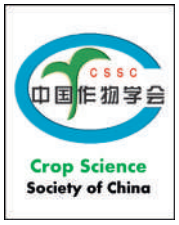
THE CROP JOURNAL 8 (2020) 943–952



Available online at www.sciencedirect.com

ScienceDirect

Functional markers developed from *TaGS3*, a negative regulator of grain weight and size, for marker-assisted selection in wheat

Wei Zhang[c,1], *Huifang Li*[b,1], *Liya Zhi*[a], *Qiannan Su*[a], *Jiajia Liu*[a], *Xiaoli Ren*[a], *Deyuan Meng*[a], *Na Zhang*[a,d,e], *Jun Ji*[a,d,e], *Xueyong Zhang*[b,*], *Junming Li*[a,d,e,*]

[a]*Center for Agricultural Resources Research, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Shijiazhuang 050022, Hebei, China*

[b]*Key Laboratory of Crop Gene Resources and Germplasm Enhancement, Ministry of Agriculture/The National Key Facility for Crop Gene Resources and Genetic Improvement, Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, China*

[c]*Hebei University of Economics and Business, Shijiazhuang 050061, Hebei, China*

[d]*State Key Laboratory of Plant Cell and Chromosome Engineering, Chinese Academy of Sciences, Beijing 100101, China*

[e]*The Innovative Academy of Seed Design, Chinese Academy of Sciences, Beijing 100101, China*

ARTICLE INFO

Article history:
Received 19 January 2020
Received in revised form 24 February 2020
Accepted 14 March 2020
Available online 7 May 2020

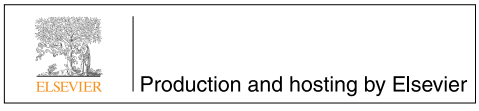
Keywords:
TaGS3
Grain yield
Haplotype
MAS
Triticum aestivum

ABSTRACT

The *TaGS3* homoeologous genes (homoeologs) located on chromosomes 7A, 4A, and 7D in hexaploid wheat were cloned. Relative expression analysis of the three *TaGS3* homoeologs revealed that the expression levels of *TaGS3-4A* and *TaGS3-7D* in developing grains were higher than that of *TaGS3-7A*. Genetic evidence showed that *TaGS3* was a negative regulator of grain weight and grain size. Fifteen polymorphic sites and five haplotypes were detected in *TaGS3-4A*. Two molecular markers were developed to distinguish the five haplotypes. Association analysis using 260 accessions from Chinese wheat mini-core collection (MCC) indicated that *TaGS3-4A* affected thousand grain weight (TGW) and grain length (GL). *HAP-4A-1* and *HAP-4A-2* were favorable haplotypes that increased TGW and GL and had undergone strong selection during domestication of wheat. In addition, interaction of the *TaGS3-4A* and *TaGS3-7D* homoeologs had significant additive effects on the grain traits. *Hap-4A-1/Hap-7D-2* was the best haplotype combination in increasing TGW and GL. The frequencies and geographic distributions of favorable *TaGS3* haplotypes among 1388 wheat accessions from worldwide sources provided clues for selection of yield-related traits. Our findings demonstrated that *TaGS3-4A* had significant effects on TGW and

* *Corresponding authors.*
E-mail addresses: zhangxueyong@caas.cn, (X. Zhang), ljm@sjziam.ac.cn. (J. Li).

Peer review under responsibility of Crop Science Society of China and Institute of Crop Science, CAAS.



[1] Wei Zhang and Huifang Li contributed equally to this work.

https://doi.org/10.1016/j.cj.2020.03.003

GL. Marker-assisted selection of *HAP-4A-1/2* combined with *HAP-7D-2* has potential to increase wheat yields.

1. Introduction

Wheat (*Triticum aestivum* L.) is a staple crop in the world providing nearly one-half of world food consumption. Improvement of the grain yield potential is the key focus of wheat breeders. Grain size and thousand grain weight (TGW) are major determinants of grain yield and grain size was the target of selection during domestication, because large seeds were generally favored due to ease of harvesting, enhanced seedling vigor and widespread market preference.

Progress in rice genomics and isolation of increasing numbers of yield-related genes in other cereals provide opportunities for homology-based cloning of genes in wheat. Cloning yield-related genes, developing functional markers and identifying favorable haplotypes in wheat should be helpful for yield improvement. A number of genes for grain weight/size that play different roles at various stages of grain development have already been cloned in wheat. For example, *TaSus2* was isolated and found to be significantly associated with thousand grain weight [1]. *TaCKX6* was reported to play a leading role in controlling cytokin in levels and affected grain weight [2]. *TaCwi-A1*, encoding the critical enzyme for sink tissue development and carbon partitioning (cell wall invertase), had a significant association with grain weight [3,4]. *TaSAP1-A1* was significantly associated with grain weight, grain number per spike, spike length and peduncle length in multiple environments [5]. *TaCYP78A3* influenced grain size by affecting the extent of integument cell proliferation [6]. *TaGS5* was found to be a positive regulator of grain size [7–9]. *TaFlo2-A1* was associated with TGW and a molecular marker was developed for marker-assisted selection [10]. *TaGW2* was significantly associated with grain width and negatively regulated TGW by controlling gene expression level during seed development [11–15]. Knockdown of *TaBT1* caused a decrease in grain size, TGW and grain total starch content [16].

OsGS3 was identified as a major gene for both grain weight and length in rice [17,18]. It also affected grain length by regulating stigma length and stigma exsertion [19]. Although the wheat orthologs *TaGS3-7D* and *TaGS3-7A* were reported to be associated with grain weight and grain length in wheat [20,21], the effects of *TaGS3-4A* on grain size and grain weight remain poorly understood. Therefore, identification of allelic variants of *TaGS3-4A* with increased grain size/weight and development of molecular markers could facilitate breeding for yield by marker-assisted selection.

The objectives of this study were to (1) determine the sequence diversity and allelic distribution of *TaGS3* homoeologs in wheat; (2) assess the expression of the three *TaGS3* homoeologous genes in different tissues; (3) develop functional markers for *TaGS3-4A* and identify the favorable haplotypes for yield improvement; and (4) examine interactions among favorable haplotypes of *TaGS3-4A* and *TaGS3-7D*. The study identified potentially important genes and functional markers for marker-assisted selection.

2. Materials and methods

2.1. Plant materials

Chinese Spring (CS) and *T. aestivum* cv. Kenong 9204 (Kn9204) were used for cloning the *TaGS3* homeologous genes in wheat. The Chinese wheat mini-core collection (MCC) consisting of 150 landraces, 88 modern cultivars and 22 introduced lines and representing more than 70% of the total genetic diversity in the Chinese wheat germplasm collection were used to detect the association of *TaGS3-4A* and *TaGS3-7D* haplotypes with grain traits [22]. Three hundred and forty-three modern Chinese cultivars (MC) [23], 478 cultivars from North America, 53 cultivars from CIMMYT, 382 cultivars from Europe, 82 cultivars from the former USSR (Former Union of Soviet Socialist Republics), and 50 cultivars from Australia were used to investigate the global distribution of *TaGS3* haplotypes [24].

2.2. Measurement of grain traits

The MCC panel was planted at Luoyang in Henan province in 2002, 2005, and 2006, Shunyi in Beijing in 2010 and Luancheng in Hebei province in 2016. A randomized block design with three replications was used in each environment, and each accession was planted in a 2-row 2 m plot with 25 cm row spacing and 40 seeds per row. Grain traits were measured by a SC-G multifunctional seed analyzer (Hangzhou Wanshen Detection Technology Co., Ltd., Hangzhou).

2.3. Cloning and characterization of *TaGS3*

Primers based on constructed sequences were designed using Primer 5.0 software (http://www.premierbiosoft.com) and synthesized by Shanghai Sangon Biotech Co., Ltd. (http://www.sangon.com/). All the PCR primers used in this study are listed in Table S1.

Full-length genomic DNA of the *TaGS3* homoeologs was obtained by PCR-directed cloning. The genomic sequence of the rice *OsGS3* gene (Os03g0407400) was used as a query for BLAST searches against the wheat sequences database in EnsemblePlants (http://plants.ensembl.org/). CS and Kn9204 were used to clone the *TaGS3* sequences. A pair of conserved primers was used to specifically amplify *TaGS3* coding sequences from the three wheat sub-genomes: A, B, and D (Table S1). The PCR products were separated by electrophoresis in agarose gels, and the target bands were extracted and cloned into the pEASY-T1 vector and transformed to DH5α competent *E. coli* cells by the heat shock method (Transgen,

Beijing, Product Code: CT101). Positive clones were selected for sequencing. PCR and isolation of positive clones were repeated at least three times.

2.4. Gene expression analyses

Total RNA was extracted using TRIzol reagent (Invitrogen, USA). The first-strand cDNA was synthesized from DNase I-treated total RNA using the PrimeScript RT Reagent Kit (TaKaRa, Japan) according to the manufacturer's instructions.

Quantitative real-time PCR was carried out in total volumes of 20 uL using SYBR PCR kit (TaKaRa) on an ABI 7500 real-time PCR system (Applied Biosystems, USA) according to the manufacturer's instructions. All expression level data obtained by quantitative real-time PCR were based on at least four biological samples on which three replications were conducted using the comparative Ct method after normalization to the *GADPH* control. The specific primer sequences are listed in Table S1.

2.5. Plasmid construction and genetic transformation

The coding sequence of *TaGS3-4A* was cloned into binary vector pJIT163 under the control of the 35S promoter. The recombinant plasmid was transformed into *T. aestivum* cv. Kn199. Positive lines were verified by PCR amplification with primers listed in Table S1.

2.6. SNP identification and functional marker development

Twenty-four cultivars (Table S2) were chosen for initial detection of sequence variation in the coding region of *TaGS3-4A*. SNPs were identified using DNAMAN software.

Two markers (Table S1) were developed to distinguish the five haplotypes of *TaGS3-4A*. Genome-specific fragments were amplified and digested by the corresponding primers and restriction enzyme. PCR products were selected for DNA sequencing and analysis in order to examine the accuracy of markers.

2.7. Statistical analyses

Statistical analyses were based on phenotypic data for grain size and grain weight in five environments. One-way ANOVA was performed in the SPSS System for Windows version 17.0 (SPSS, Inc., Chicago, IL, USA) to determine phenotypic differences between individual haplotypes and haplotype combinations and Tukey tests were conducted to determine the significance of differences.

3. Results

3.1. Cloning and chromosome location of *TaGS3* homoeologs

The rice *OsGS3* sequence (DQ355996) was used as a query against the wheat genome sequence database in the EnsemblPlants (http://plants.ensembl.org/index.html) to select potential candidate *TaGS3* genes. Three *Triticum aestivum* genes (2018 IWGSC RefSeq v1.0; TraesCS7A02G017700, start 7598370; TraesCS7D02G015000, start 733693567; and TraesCS4A02G474000, start 6483394) were identified as potential orthologs of *OsGS3*. Uniprot (https://www.uniprot.org) and ENA (https://www.ebi.ac.uk/ena) tools were used to search for the deduced protein sequences translated from the *TaGS3* homoeologs.

To clone the full-length genomic sequence of *TaGS3* in Chinese Spring a conserved primer pair yielding three fragments was designed to amplify the three *TaGS3* homoeologs. The *TaGS3* mRNA was also cloned with that primer (Table S1).

Among the three *TaGS3* homoelogous genes the length of the third intron was conserved, whereas the lengths of the fifth exon and the other three introns varied (Fig. 1-A). Forward and reverse primers were designed at exons 2 and exon 3, respectively, based on the length polymorphism of the second intron to distinguish the three *TaGS3* homoeologs simultaneously.

3.2. Characterization of *TaGS3* homoeologs

In contrast to the earlier report [20] we found that the open reading frames of *TaGS3-7A*, *TaGS3-4A*, and *TaGS3-7D* were 1983, 1936, and 2351 bp, respectively. The cDNA of both *TaGS3-7A* and *TaGS3-7D* were 510 bp and predicted to encode a polypeptide with 169 amino acids. The cDNA of *TaGS3-4A* was 513 bp and encoded 170 amino acids (Fig. 1-B). Like *OsGS3*, the three *TaGS3* homoeologs also consisted of five exons. The lengths of the second, third and fourth exons of *TaGS3* were the same as those in *OsGS3*; that is, 53, 45, and 54 bp, respectively (Fig. 1-A).

Nucleotide sequence analysis of the three *TaGS3* homoeologs showed that their identities were 97.47%. The TaGS3-4A protein contained one insertion, and there were 11 amino acid differences among the three genes (Fig. 1-B). Sequence alignments showed that TaGS3-4A and TaGS3-7D were more similar than either was to TaGS3-7A (Fig. 1-B, Fig. S1). TaGS3-7D protein showed 98% similarity to *Aegilops tauschii*, and the TaGS3 proteins showed higher similarities with those in *Hordeum vulgare* (89%), *Brachypodium distachyon* (63%), and *Oryza sativa* (62%) protein than in *Arabidopsis thaliana* (40%) (Fig. S1).

3.3. *TaGS3-4A* and *TaGS3-7D* were preferentially expressed in developing grains

Temporal and spatial expression patterns of the *TaGS3* homoeologs were investigated by quantitative real-time RT-PCR using genome-specific primers (Table S1). All three were ubiquitously expressed, but with different patterns in different tissues. There was higher expression in developing grains than in roots, stems and leaves.

Although all three genes were expressed in grains, their expression abundances were significantly different (Fig. S2). *TaGS3-4A* and *TaGS3-7D* were expressed at a much higher level than *TaGS3-7A*, and *TaGS3-4A* was preferentially expressed at the early stage of grain development. Since the haplotypes of *TaGS3-7D* were reported previously [20] we consequently focused on characterization of *TaGS3-4A* in the remainder of this study.

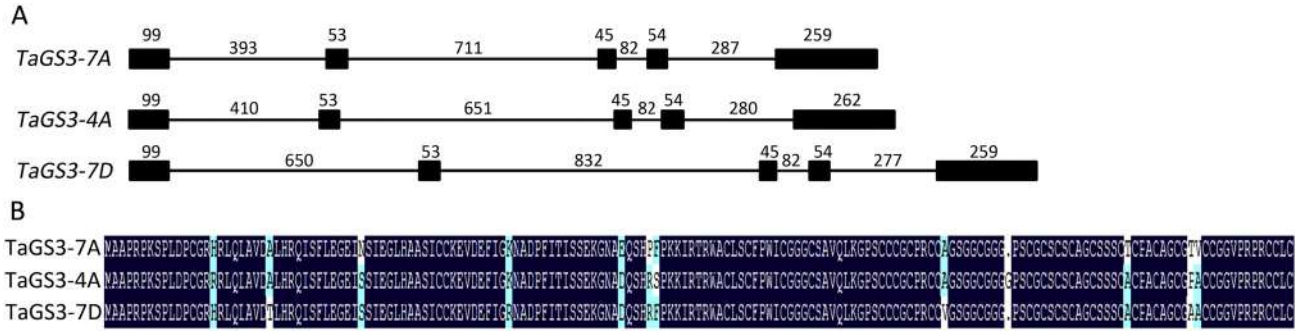


Fig. 1 – Gene structure and amino acid sequence of the three *TaGS3* homoeologs in Kn9204. Solid blocks indicate exons and lines represent introns. Numbers above both exons and introns denote size (bp).

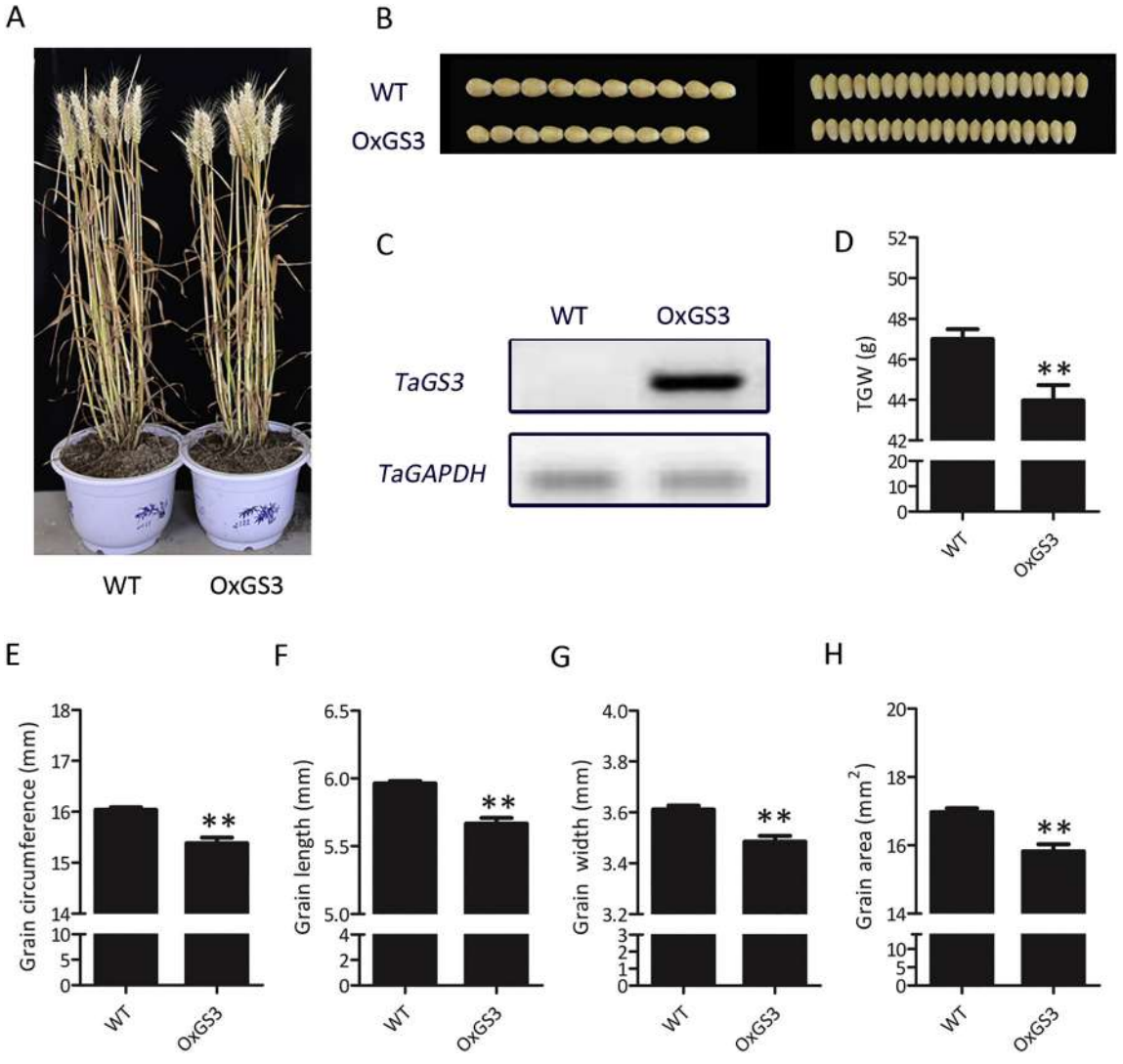


Fig. 2 – Genetic effects of *TaGS3-4A* on plant morphology and grain size. (A) Phenotypes of WT (wild type, Kn199) and *TaGS3-4A* transgenic line (OxGS3). (B) Grains of WT and OxGS3. (C) Confirmation of OxGS3 transgenic plant. (D–H) Grain trait comparisons of WT and OxGS3. Error bars indicate standard error (n = 10). **, significant difference between WT and OxGS3 at $P < 0.01$.

3.4. *TaGS3* negatively regulates grain weight and size in wheat

Transgenic *TaGS3-4A* over-expression lines in *T. aestivum* cv. Kn199 were constructed to investigate the potential function of *TaGS3* (Fig. 2-A). Twenty positive lines (T_0) were obtained and verified by PCR (Fig. 2-C). One homozygous line (T_3) was used for phenotypic analysis. Ten T_3 plants were harvested and at least 200 seeds per plant were analyzed. Compared with the wild type (WT) the transgenic plants (OxGS3) had smaller grains (Fig. 2-B), with mean reductions of 6.44% for TGW (Fig. 2-D), 4.08% for grain circumference (Fig. 2-E), 4.99% for grain length (GL) (Fig. 2-F), 3.47% for grain width (GW) (Fig. 2-G) and 6.75% for grain surface area (Fig. 2-H).

3.5. Allelic variation of *TaGS3-4A*

Fifteen polymorphic sites were detected in the *TaGS3-4A* genomic sequences of the 260 MCC accessions (Fig. S3). Eight SNPs (located at 634, 692, 710, 713, 718, 720, 1234, and 1258 bp) were present in introns and the exons had four SNPs (1718, 1761, 1791, and 1826 bp). Insertion/deletion (InDel) mutations were located at 354, 357, and 787 bp. Based on the 354 bp InDel and SNPs at 357 and 713 bp we identified five haplotypes, which were named *HAP-4A-1* to *HAP-4A-5*.

3.6. Marker development for *TaGS3-4A*

Two molecular markers were developed to distinguish the *TaGS3-4A* haplotypes. One InDel with a short repeat sequence (ACT) was at 354 bp. Based on the 3 bp repeat sequence diversity, a marker was developed to discriminate *HAP-4A-1/3*, *HAP-4A-2/4*, and *HAP-4A-5* (Fig. 3-A, B). The nucleotide polymorphism at 357 bp created a restriction enzyme recognition site for *Bsr* I (ACTGG) (Fig. 3-C, D), which was employed to develop a cleaved amplified polymorphism sequence (CAPS) marker to distinguish *HAP-4A-4* from the other four haplotypes. No restriction enzyme recognition site was found in *HAP-4A-4* (357 bp deletion G), whereas it existed in the other four haplotypes. Another restriction enzyme recognition site for *Hpy188* I (TCNGA) at 713 bp was employed to develop a CAPS marker to distinguish *HAP-4A-1* and *HAP-4A-2* from *HAP-4A-3*, *HAP-4A-4*, and *HAP-4A-5* (Fig. 3-E, F). An *Hpy188* I enzyme recognition site (TCAGA) was found in *Hap-4A-1* and *Hap-4A-2*, but not in *Hap-4A-3/4/5* (ACAGA). Thus, two markers for the three sites (354, 357, and 713 bp) distinguished all five haplotypes.

3.7. Differences in TGW and GL among *TaGS3-4A* haplotypes

To associate *TaGS3-4A* haplotypes with TGW, GL and GW, the two molecular markers were used to genotype the MCC. The MCC accessions clustered into two sub-populations

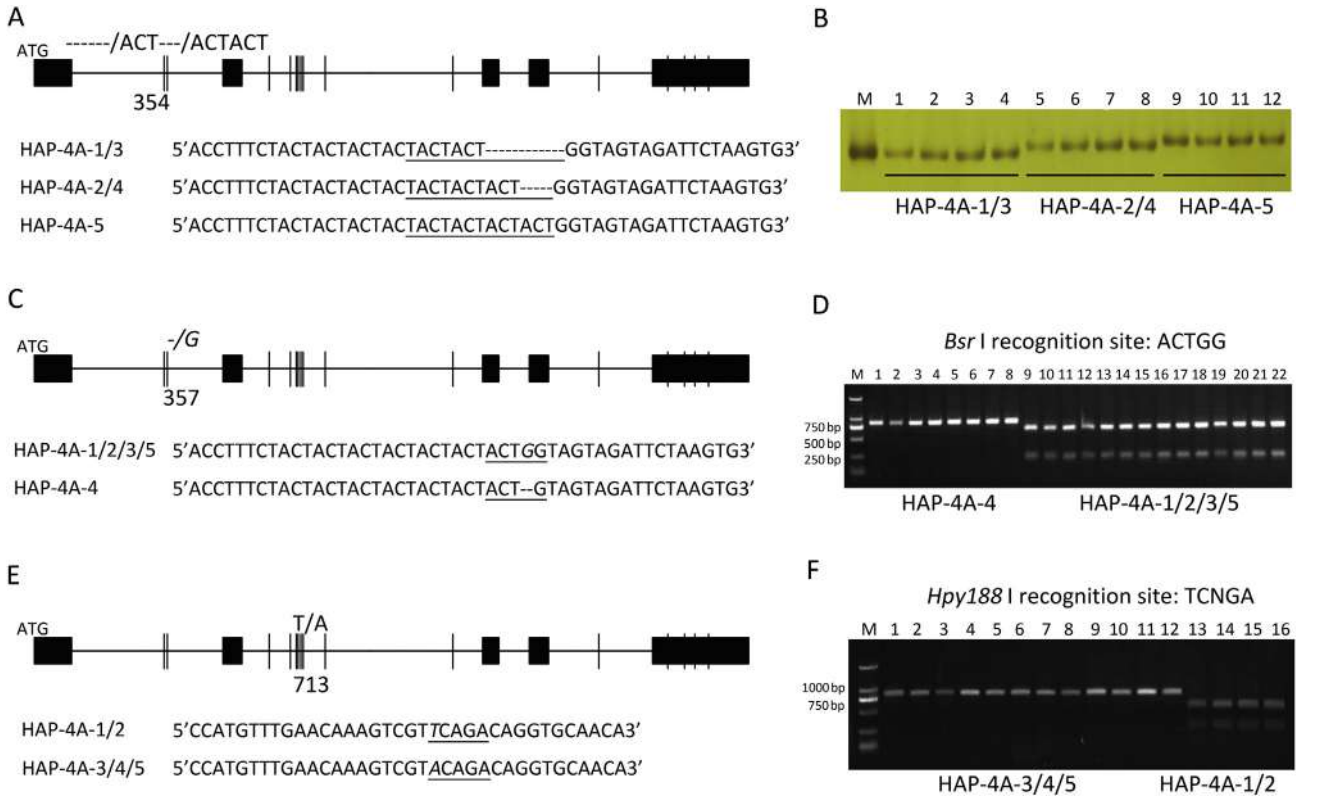


Fig. 3 – Marker development for *TaGS3-4A*. (A, B) Marker developed using nucleotide polymorphism at 354 bp. (C, D) PCR products of CAPS marker designed for SNP at 357 bp restrictively digested by *Bsr* I. (E, F) PCR products of CAPS marker based on the SNP at 713 bp restrictively digested by *Hpy188* I.

representing landraces and modern cultivars [25]. Therefore, the association analysis between *TaGS3-4A* haplotypes and grain traits took population structure into account.

Haplotype association analysis of the full MCC panel indicated that *HAP-4A-1* was quite similar to *HAP-4A-2* with positive effects on TGW and GL. Both were favorable haplotypes that increased TGW and GL whereas *HAP-4A-3* had a significantly negative effect on TGW and GL in all five environments. The TGW differences between *HAP-4A-1/2* and *HAP-4A-3* were 5.33 to 5.47, 3.79 to 3.98, 3.23 to 3.57, 3.26 to 6.17, and 4.82 to 5.12 g in 2002, 2005, 2006, 2010, and 2016, respectively (Table S3). Compared with *HAP-4A-3*, the favorable *HAP-4A-1* haplotype enhanced GL by 0.30 to 0.46 mm in four of the five environments. For the landraces, *HAP-4A-1* significantly increased TKW and GL in five and four environments, respectively. For the modern cultivars, *HAP-4A-2* has a significant influence on GL, with phenotypic differences between *HAP-4A-2* and *HAP-4A-3* being 0.47, 0.44, 0.43, and 0.49 mm in 2002, 2006, 2010, and 2016, respectively (Table S3). *HAP-4A-1* and *HAP-4A-2* had a slightly different effect on GL in landraces and modern cultivars (Table S3). TGW differences between *HAP-4A-1/2* and *HAP-4A-3* were larger in landraces than in modern cultivars. This was possibly due to greater numbers of superior alleles already selected in modern cultivars by breeding.

3.8. Additive genetic effects of favorable haplotypes at *TaGS3-4A* and *TaGS3-7D*

Association analysis was carried out on the MCC panel to determine combination effects of different haplotypes at the *TaGS3-4A* and *TaGS3-7D* loci. *HAP-7D-1* and *HAP-7D-2* were discriminated using a previously reported co-dominant marker [20]. Association analysis showed that the TGW of *HAP-7D-2* haplotype was significantly higher than that of *HAP-7D-1* in all five environments ($P < 0.01$), indicating that *HAP-7D-2* was the favorable haplotype (Table S4). The TGW differences between the two haplotypes were 3.89, 3.90, 3.94, 3.49, and 5.27 g in 2002, 2005, 2006, 2010, and 2016, respectively. The two haplotypes were also significantly different in GL, with mean differences of 0.32, 0.31, 0.25, 0.25, and 0.32 mm, respectively, in the five environments ($P < 0.01$) (Table S4). *TaGS3-7D* was also associated with GW, with the two haplotypes having significant differences in four environments ($P < 0.05$) (Table S4). The mean differences were 0.11, 0.09, 0.08, and 0.09 mm, respectively.

Combinations of favorable haplotypes of *TaGS3-4A* and *TaGS3-7D* showed strong additive effects on TGW and GL. *HAP-4A-1/HAP-7D-2* (type 1) was the best combination in increasing TGW and GL. On the contrary, the combination of *HAP-4A-3/HAP-7D-1* (type 6) manifested a significant negative effect on TGW and GL in all five environments. TGW and GL for the favorable haplotype combination *HAP-4A-1/HAP-7D-2* in the MCC panel was 7.50 to 16.01 g higher and 0.80 to 0.98 mm longer than those for *HAP-4A-3/HAP-7D-1* across environments (Table 1). However, the combination *HAP-4A-1/HAP-7D-2* showed no association with GW.

3.9. Haplotype frequencies of *TaGS3-4A* and *TaGS3-7D* in the MCC panel

The frequencies of the *TaGS3* haplotypes in the MCC were determined. For *TaGS3-4A*, the combined frequency of superior *HAP-4A-1* and *HAP-4A-2* was 28.76% in landraces, 46.43% in modern cultivars and 45.46% in introduced accessions (Fig. S4). The frequencies in modern cultivars and introduced lines were 17.67% and 16.70% higher than in landraces, respectively. These results indicate that *TaGS3-4A* underwent strong selection in breeding. For *TaGS3-7D*, the highest frequency of *HAP-7D-2* was 36.36% in introduced lines, 19.32% in modern cultivars, and 10.07% in landraces (Fig. S4). This indicated that there was far less selection for *HAP-7D-2* in Chinese breeding programs.

The frequency of the best favorable haplotype combination (*HAP-4A-1/HAP-7D-2*, type 1) was 9.09% among introduced lines, considerably higher than in both Chinese cultivars and landraces (Fig. S4). These results imply that the pyramiding effect of *TaGS3-4A/TaGS3-7D* was not selected by Chinese breeding programs.

3.10. Favorable *TaGS3* haplotypes have potential for yield improvement globally

Favorable haplotypes associated with high yield traits are subject to selection in breeding. Wheat production in China is divided into ten ecological regions [25]. The distribution of *TaGS3-4A*, *TaGS3-7D* haplotypes and *TaGS3-4A/TaGS3-7D* combinations across regions were evaluated in landraces of MCC and Chinese modern cultivars (MC) from wheat ecological regions in China. For *TaGS3-4A*, the frequencies of *HAP-4A-1* and *HAP-4A-2* were generally higher in regions X, II, V, VII and IV with combined frequencies of 56.25%, 46.77%, 42.86%, 40.00%, and 34.55%, respectively. *HAP-4A-4* was relatively frequent (33.82%) in region III and *HAP-4A-5* was comparatively frequent in regions VI and IX (51.85% and 43.75%). The highest frequency of the inferior haplotype *HAP-4A-3* was around 40% in region VII and the lowest frequency was 2.94% in region III (Fig. S5-A). For *TaGS3-7D*, the highest frequency of the favorable *HAP-7D-2* was 33.33% in region X and the lowest frequency was 9.84% in region IV (Fig. S5-C). Furthermore, our results indicated that distribution of the favorable haplotype combination (type 1) was not as high as might be predicted from the single haplotype frequencies. The frequency of type 1 haplotype combination disappeared in eight (I, II, IV, V, VI, VIII, IX, and X) of the ten regions, and the highest frequency of *HAP-4A-1/HAP-7D-2* (type 1) was only 6.67% in region VII (Fig. S6).

Similar comparisons among materials from North America, CIMMYT, Europe, former USSR, Australia and China indicated that the combined frequencies of *HAP-4A-1* and *HAP-4A-2* were generally high in Europe, North America and China at 46.97%, 40.11%, and 36.80%, respectively (Fig. S5-B). For *TaGS3-7D*, the frequency of *HAP-7D-2* was generally high in Europe and North America at 63.87% and 48.74%, respectively (Fig. S5-D). For the best favorable haplotype combination *HAP-4A-1/HAP-7D-2* (type 1), the frequency was 14.20% in Europe and 7.65% in North America (Fig. 4). The low frequency of type 1 in the global wheat collections suggests potential

Table 1 – Effect of *TaGS3-4A*/*TaGS3-7D* haplotype combinations on grain traits in five environments.

Environment	Genotype	Thousand grain weight (g)	Grain length (mm)	Grain width (mm)
02LY	*HAP-4A-1/HAP-7D-2*	46.56 ± 4.50 a(A)	7.05 ± 0.42 a(A)	3.28 ± 0.11 a
	HAP-4A-1/HAP-7D-1	35.37 ± 2.28 bc(B)	6.34 ± 0.12 b(B)	3.09 ± 0.05 ab
	HAP-4A-2/HAP-7D-2	39.38 ± 1.40 bc(AB)	6.64 ± 0.11 ab(AB)	3.16 ± 0.08 ab
	HAP-4A-2/HAP-7D-1	39.55 ± 1.04 bc(AB)	6.54 ± 0.08 b(AB)	3.17 ± 0.04 ab
	HAP-4A-3/HAP-7D-2	36.04 ± 3.79 bc(B)	6.41 ± 0.13 b(AB)	2.92 ± 0.15 b
	HAP-4A-3/HAP-7D-1	33.75 ± 0.98 c(B)	6.25 ± 0.08 b(B)	3.02 ± 0.03 ab
	HAP-4A-4/HAP-7D-1	37.02 ± 1.47 bc(B)	6.37 ± 0.09 b(B)	3.14 ± 0.04 ab
	HAP-4A-5/HAP-7D-2	40.60 ± 2.72 b(AB)	6.73 ± 0.23 ab(AB)	3.24 ± 0.09 a
	HAP-4A-5/HAP-7D-1	36.06 ± 0.98 bc(B)	6.34 ± 0.08 b(B)	2.95 ± 0.05 b
05LY	*HAP-4A-1/HAP-7D-2*	38.75 ± 4.52 a	7.27 ± 0.36 a(A)	3.14 ± 0.09 a
	HAP-4A-1/HAP-7D-1	33.52 ± 2.34 ab	6.45 ± 0.13 b(B)	3.10 ± 0.10 a
	HAP-4A-2/HAP-7D-2	36.45 ± 1.30 ab	6.64 ± 0.11 b(B)	3.12 ± 0.05 a
	HAP-4A-2/HAP-7D-1	35.25 ± 1.01 ab	6.59 ± 0.08 b(B)	3.09 ± 0.03 a
	HAP-4A-3/HAP-7D-2	33.45 ± 2.80 ab	6.50 ± 0.11 b(B)	3.08 ± 0.12 a
	HAP-4A-3/HAP-7D-1	31.25 ± 0.90 b	6.29 ± 0.08 b(B)	3.05 ± 0.03 a
	HAP-4A-4/HAP-7D-1	32.97 ± 1.36 ab	6.34 ± 0.09 b(B)	3.06 ± 0.05 a
	HAP-4A-5/HAP-7D-2	37.54 ± 2.29 ab	6.70 ± 0.13 b(AB)	3.13 ± 0.10 a
	HAP-4A-5/HAP-7D-1	32.16 ± 0.86 b	6.43 ± 0.07 b(B)	2.96 ± 0.03 a
06LY	*HAP-4A-1/HAP-7D-2*	44.78 ± 3.66 a(A)	7.26 ± 0.30 a(A)	3.22 ± 0.06 a
	HAP-4A-1/HAP-7D-1	33.86 ± 1.81 c(B)	6.42 ± 0.12 b(B)	3.19 ± 0.08 a
	HAP-4A-2/HAP-7D-2	38.05 ± 1.33 bc(AB)	6.63 ± 0.10 b(B)	3.28 ± 0.04 a
	HAP-4A-2/HAP-7D-1	37.45 ± 0.94 bc(B)	6.60 ± 0.08 b(B)	3.23 ± 0.02 a
	HAP-4A-3/HAP-7D-2	36.46 ± 2.76 bc(B)	6.59 ± 0.14 b(B)	3.16 ± 0.12 a
	HAP-4A-3/HAP-7D-1	33.77 ± 0.79 c(B)	6.30 ± 0.08 b(B)	3.16 ± 0.03 a
	HAP-4A-4/HAP-7D-1	35.83 ± 1.31 bc(B)	6.41 ± 0.09 b(B)	3.15 ± 0.04 a
	HAP-4A-5/HAP-7D-2	40.86 ± 2.53 ab(AB)	6.66 ± 0.13 b(B)	3.21 ± 0.08 a
	HAP-4A-5/HAP-7D-1	34.77 ± 0.75 c(B)	6.49 ± 0.06 b(B)	3.13 ± 0.03 a
10SY	*HAP-4A-1/HAP-7D-2*	46.28 ± 3.59 a(A)	7.51 ± 0.34 a(A)	3.13 ± 0.07 a
	HAP-4A-1/HAP-7D-1	35.35 ± 1.67 bc(B)	6.59 ± 0.12 b(B)	3.17 ± 0.08 a
	HAP-4A-2/HAP-7D-2	36.14 ± 1.45 bc(B)	6.82 ± 0.10 b(B)	3.11 ± 0.07 a
	HAP-4A-2/HAP-7D-1	36.06 ± 1.08 bc(B)	6.81 ± 0.08 b(B)	3.10 ± 0.03 a
	HAP-4A-3/HAP-7D-2	34.32 ± 3.49 bc(B)	6.72 ± 0.16 b(B)	3.02 ± 0.21 a
	HAP-4A-3/HAP-7D-1	32.66 ± 0.83 c(B)	6.57 ± 0.08 b(B)	3.05 ± 0.02 a
	HAP-4A-4/HAP-7D-1	35.69 ± 1.30 bc(B)	6.62 ± 0.08 b(B)	3.06 ± 0.04 a
	HAP-4A-5/HAP-7D-2	40.26 ± 2.29 b(AB)	6.93 ± 0.16 b(B)	3.14 ± 0.07 a
	HAP-4A-5/HAP-7D-1	34.00 ± 0.83 bc(B)	6.69 ± 0.07 b(B)	3.04 ± 0.03 a
16LC	*HAP-4A-1/HAP-7D-2*	63.10 ± 3.45 a(A)	7.77 ± 0.28 a(A)	3.71 ± 0.08 a
	HAP-4A-1/HAP-7D-1	48.29 ± 2.25 b(B)	6.79 ± 0.14 b(B)	3.49 ± 0.07 ab
	HAP-4A-2/HAP-7D-2	52.81 ± 2.07 b(B)	7.10 ± 0.13 b(B)	3.61 ± 0.06 ab
	HAP-4A-2/HAP-7D-1	52.55 ± 1.32 b(B)	7.03 ± 0.09 b(B)	3.62 ± 0.03 ab
	HAP-4A-3/HAP-7D-2	50.98 ± 4.80 b(B)	7.05 ± 0.10 b(B)	3.51 ± 0.17 ab
	HAP-4A-3/HAP-7D-1	47.09 ± 1.18 b(B)	6.79 ± 0.10 b(B)	3.46 ± 0.03 b
	HAP-4A-4/HAP-7D-1	51.47 ± 1.65 b(B)	6.90 ± 0.08 b(B)	3.58 ± 0.04 ab
	HAP-4A-5/HAP-7D-2	55.19 ± 3.31 b(AB)	7.26 ± 0.20 b(AB)	3.62 ± 0.14 ab
	HAP-4A-5/HAP-7D-1	47.55 ± 1.24 b(B)	6.91 ± 0.07 b(B)	3.47 ± 0.03 ab

02LY, Luoyang (2002); 05LY, Luoyang (2005); 06LY, Luoyang (2006); 10SY, Shunyi (2010); 16LC, Luancheng (2016).
Different capital and small letters within groups indicate significance of differences between haplotypes at $P < 0.01$ and $P < 0.05$ for each trait, respectively.
HAP-4A-1/HAP-7D-2 ($n = 5$); *HAP-4A-1/HAP-7D-1* ($n = 11$); *HAP-4A-2/HAP-7D-2* ($n = 20$); *HAP-4A-2/HAP-7D-1* ($n = 55$); *HAP-4A-3/HAP-7D-2* ($n = 5$); *HAP-4A-3/HAP-7D-1* ($n = 46$); *HAP-4A-4/HAP-7D-2* ($n = 1$); *HAP-4A-4/HAP-7D-1* ($n = 31$); *HAP-4A-5/HAP-7D-2* ($n = 8$); *HAP-4A-5/HAP-7D-1* ($n = 69$).

worldwide for its selection to improve yield based on higher TGW and larger grain size.

4. Discussion

4.1. *TaGS3 homoeologs are orthologs of OsGS3*

The rice genome encodes three non-canonical Gγ subunits (GS3, DEP1, and OsGGC2) [26,27]. TaGS3 was the non-canonical Gγ subunit identified in wheat. Consistent with *OsGS3* gene structure, all of the three *TaGS3* homoeologous genes have five exons and four introns (Fig. 1). As mentioned in the previous reports [21], *TaGS3* homoeologs were located on chromosomes 7A, 4A and 7D.

The *OsGS3* was first identified as a major gene for grain weight and grain length [17]. It explained 80%–90% of the phenotypic variation in grain weight and length and acted as a negative regulator of grain size. Comparative sequencing analysis revealed that different nucleotide sequences of *OsGS3* in different varieties affected grain size. The nucleotide differences among varieties mainly occurred in exon regions

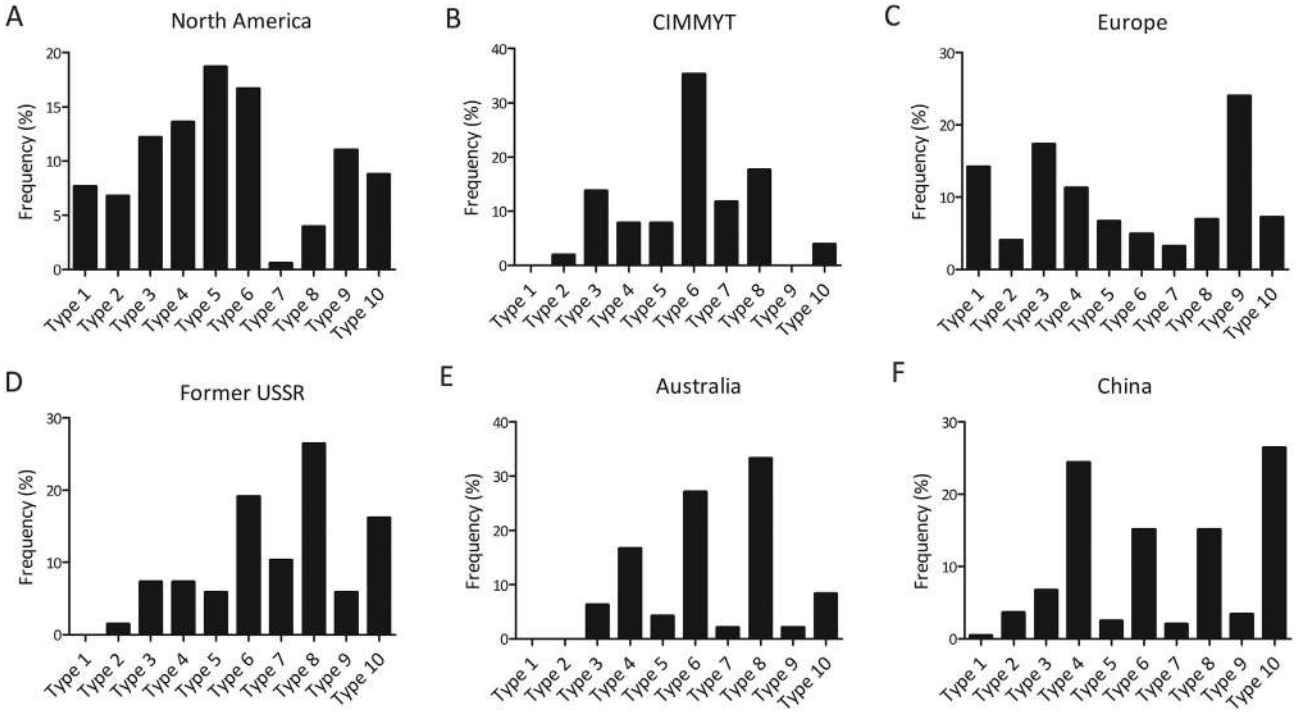


Fig. 4 – The distribution of *TaGS3-4A*/*TaGS3-7D* haplotype combination in global wheat cultivars. Type 1, *HAP-4A-1*/*HAP-7D-2*; Type 2, *HAP-4A-1*/*HAP-7D-1*; Type 3, *HAP-4A-2*/*HAP-7D-2*; Type 4, *HAP-4A-2*/*HAP-7D-1*; Type 5, *HAP-4A-3*/*HAP-7D-2*; Type 6, *HAP-4A-3*/*HAP-7D-1*; Type 7, *HAP-4A-4*/*HAP-7D-2*; Type 8, *HAP-4A-4*/*HAP-7D-1*; Type 9, *HAP-4A-5*/*HAP-7D-2*; Type 10, *HAP-4A-5*/*HAP-7D-1*. CIMMYT, International Maize and Wheat Improvement Center; Former USSR, Former Union of Soviet Socialist Republics.

and regulated grain traits by producing frameshift mutations that caused premature termination of transcription [18]. The *OsGS3-1* allele associated with medium grain length was present in the majority of widely grown *indica* varieties, and insertion of a 3 bp in the fifth exon did not change the grain length, in contrast to its effect in most temperate *japonica* varieties. However, one base pair deletion at the same site caused a frameshift and generated an allele conferring small grain. In contrast, premature termination of transcription caused by a single base substitution in the second exon greatly increased grain length [18].

Compared with *OsGS3*, only a few variations were found in the exon sequences of *TaGS3*, and sequence polymorphism mainly occurred in intron regions (Fig. S3). This phenomenon has also been found in some other wheat genes. The yield-related genes *TaDep1* and *TaSUS1-7B* had nucleotide sequence diversity in intron regions that influenced gene expression [24,28].

4.2. Additive effects of *TaGS3-4A* and *TaGS3-7D* on TGW

Favorable alleles at each locus affecting grain yield gave positive effects on phenotypic values. There were significant additive effects when favorable alleles were combined [29]. Bread wheat is an allohexaploid grass species that arose through hybridization of three related diploid grasses [30]. There are three possible evolutionary outcomes for homoeologous genes in polyploids: retention of the original or similar function, functional diversification, and gene silencing [31]. Thus, it is meaningful to clone homoeologs and distinguish their biological functions in polyploid plant species. Haplotype interaction of *TaGW2-6A* and *TaGW2-6B* showed additive effects between the favorable haplotypes and *Hap-6A-A*/*Hap-6B-1* was the best combination for increased TKW [32]. In the present study, real-time RT-PCR analysis showed that *TaGS3-4A* and *TaGS3-7D* were more highly expressed in developing grains than *TaGS3-7A* (Fig. S2), suggesting that *TaGS3-4A* and *TaGS3-7D* play more important roles in grain development. Molecular markers of *TaGS3-4A* and *TaGS3-7D* were therefore developed to facilitate selection of grain traits in breeding. Association analysis revealed that *HAP-4A-1*, *HAP-4A-2* and *HAP-7D-2* could individually increase TGW by an average 13.33%, 11.83%, and 10.91% in the five environments, respectively (Table S3, Table S4). These results indicated the potential of the favorable haplotypes to improve selected grain traits and thereby yield. The additive effects of *TaGS3-4A* and *TaGS3-7D* showed that *HAP-4A-1*/*HAP-7D-2* was a superior combination that could increase TGW by an average 34.5% across five environments (Table 1). Our study suggested that the most favorable haplotype combination was *HAP-4A-1*/*HAP-7D-2* and could be obtained by marker-assisted selection.

The geographic distribution of *TaGS3* haplotypes in the ten Chinese wheat ecological production regions showed that the

frequency of the favorable haplotypes of *TaGS3-4A* and *TaGS3-7D* were higher in region X (Fig. S5), and the favorable haplotype combination *HAP-4A-1/HAP-7D-2* was detected only in regions VII and III (Fig. S6). The imbalanced distribution suggested that combined selection of *HAP-4A-1/HAP-7D-2* had not been considered in the past breeding process. For example, in the Yellow-Huai River Valleys (region II), the major Chinese wheat production area, the frequency of *HAP-4A-1/HAP-7D-2* combination was zero compared to the individual frequencies of 1.61% and 14.70%, respectively. We assume that the favorable haplotype combination *HAP-4A-1/HAP-7D-2* confer too longer grain generating lemma and palea fail to close properly, which might cause pre-harvest sprouting. Thus, the highest frequency of *HAP-4A-1/HAP-7D-2* combination was detected in the drought northern spring wheat region (VII) (Fig. S6). This research indicates the scope for improving grain traits through marker-assisted selection and provides a useful example for understanding interaction of homeologous genes controlling complex traits in polyploid wheat.

4.3. *The favorable haplotypes of TaGS3 did not predominate in all six wheat world production regions*

TaGS3-4A and *TaGS3-7D* have undergone selection in wheat breeding (Fig. S4), but *HAP-4A-1*, *HAP-4A-2*, and *HAP-7D-2* were not predominating haplotypes of those genes in wheat cultivars (Fig. S5). Comprehensive analysis of *TaGS3-4A* and *TaGS3-7D* indicated that the type 1 haplotype combination had not undergone strong selection in wheat breeding. There may be three reasons for this. First, genetic results indicated that *TaGS3* is a negative regulator of grain weight and size (Fig. 2), and its expression level in grains was negatively related to TGW (Fig. S7). Thus, effective selection of the favorable haplotypes of *TaGS3* was lack in wheat breeding. The molecular markers developed in this study will be helpful for identifying the favorable haplotypes to improve grain traits in wheat. Second, selection of favorable *TaGS3* haplotypes is influenced by a balance of all yield-related genes during breeding. It was observed that there was no advantage in predicting yield and protein content in rye when many traits are considered for simultaneous improvement as additional traits could introduce issues in co-linearity [33]. Miralles and Slafer [34] reviewed the factors influencing grain yield and concluded that increased grain yield was associated with higher grain number but was also associated with a negative relationship between grain number and grain weight. Thus, we assume that favorable haplotypes of *TaGS3* might be neglected during selection of other yield-related traits in the early generations of breeding. Third, there are different levels of gene function regulation. Except for nucleotide sequence polymorphism, there are other avenues of regulation such as transcriptional regulation and transcript modification that may dominate development of grain traits. As a consequence, favorable *TaGS3* haplotypes might not have been effectively selected in past breeding programs. In fact, we found that alternative splicing (AS) of *TaGS3* does exist in wheat (Fig. S8). It has been reported as an important transcriptional regulatory mechanism for gene expression diversity in eukaryotes [35–37].

5. Conclusions

Expression, haplotypes, genetic effects and geographic distribution of *TaGS3* were analyzed. Transgenic results indicated that *TaGS3* was a negative regulator of grain weight and grain size. Two functional markers were developed to distinguish the five haplotypes of *TaGS3-4A* found in the test panels. Haplotypes *HAP-4A-1* and *HAP-4A-2* were associated with higher TGW and longer GL and therefore regarded as favorable. Cultivars carrying *HAP-7D-2* exhibited higher TGW and longer GL than those carrying *HAP-7D-1* in the MCC panel. The combination of favorable *TaGS3-4A* and *TaGS3-7D* haplotypes gave significant additive effects on TGW and GL. However, the favorable haplotypes *HAP-4A-1/2*, *HAP-7D-2* and the *HAP-4A-1/HAP-7D-2* combination occurred at lower frequencies in global wheat panels, thus generating opportunities for improving grain traits by marker-assisted selection. This research also provides a valuable example of understanding the interaction of alleles/haplotypes at homoeologous gene loci in control of complex traits in polyploid species.

Acknowledgments

We are grateful to Professor Robert McIntosh for his critical review of the manuscript. This work was jointly supported by Natural Science Foundation of Hebei Province (C2018207020), National Natural Science Foundation of China (31771784), Hebei Provincial Science and Technology Research and Development Project (16226320D) and China Agriculture Research System (CARS-03-01B).

Author contributions

Wei Zhang and Junming Li planned the research project. Wei Zhang, Huifang Li, Qiannan Su and Liya Zhi carried out the experiments. Xueyong Zhang, Wei Zhang, Huifang Li, Qiannan Su, Liya Zhi, Xiaoli Ren, Jiajia Liu, Deyuan Men, Na Zhang and Jun Ji collected phenotypic data. Wei Zhang wrote the manuscript. Xueyong Zhang and Junming Li revised the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cj.2020.03.003.

REFERENCES

[1] Q. Jiang, J. Hou, C. Hao, L. Wang, H. Ge, Y. Dong, X. Zhang, The wheat (*T. aestivum*) sucrose synthase 2 gene (*TaSus2*) active in endosperm development is associated with yield traits, Funct. Integr. Genom. 11 (2011) 49–61.

[2] L. Zhang, Y.L. Zhao, L.F. Gao, G.Y. Zhao, R.H. Zhou, B.S. Zhang, J.Z. Jia, *TaCKX6-D1*, the ortholog of rice *OsCKX2*, is associated with grain weight in hexaploid wheat, New Phytol. 195 (2012) 574–584.

[3] D. Ma, J. Yan, Z. He, W. Ling, X. Xia, Characterization of a cell wall invertase gene A1 on common wheat chromosome 2A and development of functional markers, Mol. Breed. 29 (2012) 43–52.
[4] Y. Jiang, Q. Jiang, C. Hao, J. Hou, L. Wang, H. Zhang, S. Zhang, X. Chen, X. Zhang, A yield-associated gene *TaCWI*, in wheat: its function, selection and evolution in global breeding revealed by haplotype analysis, Theor. Appl. Genet. 128 (2015) 131–143.
[5] J. Chang, J. Zhang, X. Mao, A. Li, J. Jia, R. Jing, Polymorphism of *TaSAP1-A1* and its association with agronomic traits in wheat, Planta 237 (2013) 1495–1508.
[6] M. Ma, Q. Wang, Z. Li, H. Cheng, Z. Li, X. Liu, W. Song, R. Appels, H. Zhao, Expression of *TaCYP78A3*, a gene encoding cytochrome P450 CYP78A3 protein in wheat (*Triticum aestivum* L.), affects seed size, Plant J. 83 (2015) 312–325.
[7] L. Ma, T. Li, C. Hao, Y. Wang, X. Chen, X. Zhang, *TaGS5-3A*, a grain size gene selected during wheat improvement for larger kernel and yield, Plant Biotechnol. J. 14 (2016) 1269–1280.
[8] S. Wang, X. Yan, Y. Wang, H. Liu, D. Cui, F. Chen, Haplotypes of the *TaGS5-A1* gene are associated with thousand-kernel weight in Chinese bread wheat, Front. Plant Sci. 7 (2016) 783.
[9] S. Wang, X. Zhang, F. Chen, D. Cui, A single-nucleotide polymorphism of *TaGS5* gene revealed its association with kernel weight in Chinese bread wheat, Front. Plant Sci. 6 (2015) 1166.
[10] M. Sajjad, X. Ma, S. Habibullah Khan, M. Shoaib, Y. Song, W. Yang, A. Zhang, D. Liu, *TaFlo2-A1*, an ortholog of rice *Flo2*, is associated with thousand grain weight in bread wheat (*Triticum aestivum* L.), BMC Plant Biol. 17 (2017) 164.
[11] J. Geng, L. Li, Q. Lv, Y. Zhao, Y. Liu, L. Zhang, X. Li, *TaGW2-6A* allelic variation contributes to grain size possibly by regulating the expression of cytokinins and starch-related genes in wheat, Planta 246 (2017) 1153–1163.
[12] Q. Li, L. Li, Y. Liu, Q. Lv, H. Zhang, J. Zhu, X. Li, Influence of *TaGW2-6A* on seed development in wheat by negatively regulating gibberellin synthesis, Plant Sci. 263 (2017) 226–235.
[13] H. Zhai, Z. Feng, X. Du, Y. Song, X. Liu, Z. Qi, L. Song, J. Li, L. Li, H. Peng, Z. Hu, Y. Yao, M. Xin, S. Xiao, Q. Sun, Z. Ni, A novel allele of *TaGW2-A1* is located in a finely mapped QTL that increases grain weight but decreases grain number in wheat (*Triticum aestivum* L.), Theor. Appl. Genet. 131 (2018) 539–553.
[14] Y. Zhang, D. Li, D. Zhang, X. Zhao, X. Cao, L. Dong, J. Liu, K. Chen, H. Zhang, C. Gao, D. Wang, Analysis of the functions of *TaGW2* homoeologs in wheat grain weight and protein content traits, Plant J. 94 (2018) 857–866.
[15] W. Wang, J. Simmonds, Q. Pan, D. Davidson, F. He, A. Battal, A. Akhunova, H.N. Trick, C. Uauy, E. Akhunov, Gene editing and mutagenesis reveal inter-cultivar differences and additivity in the contribution of *TaGW2* homoeologues to grain size and weight in wheat, Theor. Appl. Genet. 131 (2018) 2463–2475.
[16] Y. Wang, J. Hou, H. Liu, T. Li, K. Wang, C. Hao, H. Liu, X. Zhang, *TaBT1*, affecting starch synthesis and thousand kernel weight, underwent strong selection during wheat improvement, J. Exp. Bot. 70 (2019) 1513–1523.
[17] C. Fan, Y. Xing, H. Mao, T. Lu, B. Han, C. Xu, X. Li, Q. Zhang, *GS3*, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein, Theor. Appl. Genet. 112 (2006) 1164–1171.
[18] H. Mao, S. Sun, J. Yao, C. Wang, S. Yu, C. Xu, X. Li, Q. Zhang, Linking differential domain functions of the GS3 protein to natural variation of grain size in rice, Proc. Natl. Acad. Sci. U. S. A. 107 (2010) 19579–19584.
[19] N. Takano-Kai, K. Doi, A. Yoshimura, *GS3* participates in stigma exsertion as well as seed length in rice, Breed. Sci. 61 (2011) 244–250.
[20] Y. Zhang, J. Liu, X. Xia, Z. He, *TaGS-D1*, an ortholog of rice *OsGS3*, is associated with grain weight and grain length in common wheat, Mol. Breed. 34 (2014) 1097–1107.
[21] J. Yang, Y. Zhou, Y. Zhang, W. Hu, Q. Wu, Y. Chen, X. Wang, G. Guo, Z. Liu, T. Cao, H. Zhao, Cloning, characterization of *TaGS3* and identification of allelic variation associated with kernel traits in wheat (*Triticum aestivum* L.), BMC Genet. 20 (2019) 98.
[22] C. Hao, L. Wang, H. Ge, Y. Dong, X. Zhang, Genetic diversity and linkage disequilibrium in Chinese bread wheat (*Triticum aestivum* L.) revealed by SSR markers, PLoS One 6 (2011), e17279. .
[23] J. Hou, T. Li, Y. Wang, C. Hao, H. Liu, X. Zhang, ADP-glucose pyrophosphorylase genes, associated with kernel weight, underwent selection during wheat domestication and breeding, Plant Biotechnol. J. 15 (2017) 1533–1543.
[24] J. Hou, Q. Jiang, C. Hao, Y. Wang, H. Zhang, X. Zhang, Global selection on sucrose synthase haplotypes during a century of wheat breeding, Plant Physiol. 164 (2014) 1918–1929.
[25] X.Y. Zhang, C.W. Li, L.F. Wang, H.M. Wang, G.X. You, Y.S. Dong, An estimation of the minimum number of SSR alleles needed to reveal genetic relationships in wheat varieties. I. Information from large-scale planted varieties and cornerstone breeding parents in Chinese wheat improvement and production, Theor. Appl. Genet. 106 (2002) 112–117.
[26] S. Sun, L. Wang, H. Mao, L. Shao, X. Li, J. Xiao, Y. Ouyang, Q. Zhang, A G-protein pathway determines grain size in rice, Nat. Commun. 9 (2018) 851.
[27] T. Yuri, C. David, B.J. Ramón, Diversity of heterotrimeric G-protein γ subunits in plants, BMC Res. Notes 5 (2012) 608.
[28] Y. Liu, X. Xia, Z. He, Characterization of dense and erect panicle 1 gene (*TaDep1*) located on common wheat group 5 chromosomes and development of allele-specific markers, Acta Agron. Sin. 39 (2013) 589–598 (in Chinese with English abstract).
[29] F. Li, W. Wen, J. Liu, Y. Zhang, S. Cao, Z. He, A. Rasheed, H. Jin, C. Zhang, J. Yan, P. Zhang, Y. Wan, X. Xia, Genetic architecture of grain yield in bread wheat based on genome-wide association studies, BMC Plant Biol. 19 (2019) 168.
[30] The International Wheat Genome Sequencing Consortium (IWGSC), A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. Ancient Hybridizations Among the Ancestral Genomes of Bread Wheat 345 (2014) 1251788.
[31] J.F. Wendel, Genome evolution in polyploids, Plant Mol. Biol. 42 (2000) 225–249.
[32] L. Qin, C. Hao, J. Hou, Y. Wang, T. Li, L. Wang, Z. Ma, X. Zhang, Homologous haplotypes, expression, genetic effects and geographic distribution of the wheat yield gene *TaGW2*, BMC Plant Biol. 14 (2014) 107.
[33] A.W. Schulthess, Y. Wang, T. Miedaner, P. Wilde, J.C. Reif, Y. Zhao, Multiple-trait- and selection indices-genomic predictions for grain yield and protein content in rye for feeding purposes, Theor. Appl. Genet. 129 (2016) 273–287.
[34] D.J. Miralles, G.A. Slafer, Sink limitations to yield in wheat: how could it be reduced? J. Agric. Sci. 145 (2007) 139–149.
[35] W.B. Barbazuk, Y. Fu, K.M. McGinnis, Genome-wide analyses of alternative splicing in plants: opportunities and challenges, Genome Res. 18 (2008) 1382–1391.
[36] B.B. Wang, V. Brendel, Genomewide comparative analysis of alternative splicing in plants, Proc. Natl. Acad. Sci. U. S. A. 103 (2006) 7175–7180.
[37] M. Warnefors, V. Pereira, A. Eyre-Walker, Transposable elements: insertion pattern and impact on gene expression evolution in hominids, Mol. Biol. Evol. 27 (2010) 1955–1962.