Functional Conservation and Divergence among Homoeologs of *TaSPL20* and *TaSPL21*, Two SBP-Box Genes Governing Yield-Related Traits in Hexaploid Wheat[1[OPEN]]

Bin Zhang[2], Weina Xu[2], Xia Liu[2], Xinguo Mao, Ang Li, Jingyi Wang, Xiaoping Chang, Xueyong Zhang, and Ruilian Jing*

National Key Facility for Crop Gene Resources and Genetic Improvement/Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing 100081, China (B.Z., W.X., X.L., A.L., J.W., X.C., X.Z., R.J.); and Institute of Crop Germplasm Resources, Shanxi Academy of Agricultural Sciences, Key Laboratory of Crop Gene Resources and Germplasm Enhancement on Loess Plateau, Ministry of Agriculture, Taiyuan 030031, Shanxi, China (X.L.)

ORCID IDs: 0000-0002-7347-0367 (B.Z.); 0000-0002-0798-0000 (W.X.); 0000-0001-9824-916X (X.L.); 0000-0001-8600-2098 (R.J.).

Maintaining high and stable yields has become an increasing challenge in wheat breeding due to climate change. Although Squamosa-promoter binding protein (SBP)-box genes have important roles in plant development, very little is known about the actual biological functions of wheat SBP-box family members. Here, we dissect the functional conservation, divergence, and exploitation of homoeologs of two paralogous *TaSPL* wheat loci during domestication and breeding. *TaSPL20* and *TaSPL21* were highly expressed in the lemma and palea. Ectopic expressions of *TaSPL20/21* in rice exhibited similar functions in terms of promoting panicle branching but had different functions during seed development. We characterized all six *TaSPL20/21* genes located across the three homoeologous (A, B, and D) genomes. According to the functional analysis of naturally occurring variants in 20 environments, four favorable haplotypes were identified. Together, they reduced plant height by up to 27.5%, and *TaSPL21-6D-Hap*II increased 1000-grain weight by 9.73%. Our study suggests that *TaSPL20* and *TaSPL21* homoeologs underwent diversification in function with each evolving its own distinctive characteristics. During domestication and breeding of wheat in China, favorable haplotypes of each set were selected and exploited to varying degrees due to their large effects on plant height and 1000-grain weight.

Wheat (*Triticum aestivum*) is a global food crop providing calories and protein for 30% of the human population (Tanno and Willcox, 2006; Mayer et al., 2014). Due to climate change over recent decades, potential global wheat production was reduced by 5.5% relative to what would have occurred in its absence (Lobell et al., 2011). Thus, high and stable yield is always a major objective in wheat breeding.

Squamosa-promoter binding protein (SBP)-box genes play important roles in plant phase transition and flowering (Schwarz et al., 2008; Wang et al., 2009, 2014), leaf development (Shikata et al., 2009; Usami et al., 2009), plant architecture (Wang et al., 2005; Chuck et al., 2010; Jiao et al., 2010; Miura et al., 2010), fruit development (Manning et al., 2006), organ size (Wang et al., 2008), gibberellin signaling (Zhang et al., 2007), grain quality, and yield (Wang et al., 2012; Si et al., 2016). *Squamosa-promoter binding protein-like* (*SPL*) genes, first identified in *Antirrhinum majus*, were named *SBP1* and *SBP2* based on their ability to bind the promoter of the floral meristem identity gene *SQUAMOSA* (Klein et al., 1996). Since then, a large number of *SPL* family members have been functionally characterized in many plant species, including *Chlamydomonas reinhardtii* (Kropat et al., 2005), *Physcomitrella patens* (Riese et al., 2008), Arabidopsis (*Arabidopsis thaliana*; Cardon et al., 1997), maize (*Zea mays*; Wang et al., 2005), rice (*Oryza sativa*; Jiao et al., 2010), tomato (*Solanum lycopersicum*; Manning et al., 2006), *Populus x canadensis* (Wang et al., 2011), and *Betula pendula* (Lännenpää et al., 2004). In rice, *OsSPL13* positively regulates cell size in the grain hull resulting in enhanced grain length and yield (Si et al., 2016); *OsSPL14* regulates plant architecture by controlling shoot branching during vegetative growth and panicle development (Jiao et al., 2010; Miura et al., 2010); and

[1] The work was supported by The National Natural Science Foundation of China (31461143024) and The National Key Research and Development Program of China (2016YFD0100605).

[2] These authors contributed equally to the article.

* Address correspondence to jingruilian@caas.cn.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Ruilian Jing (jingruilian@caas.cn).

B.Z. and R.J. conceived and designed the experiments; B.Z., W.X., X.L., X.M., A.L., J.W., X.C., and X.Z. performed the experiments; B.Z, W.X., and X.L. analyzed the data; B.Z. and R.J. contributed to the writing of the manuscript; B.Z and R.J. revised the manuscript.

[OPEN] Articles can be viewed without a subscription.
www.plantphysiol.org/cgi/doi/10.1104/pp.17.00113

overexpression of *OsSPL16* promoted cell proliferation and grain filling, resulting in increased grain width and yield (Wang et al., 2012). However, apart from bioinformatics and gene expression information (Zhang et al., 2014a; Wang et al., 2015), very little is known about the biological functions of SBP-box family members in common wheat. Ten wheat *TaSPL* genes have been cloned and classified into five groups (G1–G5; Zhang et al., 2014a). Among them, *TaSPL20* and *TaSPL21* (hereafter referred to as *TaSPL20/21*) in G1 do not cluster with *OsSPL13/14/16* and are abundantly expressed in the shoot apical meristem and spike (Zhang et al., 2014a).

Wheat (2n = 6x = 42; genomic code AABBDD) underwent two separate allopolyploidization events (Marcussen et al., 2014; Mayer et al., 2014); the first combined the genomes of two wild diploid species, *Triticum urartu* (2n = 2x = 14; AA) and probably *Aegilops speltoides* (2n = 2x = 14; BB), resulting in the allotetraploid *Triticum dicoccoides* (2n = 4x = 28; AABB). Wild emmer was domesticated to cultivated emmer (*Triticum dicoccum*), which subsequently hybridized with the diploid grass species *Aegilops tauschii* (2n = 2x = 14; DD) to form modern hexaploid wheat. Thus, in addition to paralogous genes (paralogs) that arise by duplication and characterize diploid species, the hexaploid wheat genome contains triplicated orthologous genes (homoeologs) from the ancestral diploid species. These genes either retained their original functions or underwent functional divergence, including silencing, through mutation (Wendel, 2000; Pfeifer et al., 2014). Up to 46% of homoeologs in hexaploid wheat have been silenced, neofunctionalized, or subfunctionalized over 1.5 million years of evolution (Pont et al., 2011).

Here, we address an interesting question about the biological functions and fates of homoeologs of paralogous genes *TaSPL20* and *TaSPL21* in modern hexaploid wheat genotypes. In addition, where and how different haplotypes of the paralogous series *TaSPL20* and *TaSPL21* underwent variation during domestication and breeding were also investigated in this study. We found both *TaSPL20* and *TaSPL21* were highly expressed in lemmas and paleas during early wheat spike developmental stages. Ectopic expressions of *TaSPL20* and *TaSPL21* in rice exhibited similar functions in terms of promoting panicle branching but had different functions during seed development. We characterized the three homoeologs of each gene, located on group 6 and group 7 chromosomes, respectively. Functional analyses of naturally occurring variants at each locus suggested that the evolutionary histories of the two gene sets were quite different. During domestication and breeding of wheat, favorable haplotypes of each set were selected and exploited due to their large effects on plant height (PH) and 1000-grain weight (TGW).

RESULTS

TaSPL20 and *TaSPL21* Are Highly Expressed in the Lemma and Palea during Early Spike Development

TaSPL20/21 are abundantly expressed in shoot apical meristems and young spikes, and show tissue-specific expression patterns associated with spike development (Zhang et al., 2014a). To monitor their detailed expression patterns based on conserved regions among each set of homoeologs, we sampled stamens, pistils, lemmas, paleas, and pedicels from wheat spikes with lengths of 3, 4, 6, 8, and 10 cm, and at the flowering stage. Both *TaSPL20* and *TaSPL21* members displayed specific expression patterns in lemmas and paleas at early spike development (3–6 cm spikes), suggesting that both groups may regulate wheat spike development (Fig. 1A).

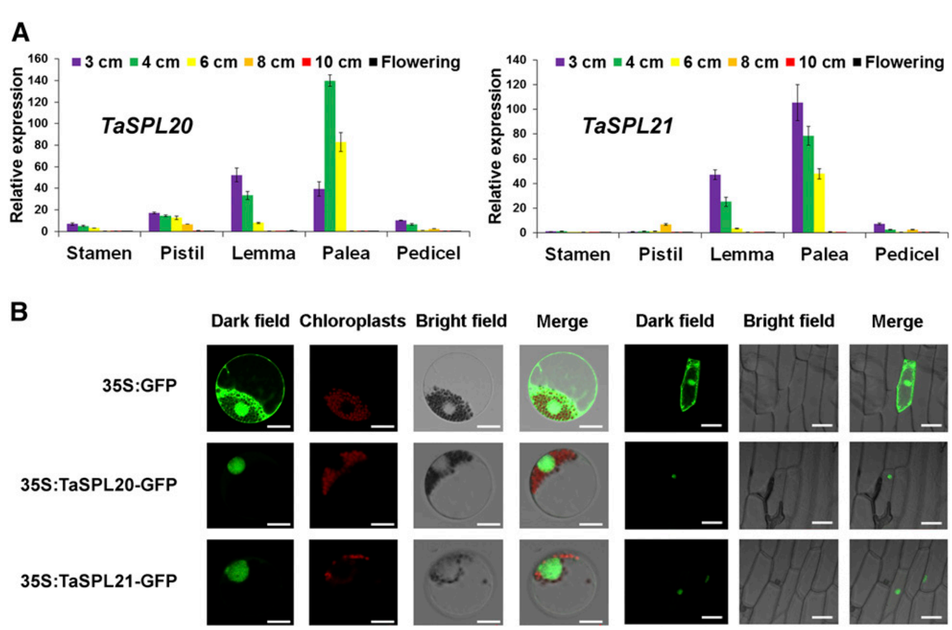

Figure 1. Expression patterns and subcellular localizations of *TaSPL20* and *TaSPL21* in wheat. A, Expression patterns of *TaSPL20* and *TaSPL21* in wheat. The 3, 4, 6, 8, and 10 cm and flowering indicate the stamens, pistils, lemmas, paleas, pedicels collected from the wheat spikes with lengths of 3, 4, 6, 8, and 10 cm and at the flowering stage. The *Actin* gene was used as an internal control. Error bars denote ± SE. B, Subcellular localizations of *TaSPL20* and *TaSPL21* in wheat. The vector control (35S:GFP) and fusion proteins (35S:TaSPL20-GFP and 35S:TaSPL21-GFP) were each introduced into wheat protoplasts (left) and onion epidermal cells (right). GFP was observed with a laser scanning confocal microscope. Scale bars = 20 μm for wheat protoplasts and 100 μm for onion epidermal cells.

SBP-box genes encode a plant-specific family of transcription factors and contain bipartite nuclear localization signals. Based on coding regions of the D genome members, we determined the subcellular localizations of TaSPL20 and TaSPL21 in wheat protoplasts and living onion epidermal cells by transient expression. GFP signals were observed in entire cells of the controls, whereas TaSPL20-GFP and TaSPL21-GFP fusion proteins were exclusively localized in the nuclei, indicating that TaSPL20 and TaSPL21 function as transcription factors (Fig. 1B).

Ectopic Expression of *TaSPL20* and *TaSPL21* in Rice Promotes Panicle Branching and Influences Seed Development

We generated transgenic rice lines containing *TaSPL20/21* coding regions from the D genome under control of ubiquitin promoter in order to assess *TaSPL20* and *TaSPL21* functions. Under field conditions, five and three representative homozygous lines overexpressing *TaSPL20* (*TaSPL20-OE*) and *TaSPL21* (*TaSPL21-OE*) were obtained for detailed analysis (Figs. 2 and 3; Supplemental Fig. S1). Compared to wild type, lines possessing *TaSPL20-OE* (L1–L5) and *TaSPL21-OE* (L1–L3; Fig. 2A) produced more primary branches (21–34%; Fig. 2C), more secondary branches (19–23%; Fig. 2D), higher grain numbers (24–27%; Fig. 2E), and longer panicles (8%; Fig. 2F). Since two of eight transgenic lines (*TaSPL20-OE* line 5 and *TaSPL21-OE* line 1) showed slightly increased tiller numbers (Fig. 2G), *TaSPL20* and *TaSPL21* maybe weakly regulated tiller numbers during vegetative stage besides their functions in panicle branching during reproductive development.

The potential functions of *TaSPL20* and *TaSPL21* during seed development were also investigated. *TaSPL20-OE* lines produced larger seeds (Fig. 3A) as measured by increased TGW (Fig. 3D), seed surface area (Fig. 3E), grain length (Fig. 3, B and F), and grain width (Fig. 3, C and G) compared to wild type. *TaSPL21-OE* lines displayed different seed phenotypes; seed size (Fig. 3, A and E) was unchanged and TGW (Fig. 3D) was significantly reduced.

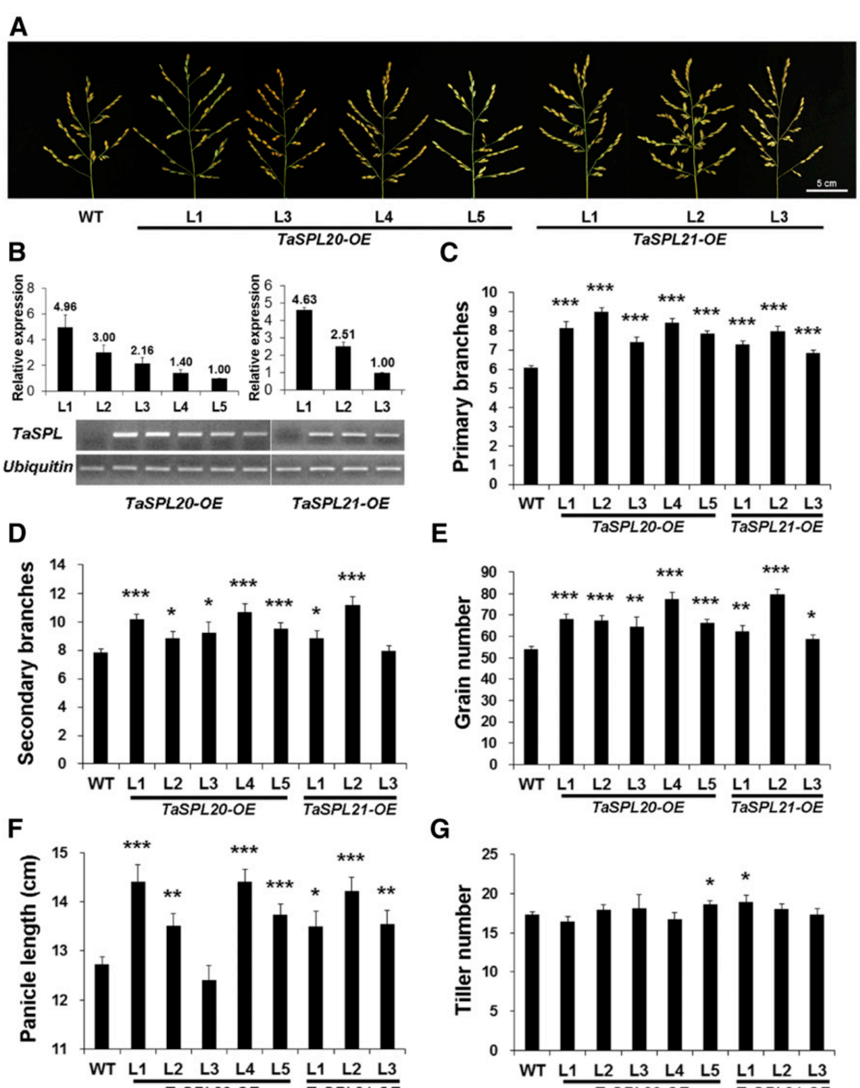

Figure 2. Analysis of panicle morphologies of *TaSPL20-OE* and *TaSPL21-OE* lines. A, Comparison of panicle morphologies; B, relative expression levels of *TaSPL20* and *TaSPL21* in transgenic rice lines; C, primary panicle branching; D, secondary panicle branching; E, grain number per main panicle; F, panicle lengths; and G, tiller numbers were measured. WT, Wild type; L1, L2, L3, L4, and L5 refer to transgenic lines. The error bars denote ± SE; $*P < 0.05$, $**P < 0.01$, and $***P < 0.001$.

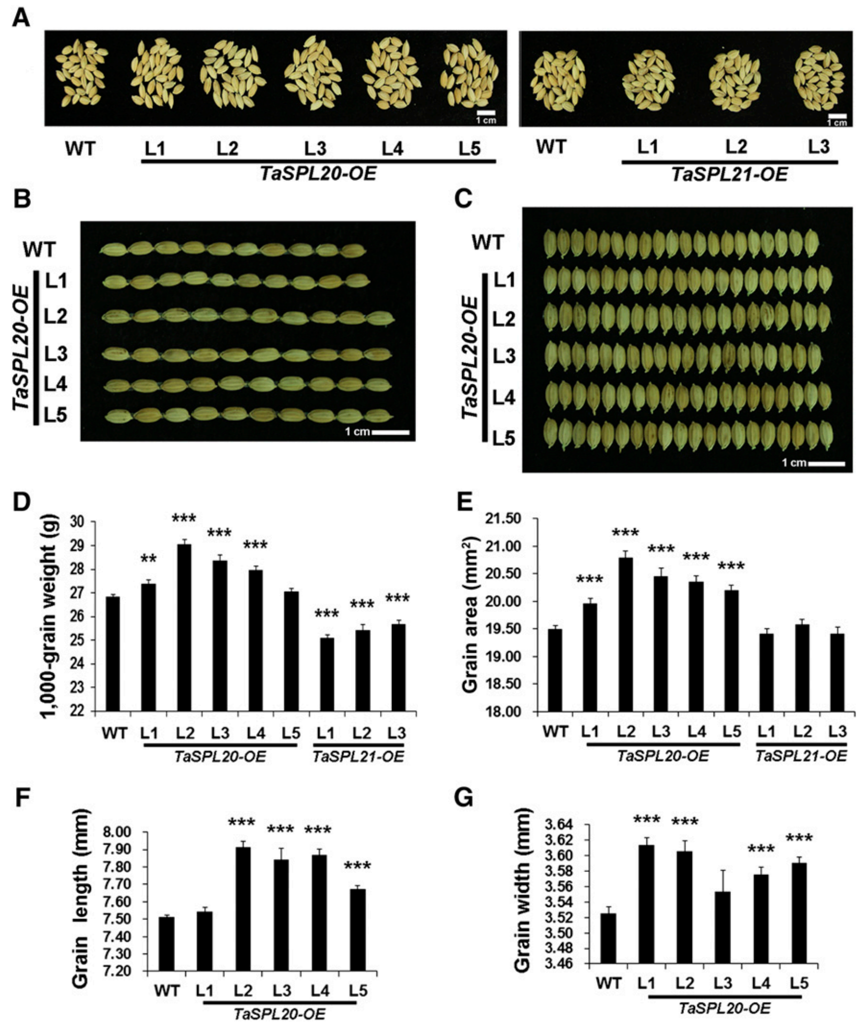

Figure 3. Analysis of grain size in five *TaSPL20-OE* and three *TaSPL21-OE* rice lines. A to C, Appearance of the grains in wild-type and transgenic rice plants. Comparisons of TGW (D), seed area (E), grain length (F), and grain width (G). WT, Wild type; L1, L2, L3, L4, and L5 are transgenic lines. Error bars denote ± SE; $^{*}P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$.

The results thus indicated that both *TaSPL20* and *TaSPL21* affected panicle branching during reproductive development in a similar way, but their effects on seed development were different.

Chromosome Mapping, Naturally Occurring Variants, and Development of Functional Markers for *TaSPL20* and *TaSPL21* Homoeologs

Using the D genome coding regions of *TaSPL20* and *TaSPL21* as queries, the 5′ and 3′ flanking regions of the respective homoeologs were identified by BLASTN searches against the draft genome databases of the A (Ling et al., 2013) and D (Jia et al., 2013) genome progenitors and wheat variety Chinese Spring (Wilkinson et al., 2012). Genome-specific primer pairs for *TaSPL20* (*TaSPL20-7A*, *-7B*, and *-7D*) and *TaSPL21* (*TaSPL21-6A*, *-6B*, and *-6D*) homoeologs were designed based on polymorphisms in the genome sequences. The genomic fragments, including 5′ and 3′ flanking regions, were 2827, 2720 and 4740 bp, and 3463, 2089 and 4322 bp, respectively (Fig. 4B). The chromosomal locations of *TaSPL20* and *TaSPL21* homoeologs were determined using Chinese Spring nullisomic-tetrasomic lines, a series of lines with each missing one pair of chromosomes that is replaced by an extra pair of homoeologous chromosomes. *TaSPL20* homoeologs were located on group 7 chromosomes (7A, 7B, and 7D; Fig. 4A), and *TaSPL21* homoeologs were located on group 6 chromosomes (6A, 6B, and 6D; Fig. 4A).

We analyzed the entire genomic fragment of each gene in 37 cultivars showing wide variations in tiller number and spike-related traits to detect sequence variations (Supplemental Table S1; Zhang et al., 2015b). Three haplotypes of *TaSPL20-7A* were characterized by eight variations, the first three of which preceded the ATG start codon. A cleaved amplified polymorphic sequence (CAPS) marker was developed based on InDel1 (insertion and deletion, InDel) and SNP5 to identify the three haplotypes simultaneously (Fig. 4C). As no nucleotide variation was detected in the genomic region of *TaSPL20-7B*, this locus was excluded from further research. Two and three SNPs occurred in the 5′ flanking regions of *TaSPL20-7D* (Fig. 4D) and *TaSPL21-6A* (Fig. 4E), respectively. A derived cleaved amplified polymorphic sequence (dCAPS) marker and a CAPS marker were

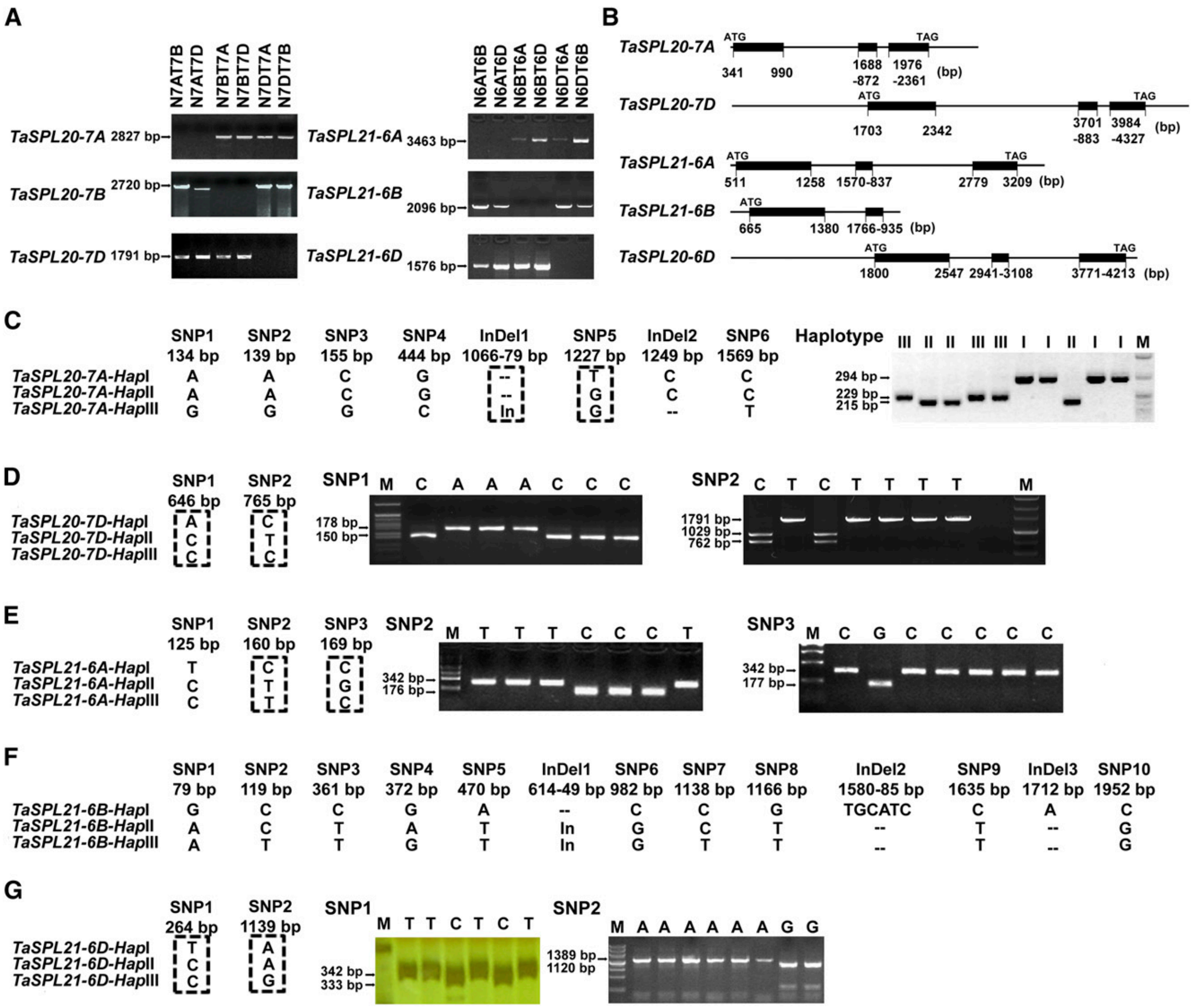

Figure 4. Chromosome location, gene structure, haplotypes, and functional markers of *TaSPL20-7A*, *TaSPL20-7D*, *TaSPL21-6A*, *TaSPL21-6B*, and *TaSPL21-6D*. A, *TaSPL20* homoeologs were located on group 7 chromosomes (left), and *TaSPL21* homoeologs were located on group 6 chromosomes (right) using Chinese Spring nullisomic-tetrasomic lines. B, Schematic diagram of *TaSPL20* and *TaSPL21* homoeologs. Gene structures of *TaSPL20-7D*, *TaSPL21-6A*, and *TaSPL21-6D* were determined by aligning genomic sequences and corresponding coding sequences; *TaSPL20-7A* and *TaSPL21-6B* were predicted by FGENESH (Softberry, http://www.softberry.com). Exons are indicated by black boxes; flanking regions and introns are indicated by solid lines. C, Three haplotypes were identified for *TaSPL20-7A*. A CAPS marker was developed based on InDel and SNP5 with restriction endonuclease *Nco*I, which cleaved the sequence only when the SNP5 site was G. After enzyme digestion, the main PCR products of *Hap*I, *Hap*II, and *Hap*III were 294, 215, and 229 bp, respectively. D, Three haplotypes were identified for *TaSPL20-7D*. dCAPS marker was developed based on SNP1, which cleaved the sequence only when the SNP1 site was C. A CAPS marker was developed based on SNP2 with restriction endonuclease *Bam*HI, which cleaved the sequence only when the SNP2 site was C. E, Three haplotypes were identified for *TaSPL21-6A*. One CAPS marker was developed based on SNP2 using restriction endonuclease *Bgl*I, which cleaved the sequence only when the SNP2 site was C. The other CAPS marker was developed based on SNP3 with restriction endonuclease *Bse*DI/*Sec*I, which cleaved the sequence only when the SNP3 site was G. F, Three haplotypes were identified for *TaSPL21-6B*. Haplotypes were detected by direct sequencing. G, Three haplotypes were identified in *TaSPL21-6D*. One CAPS marker was developed based on SNP1 using restriction endonuclease *Nco*I, which cleaved the sequence only when the SNP1 site was C. The second CAPS marker was based on SNP2 using restriction endonuclease *Ava*II, which cleaved the sequence only when the SNP2 site was G. M, DNA marker. The sizes of PCR products are shown on the left.

developed for *TaSPL20-7D* based on SNP1 and SNP2 (Fig. 4D), and two CAPS markers were developed for *TaSPL21-6A* based on SNP2 and SNP3 (Fig. 4E). Three haplotypes in *TaSPL21-6B* formed by 10 SNPs and three InDel (six preceded the ATG start codon) were detected by direct sequencing (Fig. 4F). Two SNPs in the 5′ flanking regions of *TaSPL21-6D* formed three haplotypes (Fig. 4G). Two CAPS markers were developed based on SNP1 and SNP2

(Fig. 4G). Thus, based on functional markers or direct sequencing, we distinguished haplotypes of all *TaSPL20* and *TaSPL21* homoeologs.

Homoeolog-Specific Functions of *TaSPL20* and *TaSPL21* Affecting Yield-Related Traits

In order to investigate the effects of different *TaSPL20/21* haplotypes, we performed an association analysis of each haplotype and six yield-related traits (PH; number of spikes per plant, NSP; spike length, SL; total number of spikelets per spike, TNSS; number of grains per spike, NGS; and TGW) using 262 accessions (population 1; Supplemental Table S2). We used a general linear model to account for population structure (Q). The accessions in population 1 represented cultivars mainly from two major Chinese wheat production areas, the northern Winter Wheat and Yellow and Huai River Valleys Facultative Wheat Zones.

In 20 environments (4 years × 2 sites × well-watered, drought, and heat stress conditions), there was no significant association between *TaSPL20-7A* and any of the six traits (Fig. 5A; Supplemental Table S3). *TaSPL20-7D* was significantly associated with PH in all environments except E3 and E8 and weakly associated with TGW (12 environments; Fig. 5B; Supplemental Table S4); *TaSPL21-6A* was strongly associated with PH in all environments and weakly associated with SL (10 environments; Fig. 5C; Supplemental Table S5); *TaSPL21-6B* showed weak associations with PH (12 environments), TNSS (11 environments), NGS (10 environments; Fig. 5D; Supplemental Table S6); and *TaSPL21-6D* was strongly associated with PH and TGW in 19 environments (Fig. 5E; Supplemental Table S7). For PH, the phenotypic variation explained by *TaSPL20-7D*,

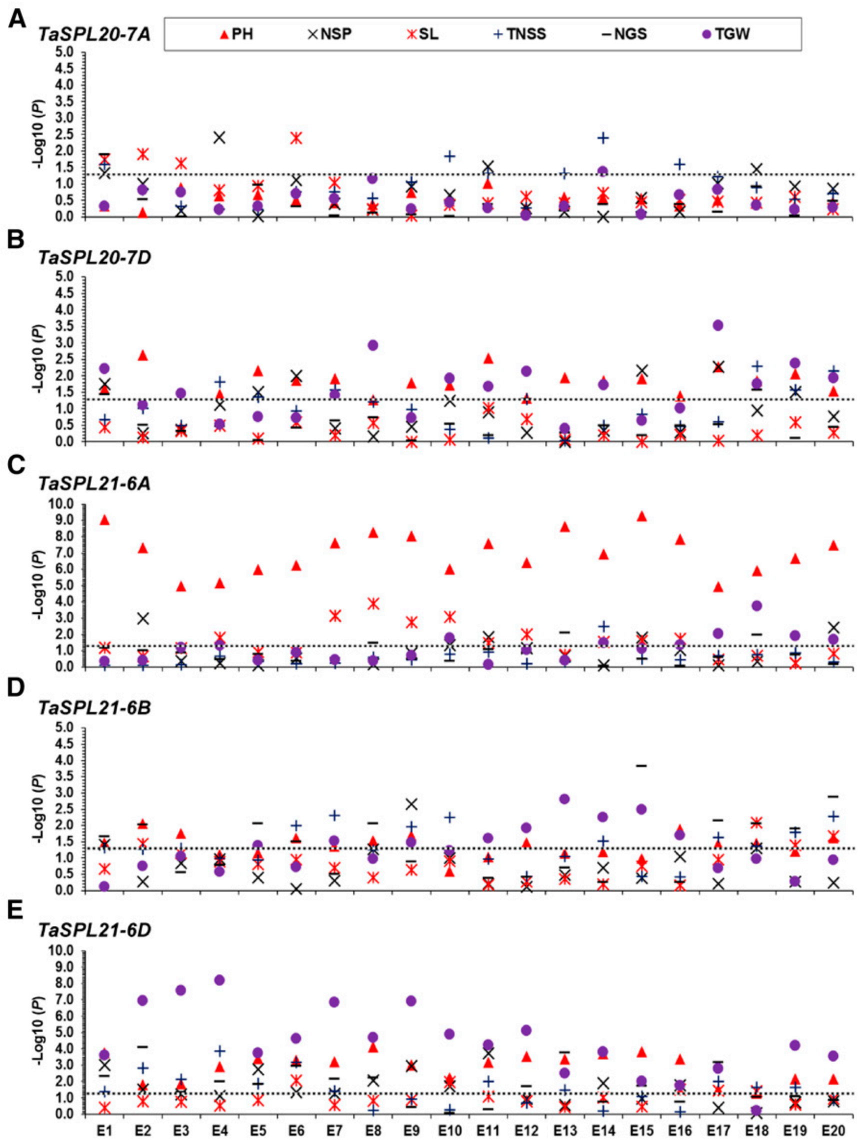

Figure 5. Haplotype analysis of *TaSPL20* and *TaSPL21* homoeologs using agronomic trait data from 20 environments. NSP, Number of spikes per plant; SL, spike length; TNSS, total number of spikelets per spike; NGS, number of grains per spike. E1 to E20 indicate the environments at Changping in 2009 under DS and WW conditions, Shunyi in 2009 under DS, WW, DS + HS, and WW + HS, Changping in 2010 under DS and WW, Shunyi in 2010 under DS, WW, DS + HS, and WW + HS, Shunyi in 2011 under DS, WW, DS + HS, and WW + HS, Changping in 2012 under DS and WW, and Shunyi in 2012 under DS and WW, respectively. Negative $\log_{10}$-transformed *P* values are plotted. Black horizontal dotted line indicates the threshold value for significant associations ($P < 0.05$).

TaSPL21-6A, and *TaSPL21-6D* haplotypes ranged from 2.48 to 4.64%, 8.68 to 15.47%, and 2.68 to 7.26%, respectively. The phenotypic variation for TGW explained by *TaSPL21-6D* haplotypes ranged from 3.05% to 13.95%.

Naturally Occurring Variants of *TaSPL20/21* Homoeologs Related to PH and TGW and Functional Buffering Effects

Since *TaSPL20-7D*, *TaSPL21-6A*, and *TaSPL21-6D* were strongly associated with PH, and *TaSPL21-6D* was strongly associated with TGW, we compared PH or TGW among accessions with favorable haplotypes and those with other alleles across all 20 environments (Fig. 6, B–E). *TaSPL20-7D-Hap*I, *TaSPL21-6A-Hap*II, and *-6D-Hap*III exhibited the lowest PH among their corresponding haplotype sets, whereas *TaSPL20-7D-Hap*II, *TaSPL21-6A-Hap*I, and *TaSPL21-6D-Hap*I were associated with tallest PH (Fig. 6, B–D). The phenotypic differences between contrasting groups were 9.67 to 14.71 cm for *TaSPL20-7D*, 13.73 to 23.21 cm for *TaSPL21-6A*, and 12.33 to 27.99 cm for *TaSPL21-6D*. Thus, *TaSPL20-7D-Hap*I, *TaSPL21-6A-Hap*II, and *TaSPL21-6D-Hap*III reduced PH by about 13.00, 18.96, and 21.68% compared to the contrasting tallest haplotype. *TaSPL21-6D-Hap*II had significantly higher TGW than *TaSPL21-6D-Hap*I, and the phenotypic differences between them ranged from 1.83 g to 4.96 g, or about 9.73% compared to the TGW for *TaSPL21-6D-Hap*I (Fig. 6E). Thus, *TaSPL20-7D-Hap*I and *TaSPL21-6A-Hap*II and *-6D-Hap*III represent favorable haplotypes able to reduce PH, and *TaSPL21-6D-Hap*II and *-6D-Hap*III (*Hap*II&III) increase TGW. We also evaluated the buffering effects on PH by *TaSPL20/21* homoeologs by combinations of favorable haplotypes (Fig. 6A). Combination of *TaSPL20-7D-Hap*I with *TaSPL21-6A-Hap*II and *-6D-Hap*II&III reduced PH by up to 27.5% from 98.56 cm (those without any favorable haplotype) to 71.48 cm (those with triple favorable haplotypes).

*TaSPL20-7D-Hap*I, *TaSPL21-6A-Hap*II, and *TaSPL21-6D-Hap*II&III Were Positively Selected in Wheat Breeding

Artificial selection leaves strong footprints in genomes, manifesting progressive accumulation of favorable haplotypes (Barrero et al., 2011). To determine whether favorable haplotypes of *TaSPL20/21* homoeologs were selected during wheat breeding, we assessed frequency changes of *TaSPL20-7D-Hap*I, *TaSPL21-6A-Hap*II, *TaSPL21-6D-Hap*II, and *Hap*III haplotypes over time in population 1 and two other populations. Population 2 consisting of 157 landraces (Supplemental Table S8) came mainly from the Chinese wheat mini-core collection, representing more than 70% of the genetic diversity of the full Chinese germplasm collection, and population 3 (Supplemental Table S9) was from the Chinese wheat core collection (Hao et al., 2008, 2011).

Based on phenotypic data from population 1 in 20 environments, PH declined from landraces to modern cultivars and continually fell in modern cultivars from pre-1960 to the 2000s. TGW showed the opposite trend (Fig. 7G). For PH, the relative values among haplotypes at each locus were *TaSPL20-7D-Hap*I < *Hap*III < *Hap*II (Fig. 6B), *TaSPL21-6A-Hap*II < *Hap*III < *Hap*I (Fig. 6C), *TaSPL21-6D-Hap*III&II < *Hap*I (Fig. 6D); and for TGW, the relative values were *TaSPL21-6D-Hap*III&II > *Hap*I (Fig. 6E). Frequencies of favorable haplotypes in populations 1, 2, and 3 for *TaSPL20-7D-Hap*I (Fig. 7, A and B), *TaSPL21-6A-Hap*II (Fig. 7, C and D), and *TaSPL21-6D-Hap*II&III (Fig. 7, E and F) gradually increased from 0%/1.31%, 0%/28.57%, 20.00%/11.46% in landraces to 19.15%/20.59%, 24.49%/52.94%, 78.72%/82.35% in the 2000s/1990s (the numbers left of the slash refer to frequencies in population 1; those right of a slash refer to frequencies in populations 2 and 3). The frequencies of *TaSPL20-7D-Hap*II, *TaSPL21-6A-Hap*I, and *TaSPL21-6D-Hap*I decreased in a corresponding manner. For the *TaSPL21-6A* (Fig. 7, C and D) and *TaSPL21-6D* (Fig. 7, E and F) loci, there were sharp increases/decreases in frequencies from landraces to the 1960s, suggesting that selection occurred at the very beginning of modern wheat breeding programs. In particular, the high frequency of *TaSPL21-6D-Hap*II in modern cultivars released after the 1960s suggests that this haplotype had been widely selected in wheat breeding (Fig. 7, E and F). *TaSPL21-6D-Hap*III, another important haplotype along with *TaSPL20-7D-Hap*I and *TaSPL21-6A-Hap*II, could have potential for wheat improvement.

Geographic Distribution of *TaSPL20-7D*, *TaSPL21-6A*, and *TaSPL21-6D* Haplotypes across the 10 Chinese Wheat Production Zones

Wheat production in China is divided into ten major agro-ecological production zones based on ecological conditions, variety type, and planting time (He et al., 2001; Zhuang, 2003). The distributions of *TaSPL20/21* haplotypes were investigated in both landraces (population 2) and modern cultivars (population 3 comprising 348 accessions) from each zone (Fig. 8). Comparisons of landraces and modern cultivars across the ten zones revealed the detailed selection histories of *TaSPL20-7D-Hap*I (Fig. 8, A and B), *TaSPL21-6A-Hap*II (Fig. 8, C and D), and *TaSPL21-6D-Hap*II&III (Fig. 8, E and F), including their origins and expansion paths during wheat breeding. *TaSPL20-7D-Hap*I first appeared in landraces in zone IV and rapidly spread to nearby areas (zones I, II, III, V, VIII, IX, and X) in current modern cultivars (Fig. 8, A and B). *TaSPL21-6A-Hap*II was relatively frequent in zones III, IV, and V where autumn-planted spring wheat is grown and frequencies generally increased in modern cultivars (Fig. 8, C and D). *TaSPL21-6D-Hap*II was present in landraces only in zone VI and *TaSPL21-6D-Hap*III was not detected in landraces. *TaSPL21-6D-Hap*II&III overwhelmingly replaced *TaSPL21-6D-Hap*I in modern cultivars in almost all zones (Fig. 8, E and F). Geographic distribution patterns further demonstrated that favorable haplotypes were strongly selected, but to different degrees, across the ten production zones.

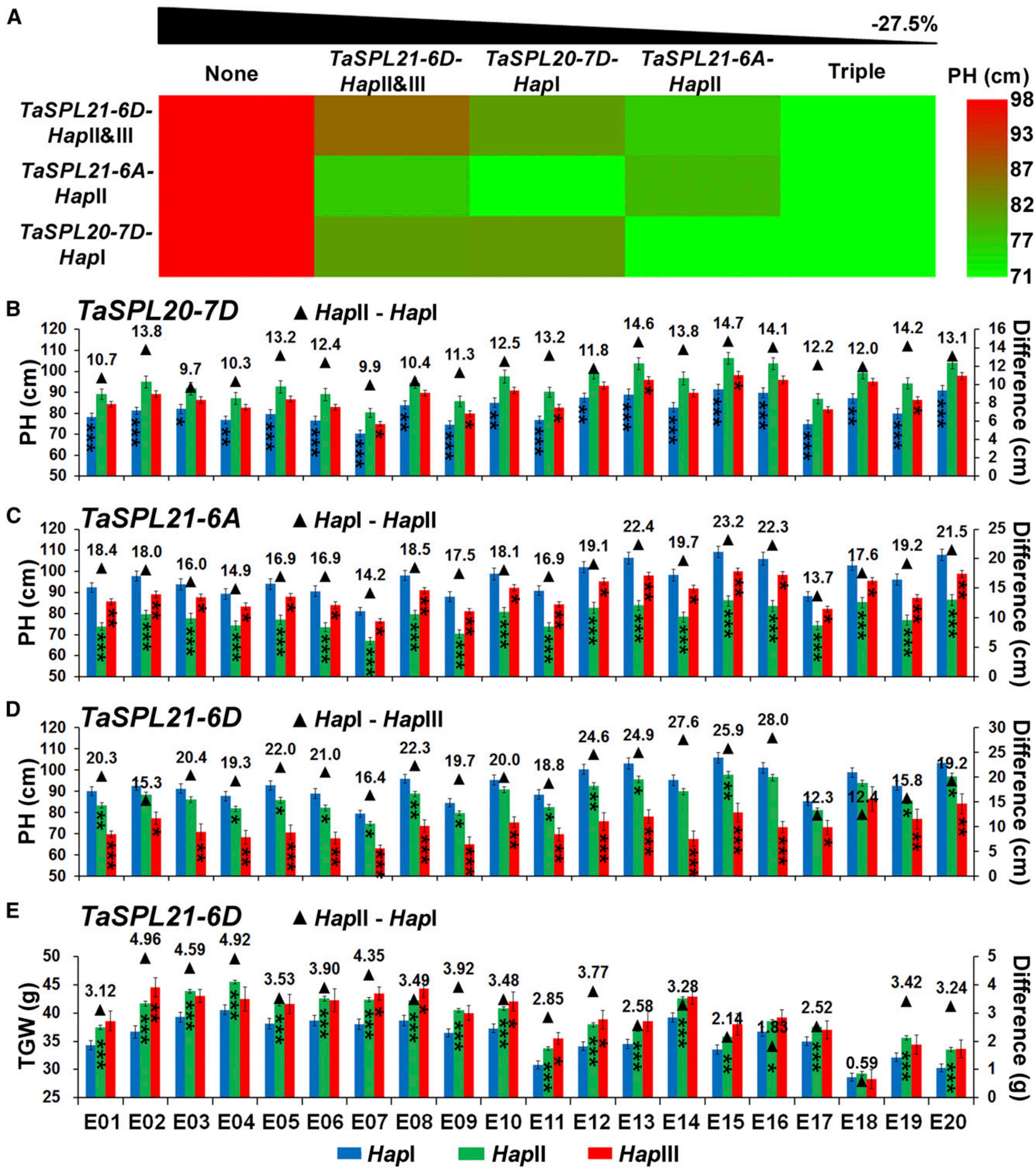

Figure 6. Phenotypic comparisons of large-effect haplotypes of *TaSPL20/21* homoeologs in 20 environments. A, Heatmap view of functional buffering effects on PH by combining *TaSPL20-7D-Hap*I, *TaSPL21-6A-Hap*II, and *TaSPL21-6D-Hap*II&III based on phenotypic data from 20 environments. None, no favorable haplotype; Triple, combination of three favorable haplotypes. B to D, Statistical analyses were performed by comparing haplotypes with the tallest PH and others. E, Statistical analysis compared haplotypes with the lowest TGW and others. $^{*}P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$. The error bars denote ± SE. E1 to E20, see legend of Fig. 5 for abbreviations.

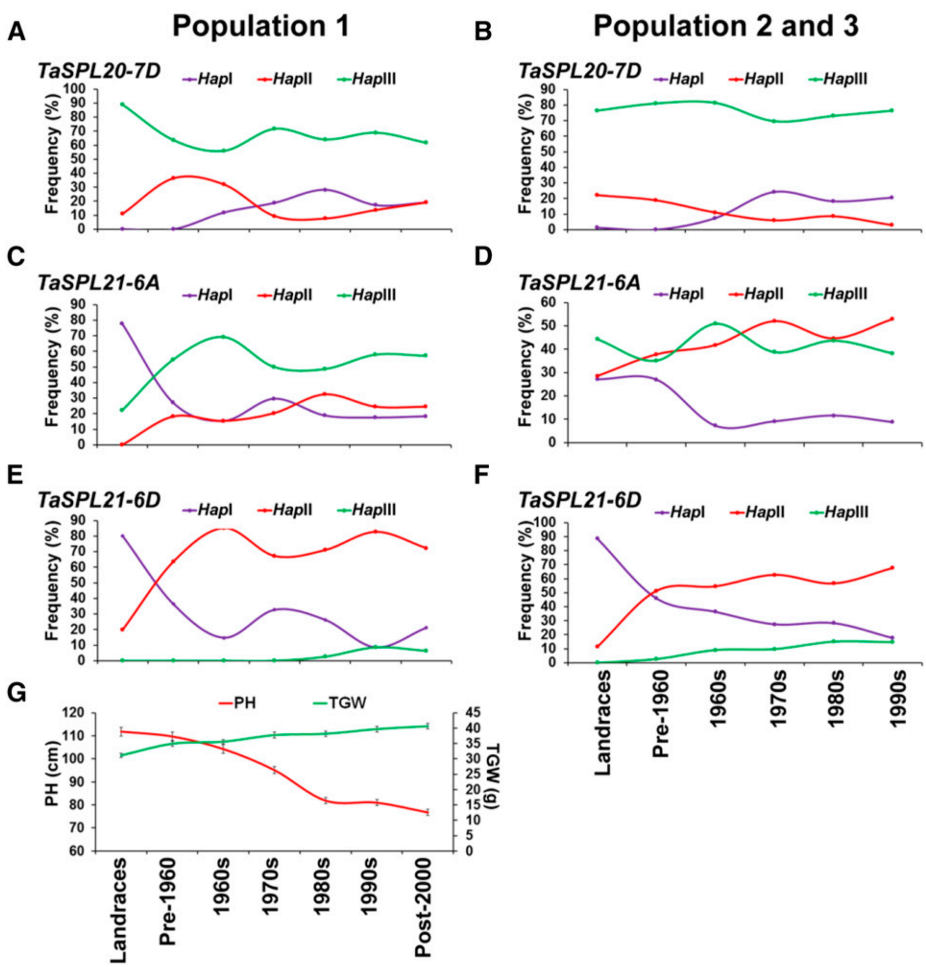

Figure 7. Favorable haplotypes of *TaSPL20/21* homoeologs selected in wheat breeding. Frequency of *TaSPL20-7D* (A and B), *TaSPL21-6A* (C and D), and *TaSPL21-6D* (E and F) haplotype changes over decades in landraces and modern cultivars. Changes in PH and TGW in population 1 (262 accessions) grown in 20 environments (G). Frequencies in population 1 (A, C, and E); verification of frequency distributions in populations 2 and 3 (157 landraces and 348 modern cultivars; B, D, and F). In population 1, 10 landraces are represented; 11, 27, 55, 40, 59, and 50 accessions were released pre-1960s, 1960s, 1970s, 1980s, 1990s, and post-2000, respectively; 10 accessions with unknown release dates were excluded. In population 3, there were 37, 55, 102, 106, and 34 accessions released in pre-1960s, 1960s, 1970s, 1980s, and 1990s, respectively; 14 accessions with unknown release dates were excluded. Bars indicate 2 × SE.

Evolutionary Origins of the *TaSPL20-7D*, *TaSPL21-6A*, and *TaSPL21-6D* Haplotypes

To obtain further insights on the origins of the *TaSPL20-7D*, *TaSPL21-6A*, and *TaSPL21-6D* haplotypes, we evaluated these loci in diploid, tetraploid, and hexaploid wheat and related species. These included 37 accessions of common wheat and 26 accessions of wild wheat species, viz. seven *T. urartu* (AA-genome), nine *T. dicoccoides* (AABB-genome), and 10 *Ae. tauschii* (DD-genome) accessions. The nucleotide diversities (π) across the entire *TaSPL20-7D* region were 0.01120 and 0.00024 in *Ae. tauschii* and common wheat, respectively (Fig. 9A), and the entire *TaSPL21-6D* region 0.00333 and 0.00012 in *Ae. tauschii* and common wheat (Fig. 9C). Thus, compared to *Ae. tauschii* π values at *TaSPL20-7D* and *TaSPL21-6D* in common wheat were significantly less, reflecting the bottleneck effect of the second polyploidization.

The π values for the entire *TaSPL21-6A* region were 0.00000, 0.00013, and 0.00039 in *T. urartu*, *T. dicoccoides*, and common wheat, respectively (Fig. 9B). In seven AA-genome accessions across the full 3463 bp length, we detected no sequence variation, indicating that *TaSPL21-6A* was very conserved in *T. urartu*. However, wide variations were found between the sequences in *T. urartu* and common wheat (Fig. 9D), and there were only minor differences between those in *T. dicoccoides* and common wheat (Fig. 9E). These findings suggest *TaSPL21-6A* underwent strong selection at the first wheat ployploidization event.

In regard to polymorphic sites at *TaSPL21-6A* in common wheat we detected one haplotype (*Hap*III*-like*) in *T. urartu*, and two haplotypes (*Hap*II*-like* and *Hap*III*-like*) in *T. dicoccoides*. The favorable *TaSPL21-6A-Hap*II haplotype likely arose after the first wheat ployploidization event. Similarly, for *TaSPL20-7D* and *TaSPL21-6D*, we detected only the *TaSPL20-7D-HapIII-like* and *TaSPL21-6D-HapI-like* alleles in *Ae. tauschii*. In this case, the favorable haplotypes *TaSPL20-7D-Hap*I and *TaSPL21-6D-Hap*II&III likely arose after the second ployploidization event.

DISCUSSION

SBP-box genes exist only in plants and originated before the divergence of the green algae and ancestors of land plants (Guo et al., 2008). Following differentiation of developmental characteristics in land plants, SBP-box genes might possess ancient and neofunctional

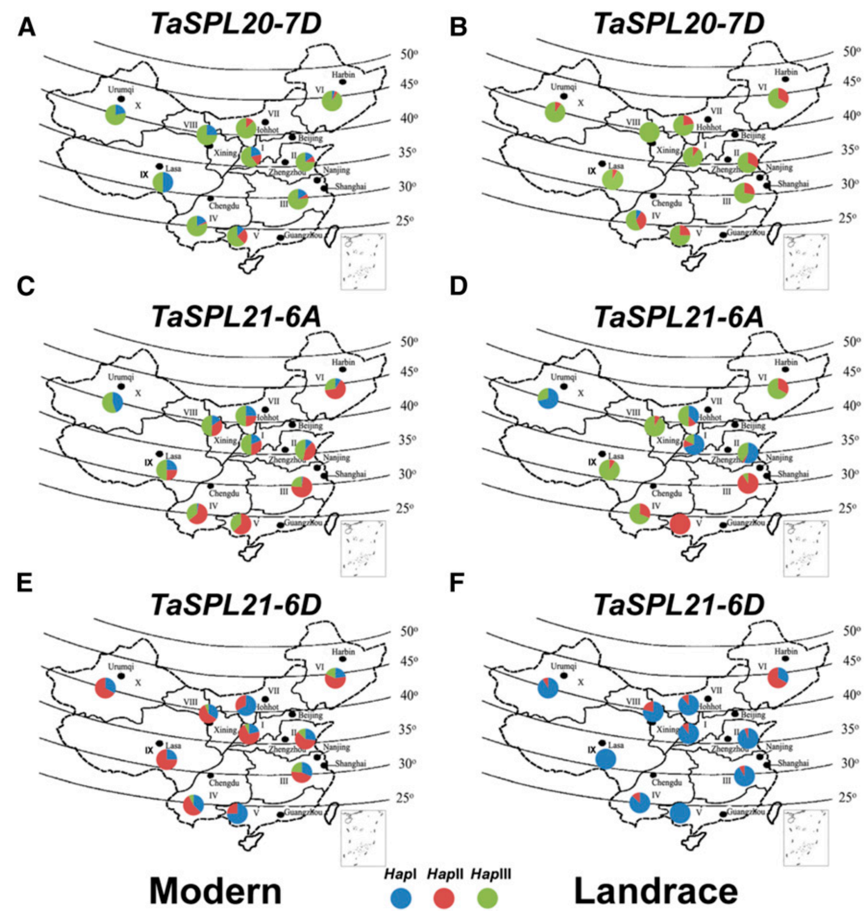

Figure 8. Geographic distributions of *TaSPL20-7D*, *TaSPL21-6A*, and *TaSPL21-6D* haplotypes in Chinese wheat landraces (157 accessions) and modern cultivars (348 accessions) across the 10 Chinese wheat ecological production zones. Distributions of *TaSPL20-7D* (A), *TaSPL21-6A* (C), and *TaSPL21-6D* (E) haplotypes in 348 modern cultivars. Distributions of *TaSPL20-7D* (B), *TaSPL21-6A* (D), and *TaSPL21-6D* (F) haplotypes in 157 landraces. I, Northern Winter Wheat Zone; II, Yellow and Huai River Valleys Facultative Wheat Zone; III, Middle and Lower Yangtze Valleys Autumn-Sown Spring Wheat Zone; IV, Southwestern Autumn-Sown Spring Wheat Zone; V, Southern Autumn-Sown Spring Wheat Zone; VI, Northeastern Spring Wheat Zone; VII, Northern Spring Wheat Zone; VIII, Northwestern Spring Wheat Zone; IX, Qinghai-Tibetan Plateau Spring-Winter Wheat Zone; X, Xinjiang Winter-Spring Wheat Zone. The maps were generated using Mapinfo Professional software.

evolutionary patterns (Zhang et al., 2015a). In seed plants, SBP-box genes have various degrees of similarity between different family members because of duplications at both the gene and genome levels (Riese et al., 2007). Based on the phylogenetic tree and expression patterns, *TaSPL20* and *TaSPL21* are considered to be paralogous genes. The functions of homoeologous genes at the *TaSPL20*/*TaSPL21* loci were characterized at two levels, the protein level (transgenic overexpression lines) and nucleotide level (haplotypes). We described the effects of selection and application of favorable haplotypes on PH and TGW during wheat domestication and breeding.

AtSPL3, *AtSPL4*, and *AtSPL5*, closely related members in Arabidopsis, promote vegetative phase change and flowering (Wu and Poethig, 2006). *AtSPL9* and *AtSPL15*, considered to be paralogous genes, act redundantly in regulating plastochron length, juvenile-to-adult phase transition, and organ size (Schwarz et al., 2008; Wang et al., 2008). Loss of *AtSPL2* function slightly enhanced the phenotype of *atspl9 atspl15* double mutants (Schwarz et al., 2008). Another pair of paralogs, *AtSPL10* and *AtSPL11*, control leaf shape and epidermal traits (Shikata et al., 2009). Duplicated SBP-box gene pairs have also been found in rice, such as *OsSPL3* and *OsSPL12*, *OsSPL4* and *OsSPL11*, *OsSPL5* and *OsSPL10* (Guo et al., 2008; Zhang et al., 2014a). In common wheat, we found that *TaSPL20* and *TaSPL21* were highly expressed in the lemma and palea during early spike development (Fig. 1A), and their proteins were exclusively localized in the nucleus (Fig. 1B). Ectopic expression analysis of *TaSPL20* or *TaSPL21* in rice showed similar functions by increasing the number of primary branches (21–34%; Fig. 2C), secondary branches (19–23%; Fig. 2D), grain number (24–27%; Figure 2E), and panicle length (~8%; Fig. 2F). Overexpression of *TaSPL20* produced larger seeds, whereas overexpression of *TaSPL21* did not. *TaSPL20* and *TaSPL21* were located on group 7 and 6 chromosomes, respectively (Fig. 4A). Their proteins share conserved functions in promoting branching at the reproductive stage but display divergent functions during seed development.

A key factor in wheat becoming a global food crop was its adaptation to a wide range of environmental conditions, largely attributed to its allohexaploid genomic plasticity and wide genetic variation (Feldman and Levy, 2012). Allopolyploidization events in wheat did not lead to functional dominance of one subgenome over the others (Mayer et al., 2014). Instead, it underwent functional or genetic diploidizaiton in two ways; the first was rapid genomic change (revolutionary change) through generating genetic or epigenetic

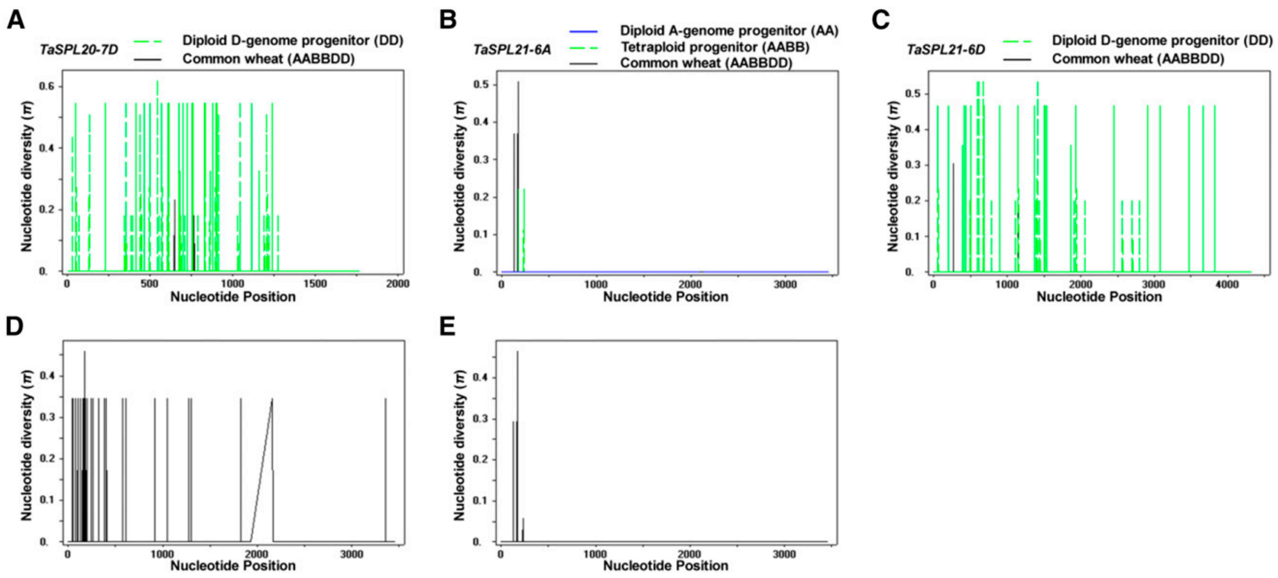

Figure 9. Nucleotide diversities (π) in *TaSPL20-7D* (A), *TaSPL21-6A* (B), and *TaSPL21-6D* (C) haplotypes between diploid, tetraploid, and hexaploid species. Seven AA-genome (*T. urartu*), 9 AABB-genome (*T. dicoccoides*), 10 DD-genome (*Ae. tauschii*), and 37 common wheat accessions were used. D, π at *TaSPL20-7D* in the *T. urartu* and common wheat gene pools; E, π at *TaSPL20-7D* in *T. dicoccoides* and common wheat gene pools.

alterations during or soon after allopolyploidization; the second was slow genome change (evolutionary change) through subfunctionalization (evolution of partitioned functions among ancestral alleles or homoeoalleles), neofunctionalization (evolution of novel functions in specific alleles or homoeoalleles), new allelic variations, and variation in dosage effects through copy number variation (Feldman and Levy, 2005; Feldman et al., 2012). For example, some genes located in the A or B subgenomes, that are silenced in allohexaploid wheat, are expressed in extracted allotetraploids (AABB) but silenced again upon adding the D subgenome (Zhang et al., 2014b). This transregulation is likely a transitory, compensatory phenomenon (Pont et al., 2013). In the course of wheat evolution, more stable cis-acting regulations were established at the allohexaploid level (Zhang et al., 2014b). In a similar way, *TaSPL20* and *TaSPL21* homoeologs underwent diversification in function with each evolving its own distinctive characteristics according to the functional analysis of naturally occurring variants (Figs. 5 and 10). *TaSPL20-7A* had no functional effect on six yield-related traits; *TaSPL20-7D* had a moderate effect on PH and was weakly associated with TGW; *TaSPL21-6D* had strong effects on PH and TGW; the effect of *TaSPL21-6A* on PH was stronger than *TaSPL20-7D* and *TaSPL21-6D*; *TaSPL21-6B* showed weak effects on PH, total number of spikelets per spike, and number of grains per spike. Thus, through evolution, domestication, and breeding, homoeologous genes at the paralogous *TaSPL20* and *TaSPL21* loci became functionally different: *TaSPL20-7A* was silenced, *TaSPL21-6A* underwent subfunctionalization; the function of *TaSPL21-6D* was enhanced; and *TaSPL20-7D* and *TaSPL21-6B* likely underwent subfunctionalization and were silenced, respectively (Fig. 10).

Wheat was domesticated in the Fertile Crescent ~10,000 years ago and underwent two rounds of polyploidization. Due to duplicated functionality, sequence insertions/deletions and point mutations led to buffering effects rather than lethality or sublethality, providing the plasticity to generate adaptive variants (Comai, 2005; Dubcovsky and Dvorak, 2007). Such mutations account for genetic diversity, and some of them took place exclusively in the allopolyploid background (Dvorak et al., 2004; Feldman and Levy, 2012). In addition, domestication led to selection and spread of favorable genes or alleles affecting important agronomic traits (Clark et al., 2004). For example, in rice (a diploid species) knockout of *GPC* genes results in almost complete sterility, whereas in tetraploid wheat, *GPC-B1* mutation led only to a small change in maturity, and in hexaploid wheat editing of one of the three homoeologs had even more subtle effects (Dubcovsky and Dvorak, 2007). Another example is the deletion within the upstream regulatory region of *Ppd-D1* that causes photoperiod insensitivity (Faure et al., 2007). In this study, 8, 0, 2, 3, 13, and 2 variations were identified in *TaSPL20-7A*, *TaSPL20-7B*, *TaSPL20-7D*, *TaSPL21-6A*, *TaSPL21-6B*, and *TaSPL21-6D*, and 3, 0, 3, 3, 3, and 3 haplotypes were detected in common wheat based on the respective DNA sequences (Fig. 4). Among them, we identified several favorable haplotypes across a range of environmental conditions (4 years × 2 sites × well-watered,

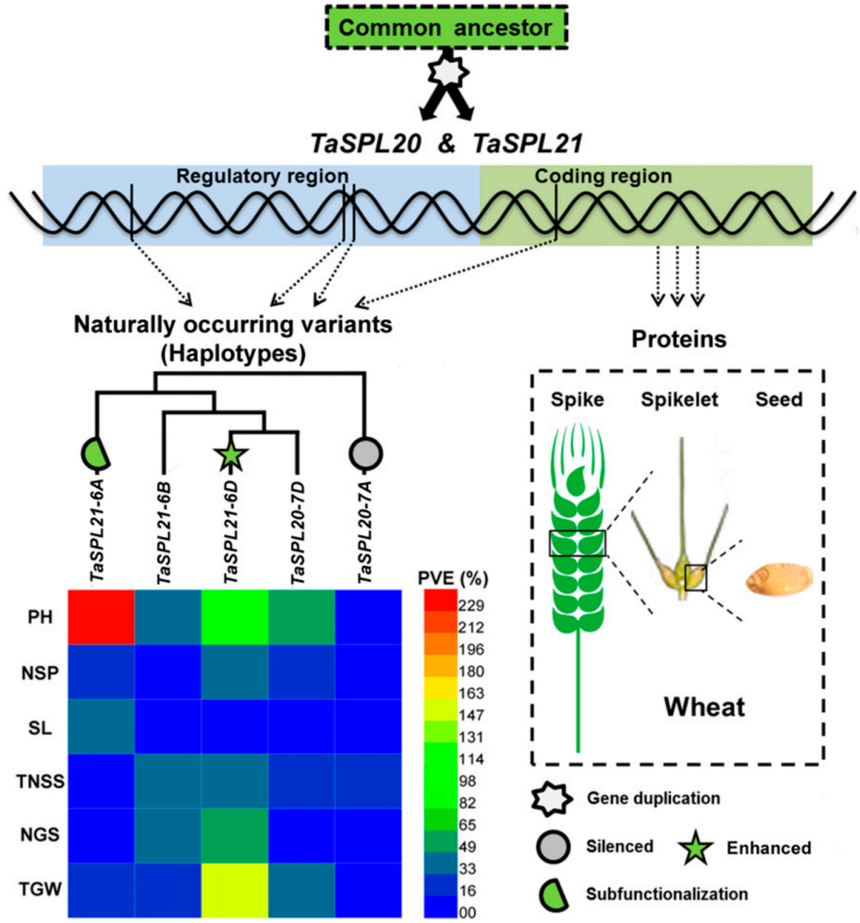

Figure 10. Model for functional conservation and divergence among homoeologous genes at the *TaSPL20*/*TaSPL21* loci governing yield-related traits in hexaploid wheat. The functions of homoeologous genes at the *TaSPL20*/*TaSPL21* loci were characterized at two levels: the protein level (right) and nucleotide variation/haplotype level (left). The protein functions of the *TaSPL20/TaSPL21* loci in wheat were deduced from the phenotypes of ectopic expression in rice (shown in a dotted rectangle), i.e. both *TaSPL20* and *TaSPL21* may similarly affect panicle branching and function during wheat spike and spikelet development; they may also affect seed development but with different functions in wheat. According to the functional analysis of naturally occurring variants, *TaSPL20* and *TaSPL21* homoeologs display diverse functions, with each having distinctive characteristics after evolution, domestication, and breeding, i.e. the function of *TaSPL20-7A* was silenced, *TaSPL21-6A* underwent subfunctionalization, the function of *TaSPL21-6D* was enhanced, *TaSPL20-7D* and *TaSPL21-6B* likely underwent subfunctionalization and was silenced, respectively. Heat map summarizes accumulative PVE (phenotypic variation explained) values of *TaSPL20-7A*, *TaSPL20-7D*, *TaSPL21-6A*, *TaSPL21-6B*, and *TaSPL21-6D* haplotypes in 20 environments. Vertical line, naturally occurring variants; PH, plant height; NSP, number of spikes per plant; SL, spike length; TNSS, total number of spikelets per spike; NGS, number of grains per spike; TGW, 1000-grain weight.

drought, and heat-stress conditions): *TaSPL20-7D-Hap*I, *TaSPL21-6A-Hap*II, and *TaSPL21-6D-Hap*II&III reduced PH by 13.00%, 18.96%, and 21.68%, respectively, compared to haplotypes with the highest PH, and *TaSPL21-6D-Hap*II increased TGW by 9.73% compared to *Hap*I. The combined effect of *TaSPL20/21* homoeologs reduced PH by as much as 27.5%, from 98.56 to 71.48 cm (Fig. 6). The favorable haplotype *TaSPL21-6A-Hap*II was likely derived after the first ployploidization, whereas *TaSPL20-7D-Hap*I and *TaSPL21-6D-Hap*II&III likely evolved after the second ployploidization. Evolutionary analysis, geographic distribution, and frequency change data suggest that favorable haplotypes at the *TaSPL20* and *TaSPL21* loci underwent selection but at different degrees during wheat domestication and breeding.

CONCLUSION

Wheat has become one of the most important crops in the world and experienced two rounds of polyploidization. Drought, heat, and other abiotic stresses greatly affect wheat productivity. High grain yield and yield stability are always major objectives in wheat improvement. In this study, we presented a comprehensive analysis of the biological functions and fates of triplicated homoeologs of paralogous genes *TaSPL20* and *TaSPL21* in modern hexaploid wheat genotypes. Their proteins share conserved functions in promoting panicle branching and influencing seed development. Functional analyses of the natural variants at the *TaSPL20* and *TaSPL21* loci suggested that the triplicated homoeologs of the two gene sets underwent diversification in function, with each evolving its own distinctive characteristics. Among them, we identified several favorable haplotypes of *TaSPL20-7D*, *TaSPL21-6A*, *TaSPL21-6D* across a range of environmental conditions, which were strongly associated with important yield-related traits, such as PH and TGW. We also observed the selection and exploitation of favorable haplotypes during wheat breeding in China. In addition, all the natural variants occurred in the 5′ flanking regions of *TaSPL20-7D*, *TaSPL21-6A*, and *TaSPL21-6D*. Further dissecting the upstream regulator of these three genes will facilitate our understanding of the molecular mechanisms of these important natural variants in regulating yield formation.

MATERIALS AND METHODS

Plant Materials and Phenotypic Assessment

Common wheat cultivar Yanzhan 4110 was used for gene cloning and expression analysis. Thirty-seven cultivars (Supplemental Table S1) showing wide variation in spike-related traits and tiller number were used to detect naturally occurring variants in target gene sequences (Zhang et al., 2015b). Twenty-nine accessions of wheat-related species were chosen for evolutionary studies, including seven diploid A-genome progenitor accessions, three diploid B-genome progenitor accessions, nine tetraploid (AABB) progenitor accessions, and 10 diploid D-genome progenitor accessions. A set of Chinese Spring nullisomic-tetrasomic lines was used for chromosome location. Three common wheat germplasm populations were used in the study: population 1 (262 accessions; Supplemental Table S2) was used for population structure and association analysis and consisted of 209 modern varieties, 43 advanced lines and 10 landraces (Zhang et al., 2013), population 2 comprised 157 landraces (Supplemental Table S8), and population 3, with 348 modern cultivars (Supplemental Table S9), was used for determination of haplotype frequencies and geographic distribution analysis.

Population 1 was planted at the Experimental Stations at Changping (116°13′E; 40°13′N) and Shunyi (116°56′E; 40°23′′N) of the Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing, over 4 years (2009–2012) for measuring six agronomic traits, viz., PH, NSP, SL, TNSS, NGS, and TGW. Two water regimes, rain-fed (drought stress, DS) and well-watered (WW), were applied at each site. The amounts of rainfall in the growing seasons were 192, 131, 180, and 158 mm, respectively. The WW plots were irrigated with 750 m^3/ha (75 mm) at the preoverwintering, booting, flowering, and grain filling stages when the amounts of rainfall were insufficient during each corresponding period. In addition, greenhouse experiments using polythene covers at flowering to increase the temperature and thereby to simulate heat stress (HS) were performed at Shunyi. During the HS period, the average highest temperature outside the greenhouse was 33°C; the average high temperature inside the greenhouse under DS condition was 43°C, whereas in the greenhouse with WW conditions the temperature was 41°C. The 20 environments (E1–E20) indicate the environments at Changping in 2009 under DS and WW conditions, Shunyi in 2009 under DS, WW, DS + HS, and WW + HS, Changping in 2010 under DS and WW, Shunyi in 2010 under DS, WW, DS + HS, and WW + HS, Shunyi in 2011 under DS, WW, DS + HS, and WW + HS, Changping in 2012 under DS and WW, and Shunyi in 2012 under DS and WW, respectively.

PCR Primers, Sequence Analysis, and Statistical Analyses

All primers were designed by Primer Premier 5.0 software (http://www.premierbiosoft.com/) and listed in Supplemental Table S10. Sequence alignment and SNP detection were carried out by DNASTAR Lasergene 7.1.0 (DNASTAR). Nucleotide diversity (π) and haplotype variations were analyzed by DnaSP 5.10 software (http://www.ub.edu/dnasp/). Genetic mapping was performed with MAPMAKER/EXP 3.0 (Lander et al., 1987).

Sample Preparation and Real-Time Quantitative PCR

Wheat (Yanzhan 4110) tissue samples included the stamens, pistils, lemmas, paleas, and pedicels were collected from the spikes with lengths of 3, 4, 6, 8, and 10 cm and at the flowering stage. Quantitative real-time PCR was carried out in triplicate with an ABI Prism 7500 system using a SYBR Green PCR Master Mix kit (TaKaRa Biotechnology). Relative gene expression levels were estimated by the $2^{-\triangle\triangle CT}$ method (Livak and Schmittgen, 2001).

Subcellular Localization

The full-length open reading frame (ORF) of *TaSPL20* and *TaSPL21* from the D genome were fused upstream of GFP in the pJIT163-GFP expression vector driven by the CaMV35S promoter. The constructs were transferred into wheat mesophyll protoplasts by the polyethylene glycol-mediated method followed by incubation at 25°C for 16 h (Yoo et al., 2007) and into live onion epidermal cells by biolistic bombardment (Helios; Bio-Rad) followed by incubation at 28°C for 36 to 48 h. Transformed cells were observed with a laser scanning confocal microscope (Leica TCS-NT).

Generation and Phenotyping of Transgenic Rice Plants

The full-length ORF of *TaSPL20* and *TaSPL21* from the D genome were amplified and inserted into pCUbi1390 vectors. The constructs were transferred into rice (*Oryza sativa* ssp. *japonica*) cv Kitaake by *Agrobacterium*-mediated transformation (Hiei et al., 1994). Putative transformants were examined by 0.1% hygromycin and subsequently confirmed by PCR genotyping. Homozygous lines overexpressing *TaSPL20* and *TaSPL21* and wild type were grown in an area dedicated to transgenic plants at the Institute of Crop Science Experimental Station (116°28′E; 39°48′N) in Langfang. More than 50 individuals for each line were used for phenotypic assays, which included tiller number, number of primary and secondary panicle branches, panicle length, grain number per main panicle, TGW, grain area, grain length, and grain width.

SNP Detection and Functional Marker Development

Genome-specific fragments were cloned into pEASY-Blunt vectors and transformed to *Escherichia coli* DH5α competent cells. Positive clones were sequenced with an ABI 3730XI 96 capillary DNA analyzer (Applied Biosystems). CAPS markers or dCAPS markers were developed based on SNPs (Supplemental Table S10). Target fragments were amplified by the corresponding primers and separated by electrophoresis in agarose gels after digestion by a specific restriction endonuclease.

Population Structure and Association Analysis

Population structure was estimated by Structure v2.3.2 using data from 209 whole-genome SSR markers (Li et al., 2015). Association mapping was conducted using the general linear model in TASSEL V2.1 that accounted for population structure (Q). Statistical analysis was conducted by SAS 8.01 software.

Accession Numbers

Sequence data from this article can be found in the GenBank database under the following accession numbers: the full-length ORF of *TaSPL20* (KF447878) and *TaSPL21* (KF447882) from the D genome. The genomic DNA sequences of *TaSPL20-7A*, *TaSPL20-7B*, *TaSPL20-7D*, *TaSPL21-6A*, *TaSPL21-6B*, and *TaSPL21-6D* have been submitted to GenBank with accession numbers KY114060, KY114061, KY114063, KY114058 (the ORF sequence, KY114059), KY114062, and KY114057, respectively.

Supplemental Data

The following supplemental materials are available.

Supplemental Figure S1. Transgenic rice lines overexpressing *TaSPL20* and *TaSPL21* were grown under field conditions.

Supplemental Table S1. Thirty-seven cultivars with wide variation in spike-related traits and tiller number.

Supplemental Table S2. Basic information of 262 accessions (names, origins, and released dates) and sequence polymorphism assays of homoeologous genes at the *TaSPL20/TaSPL21* loci.

Supplemental Table S3. *TaSPL20-7A* haplotype associations with agronomic traits in 20 environments.

Supplemental Table S4. *TaSPL20-7D* haplotype associations with agronomic traits in 20 environments.

Supplemental Table S5. *TaSPL21-6A* haplotype associations with agronomic traits in 20 environments.

Supplemental Table S6. *TaSPL21-6B* haplotype associations with agronomic traits in 20 environments.

Supplemental Table S7. *TaSPL21-6D* haplotype associations with agronomic traits in 20 environments.

Supplemental Table S8. Basic information of 157 landrace accessions (names, origins, and ecological zones) and sequence polymorphisms of homoeologous genes at *TaSPL20/TaSPL21* loci.

Supplemental Table S9. Basic information of 348 accessions (names, origins, ecological zones, and released dates) and sequence polymorphisms of homoeologous genes at *TaSPL20/TaSPL21* loci.

Supplemental Table S10. Primers used in this study.

ACKNOWLEDGMENTS

We thank Professor Robert A. McIntosh (Plant Breeding Institute, University of Sydney, New South Wales, Australia) for revising the manuscript.We also thank our colleagues Dr. Jianmin Wan and his group members (Institute of Crop Science, Chinese Academy of Agricultural Sciences) for help in rice transformation.

Received January 30, 2017; accepted April 15, 2017; published April 19, 2017.

LITERATURE CITED

Barrero RA, Bellgard M, Zhang X (2011) Diverse approaches to achieving grain yield in wheat. Funct Integr Genomics **11:** 37–48

Cardon GH, Höhmann S, Nettesheim K, Saedler H, Huijser P (1997) Functional analysis of the *Arabidopsis thaliana* SBP-box gene *SPL3*: A novel gene involved in the floral transition. Plant J **12:** 367–377

Chuck G, Whipple C, Jackson D, Hake S (2010) The maize SBP-box transcription factor encoded by *tasselsheath4* regulates bract development and the establishment of meristem boundaries. Development **137:** 1243–1250

Clark RM, Linton E, Messing J, Doebley JF (2004) Pattern of diversity in the genomic region near the maize domestication gene *tb1*. Proc Natl Acad Sci USA **101:** 700–707

Comai L (2005) The advantages and disadvantages of being polyploid. Nat Rev Genet **6:** 836–846

Dubcovsky J, Dvorak J (2007) Genome plasticity a key factor in the success of polyploid wheat under domestication. Science **316:** 1862–1866

Dvorak J, Yang Z-L, You FM, Luo M-C (2004) Deletion polymorphism in wheat chromosome regions with contrasting recombination rates. Genetics **168:** 1665–1675

Faure S, Turner A, Beales J, Higgins J, Laurie D (2007) Photoperiodic control of flowering time in barley and wheat. Plant and Animal Genome XV, San Diego, CA

Feldman M, Levy AA (2005) Allopolyploidy—a shaping force in the evolution of wheat genomes. Cytogenet Genome Res **109:** 250–258

Feldman M, Levy AA (2012) Genome evolution due to allopolyploidization in wheat. Genetics **192:** 763–774

Feldman M, Levy AA, Fahima T, Korol A (2012) Genomic asymmetry in allopolyploid plants: Wheat as a model. J Exp Bot **63:** 5045–5059

Guo AY, Zhu QH, Gu X, Ge S, Yang J, Luo J (2008) Genome-wide identification and evolutionary analysis of the plant specific SBP-box transcription factor family. Gene **418:** 1–8

Hao C, Dong Y, Wang L, You G, Zhang H, Ge H, Jia J, Zhang X (2008) Genetic diversity and construction of core collection in Chinese wheat genetic resources. Chin Sci Bull **53:** 1518–1526

Hao C, Wang L, Ge H, Dong Y, Zhang X (2011) Genetic diversity and linkage disequilibrium in Chinese bread wheat (*Triticum aestivum* L.) revealed by SSR markers. PLoS One **6:** e17279

He ZH, Rajaram S, Xin ZY, Huang GZ (2001) A History of Wheat Breeding in China. CIMMYT, Mexico City, Mexico.

Hiei Y, Ohta S, Komari T, Kumashiro T (1994) Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA. Plant J **6:** 271–282

Jia J, Zhao S, Kong X, Li Y, Zhao G, He W, Appels R, Pfeifer M, Tao Y, et al, International Wheat Genome Sequencing Consortium (2013) *Aegilops tauschii* draft genome sequence reveals a gene repertoire for wheat adaptation. Nature **496:** 91–95

Jiao Y, Wang Y, Xue D, Wang J, Yan M, Liu G, Dong G, Zeng D, Lu Z, Zhu X, et al (2010) Regulation of *OsSPL14* by OsmiR156 defines ideal plant architecture in rice. Nat Genet **42:** 541–544

Klein J, Saedler H, Huijser P (1996) A new family of DNA binding proteins includes putative transcriptional regulators of the *Antirrhinum majus* floral meristem identity gene *SQUAMOSA*. Mol Gen Genet **250:** 7–16

Kropat J, Tottey S, Birkenbihl RP, Depège N, Huijser P, Merchant S (2005) A regulator of nutritional copper signaling in *Chlamydomonas* is an SBP domain protein that recognizes the GTAC core of copper response element. Proc Natl Acad Sci USA **102:** 18730–18735

Lännenpää M, Jänönen I, Hölttä-Vuori M, Gardemeister M, Porali I, Sopanen T (2004) A new SBP-box gene *BpSPL1* in silver birch (*Betula pendula*). Physiol Plant **120:** 491–500

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newberg LA (1987) MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics **1:** 174–181

Li W, Zhang B, Li R, Chang X, Jing R (2015) Favorable alleles for stem water-soluble carbohydrates identified by association analysis contribute to grain weight under drought stress conditions in wheat. PLoS One **10:** e0119438

Ling H-Q, Zhao S, Liu D, Wang J, Sun H, Zhang C, Fan H, Li D, Dong L, Tao Y, et al (2013) Draft genome of the wheat A-genome progenitor *Triticum urartu*. Nature **496:** 87–90

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{(-\Delta\Delta C(T))}$ method. Methods **25:** 402–408

Lobell DB, Schlenker W, Costa-Roberts J (2011) Climate trends and global crop production since 1980. Science **333:** 616–620

Manning K, Tör M, Poole M, Hong Y, Thompson AJ, King GJ, Giovannoni JJ, Seymour GB (2006) A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. Nat Genet **38:** 948–952

Marcussen T, Sandve SR, Heier L, Spannagl M, Pfeifer M, Jakobsen KS, Wulff BB, Steuernagel B, Mayer KF, et al, International Wheat Genome Sequencing Consortium (2014) Ancient hybridizations among the ancestral genomes of bread wheat. Science **345:** 1250092

Mayer KF, Rogers J, Doležel J, Pozniak C, Eversole K, Feuillet C, Gill B, Friebe B, Lukaszewski AJ, et al, International Wheat Genome Sequencing Consortium (IWGSC) (2014) A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. Science **345:** 1251788

Miura K, Ikeda M, Matsubara A, Song X-J, Ito M, Asano K, Matsuoka M, Kitano H, Ashikari M (2010) *OsSPL14* promotes panicle branching and higher grain productivity in rice. Nat Genet **42:** 545–549

Pfeifer M, Kugler KG, Sandve SR, Zhan B, Rudi H, Hvidsten TR, Mayer KF, Olsen O-A, International Wheat Genome Sequencing Consortium (2014) Genome interplay in the grain transcriptome of hexaploid bread wheat. Science **345:** 1250091

Pont C, Murat F, Confolent C, Balzergue S, Salse J (2011) RNA-seq in grain unveils fate of neo- and paleopolyploidization events in bread wheat (*Triticum aestivum* L.). Genome Biol **12:** R119

Pont C, Murat F, Guizard S, Flores R, Foucrier S, Bidet Y, Quraishi UM, Alaux M, Doležel J, Fahima T, et al (2013) Wheat syntenome unveils new evidences of contrasted evolutionary plasticity between paleo- and neoduplicated subgenomes. Plant J **76:** 1030–1044

Riese M, Höhmann S, Saedler H, Münster T, Huijser P (2007) Comparative analysis of the SBP-box gene families in *P. patens* and seed plants. Gene **401:** 28–37

Riese M, Zobell O, Saedler H, Huijser P (2008) SBP-domain transcription factors as possible effectors of cryptochrome-mediated blue light signalling in the moss *Physcomitrella patens*. Planta **227:** 505–515

Schwarz S, Grande AV, Bujdoso N, Saedler H, Huijser P (2008) The microRNA regulated SBP-box genes *SPL9* and *SPL15* control shoot maturation in *Arabidopsis*. Plant Mol Biol **67:** 183–195

Shikata M, Koyama T, Mitsuda N, Ohme-Takagi M (2009) *Arabidopsis* SBP-box genes *SPL10*, *SPL11* and *SPL2* control morphological change in association with shoot maturation in the reproductive phase. Plant Cell Physiol **50:** 2133–2145

Si L, Chen J, Huang X, Gong H, Luo J, Hou Q, Zhou T, Lu T, Zhu J, Shangguan Y, et al (2016) *OsSPL13* controls grain size in cultivated rice. Nat Genet **48:** 447–456

Tanno K, Willcox G (2006) How fast was wild wheat domesticated? Science **311:** 1886

Usami T, Horiguchi G, Yano S, Tsukaya H (2009) The *more and smaller cells* mutants of *Arabidopsis thaliana* identify novel roles for *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE* genes in the control of heteroblasty. Development **136:** 955–964

Wang JW, Czech B, Weigel D (2009) miR156-regulated SPL transcription factors define an endogenous flowering pathway in *Arabidopsis thaliana*. Cell **138:** 738–749

Wang B, Geng S, Wang D, Feng N, Zhang D, Wu L, Hao C, Zhang X, Li A, Mao L (2015) Characterization of *Squamosa Promoter Binding Protein-LIKE* genes in wheat. J Plant Biol **58:** 220–229

Wang H, Nussbaum-Wagler T, Li B, Zhao Q, Vigouroux Y, Faller M, Bomblies K, Lukens L, Doebley JF (2005) The origin of the naked grains of maize. Nature **436:** 714–719

Wang JW, Park MY, Wang LJ, Koo Y, Chen XY, Weigel D, Poethig RS (2011) miRNA control of vegetative phase change in trees. PLoS Genet **7:** e1002012

Wang JW, Schwab R, Czech B, Mica E, Weigel D (2008) Dual effects of miR156-targeted *SPL* genes and *CYP78A5/KLUH* on plastochron length and organ size in *Arabidopsis thaliana*. Plant Cell **20:** 1231–1243

Wang Y, Wu F, Bai J, He Y (2014) *BrpSPL9* (*Brassica rapa* ssp. *pekinensis SPL9*) controls the earliness of heading time in Chinese cabbage. Plant Biotechnol J **12:** 312–321

Wang S, Wu K, Yuan Q, Liu X, Liu Z, Lin X, Zeng R, Zhu H, Dong G, Qian Q, et al (2012) Control of grain size, shape and quality by *OsSPL16* in rice. Nat Genet **44:** 950–954

Wendel JF (2000) Genome evolution in polyploids. Plant Mol Biol **42:** 225–249

Wilkinson PA, Winfield MO, Barker GL, Allen AM, Burridge A, Coghill JA, Edwards KJ (2012) CerealsDB 2.0: An integrated resource for plant breeders and scientists. BMC Bioinformatics **13:** 219

Wu G, Poethig RS (2006) Temporal regulation of shoot development in *Arabidopsis thaliana* by *miR156* and its target *SPL3*. Development **133:** 3539–3547

Yoo SD, Cho YH, Sheen J (2007) *Arabidopsis* mesophyll protoplasts: A versatile cell system for transient gene expression analysis. Nat Protoc **2:** 1565–1572

Zhang SD, Ling LZ, Yi TS (2015a) Evolution and divergence of SBP-box genes in land plants. BMC Genomics **16:** 787

Zhang B, Liu X, Xu W, Chang J, Li A, Mao X, Zhang X, Jing R (2015b) Novel function of a putative *MOC1* ortholog associated with spikelet number per spike in common wheat. Sci Rep **5:** 12211

Zhang B, Liu X, Zhao G, Mao X, Li A, Jing R (2014a) Molecular characterization and expression analysis of *Triticum aestivum* squamosa-promoter binding protein-box genes involved in ear development. J Integr Plant Biol **56:** 571–581

Zhang Y, Schwarz S, Saedler H, Huijser P (2007) SPL8, a local regulator in a subset of gibberellin-mediated developmental processes in *Arabidopsis*. Plant Mol Biol **63:** 429–439

Zhang B, Shi W, Li W, Chang X, Jing R (2013) Efficacy of pyramiding elite alleles for dynamic development of plant height in common wheat. Mol Breed **32:** 327–338

Zhang H, Zhu B, Qi B, Gou X, Dong Y, Xu C, Zhang B, Huang W, Liu C, Wang X, et al (2014b) Evolution of the BBAA component of bread wheat during its history at the allohexaploid level. Plant Cell **26:** 2761–2776

Zhuang Q (2003) Chinese Wheat Improvement and Pedigree Analysis. Agricultural Press, Beijing, China