-kb Fy has a longer repetitive domain than any currently identified HMW-GSs. This suggests that this subunit may promote the formation of larger, more stable glutenin polymers than other x- and y-genes with smaller repetitive domains. Our future work will focus on evaluating the wheat quality improvement potential of this larger HMW-GS by introducing it into wheat transgenically.

Thus, a novel 4.0-kb Fy, the largest HMW-GS to date, was characterized from *Er. distans* PI 193264. An extra-long repetitive domain comprising 1,279 AA residues was mainly responsible for its larger mass. The novel 4.0-kbFy shed a new light on the study of HMW-GS gene evolution, and also provided a novel candidate gene for wheat quality improvement.

Acknowledgements

The authors acknowledge financial support from the Ministry of Science and Technology of China (2016YFD0100500), the National Natural Science Foundation of China (U1403185, 31771783) and the Key Fund Project of the Sichuan Provincial Department of Education (15ZA0021).

References

Békés, F., Anderson, O.D., Gras, P.W., Gupta, R.B., Tam, A.S., Wrigley, C.W., Appels, R. 1994. The contribution to mixing properties of 1D HMW glutenin subunits expressed in a bacterial system. In: Henry, R.J., Ronalds, J.A. (eds) Improvement of cereal quality by genetic engineering. Plenum Press, New York. pp. 97–103.

Dai, S.F., Pu Z.J., Liu D.C., Wei, Y.M., Zheng, Y.L., Hu, X.K. Yan, Z.H. 2013. Characterization of novel HMW-GS in two diploid species of *Eremopyrum*. Gene **519:**55–59.

D'Ovidio, R., Lafiandra, D., Porceddu, E. 1996. Identification and molecular characterization of a large insertion within the repetitive domain of a high-molecular-weight glutenin subunit gene from hexaploid wheat. Theor. Appl. Genet. **93:**1048–1053.

Frederiksen, S. 1991. Taxonomic studies in *Eremopyrum* (Poaceae). Nord. J. Bot. **11:**271–285.

Jiang Q.T., Ma J., Wei Y.M., Liu Y.X., Lan X.J., Dai, S.F., Lu, Z.X., Zhao, S., Zhao, Q.Z., Zheng, Y.L. 2012. Novel variants of HMW glutenin subunits from *Aegilops* section Sitopsis species in relation to evolution and wheat breeding. BMC Plant Biology **12:**73.

Juhász, A., Larroque, O.R., Tamás, L., Hsam, S.L.K., Zeller, F.J., Békés, F., Bedő, Z. 2003. Bánkfúti 1201 – an old Hungarian wheat variety with special storage protein composition. Theor. Appl. Genet. **107:**697–704.

Lawrence, G.J., Shepherd, K.W. 1981. Chromosomal location of genes controlling seed protein in species related to wheat. Theor. Appl. Genet. **59:**25–31.

Li, F., Jiang, X.L., Wei, Y.F., Xia, G.M., Liu, S.W. 2012. Characterization of a novel type of HMW subunit of glutenin from *Australopyrum retrofractum*. Gene **492:**65–70.

Liu, S.W., Zhao, F., Gao, X., Chen, F.G., Xia, G.M. 2010. A novel high molecular weight glutenin subunit from *Australopyrum retrofractum*. Amino Acids. **39:**385–392.

Liu, Z., Yan, Z., Wan, Y., Liu, K., Zheng, Y., Wang, D. 2003. Analysis of HMW glutenin subunits and their coding sequences in two diploid *Aegilops* species. Theor. Appl. Genet. **106:**1368–1378.

Payne, P.I. 1987. Genetics of wheat storage proteins and the effect of allelic variation on breadmaking quality. Annu. Rev. Plant Physiol. **38:**141–153.

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

Popineau, Y., Cornec, M., Lefebvre, J., Marchylo, B. 1994. Influence of high Mr glutenin subunits on glutenin polymers and rheological properties of gluten and gluten subfractions of near-isogenic lines of wheat Sicco. J. Cereal Sci. **19:**231–241.

Rasheed, A., Xia, X.C., Yan, Y.M., Appels, R., Mahmood, T., He, Z.H. 2014. Wheat seed storage proteins: Advances in molecular genetics, diversity and breeding applications. J. Cereal Sci. **60:**11–24.

Shewry, P.R., Halford, N.G., Tatham, A.S., Popineau, Y., Lafiandra, D., Belton, P.S. 2003. The high molecular weight subunits of wheat glutenin and their role in determining wheat processing properties. Adv. Food Nutr. Res. **45:**219–302.

Sun, X., Hu, S.L., Liu, X., Qian, W.Q., Hao, S.T., Zhang, A.M., Wang, D.W. 2006. Characterization of the HMW glutenin subunits from *Aegilops searsii* and identification of a novel variant HMW glutenin subunit. Theor. Appl. Genet. **113:**631–641.

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, J.D., Higgins, D.G., Gibson, T.J. 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.

Wan, Y.F., Yan, Z.H., Liu, K.F., Zheng, Y.L., D'Ovidio, R., Shewry, P.R. Halford, N.G., Wan, D.W. 2005. Comparative analysis of the D genome-encoded high-molecular weight subunits of glutenin. Theor. Appl. Genet. **111:**1183–1190.

Wang, X.Q., Yan, Z.H., Dai, S.F., Liu, D.C., Wei, Y.M., Zheng, Y.L. 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, Z.H., Wei, Y.M., Wang, J.R., Liu, D.C., Dai, S. F., Zheng Y.L. 2006 Characterization of two HMW glutenin subunit genes from *Taeniatherum* Nevski. Genetica **127:**267–276.

Zhang, Y.Z., Li, X.H., Wang, A.L., An, X.L., Zhang, Q., Pei, Y.H., Gao, L.Y., Ma, W.J., Appels, R., Yan, Y.M. 2008. Novel x-type high-molecular-weight glutenin genes from *Aegilops tauschii* and their implications on the wheat origin and evolution mechanism of Glu-D1-1 proteins. Genetics **178:**23–33.

Electronic Supplementary Material (ESM)

Electronic Supplementary Material (ESM) associated with this article can be found at the website of CRC at https://akademiai.com/loi/0806

Electronic Supplementary *Figure S1*. PCR amplification of the ORF of the 4.0-kb Fy from DNA clones to remove signal peptide