Identification and Characterization of High-Molecular-Weight Glutenin Subunits from *Agropyron intermedium*

Shuanghe Cao¹, Zhixin Li^{1,2}, Caiyan Gong¹, Hong Xu³, Ran Yang¹, Shanting Hao¹, Xianping Wang¹, Daowen Wang¹, Xiangqi Zhang¹*

1 State Key Laboratory of Plant Cell and Chromosome Engineering, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China, 2 College of Agriculture, Yangtze University, Jingzhou, Hubei, China, 3 College of life sciences, Northwest Sci-Tech University of Agriculture and Forestry, Yangling, Shanxi, China

Abstract

High-molecular-weight glutenin subunit (HMW-GS) is a primary determinant of processing quality of wheat. Considerable progress has been made in understanding the structure, function and genetic regulation of HMW-GS in wheat and some of its related species, but less is known about their orthologs in Agropyron intermedium, a useful related species for wheat improvement. Here seven HMW-GSs in Aq. intermedium were identified using SDS-PAGE and Western blotting experiments. Subsequently, the seven genes (Glu- $1Aix1 \sim 4$ and Glu- $1Aiy1 \sim 3$) encoding the seven HMW-GSs were isolated using PCR technique with degenerate primers, and confirmed by bacterial expression and Western blotting. Sequence analysis indicated that the seven Ag. intermedium HMW-GSs shared high similarity in primary structure to those of wheat, but four of the seven subunits were unusually small compared to the representatives of HMW-GS from wheat and two of them possessed extra cysteine residues. The alignment and clustering analysis of deduced amino acid sequences revealed that 1Aix1 and 1Aiy1 subunits had special molecular structure, belonging to the hybrid type compounding between typical xand y-type subunit. The xy-type subunit 1Aix1 is composed of the N-terminal of x-type and C-terminal of y-type, whereas yxtype subunit 1Aiy1 comprises the N-terminal of y-type and C-terminal of x-type. This result strongly supported the hypothesis of unequal crossover mechanism that might generate the novel coding sequence for the hybrid type of HMW-GSs. In addition to the aforementioned, the other novel characteristics of the seven subunits were also discussed. Finally, phylogenetic analysis based on HMW-GS genes was carried out and provided new insights into the evolutionary biology of Ag. intermedium.

Citation: Cao S, Li Z, Gong C, Xu H, Yang R, et al. (2014) Identification and Characterization of High-Molecular-Weight Glutenin Subunits from Agropyron intermedium. PLoS ONE 9(2): e87477. doi:10.1371/journal.pone.0087477

Editor: Mark Gijzen, Agriculture and Agri-Food Canada, Canada

Received September 8, 2013; Accepted December 23, 2013; Published February 4, 2014

Copyright: © 2014 Cao et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by National Nature Science Foundation of China (grant no. 30930059), Ministry of Science and Technology of China (grant nos. 2009CB118300, 2011BAD07B02), and Ministry of Agriculture of the People's Republic of China (grant no. 2013ZX08009-003-004). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: xqzhang@genetics.ac.cn

Introduction

In common wheat (Triticum aestivum L.) and related species, highmolecular-weight glutenin subunit (HMW-GS) is an important group of seed storage proteins and shares the primary structure that includes the conserved signal peptide (which is excised from the mature HMW-GS), N- and C-terminal domains, and the central repetitive region composed of tri-, hexa- and nonapeptide motifs [1,2,3]. Mainly judging from the sizes and sequences of conserved N-terminal domains (especially, the number and location of cysteine residues), HMW-GSs are categorized into x- and y-type encoded by Glu-1-1 and Glu-1-2 in the Glu-1 locus, respectively [1,4]. In addition to wheat, many orthologous HMW-GSs have been identified and isolated from related species (reviewed in [5]). Compared to wheat, quite a few HMW-GSs from related species possess an unusual number and/or location of cysteine residues, which are involved in the formation of inter- or intra-molecular disulphide bonds, and accordingly have an important effect on the network of gluten macro polymer (GMP) [6,7,8]. Noticeably, two hybrid HMW-GSs were identified from Aegilops searsii and Triticum aestivum-Agropyron intermedium disomic

alien addition lines, respectively [9,10]. Owing to their role in determining the strength and elastic properties of the gluten complex, and their conservation in wheat and related species [1,3,7], continued mining for novel HMW-GSs from wild germplasm may be beneficial for not only improving wheat end use quality but also studying the structural variation and evolution of this important protein family.

Ag. intermedium (Host) Beauvoir (= Thinopyrum intermedium (Host) Barkwarth = Elytrigia intermedia (Host) Nevski, 2n = 6x = 42), being one of the wild relatives of wheat, possesses many desirable characteristics for wheat improvement, such as high grain protein content, resistance to many diseases caused by fungi or viruses, tolerance to low temperature, drought, moisture and salt stresses, perennial habits and high crossability with wheat, etc (reviewed in [11]). Because of its high crossability with wheat, a number of useful genes have been transferred from this species to common wheat through chromosome engineering, which has led to the development of many useful wheat germplasm, including partial amphiploids, addition, substitution and translocation lines [12,13,14,15,16]. Ag. intermedium has been one of the most important related species. Sun (1981) developed five partial amphiploids of *T. aestivum-Ag. intermedium*, i.e., octoploid *Trititrigia* Zhong1~5 (2n = 8x = 56) through crossing common wheat with *Ag. intermedium* [15]. Furthermore, He et al. (1988) developed two sets of *T. aestivum-Ag. intermedium* disomic alien addition lines (TAI-II & TAI-II series) based on Zhong 1~5 [16]. We measured the quality parameters of TAI series lines and found that some of them have good quality (data not shown). In addition, several high-quality varieties were bred and released in China, such as Shanmai150, Gaoyou503, Xiaobingmai33 and so on, using these octoploid *Trititrigia* or their derived lines. Thus, it is valuable to identify and characterize *Ag. intermedium* glutenin subunits for improving the processing property of wheat flour, considering that its quality is largely determined by the composition of glutenin subunits [1,3,7].

We previously identified and characterized a novel Ag. intermedium HMW-GS gene (Glu-1Aix1) from a T. aestivum-Ag. intermedium addition line TAI-13. In addition to belonging to the hybrid type mentioned above, the subunit 1Aix1 possesses two extra cysteine residues in the repetitive region [10]. In the present study, we identified and characterized a series of HMW-GSs and their coding sequences from Ag. intermedium and further investigated their evolutionary biology with phylogenetic analysis.

Materials and Methods

Plant Materials

The Ag. intermedium line in this study was kindly provided by Professor Mengyuan He (Northeast Normal University, Changchun, China) and its accession number is No. 33826 in N. I. Vavilov Research Institute of Plant Industry, Russia. This Ag. intermedium line has been utilized to create the chromosomal engineering lines, Zhong $1\sim5$ and TAI series, which are good "bridge" parents of wheat breeding [15,16]. Common wheat variety Chinese Spring was used as a standard for assessing the electrophoretic mobilities of the HMW-GSs from Ag. intermedium.

Glutenin Preparation, SDS-PAGE and Western Blotting

The glutenin proteins were extracted from mature seeds of Ag. intermedium and Chinese Spring according to the method described by Fu and Kovacs [17] and were separated by SDS-PAGE. The electrophoresis was carried out at 160 volts for 1.5 hour using 12% gel of Laemmli basic discontinuous system in Mini-PROTEAN[®] 3 Cell (BIO-RAD), and the proteins were stained with Coomassie Brilliant Blue R-250 [18]. The HMW-GSs were identified and confirmed by Western blotting analysis according to the method described by Liu et al. [19].

DNA Extraction and PCR

Genomic DNA was extracted from the young leaves of Ag. intermedium with CTAB method [20]. For amplifying HMW-GS genes of Ag. intermedium by genomic PCR, a pair of degenerate primers, P1 (5'-ATGGCTAAGCGGC/TTA/GGTCCTCTTT-G-3') and P2 (5'-CTATCACTGGCTG/AGCCGACAATGCG-3'), was designed according to published DNA sequences of wheat HMW-GS genes [21]. The primer P1 contained the start codon of the HMW-GS gene's ORF (open reading frame), and the primer P2 possessed the two tandem stop codons that are conserved almost in all of HMW-GS genes characterized so far. Genomic PCR was carried out using high fidelity LA Taq polymerase with GC buffer (Takara). The parameters for the PCR reaction were one step at 94°C for 5 min, followed by 35 cycles of 94°C for 1 min, 65°C for 40 sec and 72°C for 3 min, and a final extension step at 72°C for 7 min.

Isolation and Confirmation of Complete ORFs

PCR products were separated in 1% agarose gels. DNA fragments of expected sizes were recovered from agarose gels, and then were ligated into the pGEM-T vector (Promega). The competent cells of *Escherichia coli* DH10B were transformed with the ligation reactions following the standard procedures (Promega). The positive clones were selected with blue/white color screening and sequenced by a commercial company (Bioasia Biotechnology Company). The sequencing of HMW-GS gene is stumbled by tandem and interspersed repeats in its central region. In order to obtain the complete ORFs, their nest-deleted subclones were prepared according to the protocol: the purified plasmid of positive clone was digested by *NcoI*, *ApaI*, Exonuclease III and S1 Nuclease in turn, and then their products self-looped by T4 DNA

Table 1. PCR primers used in bacterial expression experiments.

Primer pair	Sequence ^a	Restriction site introduced	Mutant ORF amplified
E1	5' -ACC <u>CATATG</u> GAAGGTGAGGCCTCTG-3'	Ndel	1Ai4
	5' -CTAGAATTCCTATCACTGGCTGGCC-3'	EcoRI	
E2	5' -ACC <u>CATATG</u> GAAGGTGAGGCCTCTG-3'	Ndel	1Ai1
	5' -CTA <u>GAATTC</u> CTATCACTGGCTAGCC-3'	EcoRI	
E3	5' -ACC <u>CATATG</u> GAAGGTGAGGCCTCTG-3'	Ndel	1Ai5
	5' -CTA <u>GAATTC</u> CTATCACTGGCTGGCC-3'	EcoRI	
E4	5' -ACC <u>CATATG</u> GAAGGTGAGGCCTCTG-3'	Ndel	1Ai6
	5' -CTAGAATTCCTATCACTGGCTAGCC-3'	EcoRI	
E5	5' -ACC <u>CATATG</u> GAAGGTGAGGCCTCTA-3'	Ndel	1Ai2
	5' -CTA <u>GAATTC</u> CTATCACTGGCTGGCC-3'	EcoRI	
E6	5' -ACC <u>CATATG</u> GAAGGTGAGGCCTCTA-3'	Ndel	1Ai3
	5' -CTAGAATTCCTATCACTGGCTGGCC-3'	EcoRI	
E7	5' -ACCCATATGGAAGGTGAGGCCTCTA-3'	Ndel	1Ai7
	5' -CTAGAATTCCTATCACTGGCTGGCC-3'	EcoRI	

^aThe underlined nucleotides constitute the restriction sites listed in the third column.

doi:10.1371/journal.pone.0087477.t001



Figure 1. SDS-PAGE (A) and Western blotting (B) analysis of HMW-GSs of *Ag. intermedium.* Lane 1 shows the named HMW-GSs from common wheat variety Chinese Spring as a control. Lanes 2~7 show the HMW-GSs from six representative seeds of the *Ag. intermedium* line used in this study. The seven expressed HMW-GSs with distinct electrophoretic mobility comparing with Chinese Spring were detected by SDS-PAGE (A) and were confirmed using Western blotting experiment with polyclonal antibody specific for HMW-GSs (B). Among the seven HMW-GSs from *Ag. intermedium*, three subunits (marked with solid triangles in lane 2 of B) share comparable electrophoretic mobility with Chinese Spring, the other four subunits (marked with hollow triangles in lane 2 of B) moved faster than those HMW-GSs from Chinese Spring. doi:10.1371/journal.pone.0087477.g001

ligase were transformed into competent cells of *E. coli* DH10B. The identification and sequencing of the subclones were carried out as described above. The complete ORF sequences were determined through overlapping subclones created by the nested deletion method.

Bacterial Expression and Western Blotting Analysis

For bacterial expression of the mature HMW-GSs from Ag. intermedium, a series of PCR primers (Table 1) was designed to amplify mutant ORFs from which the coding sequences of signal peptides were removed, and appropriate restriction enzyme sites (*Ndel* and *EcoRI*) were introduced to facilitate subsequent cloning experiments. PCR parameters for amplifying the mutant ORF are identical with the described above. After cloning the mutant ORFs into the expression vector pET-30a (Invitrogen), positive constructs were selected for expressing the mature proteins of the target HMW-GSs in bacterial cells. The conditions for inducing bacterial expression of pET construct are same as detailed by Wan et al. [22]. *E. coli* BL21 (DE3) pLysS cells were transformed with the pET plasmids containing the HMW-GS genes. The positive clones were cultured in LB liquid medium containing 35 µg/mL kanamycin in a shaking incubator at 37° C until the OD₆₀₀ reached 0.6, and then the expression of target HMW-GSs were induced by adding 1 mM IPTG into the medium. After incubation overnight, the *E. coli* cells were collected by centrifugation at 5,000 g and 4°C for 10 min, and the expressed proteins were subsequently extracted using 50% (v/v) 1-propanol containing 2% (w/v) DTT. To confirm the identity of the bacterially expressed proteins directed by pET constructs, Western blotting experiments using the polyclonal antibody specific for HMW-GSs were conducted [19].

Comparison of Deduced Amino Acid Sequences and Clustering Analysis

The amino acid sequences were deduced from the HMW-GS genes of *Ag. intermedium* by ORF Finder program (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) and compared to previously published HMW-GSs through multiple alignments using DNA-MAN 5.2.2 (Lynnon BioSoft, http://www.lynnon.com/). Homology tree with observed divergency was used for clustering analysis based on the multiple alignments. Some of representative HMW-GSs from *Triticum* genus, 1Ax2* (M22208), 1Bx7 (X13927), 1Dx5



Figure 2. Amplification of complete ORFs coding for HMW-GSs of *Ag. intermedium.* Lane 1 is DNA marker III from TIANGEN and lane 2~7 shows the PCR bands from *Ag. intermedium* seedlings corresponding to the seeds of lane 2~7 in Figure 1. doi:10.1371/journal.pone.0087477.q002

(X12928), 1Ay (AJ306977, from *T. timopheevi*), 1By9 (X61026) and 1Dy10 (X12929), were selected for comparing to those of the *Ag. intermedium*.

MALDI-TOF-MS

The seed protein samples prepared from Ag. intermedium were separated using 10% SDS-PAGE with the Tris-Glycine-SDS running buffer. After Coomassie Brilliant Blue R-250 staining, the native target subunits were manually excised from the gel, followed by in gel digestion with trypsin [23]. The samples of the digested protein were analyzed in an Autoflex MALDI-TOF-MS (Matrix-Assisted Laser Desorption/Ionization Time of Flight Mass Spectrometry) (Bruker Daltonics) with a mass range of m/z 1000 to 15000 Da. The MALDI-TOF-MS measured peptide mass spectra were compared to the calculated mass spectra, which were predicted based on the trypsin-digested amino acid sequence deduced from the target gene using the bioinformatic program PeptideMass (http://web.expasy.org/peptide_mass/). The MALDI-TOF-MS analysis was repeated three times using the protein samples prepared from separate batches of seeds.

Phylogenetic Analysis

To analyze the phylogenetic relationship between Ag. intermedium and the other species in Triticeae, HMW-GS sequences of representative species in the tribe were collected from GenBank (http://www.ncbi.nlm.nih.gov/genbank/). For calculating evolutionary distances and constructing phylogenetic trees, Clustal W program [24] and the MEGA program (Version 5.2, http://www. megasoftware.net/) were used.

Results

Identification of HMW-GSs in Ag. intermedium

SDS-PAGE protein profiles of individual seeds of Ag. intermedium showed that quite a few protein bands were distributed in HMW-GS region of common wheat variety Chinese Spring (lanes $2\sim7$ in Figure 1A). For identifying HMW-GSs, Western blotting experiment with the polyclonal antibody specific for HMW-GS was carried out. The Western blotting patterns indicated that most of the seeds used in this study expressed 3 to 5 HMW-GSs (lanes $3\sim7$ in Figure 1B), but only one possessed seven HMW-GSs (lane 2 in Figure 1B). By comparing the bands detected by Western blotting, seven unique HMW-GSs with distinct electrophoretic mobilities were identified in the Ag. intermedium total seed protein extracts (more than 100 seeds tested). This means that the seven HMW-GSs expressed in the seed shown in lane 2 of Figure 1 covered those from other seeds. Thus, they probably represent all of the HMW-GSs from the three genomes of the Ag. intermedium line

Table 2. Properties of the primary structure of the HMW-GS from *Ag. intermedium* in comparison with those of the representatives of wheat HMW-GS.

Subunit	ORF ^a	N-tern	ninal	C-term	inal	Rep. de	omain	Total		Renamed subunit	Gene
		Size	Cys ^b	Size	Cys ^b	Size	Суз ^ь	Size	Cys ^b	_	
1Ai1	2442	86	3	42	1	665	0	793	4	1Aix2	Glu-1Aix2
1Ai2	2046	104	5	42	1	515	0	661	6	1Aiy1	Glu-1Aiy1
1Ai3	1908	104	5	42	1	469	1	615	7	1Aiy2	Glu-1Aiy2
1Ai4	1770	81	3	42	1	446	2	569	6	1Aix1	Glu-1Aix1
1Ai5	1500	81	3	42	1	356	0	479	4	1Aix3	Glu-1Aix3
1Ai6	1386	81	3	42	1	318	0	441	4	1Aix4	Glu-1Aix4
1Ai7	1149	105	5	42	1	215	2	362	8	1Aiy3	Glu-1Aiy3
1Ax2*	2445	86	3	42	1	666	0	794	4		
1Bx7	2367	81	3	42	1	645	0	768	4		
1Dx5	2544	89	3	42	1	696	1	827	5		
1Ay	1761	104	5	42	1	420	0	566	6		
1By9	2115	104	5	42	1	538	1	684	7		
1Dy10	1944	104	5	42	1	481	1	627	7		

^aThe nucleotide number of Open Reading Frame;

^bThe number of Cys.

doi:10.1371/journal.pone.0087477.t002



Figure 3. SDS-PAGE (A) and Western blotting (B) analysis of HMW-GSs from *Ag. intermedium* **and bacterial expression products.** Lane 1 is HMW-GSs from common wheat variety Chinese Spring, the four expressed HMW-GSs are noted on left; Lane 2 is native HMW-GS from the seed of *Ag. intermedium* same as the lane 2 in Figure 1, the seven expressed HMW-GSs are marked with triangles; Lane 3, 5, 7, 9, 11, 13 and 15 are total cell proteins from IPTG induced *E. coli* containing pET-*Glu-1Ai1*, pET-*Glu-1Ai2*, pET-*Glu-1Ai3*, pET-*Glu-1Ai5* pET-*Glu-1Ai6* and pET-*Glu-1Ai7*, respectively, whereas the dextral lanes for each of them shows the total cell proteins from their bacterial cells without induction of IPTG. The seven expressed target proteins in *E. coli*, which were detected by SDS-PAGE (marked with arrows in A) and were confirmed by Western blotting (lanes 3, 5, 7, 9, 11, 13 and 15 in B), share comparable electrophoretic mobility with those native HMW-GSs from *Ag. intermedium* (lane 2). doi:10.1371/journal.pone.0087477.g003

(E1E2X, 2n = 6x = 42) used in this study [25] and were tentatively designated as 1Ai1, 1Ai2, 1Ai3, 1Ai4 1Ai5, 1Ai6 and 1Ai7, respectively (from higher to lower molecular mass in lane 2 in Figure 1). In term of their electrophoretic mobilities, three of the seven subunits (1Ai1~3, marked with the solid triangles in Figure 1B) were comparable to the HMW-GSs from Chinese Spring, and the others (1Ai4~7, marked with hollow triangles in lane 2 of Figure 1B) were unusually small.

Isolation of HMW-GS Genes from Ag. intermedium

In genomic PCR experiments using the primers P1 and P2, tens of DNA fragments (lanes $2 \sim 7$ in Figure 2) were generated from *Ag. intermedium* individual plants corresponding to the representative seeds with different HMW-GS genotypes in Figure 1. After cloning these fragments into the plasmid vector pGEM-T, at least three positive clones for each insert were sequenced. The sequencing results at both ends showed that the positive clones of the seven inserts from the lane 2 of Figure 2 included those of the inserts from the other individual plants (lanes $3 \sim 7$ in Figure 2). Subsequently, the positive clones of seven inserts, designated as p1Ai1-2.4, p1Ai2-2.1, p1Ai3-1.9, p1Ai4-1.8, p1Ai5-1.5, p1Ai6-1.4 and p1Ai7-1.2, were completely sequenced by overlapping the subclones prepared from nested deletion method. Sequence analysis using the ORF Finder program (http://www.ncbi.nlm. nih.gov/projects/gorf/) indicated that the seven genes all possessed complete HMW-GS ORFs and were deduced to represent the ORFs of the 1Ai1, 1Ai2, 1Ai3, 1Ai4, 1Ai5, 1Ai6 and 1Ai7 subunits, respectively (lane 2 in Figure 1). According to the sequencing results, the lengths of the seven ORFs were in decreasing order 2442, 2046, 1908, 1770, 1500, 1386 and 1149 bp and encoded 793, 661, 615, 569, 479, 441 and 362 amino acid residues, respectively (Table 2). Of these, 1Ai4, 1Ai5, 1Ai6 and 1Ai7 ORFs were unusually small compared to the representatives of wheat HMW-GS genes. Furthermore, blast analysis showed that 1Ai4 ORF was completely identical to that of 1Aix1 (accession number DQ304542 in GenBank), an Ag. intermedium HMW-GS isolated from T. aestivum-Ag. intermedium addition lines TAI-13 [10], whereas the other six differed from

The subulity and subulity norm Agropyion internetion	HMW	Glutenin	Subunits	from	Agropyron	intermedium
--	-----	----------	----------	------	-----------	-------------

Α				
1Aix1	EGEASCOLOCERELEACROVVE	DERDA	28	
1Aix2	EGEAS GOLOCERELQERELKACCOVMDR	QLRDI	33	
1Aix3	EGEASGQLQCFRELEACRQVVDQ	QLRDA	28	
1Aix4 1Aiv1	EGEAS GLOCORELEASSFEACROWD COLAGWLPWSTGLOMRCCC	DIRDV	20	1Dy
1Aiy2	EGEASRCLQCFRELQESSLEACELVVDCQLAGELFWSTGLQMRCCC	LEDV	51	1Aix
1Aiy3	EGEASRQLQCPRELQESSLEACRQVVECQLAGCLPWSTGLQMRCCQ	QLRDV	51	ing
1Ay	EGEASKQLQCPRELQESSLEACRLVVDCQLASRLPWSTGLQMRCCQ	QLRDI	51	1Aiy
1Dv10	EGEASROLOCORPLOE SSLEAGROWN COLAGREPWSTGLOMRCCO		51	1D
1Ax2*	EGEASGQLQCERELQEHSLKACRQVVL	QLRDV	33	Iby
1Bx7	EGEASGQLQCEHELEACQQVVDQ	QLRDV	28	1Aiy
1Dx5	EGEAS EQLOCERELQERELKACCOVMD	QLRDI	36	
				1Ay
1Aix1	SPECRPVAVSFVARQYEQQTVVPFKCCSFYFCETTFFQQLQQRIFWGIP	TLL.R	81	1Aix
1Aix2	SPECHPVVVSFVAGQYEQQIVVPPKGGSFYFGETTEFQQLQQSILWRIP	ALL.K	86	
1Aix3	S PEORPVAVS FVARQYEQ CTW PPKCC SFYEGDAME FOOLOO RUFWCI P S DECERDVAVS FVARQYEQ CTW PDKCC SFYEGDAME FOOLOO PUFWCI D	PLL.R	81	1Aix
1Aiv1	SAKCRPVAVSCVAROYEC. TVVSPKCCSFYFCDTTHLOOLOCGIFWGTS	SCIVO	104	1Aix
1Aiy2	SAKCRPVAVSÇVARQYEÇ.TWYSPKCCSFYFCDTTHLQQLQCCIFWCTS	QVIQ	104	
1Aiy3	SANCRLIAVSCVARQYECQTAVPPKCGSFYFGDTTFFQQLQCRIFWGRS	SQIVQ	105	1Ax.
1Ay 1By9	SAKORPVALSQVARQYGQ, TAVPPKGCPTYHRDTYHLQQLQQGHFGGTS	SQIVQ	104	1 Aix
1Dy10	SAKCRSVAVSCVARCYEC. TVVPPKCCSFYFGETTELCOLOCGIFWCTS	QVIQ	104	
1Ax2*	SPECQPVGGGFVARQYECQVVVPPKGCSFYFGETTEFQQLQCSILWGIP	ALL.R	86	1Dx
1Bx7	SPECEPITVSFGTROVECOPVVPSKAGSFYFSDTTHSCOLOOMIFWGIP	ALL.R	81	1Bv
1DX5	SPECHEA AASEANGONDOOIMAEEKGESEALGEDIMEEGODOOKULMEID	ALL.K	89	IDA
P				
5				
1Aix1	SFYHVSVECCAPGLEVAKAQOPATQLSAMCRMEGGDALSASO 42			
1A1X2	SSYNVSADHOAASLKVAKAQQLAAQLEAMORLEGGLAISASO 42 SEVHUSADHOAASLKVAKAQQLAAQLEAMORLEGGLAISASO 42			
1Aix4	SFYHVSADHÇAASLKVAKAQQLAAQLFAMCRLECGDALSASQ 42			1Dv
1Aiy1	SFYHVSADQCAASLMVAKAQQLAAQLFAMCRLECSAALSASQ 42			iDy
1Aiy2	SFYHVSADQQAASPMVAKAQQPATQLFTVCRMEGCDALSASQ 42			1Aiy
1ALY3	SEVENS VDCC 7 AS DEVAKAHH DVACLETMCCMFCCDALSASO 42			
1Dy10	SEVHVSADCCAASPMVAKACCPATCLETVCEMECCEALSASC 42			1A _{1X}
1By9	NFYHVNTDCCTASPKVAKVQQPATQLFIMCFMEGGDALSASQ 42			1Ay
1Ax2*	SPHINSAPHQAASLKVAKAQQLAAQLFAMCRLEGGDALASQ 42			
1Dx5	SSYHVSVEHQAASLKVAKAQQLAAQLFAMCRLEGCEALSASQ 42			1By
				1Aix
				1Aix
				1Aix
				1Ax.
				1Dx
				1Aiy
				1Bx
				150
				1Aiy
C				
1Aix1	EC <mark>QGCQC</mark> HC <mark>FT</mark> SPQCPRQACQPGQCQQTGCSQQPGQGQQTGQAQCLG	47		
1Aix2	FCQCCQC YYFTSPQCSGQGCQIGQWLCPG	29		
1Aix3		23		1.4.1
1Ax2*	ERQGCQGYYFT SPQCSGQGQCLG	23		TAD
1Bx7	EC <mark>QGCQC</mark> YY <mark>FI</mark> SSQCSGQGHQSG	23		1Aiy
1Dx5	ECOCOCYVETSPOCPPOSCO			1.4.5
1Aiv2		29		IAIy
and a state V had	SCREACE YYERSPOC SCOCOPECCOSCCWOOPECCOOSCWOOPE SCREACE HORSPOC TCOACOPECCOISCVOOPECCOO	29 47 39		
1Aiy3	ECRECCE YYENSBOQ SGCGQDFGQCQQSGCWQQPEQGQQQSGCWQQPG ECRECCE HCENSBOQ TGCACQPGQCQQIGCVQQPEQGQQ ECCCE HCENSBOQ PRQACQPGQCQQTGCSQQPEQGQQTGQAQCPG	29 47 39 47		1Dy
1Aiy3 1Ay	CCCCCC YIEISPQC SCCCCQPCQCQCSCCWQQPCQCQQSCCWQQPC CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	29 47 39 47 35		1Dy
1Aiy3 1Ay 1Dy10 1By9		29 47 39 47 35 39 39		1Dy 1By
1Aiy3 1Ay 1Dy10 1By9	BCGCCC YY FISPOC SGCCCOPGCQQQCCCCGCCQCCCQCCCCCCCCCCCCCCCCC	29 47 39 47 35 39 39		1Dy 1By 1Ay
1Aiy3 1Ay 1Dy10 1By9	COCCCC YY FISPOC SCOCCOPCCCQSCCWQQPCQCQCSCCWQCPG CCCCCCCC HOSISPQCTCQACQPCQCQCCCCCCQCCQCQCQQC CCCCCCC HOSISPQCTCQACQPCQCQCCCCQQCCCQQCCQQCCQAQCPG CCCCCCC HFF1SCQACQPCQCQCICCAQQLCQCQ CCCCCCC HF51SCQACQPCQCQCICCVQQPCQCQQ CCCCCCC HCF1SPQCSCQACQPCQCQQICCVQPCQCQQ CCCCCCCCCCCCCCCCCQCCCQQCCQQCCQ	29 47 39 47 35 39 39		1Dy 1By 1Ay
1Aiy3 1Ay 1Dy10 1By9 1Aix1		29 47 39 47 35 39 39 93 57		1Dy 1By 1Ay 1Ais
1Aiy3 1Ay 1Dy10 1By9 1Aix1 1Aix2 1Aix3	EGGCCC YY PISPQC SCGCCQPGQCQQSCCWQQPCQGQQCCCGG EGGCCC HC PISPQC TCQAQQPCQCQCTCCVQPCQCQQ EGGCCC HC PISPQC TCQAQQPCQCQTCCVQPCQCQQ EGGCCC HC PISPQC TCQAQQPCQCQTCCVQQPCQCQQ EGGCCC HC PISPQC SCQAQQPCQCQTCCAQQTCCAQQCCGQ EGGCCC HC PISPQC SCQAQQPCQCQQTCCVQQPCQCQQ EGGCCC HC PISPQC SCQAQQPCQCQQTCCVQQPCQCQQ EGGCCC HC PISPQC SCQAQQPCQCQQTCCVQQPCQCQQ EGGCCC HC PISPQC TCQAQQPCQCQQTCCVQQPCQCQQ EGGCCC HC PISPQC TCQAQCQCQCCQQCCQQQQ EGGCCC HC PISPQC TCQAQCQCQCQCCQQQQQ EGGCCC HC PISPQC TCQAQCQCQQTCCVQQPCQQQ EGGCCC HC PISPQC TCQAQCQCQCQQQQQ EGGCCC HC PISPQC TCQAQCQCQCQQQQQQQQQ EGGCCC HC PISPQC TCQAQCQCQQCQQQQQQQQQQQQQQQQQQ EGGCCC HC PISPQC TCQAQQCQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ	29 47 39 47 35 39 39 93 57 45		1Dy 1By 1Ay 1Ais 1Ais
1Aiy3 1Ay 1Dy10 1By9 1Aix1 1Aix2 1Aix3 1Aix4	GGGCCG YY PHEPQQ SGQCQQPGQCQQGCGCQQQCQCGQQQQCPG GGGCCG HC PHEPQQ PRQAQQPGQCQQTGCVQPCQCQQ GGGCCG HC PHEPQQ PRQAQQPGQCQTGCVQPCQCQQ GGGCCG HC PHEPQQ PRQAQQPGQCQTGCVQQPCQCQQ GGGCCG HC PHEPQQ PRQAQQPCQCQTGCVQQPCQCQQTGQAQCPG GGGCCG HC PHEPQQ PRQAQQPCQCQTGCVQQPCQCQQTGQAQCPG GGGCCG HC PHEPQQ PRQAQQPCQCQTGCVQQPCQCQQ GGGCCG HC PHEPQQ PCCQC SGCCQQTGCVQQPCQCQQ GGCCCG HC PHEPQQ PCCQC SGCCQQTGCVQQPCQCQQ GGCCCG HC PHEPQQ PCCQC SGCCQC GCGCCG YP PHEPLQ PGCCQ SGCCQC SGCCQQ SGCCQC GCQCG YY PHEPPQ PGCCQ SGCCQC GCQCG YY PHEPPQ PGCCQ GCQCG YY PHEPPPQ PGCCQ GCQCG YY PHEPPPQ PGCCQ GCQCG YY PHEPPPPQ PGCCQ GCQCG YY PHEPPPPQ PGCCQ GCQCG YY PHEPPPPQ PGCCQ GCQCG YY PHEPPPPQ PGCCQ GCQC YY PHEPPPPP GCQC YY PHEPPP GCQC YY PHEPP GCQC Y PHEP	29 47 39 47 35 39 39 39 93 57 45 45		1Dy 1By 1Ay 1Ais 1Ais
1Aiy3 1Ay 1Dy10 1By9 1Aix1 1Aix2 1Aix3 1Aix4 1Ax2*		29 47 39 47 35 39 39 39 93 57 45 45 45		1Dy 1By 1Ay 1Ais 1Ais 1Ais
1Aiy3 1Ay 1Dy10 1By9 1Aix1 1Aix2 1Aix3 1Aix4 1Ax2* 1Bx7 1Dx5	CGCCCC YY PISIPOC SCCCCQ FGCCQQCSCCWQQ PCQCQQCSGQWQC PG CGCCCC HC PISPQC TCQAQCPCCCQQ TCCVQQ PCQCQQ TCQAQC PG CGCCCC HC PISPQC SCQAQCQ CCQ TCCXQQ PCQCQQ TCQAQC PG CGCCCC HC PISPQC SCQAQCQ CCQ TCCXQQ PCQCQQ TCQAQC PG CGCCCC HC PISPQC SCQAQCQ CCQ TCCXQQ PCQCQQ TCQAQC PG CGCCCC HC PISPQC SCQAQCQ CCQ CCQ CCQ CQ CQQ PCQCQQ TCQAQC PC CGCCCC HC PISPQC SCQAQCQ CCQ CCQ CQ CQ CQ CQ CQ CCQ CCQ PC CGCCCC HC PISPQC TCQAQ CQ	29 47 39 47 35 39 39 93 57 45 45 45 45 57		1Dy 1By 1Ay 1Ais 1Ais 1Ais 1Ais
1Aiy3 1Ay 1Dy10 1By9 1Aix1 1Aix2 1Aix3 1Aix4 1Ax2* 1Bx7 1Dx5 1Aiy1	CGCCCC YY 113 PQC SGCCCQ PGCCQ PGCCQ SGCWQ PEQGQQ SGCWQ PEG CGCCCC HC 113 PQC TCQAQCPCQCQ IGCVQ PEQGQQ TGQAQCPG CGCCCC HC 113 PQC TGQAQCPCQCQ IGCVQ PEQGQQ	29 47 39 47 35 39 39 93 57 45 45 45 57 75		1Dy 1By 1Ay 1Ais 1Ais 1Ais 1Aix 1Aiy 1Bx
1Aiy3 1Ay 1Dy10 1By9 1Aix1 1Aix2 1Aix3 1Aix4 1Ax2 1Aix4 1Bx7 1Dx5 1Aiy1 1Aiy2	GGCCC YY JISPQC SGCCCQPGQCQQSGCWQQPCQGQQQSGQWQCPG GGCCC HCJISPQC TCQAQQPCQCQQIGCVQPCQCQQTGQAQQPG GGCCC HCJISPQC TCQAQQPCQCQQIGCVQQPCQCQQTGQAQQPG GGCCC HCJISPQC TQAQQPCQCQQIGCVQQPCQCQQTGQAQQPG GGCCC HCJISPQC TQAQQPCQCQQIGCVQQPCQCQQTGQAQQPG GGCCC HCJISPQC TQAQQPCQCQQIGCVQQPCQCQQ GGCCC HCJISPQC TQAQQPCQCQQIGCVQQPCQCQQ GGCCC HCJISPQC TQAQQPCQCQQIGCVQPCGQQ GGCCC HCJISPQC TGQAQQPCQCQQIGCVQPCQCQQ GGCCC HCJISPQC TGQAQQPCQCQQIGCVQPCQCQQ GGCCC HCJISPQC TGQAQQPCQCQQIGCVQPCQCQQC GGCCC HCJISPQC TGQAQQPCQCQQIGCVQPCQCQCGC GGCCC HCJISPQC TGQAQQPCQCQQIGCVQPCQCQCGC GGCCC HCJISPQC TGQAQQPCQCQQIGCVQPCQCQCGC GGQCC YPPTSIC FGCGC GGQC YPPTSIC FGCGC GGQC YPPTSIC FGCGC GGQC YPPTSIC FGCGC GGQC YPTSIC FGCGC GGQC YPTSIC FGCGC GGQC YPTSIC FGCGC GGCC YPTSIC FGCGC GGCC YPTSIC FGCGC GGC YPTSI	29 47 39 47 35 39 39 39 57 45 45 45 57 75		1Dy 1By 1Ay 1Ais 1Ais 1Ax 1Ax 1Ax
lAiy3 lAiy3 lAy lDy10 lBy9 lAix1 lAix2 lAix3 lAix4 lAx2* lBx7 lDx5 lAiy1 lAiy2 lAiy3 lAy	GGCCC YY PISPOR SGCCCQPGQCQQSGCWQQPCQGQQQSGCWQPFG GGCCC HCSISPQCTCQAQCFCQCQCCCCQQCCCQQCCQQQ GGCCC HCSISPQCTCQAQCFCQCQCCCQQCGQQQCGQQQCGQAQQCFG GGCCC HCSISPQCTQQAQCGQQCGQQCGQQQCGQQQC GGCCC HCSISPQCTQQAQCGQQCGQQCGQQQCGQQQ GGCCC HCSISPQCTQQAQCGQQCGQQCGQQQCGQQQ GGCCC HCSISPQCTQQAQCGQQCGQQCGQQQQ GGCCC HCSISPQCTQQAQQCQQCGQQCGQQQ GGCCC HCSISPQCTQQAQQCQQCGQQQC GGCCC HCSISPQCTQQAQQCQCGQQQCGQQQ GGCCC HCSISPQCTQQAQQCQCGQQCGCQQQC GGQCCY PTSLC AGCGCQ SGQCQC GGQCCY PTSSLC AGCGCQ GGQCCY PTSSLC AGCGQCGQCGQC GGQCY PTSSLC AGCGQCGQCGQC GGQCY PTSSLC AGCGQC SGQCQC GGQCY PTSSLC AGCGQ SGQCQC GGQCY PTSSLC AGCGQ SGQCQC GGQCY PTSSLC AGCGQ SGQCQC GYN GYN GYN PTSSLC AGCGC SGCQCQC GYN PTSSLC AGCGC SGCQCQC GYN PTSSLC AGCGCY SGCQCY FGCGRUPCHCQCQSGCCQC GYN PTSSLC AGCGCY SGCQCY FGCGRUPCHCQCQSGCCQCY D GGQCYY PTSSLC AGCGCY FGCGCY FGCGRUPCHCQCQSGCCQCY D GGQCYY PTSSLC AGCGCY FGCGCY FGCGCUCSGCQCY CONC GYN PTSSLC AGCGCY FGCGCY FGCGCUCSGCCQCY D GGCGYY FTSSLC AGCGCY FGCGCY FGCGCUCSGCCY TD	29 479 479 39 39 39 39 57 45 45 57 51 93 75		1Dy 1By 1Ay 1Ais 1Ais 1Ax 1Aiy 1Bx 1Ais
lAiy3 lAiy3 lAy lDy10 lBy9 lAix1 lAix2 lAix3 lAix4 lAx2* lBx7 lDx5 lAiy1 lAiy3 lAiy3 lAy lDy10	GGCCC YY PISPOR SGCCCQPGQCQQSGCWQQPCQGQQCQCQPG GGCCC HCSPBQCTCQAQCPCQCQCCCCQQCCCQQCCQQQCCQQQCCQQC GGCCC HCSPBQCTCQAQCPCQCQCCCQQCCQQCQQCQQQCCQAQQC GGCCC HSPA SGQAQCPCQCQCCQCCQCQCQQCCQQCCQAQQCCQAQQC GGCCC HCSPBQCTCQAQCQCQCCQQCQQCCQQQC GGCCC HCSPBPQTCQAQCPCQCQCCCQCQCCQQQQ GGCCC HCSPBPQTCQAQQCQCQQCCQQQQ GGCCC HCSPBPQTCQAQQCQCQQCCCCCCC GCQCCYPTSLC AGCGCQSCQQCCQQCCCCCCCC GCQCCYPTSPC FGCCQ GCQCCYPTSPC FGCCQ GCQCCYPTSPC FGCCQ GCQCYPTSPC FGCCQ GCQCCYPTSPC FGCCQ GCQCCYPTSPC FGCCQ GCQCCYPTSPC FGCCQ GCQCYPTSPC FGCCQC GCQCYPTSPC FGCCQC GCQCYPTSPC FGCCQC GCQCYPTSPC FGCCQC GCQCYPTSPC FGCCQC GCQCYPTSPSPC FGCCQC	29 47 39 47 35 39 93 57 45 45 45 57 75 83 75 83 75		1Dy 1By 1Ay 1Ais 1Ais 1Ax 1Ais 1Bx 1Bx 1Ais



Figure 4. Alignments and clustering analyses based on N-terminal (A), C-terminal (B) and the last 93 residues in the repetitive region (C) of the HMW-GSs from Ag. intermedium and several representative HMW-GSs from Triticum genus. Noticeably, the subunit

1Aix1 possesses a N-terminal clustered to x-type subunits (A) and a C-terminal and last part of repetitive region clustered to y-type subunits (B and C). Conversely, the subunit 1Aiy1 possesses a N-terminal more similar to y-type subunits (A) and a C-terminal and last part of repetitive region more similar to x-type subunits (B and C). The default parameters were used for full alignment and clustering analysis of sequences by aid of DNAMAN version 5. 2. 2.

doi:10.1371/journal.pone.0087477.g004

previously published HMW-GSs (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

Expression in *E. coli* and Western Blotting Analysis of HMW-GSs

To confirm the cloned Ag. intermedium HMW-GS genes, they were subjected to bacterial expression and Western blotting analysis. For bacterial expression, the nucleotide sequence encoding the signal peptide was removed from the cloned ORF by PCR mutagenesis. This ensured the synthesis of the mature proteins of the cloned genes, which should possess comparable electrophoretic mobility with those of the native subunits from the seeds. Seven expression constructs, pET-Glu-1Ai1, pET-Glu-1Ai2, pET-Glu-1Ai3, pET-Glu-1Ai4, pET-Glu-1Ai5 pET-Glu-1Ai6 and pET-Glu-1Ai7, were created for expressing the mature proteins of the 1Ai1, 1Ai2, 1Ai3, 1Ai4, 1Ai5, 1Ai6 and 1Ai7 subunits in E. coli cells, respectively. In SDS-PAGE analysis showed that the electrophoretic mobility of the seven proteins directed by the seven expression constructs above, (lanes 3, 5, 7, 9, 11, 13 and 15 in Figure 3A shown by triangles) were in accordance with their native subunits extracted from the seed (lane 2 in Figure 3A), respectively. Furthermore, the seven bacterially expressed subunits were exhibited strong reactions to the polyclonal antibody specific for HMW-GSs in Western blotting (Figure 3B).

Characteristics of the Amino Acid Sequences Deduced from *Ag. intermedium* HMW-GS Genes

Analysis of amino acid sequences derived from the cloned seven genes showed that HMW-GSs in Ag. intermedium possessed the same primary structure as known HMW-GSs, with the conserved signal peptide (which was excised from the mature proteins), Nand C-terminal domains, and the central repetitive region composed of tri-, hexa- and nona-peptides motifs [3]. Some of their properties, in comparison to those of the representatives of wheat HMW-GSs, are summarized in Table 2. There were 3 or 5 cysteine residues in their conserved N-terminal region and 1 cysteine residue in their conserved C-terminal region, and 1Ai4 and 1Ai7 had at least one extra cysteine residue in their repetitive regions compared to the representatives of wheat HMW-GSs (Table 2). Judging from the size of conserved N-terminal domains (especially the number of the cysteine residues), 1Ai1, 1Ai4, 1Ai5 and 1Ai6 belonged to x-type, whereas 1Ai2, 1Ai3 and 1Ai7 fell into y-type. The ORF of 1Ai4 was completely identical to Glu-1Aix1 (DQ304542), an Ag. intermedium HMW-GS gene isolated previously from the T. aestivum-Ag. intermedium addition lines TAI-13, so it still follows the previous name [10]. Here we designated the genes of 1Ai1, 1Ai2, 1Ai3, 1Ai5, 1Ai6 and 1Ai7 as *Glu-1Aix2*, Glu-1Aiy1, Glu-1Aiy2, Glu-1Aix3, Glu-1Aix4 and Glu-1Aiy3, respectively, according to their size and attributive types. Accordingly, the HMW-GSs 1Ai1~7 in Figure 1 were renamed 1Aix2, 1Aiy1, 1Aiy2, 1Aix1, 1Aix3, 1Aix4 and 1Aiy3, respectively (Table 2). The six unique HMW-GS genes cloned in this study, Glu-1Aix2, Glu-1Aiy1, Glu-1Aiy2, Glu-1Aix3, Glu-1Aix4 and Glu-1Aiy3 have been submitted to GenBank with accession numbers, EF105400, EF105401, EF105402, EF105403, EF105404 and EF105405, respectively.

Except for 1Aiy1 and 1Aix1, the classificatory affiliations of the other subunits were confirmed by alignment and clustering

analysis based on the conserved N- and C-terminal domains (Figure 4A and 4B). In the case of 1Aiy1, its N-terminal domain displayed higher similarities to those of y-type subunits, whereas a reverse pattern was found in its C-terminal domain (designated as yx-type). Conversely, 1Aix1 comprised the N-terminal of x-type and C-terminal of y-type (designated as xy-type), just as identified previously [10]. Considering that the sizes of the C-terminal amino acid sequences are too short to be used reliably in alignment and clustering analysis (Figure 4B), the near C-terminal portion of the repetitive domain of the two hybrid subunits were compared with those of the representatives of HMW-GSs (Figure 4C). The alignment analysis showed the near C-terminal portion of the repetitive domain of 1Aiy1 displayed higher similarities to those of x-type subunits than y-type subunits, whereas that of 1Aix1 showed higher similarities to those of y-type subunits than x-type subunits (left panel in Figure 4C). Consequently, in the cladogram developed from near C-terminal domain sequences, 1Aiy1 clustered with x-type subunits whereas 1Aix1 fell into the group of y-type subunits (right panel in Figure 4C), which is in line with the result from their C-terminal domain sequences. Moreover, the nonapeptide GHCPTSPQQ, which has not been found in any of the previously reported x-type subunits but existed in near Cterminal portion of the repetitive domain of the majority of the ytype subunit characterized so far, was present in the repetitive domain of 1Aix1 (left panel in Figure 4C). Conversely, 1Aiy1 lacks GHCPTSPQQ-like nonapeptide in its repetitive domain. These results further validated that both subunits belong to the hybrid types.

Here the native 1Aiy1 subunit also was subjected to MALDI-TOF-MS analysis which is effective in gaining structural information of the HMW-GSs directly isolated from seeds [26,27]. By repeated MALDI-TOF-MS experiments, the molecular mass of sixteen peptides from 1Aiy1 could be reliably determined (Table 3), and their molecular mass values matched closely with the calculated values (columns 1 and 2 in Table 3). Together, the sixteen peptides covered the complete portion of the amino acid sequence (minus the signal peptide) deduced from *Glu-1Aiy1* (column 3 in Table 3). Consequently, the peptide mass fingerprint of native 1Aiy1 indicated that the deduced amino acid sequence from *Glu-1Aiy1* was accurate and really to be a true representation of the native subunit.

Phylogenetic Analysis

Along with the development of molecular biology, the study of evolutionary biology demands for more elaborate and comprehensive data. Comparative genetics demonstrated that chromosome group 1 is very well conserved in *Triticeae* [28]. Moreover, the gene-rich region in *Glu-1* loci of chromosome group 1 indicated very good microcolinearity [29]. This region is therefore a good target to study evolution events at the molecular level. Recently, the genomic evolution of common wheat was surveyed through the comparative genomics of *Glu-1* loci [29,30]. In addition, a lot of research showed that the coding sequences for HMW-GSs from *Glu-1* loci are phylogenetically informative and even were used to estimate the divergence time of genomes [9,31,32,33]. Most importantly, a large amount of information about HMW-GS genes from *Triticeae* is already available from other studies. Here we created a phylogenetic tree (Figure 5) based on the conserved

Measured mass(M+H) ⁺ Location Missed cleave Peptide subset mass(M+H) ⁺ Calculated mass(M+H) ⁺ Location Missed cleave using the use 1406.118 1405.685 1-12 2 EGESSFL 1298.749 1298.568 13-23 0 ELGESSFL 2313.18 24-43 0 QVVDQU 1286.759 1250.598 3-4-43 0 QVVDQU 1186.176 1185.652 55-65 0 QVDQU 1186.176 1185.652 55-65 0 QVDQU 1186.176 1185.652 66-75 0 QCGUNC 1186.176 1185.652 2238.069 115-136 0 QGSYYPG 2287.455 2288.069 115-136 0 QCGUNC MGEGGG 287.455 2288.069 115-136 0 QGSYYPG 287.455 2288.069 137-179 0 QGSYYPG 287.455 2288.069 137-133 0<	
112 2 EGEASRKL 106.118 1405.685 1-12 2 EGEASRKL 1298.749 1298.568 13-23 0 ELGESSFK 1298.749 1298.568 13-23 0 ELGESSFK 1298.749 1298.568 13-23 0 ELGESSFK 1250.554 1250.588 44-54 1 CCQURD 1186.176 1185.652 55-65 0 QWDOOL 1186.176 1185.652 55-65 0 QYDOOL 1186.176 1185.652 66-75 0 QYDOOL 1186.176 1178.605 66-75 0 QYDOOL 1186.145 4771.798 4771.798 0 QFGOOC 2287.455 2288.069 115-136 0 QGSYYPG 2787.458 4770.21 137-179 0 QGSYYPG 2854.134 6852.202 137-179 0 QGNOGO 10014.945 10016.624 241-332 0 QGNOGO	Peptide sequence predicted Missed cleavge using the bioinformatic program Peptide Mass (with trypsin digestion)
1406.118 1405.685 1-12 2 EGEASRIL 1298.749 1298.568 13-23 0 ELGESSFEJ 1298.749 1298.568 13-23 0 ELGESSFEJ 1298.745 2313.18 24-43 0 QVDQQL 1250.558 1250.598 44-54 1 CCQURD 1250.156 1186.176 1185.652 55-65 0 QVDQQL 1186.176 1185.652 55-65 0 CRPVAVSC 1186.176 1185.652 55-65 0 QYDQUL 1186.176 1185.652 55-65 0 QYEQTVSC 2287.455 2288.069 115-136 0 QGSYPGG 2287.455 2288.069 137-179 0 QGSYPGG 2287.455 2288.069 137-179 0 QGSYPG 2171.798 4770.1 137-179 0 QGSYPG 285.4134 6852.202 180-240 0 QGSYPG 10014.945 10016.624 241-3	
12-98.749 12-98.568 13-23 0 ELOESSFE 2313.545 2313.18 24-43 0 OWDOQL 213.545 2313.18 24-43 0 OWDOQL 1250.654 1250.598 44-54 1 CQUDQL 1186.176 1185.652 55-65 0 OWDQQL 1186.176 1185.652 55-65 0 CRPVAVSC 1186.176 1185.652 55-65 0 OYDQUL 1186.176 1185.652 55-65 0 OYDQUL 1186.176 1178.605 56-75 0 OYDQUL 2451.455 2288.069 115-136 0 OSTOPOR 2471.798 4770.21 137-179 0 OGSTOPOR 2471.798 4770.21 137-179 0 OGSTOPOR 2471.798 4770.21 137-179 0 OGSTOPOR 241.45 186-240 0 0 OGSTOPOR 258.434 6852.202 188-240 0	2 EGEASRKLQCER
213.545 2313.18 24-43 0 QWDQQL 1250.654 1250.598 44-54 1 CCQQIRD 1186.176 1185.652 55-65 0 CRPVASC 1186.176 1185.652 55-65 0 CRPVASC 1186.176 1185.652 55-65 0 CRPVASC 1186.176 1185.652 56-75 0 CRPVASC 1178.789 1178.665 66-75 0 OYEQTVSC 2451.455 4250.046 76-114 0 CGSTYPG 2587.455 2288.069 115-136 0 OYEQTVSC 2471.798 4770.21 137-179 0 OGSTYPSC 2587.455 2288.069 1157-136 0 OGSTYPSC 2654.134 6852.202 1337-179 0 OGSTYPSC 210016.624 10016.624 137-1332 0 OGSTSYPSC 13765.658 13762.182 137-1332 0 SWOGSC 13765.658 13762.182 333-457	0 ELQESSFEACR
1250.654 1250.598 44-54 1 CCQOLRD 1186.176 1185.652 55-65 0 CRPVASC 1178.789 1185.652 55-65 0 CRPVASC 1178.789 1178.605 66-75 0 CRPVASC 2171.789 1178.605 66-75 0 CRPVASC 2287.455 2288.069 115-136 0 GGSFYPGG 2717.798 4770.21 137-179 0 QGPGGC 4771.798 4770.21 137-179 0 QGPGGC 10014.945 10016.624 180-240 0 QGPGGC 10014.945 10016.624 241-332 0 QGPGGC 13765.658 13762.182 333-457 0 QGPGGC 13765.658 13762.182 333-457 0 GGSYPT 13765.658 13762.182 333-457 0 GGSYPT 13765.658 13762.182 333-457 0 GGSYPT 14046.246 14045.376 645.560	0 QVVDQQLAGWLPWSTGLQMR
1186.176 1185.652 55-65 0 СRPVAVSC 1178.789 1178.605 66-75 0 QYEQTVVS 1178.789 1178.605 66-75 0 QYEQTVVS 1178.789 1178.605 66-75 0 QYEQTVVS 4251.455 4250.046 76-114 0 GGSYYPG 4251.455 4250.046 76-114 0 QGSYYPG 4251.455 137-179 137-179 0 QGSYYPG 4771.798 4770.21 137-179 0 QGSYYPG 4771.798 4770.21 137-179 0 QGSYYPG 4771.798 4770.21 1360-240 0 QGSYYPG 4771.798 10016.624 180-240 0 QGYAPG 10014.945 10016.624 241-332 0 QGYAPG 13765.658 13762.182 333-457 0 QHGGGG 13765.658 13762.182 333-457 0 GGGGGY 13765.658 13762.182 333-457	1 CCQQLRDVSAK
1178.789 1178.605 66-75 0 QYEOTWG 4251.455 4250.046 76-114 0 GG5FYGE 4251.455 4250.046 76-114 0 GG5FYGE 4251.455 2288.069 115-136 0 GG5FYFGE 2287.455 2288.069 115-136 0 GG5FYFGE 4771.798 4770.21 137-179 0 QG5GQGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	0 CRPVAVSQVAR
4251455 4250.046 76-114 0 6655Y0E 2287.455 2288.069 115-136 0 055YYPG 2287.455 2288.069 115-136 0 055YYPG 4771.798 4770.21 137-179 0 0 05YYPG 6854.134 6852.202 180-240 0 0 00YOGG 10014.945 10016.624 241-332 0 0 00YOGG 13765.658 13762.182 333-457 0 0 5WGGGGG 13765.658 13762.182 333-457 0 0 00FGGGG 481.774 5479.452 458-507 0 0 60GGYPT 481.774 1406.246 14045.376 0 60GGYPT 1400.045 1400.645 1401.73 638-650 0 0 60GGGGG	0 QYEQTVVSPK
2287.455 2288.069 115-136 0 QGSYYPG 4771.798 4770.21 137-179 0 WOEPGGG 6854.134 6852.202 137-179 0 WOEPGGG 6854.134 6852.202 180-240 0 WOEPGGG 10014.945 10016.624 241-332 0 QQPVGGG 13765.658 13762.182 333-457 0 QGGYYPT 5481.774 5479452 458-507 0 GQGGYYPT 14046.246 14045.376 568-637 0 GQGGYYPT 1400.645 1400.713 638-650 0 GQGGYYPT	0 GGSFYPGETTPLQQLQQGIFWGTSSQTVQGYYPSVTSPR
4771.798 4770.21 137-179 0 WQEPGGG 6854.134 6852.202 180-240 0 QQPVGG 10014.945 10016.624 180-240 0 QQPVGG 1376.658 13762.182 241-332 0 SWGGGG 13765.658 13762.182 333-457 0 PTEPHQFC 13765.658 13762.182 333-457 0 QGGGG 413765.658 13762.182 333-457 0 QGGGG 4140.6546 14045.376 458-507 0 GQGGGY 4100.645 14045.376 568-637 0 GQGGYY 4100.645 1400.713 638-650 0 GQUAAC	0 QGSYYPGQASPQQPGQGQPGK
6854.134 6852.202 180-240 0 QQPVGGG 10014.945 10016.624 241-332 0 SWGGGGG 13765.658 13762.182 333-457 0 SWGGGGG 13765.658 13762.182 333-457 0 QPPGGG 5481.774 5479.452 458-507 0 GQGGGGG 14046.246 14045.376 508-637 0 GQGGGG 1400.645 1400.713 638-650 0 GQGGAY	0 WQEPGQGQQGYYPTSLQQPGQGQQGQQGQQGQYPTSPQHPGQR
10014.945 10016.624 241–332 0 SWGQGQ 13765.658 13762.182 333-457 0 <t< td=""><td>0 QQPVQGQQIGQGQQPEQGQQPGQWQQGYYPTSPQQPGQQQQGQQPGQQGQQGYYLIFSR</td></t<>	0 QQPVQGQQIGQGQQPEQGQQPGQWQQGYYPTSPQQPGQQQQGQQPGQQGQQGYYLIFSR
13765.658 13762.182 333-457 0 QHPGQGG 5479.182 333-457 0 QHPGQGG 5481.774 5479.452 458-507 0 GQCYVPT 5481.774 5479.452 458-507 0 GQCGV 14046.246 14045.376 508-637 0 OPGQGV 1400.6245 1400.713 638-650 0 AQQLAAC	0 PTFPHQPGQGQQGQQGQQQQQQQ PTFPHQPGQGQQQQQ
5481.774 5479.452 458-507 0 GQSGGW 14046.246 14045.376 508-637 0 QPGQGY 1400.645 1400.713 638-650 0 AQQLAAG	0 QHPGQGQQGYYPTFSQQPGQWQQSGQGHQGYYPTSPQQSGQGQQGYYPTFSAAPGQGYYPTFSHQSGQQQPGQQQQ GQGYYPTFPQLPGQGQQGYYPTYPQQSGQQQPGQWQQSGQWQQPGR
14046.246 14045.376 508-637 0 QPGQGYY 6QWQQPt 6QWQQPt 6QWQQPt 6QWQQPt 1400.645 1400.713 638-650 0 AQQLAAC	0 GQQSGQWQQPGQGQQSGQWQQPGQGQQGYYPTSPQQSGQGQQGYYPTSPR
1400.645 1400.713 638-650 0 AQQLAAQ	0 QPGQGYYPTFPHQSGQGQQPGQGQQPGQGQYYPTFPQPPGQGQQGYYPTSPQQSGQQQGQQGQQSGQWQQPGQGQQS GQWQQPGQQQGYTTFPQQPGQGQQLGQEQQGYNSPYHVSAEQQAASLMVAK
	0 AQQLAQLPAMCR
1035.351 1035.516 651–661 0 LEGSAALS	0 LEGSAALSASQ

ч



Figure 5. Phylogenetic tree of *Thinopyrum intermedium* (= *Ag. intermedium*) and some representative HMW-GSs from *Triticeae*.

This phylogenetic tree was constructed with Maximum Likelihood Estimation method based on the nucleotide sequences encoding signal peptide and N-terminal conserved region of HMW-GSs plus the next three repeat units, one dodecapeptide, one undecapeptide and one hexapeptide repeat. D-hordein from barley was used as outgroup. The species names of HMW-GS genes in this figure are consistent with their accession names in GenBank, so here we replaced *Agropyron intermedium* with *Thinopyrum intermedium*. doi:10.1371/journal.pone.0087477.q005

sequences encoding the signal peptide and N-terminal region of *Ag. intermedium* HMW-GSs and the representative HMW-GSs from the major genomes in the tribe *Triticeae*, such as A, B, C, D, E, G, H, K, M, N, P, R, S, St, T, U, V etc [34]. In the phylogenetic tree, regardless of x- or y-types, the HMW-GSs from *Ag. intermedium* have closest distance with those from *Leymus* genus and *Thinopyrum elongatum* (= *Agropyron elongatum* = *Elytrigia elongata*), (subgroups A, B, C, D and E in Figure 5). Only subunit 1Aix2 has also close relationship with those from *Aegilops* with S^b or D genome (subgroup D in Figure 5). In contrast, none of *Ag. intermedium* HMW-GS was classified into close branch with those from *Thinopyrum bessaracicum* (E^b), *Thinopyrum junceum* (J) and *Pseudoroegneria* species (having St genome).

Discussion

The HMW-GSs and their coding genes from an Ag. intermedium line were comprehensively analyzed in this study. SDS-PAGE and Western blotting analysis of seed proteins indicated that the Ag. intermedium line aggregately expressed seven unique HMW-GSs. The putative genes encoding the seven subunits were isolated from genomic DNA with PCR method. The further experiments on bacterial expression and Western blotting analysis confirmed that the seven genes were active, leading to the synthesis of HMW-GSs in the seeds of the Ag. intermedium line. Amino acid sequence comparisons showed that the seven HMW-GSs from Ag. intermedium possessed high similarity to those from wheat (Figure 4), indicating that the Glu-1 loci of Ag. intermedium were closely related to those of wheat. However, the HMW-GSs 1Aix1, 1Aix3, 1Aix4 and 1Aiy3 subunits from Ag. intermedium are substantially smaller than those from Chinese Spring. Of them, 1Aiy3 probably is one of the smallest HMW-GSs published so far. Judging from the sizes of their HMW-GS genes (approximately



Figure 6. Illustration for the developmental mechanism of two hybrid HMW-GSs based on unequal double crossover hypothesis. The broken line box indicates the double crossover region. The xy and yx represent the hybrid subunit with 5'region of x-type and 3'region of y-type and the hybrid subunit with 5'region of y-type and 3'region of x-type, respectively.

doi:10.1371/journal.pone.0087477.g006

 $1.15\sim2.44$ Kb), the degree of diversity among the three genomes (E₁E₂X) of *Ag. intermedium* is considerably higher than that of wheat, from which the sizes of HMW-GS genes are about 2.0~2.5 Kb. In addition, there is large variation of genotypes even in a single *Ag. intermedium* line (Figure 1 and 2) because *Ag. intermedium* is in general cross-pollinated. Thus, *Ag. intermedium* may supply abundant variation of HMW-GSs for the quality improvement of wheat. Especially for these unusually small subunits, it may be interesting to research their effect on end-use quality of wheat.

It is well known that the HMW-GSs from not only common wheat but also its relatives are encoded by the Glu-1 locus containing the Glu-1-1 (x-type) and Glu-1-2 (y-type) genes. For common wheat with three genomes (A, B and D), there generally are $3\sim 5$ HMW-GSs for each line due to *Glu-1A-2* (v-type) gene is inactivated in most if not all cases (reviewed in [7]). Similarly, there are $3\sim 5$ HMW-GSs in most of seeds from the Ag. intermedium line (Figure 1). This result, together with previous findings, suggested that there was one genetic locus of Glu-1 on each chromosome 1 from three genomes of Ag. intermedium [16,35,36]. However, seven HMW-GSs were observed in one of the seeds from the Ag. intermedium line used in this study (lane 2 in Figure 1). We speculated that there was at least one heterozygous Glu-1 considering that Ag. intermedium was a naturally cross pollinated species. Based on the alignment and clustering analysis of N- and C-terminals as well as the portions of repetitive region near Cterminals (Figure 4), the identity of Glu-1Aix3 and Glu-1Aix4 was 100%, which suggested that the two genes likely were alleles of the heterozygous Glu-1.

A hybrid HMW-GS (GenBank accession No. AY611724) was isolated from *Ae. searsii*, and it was hypothesized that unequal crossover between the two *Glu-1* loci located on the sister chromatids of the homologous chromosomes during meiosis might have led to the formation of the novel hybrid ORF [9]. Following this mechanism underlying the evolution of the subunit, this would generate two hypothetical recombinant *Glu-1* loci, which belong to xy-type and yx-type, respectively (Figure 6). In our study, two hybrid subunits, xy-type subunit 1Aix1 and yx-type subunit 1Aiy1, were identified from one *Ag. intermedium* line, which strongly supported the hypothetic mechanism above. In addition to the configuration shift of the subunits, the size change of the subunits might be resulted from the unequal crossover (Figure 6).

In previous literature on HMW-GSs, quite a few research results showed that cysteine residues are involved in the formation of inter- or intra-molecular disulphide bonds [6,7,8] and in addition there is little doubt that this network of GMP is stabilized by inter-chain disulphide bonds [8]. Most importantly, it has been shown that the 1Dx5 subunit with an additional cysteine residue in its repetitive region exerts a positive influence on bread-making quality of flour [37,38]. Considering these distinctive

References

- Larwance GJ, Shepherd KW (1981) Chromsomal location of genes controlling seed protein in species realted to wheat. Theor Appl Genet 59: 25–31.
- Payne PI (1987) Genetics of wheat storage proteins and the effect of allelic variation on bread-making quality. Annu Rev of Plant Physiol Plant Mol Biol 38: 141–153.
- Shewry PR, Tatham AS, Barro F, Barcelo P, Lazzeri P (1995) Biotechnology of breadmaking: unraveling and manipulating the multi-protein gluten complex. Nat Biotechnol 13: 1185–1190.
- 4. Halford NG, Forde J, Anderson OD, Greene FC, Shewry PR (1987) The nucleotide and deduced amino-acid-sequences of an HMW glutenin subunit gene from chromosome 1B of bread wheat (*Triticum Aestizum L.*) and comparison with those of genes from chromosomes 1A and 1D. Theor Appl Genet 75: 117–126.

characteristics on the location and/or number of cysteine residue in 1Aix1, 1Aiy1 and 1Aiy3 (Table 2) from *Ag. intermedium*, it is quite necessary to identify their effect on the quality of flour.

During the recent decades, cytogenetic research showed that Ag. intermedium is an autoallo-hexaploid species designated with genomes E1E1E2E2XX, where the E1, E2 and X genomes are known to be related to the E^e genome (similar to J genome [34]) of diploid Th. elongatum, the E^b genome of Th. bessarabicum and the St genome of Ps. stipifolia, respectively [11,25,39,40]. However, our phylogenetic analysis based on HMW-GS genes revealed the two of the three species (*Th. bessarabicum* and *Ps. stipifolia*) are not close to Ag. intermedium in term of evolution (Figure 5). Considering Ag. intermedium is in general cross pollinated and that there is a wide variability within a single accession, the Ag. intermedium accession in this study probably is different from those in the cytogenetics studies. Thus, our data only displayed the putative origin of the specific Ag. intermedium accession in our study. In our phylogenetic analysis, another remarkable point is that some Aegilops species with D (Ae. tauschii) or S^b (Ae. bicornis) genome probably are involved in the origin of Ag. intermedium (subgroup D of Figure 5). In addition, recent research showed that all of the Ag. intermediumwheat addition lines with alien HMW-GS lost the activity of Glu-D1 [41], so at least one donor of Ag. intermedium is exchangeable with wheat D group chromosomes, i.e. a wheat D group-like genome probably is involved in the origin of Ag. intermedium. Furthermore, Ag. intermedium is known commonly as "intermedium wheatgrass" because of the similarity of their seed "heads or ears" to wheat. Most important is their high crossability with wheat, which is very helpful for germplasm innovation. As such, it is inferred that the common or similar genomic donors of wheat and Ag. intermedium probably contribute to their similar phenotypes and high crossability. For Leynus, with two basic genomes Ns and Xm [34], some of its species are the most likely direct donors of the Ag. intermedium based on the phylogenetic tree (subgroups B, C and E of Figure 5). In all, our phylogenetic analysis expanded the understanding on the origin of Ag. intermedium. However, more evidence is necessary to confirm our phylogenetic result using cytogenetics and/or comparative genomics approaches.

Acknowledgments

We would like to thank Dr. Bujun Shi (Australian Centre for Plant Functional Genomics, the University of Adelaide, Australia) for constructive suggestions on this manuscript.

Author Contributions

Conceived and designed the experiments: SC DW XZ. Performed the experiments: SC HX CG SH. Analyzed the data: SC XZ. Wrote the paper: SC XZ. Experimental material identification and preparation: ZL RY XW.

- Jiang QT, Wei YM, Wang F, Wang JR, Yan ZH, et al. (2009) Characterization and comparative analysis of HMW glutenin *IAy* alleles with differential expressions. BMC Plant Biol 9: 16.
- Altschuler Y, Galili G (1994) Role of conserved cysteines of a wheat gliadin in its transport and assembly into protein bodies in *Xenopus* oocytes. J Biol Chem 269: 6677–6682.
- Shewry PR, Halford NG (2002) Cereal seed storage proteins: structures, properties and role in grain utilization. J Exp Bot 53: 947–958.
- Shewry PR, Tatham AS (1997) Disulphide bonds in wheat gluten proteins. J Cereal Sci 25: 207–227.
- Sun X, Hu S, Liu X, Qian W, Hao S, et al. (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.

- Cao S, Xu H, Li Z, Wang X, Wang D, et al. (2007) Identification and characterization of a novel Ag. intermedium HMW-GS gene from *T. aestivum-Ag.* intermedium addition lines TAI-I series. J Cereal Sci 45: 293–301.
- Tang S, Li Z, Jia X, Larkin PJ (2000) Genomic in situ hybridization (GISH) analyses of *Thinopyrum intermedium*, its partial amphiploid Zhong 5, and diseaseresistant derivatives in wheat. Theor Appl Gene 100: 344–352.
- Gupta PK (1972) Cytogenetic evolution in *Triticinae*: homologous relationships. Genetica 43: 504–530.
- Cauderon Y, Saigne B, Dauge M (1973) The resistance to wheat rusts of Ag. intermedium and its use in wheat improvement. In: Sears ER, Sears LMS (eds) Proceedings of fourth international wheat genetics symposium Columbia Missouri, USA. pp401–407.
- Sharma HC, Gill BS (1983) Current status of wide hybridization in wheat. Euphytica 32: 17–31.
- Sun SC (1981) The approach and methods of breeding new varieties and new species from *Agrotriticum* hybrids. Acta Agronomica Sinica 7: 51–58.
- He MY, Xu ZY, Zou MQ, Zhang H, Chen DW, et al. (1988) The establishment of the two sets of alien addition lines of wheat-wheatgrass. Sci China (Series B) 32: 695–705.
- Fu BX, Kovacs MIP (1999) Rapid single-step procedure for isolating total glutenin proteins of wheat flour. J Cereal Sci 29: 113–116.
- Laemmli UK (1970) Cleavage of structural proteins during assembly of head of bacteriophage T4. Nature 227: 680–685.
- Liu Z, Yan Z, Wan Y, Liu K, Zheng Y, et al. (2003) Analysis of HMW glutenin subunits and their coding sequences in two diploid *Aegilops* species. Theor Appl Genet 106: 1368–1378.
- Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res 8: 4321–4325.
- Xie RL, Wan YF, Zhang Y, Wang DW (2001) HMW glutenin subunits in multiploid *Aegilops* species: composition analysis and molecular cloning of coding sequences. Chinese Sci Bull 46: 309–313.
- Wan Y, Wang D, Shewry PR, Halford NG (2002) Isolation and characterization of five novel high molecular weight subunit of glutenin genes from *Triticum* timopheevi and Aegilops cylindrica. Theor Appl Genet 104: 828–839.
- Kumarathasan P, Mohottalage S, Goegan P, Vincent R (2005) An optimized protein in-gel digest method for reliable proteome characterization by MALDI-TOF-MS analysis. Anal Biochem 346: 85–89.
- Thompson JD, Higgins DG, Gibson TJ (1994) Clustal W: 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.
- Friebe B, Mukai Y, Gill BS, Cauderon Y (1992) C-banding and in situ hybridization analyses of Agropyron intermedium, a partial wheat ×Ag. intermedium amphiploid, and 6 derived chromosome addition lines. Theor Appl Genet 84: 899–905.
- Cunsolo V, Foti S, Saletti R, Gilbert S, Tatham AS, et al. (2002) Investigation and correction of the gene-derived sequence of glutenin subunit 1Dx2 by matrixassisted laser desorption/ionisation mass spectrometry. Rapid Commun Mass Sp 16: 1911–1918.

- Cunsolo V, Foti S, Saletti R, Gilbert S, Tatham AS, et al. (2003) Structural studies of glutenin subunits IDy10 and IDy12 by matrix-assisted laser desorption/ionisation mass spectrometry and high-performance liquid chromatography/electrospray ionisation mass spectrometry. Rapid Commun Mass Sp 17: 442–454.
- Van Deynze AE, Dubcovsky J, Gill KS, Nelson JC, Sorrells ME, et al. (1995) Molecular-genetic maps for group-1 chromosomes of *Triticeae* species and their relation to chromosomes in rice and oat. Genome 38: 45–59.
- Gu YQ, Salse J, Coleman-Derr D, Dupin A, Crossman C, et al. (2006) Types and rates of sequence evolution at the high-molecular-weight glutenin locus in hexaploid wheat and its ancestral genomes. Genetics 174: 1493–1504.
- Gao SC, Gu YQ, Wu JJ, Coleman-Derr D, Huo NX, et al. (2007) Rapid evolution and complex structural organization in genomic regions harboring multiple prolamin genes in the polyploid wheat genome. Plant Mol Biol 65: 189– 203.
- De Bustos A, Jouve N (2003) Characterisation and analysis of new HMWglutenin alleles encoded by the *Glu-R1* locus of *Secale cereale*. Theor Appl Genet 107: 74–83.
- Blatter RHE, Jacomet S, Schlumbaum A (2004) About the origin of European spelt (*Triticum spelta* L.): allelic differentiation of the HMW glutenin B1–1 and A1–2 subunit genes. Theor Appl Genet 108: 360–367.
 Li XH, Wang AL, Xiao YH, Yan YM, He ZH, et al. (2008) Cloning and
- Li XH, Wang AL, Xiao YH, Yan YM, He ZH, et al. (2008) Cloning and characterization of a novel low molecular weight glutenin subunit gene at the *Glu-A3* locus from wild emmer wheat (*Triticum turgidum* L. var. dicoccoides). Euphytica 159: 181–190.
- Dong YC (2006) Wheat. In: Dong YC, Zheng DS (eds) Crops and their wild relatives in China. China Agroculture Press, Beijing. pp141–216.
- Forster BP, Reader SM, Forsyth SA, Koebner RMD, Miller TE, et al. (1987) An assessment of the homeology of 6 Agropyron intermedium chromosomes added to wheat. Genet Res 50: 91–97.
- Gao MJ, Hao S, He MY, Bu XL (1993) Isozymic studies of the homoeology between *Triticum aestiucm* and *Elytrigia intermedia* chromosomes in Set I wheatwheatgrass alien addition lines. Acta Genetica Sinica 20: 462–467.
- Gupta RB, Macritchie F (1994) Allelic variation at glutenin subunit and gliadin loci, *Glu-1*, *Glu-3* and *Gli-1* of common wheats. 2. Biochemical basis of the allelic effects on dough properties. J Cereal Sci 19: 19–29.
- Margiotta B, Colaprico G, Dovidio R, Lafiandra D (1993) Characterization of high Mr subunits of glutenin by combined chromatographic (RP-HPLC) and electrophoretic separations and Restriction-Fragment-Length-Polymorphism (RFLP) analyses of their encoding genes. J Cereal Sci 17: 221–236.
- Chen Q, Conner RL, Laroche A, Thomas JB (1998) Genome analysis of *Thinopyrum intermedium* and *Thinopyrum ponticum* using genomic in situ hybridization. Genome 41: 580–586.
- Dvorak J (1981) Genome relationships among Elytrigia (= Agropyron) elongata, Elytrigia stipifolia, Elytrigia elongata 4x, Elytrigia caespitosa, Elytrigia intermedia, and Elytrigia elongata 10x. Can J Genet Cytol 23: 481–492.
- Niu ZX, Klindworth DL, Wang RRC, Jauhar PP, Larkin PJ, et al. (2011) Characterization of HMW glutenin subunits in *Thinopyrum intermedium*, *Th. bessarabicum*, *Lophopyrum elongatum*, *Aegilops markgrafii*, and their addition lines in wheat. Crop Sci 51: 667–677.