



Article

Molecular Characterization and SNP-Based Molecular Marker Development of Two Novel High Molecular Weight Glutenin Genes from *Triticum spelta* L.

Yuemei Cao [†], Junwei Zhang [†] , Ruomei Wang, Haocheng Sun and Yueming Yan ^{*}

Beijing Key Laboratory of Plant Gene Resources and Biotechnology for Carbon Reduction and Environmental Improvement, College of Life Science, Capital Normal University, Beijing 100048, China

^{*} Correspondence: yanym@cnu.edu.cn

[†] These authors contributed equally to this work.

Abstract: Spelt wheat (*Triticum spelta* L., $2n=6x=42$, AABBDD) is a valuable source of new gene resources for wheat genetic improvement. In the present study, two novel high molecular weight glutenin subunits (HMW-GS) 1Ax2.1* at *Glu-A1* and 1By19* at *Glu-B1* from German spelt wheat were identified. The encoding genes of both subunits were amplified and cloned by allele-specific PCR (AS-PCR), and the complete sequences of open reading frames (ORF) were obtained. 1Ax2.1* with 2478 bp and 1By19* with 2163 bp encoded 824 and 720 amino acid residues, respectively. Molecular characterization showed that both subunits had a longer repetitive region, and high percentage of α -helices at the N- and C-termini, which are beneficial for forming superior gluten macropolymers. Protein modelling by AlphaFold2 revealed similar three-dimensional (3D) structure features of 1Ax2.1* with two x-type superior quality subunits (1Ax1 and 1Ax2*) and 1By19* with four y-type superior quality subunits (1By16, 1By9, 1By8 and 1By18). Four cysteine residues in the three x-type subunits (1Ax2.1*, 1Ax1 and 1Ax2*) and the cysteine in intermediate repeat region of y-type subunits were not expected to participate in intramolecular disulfide bond formation, but these cysteines might form intermolecular disulfide bonds with other glutenins and gliadins to enhance gluten macropolymer formation. The SNP-based molecular markers for 1Ax2.1* and 1By19* genes were developed, which were verified in different F2 populations and recombination inbred lines (RILs) derived from crossing between spelt wheat and bread wheat cultivars. This study provides data on new glutenin genes and molecular markers for wheat quality improvement.

Keywords: spelt wheat; HMW-GS; three-dimensional structure; SNP marker; quality improvement



Citation: Cao, Y.; Zhang, J.; Wang, R.; Sun, H.; Yan, Y. Molecular Characterization and SNP-Based Molecular Marker Development of Two Novel High Molecular Weight Glutenin Genes from *Triticum spelta* L.. *Int. J. Mol. Sci.* **2022**, *23*, 11104. <https://doi.org/10.3390/ijms231911104>

Academic Editors: Jinsong Bao and Jianhong Xu

Received: 26 August 2022

Accepted: 11 September 2022

Published: 21 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Wheat (*Triticum aestivum* L., $2n=6x=42$, AABBDD) is one of the three major food crops globally with a long history of cultivation, accounting for almost 35% of the human population's dietary needs [1]. Wheat is an allohexaploid species, and has a huge genome (up to 17 GB) containing A, B, and D subgenomes and a large amount of repetitive sequences [2]. Wheat flour can be processed into various foods such as bread, noodles, cookies as well as non-food products because of its versatile functional features and excellent textural properties, which are closely related to the structure of gluten and interactions within the protein complex in grains [3,4].

The gluten proteins include water-insoluble monomeric gliadins and polymeric glutenins, which are considered as the largest protein molecule in nature [5]. Gliadins act as a "plasticizer" to endow dough extensibility or viscous flow. The glutenin proteins include high and low molecular weight glutenin subunits (HMW-GS, LMW-GS), of which HMW-GS serve as "backbone" for interactions with other glutenin subunits and confer dough elasticity or strength [6,7]. HMW-GS are encoded by two tightly linked paralogous genes (larger x- and smaller y-type) located at the *Glu-A1*, *Glu-B1* and *Glu-D1* loci on the long

arm of the chromosome 1A, 1B and 1D, respectively. Theoretically, individual hexaploid wheat cultivars can express 6 different HMW-GS, but usually 3~5 HMW-GS are present due to gene silencing [8,9]. The allelic variations at *Glu-1* loci are closely associated with flour processing quality [10]. For example, the subunit pairs and subunits 1Dx5+1Dy10, 1Bx7+1By8, 1Bx13+1By16, 1Bx17+1By18, 1Ax1, 1Ax2* have positive effects on gluten quality whereas 1Dx2+1Dy12 and 1Bx20 are related to poor dough strength [11–13].

Typical HMW-GS have three structural domains: a long and variable central repeat region conferring elasticity to protein molecules, and flanked by the short and conservative N- and C-terminals, in which most or all cysteine residues are present [14,15]. The repeat region is rich in β -turns flanked by spherical conservative regions formed by α -helices [16]. Most of the cysteine residues form intra-chain disulfide bonds, some of them form inter-chain disulfide bonds and directly affect dough formation and rheological characteristics [17–19]. Generally, the β -turn structure confers protein molecules with significant deformation resistance [20]. In particular, the central repeat domain can be folded in the action of protein disulfide isomerase, which is the basis of dough extension [21]. However, although different methods such as homology modelling and fold identification algorithm were used for deciphering the structure features of HMW-GS [22,23], the detailed 3D structures of wheat glutenin subunits and gluten proteins are still not clear due to the complexity of the protein compositions and the difficulty of crystallization [24–26]. The recently developed protein predicting tool AlphaFold has provided the possibility to dissect the 3D structure of glutenin subunits, which is considered as the first computational method that can regularly predict protein structures with atomic accuracy even in cases in which no similar structure is known [27–29]. The recent study has shown that AlphaFold2 using the transformer design can obtain a high accuracy and is easy for 3D structure prediction of wheat MATE proteins [30].

In the past decades, a lot of studies have showed that the allelic variations at *Glu-1* in bread wheat are limited. However, more extensive *Glu-1* allelic variations are present in wheat related species such as spelt wheat [31,32], *T. dicoccum* [33], *Ae. tauschii* [34,35], *Ae. longissima* [36,37], *Ae. speltoides* and *Ae. kotschyi* [38]. Spelt wheat (*T. spelta* L., $2n=6x=42$, AABBDD) is closely related to common wheat and belongs to the same species with few differences among the subspecies [39–41]. As the first-class gene resource of wheat, spelt wheat is rich in nutritional value, various disease resistance factors, tolerance for various abiotic stresses and can be used for wheat genetic improvement through introducing the beneficial gene from spelt wheat [42–45]. In particular, some novel HMW-GS and subunit combinations in European spelt wheat varieties were identified such as 1Ax2.1*, 1Bx13*+1By19* and 1Bx6.1+1By22.1 [31]. These new allelic variations can be used as potential gene resources for improving wheat gluten quality. However, the molecular characterization of these novel glutenin genes and their application values for wheat quality improvement are still unknown.

To create novel germplasm for the improvement of wheat quality, we amplified and cloned the encoding genes of 1Ax2.1* and 1By19* subunits from European spelt wheat varieties by AS-PCR, and their molecular characterization and phylogenetic relationships among HMW-GS genes were investigated. Meanwhile, AlphaFold modeling was used to reveal the 3D structural features of 1Ax2.1* and 1By19* subunits and previously characterized superior HMW subunits. On the basis of the sequence variations of promoter region, the specific SNP-based markers for 1Ax2.1* and 1By19* genes were developed and validated by using a wide range of wheat cultivars, F2 populations and RILs from different crossings between spelt wheat and bread wheat cultivars. The results are discussed in relation to provide new gene resources and molecular markers for the improvement of wheat's breadmaking quality.

2. Results and Discussion

2.1. Identification of Novel HMW-GS in Spelt Wheat

HMW-GS compositions at *Glu-1* loci in Spelt 137 and Spelt 6 were identified by SDS-PAGE (Figure 1). The results showed that Spelt 137 contained 1Ax2.1*, 1Bx13+1By22* and 1Dx2+1Dy12 subunits at the *Glu-A1*, *Glu-B1* and *Glu-D1* locus, respectively. The new x-type subunit 1Ax2.1* encoded by *Glu-A1* located between 1Ax1 and 1Ax2* and had a lower electrophoretic mobility and higher molecular weight than 1Ax2* (Figure 1A). In addition, one new y-type subunit 1By22* was present at *Glu-B1* locus, which had a similar electrophoretic mobility and molecular weight with 1By18 subunit. The HMW-GS compositions in Spelt 6 were null at *Glu-A1*, 1Bx13*+1By19* at *Glu-B1* and 1Dx2+1Dy12 at *Glu-D1*. The new subunits 1Bx13* and 1By19* encoded by *Glu-B1* showed a similar electrophoretic mobility and molecular weight with 1Bx13 and 1By16 subunit, respectively (Figure 1B). The above two new subunits in spelt wheat were firstly found in our previous study [31]. These new allelic variations at *Glu-1* loci provide potential gene resources for wheat quality improvement.

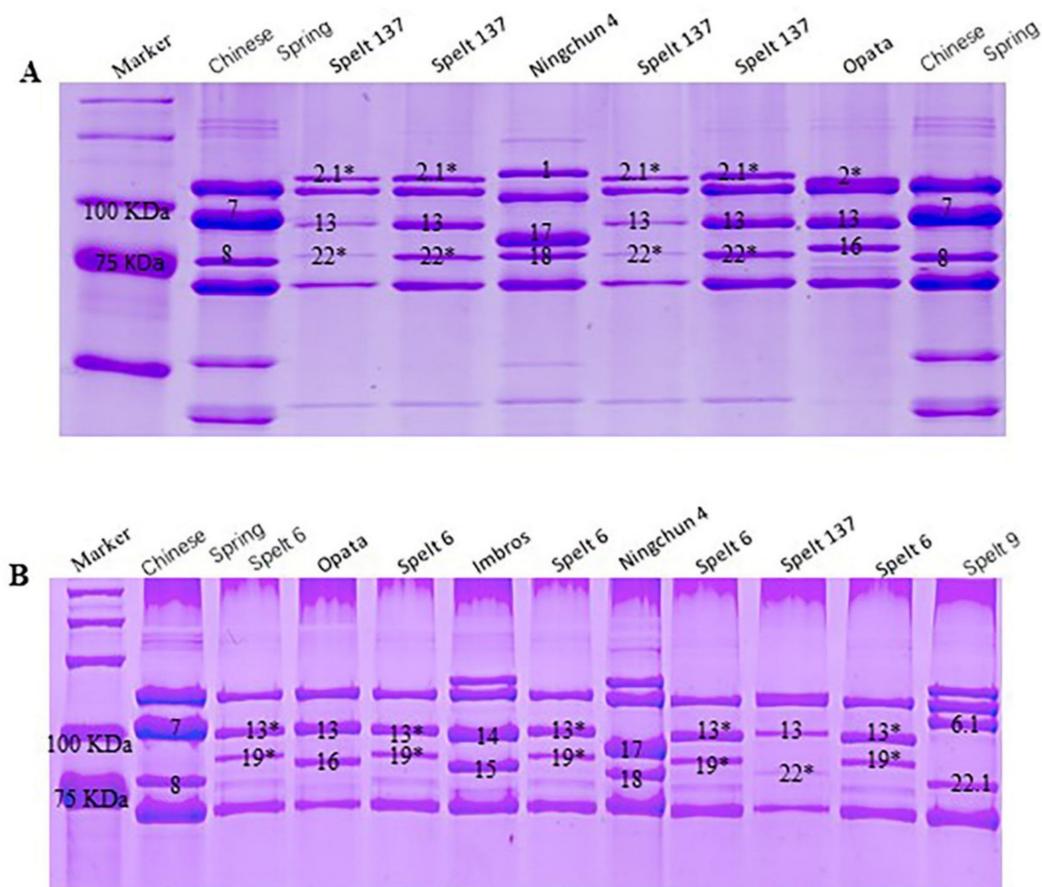


Figure 1. Identification of HMW-GS in European spelt wheat varieties Spelt 137 (A) and Spelt 6 (B) by SDS-PAGE. The HMW-GS in wheat is marked on the gel image with numbers or numbers with asterisks.

2.2. Molecular Cloning and Characterization of 1Ax2.1* and 1By19* Genes from Spelt Wheat

Two specific amplification fragments with about 2500 bp from Spelt 137 and 2200 bp from Spelt 6 were obtained by AS-PCR (Figure S1), which were corresponding to the sizes of x-type and y-type HMW-GS genes, respectively. Both PCR fragments were collected, cloned and sequenced, and their complete coding sequences with typical structural characteristics of previously characterized HMW-GS genes were obtained and designated as 1Ax2.1* and 1By19*. Both 1Ax2.1* and 1By19* gene sequences were deposited in GenBank with the accession numbers of MK395158 and MK395159, respectively. The nucleotide

sequences of *1Ax2.1** (2478 bp) and *1By19** (2163 bp) genes encoded 824 and 720 amino acid residues, respectively.

The deduced amino acid sequence of *1Ax2.1** gene showed four distinct domains as other x-type HMW-GSs in Figure 2: a signal peptide of 21 amino acid residues, an N-terminal sequence of 86 amino acid residues followed by a central repetitive domain of 672 residues, a C-terminal domain of 42 residues. The repeat region of *1Ax2.1** subunit consisted of tandem and interspersed repeats of 33 hexapeptide motifs (consensus PGQGQQ), 10 nonapeptide motifs (consensus GYYPTSPQQ and GYYPTSLQQ) and 23 tripeptide motifs (consensus GQQ). The deduced amino acid sequences of both genes were aligned with other HMW-GSs from wheat and related species to compare their sequence characteristics. Compared with the superior subunit *1Ax2** [46], *1Ax2.1** had 6 and 9 amino acid insertions, 6 amino acid deletions and 22 single amino acid substitutions, which ultimately increased the length of the repeat region. Differences in subunit size were mainly caused by the sequences of repetitive polypeptide motifs caused by base deletion or insertion [15,47]. It is known that the longer and more regular repeating units are usually more conducive to superior gluten quality [15,36]. *1Ax2.1** contained a longer repetitive domain with additional two hexapeptides (consensus PGQGQQ) than the superior quality subunit *1Ax2**. Besides, *1Ax2.1** had 6 amino acids deletion in repeat domain compared with the superior quality *1Ax1*, but the tripeptides (GQQ), hexapeptides (PGQGQQ) and nonapeptides (GYYPTSPQQ) were highly consistent. And just like *1Ax2.1**, the number and distribution of cysteine residues of these subunits were highly conserved between them, including sites 31, 46, 61 at the N-terminal and site 812 or 818 or 803 at the C-terminal.

Similarly, *1By19** gene encoded 720 amino acid residues, and showed a similar primary structures with y-type HMW-GSs in Figure 3, including a signal peptide of 21 amino acid residues, an N-terminal of 104 amino acid residues, a repeat region of 553 amino acid residues and a C-terminal of 42 amino acid residues. The repeat region contained 8 nona- and 34 hexapeptides. In addition, *1Ax2.1** subunit contained four conserved cysteine residues (three at the N-terminus and one at the C-terminus) while *1By19** subunits had seven conserved cysteine residues (five at the N-terminus, one at the repetitive domain, and one at the C-terminus). Likewise, the repeat unit in *1By19** accounted for about 50% of the total amino acids in the repeat region, in which the number of peptide segment PGQGQQ was equivalent to the y-type superior quality subunit *1Dy10*. Overall, *1By19** had the same length of amino acid sequence as the superior quality subunits *1By8* and *1By18*. But the amino acid length of *1By19** located between the superior quality subunits *1By9* and *1By16* due to the insertion and deletion of amino acids in the middle repeat region. Besides, these subunits all contained the same number of nonapeptides (consensus GYYPTSLQQ), but *1By19** contained 1~2 additional hexapeptides (consensus PGQGQQ). And each of these subunits contained seven cysteine residues, including site 606 or 591 or 624 near the C-terminus flanked by sites 31, 43, 65, 66 and 76 at N-terminus and site 708 or 693 or 726 at the C-terminus.

2.3. SNP and InDel Variations in *1Ax2.1** and *1By19** Genes

Sequence variations resulting from point mutation and insertion/deletions (InDels) are the most common cause of wheat storage protein variation [48,49]. As listed in Table 1, 11 SNPs in *1Ax2.1** gene were present at different positions in N-terminal domain and in repeat region. Among them, 8 SNPs were nonsynonymous and resulted in the changes of corresponding amino acid residues of H-Y, L-S/P/G, E-G/T/Q, R-G, R-G/L, A-P/G, L-S/P. The remaining three SNPs at positions 198 bp (T-C), 288 bp (G-A) and 1302 bp (T-A/C) were synonymous and did not cause amino acid residue changes. One nonsynonymous SNP variation in *1By19** gene was detected: A/G-T transversion at the positions of 1137 bp, which led to the amino acid residue change (Q/E/G-H). No InDel variations were found in *1By19** gene.

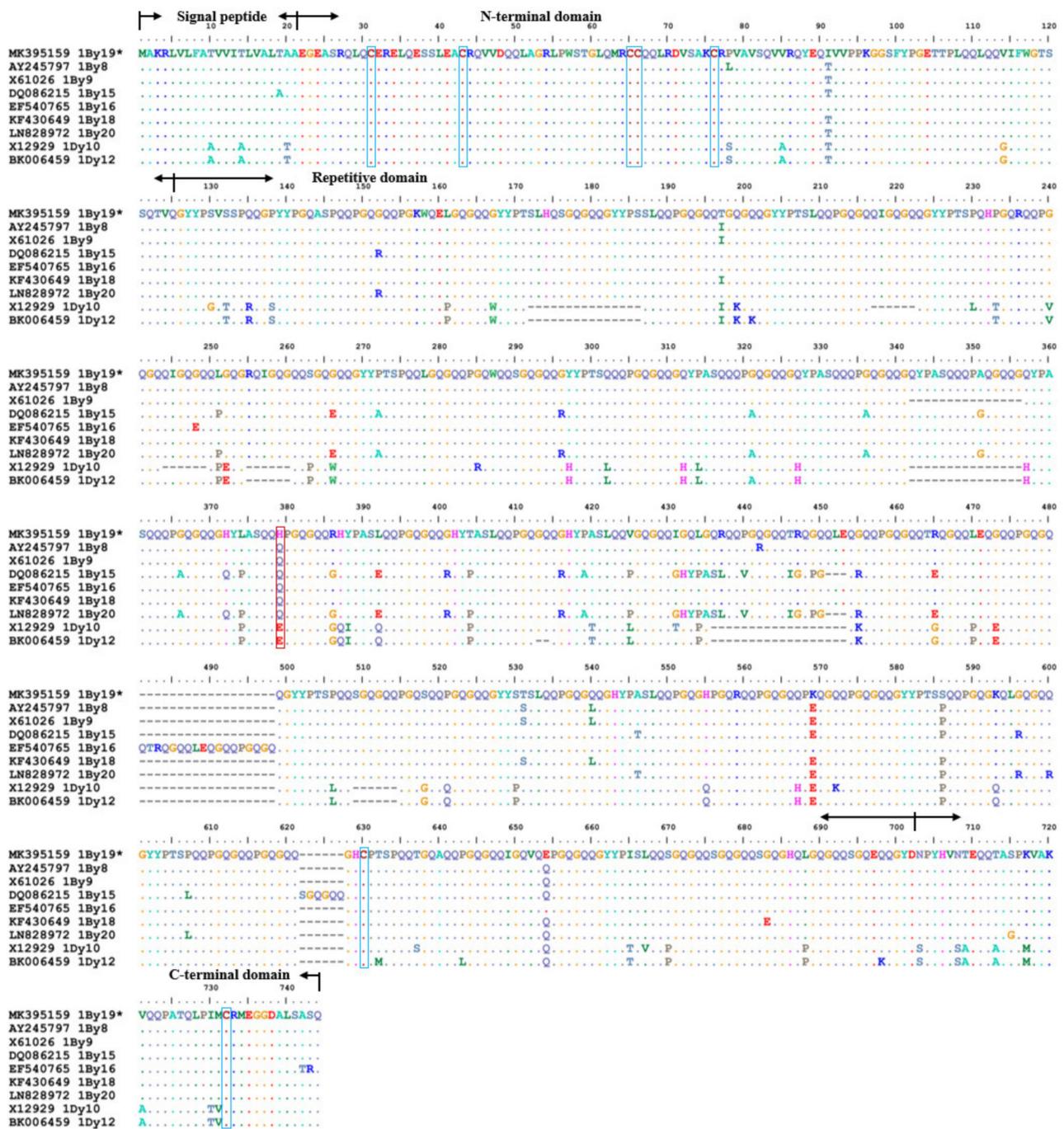


Figure 3. Comparison of the deduced amino acid sequences of *1By19** gene from spelt wheat with other eight y-type subunit genes. The blue boxes indicate the cysteine residues, red boxes indicate differences of amino acid.

Table 1. Positions of SNPs identified between *1Ax2.1** and other x-type HMW-GS genes.

Positions (bp)	198	288	325	1076	1302	1448	1696	1777	1813	1815	2066
<i>1Ax2.1*</i>	T	G	C	T	T	A	A	A	G	G	T
15 other x-type HMW-GS genes	C	A	T	C/-	A/C	G	G	G/-	C/T	A	C

2.4. Verification of the Cloned 1Ax2.1* and 1By19* Genes from Spelt Wheat by Tandem Mass Spectrometry Analysis

The corresponding HMW-GS on the SDS-PAGE gel were collected and digested, and then identified by MALDI-TOF/TOF-MS (Table 2). The results showed that 1Ax2.1* and the x-type subunit MK395158 could well match in five peptide segments with 35 (45–79), 24 (195–218), 49 (357–405), 6 (725–730) and 24 (792–815) amino acid residues, respectively (Figure 2). 1By19* was well matched to the y-type subunit MK395159 in one peptide segment of 43 residues (QLQCERELQESSLEACRQVVDDQQLAGRLPWSTGLQMRCCQQLR) at the position of 28–70 amino acid (Figure 3). All peptide segments identified by tandem mass spectrometry acquired a high protein score (C.I.%) of 100, indicating a high reliability of MS/MS analysis. These results further verified the validity of the cloned 1Ax2.1* and 1By19* genes.

Table 2. Identification of 1Ax2.1* and 1By19* subunits by MALDI-TOF/TOF-MS.

HMW-GS	Identified Protein	Accession No	Tryptic Fragments Identified by MS Data	Positions	Matched Peptides	Protein Score C. I. %
1Ax2.1*	x-type	MK395158	QVVDQQLRDVSPECQPVGGG	45–79	5	100
			PVARQYEQVVPK			
			QQPGQQQLRQGGQQSGGQPR	195–218		
			QQDQSGGQQPGQRQPGY	357–405		
			YSTSPQQLGQGQPRYPYPTSPQQPGQEQQPR	725–730		
QWLQPR	792–815					
1By19*	y-type	MK395159	AQQLAAQLPAMCRLEGGDALLASQ QLQCERELQESSLEACRQVVDD QQLAGRLPWSTGLQMRCCQQLR	28–70	1	100

2.5. Phylogenetic Analysis of Ax2.1* and 1By19* Genes

The polymorphism of glutenin subunits is beneficial for understanding the origin and evolution of spelt wheat. The early study on the allelic variations at *Glu-1* and *Glu-3* in spelt wheat varieties supported the hypothesis of secondary origin of spelt wheat from hybridizing between cultivated emmer (*T. dicoccum*, AABB) and club wheat (*T. aestivum* subsp. *compactum*, AABBDD) [31]. Here, the complete encoding sequences of the cloned 1Ax2.1* and 1By19* genes from spelt wheat and other 27 HMW-GS genes from common wheat and related species were used to construct a neighbor-joining phylogenetic tree (Figure 4). The results showed that the x-type and y-type HMW-GS genes in the phylogenetic tree were clearly separated into two clades. In the x-type gene clade, 1Bx genes together with two Sx genes from S genome were classified into a subgroup while the Dx and Ax genes were classified into another subgroup. In particular, 1Ax2.1* gene from spelt wheat showed a close phylogenetic relationship with 1Ax1, 1Ax2* and 1Ax2.1 genes. In the y-type gene clade, the Dy and By genes were classified into different subgroups while 1By19* gene from spelt wheat was more closed to 1By16, 1By9, 1By8 and 1By18 genes.

To further explore the evolutionary relationships between 1Ax2.1* and 1By19* and other HMW-GS genes, their divergence times (million years ago, MYA) were estimated and the results are shown in Table S3. In general, the divergent times between x-type subunit genes between y-type subunit genes at each locus were about 0–13 MYA and 0.2–6.0 MYA while the divergence between x- and y-type subunit genes in each locus occurred about 14–17 MYA. The divergence among 1Ax2.1* and 1Ax1 and 1Ax2* genes occurred more recently, at about 0.92–1.00 MYA while earlier divergence between 1Ax2.1* and 1Ax2.1 occurred at 2.31 MYA. 1By19* displayed a high similarity with 1By16 and they were diverged more recently at 0.31 MYA. It is known that the hexaploid wheat is originated from the hybridation between tetraploid wheat and *T. tauschii* that occurred at about ten-thousand years ago. Thus, 1Ax2.1* and 1By19* genes should emerge before spelt wheat formation.

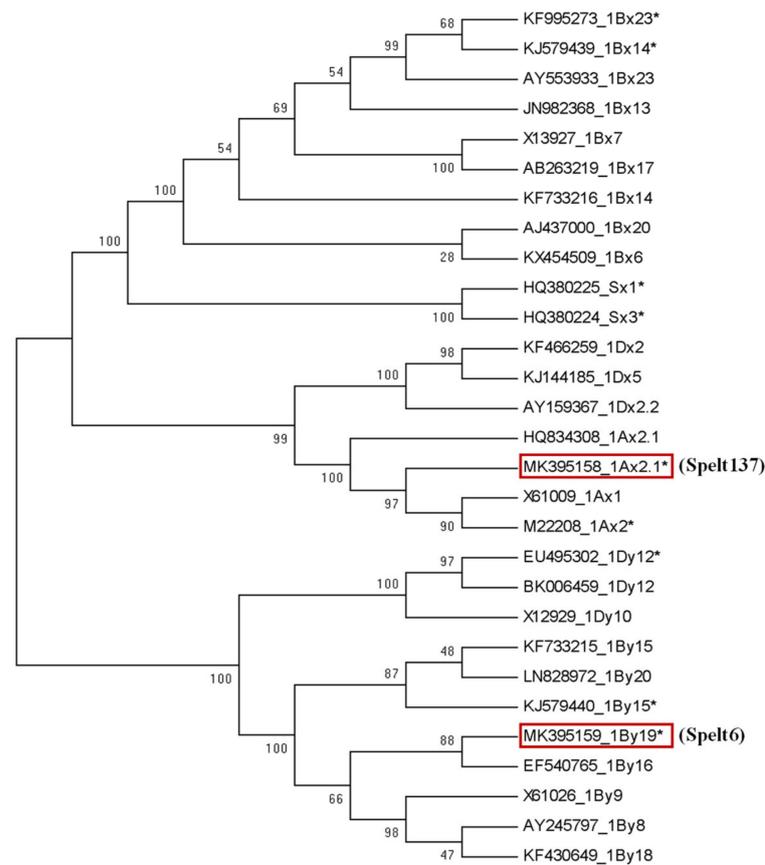


Figure 4. Phylogenetics of *1Ax2.1** and *1By19** genes and other 27 HMW-GS genes from A, B and D genomes of common wheat and related species. Bootstrap value $\geq 28\%$ is indicated above or below the branches. The red frame indicates *1Ax2.1** and *1By19** genes, respectively. The 27 HMW-GS genes include X61009, M22208, KX454509, X13927, JN982368, KF733216, KJ579439, AB263219, AY553933, KF466259, AY159367, KJ144185, X61026, KF733215, KJ579440, EF540765, KF430649, X12929, BK006459 from *T. aestivum*; HQ380225 and HQ380224 from *Ae. speltoides*; KF995273 from *Triticum turgidum*; EU495302 from *T. aestivum* subsp. *Yunnanense*; AJ437000, LN828972 and AY245797 from *T. turgidum* subsp. *Durum*; HQ834308 from *T. monococcum* subsp. *Monococcum*.

2.6. Secondary Structure and 3D Structure Analysis of *1Ax2.1** and *1By19** Protein Subunits

The secondary structure characteristics of glutenin proteins can be used to uncover the molecular mechanisms of gluten quality formation [50]. In general, higher content of β -sheets and α -helix in glutenin subunits is helpful for forming good gluten structure and superior breadmaking quality [51,52]. In this study, the secondary structures of *1Ax2.1** and *1By19** subunits and other eight wheat HMW-GS (*1Ax2**, *1Ax1*, *1By8*, *1By9*, *1By15*, *1By16*, *1By18* and *1By20*) were predicted by the PSIPRED server and a comparative analysis for their secondary structure features was performed. Among them, *1Ax2**, *1Ax1*, *1By8*, *1By15*, *1By16* and *1By18* are considered as superior quality subunits whereas *1By20* is a poor quality subunit [11–13]. As shown in Table 3, *1Ax2.1** subunit had seven α -helices (9.58%) and four β -strands (0.82%), better than the superior quality subunits *1Ax2** with the same seven α -helices (9.33%) and two β -strands (0.98%), and *1Ax1* with six α -helices (9.40%) and without β -strands. *1By19** subunit and other six y-type subunits generally contained eight α -helices with different percentages, but no β -strands were present. *1By19** had 12.36% α -helices, higher than *1By15* (12.03%), *1By16* (11.65%), *1By18* (12.22%), and *1By20* (11.85%), and lower than *1By8* (16.81%) and *1By9* (13.19%).

Table 3. Secondary structure prediction of 1Ax2.1* and 1By19* subunits and other eight wheat HMW-GS.

HMW-GS (Accession No.)	Type	Structure Motifs	Content (%)	Total	Amino Acid Length
1Ax2.1* (MK395158)	x-type	α -helix	9.58	7	824
		β -strand	0.82	4	
1Ax2* (M22208)	x-type	α -helix	9.33	7	815
		β -strand	0.98	2	
1Ax1 (X61009)	x-type	α -helix	9.40	6	830
		β -strand	0.00	0	
1By19* (MK395159)	y-type	α -helix	12.36	8	720
		β -strand	0.00	0	
1By8 (AY245797)	y-type	α -helix	16.81	8	720
		β -strand	0.00	0	
1By9 (X61026)	y-type	α -helix	13.19	8	705
		β -strand	0.00	0	
1By15 (DQ086215)	y-type	α -helix	12.03	6	723
		β -strand	0.00	0	
1By16 (EF540765)	y-type	α -helix	11.65	8	738
		β -strand	0.00	0	
1By18 (KF430649)	y-type	α -helix	12.22	8	720
		β -strand	0.00	0	
1By20 (LN828972)	y-type	α -helix	11.85	8	717
		β -strand	0.00	0	

The 3D structures of 1Ax2.1* and 1By19* subunits and other six superior quality subunits 1Ax2*, 1Ax1, 1By8, 1By9, 1By16 and 1By18 were further predicted by AlphaFold2 (Figure 5). As a deep learning algorithm, AlphaFold2 can visualize, analyze and interpret 3D protein structures basing on amino acid sequences [53]. The results showed that Alphafold2 predictions with five cycles for eight HMW-GS were almost identical, including three indicators of alignment errors, per-residue confidence scores, matching templates (Figure S2). Eight subunits had different confidence scores in 3D structure models after five cycles of prediction, which were divided into five different ranks according to their scores (Figure S3). The highest level for eight subunits is shown in Figure 5, and the confidence score for 1Ax2.1* was 33.9, slightly close to 1Ax1 (34.9) and 1Ax2* (34.4). The confidence score for 1By19* was 36.9, slightly close to 1By16 (37.4), 1By9 (37.7), 1By8 (37.2) and 1By18 (37.0). All subunits predicted had a typical 3D structure, including two non-repetitive structural domains (N-terminal and C-terminal) with a repetitive fragment in the middle and a large number of α -helices at the N- and C-termini. In general, three x-type subunits (1Ax2.1*, 1Ax1 and 1Ax2*), and five y-type subunits (1By19*, 1By16, 1By9, 1By18 and 1By8) showed a similar 3D structural characteristics, respectively. It is worth noting that the number and distribution of Cys residues play an important role in determining the formation of glutenin polymer and the subsequent rheological parameters of dough [54–56]. Particularly, three x-type subunits (1Ax2.1*, 1Ax1 and 1Ax2*) contained four cysteine residues (three in the N-terminal domain and one in the C-terminal domain), which did not form intramolecular disulfide bonds within the subunits. Five y-type subunits (1By19*, 1By16, 1By18, 1By8 and 1By9) contained seven cysteine residues (five in the N-terminal domain, one in the C-terminal domain and one in the intermediate repeating structural domain), and formed three disulfide bonds: Cys 31~76, 43~65 and 66~706 in 1By19*, Cys 31~76, 43~65 and 66~726 in 1By16, Cys 31~76, 43~65 and 66~693 in 1By9, Cys 31~76, 43~65 and 66~708 in 1By8, and Cys 31~76, 43~65 and 66~708 in 1By8. However, the cysteines at the intermediate repeat region were unable to form a disulfide bond (Figure 5). Meanwhile, 1Ax2.1* and 1By19* subunits contained a large number of α -helices at the N- and C-termini, consistent with the above mentioned secondary structure prediction results as well as previous report [57,58]. Although the cysteines in 1Ax2.1*, 1Ax1 and 1Ax2*

subunits and three cysteines in 1By19*, 1By16, 1By9, 1By8 and 1By18 subunits did not form intramolecular disulfide bonds, they might form intermolecular disulfide bonds with other glutenins or gliadins to further form gluten macropolymers [59].

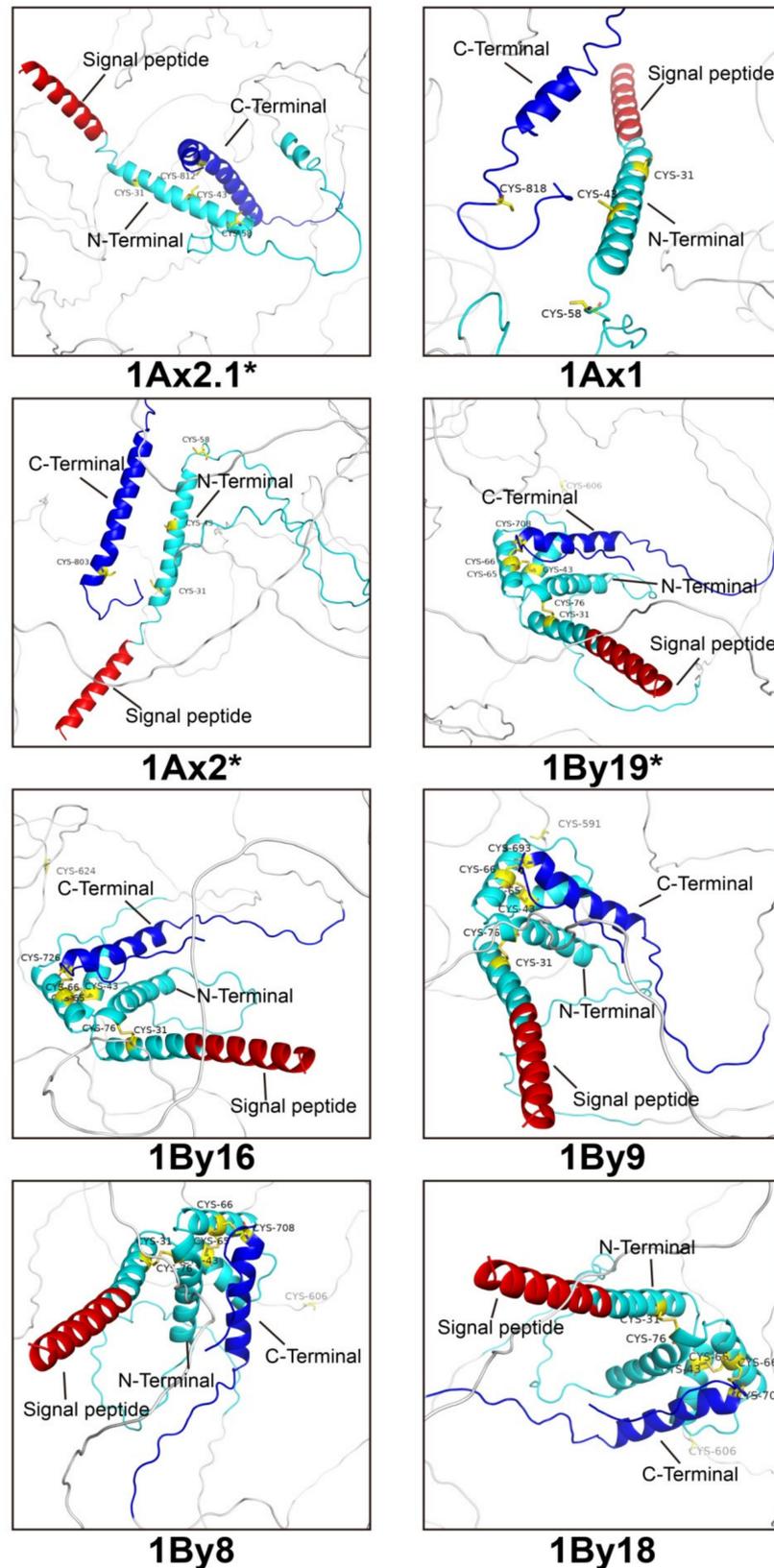


Figure 5. Localized view of the 3D structures of eight wheat HMW-GS by AlphaFold2.

2.7. Development and Validation of SNP-Based Molecular Markers for 1Ax2.1* and 1By19* Genes

A new marker type named SNP, is only a bi-allelic type of marker in nature with low expectation of heterozygosity [60,61]. In the past years, a number of SNPs-based molecular markers for glutenin genes have been developed such as 1Dx2 and 1Dx5 alleles [62], 1By18 [49], 1Slx2.3* and 1Sly16* [36]. These markers can be for rapid improvement of wheat gluten quality. Herein, we amplified and cloned the upstream promoter sequences of 1Ax2.1* and 1By19* genes. Then BioEdit7.0 was used to compare their promoter sequence differences with other x- and y-type HMW-GS genes and the specific SNP variations were identified. Ultimately, we used these SNP variations to develop SNP-based molecular markers for 1Ax2.1* and 1By19* genes.

The specific primers were designed (Table S4), and then used to amplify the upstream promoter sequences of 1Ax2.1* and 1By19* genes, and two specific fragments 905 bp and 728 bp from 1Ax2.1* and 1By19* genes were obtained, respectively (Figure S4). After cloning and sequencing, the typical promoter sequences of 1Ax2.1* and 1By19* genes were obtained, which were used for sequence alignment with the promoter sequences of other 12 x- and y-type HMW-GS genes deposited in GenBank (Figure S5 and S6). Two specific SNP sites in the upstream promoter region of each gene were identified: −377 bp and −203 bp in 1Ax2.1* and −591 bp and −56 bp in 1By19*. According to these specific SNP sites, two pairs of specific primers (2.1*F/R and 19*F/R) were designed (Table S5) and used to develop SNP-based molecular markers of 1Ax2.1* and 1By19* genes as showed in Figures S5 and S6 via combining with SDS-PAGE identification. The results from 47 bread wheat and spelt wheat varieties with different *Glu-1* allelic variations (Table S1) showed that one 209 bp specific fragment was amplified by using 2.1*F/R primer pair from the varieties containing 1Ax2.1* subunit whereas no amplified products were obtained in the varieties without 1Ax2.1* subunit (Figure 6A,B). Similarly, one 570 bp specific fragment was amplified by the 19*F/R primer pair from varieties with 1By19* subunit, which was absent in the varieties without 1By19* subunit (Figure 6C,D). These results were well consistent with the SDS-PAGE identification. Further collection and sequencing of both amplified fragments also showed a consistence with their amplified regions as shown in Figures S5 and S6.

To further validate the developed SNP markers, two pairs of specific primers were used to amplify the promoter sequences of the target genes in different hybrid populations and RILs. Meanwhile, SDS-PAGE identification was conducted to verify the PCR results. For the marker validation, 200–250 grains from F2 generation populations and 50–80 grains from each RIL were detected. In the F2 populations of Spelt 137×Zhongmai 175, Spelt 137×Ningchun 4 and Spelt 137×Zhongmai 8601 crosses, 2.1*F/R primer could specifically amplify the 209 bp fragment in the F2 grains containing 1Ax2.1* subunit, and no amplified products were produced in the F2 grains without 1Ax2.1* subunit. At the same time, three F2 populations of Spelt 6×Zhongmai 175, Spelt 6×Ningchun 4 and Spelt 6×Zhongmai 8601 crosses were used to verify the molecular marker of 1By19* gene and the similar results were obtained. The F2 grains having 1By19* subunit showed the 570 bp specific fragment when amplified by 19*F/R primer pair (Figure 7). Furthermore, seven RILs from Spelt 137×Ningchun 4 crossing and six RILs from Spelt 6×Ningchun 4 crossing were further used to verify the developed molecular markers for 1Ax2.1* and 1By19* genes. As expected, all RILs containing 1Ax2.1* and 1By19* subunits could exhibit 209 bp and 570 bp specific fragments, respectively (Figure 8). These results were well consistent with SDS-PAGE identification, confirming that the developed SNP-based molecular markers have high specificity and accuracy for identifying 1Ax2.1* and 1By19* genes. Therefore, these SNP-based markers are expected to be used for breadmaking quality improvement via marker-assisted selection during wheat breeding program.

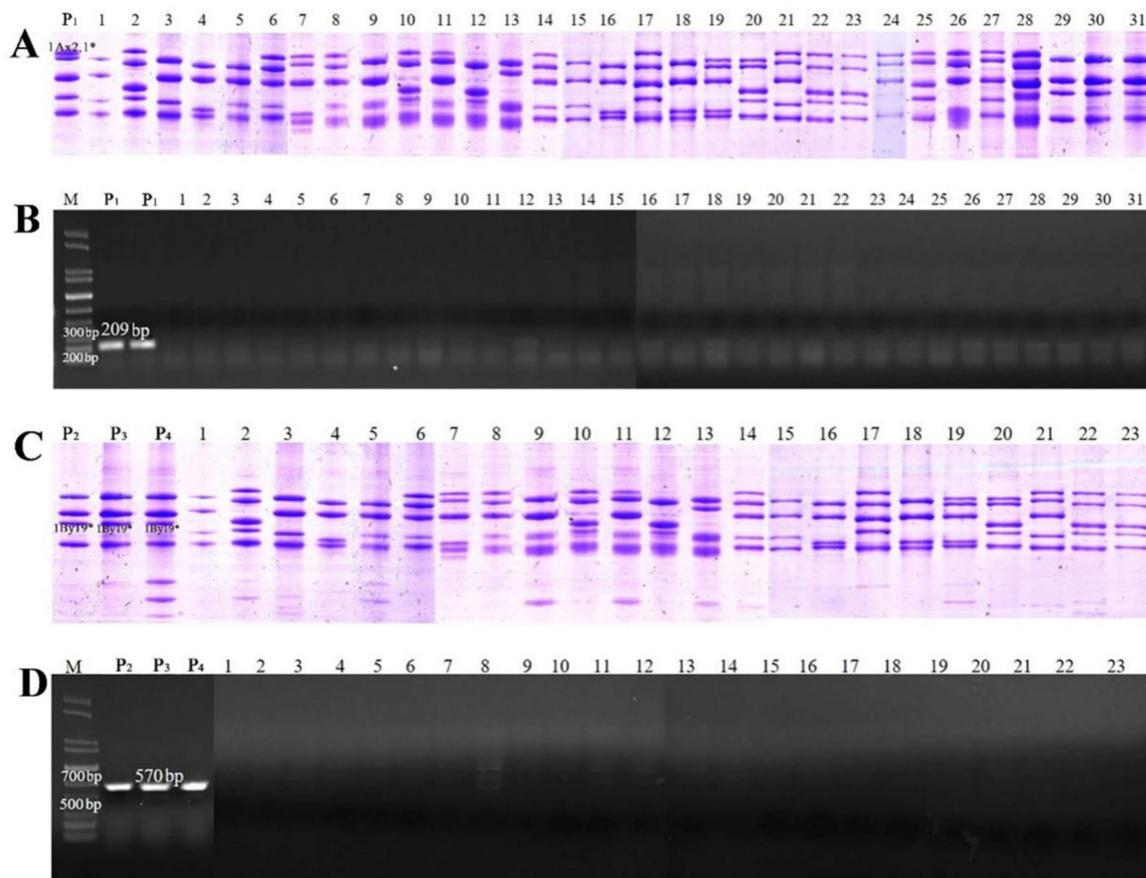


Figure 6. Development of the SNP-based molecular markers for *1Ax2.1** and *1By19** genes. (A) Identification of *1Ax2.1** subunit in different wheat varieties by SDS-PAGE. (B) PCR amplification of *1Ax2.1** gene from different wheat varieties by the primer 2.1*F/R. The 209 bp specific amplified fragment is indicated. (C) Identification of *1By19** subunit in different wheat varieties by SDS-PAGE. (D) PCR amplification of *1By19** gene from different wheat varieties by the primer 19*F/R. The 570 bp specific amplified fragment is indicated. The varied numbers in the figure are the same as those in Table S1. P₁: Spelt 137; P₂: Spelt 6; P₃: Spelt 20; P₄: Spelt 24.

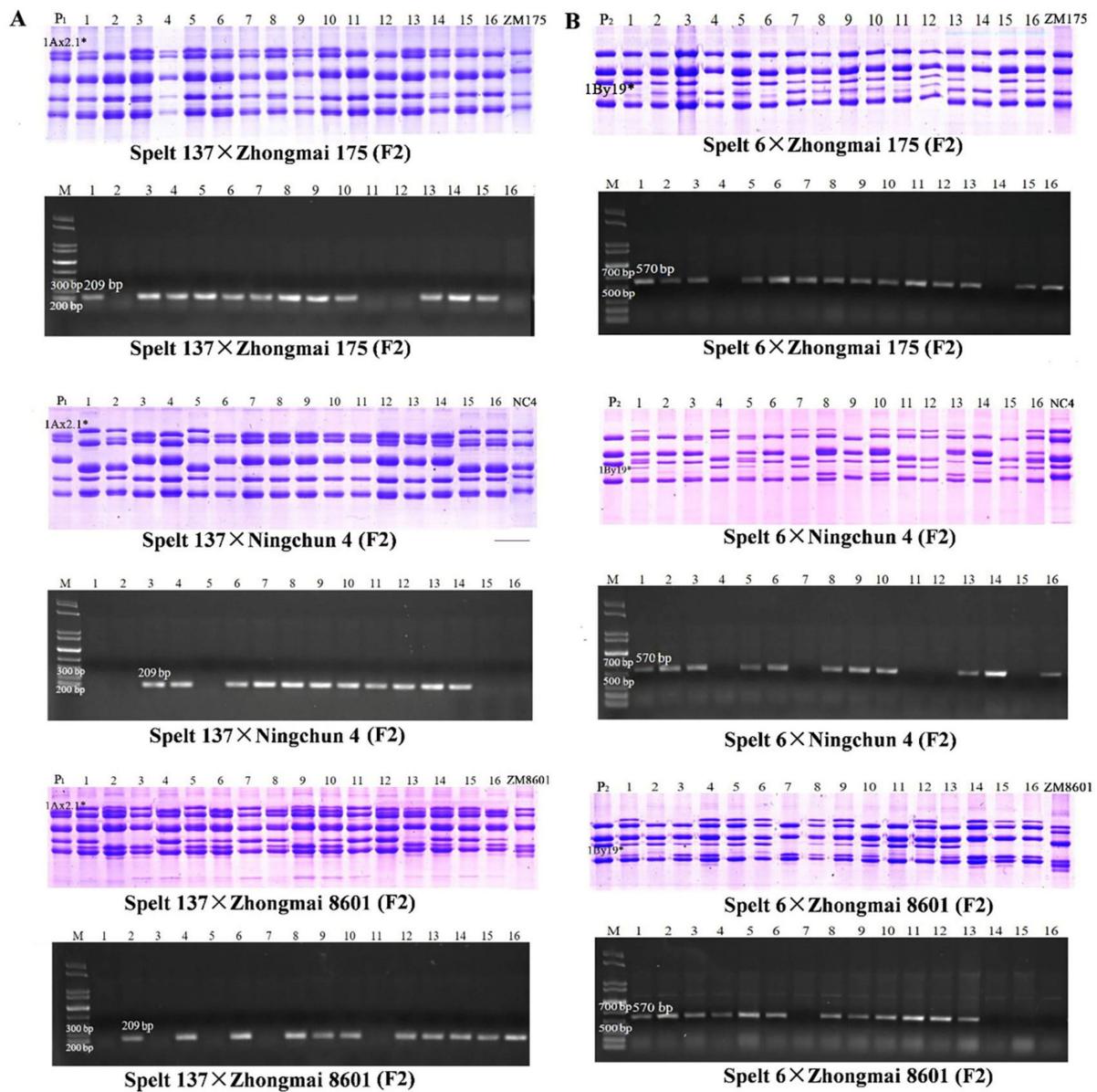


Figure 7. Validation of the SNP-based molecular marker for *1Ax2.1** (A) and *1By19** (B) genes in the F2 populations from crossing between spelt wheat and bread wheat varieties. *1Ax2.1** and *1By19** subunits and their special amplification fragments 209 bp and 570 bp are indicated. P₁: Spelt 137; P₂: Spelt 6; ZM175: Zhongmai 175; NC4: Ningchun 4; ZM8601: Zhongmai 8601.

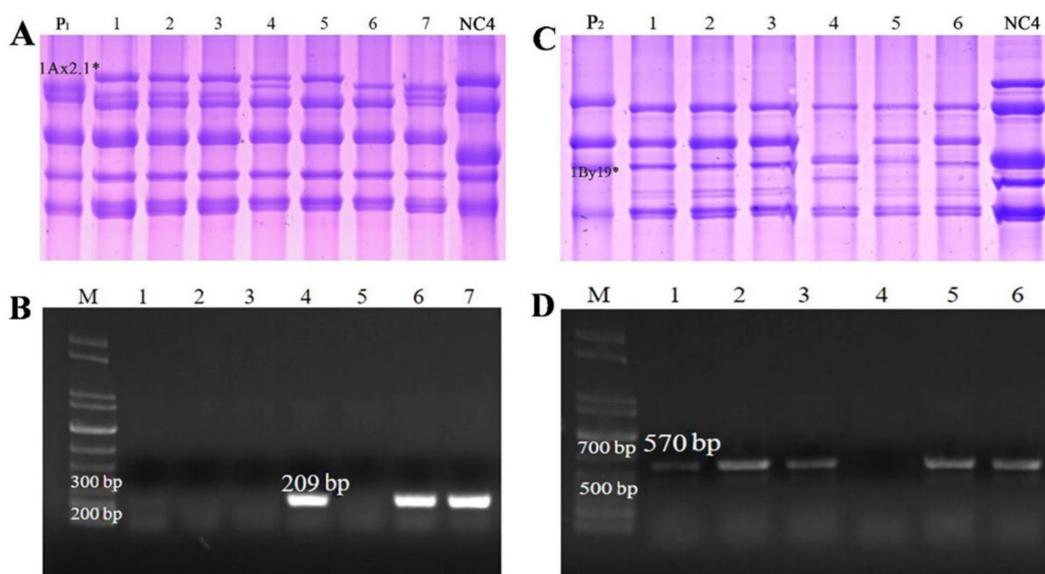


Figure 8. Validation of the SNP-based molecular markers for *1Ax2.1** and *1By19** genes in the RILs derived from Spelt 137×Ningchun 4 and Spelt 6×Ningchun 4 crosses. (A) HMW-GS from seven RILs (Lanes 1–7) from Spelt 137×Ningchun 4 identified by SDS-PAGE; (B) PCR amplification of *1Ax2.1** gene from seven RILs (Lanes 1–7) of Spelt 137×Ningchun 4; (C) HMW-GS from six RILs (lanes 1–6) from Spelt 6×Ningchun 4 identified by SDS-PAGE; (D) PCR amplification of *1By19** gene from six RILs (lanes 1–6) from Spelt 6×Ningchun 4. *1Ax2.1** and *1By19** subunits and the special amplification fragments 209 bp and 570 bp are indicated. P₁: Spelt 137; P₂: Spelt 6; NC4: Ningchun 4.

3. Materials and Methods

3.1. Plant Materials

The materials used in this work included two European spelt wheat varieties Spelt 6 and Spelt 137, 45 bread wheat and spelt wheat varieties with different allelic variations at *Glu-1* loci (Table S1), six F₂ generation populations from Spelt 137×Zhongmai 175, Spelt 137×Ningchun 4, Spelt 137×Zhongmai 8601, Spelt 6×Zhongmai 175, Spelt 6×Ningchun 4, and Spelt 6×Zhongmai 8601, seven and six F₆ RILs, respectively, derived from crossing between Spelt 137×Ningchun 4 and Spelt 6×Ningchun 4 via consecutive self-crossing combining with screening and identification. All spelt wheat varieties were collected from Plant Breeding Institute, Technical University of Munich, Germany.

3.2. HMW-GS Extraction and SDS-PAGE

The extraction of seed HMW-GS was carried out with referred to the previous study with minor modifications [63]. Firstly, 70% ethanol (*v/v*) and 55% isopropanol (*v/v*) were sequentially added to homogenized seeds to remove albumins, globulins and gliadins. Then, glutenins were extracted by a commonly used glutenin extraction buffer (50% isopropanol, 80 μL Tris-HCl, pH 8.0) with 1% dithiothreitol (*w/v*) and 1.4% 4-vinylpyridine (*w/v*). The final extracted glutenin proteins were used for SDS-PAGE based on the reported method [34].

3.3. DNA Isolation, AS-PCR Amplification and Sequencing

The wheat seedling leaves were used to extract genomic DNA according to the improved CTAB method [37]. Two pair of AS-PCR primers were designed according to the published HMW-GS gene sequences and listed in Table S2. The high-fidelity polymerases (Vazyme Biotech, Nanjing, China) were used for amplifying the complete coding sequence of HMW-GS. PCR reaction was performed by CFX96 Real Time system (Bio-Rad Laboratories) programmed at an initial denaturation at 95 °C for 3 min followed by 35 cycles of 95 °C for 15 s, 61 °C for 15 s, 72 °C for 150 s and finally extended at 72 °C for 5 min. The

PCR products were separated by 1% agarose gel in Tris-acetic acid-EDTA buffer and the fragments of expected size were collected and purified by using Gel Extraction Kit (Omega, Bienne, Switzerland), and then the purified products were ligated into pMD18-T vector (TaKaRa Biotechnology, Dalian, China) and transformed into Trans-2-blue competent cells. Three positive clones were randomly selected to reduce the sequencing error, and then sequenced by TaKaRa Biotechnology, Dalian, China.

3.4. Sequence Alignment and SNP/InDel Identification

Multiple sequences alignments of the cloned HMW-GS genes and previously characterized x-type and y-type subunit genes were conducted by using Bioedit (version 7.0, Tom Hall, Scotts Valley, USA) SNPs and InDels variations in HMW-GS genes were identified. 15 x-type and 12 y-type HMW-GS genes deposited in GenBank are as follows: 1Ax1 (X61009), 1Ax2* (M22208), 1Ax2.1 (HQ834308), 1Bx6 (KX454509), 1Bx7 (X13927), 1Bx13 (JN982368), 1Bx14 (KF733216), 1Bx17 (AB263219), 1Bx20 (AJ437000), 1Bx23 (AY553933), 1Dx2 (KF466259), 1Dx2.2 (AY159367), 1Dx5 (KJ144185), Sx1* (HQ380225), and Sx3* (HQ380224), 1Ay (FJ404595), 1By8 (AY245797), 1By9 (X61026), 1By15 (KF733215), 1By15* (KJ579440), 1By16 (EF540765), 1By18 (KF430649), 1By20 (LN828972), 1Dy10 (X12929), 1Dy12 (BK006459), Sy9* (HQ380223), and Sy18* (HQ380222).

3.5. MALDI-TOF/TOF-MS

According to the previously described method [49,51], the corresponding HMW subunits bands on SDS-PAGE gel were excised, and then digested with trypsin. HMW-GS identification of tandem mass spectrometry was performed by using EASY-nLC 1000 (Thermo/Finnigan, San Jose, CA, USA) equipped with Orbitrap Q Exactive mass spectrometry (Thermo/Finnigan).

3.6. Construction of Phylogenetic Tree and Estimation of Divergence Time

The full-length homologous nucleotide sequences were aligned using the Clustal W program, and the alignment file was used to construct phylogenetic tree according to the complete coding regions of HMW-GS by using software MEGA (version 6.0, Koichiro Tamura et al., Auckland, New Zealand). An evolution rate of 6.5×10^{-9} substitution/site year was used to estimate the divergence times of HMW-GS genes based on the reported method [64].

3.7. Secondary Structure and 3D Structure Prediction of HMW-GS

Secondary structure prediction of the deduced amino acid sequences from the cloned HMW-GS genes were carried out by PSIPRED (<http://bioinf.cs.ucl.ac.uk/psipred/psiform.html>, accessed on 10 May 2022) [65,66]. The 3D structures of HMW-GS were predicted by AlphaFold2 [67,68], and then editing was carried out by Pymol software (version 1.7.4, Schrödinger, Warren Lyford DeLano, New York City, USA).

3.8. Development and Validation of SNP-Based Molecular Markers

According to the SNP variations in the upstream promoter sequences of the coding region in the cloned HMW-GS genes (Table S4), two pairs of AS-PCR primers were designed (Table S5). Genomic DNA was extracted using the Trans Fast Taq DNA Polymerase system (TransGen Biotech, Beijing, China). PCR cycles consisted of 94 °C for 3 min for activation, followed by 35 cycles of 94 °C for 5 s, 58 °C/55 °C (Table S6) for 15 s, 72 °C for 2/4 s, and finally ended at 72 °C for a 5 min extension step. The materials used for marker development and verification were described in the section of plant materials.

4. Conclusions

Two novel HMW-GS 1Ax2.1* and 1By19* were identified in European spelt wheat, and their complete encoding gene sequences of corresponding genes were cloned and sequenced. Molecular characterization showed that both subunits had longer and more

regular repeating units, and a high percentage of α -helices and β -sheets. AlphaFold2 prediction showed that 1Ax2.1* and 1By19* showed similar 3D structural characteristics with three superior quality x-type subunits (1Ax2.1*, 1Ax1 and 1Ax2*) and four superior quality y-type subunits (1By16, 1By9, 1By8 and 1By18), respectively. Both subunits contained a large number of α -helices at the N- and C-termini. In particular, four cysteine residues in three x-type subunits (1Ax2.1*, 1Ax1 and 1Ax2*) and the cysteines in the intermediate repeat region were unable to form intramolecular disulfide bonds, but these cysteines might form intermolecular disulfide bonds with other glutenins and gliadins to enhance gluten macropolymer formation. The SNP-based molecular markers for 1Ax2.1* and 1By19* genes were developed according to the SNP variations in the promoter regions and verified in different F2 generation populations as well as different recombination inbred lines derived from crossing between spelt wheat and bread wheat cultivars. These molecular markers have shown a high reliability and have a good application prospect for improving bread making through marker-assisted selection.

Supplementary Materials: The supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms231911104/s1>.

Author Contributions: Y.C., J.Z. and R.W. performed all of the experiments, data analysis and wrote the paper. H.S. performed protein identification and data analysis. Y.Y. designed and supervised experiments. All authors have read and agreed to the published version of the manuscript.

Funding: This research was financially supported by a grant from the National Natural Science Foundation of China (31971931).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: This manuscript has no financial or non-financial competing interests.

Abbreviations

HMW-GS	High molecular weight glutenin subunit
LMW-GS	Low molecular weight glutenin subunit
MATE	Multidrug and toxin efflux
PCR	Polymerase chain reaction
AS-PCR	Allele specific polymerase chain reaction
SNP	Single nucleotide polymorphism
ORF	Open reading frame
InDels	Insertion/Deletion variations
CTAB	Hexadecyl trimethyl ammonium bromide
RILs	Recombinant inbred lines
MYA	Million years ago
SDS-PAGE	Sodium dodecyl sulfate polyacryl amide gel electrophoresis
MALDI-TOF-MS	Matrix-assisted laser desorption ionization tandem time of flight mass spectrometry
	Molecular weight

References

1. Kumar, R.; Mamrutha, H.M.; Kaur, A.; Venkatesh, K.; Sharma, D.; Singh, G.P. Optimization of Agrobacterium-mediated transformation in spring bread wheat using mature and immature embryos. *Mol. Biol. Rep.* **2019**, *46*, 1845–1853. [[CrossRef](#)] [[PubMed](#)]
2. Spannagl, M.; Martis, M.M.; Pfeifer, M.; Nussbaumer, T.; Mayer, K.F. Analysing complex Triticeae genomes—concepts and strategies. *Plant Methods* **2013**, *9*, 1–9. [[CrossRef](#)] [[PubMed](#)]
3. Kumar, A.; Kapoor, P.; Chunduri, V.; Sharma, S.; Garg, M. Potential of *Aegilops sp.* for improvement of grain processing and nutritional quality in wheat (*Triticum aestivum*). *Front. Plant Sci.* **2019**, *10*, 308. [[CrossRef](#)]
4. Biesiekierski, J.R. What is gluten? *J. Gastroenterol. Hepatol.* **2017**, *32*, 78–81. [[CrossRef](#)] [[PubMed](#)]
5. Wrigley, C.W. Giant proteins with flour power. *Nature* **1996**, *381*, 738–739. [[CrossRef](#)]

6. Mioduszeowski, Ł.; Cieplak, M. Viscoelastic properties of wheat gluten in a molecular dynamics study. *PLoS Comput. Biol.* **2021**, *17*, e1008840. [[CrossRef](#)] [[PubMed](#)]
7. Bonilla, J.C.; Erturk, M.Y.; Schaber, J.A.; Kokini, J.L. Distribution and function of LMW glutenins, HMW glutenins, and gliadins in wheat doughs analyzed with 'in situ' detection and quantitative imaging techniques. *J. Cereal Sci.* **2020**, *93*, 102931. [[CrossRef](#)]
8. Payne, P.I.; Holt, L.M.; Law, C.N. Structural and genetical studies on the high-molecular-weight subunits of wheat glutenin. *Theor. Appl. Genet.* **1981**, *60*, 229–236. [[CrossRef](#)]
9. Law, C.N.; Payne, P.I. Genetical aspects of breeding for improved grain protein content and type in wheat. *J. Cereal Sci.* **1983**, *1*, 79–93. [[CrossRef](#)]
10. Li, Y.; Fu, J.; Shen, Q.; Yang, D. High-molecular-weight glutenin subunits: Genetics, structures, and relation to end use qualities. *Int. J. Mol. Sci.* **2020**, *22*, 184. [[CrossRef](#)]
11. Branlard, G.; Dardevet, M. Diversity of grain protein and bread wheat quality: II. Correlation between high molecular weight subunits of glutenin and flour quality characteristics. *J. Cereal Sci.* **1985**, *3*, 345–354. [[CrossRef](#)]
12. Brönneke, V.; Zimmermann, G.; Killermann, B. Effect of high molecular weight glutenins and D-zone gliadins on bread-making quality in German wheat varieties. *Cereal Res. Commun.* **2000**, *28*, 187–194. [[CrossRef](#)]
13. Lama, S.; Kabir, M.R.; Akhond, M.A.Y. Biochemical and molecular characterization of Bangladeshi wheat varieties for bread-making quality. *Plant Tissue Cult. Biotechnol.* **2018**, *28*, 57–68. [[CrossRef](#)]
14. Pang, B.S.; Zhang, X.Y. Isolation and molecular characterization of high molecular weight glutenin subunit genes *1Bx13* and *1By16* from hexaploid wheat. *J. Integr. Plant Biol.* **2008**, *50*, 329–337. [[CrossRef](#)] [[PubMed](#)]
15. Shewry, P.R.; Halford, N.G.; Tatham, A.S. The high molecular weight subunits of wheat glutenin. *J. Cereal Sci.* **1992**, *15*, 105–120. [[CrossRef](#)]
16. Tatham, A.S.; Drake, A.F.; Shewry, P.R. Conformational studies of synthetic peptides corresponding to the repetitive region of the high molecular weight (HMW) glutenin subunits of wheat. *J. Cereal Sci.* **1990**, *11*, 189–200. [[CrossRef](#)]
17. Pézolet, M.; Bonenfant, S.; Dousseau, F.; Popineau, Y. Conformation of wheat gluten proteins Comparison between functional and solution states as determined by infrared spectroscopy. *FEBS Lett.* **1992**, *299*, 247–250. [[CrossRef](#)]
18. Lafiandra, D.; Shewry, P.R. Wheat Glutenin polymers 2, the role of wheat glutenin subunits in polymer formation and dough quality. *J. Cereal Sci.* **2022**, *106*, 103487. [[CrossRef](#)]
19. Huebner, F.R.; Bietz, J.A.; Wall, J.S. Disulfide bonds: Key to wheat protein functionality. *Adv. Exp. Med. Biol.* **1977**, *86A*, 67–88.
20. Tatham, A.S.; Mifflin, B.J.; Shewry, P.R. The beta-turn conformation in wheat gluten proteins: Relationship to gluten elasticity. *Cereal Chem.* **1985**, *62*, 405–442.
21. Gao, J.H.; Yu, P.X.; Liang, H.R.; Fu, J.H.; Luo, Z.Y.; Yang, D. The wPDI Redox Cycle coupled conformational change of the repetitive domain of the HMW-GS 1Dx5-A computational study. *Molecules* **2020**, *25*, 4393. [[CrossRef](#)] [[PubMed](#)]
22. Cazalis, R.; Aussenac, T.; Rhazi, L.; Marin, A.; Gibrat, J.F. Homology modeling and molecular dynamics simulations of the N-terminal domain of wheat high molecular weight glutenin subunit 10. *Protein Sci.* **2003**, *12*, 34–43. [[CrossRef](#)]
23. Robert, A.; Barkoutsos, P.K.; Woerner, S.; Tavernelli, I. Resource-efficient quantum algorithm for protein folding. *npj Quantum Inform.* **2021**, *7*, 1–5. [[CrossRef](#)]
24. Magoshi, J.; Becker, M.; Han, Z.; Nakamura, S. Thermal properties of seed proteins. *J. Therm. Anal. Calorim.* **2002**, *70*, 833–839. [[CrossRef](#)]
25. Calucci, L.; Pinzino, C.; Capocchi, A.; Galleschi, L.; Ghiringhelli, S.; Saviozzi, F.; Zandomeneghi, M. Structure and dynamics of high molecular weight glutenin subunits of durum wheat (*Triticum durum*) in water and alcohol solutions studied by electron paramagnetic resonance and circular dichroism spectroscopies. *J. Agric. Food Chem.* **2001**, *49*, 359–365. [[CrossRef](#)] [[PubMed](#)]
26. Porath, J. From gel filtration to adsorptive size exclusion: Communications from the 11th international conference on methods in protein structure analysis. *J. Protein. Chem.* **1997**, *16*, 463–468. [[CrossRef](#)]
27. Bryant, P.; Pozzati, G.; Elofsson, A. Improved prediction of protein-protein interactions using alphafold2. *Nat. Commun.* **2022**, *13*, 1265. [[CrossRef](#)]
28. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Hassabis, D. Highly accurate protein structure prediction with alphafold. *Nature* **2021**, *596*, 583–589. [[CrossRef](#)]
29. Zweckstetter, M. NMR hawk-eyed view of AlphaFold2 structures. *Protein Sci.* **2021**, *30*, 2333–2337. [[CrossRef](#)]
30. Duan, W.; Lu, F.; Cui, Y.; Zhang, J.; Du, X.; Hu, Y.; Yan, Y. Genome-wide identification and characterisation of wheat MATE genes reveals their roles in aluminium tolerance. *Int. J. Mol. Sci.* **2022**, *23*, 4418. [[CrossRef](#)]
31. Yan, Y.; Hsam, S.L.K.; Yu, J.Z.; Jiang, Y.; Ohtsuka, I.; Zeller, F.J. HMW and LMW glutenin alleles among putative tetraploid and hexaploid European spelt wheat (*Triticum spelta* L.) progenitors. *Theor. Appl. Genet.* **2003**, *107*, 1321–1330. [[CrossRef](#)] [[PubMed](#)]
32. An, X.; Li, Q.; Yan, Y.; Xiao, Y.; Hsam, S.L.K.; Zeller, F.J. Genetic diversity of European spelt wheat (*Triticum aestivum* ssp. *spelta* L. em. Thell.) revealed by glutenin subunit variations at the *Glu-1* and *Glu-3* loci. *Euphytica* **2005**, *146*, 193–201. [[CrossRef](#)]
33. Li, Q.Y.; Yan, Y.M.; Wang, A.L.; An, X.L.; Zhang, Y.Z.; Hsam, S.L.K.; Zeller, F.J. Detection of HMW glutenin subunit variations among 205 cultivated emmer accessions (*Triticum turgidum* ssp. *dicoccum*). *Plant Breed.* **2006**, *125*, 120–124. [[CrossRef](#)]
34. Yan, Y.; Hsam, S.L.K.; Yu, J.; Jiang, Y.; Zeller, F.J. Allelic variation of the HMW glutenin subunits in *Aegilops tauschii* accessions detected by sodium dodecyl sulphate (SDS-PAGE), acid polyacrylamide gel (A-PAGE) and capillary electrophoresis. *Euphytica* **2003**, *130*, 377–385. [[CrossRef](#)]

35. Zhang, Y.; Li, X.; Wang, A.; An, X.; Zhang, Q.; Pei, Y.; Gao, L.; Ma, W.; Appels, R.; Yan, Y. 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* **2008**, *178*, 23–33. [[CrossRef](#)]
36. Cui, D.; Wang, J.; Li, M.; Lu, Y.; Yan, Y. Functional assessment and SNP-based molecular marker development of two 1S1-encoded HMW glutenin subunits in *Aegilops longissima* L. *Mol. Breed.* **2019**, *39*, 1–15. [[CrossRef](#)]
37. Hu, J.; Wang, J.; Deng, X.; Yan, Y. Cloning and characterization of special HMW glutenin subunit genes from *Aegilops longissima* L. and their potential for wheat quality improvement. *3 Biotech.* **2019**, *9*, 1–13. [[CrossRef](#)] [[PubMed](#)]
38. Ma, C.; Yang, Y.; Li, X.; Ge, P.; Guo, G.; Subburaj, S.; Zeller, F.J.; Hsam, S.L.K.; Yan, Y. Molecular cloning and characterization of six novel HMW-GS genes from *Aegilops speltoides* and *Aegilops kotschyi*. *Plant Breed.* **2013**, *132*, 284–289. [[CrossRef](#)]
39. Pu, Z.E.; Liu, Y.X.; Li, W. Origin and genetics of spelt wheat (*Triticum spelta* L.) and its utilization in the genetic improvement of common wheat (*Triticum aestivum* L.). *J. Plant. Genet. Resour.* **2009**, *10*, 475–479.
40. Liu, M.; Zhao, Q.; Qi, F.; Stiller, J.; Tang, S.; Miao, J.; Liu, C. Sequence divergence between spelt and common wheat. *Theor. Appl. Genet.* **2018**, *131*, 1125–1132. [[CrossRef](#)]
41. Campbell, K.G. Spelt: Agronomy, genetics, and breeding. *Plant Breed. Rev.* **1997**, *15*, 187–213.
42. Frakolaki, G.; Giannou, V.; Topakas, E.; Tzia, C. Chemical characterization and breadmaking potential of spelt versus wheat flour. *J. Cereal Sci.* **2018**, *79*, 50–56.
43. Zeven, A.C.; Turkensteen, L.J.; Stubbs, R.W. Spelt (*Triticum spelta* L.) as a possible source of race-non-specific resistance to yellow rust (*Puccinia striiformis* West.). *Euphytica* **1968**, *17*, 381–384. [[CrossRef](#)]
44. Piergiovanni, A.R.; Rizzi, R.; Pannacciulli, E.; Gatta, C.D. Mineral composition in hulled wheat grains: A comparison between emmer (*Triticum dicoccon* Schrank) and spelt (*T. spelta* L.) accessions. *Int. J. Food Sci. Nutr.* **1997**, *48*, 381–386. [[CrossRef](#)]
45. Akel, W.; Thorwarth, P.; Mirdita, V.; Weissman, E.A.; Liu, G.; Würschum, T.; Longin, C.F.H. Can spelt wheat be used as heterotic group for hybrid wheat breeding? *Theor. Appl. Genet.* **2018**, *131*, 973–984. [[CrossRef](#)]
46. Halford, N.G.; Field, J.M.; Blair, H.; Urwin, P.; Moore, K.; Robert, L.; Thompson, R.; Flavell, R.B.; Shewry, P.R. Analysis of HMW glutenin subunits encoded by chromosome 1A of bread wheat (*Triticum aestivum* L.) indicates quantitative effects on grain quality. *Theor. Appl. Genet.* **1992**, *83*, 373–378. [[CrossRef](#)] [[PubMed](#)]
47. D'Ovidio, R.; Lafiandra, D.; Porceddu, E. Identification and molecular characterization of a large insertion within the repetitive domain of a high-molecular-weight glutenin subunit gene from hexaploid wheat. *Theor. Appl. Genet.* **1996**, *93*, 1048–1053. [[CrossRef](#)]
48. Yan, Y.M.; Jiang, Y.; An, X.L.; Pei, Y.H.; Li, X.H.; Zhang, Y.Z.; Wang, A.L.; He, Z.H.; Xia, X.C.; Bekes, F.; et al. Cloning, expression and functional analysis of HMW glutenin subunit *1By8* gene from Italy pasta wheat (*Triticum turgidum* L. ssp. *durum*). *J. Cereal Sci.* **2009**, *50*, 398–406. [[CrossRef](#)]
49. Liang, X.; Zhen, S.; Han, C.; Chang, W.; Yan, Y. Molecular characterization and marker development for hexaploid wheat-specific HMW glutenin subunit *1By18* gene. *Mol. Breed.* **2015**, *35*, 221.
50. Fevzioglu, M.; Ozturk, O.K.; Hamaker, B.R.; Campanella, O.H. Quantitative approach to study secondary structure of proteins by FT-IR spectroscopy, using a model wheat gluten system. *Int. J. Biol. Macromol.* **2020**, *164*, 2753–2760. [[CrossRef](#)]
51. Ruomei, W.; Junwei, Z.; Fei, L.; Nannan, L.; Prodanović, S.; Yueming, Y. Cloning and molecular characterization of two novel LMW-m type glutenin genes from *Triticum spelta* L. *Genetika* **2021**, *53*, 141–155.
52. Wang, Y.; Zhen, S.; Luo, N.; Han, C.; Lu, X.; Li, X.; Xia, X.; He, Z.; Yan, Y. Low molecular weight glutenin subunit gene *Glu-B3h* confers superior dough strength and breadmaking quality in wheat (*Triticum aestivum* L.). *Sci. Rep.* **2016**, *6*, 27182. [[CrossRef](#)] [[PubMed](#)]
53. Skolnick, J.; Gao, M.; Zhou, H.; Singh, S. AlphaFold 2: Why it works and its implications for understanding the relationships of protein sequence, structure, and function. *J. Chem. Inf. Model.* **2021**, *61*, 4827–4831. [[CrossRef](#)] [[PubMed](#)]
54. Lutz, E.; Wieser, H.; Koehler, P. Identification of disulfide bonds in wheat gluten proteins by means of mass spectrometry/electron transfer dissociation. *J. Agric. Food Chem.* **2012**, *60*, 3708–3716. [[CrossRef](#)] [[PubMed](#)]
55. Shewry, P.; Gilbert, S.; Savage, A.; Tatham, A.S.; Wan, Y.F.; Belton, P.S.; Wellner, N.; D'Ovidio, R.; Békés, F.; Halford, N.G. Sequence and properties of HMW subunit IBx20 from pasta wheat (*Triticum durum*) which is associated with poor end use properties. *Theor. Appl. Genet.* **2003**, *106*, 744–750. [[CrossRef](#)] [[PubMed](#)]
56. Pirozi, M.R.; Margiotta, B.; Lafiandra, D.; MacRitchie, F. Composition of polymeric proteins and bread-making quality of wheat lines with allelic HMW-GS differing in number of cysteines. *J. Cereal Sci.* **2008**, *48*, 117–122. [[CrossRef](#)]
57. Shewry, P.R.; Halford, N.G. Cereal seed storage proteins: Structures, properties and role in grain utilization. *J. Exp. Bot.* **2002**, *53*, 947–958. [[CrossRef](#)]
58. Wang, N.; Ma, S.; Li, L.; Zheng, X. Aggregation characteristics of protein during wheat flour maturation. *J. Agric. Food Chem.* **2019**, *99*, 719–725. [[CrossRef](#)]
59. Andrews, J.L.; Skerritt, J.H. Quality-related epitopes of high Mr subunits of wheat glutenin. *J. Cereal Sci.* **1994**, *19*, 219–229. [[CrossRef](#)]
60. Ramesh, P.; Mallikarjuna, G.; Sameena, S.; Kumar, A.; Gurulakshmi, K.; Reddy, B.V.; Sekhar, A.C. Advancements in molecular marker technologies and their applications in diversity studies. *J. Biosci.* **2020**, *45*, 1–15. [[CrossRef](#)]
61. Ravel, C.; Faye, A.; Ben-Sadoun, S.; Ranoux, M.; Gérard Branlard. SNP markers for early identification of high molecular weight glutenin subunits (HMW-GSs) in bread wheat. *Theor. Appl. Genet.* **2020**, *133*, 751–770. [[CrossRef](#)] [[PubMed](#)]

62. Schwarz, G.; Sift, A.; Wenzel, G.; Mohler, V. DHPLC scoring of a SNP between promoter sequences of HMW glutenin x-type alleles at the Glu-D1 locus in wheat. *J. Agric. Food Chem.* **2003**, *51*, 4263–4267. [[CrossRef](#)] [[PubMed](#)]
63. 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.* **1996**, *24*, 73–78. [[CrossRef](#)]
64. 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* **1999**, *42*, 296–307. [[CrossRef](#)] [[PubMed](#)]
65. Buchan, D.W.; Jones, D.T. The PSIPRED protein analysis workbench: 20 years on. *Nucleic Acids Res.* **2019**, *47*, W402–W407. [[CrossRef](#)]
66. Buchan, D.W.; Ward, S.M.; Lobley, A.E.; Nugent, T.C.O.; Bryson, K.; Jones, D.T. Protein annotation and modelling servers at University College London. *Nucleic Acids Res.* **2010**, *38*, W563–W568. [[CrossRef](#)]
67. Mirdita, M.; Schütze, K.; Moriwaki, Y.; Heo, L.; Ovchinnikov, S.; Steinegger, M. ColabFold: Making protein folding accessible to all. *Nat. Methods* **2022**, *19*, 679–682. [[CrossRef](#)]
68. Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Velankar, S. AlphaFold Protein Structure Database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **2022**, *50*, D439–D444. [[CrossRef](#)]