



Cloning and Characterization of *TaTGW-7A* Gene Associated with Grain Weight in Wheat via SLAF-seq-BSA

Ming-Jian Hu^{1†}, Hai-Ping Zhang^{1†}, Kai Liu¹, Jia-Jia Cao¹, Sheng-Xing Wang¹, Hao Jiang¹, Zeng-Yun Wu¹, Jie Lu¹, Xiao F. Zhu¹, Xian-Chun Xia^{1,2}, Gen-Lou Sun^{1,3}, Chuan-Xi Ma¹ and Cheng Chang^{1*}

¹ College of Agronomy, Anhui Agricultural University – Key Laboratory of Wheat Biology and Genetic Improvement on Southern Yellow & Huai River Valley, The Ministry of Agriculture, Hefei, China, ² National Wheat Improvement Center/The National Key Facility for Crop Gene Resources and Genetic Improvement, Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing, China, ³ Department of Biology, Saint Mary's University, Halifax, NS, Canada

OPEN ACCESS

Edited by:

Soren K. Rasmussen,
University of Copenhagen, Denmark

Reviewed by:

Meixue Zhou,
University of Tasmania, Australia
Elisabetta Mazzucotelli,
Agricultural Research Council (CRA),
Italy

*Correspondence:

Cheng Chang
changtgw@126.com

† These authors have contributed
equally to this work.

Specialty section:

This article was submitted to
Crop Science and Horticulture,
a section of the journal
Frontiers in Plant Science

Received: 07 October 2016

Accepted: 01 December 2016

Published: 20 December 2016

Citation:

Hu M-J, Zhang H-P, Liu K, Cao J-J,
Wang S-X, Jiang H, Wu Z-Y, Lu J,
Zhu XF, Xia X-C, Sun G-L, Ma C-X
and Chang C (2016) Cloning
and Characterization of *TaTGW-7A*
Gene Associated with Grain Weight
in Wheat via SLAF-seq-BSA.
Front. Plant Sci. 7:1902.
doi: 10.3389/fpls.2016.01902

Thousand-grain weight (TGW) of wheat (*Triticum aestivum* L.) contributes significantly to grain yield. In the present study, a candidate gene associated with TGW was identified through specific-locus amplified fragment sequencing (SLAF-seq) of DNA bulks of recombinant inbred lines (RIL) derived from the cross between Jing 411 and Hongmangchun 21. The gene was located on chromosome 7A, designated as *TaTGW-7A* with a complete genome sequence and an open reading frame (ORF). A single nucleotide polymorphism (SNP) was present in the first exon between two alleles at *TaTGW-7A* locus, resulting in a Val to Ala substitution, corresponding to a change from higher to lower TGW. Cleaved amplified polymorphic sequence (CAPS) (*TGW7A*) and InDel (*TG9*) markers were developed to discriminate the two alleles *TaTGW-7Aa* and *TaTGW-7Ab* for higher and lower TGW, respectively. A major QTL co-segregating with *TaTGW-7A* explained 21.7–27.1% of phenotypic variance for TGW in the RIL population across five environments. The association of *TaTGW-7A* with TGW was further validated in a natural population and Chinese mini-core collections. Quantitative real-time PCR revealed higher transcript levels of *TaTGW-7Aa* than those of *TaTGW-7Ab* during grain development. High frequencies of the superior allele *TaTGW-7Aa* for higher TGW in Chinese mini-core collections (65.0%) and 501 wheat varieties (86.0%) indicated a strong and positive selection of this allele in wheat breeding. The molecular markers *TGW7A* and *TG9* can be used for improvement of TGW in breeding programs.

Keywords: common wheat, gene expression, gene-specific marker, IAA, SLAF-seq

Abbreviations: ANOVA, analyses of variance; CAPS, cleaved amplified polymorphic sequence; DPA, days post anthesis; GL, grain length; GLM, general linear model; GW, grain width; IGPS, indole-3-glycerol-phosphate synthase; ORF, open reading frame; QTL, quantitative trait locus/loci; RIL, recombinant inbred lines; RT-PCR, real-time quantitative PCR; SLAF-seq, specific-locus amplified fragment sequencing; SNP, single nucleotide polymorphism; SSR, simple sequence repeat; Super-BSA, combination of SLAF-seq and BSA analysis; TGW, thousand-grain weight.

INTRODUCTION

Thousand-grain weight is the most stable component of wheat grain yield (Giura and Saulescu, 1996). Improvement of TGW is one of the main targets in wheat breeding. Previous studies showed that wheat TGW is controlled by polygenes (Giura and Saulescu, 1996; Huang et al., 2003; Kuchel et al., 2007). Many QTL for TGW have been identified on almost all wheat chromosomes (Huang et al., 2003, 2006; Quarrie et al., 2005; Breseghello and Sorrells, 2006; Sun et al., 2009; Lopes et al., 2013; Simmonds et al., 2014; Su et al., 2016). However, only a few major QTL have been cloned due to large genome size and limited genome sequence information in wheat. The homology-based cloning of wheat genes is a common approach based on the collinearity in cereal crops (Choi et al., 2004; Gupta et al., 2008; Hu et al., 2016). *TaGW2*, encoding an E3 ubiquitin ligase and associated with GW and TGW, was cloned based on the homologous rice gene *GW2* (Su et al., 2011). Similarly, *TaCwi-A1* associated with TGW (Ma et al., 2012; Jiang et al., 2015), *TaGASR7-A1* for GL and yield (Dong et al., 2014), *TaGS-D1* for GL and TGW (Zhang et al., 2014), and *TaTGW6* encoding an indole-3-acetic acid (IAA)-glucose hydrolase (Hu et al., 2016) have been cloned based on rice gene sequences.

Homology-based cloning is an efficient approach for isolation of wheat genes following known rice gene sequences. But obviously, it is not feasible for identification of unknown functional genes. The genetic network controlling TGW is very complex, and the molecular regulatory mechanisms may differ among diverse germplasm resources (Browne et al., 2006). Identification of more TGW genes will not only accelerate multi-gene pyramiding in wheat breeding, but it is also important for further elucidating the molecular mechanism of yield formation.

Specific-locus amplified fragment sequencing (SLAF-seq) is a high-throughput strategy for a large scale of SNP discovery and genotyping based on next generation sequencing (Sun et al., 2013). Bulk segregant analysis (BSA) is an efficient approach to detect major QTL by genotyping two DNA bulks from artificial population lines with contrasting phenotypes (Quarrie et al., 1999; Pomraning et al., 2011). Candidate genes for targeting traits can be identified rapidly using super-BSA. This approach was successfully used in detection of QTL, fine mapping of candidate genes and development of molecular markers in high plants (Trick et al., 2012; Qin et al., 2015; Xu et al., 2015a,b). In our previous studies, several genes for TGW, including *TaCKOX6a02* (Lu et al., 2015), *Tackx4* (Chang et al., 2015), and *TaTGW6* (Hu et al., 2016), have been identified using RIL from a cross between Jing 411 and Hongmangchun 21. The two parents, Jing 411 and Hongmangchun 21, showed large differences in grain size, TGW, and grain filling rate. In this study, therefore, the objectives were to (i) use super-BSA and bio-information analysis to identify candidate genes for TGW, (ii) and develop functional markers for marker-assisted selection in wheat breeding.

MATERIALS AND METHODS

Plant Materials and Field Trials

Jing 411 is a wheat variety with high TGW (47.6 g on average over five cropping seasons); while Hongmangchun 21 is a wheat landrace with low TGW (19.7 g on average over five cropping seasons). An F₈ RIL population (Pop 1, 150 lines kindly provided by Prof. Shihe Xiao at Chinese Academy of Agricultural Sciences) derived from the Jing 411 × Hongmangchun 21 cross was used for candidate gene mapping (**Supplementary Table S1**). Two hundred and forty-four wheat varieties, including 17 Chinese landraces and 227 modern varieties (Pop 2, **Supplementary Table S2**), and 257 accessions from Chinese wheat mini-core collections (Pop 3, kindly provided by Prof. Jizeng Jia at Chinese Academy of Agricultural Sciences) were used to validate the gene-specific marker (Hao et al., 2008) (**Supplementary Table S3**). Five hundred and one wheat varieties, including historical varieties, modern varieties, advanced lines and landraces, collected from seven wheat regions of China (Pop 4, **Supplementary Table S4**), and 180 varieties with detailed information of origins among Chinese wheat mini-core collections mentioned above were used for characterizing allelic distributions of target gene. The Pop 1 was grown at the Experimental Station of Anhui Agricultural University (Hefei, Anhui, China: 31°58'N, 117°240'E) in 2007–2008, 2010–2011, 2011–2012, 2013–2014, and 2014–2015 cropping seasons; Pop 2 was planted in 2011–2012, 2012–2013, and 2013–2014 cropping seasons, and Pop 3 in 2014–2015 cropping season. Field trials were conducted in randomized complete blocks with two replications, with double 4-m rows spaced 25 cm apart. Field management followed local agricultural practice. The flowering time was recorded for each line.

Grain Size and Thousand Grain Weight Assay

The TGW was evaluated by weighing 1000 grains in triplicate each plot and the averaged data were used for subsequent analysis. Three hundred grains were measured to get average GL and GW with the SC-G wheat grain appearance quality image analysis system developed by the Hangzhou WSeen Detection Technology Co. Ltd., Hangzhou, China (Yin et al., 2015).

DNA and RNA Extraction

Genomic DNA was extracted from seeds following Kang et al. (1998). RNA was isolated from seeds of six low- and six high-TGW wheat varieties using a MiniBEST Plant RNA Extraction Kit (Takara). Immature grains collected at 10 days, 20 days, and 30 DPA, and mature seeds were used in these analyses. The cDNA was synthesized using Promega GoScript (Promega, Beijing, China) according to the manufacturer's instructions.

SLAF Library Construction and High-Throughput Sequencing

Two DNA bulks of RILs used for BSA were formed from 30 high- and 30 low-TGW lines, respectively. The SLAF

library was constructed following previous studies (Trick et al., 2012; Sun et al., 2013; Xu et al., 2015b) with minor modifications. In brief, we used *Triticum aestivum* as reference genome for a preliminary SLAF experiment to determine conditions and appropriate restriction enzymes. Genomic DNA of two bulks and parents were incubated with *RsaI* (New England Biolabs, Nanjing, China). A single-nucleotide (A) overhang was added to the digested fragments with a Klenow Fragment (3'→5' exo-) (New England Biolabs, Nanjing, China) and dATP at 37°C, then PAGE-purified duplex tag-labeled sequencing adapters (Life Technologies, Gaithersburg, MD, USA) were ligated to the A-tailed DNA with T4 DNA ligase. PCR was performed using diluted restriction-ligation samples, dNTP, Q5® High-Fidelity DNA Polymerase, and PAGE-purified PCR primers 5'-AATGATACGGCGACCACCGA-3' and 5'-CAAGCAGAAGACGGCATACG-3' (Life Technologies, Gaithersburg, MD, USA).

The PCR products were purified using Agencourt AMPure XP beads (Beckman Coulter, High Wycombe, UK) and pooled. The pooled sample was separated by electrophoresis on a 2.0% agarose gel. Fragments with 464–484 bp were excised, and purified using a QIAquick Gel Extraction Kit (Qiagen, Beijing, China). The gel-purified products were sequenced using the Illumina HiSeq™ 2500 platform (Illumina Inc., San Diego, CA, USA) according to the manufacturer's recommendations by the Biomarker Technologies Co, LTD (Beijing, China¹).

SLAF Analysis and Gene Prediction

The SLAF-seq data were processed as described by Sun et al. (2013). In brief, low-quality reads with a quality score <Q30 (means a sequencing quality score of 30, indicating a 0.1% chance of error, and 99.9% confidence) were filtered out. In the present study, SLAFs from two parents with sequencing depth <5X were removed. Additionally, the polymorphic SLAFs with more than three SNP were also filtered out. The rest polymorphic SLAF tags were clustered and mapped onto the wheat reference genome based on similarity using BLAT and short oligonucleotide alignment program (SOAP) (Kent, 2002; Li et al., 2008). Sequences with more than 90% similarity were grouped into one SLAF locus. The sequence error rate was estimated using *Arabidopsis thaliana* genome sequence data as the control (Davey et al., 2013). Finally, the SLAF-tag with same genotype as female parent (Jing 411) or the male (Hongmangchun 21) was considered as a SLAF marker and used for subsequent association analysis.

Association analysis was conducted using the SNP_index method as described by Hill et al. (2013). The $\Delta(\text{SNP_index})$ was calculated according to Hill et al. (2013) and Xu et al. (2015c), $\Delta(\text{SNP_index})$ value was calculated as follows: $\text{SNP_index}(ab) = \text{Mab}/(\text{Pab}+\text{Mab})$, $\text{SNP_index}(aa) = \text{Maa}/(\text{Paa}+\text{Maa})$, and $\Delta(\text{SNP_index}) = \text{SNP_index}(aa) - \text{SNP_index}(ab)$; where Maa and Mab mean the depth of aa and ab populations derived from M, respectively, Paa and Pab mean the depth of aa and ab populations derived from P, respectively. The closer

marker is associated with phenotype while the closer or equal $\Delta(\text{SNP_index})$ is to 1. The threshold value of associated SLAFs was determined when the value $\geq \Delta(\text{SNP_index})$ of 99.99% of SLAFs. In this study, the threshold was 1 based on the analysis of **Supplementary Table S5** (Hill et al., 2013; Xu et al., 2015c). Association region with three or more consecutive polymorphic SLAF tags with $\Delta(\text{SNP_index}) = 1$ was identified as a hot region related to TGW on wheat chromosome (Zhang et al., 2015).

The sequences of tags highly significantly associated with target traits in hot regions were extended 50 kb forward and reverse according to the sequences' information of wheat BAC library of the Biomarker Technologies Co, LTD (Beijing, China¹) and international GenBank including NCBI², Wheat-urgi³, and Cerealsdb⁴. And then, these extended sequences were used for functional annotation and gene/transcript set/pathway enrichment analyses, based on the wheat genome information (Kanehisa, 2002; Gene Ontology Consortium, 2004; Subramanian et al., 2005; Brenchley et al., 2012; Zhang et al., 2015, 2016). Meanwhile, these SLAFs were also used for BLAST³ to obtain the homology sequences in wheat.

Candidate Gene Cloning and Development of CAPS and InDel Markers

Eight sets of primer pairs were designed using Primer Premier 5 software based on the sequence of 7AS_4248784 and its transcript Traes_7AS_378A12AA9.1 on 7AS, which were obtained by BLAST and functional annotation and gene/transcript set/pathway enrichment analyses of SLAF65386 at a hot region on chromosome 7A. These primer pairs were used for amplifying genomic DNA and cDNA sequences of candidate gene, as well as developing InDel and CAPS markers (**Supplementary Table S6**).

The PCR reactions (total volume, 10 μL) included 0.25 μM each primer, 0.25 mM dNTP, 0.5 unit LA *Taq*, 1 μL 10 \times LA PCR buffer (Takara, Dalian, China) and 100 ng of genomic DNA. The PCR conditions were 95°C for 5 min, followed by 36 cycles of 95°C for 45 s, annealing (55–62 °C) for 50 s, and extension at 72°C for 1 min, with a final extension at 72°C for 12 min. Amplified PCR fragments were separated on a 6% denaturing polyacrylamide gel or a 1.5% agarose gel. For CAPS markers, PCR products were digested with *BsmAI* (NEB), *BstNI*, *MluCI*, or *AluI* according to the manufacturer's directions, labelled with fluorochrome, and then, detected on a 2.0% agarose gel.

For candidate gene cloning, the targeting PCR fragments were recovered and cloned into the pEASY-T5 simple vector and transformed into T1 competent cells (TransGen Biotech, Beijing, China). Targeting gene sequencing was conducted in an Applied Biosystems 3500 genetic analyzer (Applied Biosystems, Shanghai, China) according to the manufacturer's instructions. Sequence alignments among different wheat varieties were performed using

¹<http://www.biomarker.com.cn/>

²<https://www.ncbi.nlm.nih.gov>

³<http://wheat-urgi.versailles.inra.fr/>

⁴www.cerealsdb.uk.net

DNAMAN 6.0⁵. The core elements of promoter were predicted by the TSSP program⁶.

Validation of Gene Effect by QTL and Association Analysis

QTL mapping for TGW was performed with SSR markers in Pop 1 (Somers et al., 2004⁷). The polymorphic markers between two parents and bulks were used for genotyping the entire Pop 1. The effect of candidate gene on grain weight was estimated by QTL analysis using composite interval mapping (QTL IciMapping v.4.0 software⁸). QTL was declared as significant in a LOD threshold of 3.0. As for *TaTGW-7A* linkage analysis with 61 SSR markers, two gene-specific markers and four CAPS markers on 7A, the LOD value is 5.0.

CAPS (*TGW7A*) and InDel (*TG9*) markers derived from the candidate gene were also validated by association analysis in Pop 2 and Pop 3.

Fluorescent Real-Time Quantitative RT-PCR

Analysis of *TaTGW-7A* transcript levels was performed using 2 × probe qPCR Mix (with ROX) (SinoGene, Beijing, China). The primer sets for *TaTGW-7A* (FQ), *Actin* (Actin, control) and the probe sequences were shown in **Supplementary Table S6**. Each sample was analyzed in triplicate. The relative qualification $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001) was used to calculate *TaTGW-7A* transcript levels with *Actin* as the endogenous control and Wangshuibai as a reference sample

IAA Purification and Quantification

In this study, IAA was quantified in immature seeds (1 g) at 20 and 30 DPA from six low- and six high-TGW varieties with three repeats. These seeds were frozen immediately in liquid nitrogen, homogenized in 80% methanol with a mortar and pestle, and then the mixture was stored at -20°C . To analyze IAA content, the mixture was stirred overnight at 4°C , and impurities were removed by centrifugation at $10,000 \times g$ for

⁵<http://www.lynnon.com>

⁶<http://www.softberry.com>

⁷<http://wheat.pw.usda.gov>

⁸<http://www.isbreeding.net>

20 min. The supernatant was passed through a 0.22- μm millipore filter and then pre-equilibrated with 5 mL of 100% methanol and 70% methanol, respectively, using a C₁₈ Waters Sep Pak cartridge (Waters, Milford, MA, USA) (Hu et al., 2016). The filtrate was collected and analyzed by high performance liquid chromatography (HPLC).

The HPLC was equipped with an Eclipse XDB-C₁₈ column (250 mm × 4.6 mm, 5 μm , Agilent, Palo Alto, CA, USA). The mobile phase consisted of solution A (methanol) and solution B (acetic acid solution, pH 3.6). The elution profile was for 38 to 42.2% A in B (linear gradient) from 0 to 9 min, for 42.2 to 38% A in B (linear gradient) from 9 to 10 min, and for 38% A in B (linear gradient) from 10 to 25 min. The flow rate was 1 mL min⁻¹ and the UV-absorbance detector was set to 250 nm. An external standard method was used to calculate the IAA concentration according to the manufacturer's instructions⁹.

Statistical Analysis

The significant difference of effects between the two alleles of *TaTGW-7A* on TGW was analyzed by ANOVA, and estimation of candidate gene effect on phenotypic variation was performed with GLM. These statistical analyses were all finished via SPSS 20 statistical software¹⁰.

RESULTS

Analysis of BSA-SLAF-seq Data

In total, 132,530 SLAF tags were acquired, and 11,571 were polymorphic between two parents and between DNA bulks. Nine hundred and forty SLAF tags were significantly correlated with TGW by SNP_index analysis, and 938 were screened and positioned onto wheat chromosomes (**Supplementary Tables S7 and S8**). In order to detect the hot regions carrying candidate target genes, these SLAF tags were also analyzed by association-region analysis. The results indicated that twenty-five polymorphic SLAF tags closely related to TGW were identified on six hot regions, and located on chromosomes 2A, 3A, 4B, 6A, 7A, and 7B, respectively, through BLAT analysis (**Table 1**).

⁹<http://www.agilent.com/home>

¹⁰<http://www-01.ibm.com/software/analytics/spss/>

TABLE 1 | Twenty-five polymorphic SLAFs in six hot-regions associated with TGW.

| Chromosome ID | Start | End | Genetic distance/cM | Associated Marker number | Transcripts |
|---------------|--------|--------|---------------------|--------------------------|-------------|
| 2A | 68.14 | 68.14 | 0 | 3 | 0 |
| 3A | 60.18 | 60.18 | 0 | 3 | 0 |
| 4B | 65.7 | 65.7 | 0 | 3 | 0 |
| 6A | 60.96 | 60.96 | 0 | 3 | 0 |
| 7A1 | 93.26 | 93.26 | 0 | 3 | 1 |
| 7A2 | 98.92 | 98.92 | 0 | 3 | 0 |
| 7B1 | 161.06 | 163.78 | 2.72 | 3 | 0 |
| 7B2 | 168.73 | 170.36 | 1.63 | 4 | 1 |
| Total | | | | 25 | 2 |

Particularly, two hot regions on 7A (six SLAF tags) and 7B (seven SLAF tags) carried more than half of TGW-associated SLAF tags (Table 1). Two full-length ORFs were also found in the two hot regions by alignment of SLAF tags sequences with wheat genome sequences. Therefore, the 13 SLAF tags on 7A and 7B were selected for subsequent analysis.

Of the 13 SLAF tags, SLAF65386 in the hot region of 7AS was 100% identical to the sequence 7AS_4248784 containing a full-length ORF, Traes_7AS_378A12AA9.1, through BLAST analysis in the Genetic and Genomic Information System Database¹¹ (Supplementary Figures S1 and S2). The gene corresponding to the transcript Traes_7AS_378A12AA9.1 was designated as *TaTGW-7A*. Simultaneously, the sequence of SLAF65386 was extended 50 kb forward and reverse, and used for functional annotation and gene/transcript set/pathway enrichment analyses. The results indicated the transcript of *TaTGW-7A* (Traes_7AS_378A12AA9.1) could be annotated in the COG database, GO database, and KEGG database, respectively (Supplementary Table S9). The gene was predicted to encode a protein with a TIM barrel fold the same as IGPS. Therefore, *TaTGW-7A* could be regarded as a candidate gene for TGW on chromosome 7AS. In the present study, the putative gene associated with TGW on 7B was not obtained (data not shown).

Genome-Wide QTL Location for TGW Using SSR Markers

In this study, several stable QTL for TGW were mapped on chromosomes 1B, 3A, 3D, 4B, 4D, 5A, 6A, 6B, 7A, and 7B (Supplementary Figure S3). These loci could explain 5.7–28.9% phenotypic variations of TGW in different environments (Supplementary Table S10). Moreover, the loci on 3A, 4B, 6A, 7A, and 7B could be detected both by SNP_index analysis and

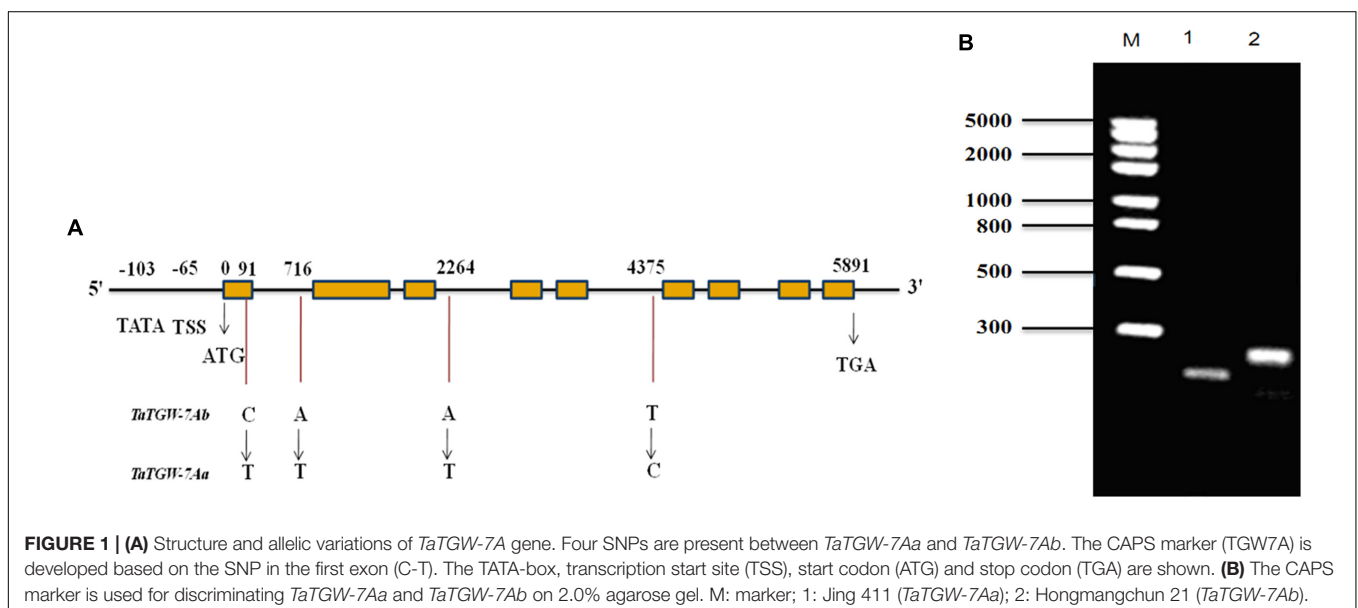
QTL mapping. Thereinto, the loci on 7A and 7B generally had higher LOD values than other loci in different environments. The result of QTL mapping was generally consistent with hot regions analysis for TGW through comparing the detected loci both by super-BSA and linkage analysis.

Cloning and Characterization of *TaTGW-7A*

The primer sets Q29, Q35, Q37, and SR were designed based on the sequences of 7AS_4248784 and Traes_7AS_378A12AA9.1 to amplify the full-length gDNA and cDNA sequences of *TaTGW-7A* from Jing 411 and Hongmangchun 21, respectively. The amplified fragments by primer sets 29P and 31P were used for assembling the full-length gDNA of *TaTGW-7A* (Supplementary Table S6). All amplified products were recovered, cloned, sequenced and assembled. Sequence analysis revealed that the Jing 411 allele had a full-length gDNA of 5891 bp, containing an ORF of 2241 bp with nine exons and eight introns (Figure 1A). The gene encodes a peptide of 746 amino acids with a predicted molecular weight of 37.7 kDa. Four nucleotide substitutions were present between Jing 411 and Hongmangchun 21 alleles (Supplementary Figure S4). Only one SNP in the first exon of *TaTGW-7A* was found consistently between the high- and low-TGW genotypes (Supplementary Table S11). This mutation resulted in a change from an alanine codon (GCC) in the low-TGW to a valine codon (GTC) in the high-TGW varieties. The TATA box was located at –103 position, and the transcription start site (TSS) was located at –65 position from the initiation codon (Figure 1A).

A signal peptide with a cleavage site located at positions 28–29 was predicted by the SignalP program (Supplementary Figure S5). Analysis of domain architectures in NCBI indicated that the functional domain of *TaTGW-7A* was an N-terminal domain of sigma 54-dependent transcriptional activators, the same as

¹¹<http://wheat-urgi.versailles.inra.fr/>



IGPS¹² involved in the tryptophan biosynthetic pathway, which can catalyze the conversion of 1-[(2-carboxyphenyl) amino]-1-deoxyribulose 5-phosphate to indole-3-glycerol phosphate in IAA biosynthesis process (Supplementary Figures S6 and S7). The complete mRNA sequence of *TaTGW-7A* has been submitted to NCBI (KT582299).

Development and Validation of *TaTGW-7A* Functional Marker

A CAPS marker (*TGW7A*) was developed based on the SNP in the first exon between two alleles of *TaTGW-7A*. A 250-bp fragment was amplified from both Jing 411 and Hongmangchun 21 using the primer set MQ (Supplementary Table S6). After digestion with *BsmAI* at the SNP site (NEB, Nanjing, China), two fragments of 250 bp vs. 196 bp clearly distinguished between the Jing 411 and Hongmangchun 21 alleles, designated as *TaTGW-7Aa* and *TaTGW-7Ab*, respectively (Figure 1B). We further sequenced approximately 3 kb downstream of 7AS_4248784, and detected a 5-bp insertion-deletion (InDel) polymorphism between *TaTGW-7Aa* and *TaTGW-7Ab*. An InDel marker (*TG9*) was also developed, co-segregating with *TGW7A* in the RIL population. In addition, other four CAPS markers were designed based on polymorphic SLAFs on chromosome 7A (Supplementary Table S6), designated as *SLAF49035*, *SLAF28300*, *SLAF133263*, and *SLAF6258*, respectively.

Three populations (Pop 1, Pop 2, and Pop 3) were used to validate the markers *TGW7A* and *TG9*, which were significantly associated with TGW in all environments, and the markers also showed significant association with GL and GW in Pop 1 and Pop 3 (Table 2).

¹²<http://www.ncbi.nlm.nih.gov/>

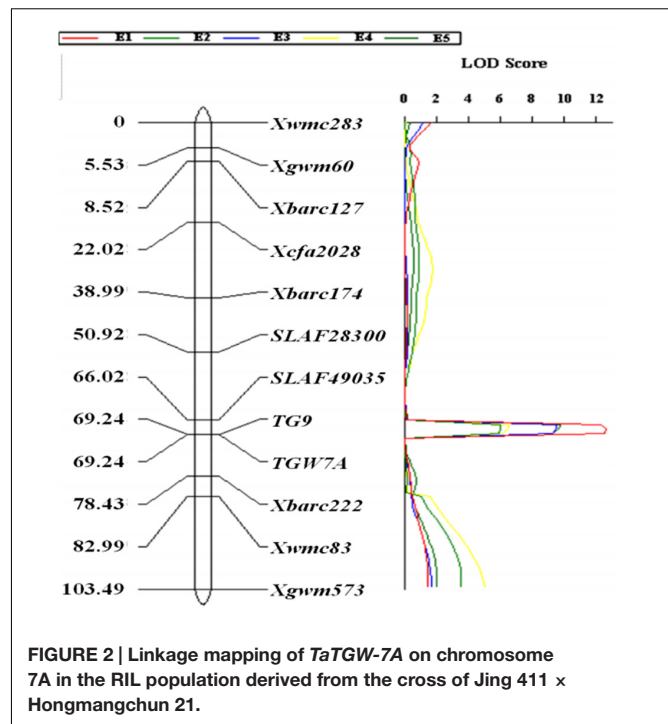


FIGURE 2 | Linkage mapping of *TaTGW-7A* on chromosome 7A in the RIL population derived from the cross of Jing 411 x Hongmangchun 21.

Validation of *TaTGW-7A* by Linkage Analysis

The gene-specific markers *TGW7A* and *TG9*, two CAPS (*SLAF49035* and *SLAF28300*) as well as eight SSR markers were integrated into a linkage map for chromosome 7A, spanning a genetic distance of 103.49 cM (Figure 2). QTL analysis

TABLE 2 | Association analysis of grain traits with *TaTGW-7A* genotypes in Pop 1 (RIL population), Pop 2 (natural population), and Pop 3 (Chinese mini-core collections).

| Year | Grain traits ^a | <i>TaTGW-7Aa</i> (Mean ± SD) | <i>TaTGW-7Ab</i> (Mean ± SD) | <i>F</i> | PVE (%) |
|--|---------------------------|---------------------------------|---------------------------------|----------|---------|
| Chinese wheat mini-core collections | | | | | |
| 2015 | TGW(g) | 34.14 ± 9.27 | 27.35 ± 6.99 | 36.92** | 12.3 |
| | GL(mm) | 6.50 ± 0.57 | 5.94 ± 0.46 | 64.62** | 19.9 |
| | GW(mm) | 3.11 ± 0.30 | 2.89 ± 0.28 | 32.31** | 10.9 |
| Natural population | | | | | |
| 2012 | TGW(g) | 38.77 ± 4.50 | 34.52 ± 7.17 | 21.94** | 8.9 |
| 2013 | TGW(g) | 42.15 ± 4.61 | 37.08 ± 7.65 | 29.12** | 12.4 |
| 2014 | TGW(g) | 44.20 ± 4.18 | 40.38 ± 6.84 | 20.22** | 7.6 |
| RIL population | | | | | |
| 2008 | TGW(g) | 41.24 ± 10.73 | 30.14 ± 5.98 | 52.56** | 26.2 |
| 2011 | TGW(g) | 38.16 ± 9.05 | 29.91 ± 4.95 | 41.00** | 21.7 |
| 2012 | TGW(g) | 38.83 ± 9.79 | 29.92 ± 4.14 | 43.77** | 22.8 |
| 2014 | TGW(g) | 38.42 ± 11.04 | 28.37 ± 4.38 | 44.26** | 23.0 |
| 2015 | TGW(g) | 34.71 ± 8.10 | 25.84 ± 5.40 | 54.88** | 27.1 |
| | GL(mm) | 6.51 ± 0.73 | 5.78 ± 0.30 | 51.86** | 25.9 |
| | GW(mm) | 3.18 ± 0.33 | 2.89 ± 0.15 | 40.93** | 21.7 |

^aTGW, thousand-grain weight; GL, grain length; GW, grain width; PVE, phenotypic variance explained by *TaTGW-7A*. Significant difference of TGW between the two alleles at the levels of 0.05 and 0.01 was marked with * and **, based on *F*-test. The two markers, *TGW7A* and *TG9* have same results to discriminate alleles of *TaTGW-7A*.

revealed that a stable QTL at *TaTGW-7A* locus was located in the interval of *SLAF49035* and *TGW7A/TG9* on chromosome 7AS, and explained 21.7–27.1% of the TGW variation in the RIL population across five environments (Table 2). In the genome-wide QTL mapping using SSR markers, a main locus was also detected between the *Xbarc174* and *Xbarc222*, which covered the interval of *SLAF49035* and *TGW7A/TG9* (Figure 2; Supplementary Figure S3). In addition, the *TaTGW-7A* locus was located on C-7AS-8 (0.45) according to the physical map of common wheat, as shown in Supplementary Figure S8 (Qi et al., 2004; Sourdille et al., 2004).

TaTGW-7A Expression

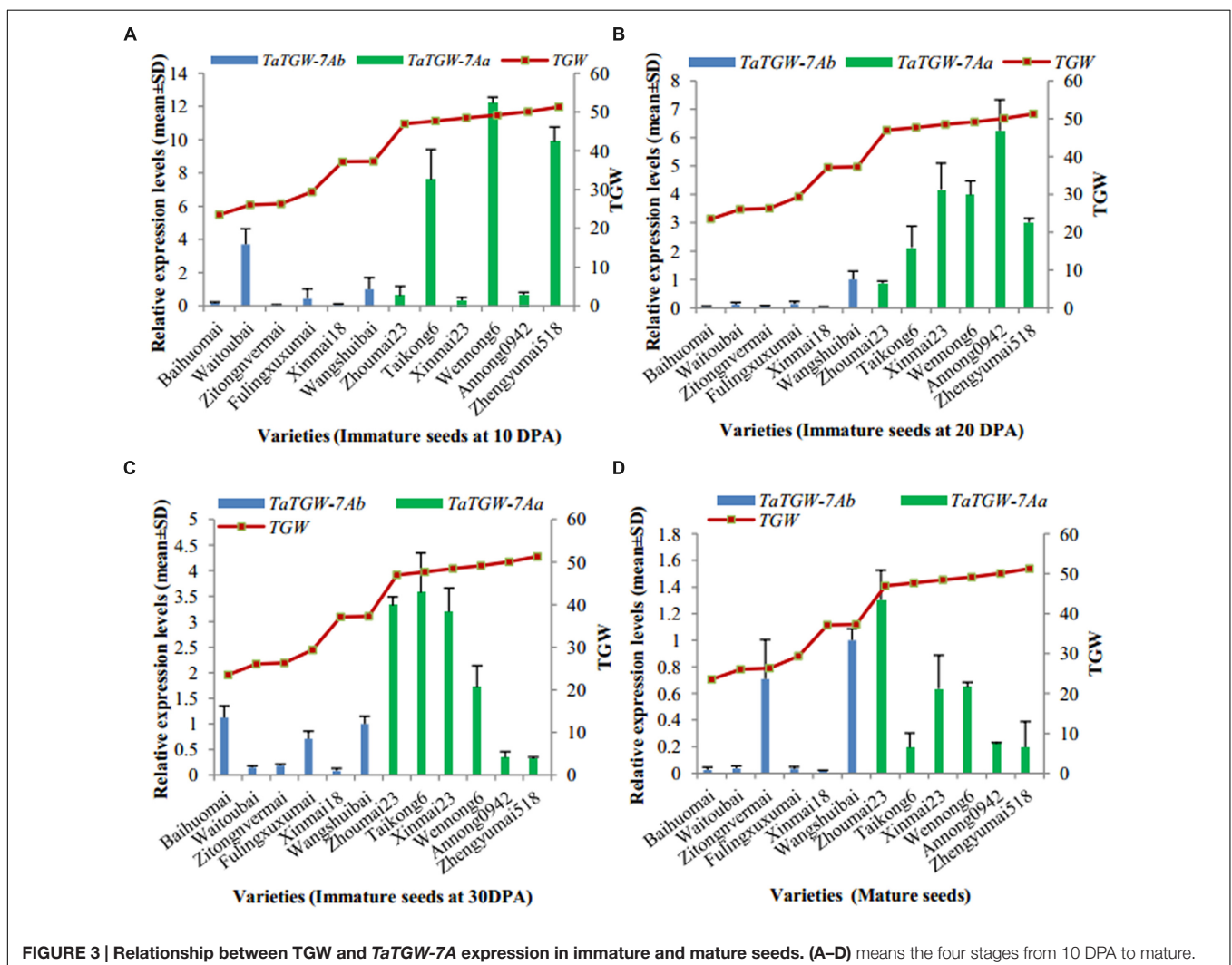
Immature (10, 20, and 30 DPA) and mature seeds were used to analyze the expression level of *TaTGW-7A* (Supplementary Table S6). Generally, the relative transcript level of *TaTGW-7A* in immature seeds at 20 and 30 DPA was associated with TGW. The average transcript level of *TaTGW-7Aa* for high TGW was generally higher than those of *TaTGW-7Ab* for low TGW (Figure 3).

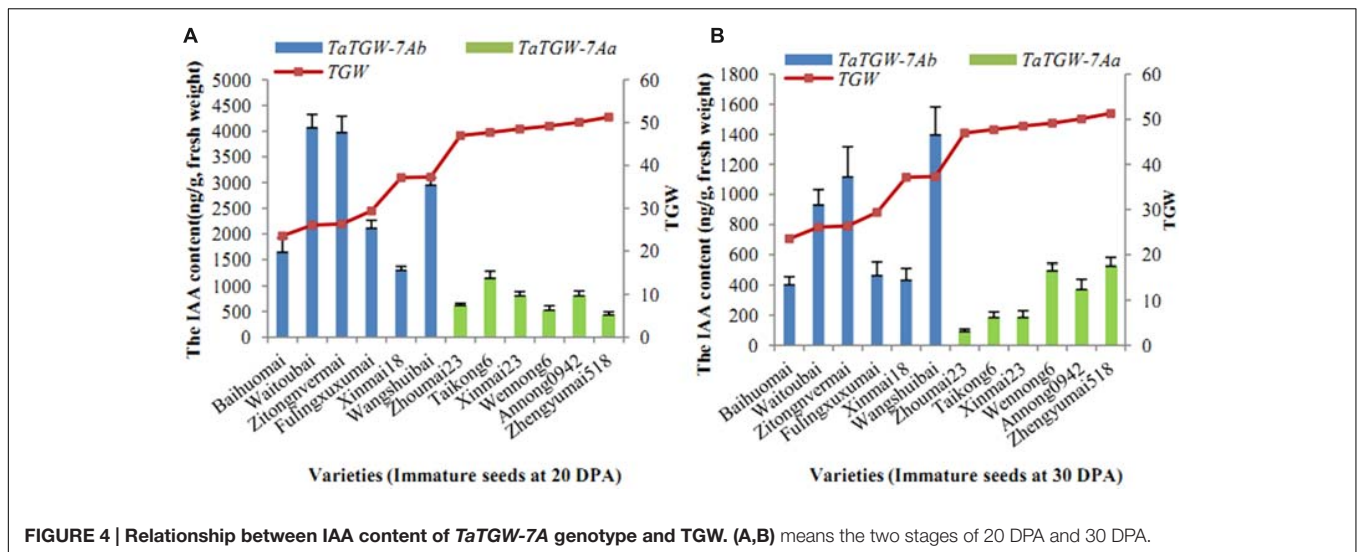
Analysis of IAA Content

Six varieties with low and six with high TGW were sampled at 20, 30 DPA and IAA contents were determined. The results showed that IAA content was related to TGW. The average level of IAA in wheat varieties with *TaTGW-7Ab* was generally higher than those in *TaTGW-7Aa* genotypes at 20 and 30 DPA (Figure 4).

Frequency Distributions of *TaTGW-7A* Alleles

A high frequency of *TaTGW-7Aa* was found in Chinese mini-core collections (Pop3, 59%), and in most of Chinese wheat production areas, including the Low and Middle Yangtze River Valley winter wheat region (62%), Southern winter wheat region (67%), Northeastern spring wheat region (100%), Northern spring wheat region (89%), Northwestern spring wheat region (62%), Qinghai-Tibet spring-winter wheat region (86%), and Xinjiang winter-spring wheat region (100%) (Supplementary Figure S9; Supplementary Tables S12 and S13). In contrast, the *TaTGW-7Ab* allele showed higher frequencies in Northern winter wheat production area (61%), and Southern winter wheat





production area (51%). In Pop 4, *TaTGW-7Aa* was a predominant allele distributed in most of wheat areas, including Northern winter wheat region (89%), Yellow and Huai River valley winter wheat region (90%), Low and middle Yangtze River valley winter wheat region (88%), Southwestern winter wheat region (74%), Northeastern spring wheat region (77%), Northwestern spring wheat region (81%) and Xinjiang winter-spring wheat region (67%) (Supplementary Table S13).

DISCUSSION

Application of Super-BSA and Bio-information to Isolate Genes in Common Wheat

Wheat is an allohexaploid species with extremely large genomes and a detailed physical map has not been available for most of wheat chromosomes. Moreover, wheat genome contains a high proportion (>80%) of repeated sequences (Eilam et al., 2007; Wicker et al., 2011), making map-based cloning difficult and time consuming. The SLAF-seq for marker development is highly efficient, inexpensive, and applicable for analyses of large populations; hence, it has been proposed as an effective method for marker-assisted breeding (Sun et al., 2013). SLAF-seq in combination with BSA is a simple and efficient approach to detect candidate genes and develop gene-specific markers (Trick et al., 2012; Chen et al., 2013; Qin et al., 2015; Xu et al., 2015c). However, this approach has not been frequently used for marker development and isolation of functional genes in common wheat.

In this study, we identified one candidate gene *TaTGW-7A* associated with TGW on the hot region of 7AS through super-BSA analysis. In addition, eight genes were predicted by functional annotation and gene/transcript set/pathway enrichment analyses of six SLAF tags on the hot region of 7A (Supplementary Table S9); *TaTGW-7A* was more significantly associated with TGW by linkage analysis, compared with the other SLAF tags on 7A (Figure 2). Moreover, the mapping

analysis of *TaTGW-7A* was generally consistent with raw QTL mapping using SSR markers (Figure 2; Supplementary Figure S3). Thus, it is worthy of further focusing the SLAF65386 in the hot region of 7AS, corresponding to Traes_7AS_378A12AA9.1 with full-length ORF. In addition, identification of *TaTGW-7A* can be considered as an effective and accurate application of super-BSA in development of molecular markers and isolation of candidate genes from wheat. However, for polyploid plants such as common wheat, it is very difficult to detect minor genes by BSA-SLAF-seq, and SLAF-seq cannot cover the whole genome for gene discovery. For example, several QTL for TGW were detected on 1B, 3A, 3D, 4B, 4D, 5A, 6A, 6B, 7A, and 7B, respectively; however, we only identified a candidate gene, *TaTGW-7A*, on 7A by super-BSA and bio-information analysis. Another hot region associated with TGW on 7B probably comprised candidate genes; however, we could not identify them via super-BSA. This is possibly because these SLAF tags do not exist in or have a long distance from the candidate gene and no enough information of genome sequence for analysis in common wheat, compared with rice. In addition, SLAF tags on 2A, 3A, 4B, and 6A also showed close association with TGW, which are also worthy of further developing CAPS markers. Therefore, other methods should also be used for identifying functional genes in wheat. In our previous studies, several genes including *TaCKOX6a02* on 3DS (Lu et al., 2015), *Tackx4* on 3AL (Chang et al., 2015), and *TaTGW6* on 4AL (Hu et al., 2016) associated with TGW were identified based on the homology analysis of rice genes.

Characterization and Putative Molecular Mechanism of *TaTGW-7A*

The *TaTGW-7A* gene contained nine exons and eight introns. Four SNPs were detected between two alleles, *TaTGW-7Aa* and *TaTGW-7Ab*, three of which were in introns and one in exon. The SNP in the first exon of *TaTGW-7A* resulted in an amino acid change from Val to Ala at residue 31, corresponding to a change from a high-TGW to a low-TGW phenotype. This suggested the point mutation possibly

affected protein stability because of changes in the secondary structure¹³. Two alpha helices and one beta sheet occurred at the amino acid mutation site of *TaTGW-7Ab*, compared with *TaTGW-7Aa*. The changes in secondary structure possibly influence the function of *TaTGW-7A* (Cheng et al., 2006). The deduced amino acid sequences showed the presence of highly conserved TIM-br_sig_trns and UPF0261 functional domains, forming a TIM barrel fold structure, the same as IGPS¹⁴, which was a key enzyme involved in tryptophan synthesis and ultimately affected IAA biosynthesis. The domain was found as an N-terminal domain of sigma 54-dependent transcriptional activators (enhancer-binding proteins), suggesting a potential role in signal recognition/reception and signal transduction (Bush and Dixon, 2012).

To explore the potential molecular mechanism of *TaTGW-7A*, we analyzed transcript expression of *TaTGW-7A* and corresponding IAA contents in wheat varieties. The transcript level of *TaTGW-7A* was significantly correlated with TGW in immature seeds at 20 and 30 DPA, and the IAA content was negatively correlated with TGW. Compared with *TaTGW-7Ab* genotype, *TaTGW-7Aa* corresponded to higher TGW and lower IAA content. It was assumed that *TaTGW-7A* was required for synthetic regulation of IAA. The higher expression of *TaTGW-7Aa* possibly reduced the biosynthesis of tryptophan and IAA.

Ishimaru et al. (2013) reported that *TGW6*, encoding an IAA-glucose hydrolase, could increase GL and TGW in the Indian rice landrace Kasalath. The *tgw6* allele in Nipponbare affected the timing of the transition from syncytial to cellular phases by controlling IAA supply through IAA-glucose hydrolase activity, and consequently limited cell number and GL. In our previous study, *TaTGW6* gene on chromosome 4AL was cloned using a comparative genetics approach (Hu et al., 2016). Our results revealed that low IAA contents in immature seeds at middle (20 DPA) and late grain filling stages (30 DPA) were significantly associated with low transcript levels of *TaTGW6*, corresponding to high TGW. Our previous study showed that the IAA content in immature seeds at the middle and late stages of grain filling (20–30 DPA) was significantly related to grain weight (Hu et al., 2016). Taken together, these results suggest that grain weight is negatively regulated by IAA at the middle to late stages of grain filling. Different from *TGW6* of rice and *TaTGW6* of wheat, *TaTGW-7A* positively regulates the grain weight of wheat. It is noteworthy that *TaTGW-7A*, *TGW6*, and *TaTGW6* are all associated with grain weight and involved in IAA biosynthesis and metabolism. These findings suggest that IAA level has a close relationship with grain development. Nevertheless, elucidation of the molecular mechanism underlying *TaTGW-7A* regulation of grain development will require more detailed research.

QTL for Grain Weight on Chromosome 7A

In previous studies, QTL for grain weight have been detected on almost all 21 wheat chromosomes in different segregating

populations (Shah et al., 1999; Kato et al., 2000; Ammiraju et al., 2001; Börner et al., 2002; Groos et al., 2003; Huang et al., 2003, 2006; Quarrie et al., 2005; Breseghello and Sorrells, 2006; Kumar et al., 2006; Li et al., 2007; Liao et al., 2008; Sun et al., 2009; Lopes et al., 2013; Simmonds et al., 2014). Several QTL for TGW or grain size on chromosome 7A have been reported (Huang et al., 2003; Kumar et al., 2006, 2007; Kuchel et al., 2007; Su et al., 2016). However, no TGW gene on chromosome 7A has been cloned so far. In the present study, a major QTL at *TaTGW-7A* locus for TGW, GL and GW, was also detected on chromosome 7AS linked to *Xbarc222*, and it was located to chromosome C-7AS-8 (0.45) based on the reports of Qi et al. (2004) and Sourdille et al. (2004). This location differs from those of the previously reported QTL clusters (Kumar et al., 2007). Therefore, *TaTGW-7A* is likely to be a novel gene associated with TGW in wheat.

Strong Positive Selection of *TaTGW-7Aa* in Wheat Breeding

As an important agronomical trait, high TGW is always the target of selection in crop improvement; therefore, underlying genes involved in TGW variation have been selected during the long-term breeding process. For example, *TaGW2-6A* (*Hap-6A-A*) associated with TGW has undergone positive selection in wheat breeding (Su et al., 2011). Similarly, *TaGS-D1* associated with TGW and GL showed a high frequency of distribution in Chinese wheat varieties and underwent strong positive selection during the breeding process (Zhang et al., 2014). *TaGS5-3A*, a gene underlying grain size, also underwent strong artificial selection during wheat polyploidization events (Ma et al., 2016).

In this study, the superior allele *TaTGW-7Aa* showed a significantly higher frequency than *TaTGW-7Ab* both in Chinese mini-core collections and 501 wheat varieties, especially in modern wheat varieties. In addition, *TaTGW-7Aa* was found to be widely distributed among 10 ecological wheat production zones of China, indicating that this allele has been strongly selected in Chinese wheat breeding. The *TGW7A* and *TG9* markers developed in this study can be used directly for TGW improvement in wheat breeding. In particular, the InDel marker *TG9* segregating with *TaTGW-7A* is more easily used for screening desirable lines for the multi-gene pyramiding in wheat breeding.

CONCLUSION

TaTGW-7A, an IGPS-like gene, was cloned and characterized via super-BSA and bio-information analysis. Two alleles, *TaTGW-7Aa* and *TaTGW-7Ab*, could be discriminated by CAPS marker *TGW7A* and InDel marker *TG9* designed based on the sequences of *TaTGW-7A*. The gene was located on chromosome 7AS and 7.6–27.1% of phenotypic variations of TGW could be explained by the locus in different environments and genetic backgrounds. *TaTGW-7Aa* showed higher transcript levels and corresponded

¹³<http://linux1.softberry.com/berry.phtml>

¹⁴<http://www.ncbi.nlm.nih.gov/>

to significantly higher TGW but lower IAA content during grain development, compared with *TaTGW-7Ab*. *TaTGW-7Aa* has been strongly and positively selected in wheat breeding in China. The molecular markers *TGW7A* and *TG9* can be used for improvement of TGW in breeding programs.

ETHICAL STANDARDS

The experiments conducted in this study comply with the current laws of China.

AUTHOR CONTRIBUTIONS

CC and C-XM initiated the project and designed the experiment; M-JH, H-PZ, and CC conducted gene cloning, marker development, mapping analysis, and prepared the manuscript; M-JH, XZ, Z-YW, HJ, J-JC, JL, and S-XW performed all tests of grain size, weight and IAA content; X-CX and G-LS assisted in writing the paper; all authors read and approved the final manuscript.

FUNDING

This work was supported by the China Agriculture Research System (CARS-03), the Promotion project of high education of Anhui province “Team construction of high level teacher of crop discipline,” the introduced leading talent research team for Universities in Anhui Province, Anhui province natural science foundation of China (1508085MC57), Wheat genetics and breeding research platform innovation team of Anhui’s university, and Jiangsu Collaborative Innovation Center for Modern Crop Production (JCIC-MCP).

REFERENCES

- Ammiraju, J. S. S., Dholakia, B. B., Santra, D. K., Singh, H., Lagu, M. D., Tamhankar, S. A., et al. (2001). Identification of inter simple sequence repeat (ISSR) markers associated with seed size in wheat. *Theor. Appl. Genet.* 102, 726–732. doi: 10.1007/s001220051703
- Börner, A., Schumann, E., Fürste, A., Cöster, H., Leithold, B., Röder, M., et al. (2002). Mapping of quantitative trait loci determining agronomic important characters in hexaploid wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 105, 921–936. doi: 10.1007/s00122-002-0994-1
- Brenchley, R., Spannagl, M., Pfeifer, M., Barker, G. L. A., D’Amore, R., Allen, A. M., et al. (2012). Analysis of the bread wheat genome using whole-genome shotgun sequencing. *Nature* 491, 705–710. doi: 10.1038/nature11650
- Breseghele, F., and Sorrells, M. E. (2006). Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. *Genetics* 172, 1165–1177. doi: 10.1534/genetics.105.044586
- Browne, R. A., Bequain, C., and Mascher, F. (2006). Components of partial disease resistance detected using a detached leaf assay in CIMMYT *Fusarium* head blight resistant wheat germplasm. *Eur. J. Plant Pathol.* 116, 199–210. doi: 10.1007/s10658-006-9050-3
- Bush, M., and Dixon, R. (2012). The role of bacterial enhancer binding proteins as specialized activators of σ^{54} -dependent transcription. *Microbiol. Mol. Biol. Rev.* 76, 497–529. doi: 10.1128/MMBR.00006-12
- Chang, C., Lu, J., Zhang, H. P., Ma, C. X., and Sun, G. L. (2015). Copy number variation of cytokinin oxidase gene *Tackx4* associated with grain weight and chlorophyll content of flag leaf in common wheat. *PLoS ONE* 10:e0145970. doi: 10.1371/journal.pone.0145970
- Chen, S., Huang, Z., Dai, Y., Qin, S., Gao, Y., Zhang, L., et al. (2013). The development of 7E chromosome-specific molecular markers for *Thinopyrum elongatum* based on SLAF-seq technology. *PLoS ONE* 8:e65122. doi: 10.1371/journal.pone.0065122
- Cheng, J., Randall, A., and Baldi, P. (2006). Prediction of protein stability changes for single-site mutations using support vector machines. *Proteins* 62, 1125–1132. doi: 10.1002/prot.20810
- Choi, H. K., Mun, J. H., Kim, D. J., Zhu, H. Y., Baek, J. M., Mudge, J., et al. (2004). Estimating genome conservation between crop and model legume species. *Proc. Natl. Acad. Sci. U.S.A.* 101, 15289–15294. doi: 10.1073/pnas.0402251101
- Davey, J. W., Cezard, T., Fuentes-Utrilla, P., Eland, C., Gharbi, K., and Blaxter, M. K. (2013). Special features of RAD sequencing data: implications for genotyping. *Mol. Ecol.* 22, 3151–3164. doi: 10.1111/mec.12084
- Dong, L. L., Wang, F. M., Liu, T., Dong, Z. Y., Li, A. L., Jing, R. L., et al. (2014). Natural variation of *TaGASR7-A1* affects grain length in common wheat under multiple cultivation conditions. *Mol. Breed.* 34, 937–947. doi: 10.1007/s11032-014-0087-2

ACKNOWLEDGMENTS

We thank Prof. Shi-He Xiao and Prof. Ji-Zeng Jia for kindly providing a RIL population, and Chinese micro-core wheat collections, respectively.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpls.2016.01902/full#supplementary-material>

TABLE S1 | Grain traits of the two parents, RIL population (Pop 1).

TABLE S2 | Grain weight and allelic variants in Pop 2.

TABLE S3 | Grain traits and allelic variations of 257 Chinese wheat mini-core collections (Pop 3).

TABLE S4 | Allelic variation and geographic distribution for 501 accessions (Pop 4).

TABLE S5 | All polymorphic SLAF with detal_index.

TABLE S6 | Primers for quantitative RT-PCR, functional markers and amplification of the full-length gDNA and cDNA.

TABLE S7 | Significant association SLAF marker.

TABLE S8 | Significantly associated SLAF tags on genome of common wheat.

TABLE S9 | Genes predicted to associate with grain development.

TABLE S10 | QTLs for TGW across five environments.

TABLE S11 | Grain weight and allele variations of twelve varieties.

TABLE S12 | The proportions of two *TaTGW-7A* alleles in RIL (Pop 1), natural population (Pop 2), and Chinese wheat mini-core collections (Pop 3).

TABLE S13 | Geographic distribution of *TaTGW-7A* alleles in Pop 3 and Pop 4.

- Eilam, T., Anikster, Y., Millet, E., Manisterski, J., Sagi-Assif, O., and Feldman, M. (2007). Genome size and genome evolution in diploid Triticeae species. *Genome* 50, 1029–1037. doi: 10.1139/G07-083
- Gene Ontology Consortium (2004). The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res.* 32, 258–261. doi: 10.1093/nar/gkh036
- Giura, A., and Saulescu, N. N. (1996). Chromosomal location of genes controlling grain size in a large grained selection of wheat (*Triticum aestivum* L.). *Euphytica* 89, 77–80. doi: 10.1007/BF00015722
- Groos, C., Robert, N., Bervas, E., and Charmet, G. (2003). Genetic analysis of grain protein-content, grain yield and thousand-kernel weight in bread wheat. *Theor. Appl. Genet.* 106, 1032–1040. doi: 10.1007/s00122-002-1111-1
- Gupta, P. K., Mir, R. R., Mohan, A., and Kumar, J. (2008). Wheat genomics: present status and future prospects. *Int. J. Plant Genomics* 2008, 1–22. doi: 10.1155/2008/896451
- Hao, C. Y., Dong, Y. C., Wang, L. F., You, G. X., Zhang, H. N., Ge, H. M., et al. (2008). Genetic diversity and construction of core collection in Chinese wheat genetic resources. *Chin. Sci. Bull.* 53, 1518–1526. doi: 10.1007/s11434-008-0212-x
- Hill, J. T., Demarest, B. L., Bisgrove, B. W., Gorski, B., Su, Y. C., and Yost, H. J. (2013). MMAPP: mutation mapping analysis pipeline for pooled RNA-seq. *Genome Res.* 23, 687–697. doi: 10.1101/gr.146936.112
- Hu, M. J., Zhang, H. P., Chao, J. J., Zhu, X. F., Wang, S. X., Jiang, H., et al. (2016). Characterization of an IAA-glucose hydrolase gene TaTGW6 associated with grain weight in common wheat (*Triticum aestivum* L.). *Mol. Breed.* 36, 1–11. doi: 10.1007/s11032-016-0449-z
- Huang, X. Q., Cloutier, S., Lycar, L., Radovanovic, N., Humphreys, D. G., Noll, J. S., et al. (2006). Molecular detection of QTLs for agronomic and quality traits in a doubled haploid population derived from two Canadian wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 113, 753–766. doi: 10.1007/s00122-006-0346-7
- Huang, X. Q., Cöster, H., Ganai, M. W., and Röder, M. S. (2003). Advanced backcross QTL analysis for the identification of quantitative trait loci alleles from wild relatives of wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 106, 1379–1389. doi: 10.1007/s00122-002-1179-7
- Ishimaru, K., Hirotsu, N., Madoka, Y., Murakami, N., Hara, N., Onodera, H., et al. (2013). Loss of function of the IAA-glucose hydrolase gene TGW6 enhances rice grain weight and increases yield. *Nat. Genet.* 45, 707–711. doi: 10.1038/ng.2612
- Jiang, Y., Jiang, Q., Hao, C., Hou, J., Wang, L., Zhang, H., et al. (2015). A yield-associated gene TaCWI in wheat: its function, selection and evolution in global breeding revealed by haplotype analysis. *Theor. Appl. Genet.* 128, 131–143. doi: 10.1007/s00122-014-2417-5
- Kanehisa, M. (2002). The KEGG database. *Novartis Found. Symp.* 247, 91–103. doi: 10.1002/0470857897.ch8
- Kang, H. W., Cho, Y. G., Yoon, U. H., and Eun, M. Y. (1998). A rapid DNA extraction method for RFLP and PCR analysis from a single dry seed. *Plant Mol. Biol. Rep.* 16, 1–9. doi: 10.1023/A:100741860609
- Kato, K., Miura, H., and Sawada, S. (2000). Mapping QTLs controlling grain yield and its components on chromosome 5A of wheat. *Theor. Appl. Genet.* 101, 1114–1121. doi: 10.1007/s001220051587
- Kent, W. J. (2002). BLAT—the BLAST-like alignment tool. *Genome Res.* 12, 656–664. doi: 10.1101/gr.229202
- Kuchel, H., Williams, K. J., Langridge, P., Eagles, H. A., and Jefferies, S. P. (2007). Genetic dissection of grain yield in bread wheat. I. QTL analysis. *Theor. Appl. Genet.* 115, 1029–1041. doi: 10.1007/s00122-007-0629-7
- Kumar, N., Kulwal, P. L., Balyan, H. S., and Gupta, P. K. (2007). QTL mapping for yield and yield contributing traits in two mapping populations of bread wheat. *Mol. Breed.* 19, 163–177. doi: 10.1007/s11032-006-9056-8
- Kumar, N., Kulwal, P. L., Gaur, A., Tyagi, A. K., Khurana, J. P., Khurana, P., et al. (2006). QTL analysis for grain weight in common wheat. *Euphytica* 151, 135–144. doi: 10.1007/s10681-006-9133-4
- Li, R., Li, Y., Kristiansen, K., and Wang, J. (2008). SOAP: short oligonucleotide alignment program. *Bioinformatics* 24, 713–714. doi: 10.1093/bioinformatics/btn025
- Li, S. S., Jia, J. Z., Wei, X. Y., Zhang, X. C., Li, L. Z., Chen, H. M., et al. (2007). A intervarietal genetic map and QTL analysis for yield traits in wheat. *Mol. Breed.* 20, 167–178. doi: 10.1007/s11032-007-9080-3
- Liao, X. Z., Wang, J., Zhou, R. H., Ren, Z. L., and Jia, J. Z. (2008). Mining favorable alleles of QTLs conferring thousand-grain weight from synthetic wheat. *Acta Agron. Sin.* 34, 1877–1884. doi: 10.1016/S1875-2780(09)60012-2
- Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2-DDCT method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262
- Lopes, M. S., Reynolds, M. P., McIntyre, C. L., Mathews, K. L., Jalal Kamali, M. R., Mossad, M., et al. (2013). QTL for yield and associated traits in the Seri/Babax population grown across several environments in Mexico, in the West Asia, North Africa, and South Asia regions. *Theor. Appl. Genet.* 126, 971–984. doi: 10.1007/s00122-012-2030-4
- Lu, J., Chang, C., Zhang, H. P., Wang, S. X., Sun, G. L., Xiao, S. H., et al. (2015). Identification of a novel allele of TaCKX6a02 associated with grain size, filling rate and weight of common wheat. *PLoS ONE* 10:e0144765. doi: 10.1371/journal.pone.0144765
- Ma, D. Y., Yan, J., He, Z. H., Wu, L., and Xia, X. C. (2012). Characterization of a cell wall invertase gene TaCwi-A1 on common wheat chromosome 2A and development of functional markers. *Mol. Breed.* 29, 43–52. doi: 10.1007/s11032-010-9524-z
- Ma, L., Li, T., Hao, C. Y., Wang, Y. Q., Chen, X. H., and Zhang, X. Y. (2016). TaGS5-3A, a grain size gene selected during wheat improvement for larger kernel and yield. *J. Plant Biotechnol.* 14, 1269–1280. doi: 10.1111/pbi.12492
- Pomraning, K. R., Smith, K. M., and Freitag, M. (2011). Bulk segregant analysis followed by high-throughput sequencing reveals the *Neurospora* cell cycle gene, *ndc-1*, to be allelic with the gene for ornithine decarboxylase, *spe-1*. *Eukaryot. Cell* 10, 724–733. doi: 10.1128/EC.00016-11
- Qi, L. L., Echaliier, B., Chao, S., Lazo, G. R., Butler, G. E., Anderson, O. D., et al. (2004). A chromosome bin map of 16,000 expressed sequence tag loci and distribution of genes among the three genomes of polyploid wheat. *Genetics* 168, 701–712. doi: 10.1534/genetics.104.034868
- Qin, D. D., Dong, J., Xu, F. C., Guo, G. G., Ge, S. T., Xu, Q., et al. (2015). Characterization and fine mapping of a novel barley stage green-reversible Albino gene (HvSGRA) by bulked segregant analysis based on SSR assay and specific length amplified fragment sequencing. *BMC Genomics* 16:838. doi: 10.1186/s12864-015-2015-1
- Quarrie, S. A., Jancic, V. L., Kovacevic, D., Steed, A., and Pekic, S. (1999). Bulk segregant analysis with molecular markers and its use for improving drought resistance in maize. *J. Exp. Bot.* 50, 1299–1306. doi: 10.1093/jxb/50.337.1299
- Quarrie, S. A., Steed, A., Calestani, C., Semikhodskii, A., Lebreton, C., Chinoy, C., et al. (2005). A high-density genetic map of hexaploid wheat (*Triticum aestivum* L.) from the cross Chinese Spring × SQ1 and its use to compare QTLs for grain yield across a range of environments. *Theor. Appl. Genet.* 110, 865–880. doi: 10.1007/s00122-004-1902-7
- Shah, M. M., Gill, K. S., Baenziger, P. S., Yen, Y., Kaeppeler, S. M., and Ariyaratne, H. M. (1999). Molecular mapping of loci for agronomic traits on chromosome 3A of bread wheat. *Crop Sci.* 39, 1728–1732. doi: 10.2135/cropsci1999.3961728x
- Simmonds, J., Scott, P., Leverington-Waite, M., Turner, A. S., Brinton, J., Korzun, V., et al. (2014). Identification and independent validation of a stable yield and thousand grain weight QTL on chromosome 6A of hexaploid wheat (*Triticum aestivum* L.). *BMC Plant Biol.* 14:191–203. doi: 10.1186/s12870-014-0191-9
- Somers, D. J., Isaac, P., and Edwards, K. (2004). A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 109, 1105–1114. doi: 10.1007/s00122-004-1740-7
- Sourdille, P., Singh, S., Cadalen, T., Brown-Guedira, G. L., Gay, G., Qi, L., et al. (2004). Microsatellite-based deletion bin system for the establishment of genetic-physical map relationships in wheat (*Triticum aestivum* L.). *Funct. Integr. Genomics* 4, 12–25. doi: 10.1007/s10142-004-0106-1
- Su, Z. Q., Hao, C. Y., Wang, L. F., Dong, Y. C., and Zhang, X. Y. (2011). Identification and development of a functional marker of TaGW2 associated with grain weight in bread wheat. *Theor. Appl. Genet.* 122, 211–223. doi: 10.1007/s00122-010-1437-z
- Su, Z. Q., Jin, S. J., Lu, Y., Zhang, G. R., Chao, S. M., and Bai, G. H. (2016). Single nucleotide polymorphism tightly linked to a major QTL on chromosome 7A for both kernel length and kernel weight in wheat. *Mol. Breed.* 36, 1–11. doi: 10.1007/s11032-016-0436-4
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: a knowledge-based approach

- for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* 102, 15545–15550. doi: 10.1073/pnas.0506580102
- Sun, X., Liu, D., Zhang, X., Li, W., Liu, H., Hong, W., et al. (2013). SLAF-seq: an efficient method of large-scale de novo SNP discovery and genotyping using high-throughput sequencing. *PLoS ONE* 8:e58700. doi: 10.1371/journal.pone.0058700
- Sun, X. Y., Wu, K., Zhao, Y., Kong, F. M., Han, G. Z., Jiang, H. M., et al. (2009). QTL analysis of kernel shape and weight using recombinant inbred lines in wheat. *Euphytica* 165, 615–624. doi: 10.1007/s10681-008-9794-2
- Trick, M., Adamski, N. M., Mugford, S. G., Jiang, C. C., Febrer, M., and Uauy, C. (2012). Combining SNP discovery from next-generation sequencing data with bulked segregant analysis (BSA) to fine-map genes in polyploid wheat. *BMC Plant Biol.* 12:14. doi: 10.1186/1471-2229-12-14
- Wicker, T., Mayer, K. F., Gundlach, H., Martis, M., Steuernagel, B., Scholz, U., et al. (2011). Frequent gene movement and pseudogene evolution is common to the large and complex genomes of wheat, barley, and their relatives. *Plant Cell* 23, 1706–1718. doi: 10.1105/tpc.111.086629
- Xu, F. F., Sun, X., Chen, Y. L., Huang, Y., Tong, C., and Bao, J. S. (2015a). Rapid identification of major QTLs associated with rice grain weight and their utilization. *PLoS ONE* 10:e0122206. doi: 10.1371/journal.pone.0122206
- Xu, X., Xu, R., Zhu, B., Yu, T., Qu, W., Lu, L., et al. (2015b). A high-density genetic map of cucumber derived from specific length amplified fragment sequencing (SLAF-seq). *Front. Plant Sci.* 5:768. doi: 10.3389/fpls.2014.00768
- Xu, X. W., Lu, L., Zhu, B. Y., Xu, Q., Qi, X. H., and Chen, X. H. (2015c). QTL mapping of cucumber fruit flesh thickness by SLAF-seq. *Sci. Rep.* 5:15829. doi: 10.1038/srep15829
- Yin, C. B., Li, H. H., Li, S. S., Xu, L. D., Zhao, Z. G., and Wang, J. K. (2015). Genetic dissection on rice grain shape by the two-dimensional image analysis in one japonica × indica population consisting of recombinant inbred lines. *Theor. Appl. Genet.* 128, 1969–1986. doi: 10.1007/s00122-015-2560-7
- Zhang, P., Zhu, Y. Q., Wang, L. L., Chen, L. P., and Zhou, S. J. (2015). Mining candidate genes associated with powdery mildew resistance in cucumber via super-BSA by specific length amplified fragment (SLAF) sequencing. *BMC Genomics* 16:1058. doi: 10.1186/s12864-015-2041-z
- Zhang, Y. J., Liu, J. D., Xia, X. C., and He, Z. H. (2014). TaGS-D1, an ortholog of rice OsGS3, is associated with grain weight and grain length in common wheat. *Mol. Breed.* 34, 1097–1107. doi: 10.1007/s11032-014-0102-7
- Zhang, Z., Shang, H. H., Shi, Y. Z., Huang, L., Li, J. W., Ge, Q., et al. (2016). Construction of a high-density genetic map by specific locus amplified fragment sequencing (SLAF-seq) and its application to quantitative trait loci (QTL) analysis for boll weight in upland cotton (*Gossypium hirsutum.*). *BMC Plant Biol.* 16:79. doi: 10.1186/s12870-016-0741-4

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Hu, Zhang, Liu, Cao, Wang, Jiang, Wu, Lu, Zhu, Xia, Sun, Ma and Chang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.