

SCIENTIFIC REPORTS



OPEN

Low molecular weight glutenin subunit gene *Glu-B3h* confers superior dough strength and breadmaking quality in wheat (*Triticum aestivum* L.)

Received: 02 March 2016

Accepted: 13 May 2016

Published: 07 June 2016

Yaping Wang^{1,*}, Shoumin Zhen^{1,*}, Nana Luo¹, Caixia Han¹, Xiaobing Lu¹, Xiaohui Li¹, Xianchun Xia², Zhonghu He² & Yueming Yan¹

Low molecular weight glutenin subunit is one of the important quality elements in wheat (*Triticum aestivum* L.). Although considerable allelic variation has been identified, the functional properties of individual alleles at *Glu-3* loci are less studied. In this work, we performed the first comprehensive study on the molecular characteristics and functional properties of the *Glu-B3h* gene using the wheat cultivar CB037B and its *Glu-B3* deletion line CB037C. The results showed that the *Glu-B3h* deletion had no significant effects on plant morphological or yield traits, but resulted in a clear reduction in protein body number and size and main quality parameters, including inferior mixing property, dough strength, loaf volume, and score. Molecular characterization showed that the *Glu-B3h* gene consists of 1179 bp, and its encoded B-subunit has a longer repetitive domain and an increased number of α -helices, as well as higher expression, which could contribute to superior flour quality. The SNP-based allele-specific PCR markers designed for the *Glu-B3h* gene were developed and validated with bread wheat holding various alleles at *Glu-B3* locus, which could effectively distinguish the *Glu-B3h* gene from others at the *Glu-B3* locus, and have potential applications for wheat quality improvement through marker-assisted selection.

Wheat (*Triticum aestivum* L.), one of the three major cereal crops in the world, is a critical source of energy and nutrients in the human diet and has excellent processing characteristics. Wheat dough is used to make various food products including bread, noodles, cakes, and biscuits¹. The seed storage proteins in wheat consist of monomeric gliadins and polymeric glutenins that determine the extensibility and elasticity of dough, respectively^{2,3}. According to their mobility, as determined by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), polymeric glutenins are subdivided into high and low molecular weight glutenin subunits (HMW-GS and LMW-GS, respectively), of which, LMW-GS accounts for ~60% of the glutenins and primarily determines dough strength and viscosity, thus playing a significant role in flour processing quality^{4,5}. Some studies have shown that the effects of LMW-GS on both dough resistance and dough extensibility are more favorable than the effects of HMW-GS^{6,7}.

LMW-GS is encoded by *Glu-A3*, *Glu-B3*, and *Glu-D3* loci on the short arms of chromosomes 1A, 1B, and 1D, respectively, and these loci are linked to the *Gli-1* locus, encoding gliadins^{2,8}. The molecular structure of LMW-GS contains four typical regions: (1) a signal peptide containing 20 amino acids removed in the maturation process, (2) a short N-terminal region with 13 amino acids containing one cysteine, (3) a repetitive domain rich with glutamine containing 70–186 amino acids as the variable region of gene size, and (4) a C-terminus rich with cysteine and glutamine. The C-terminus has three regions: (1) a cysteine-rich structure containing five cysteines, (2) a domain rich with glutamine, with only one cysteine and some tandem glutamines, and (3) a conserved region of the C-terminus with the last cysteine^{9,10,11}.

¹College of Life Science, Capital Normal University, Beijing 100048, China. ²Institute of Crop Science, Chinese Academy of Agricultural Sciences (CAAS), Beijing 100081, China. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to Y.Y. (email: yanym@cnu.edu.cn)

According to the molecular weight of the subunits, LMW-GS could be classified into B, C, and D type subunits, of which, B type subunits are the main type of LMW-GS and belong to alkaline proteins. The isoelectric point of the C type subunits varies from weakly acidic to strongly basic, while the D type subunits possess a low percentage of total LMW-GS and are only expressed in certain cultivars^{8,12}. LMW-GS is classified into three sub-classes (LMW-m, LMW-s, and LMW-i) based on their N-terminal amino acid sequences, in which m, s, and i represent methionine, serine, and isoleucine, respectively, the first amino acid residues of the mature proteins^{13,14}. The N-terminus of LMW-m contains METSHIPGL-, METSRIPGL-, and METSCIPGL-. The N-terminus of LMW-s is SHPGL-, while the LMW-i type lacks the typical N-terminal sequence, following the signal peptide is the repetitive domain ISQQQ^{13,15,16}.

The genes encoding LMW-GS do not contain an intron, which was estimated to have 10–15 or 30–40 copies in hexaploid wheat^{10,17} due to extensive allelic variation present at *Glu-3* loci. Gupta and Shepherd¹⁸ identified and named 20 alleles at *Glu-3* in hexaploid wheat, including six at *Glu-A3*, nine at *Glu-B3*, and five at *Glu-D3*. These alleles were shown to have differential effects on wheat processing quality. Ikeda *et al.*¹⁹ isolated several LMW-GS genes and classified them into 12 groups in a soft wheat cultivar. Dong *et al.*¹ also identified four, three, and seven LMW-GS genes at the *Glu-A3*, *Glu-B3*, and *Glu-D3* loci, respectively, in Xiaoyan 54.

To date, studies on the functional properties of individual alleles are still challenging due to the lack of LMW-GS mutants at the *Glu-3* locus and the difficulty of wheat genetic transformation. Previous studies were performed mainly using near-isogenic lines (NILs) to understand the relative importance of different *Glu-3* alleles for wheat quality. The effects of different *Glu-3* loci on processing qualities have been ranked in bread wheat. *Glu-A3* and *Glu-B3* alleles are supposedly more important than *Glu-D3* alleles in terms of wheat processing qualities. Zhang *et al.*²⁰ studied the function of 18 LMW-GS alleles in bread wheat using Aroona NILs, and found that *Glu-A3e* exhibited the worst performance with respect to almost all quality properties. *Glu-B3b*, *Glu-B3g*, and *Glu-B3i* were more highly correlated with superior breadmaking quality than the other *Glu-B3* alleles, whereas the *Glu-D3* alleles showed no significant effects on bread wheat quality. Jin *et al.*²¹ used a set of Aroona NILs of bread wheat to clarify the contribution of each allele to processing quality, and showed that alleles *Glu-A3b*, *Glu-A3d*, *Glu-B3g*, and *Glu-B3f* make significant contributions to mixograph properties. More recently, Zhen *et al.*²² used a natural mutation at the *Glu-A3* locus to study the function of the *Glu-A3a* allele in bread wheat, and found that *Glu-A3a* significantly affects dough strength and breadmaking quality. However, functional studies on most other *Glu-3* alleles have not been performed.

In the current study, we used the *Glu-B3* deletion line, which was derived from mutation screening during immature embryo culture for spring wheat CB037B (*Triticum aestivum* L., $2n = 6 \times = 42$, AABBDD) genetic transformation research in our laboratory, to conduct the first comprehensive study on the molecular characteristics and functional properties of the LMW-GS allele *Glu-B3h* using proteomic and molecular biology approaches. SDS-PAGE, Two-dimensional electrophoresis (2-DE), reversed-phase ultra performance liquid chromatography (RP-UPLC) and liquid chromatography-tandem mass spectrometry (LC-MS/MS) were used to separate and characterize the *Glu-B3h* encoded B-subunit, sequence-tagged site polymerase chain reaction (STS-PCR) were used to confirm the deletion of *Glu-B3h* in CB037C and allele-specific PCR (AS-PCR) is to amplify the *Glu-B3h* gene in CB037B. Light microscopy observation and scanning electron microscope (SEM) were designed to analyze the protein body (PB) number and size, the function properties of LMW-GS *Glu-B3h* were measured by Mixograph and Extenograph, a variety of bioinformatics method including sequence alignment, secondary structures prediction and phylogenetic analysis were performed to understand the molecular characteristic of *Glu-B3h*, considering the superior processing properties of *Glu-B3h* encoded LMW-GS, the single nucleotide polymorphism (SNP)-based molecular markers for the *Glu-B3h* gene were developed and validated. Our results demonstrate that *Glu-B3h* has important effects on dough strength and breadmaking quality and reveal its potential value for contributing to wheat quality improvement.

Results

Agronomic traits and yield performance of CB037B and CB037C at three locations. The results from three locations (Beijing, Yinchuan, and Xining) showed that CB037B and CB037C have relatively high purity and genetic stability (Supplementary Fig. S1 and Table S1). Plant morphological characteristics, main agronomic traits, and yield performance of both cultivars, including plant and spike morphology, grain weight, and yield, were not significantly different (Supplementary Table S1). Thus, both cultivars were highly consistent without regarding growth and development at the different locations. However, the CB037B developed in Beijing, Yinchuan and Xining showed significant difference in agronomic traits and yield performance, and the same trend also appeared in the CB037C which were planted in these three locations, which demonstrating that environmental factors have significant effects on agronomic traits and yield performance.

Separation and characterization of the *Glu-B3h* encoded B-subunit. The glutenin composition of CB037B and CB037C was separated and identified by SDS-PAGE (Fig. 1a). Both cultivars had the same HMW-GS composition at the *Glu-1* locus (1Ax1, 1Bx17 + 1By18, 1Dx5 + 1Dy10). At the *Glu-3* locus, CB037B had *Glu-A3a*, *Glu-B3h*, and *Glu-D3d*. CB037C had the same *Glu-3* composition as CB037B, but *Glu-B3h* encoding one abundant LMW-B subunit was absent. 2-DE analysis revealed that *Glu-B3h* encodes four protein components (spots 1, 2, 3 and 4 in Fig. 1b). RP-UPLC further confirmed the absence of one abundant LMW B-subunit encoded by *Glu-B3h* in CB037C, which eluted after 16.44 min (Fig. 1c). The B-subunit peak had the highest expression and accounted for 16.6% of the total glutenin protein content and 30.8% of the total LMW-GS content in CB037B.

The *Glu-B3h* encoded protein band on the SDS-PAGE gel (Fig. 1a) and four protein spots on the 2-DE gel (spots 1, 2, 3 and 4 in Fig. 1b) were collected and digested by trypsin, and then analyzed by LC-MS/MS. Based on these results, four peptide sequences (13.52% coverage) were matched with the *GluB3-3* gene (AC number EU369717), encoding the LMW B-subunit in the *T. aestivum* L. cultivar Aroona-B3h, of which

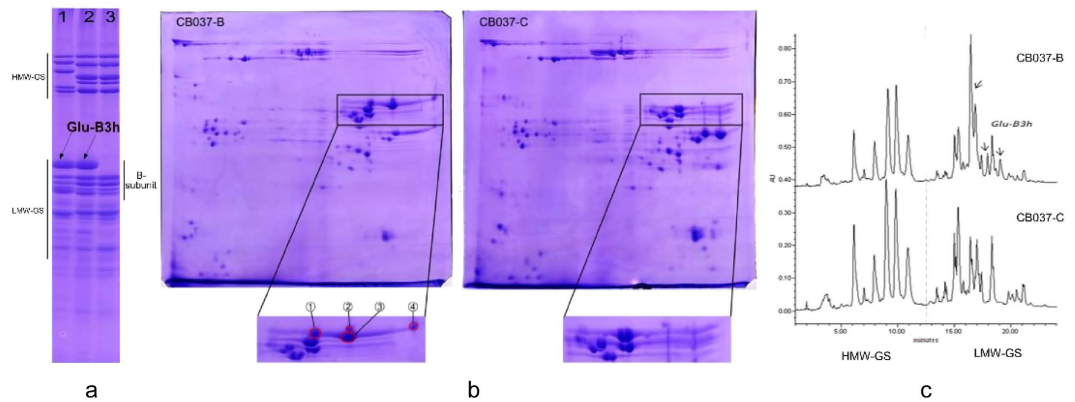


Figure 1. Identification of *Glu-B3h* from CB037B and *Glu-B3h* deletion line CB037C by SDS-PAGE (a), 2-DE (b) and RP-UPLC (c). (a) The *Glu-B3h* encoded B-subunit as well as LMW-GS and HMW-GS were indicated. 1. Aroona B3h; 2. CB037B; 3. CB037C. (b) Four protein spots differentially expressed between CB037B and CB037C were marked by 1, 2, 3 and 4. (c) Three protein peaks encoded by *Glu-B3h* as well as glutenin subunits in CB037B and CB037C were indicated.

“K.VFLQQQCSPVAM*PQSLAR.S” was completely consistent with the 236–255 amino acid sequence of the *GluB3-3* coding protein (Supplementary Table S2). Thus, the deleted B-subunit in CB037C was determined to be encoded by *Glu-B3h*.

Confirmation of *Glu-B3h* locus deletion in CB037C by STS-PCR markers. STS-PCR was used to confirm whether the *Glu-B3* locus was absent or silent in CB037C. The *Glu-B3h* gene at the *Glu-B3* locus in CB037B and CB037C was amplified using a pair of STS primers. One clear fragment of 1022 bp was amplified in CB037B and Aroona-B3h with the *Glu-B3h* gene, but no products were obtained from CB037C or other cultivars lacking the *Glu-B3h* encoded B-subunit (Supplementary Fig. S2). One amplified 1022 bp fragment was collected and sequenced, and the results indicated that it was identical to the sequence from 114 to 1136 bp of the *Glu-B3h* gene (Supplementary Fig. S3). These results confirm that the *Glu-B3* locus was deleted in CB037C.

Protein body comparison of CB037B and CB037C during grain development. Seed storage proteins are synthesized in the endoplasmic reticulum and are transferred and accumulate in the PB after being processed in the Golgi apparatus. PB differences between CB037B and CB037C during grain development were observed by different microscopy techniques (Fig. 2). Light microscopy observation of transverse slices of the grain endosperm showed that PB formation and developmental features were similar in both cultivars, but the number and sizes of PBs were significantly different (Fig. 2a). At 8 days post anthesis (DPA), starch granules were clearly visible and a handful of PBs were formed, while at 11–17 DPA, the number and size of the PBs increased quickly and some gathered PB fusion for glutenin macropolymers (GMP) formation. The PBs were fully combined to form the uniform protein matrix at 26 DPA in both cultivars. However, deletion of the *Glu-B3h* encoded B-subunit resulted in a reduction in PB number and size. SEM further verified that CB037B had more and larger PBs than CB037C (Fig. 2b), consistent with the results from light microscopy (Fig. 2a).

Functional properties of the *Glu-B3h* gene. Comparison of the two cultivars for their main quality traits over three locations showed that the deletion of LMW-GS *Glu-B3h* significantly reduced gluten strength and breadmaking quality (Table 1). In general, total grain protein contents of CB037B and CB037C were not significantly different, but GMP contents during grain development were significantly reduced in CB037C (Fig. 3a,b).

Main Mixograph and Extensograph parameters reflecting dough kneading resistance were also significantly lower in CB037C than in CB037B. These parameters included tolerance index, development time, stability, max resistance, and extensibility. Finally, the reduction of these mixing properties led to a significant reduction in loaf volume and score (Fig. 3c; Table 1). These results were highly consistent over the three locations. In addition, the locations had greater differences for quality parameters, indicating the environmental effects on gluten quality formation.

Molecular characterization of the *Glu-B3h* gene. According to the results from tandem mass spectrometry and coding sequences of *GluB3-3*, a pair of AS-PCR primers, LB3F and LB3R, was designed and synthesized, and used to amplify the *Glu-B3h* gene from CB037B. One clearly amplified band of approximately 1300 bp was present in CB037B (Supplementary Fig. S4). After collecting, cloning, and sequencing, a complete 1179 bp ORF was obtained, corresponding to the size typical of LMW-GS genes that range from 900 bp to 1200 bp. Sequence alignment showed that the amplified sequences had the basic structural characteristics of LMW-GS genes and no internal stop codons. After blasting this gene in NCBI, we found that the cloned nucleotide sequence was completely consistent with the *GluB3-3* gene from the wheat cultivar Aroona-B3h (AC number EU369717), therefore, it was determined to be *Glu-B3h* gene.

The deduced N-terminal sequence of the *Glu-B3h* gene was MENSHPGL-. Since the peptide MEN was likely to be removed from the original protein²³, the first amino acid of the mature protein of *Glu-B3h* was serine,

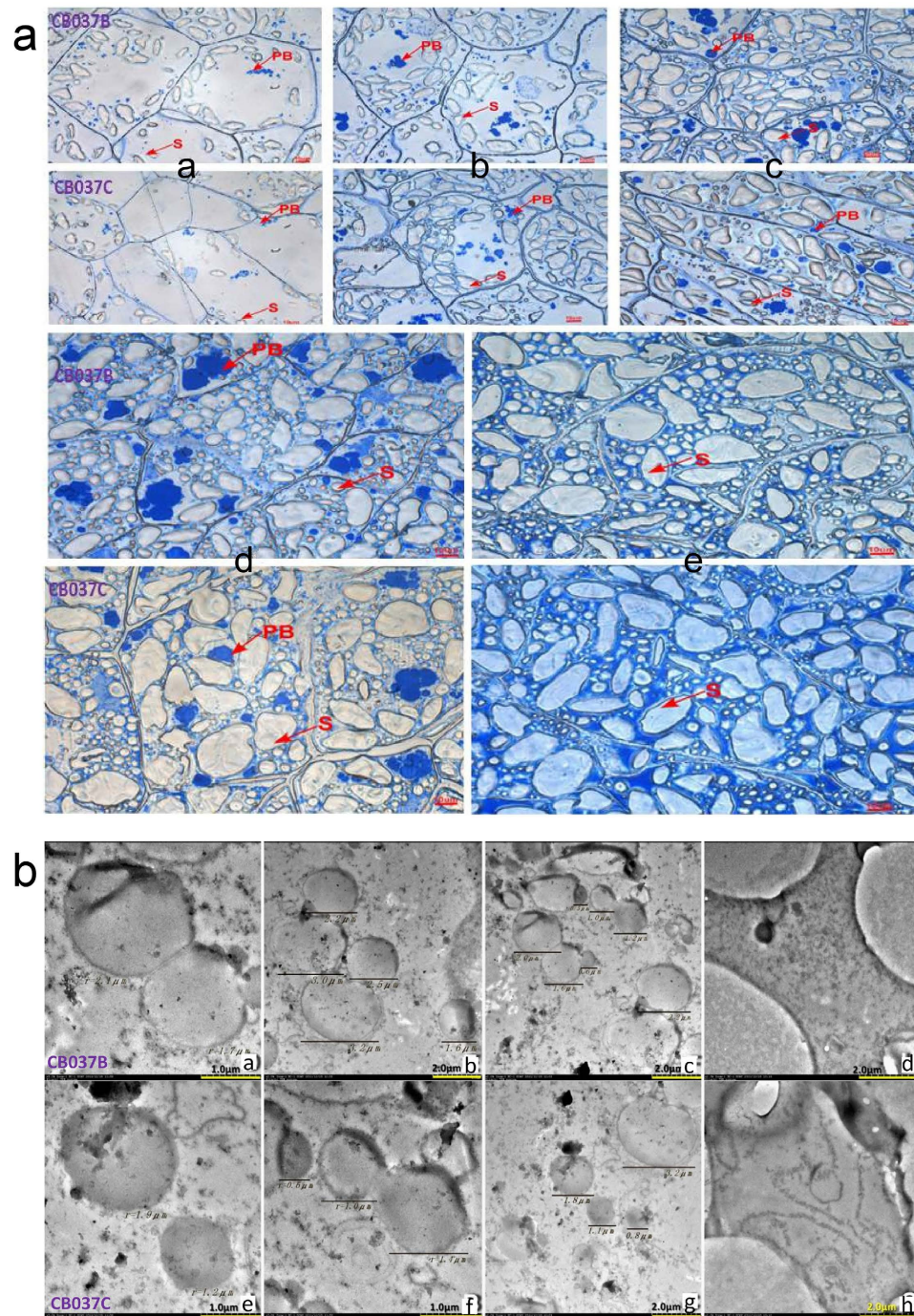


Figure 2. Developmental changes of PB during grain development of CB037B and CB037C. (a) Light microscopy observation of massicot blue staining on sections of protein bodies at 8, 11, 14, 17, 26 DPA. Numbers a, b, c, d and e correspond to the 8, 11, 14, 17 and 26 DPA; PB: protein bodies; S: starch; Bars: 10 μm. (b) Scanning electron microscopy observation of dynamic changes about endosperm structures from 4-day intervals after flowering during grain development.

indicating that this protein belongs to the LMW-s type subunit. The complete coding sequence of *Glu-B3h* was aligned with 20 known LMW-s genes from wheat to detect SNPs and insertions/deletions (InDels) variation, indicating that *Glu-B3h* had six SNPs at different positions. Two of the SNPs were at 972 bp and 1110 bp, and four were at 133 bp, 174 bp, 1117 bp, and 1172 bp, belonging to synonymous and nonsynonymous SNPs (Table 2). In particular, *Glu-B3h* had a long insertion fragment of 42 bp at position 464–505, making it longer than other LMW-GS genes.

The deduced amino acid sequences of the *Glu-B3h* gene included four clear regions of the typical primary structure of LMW-GS (Supplementary Fig. S5). The repetitive domain of *Glu-B3h* contained 15 hexapeptides (consensus

Growing location	Glu-B3h	Total protein(%)	Wet glutenin(%)	GMP content**	Tolerance index (FU)	Development time (min)	Stability (min)	Max resistance (B.U.)	Extensibility (mm)	Loaf volume (ml ³)	Loaf score
Yinchuan	CB037-B (+)	15.52 ± 0.1ab	58.45 ± 0.52a	8.75 ± 0.07a 1943.9 ± 36.2a	149 ± 2.86a	7.78 ± 0.08a	12.45 ± 0.11a	523 ± 5.51a	188 ± 1.17a	753 ± 6.17a	66 ± 0.56a
	CB037-B (-)	15.25 ± 0.08ab	56.52 ± 0.48a	8.16 ± 0.08b 1855.7 ± 38.3b	115 ± 2.74b	5.46 ± 0.06b	7.74 ± 0.09b	441b ± 4.12b	152 ± 1.11b	658 ± 6.74b	60 ± 0.45b
Beijing	CB037-B (+)	14.86 ± 0.08a	51.47 ± 0.42b	7.66 ± 0.08c 1885.9 ± 36.2c	145 ± 2.51a	7.42 ± 0.07a	11.85 ± 0.12a	483 ± 4.11c	173 ± 1.71a	746 ± 7.21a	64 ± 0.77a
	CB037-B (-)	14.92 ± 0.09a	50.82 ± 0.49b	6.50 ± 0.06d 1792.7 ± 36.2d	105 ± 1.78b	5.97 ± 0.06b	7.16 ± 0.07b	421b ± 4.32d	148 ± 1.53b	628 ± 6.12b	58 ± 0.21b
Xining	CB037-B (+)	15.87 ± 0.11b	62.25 ± 0.75c	8.86 ± 0.09a 1838.6 ± 31.7c	151 ± 2.36a	8.25 ± 0.09a	12.75 ± 0.11a	505 ± 5.01e	182 ± 1.73a	763 ± 7.44a	67 ± 0.63a
	CB037-B (-)	15.76 ± 0.12b	63.73 ± 0.88c	8.19 ± 0.07b 1784.6 ± 31.8d	118 ± 1.96b	5.46 ± 0.07b	7.96 ± 0.11b	457 ± 4.45f	159 ± 1.66b	605 ± 6.53b	52 ± 0.47b

Table 1. Comparison of main quality parameters between CB037-B with Glu-B3h and CB037-C without Glu-B3h from three growing locations*. *Different letters indicate significance level at $P=0.05$. **Upper values (%) and under values (Peak area, 1000 AU/S) were measured by the method of Weegels *et al.*⁵² and SE-HPLC, respectively.

LMW-GS	464–505	133	174	972	1110	1117	1172
EU369717	AACAACCGACTACTACCGCAACAACCACCATTTCGCGAGCAAC	A	C	A	T	T	G
20 other LMW-s genes	—	T	A	G	C	C	C

Table 2. The positions of SNPs and InDels identified between Glu-B3h and other LMW-s genes.

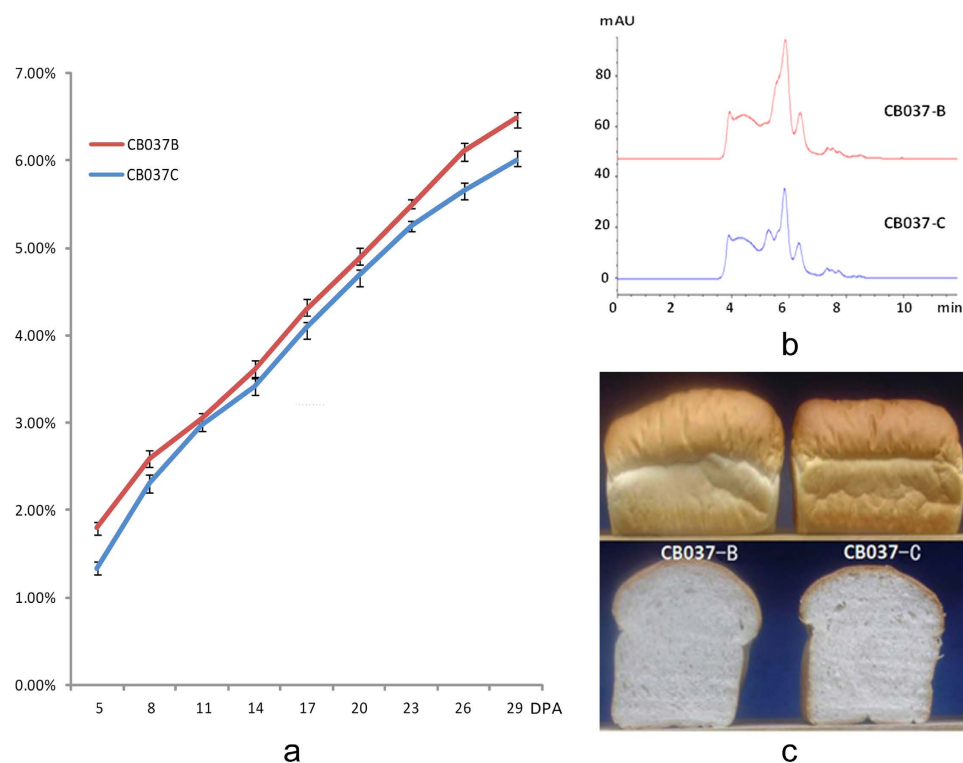


Figure 3. The comparison of GMP contents and pan bread between CB037B and CB037C. (a) GMP content determination of CB037B and CB037C at 5, 8, 11, 14, 17, 20, 23, 26, 29 DPA. (b) Separation and identification of GMP in the CB037B and CB037C by SE-HPLC. (c) Pan bread appearance of CB037B and CB037C.

PPFSQQ and QQPVLV) and two nonapeptides (QQPSFSQQQ), which accounted for 57% of the repetitive domain and were mainly responsible for the length expansion and general hydrophilic characteristics of LMW-GS.

LMW-GS	Structure motifs	Contents (%)	Total	Dispersal in every region				
				N-terminal domain	Repetitive domain	C-ter domain I	C-ter domain II	C-ter domain III
AY542896	α -helix	2.72	1	–	–	–	–	1
	β -strand	1.09	2	–	–	1	–	1
AY831866	α -helix	5.98	3	–	–	2	–	1
	β -strand	1.09	2	–	–	–	1	1
AY724436	α -helix	–	–	–	–	–	–	–
	β -strand	1.32	2	–	–	–	–	2
FJ824789	α -helix	9.68	4	–	–	2	2	–
	β -strand	–	–	–	–	–	–	–
JX877832	α -helix	9.38	4	–	–	3	1	–
	β -strand	–	–	–	–	–	–	–
<i>Glu-B3h</i>	α -helix	9.95	4	–	–	3	1	–
	β -strand	0.54	1	–	–	–	–	1

Table 3. The secondary structure prediction of the six deduced LMW-GS.

Secondary structure analysis of the *Glu-B3h* encoded protein. Secondary structures of the deduced mature glutenin subunits of *Glu-B3h* and five other LMW-GS from bread wheat (AY542896, AY831866, AY724436, FJ824789, and JX877832) were predicted on the PSIPRED server, and a comparative analysis was performed (Table 3). The results showed that the α -helices and β -strands were relatively conserved in repetitive domain. C-terminal and dispersed in a normal configuration in repetitive domain. The *Glu-B3h* encoded B-subunit (AC number EU369717) contained four α -helices located mainly at C-terminal I, and one β -strand in the conserved C-terminal region. Comparative analysis showed that the number of α -helices in the *Glu-B3h* encoded protein was much higher than AY542896, which was confirmed to have positive effects on wheat bread quality²⁴. These results indicate that the *Glu-B3h* encoded subunit could be associated with superior gluten quality and may be responsible for the significant reduction in dough strength and breadmaking quality of CB037C.

Phylogenetic analysis among *Glu-3* genes. To understand the phylogenetic relationship among the LMW-GS genes at the *Glu-3* locus, the coding sequences of the *Glu-B3h* gene and an additional 18 LMW-GS genes from wheat and related species were used to construct a homology tree using MEGA6 software. This tree included eight LMW-s (AB119007, AB164416, EU189088, AB262661, EU369700, EU369722, EU189095, and DQ357058), five LMW-m (GQ892576, GQ892588, KC222115, KC222119, and KC222107), and five LMW-i (DQ307387, HE647817, AY542896, GQ150532, and EU189087) type genes from *T. aestivum*, *T. durum*, *T. turgidum*, and *T. timopheevii*. These results indicate that LMW-i type genes underwent greater divergence during evolution and were clustered into a separate clade, while LMW-m and LMW-s type genes were separated into another clade in which two subgroups were present. *Glu-B3h* showed a close relationship with other LMW-s genes, particularly with EU369722 from bread wheat (Fig. 4).

The divergence timing of *Glu-B3h* and an additional 11 LMW-GS genes were calculated using MEGA6.0 software to further investigate the evolutionary relationships among the LMW-GS genes. LMW-s was more related to LMW-m than LMW-i, their divergence occurred around 6–8 million years ago (MYA; Supplementary Table S3). LMW-i type genes diverged much earlier, at about 11–14 MYA, suggesting that LMW-i type genes are the most original in the LMW-GS family.

Development and validation of SNP-based molecular markers for the *Glu-B3h* gene.

Based on the SNPs detected in the *Glu-B3* gene, a pair of primers specific for *Glu-B3h* (*Glu-B3h* F: CCACCA CAACAAACATTAA, *Glu-B3h* R: TGCCCGAGTTG CTGTTGT) was designed and tested in wheat cultivars with various allelic compositions at the *Glu-B3* locus, as identified by SDS-PAGE (Supplementary Table S4 and Fig. S6). PCR amplification showed that a fragment of 881 bp was present in cultivars with the *Glu-B3h* gene, while no products were obtained from the other wheat genotypes with other *Glu-B3* alleles (Fig. 5). The SNP-based molecular markers developed were further validated using different cultivars (Fig. 5a), two F₂ populations from CB037B \times Ningchun 4 and CS-1 S¹/1B \times CB037B, two recombinant inbred lines (RILs) from CB037B \times Ningchun 4 and CS-1 S¹/1B \times CB037B (Fig. 5b), and eight Aroona NILs (Fig. 5c). These molecular markers have potential applications for wheat quality improvement through marker-assisted selection.

Discussion

In this study, we used the *Glu-B3* deletion mutant CB037C to investigate the molecular characteristics and functional properties of the *Glu-B3h* gene. Our results provide new insights into the allelic variation mechanism and molecular basis of gluten quality formation. The *Glu-B3h* gene, with specific structural features, showed potential value for wheat gluten quality improvement through use of SNP-based molecular markers. Here, we focused on several key issues pertaining to the *Glu-B3h* gene.

Our results show that *Glu-B3h* encodes an LMW-s type B-subunit and has close phylogenetic relationships with other LMW-s genes. Both LMW-m and LMW-s possess similar structures and close evolutionary relationships at the amino acid sequence level²⁵. Thus, the LMW-s type gene is considered a derivative of LMW-m type subunits,

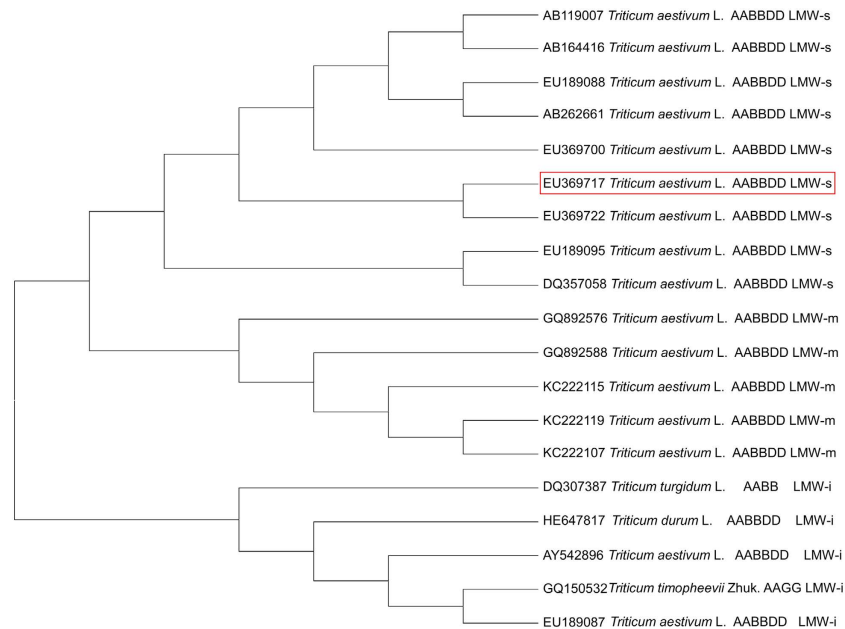


Figure 4. Phylogenetic tree of 19 LMW-GS constructed by the complete coding DNA sequences. 19 LMW-GS genes named EU369717, AB119007, AB164416, EU189088, AB262661, EU369700, EU369722, EU189095, DQ357058, GQ892576, GQ892588, KC222115, KC222119, KC222107, DQ307387, HE647817, AY542896, GQ150532 and EU189087.

and its divergence occurred ~ 7.81 MYA²⁶. In the present study, the divergence time estimation for the three types of LMW-GS genes showed that LMW-s and LMW-m genes diverged at about 6–8 MYA (Supplementary Table S3), which is consistent with the findings by Li *et al.*²⁶. The divergence time of LMW-m and LMW-i was about 12–14 MYA, which is consistent with that described by Wang *et al.*²³. It is known that the primary genetic mechanisms for allelic variation at the *Glu-1* and *Glu-3/Gli-2* loci mainly involve point mutations, unequal crossing-over, slip-mismatching, and intra-chromosomal illegitimate recombination^{8,26,27}. The extensive allelic variation of storage proteins mainly results from SNPs and InDels. In particular, *Glu-B3h* has six SNPs and a long insertion of 464–505 residues, which has led to a longer repetitive domain and larger gene size (Table 2). Considering the genome expansion and contraction from unequal crossing-over and illegitimate recombination^{26,28}, it is likely that the *Glu-B3h* gene originated from unequal crossing-over or an illegitimate recombination event that may have occurred at 6.4–6.8 MYA. SNP variation gradually accumulated during the evolutionary process.

It is known that both expression levels and structural features of glutenin subunits are closely related to gluten quality²⁹. The over-expressed HMW-GS 1Bx7^{OE} was confirmed to have positive effects on dough strength³⁰. LMW-GS accounted for approximately 60% of the total protein in mature seeds⁴. LMW B-subunits are the most abundant among the LMW-GS in mature grains and have the greatest impact on wheat processing qualities⁸. Our results from RP-UPLC indicate that the *Glu-B3h* encoding subunit has an abundant protein peak, and its content accounts for 30.8% of the total LMW-GS in CB037B (Fig. 2c). Thus, the deletion of *Glu-B3h* led to a significant decrease in GMP content, dough strength, and breadmaking quality (Table 1; Fig. 5). LMW-GS participates in the formation of gluten macropolymers through intra- and inter-molecular disulfide bonds. Thus, disulfide bonds play important roles in determining the structure and properties of gluten proteins³¹. The different amounts and distribution of cysteine residues are closely related to the formation of secondary protein structure and dough quality, in which the first and seventh cysteine residues participate in forming the intermolecular disulfide bond, while the remaining residues are involved in the formation of intra-molecular disulfide bonds¹³. Helix-helix interactions could guide the formation of the intra-molecular disulfide bonds, thus, more α -helices might contribute to superior dough quality³². A previous study confirmed that AY542896 has positive effects on dough quality in bread wheat¹⁴. The secondary structure of *Glu-B3h* contains four α -helices, which was much higher than the one α -helix of AY542896 (Table 3), and may contribute to superior dough strength and breadmaking quality (Table 1).

The size and structural features of repetitive domains also have important effects on gluten quality. In general, LMW-GS genes are 900–1200 bp in length¹⁰ and contain a long repetitive domain that facilitates the formation of more α -helices and β -strands, and confers superior gluten structure and breadmaking quality^{32,33}. Our results showed that *Glu-B3h* contains 1179 bp with a large insert fragment at nucleotides 464–505 that represent 14 amino acids (Table 2), which is longer than most other LMW-GS genes. In addition, the repetitive domain of *Glu-B3h* contains 15 hexapeptides (consensus PPFSSQQ and QQPVLV) and 2 nonapeptides (QQPSFSQQQ), which account for 57% of the repetitive domain that is mainly responsible for the length of LMW-GS and contributes to superior dough quality³⁴. These structural features may thus contribute to superior gluten quality. A recent report showed that the introduction of *Glu-B3h* into the cultivar Yumechikara made it possible to breed cultivars

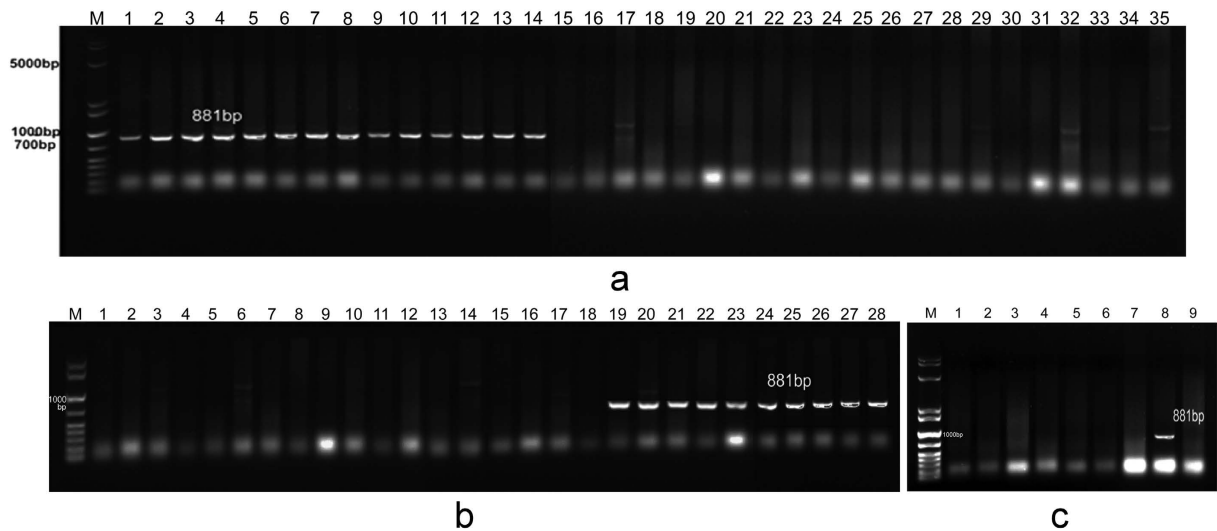


Figure 5. Development and validation of SNP-based molecular marker for *Glu-B3h*. (a) PCR amplification from different wheat cultivars: 1. CB037B. 2–35 corresponds to 26–59 in Table S4. (b) PCR amplification from RILs from CB037B × Ningchun 4 (1–9, 18–22) and CS-1S(1B) × CB037B (10–17, 23–27). 28. CB037B. (c) PCR amplification from Aroona NILs. 1. Aril20-1, 2. Aril21-2, 3. Aril23-4, 4. Aril24-3, 5. Aril26-1, 6. Aril27-6, 7. Aril29-4, 8. Aril28-4, 9. Aril30-1. M is marker (10 kb, 8 kb, 5 kb, 2 kb, 1.6 kb, 1 kb, 700 bp, 500 bp, 400 bp, 300 bp, 200 bp, 100 bp).

with good gluten and breadmaking quality³⁵. This further confirms that *Glu-B3h* has potential applications for improving wheat gluten quality.

As a traditional method widely used for LMW-GS identification, SDS-PAGE has disadvantages in distinguishing some subunits with similar mobilities. It is also time-consuming and requires the use of toxic reagents. However, marker-assisted selection has become more effective for screening for superior genes. The development of molecular markers for the *Glu-3* locus is important for improving wheat quality. Wang *et al.*³⁶ developed 10 allele-specific PCR markers based on SNPs and used them to discriminate the *Glu-B3* subunits. Based on sequence alignment of 13 LMW-GS genes previously identified, Zhang *et al.*³⁷ developed a new molecular marker system to identify the LMW-GS gene family. Several STS markers for *Glu-3* subunits were developed and identified in more than 100 wheat cultivars^{38,39,40}. Ikeda *et al.*⁴¹ developed 12 markers to differentiate 12 groups of LMW-GS genes of Norin 61. However, although considerable work on gene cloning and marker discovery has been performed, the superior LMW-GS genes and effective molecular markers widely used for wheat quality improvement are still limited. This may have resulted from the challenge of functional studies of the *Glu-3* gene and its highly repetitive sequences affecting the development and application of effective molecular markers.

SNPs are considered to be effective third-generation molecular markers that are powerful tools in marker-assisted breeding⁴². PCR-based molecular markers are generally fast to use, cost-efficient, and are subjected to few restrictions, which could provide a powerful tool for high-throughput selection during marker-assisted selection for wheat quality. In the present study, we developed SNP-based AS-PCR markers for the *Glu-B3h* gene and validated them in different wheat cultivars, F₂ populations, NILs, and RILs (Supplementary Table S4; Fig. 5). By using these markers, breeders can now efficiently select this desirable subunit in early generations of a wheat quality breeding program.

Materials and Methods

Plant materials and field trials. Plant materials used in this study included CB037B (*Triticum aestivum* L., 2n = 6 × = 42, AABBDD) and its *Glu-B3h* deletion line CB037C derived from mutation screening during immature embryo culture for wheat genetic transformation research in our laboratory, 77 wheat cultivars and lines with different LMW-GS allele compositions, 25 Aroona NILs, two F₂ cross populations and 10 RILs. All materials used in this study were listed in Supplementary Table S4.

To estimate the performance of agronomic traits, yield and quality properties at different growing environments, CB037B and CB037C were planted at three locations of wheat production areas of north China (Beijing, Yinchuan and Xining) in 2015. Field trials were performed in randomized block design with three biological replicates (each plot with 30 m²). The cultivation and management were same as local field cultivation conditions.

Grain developmental changes, agronomic trait and yield measurement. The grain dynamic changes from 4 to 24 DPA with two-day intervals were observed. The mature wheat plants from each plot were harvested and the main agronomic traits and grain yield were measured, including tiller number per plant, plant height, main ear length, effective ears per plant, ear grain number, thousand grain weight and grain yield (GY, kg/ha.).

Glutenin extraction, SDS-PAGE and 2-DE. Seed glutenin extraction and SDS-PAGE were based on Mackie *et al.*⁴³ and Yan *et al.*⁴⁴. A 15 μ L aliquot of the extract was loaded onto a SDS-PAGE gel. After electrophoresis with 12% gel at 15 mM for 2.5 h, the gel was stained for 30 min with Coomassie Brilliant Blue (CBB) R-250/G-250 (4:1) 10% (v/v) carbinol, 50% (v/v) acetic acid, and subsequently destained in distilled water.

For 2-DE, the first dimension was performed by an Ettan™ IPG-phor II™ system (GE Healthcare, USA) using 18 cm strips (pH 6–11). The IEF rehydration solution was 7 M urea, 2 M thiourea and 4% CHAPS. The rehydrate condition was 30 V at 20 °C for 12 h while the IEF condition was 300 V for 1 h, 500 V for 1 h, 1000 V for 1 h, 3000 V for 1 h, and 8000 V to 80,000 V for 10 h. The second dimension was performed on a 12% acrylamide gradient. After electrophoresis, the 2-DE gels were stained within colloidal CBB R-250/G-250 (4:1) and analyzed by using ImageMaster™ 2-D platinum software version 5.0 (Amersham Bioscience, Swiss Institute of Bioinformatics, Geneva, Switzerland, 2003) based on Lv *et al.*⁴⁵. Three biological replicates were performed.

RP-UPLC. RP-UPLC was performed on an Agilent 1100 using a Zorbax 300SB-C18 column (300 Å pore size and 5 mm particle size) based on the recent report⁴⁶.

LC-MS/MS. The expected band on SDS-PAGE gel and 2-DE spots were excised and digested with trypsin according to Jin *et al.*⁴⁷. The digested proteins (0.5 ml) were subject to LC-MS/MS identification by a Waters SYNAPT High Definition Mass Spectrometry™ mass spectrometer. Then the LC-MS/MS data were analyzed by software BioworksBrowser 3.3.

DNA extraction and STS-PCR. Total genomic DNA was extracted from dry seeds according to McDonald *et al.*⁴⁸ and An *et al.*²⁴ with minor modifications. The molecular marker identification of *Glu-B3h* was performed using STS-PCR marker as previously report by Wang *et al.*³⁶. The primer sets are SB8F: CCACCACAACAAACATTA, and SB8R: GTGGTGGTTCTATACAACGA. The PCR cycling conditions were 94 °C for 4 min, followed by 35 cycles of 94 °C for 35 s, 60 °C for 35 s, 72 °C for 90 s, and a final extension at 72 °C for 8 min.

Light microscopy and SEM observation. Transverse slices with approximately 1 mm thick were cut from the wheat grains during each of the four periods, and fixed, rinsed, dehydrated, infiltrated, and polymerized by the series of steps outlined in Arcalis *et al.*⁴⁹ (2004). For light microscopy observation, sections with approximately 800 nm thick were cut, collected on mesh nickel grids, and stained with toluidine blue. The dynamic changes of endosperm structures from 4-day intervals after flowering during grain development were observed by SEM according to López-Merino *et al.*⁵⁰ Collected developing grains were immediately fixed in the solution containing 44.5% ethanol, 1.85% methanol and 6% glacial acetic acid for 60 min followed by an overnight treatment at 4 °C, then transferred into 70% ethanol and stored at 4 °C prior to analysis. All samples were dehydrated sequentially through an ethanol concentration series (50%, 70%, 85%, 95% and 100% v/v) with 1 h incubation in each solution. After thoroughly dried, endosperm structures were observed by S-4800 FESEM machine (Hitachi, Japan).

Gluten quality testing. Main quality parameters of mature grains from CB037B and CB037C from three locations were tested with three biological replications according to Zhen *et al.*²². Total protein content (%N 5.7, 14% moisture basis) was determined by nitrogen combustion analysis with a LECO (Model FP analyzer, St. Joseph, MI) calibrated against EDTA. The extraction of GMP and content measurement were based on Xu *et al.*⁵¹ and Weegels *et al.*⁵². Separation of GMP by size exclusion-high performance liquid chromatography (SE-HPLC) were performed on a Phenomenex BIOSEPSEC 4000 column in acetonitrile buffer of 0.05% (v/v) trifluoroacetic acid and 0.05% (v/v) acetonitrile with a running time of 10 m (2 ml/m flow rate) according to Rakszegia *et al.*⁵³.

Mixograph parameters were measured with AAC54–40 A method by using 10 g electronic mixograph from America National manufacturing company. Extensograph parameters were tested based on He *et al.*⁵⁴. For breadmaking quality testing, flour samples (100 g, 14% moisture base) and dry yeast were mixed at a ratio of 100:1 (w:w) together with water to develop the dough by using a laboratory Pin Mixer (National Manufacturing Company, Lincoln, NE). After that, the proofed dough was steamed in a steamer with boiling water for 15 min and cooled in the steamer for 15 min, finally took it out and cooled for 1 h at room temperature. Pan bread score included loaf volume (weighting 36), oven spring (10), appearance (10), texture (15), structure (15) and color (14) based on previous report⁵⁵.

AS-PCR. A pair of AS-PCR primer LB3F and LB3R was synthesized to amplify the full-length of *Glu-B3h* and part of its upstream and downstream sequence according to the results of tandem mass spectrometry and coding sequences of *GluB3-3* from GenBank (AC number EU369717). The primer sequences were LB3F: 5'-CATCACAAG CACAAGCATCAA-3', LB3R: 5'-CATATCCATCGACTAAACAAA-3', synthesized by Sangong Inc., China. PCR amplifications were performed in 50 μ L reaction volumes containing 2.5 U La Taq polymerase (TaKaRa), 100 ng of template DNA, 25 μ L of 2 \times GC buffer I (MgCl₂⁺ plus), 0.4 mM dNTP, 0.5 μ M of each primer, and added to 50 μ L with double distilled H₂O. The reactions were carried out in a PTC-100 (MJ Research, Watertown, MA, USA) thermocycler using the following protocol: 94 °C for 2 min, followed by 35 cycles of 94 °C for 45 s, 57 °C for 50 s and 72 °C for 1.5 min, a finally extension at 72 °C for 10 min.

Molecular cloning, DNA sequencing and sequence alignment. PCR products were separated on 1.2% agarose gels in Tris-acetic acid-EDTA buffer and the expected fragments were purified from the gels using a Quick DNA extraction kit (Tiangen, Beijing, China). Subsequently, purified products were ligated into a PMD18–T Easy vector (TaKaRa, Dalian, China) and transformed into cells of *E. coli* strain DH5 α according to Li *et al.*⁵⁶. DNA sequencing was performed with three clones by Sino Geno Max, Beijing, China. And the

sequenced gene was blast in NCBI (<http://www.ncbi.nlm.nih.gov/>). Multiple sequence alignment of LMW-GS was performed by Bioedit 7.0.1.1.

SNPs and InDels identification and secondary structure prediction. SNPs and InDels present in LMW glutenin genes were identified using Bioedit 7.0.1.1. Prediction of secondary structure of LMW-GS genes was conducted by PSIPRED server (<http://bioinf.cs.ucl.ac.uk/psipred/>).

Phylogenetic tree construction. Clustal W program and MEGA 6.0 were used to construct phylogenetic tree and estimation of divergence times among LMW-GS genes. The Clustal W program was used to make a multiple alignment with homologous nucleotide sequences, the alignment file was converted to software MEGA 6.0 by the complete coding regions of LMW-GS genes with bootstrap values 1,000 replicates according to Johal *et al.*⁵⁷. The divergence times of LMW-GS genes were estimated by using MEGA 6.0 with the evolution rate as 6.5×10^{-9} substitution/site year according to Allaby *et al.*⁵⁸.

Development and validation of SNP-based molecular markers for Glu-B3h. The specific primers for *Glu-B3h* gene were designed based on the SNP variations and used to develop molecular markers and then validated using different genotypes, Aroona NILs, F₂ populations and RILs as shown in Supplementary Table S4. DNA extraction was as described above. PCR cycling conditions were 94 °C for 4 min, followed by 35 cycles of 94 °C for 35 s, 57 °C for 30 s, 72 °C for 90 s, and a final extension at 72 °C for 8 min. A total of 5–10 seeds for each cultivar, NIL and RIL, and 200–250 seeds from F₂ populations were tested.

References

- Dong, L. *et al.* New insights into the organization, recombination, expression and functional mechanism of low molecular weight glutenin subunit genes in bread wheat. *PLoS ONE* **5**, e13548 (2010).
- Payne, P. I. Genetics of wheat storage proteins and the effect of allelic variation on bread-making quality. *Annu Rev Plant Physiol.* **38**, 141–153 (1987).
- Wrigley, C. W. Giant proteins with flour power. *Nature* **381**, 738–739 (1996).
- Shewry, P. R. & Halford, N. G. Cereal seed storage proteins: structures, properties and role in grain utilization. *J Exp Bot.* **53**, 947–958 (2002).
- Tanaka, H., Shimizu, R. & Tsujimoto, H. Genetical analysis of contribution of low molecular weight glutenin subunits to dough strength in common wheat (*Triticum aestivum* L.). *Euphytica.* **141**, 157–162 (2005).
- Gupta, R. B., Singh, N. K. & Shepherd, K. W. The cumulative effect of allelic variation in LMW and HMW glutenin subunits on dough properties in the progeny of two bread wheats. *Theor Appl Genet.* **77**, 57–64 (1989).
- Cornish, G. B., Bekes, F., Allen, H. M. & Martin, D. J. Flour proteins link to quality traits in an Australian doubled haploid wheat population. *Aust J Agr Res.* **52**, 1339–1348 (2001).
- D'Ovidio, R. & Masci, S. The low-molecular-weight glutenin subunits of wheatgluten. *J Cereal Sci.* **39**, 321–339 (2004).
- Cassidy, B. G. & Dvorak, J. Molecular characterization of a low-molecular-weight glutenin cDNA clone from *Triticum durum*. *Theor Appl Genet.* **81**, 653–660 (1991).
- Cassidy, B. G., Dvorak, J. & Anderson, O. D. The wheat low molecular weight glutenin genes: characterization of six new genes and progress in understanding gene family structure. *Theor Appl Genet.* **96**, 743–750 (1998).
- D'Ovidio, R., Marchitelli, C., Ercoli, C. L. & Porceddu, E. Sequence similarity between allelic *Glu-B3* genes related to quality properties of durum wheat. *Theor Appl Genet.* **98**, 455–461 (1999).
- Thompson, S., Bishop, D. H. L., Madgwick, P., Tatham, A. S. & Shewry, P. R. High-level expression of a wheat LMW glutenin subunit using a baculovirus system. *J Agr Food Chem.* **42**, 426–431 (1994).
- Lew, E. J. L., Kuzmicky, D. D. & Kasarda, D. D. Characterization of low molecular weight glutenin subunits by reversed-phase high-performance liquid chromatography, sodium dodecyl sulfate-polyacrylamide gel electrophoresis, and N-terminal amino acid sequencing. *Cereal Chem.* **69**, 508–515 (1992).
- Cloutier, S., Rampitsch, C., Penner, G. A. & Lukow, O. M. Cloning and expression of a LMW-I glutenin gene. *J Cereal Sci.* **33**, 143–154 (2001).
- Masci, S., Lew, E. J. & Lafandra, D. Characterization of low-molecular-weight glutenin subunits in durum wheat by reversed-phase high-performance liquid chromatography and N-terminal sequencing. *Cereal Chem.* **72**, 100–104 (1995).
- Tao, H. P. & Kasarda, D. D. Two-dimensional gel mapping and N-terminal sequencing of LMW-glutenin subunits. *J Exp Bot.* **40**, 1015–1020 (1989).
- Sabelli, P. A. & Shewry, P. R. Characterization and organization of gene families at the *Gli-1* loci of bread and durum wheats by restriction fragment analysis. *Theor Appl Genet.* **83**, 209–216 (1991).
- Gupta, R. B. & Shepherd, K. W. Two-step one-dimensional SDS-PAGE analysis of LMW subunits of glutelin. *Theor Appl Genet.* **80**, 65–74 (1990).
- Ikeda, T. M., Nagamine, T., Fukuoka, H. & Yano, H. Identification of new low-molecular-weight glutenin subunit genes in wheat. *Theor Appl Genet.* **104**, 680–687 (2002).
- Zhang, X. *et al.* Composition and functional analysis of low-molecular-weight glutenin alleles with Aroona near-isogenic lines of bread wheat. *BMC Plant Biol.* **12**, 1–16 (2012).
- Jin, H. *et al.* Effects of allelic variation of HMW-GS and LMW-GS on mixograph properties and Chinese noodle and steamed bread qualities in a set of Aroona near-isogenic wheat lines. *J Cereal Sci.* **57**, 146–152 (2013).
- Zhen, S. *et al.* Deletion of the low-molecular-weight glutenin subunits allele *Glu-A3a* of wheat (*Triticum aestivum* L.) significantly reduces dough strength and breadmaking quality. *BMC Plant Biol.* **14**, 367 (2014).
- Wang, K. *et al.* Phylogenetic relationship of a new class of LMW-GS genes in the M genome of *Aegilops comosa*. *Theor Appl Genet.* **122**, 1411–1425 (2011).
- An, X. *et al.* Cloning and molecular characterization of three novel LMW-i glutenin subunit genes from cultivated einkorn (*Triticum monococcum* L.). *Theor Appl Genet.* **113**, 383–395 (2006).
- Huang, X. & Cloutier, S. Molecular characterization and genomic organization of low molecular weight glutenin subunit genes at the *Glu-3* loci in hexaploid wheat (*Triticum aestivum* L.). *Theor Appl Genet.* **116**, 953–966 (2008).
- Li, X. *et al.* A novel chimeric low-molecular-weight glutenin subunit gene from the wild relatives of wheat *Aegilops kotschy* and *Ae. Juvenalis*: evolution at the *Glu-3* loci. *Genetics* **180**, 93–101 (2008).
- Zhang, Y. *et al.* 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 (2008).
- Anderson, O. D. & Greene, F. C. The characterization and comparative analysis of high-molecular-weight glutenin genes from genomes A and B of a hexaploid bread wheat. *Theor Appl Genet.* **77**, 689–700 (1989).

29. Butow, B. J. *et al.* Molecular discrimination of Bx7 alleles demonstrates that a highly expressed high-molecular-weight glutenin allele has a major impact on wheat flour dough strength. *Theor Appl Genet.* **107**, 1524–1532 (2003).
30. Li, J., Han, C., Zhen, S., Li, X. & Yan, Y. Characterization of HMW glutenin subunit Bx7^{OE} and its distribution in common wheat and related species. *Plant Genetic Res.* **12**, 191–198 (2014).
31. Shewry, P. R. & Tatham, A. S. Disulphide bonds in wheat gluten proteins. *J Cereal Sci.* **25**, 207–227 (1997).
32. Masci, S., D'Ovidio, R., Lafiandra, D. & Kasarda, D. D. Characterization of a low-molecular-weight glutenin subunit gene from bread wheat and the corresponding protein that represents a major subunit of the glutenin polymer. *Plant Physiol.* **118**, 1147–1158 (1998).
33. Masci, S., D'Ovidio, R., Lafiandra, D. & Kasarda, D. D. A 1B-coded low-molecular-weight glutenin subunit associated with quality in durum wheats shows strong similarity to a subunit present in some bread wheat cultivars. *Theor Appl Genet.* **100**, 396–400 (2000).
34. Xiao, L. *et al.* Molecular characterization of high molecular weight glutenin allele *Glu-B1h* encoding 1Bx14 + 1By15 subunits in bread wheat (*Triticum aestivum* L.). *Crop Pasture Sci.* **65**, 215–226 (2014).
35. Ito, M. *et al.* Dough properties and bread-making quality-related characteristics of Yumehikara near-isogenic wheat lines carrying different *Glu-B3* alleles. *Breeding Sci.* **65**, 241–248 (2015).
36. Wang, L. *et al.* Characterization of low-molecular-weight glutenin subunit *Glu-B3* genes and development of STS markers in common wheat (*Triticum aestivum* L.). *Theor Appl Genet.* **118**, 525–539 (2009).
37. Zhang, X. *et al.* Development of a new marker system for identifying the complex members of the low-molecular-weight glutenin subunit gene family in bread wheat (*Triticum aestivum* L.). *Theor Appl Genet.* **122**, 1503–1516 (2011).
38. Zhao, X. *et al.* Characterization of three low-molecular-weight *Glu-D3* subunit genes in common wheat. *Theor Appl Genet.* **113**, 1247–1259 (2006).
39. Zhao, X. *et al.* Novel DNA variations to characterize low molecular weight glutenin *Glu-D3* genes and develop STS markers in common wheat. *Theor Appl Genet.* **114**, 451–460 (2007).
40. Wang, L., Li, G., Pena, R. J., Xia, X. & He, Z. Development of STS markers and establishment of multiplex PCR for *Glu-A3* alleles in common wheat (*Triticum aestivum* L.). *J Cereal Sci.* **51**, 305–312 (2010).
41. Ikeda, T. M., Araki, E., Fujita, Y. & Yano, H. Characterization of low molecular-weight glutenin subunit genes and their protein products in common wheats. *Theor Appl Genet.* **112**, 327–334 (2006).
42. Rafalski, A. Applications of single nucleotide polymorphisms in crop genetics. *Curr Opin Plant Biol.* **5**, 94–100 (2002).
43. Mackie, A. M., Sharp, P. J. & Lagudah, E. S. The nucleotide and derived amino acid sequence of a HMW glutenin gene from *Triticum tauschii* and comparison with those from the D genome of bread wheat. *J Cereal Sci.* **24**, 73–78 (1996).
44. Yan, Y. *et al.* Capillary electrophoresis separation of high molecular weight glutenin subunits in bread wheat (*Triticum aestivum* L.) and related species with phosphate-based buffers. *Electrophoresis.* **24**, 1429–1436 (2003).
45. Lv, D. *et al.* Proteome and phosphoproteome characterization reveals new response and defense mechanisms of *Brachypodium distachyon* leaves under salt stress. *Mol Cell Proteomics* **13**, 632–652 (2013).
46. Yu, Z. *et al.* Rapid characterization of wheat low molecular weight glutenin subunits by ultra performance liquid chromatography (UPLC). *J Agr Food Chem.* **61**, 4026–4034 (2013).
47. Jin, M. *et al.* Identification and molecular characterisation of HMW glutenin subunit 1By16* in wild emmer. *J Appl Genet.* **53**, 249–258 (2012).
48. McDonald, M. B., Elliot, L. J. & Sweeney, P. M. DNA extraction from dry seeds for RAPD analyses in varietal identification studies. *Seed Sci Technol.* **22**, 171–176 (1994).
49. Arcalis, E. *et al.* Unexpected deposition patterns of recombinant proteins in post-endoplasmic reticulum compartments of wheat endosperm. *Plant Physiol.* **136**, 3457–3466 (2004).
50. López-Merino, L., Leroy, S. A., Haldorsen, S., Heun, M. & Reynolds, A. Can *Triticum urartu* (Poaceae) be identified by pollen analysis? Implications for detecting the ancestor of the extinct two-grained einkorn-like wheat. *Bot J Linn Soc.* **177**, 278–289 (2015).
51. Xu, H. *et al.* Functional properties of a new low-molecular-weight glutenin-subunit gene from a bread wheat cultivar. *Theor Appl Genet.* **133**, 1295–1303 (2006).
52. Weegels, P. L., Van de Pijpekamp, A. M., Graveland, A., Hamer, R. J. & Schofield, J. D. Depolymerisation and re-polymerisation of wheat glutenin during dough processing. I. Relationships between glutenin macropolymer content and quality parameters. *J Cereal Sci.* **23**, 103–111 (1996).
53. Rakszegia, M. *et al.* Technological quality of field grown transgenic lines of commercial wheat cultivars expressing the *1Ax1* HMW glutenin subunit gene. *J Cereal Sci.* **47**, 310–321 (2008).
54. He, Z., Liu, L., Xia, X., Liu, J. & Pena, R. J. Composition of HMW and LMW glutenin subunits and their effects on dough properties, pan bread, and noodle quality of Chinese bread wheats. *Cereal Chem.* **82**, 345–350 (2005).
55. Wang, A. *et al.* Characterization of two 1D-encoded ω -gliadin subunits closely related to dough strength and pan bread-making quality in common wheat (*Triticum aestivum* L.). *J Cereal Sci.* **47**, 528–535 (2008).
56. Li, X. *et al.* Molecular cloning, heterologous expression, and phylogenetic analysis of a novel γ -type HMW glutenin subunit gene from the G genome of *Triticum timopheevii*. *Genome* **50**, 1130–1140 (2007).
57. Johal, J., Gianibelli, M. C., Rahman, S., Morell, M. K. & Gale, K. R. Characterization of low-molecular-weight glutenin genes in *Aegilops tauschii*. *Theor Appl Genet.* **109**, 1028–1040 (2004).
58. Allaby, R. G., Banerjee, M. & Brown, T. A. Evolution of the high molecular weight glutenin loci of the A, B, D, and G genomes of wheat. *Genome* **42**, 296–307 (1999).

Acknowledgements

This research was financially supported by grants from and the National Special Project for Main Crop Breeding in China (JFYS2016ZY03002016), the National Natural Science Foundation of China (31471485), the National Science Foundation of Beijing City/Key Developmental Project of Science Technology, Beijing Municipal Commission of Education (KZ201410028031) and the National Key Project for Transgenic Crops in China (2016ZX08009). The authors would like to thank Prof. Renyong Zhao from College of cereals and oil food, Henan Science and Technology University for quality testing and analysis.

Author Contributions

Y.W. and S.Z. performed the research; Y.W., S.Z., N.L., X.L., C.H. and X.L. analysis the data; and Y.Y. conceived the study, designed the research. X.X. and Z.H. helped draft the manuscript. All authors read and approved the final manuscript.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Wang, Y. *et al.* Low molecular weight glutenin subunit gene *Glu-B3h* confers superior dough strength and breadmaking quality in wheat (*Triticum aestivum* L.). *Sci. Rep.* **6**, 27182; doi: 10.1038/srep27182 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>