Theor Appl Genet (2011) 122:211–223
DOI 10.1007/s00122-010-1437-z

ORIGINAL PAPER

Identification and development of a functional marker of *TaGW2* associated with grain weight in bread wheat (*Triticum aestivum* L.)

Zhenqi Su · Chenyang Hao · Lanfen Wang · Yuchen Dong · Xueyong Zhang

Received: 28 March 2010 / Accepted: 25 August 2010 / Published online: 14 September 2010
© Springer-Verlag 2010

Abstract The *OsGW2* gene is involved in rice grain development, influencing grain width and weight. Its ortholog in wheat, *TaGW2*, was considered as a candidate gene related to grain development. We found that *TaGW2* is constitutively expressed, with three orthologs expressing simultaneously. The coding sequence (CDS) of *TaGW2* is 1,275 bp encoding a protein with 424 amino acids, and has a functional domain shared with *OsGW2*. No divergence was detected within the CDS sequences in the same locus in ten varieties. Genome-specific primers were designed based on the sequence divergence of the promoter regions in the three orthologous genes, and *TaGW2* was located in homologous group 6 chromosomes through CS nulli-tetrasomic (NT). Two SNPs were detected in the promoter region of *TaGW2-6A*, forming two haplotypes: Hap-6A-A (−593A and −739G) and Hap-6A-G (−593G and −769A). A cleaved amplified polymorphic sequence (CAPS) marker was developed based on the −593 A-G polymorphism to distinguish the two haplotypes in *TaGW2-6A*. This gene was fine mapped 0.6 cM from marker *cfd*80.2 near the centromere in a recombinant inbred line (RIL) population. Two hundred sixty-five Chinese wheat varieties were genotyped and association analysis revealed that Hap-6A-A was significantly associated with wider grains and higher one-thousand grain weight (TGW) in two crop seasons. qRT-PCR revealed a negative relationship between *TaGW2* expression level and grain width. The Hap-6A-A frequencies in Chinese varieties released at different periods showed that it had been strongly positively selected in breeding. In landraces, Hap-6A-A is mainly distributed in southern Chinese wheat regions. Association analysis also indicated that Hap-6A-A not only increased TGW by more than 3 g, but also had earlier heading and maturity. In contrast to Chinese varieties, Hap-6A-G was the predominant haplotype in European varieties; Hap-6A-A was mainly present in varieties released in the former Yugoslavia, Italy, Bulgaria, Hungary and Portugal.

Communicated by A. Graner.

Z. Su and C. Hao contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-010-1437-z) contains supplementary material, which is available to authorized users.

Z. Su · C. Hao · L. Wang · Y. Dong · X. Zhang (✉)
Key Laboratory of Crop Germplasm Resources and Utilization, Ministry of Agriculture, Chinese Academy of Agricultural Sciences, Beijing 100081, China
e-mail: xueyongz@caas.net.cn

Z. Su · C. Hao · L. Wang · Y. Dong · X. Zhang
The National Key Facility for Crop Gene Resources and Genetic Improvement, Chinese Academy of Agricultural Sciences, Beijing 100081, China

Z. Su · C. Hao · L. Wang · Y. Dong · X. Zhang
Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing 100081, China

Introduction

Bread wheat (*Triticum aestivum* L.) is one of the most important global staple foods. With an increasing world population, it is estimated that the global demand for wheat will increase by a further 40% before 2020 (Rajaram 2005). Therefore, higher yield is a predominating objective in wheat breeding programs. Grain size is a major component of grain yield in wheat. Larger grains not only directly relate to grain yield but also have favorable effects on seeding vigor and early growth, thereby promoting and

stabilizing yielding ability. Large grain size has been an important trait selected during domestication and modern wheat breeding (Botwright et al. 2002; Peng et al. 2003). It is usually represented in plant breeding practice by one-thousand grain weight (TGW), mainly determined by grain width (GW), grain length (GL) and grain thickness (GT) (Campbell et al. 1999; Dholakia et al. 2003). All three aspects are positively correlated with TGW (Breseghello and Sorrells 2006; Sun et al. 2009). Therefore, increasing any component of grain size can improve grain weight in a wheat breeding program.

Grain size in wheat is a quantitative trait controlled by quantitative trait loci (QTL), and numerous QTLs for grain size have been reported (Campbell et al. 2003; Dholakia et al. 2003; Huang et al. 2006; Kato et al. 2000; Röder et al. 2008; Sun et al. 2009). No gene associated with grain size has been isolated, and no ideal marker for grain weight has been developed for molecular-assisted selection.

Cloning yield-related genes, exploiting favorable allelic variation and developing functional markers are objectives for tailor-made improvement and marker-assisted selection. However, cloning genes directly from wheat by map-based cloning is very difficult due to its large hexaploid ($2n = 6x = 42$) genome and no whole genome sequence is available. Comparative genomics has demonstrated that the linear order of genetic markers and genes is well conserved among different grass genomes (e.g., rice, wheat, barley, maize, millet, and sorghum), providing a powerful tool for gene discovery in wheat (Choi et al. 2004; Fulton et al. 2002; Gale and Devos 1998; Gupta et al. 2008; Moore et al. 1995; Sorrells et al. 2003). Rice, barley and maize genome sequence have been used to obtain markers for high-resolution mapping, candidate gene identification and homology-based cloning in wheat (Bagge et al. 2007; He et al. 2008). Orthologs descended from a common ancestor often have conserved functions and are expected to produce similar phenotypes across species (Devos 2005). For example, comparative mapping of QTL controlling seed weight in rice, maize, and sorghum suggested that orthologous genes for seed size might be associated with domestication in these three crops (Paterson 1995). The *GAI*, *Rht-1* and *D8* orthologous dwarfing genes reduce plant height in Arabidopsis, wheat and maize, respectively (Peng et al. 1999). He et al. (2008) reported that the wheat orthologs of *Psy1* associated with grain yellow pigment content in wheat shared similar characteristics with maize. Li et al. (2010) showed that *ZmGS3* involved grain development in maize in a similar manner to the *GS3* gene in rice (Fan et al. 2006).

In rice, a model plant in the grass family, three major QTLs associated with grain size have been identified. *GS3* is a major QTL for grain length and weight (Fan et al. 2006). *GW2* on the short arm of chromosome 2 was identified from a QTL that controls grain width and weight (Song et al. 2007). *GW5/qSW5* on chromosome 5 regulating grain width and weight was isolated recently (Shomura et al. 2008; Weng et al. 2008). These identified grain-size genes in rice provide opportunities for cloning orthologous candidate genes in wheat. The function and mechanism of *GW2* controlling grain size was comprehensively studied by Song et al. (2007), who found that *GW2* encodes a RING-type protein with E3 ubiquitin ligase activity that negatively regulates grain width through control of cell division in the spikelet hull. Loss-of-function mutations in the coding sequence, or interference with the expression level of *GW2,* resulted in enhanced grain width, weight and yield (Matsuoka and Ashikari 2007). *GW2* was therefore selected as a candidate gene for isolating the wheat homolog(s). The coding sequence and promoter region of *TaGW2*, as the most important parts of the gene, were selected as the starting points.

The objectives of this study were to (1) isolate and characterize the coding and promoter regions of *TaGW2*, (2) identify sequence polymorphisms among varieties with different grain size in order to determine whether *TaGW2* has a similar function to *OsGW2* in rice, (3) detect superior alleles by association analysis and develop a functional marker for *TaGW2* to enable molecular-assisted selection in wheat, and (4) determine the distribution of functional alleles in varieties released in different years and geographical environments in China and Europe.

Materials and methods

Plant materials

A set of Chinese Spring (CS) nullisomic–tetrasomic lines and ditelosomic 6AL were used for chromosomal location of *TaGW2*.

The 428 wheat varieties used included 151 landraces and 277 modern varieties, among which 265 accessions (151 landraces and 114 modern varieties) were used for functional validation of the *TaGW2* marker. Two hundred and twenty-seven accessions (all above landraces and 76 modern varieties) among the 265 were sampled from the Chinese mini core collection representing more than 70% of the genetic diversity in Chinese wheat germplasm resources (Hao et al. 2008; Zhang et al. 2002) (Table S1). The years of release of all modern varieties are provided in Hao et al. (2006).

Three hundred and seventy-four European wheat accessions (Table S2) were obtained from the Clermont-Ferrand Genetic Resources Center (http://www.clermont.inra.fr/umr-asp) based on a project of advanced research (PRA 005) between the French National Institute for

Agronomical Research (INRA) and the Chinese Academy of Agricultural Science (CAAS). Because most of the European varieties did not mature normally at Luoyang Experimental Station, haplotype/yield-trait association analysis could not be undertaken.

Measurements of grain traits

During the 2001–2002 and 2005–2006 wheat-growing seasons, the varieties were planted at the CAAS Luoyang Experiment Station in Henan Province (111.6°E, 33.8°N). Each variety was planted in double 2-m rows spaced 25 cm apart, with 40 seeds in each row. The field management followed local practices. After harvest, 20 grains were randomly selected from each cultivar and lined up length-wise along a ruler to measure average grain length (GL), and then arranged breadth-wise to measure grain width (GW). The middle parts of ten grains were measured with vernier calipers to establish average grain thickness (GT). Two independent samples of 250 grains were weighted and the means were converted to one-thousand grain weight (TGW).

DNA extraction, primer design, PCR and sequencing

Genomic DNA was extracted from young leaves of ten days seedlings using the phenolchloroform method (Sharp et al. 1988). Primers were designed by the software Primer Premier Version 5.0 (Premier Biosoft International, Palo Alto, CA), and all primers were synthesized by Beijing Augct Biological Technology Co., Ltd (http://www.augct.com).

PCR reactions were performed in total volumes of 15 μl, including 3 pmol of each primer, 120 μM of each dNTP, 80 ng genomic DNA, 0.75 unit *LaTaq* and 7.5 μl 2× GC buffer [TaKaRa Biotechnology (Dalian) Co. Ltd, Product Code: DRR20AG]. The PCR procedure was 95°C for 3 min, followed by 32 cycles of 95°C for 30 s, annealing (55–60°C) for 30 s, and extension at 72°C (30 s to 3 min), with a final extension of 72°C for 10 min. The annealing temperatures and extension times depended on the primer sets and the lengths of the expected PCR products. The PCR products were separated by electrophoresis in agarose gels, and the target bands were recovered and cloned into the pEASY-T1 simple vector and transformed to DH5α competent *E. coli* cells by the heat shock method (Beijing TransGen Biotech Co., Ltd Product Code: CT111). Positive clones were selected for sequencing by ABI 3730XI DNA Analyzer. To guarantee sequence accuracy, the PCR reactions and DNA sequencing were repeated at least three times.

Isolation of cDNA and the promoter region of *TaGW2*

The cDNA sequence of *OsGW2* (GenBank: Accession EF447275.1) was used for a blast search against wheat EST in GenBank. All ESTs with high similarity to *OsGW2* cDNA were assembled to a putative *TaGW2* cDNA using the SeqMan program of the DNAStar software package (DNAStar Lasergene 7.1). The methods of RNA extraction and reverse transcription were similar to Guo et al. (2010). RNA samples were extracted using trizol reagent from immature seeds 5–25 days after flowering, seedling leaves, flag leaves and young ears. The cDNA first strand was synthesized using M-MLV transcriptase (Invitrogen) according to the manufacturer's instructions. Then, 1 μl of reverse transcript product was used for PCR to clone the putative *GW2* wheat cDNA.

To obtain the promoter region of *TaGW2*, an Aibai-Chinese Spring bacterial artificial chromosome (BAC) library (Kong et al., unpublished) was screened by the PCR-based method. The primer of BAC screening was designed according to the *TaGW2* cDNA sequence and *OsGW2* (to avoid the exon–intron junctions). The primer set was tested on CS genomic DNA before BAC screening, and the product was sequenced to confirm its specificity. The DNA of selected single BAC clones was isolated with a Large Construct Kit (Qiagen) for direct sequencing by primer walking to obtain the 5′ flanking promoter sequence of *TaGW2*. The promoter elements were identified using the TSSP program (http://www.softberry.com). According to BAC sequencing results, primers were designed for cloning the *TaGW2* promoter.

Real-time quantitative reverse transcription PCR for *TaGW2*

mRNA from immature seeds (10 days after flowering) were used for *TaGW2* expression analysis. DNA was removed by digestion with DNAaseI (Fermentas) before reverse transcription. The method of cDNA first strand synthesis was as described. The expression analysis of *TaGW2* was performed with SYBR Premix ExTaq (TaKaRa Biotechnology (Dalian) Co. Ltd, Product Code: DRR041A), and experiments were performed according to the manufacturer's instructions. The primer sets TaGW2-4 and TaGW2-5 (Table 1) were used for amplification of *TaGW2* and the actin gene, respectively. Three replications were performed to get the average and standard deviation of expression level. The relative qualification $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen 2001) were used to calculate *TaGW2* expression levels with actin gene as endogenous control and Zhongyou 9507 as a reference variety.

Table 1 Plant materials used for determining sequence differences in *TaGW2*

Type	Accession no.	Name	GW (mm)	TGW (g)	Genotype
Group I	ZM000215	S1: Baimangmai	2.98	32.95	Hap-6A-G
Slim-grain varieties	ZM013873	S2: Jimai 19	3.15	34.04	Hap-6A-G
	ZM005188	S3: Hongdongmai	2.88	29.68	Hap-6A-G
	ZM004422	S4: Mahuaban	2.68	23.06	Hap-6A-G
	ZM002685	S5: Sanyuehuang	2.98	25.69	Hap-6A-G
Group II	ZM022727	W1: Laizhou 953	3.53	50.93	Hap-6A-A
Wide-grain varieties	ZM025358	W2: Zhongyou 9507	3.50	54.21	Hap-6A-A
		W3: Xuzhou 22	3.68	53.82	Hap-6A-A
	ZM022308	W4: Lankao 906	3.44	49.76	Hap-6A-A
		W5: Zhengmai 9023	3.11	43.15	Hap-6A-A

Identification and development of a functional marker

Ten varieties including five with high and five with low grain weight and width were initially chosen for detecting sequence differences in the *TaGW2* cDNA and promoter regions (Table 1; Fig. 4a). The detected diversities among these varieties were assumed to be associated with yield-related traits, and the functions of these polymorphisms were validated on 265 Chinese wheat varieties by association analysis between genotypes (CAPS maker of *TaGW2*) and phenotypes.

The divergences in *TaGW2* were transformed to cleaved amplified polymorphism sequence (CAPS) markers which can be easily detected by agarose gel electrophoresis. In order to obtain high-quality PCR products for digestion and to avoid interference between the homologous sequences from orthologs in hexaploid wheat, the CAPS were performed as follows: the first step was to amplify the genome-specific *TaGW2* allele in different genotypes, then to obtain a smaller target fragment containing the polymorphic site through a second PCR. Finally, the products of the second PCR were digested with *Taq*I (Fermentas) according to the manufacturer's directions and the digested segments were separated on 2% agarose gels with EB.

Results

Cloning and characterization of *TaGW2* cDNA

The NCBI database search showed that seven wheat ESTs and one barley EST were very similar to *OsGW2* cDNA. These ESTs were assembled as a putative *TaGW2* cDNA. The primer set named TaGW2-1 (Table 2) was designed according to the putative sequence, and a 1,275-bp fragment was amplified in CS cDNA of immature grain and seedling leaves. Further analysis demonstrated that there were three very similar sequences in the PCR products, and the similarity of each sequence to putative *TaGW2* was more than 98%. The coding sequences (CDS) of the *TaGW2* homologs to *OsGW2* (GenBank: Accession EF447275.1) were ~87%, and the identities of their deduced amino acid sequences were ~88%. Although we could not discriminate the genome specificities of the three genes based on the cDNA sequences alone, there were clearly three sequences of *TaGW2* cDNA, suggesting simultaneous expression of *TaGW2* orthologs in the A, B and D genomes.

Each *TaGW2* cDNA contained one ORF, and was predicted to encode 424 amino acids with a molecular mass of ~47.2 kDa. The predicted protein sequences of the *TaGW2* genes showed that each possessed a RING-domain of 43 amino acids in the N terminus, similar to the *GW2* protein in rice (Song et al. 2007). *TaGW2* expression was detected in immature seed (5–25 days after flowering), seedling leaves, flag leaves and young ears (at 10 and 40 mm) (Fig. 1), and matching ESTs were found in the NCBI database derived from tissues, such as root, dormant embryo and crown. This suggested that the *TaGW2* genes were constitutively expressed, consistent with *OsGW2*. Based on the DNA sequence and predicted protein of *TaGW2*, we concluded that the *TaGW2* were orthologs of *OsGW2*.

Isolation of the promoter region of *TaGW2* in Chinese Spring

Primer set TaGW2-2 (Table 2) was designed and used for screening the Aibai-Chinese Spring BAC library. The PCR product was ~2.7 kb. Four positive BAC pools were identified, and BAC 345F12 which contains 3,000 clones was selected to isolate a single clone of *TaGW2*. To obtain the 5′ flanking sequence of *TaGW2*, primers PR-31, PR-469, PR-804, PR-1260 and PR-1772 were designed for genome walking (Table 2), and about 2.2 kb of upstream coding sequences were acquired. The core elements of the promoter were predicted with the TSSP program (http://www.softberry.com), and the TATA box was identified at

Table 2 Primers used in this study

Primer set	Primer sequence (5′–3′)	Amplified target	Size of PCR product (bp)[a]
TaGW2-1	Forward: ATGGGGAACAGAATAGGAGGGAGGA	*TaGW2*	1,275
	Reverse: TTACAACCATGCCAACCCTTGCGTG		
TaGW2-2	Forward: ATGGGGAACAGAATAGGAGGGAGGA	BAC screening	2,732
	Reverse: CGAGTATGCCTAGAATGGAAAGAC		
TaGW2-3	Forward: CGTTACCTCTGGTTTGGGTGTCGTG	Promoter region	1,635
	Reverse: GCGGCACTCTACGGCAGAACAAAT		
Hap-6A-P1	Forward: CGTTACCTCTGGTTTGGGTGTCGTG	A genome-specific	949
	Reverse: CACCTCTCGAAAATCTTCCCAATTA		
TaGW2-6B	Forward: GTGGTGAACATAGCAAATTGATTACAT	B genome-specific	1,275
	Reverse: TTGCGTAGCTTCTTCTGGTCGATAT		
TaGW2-6D	Forward: AAAAATTGATGAGGAAAGGACATCATACA	D genome-specific	751
	Reverse: TGCGTAGCTTCTTCTGGTCGATATCCA		
Hap-6A-P2	Forward: GAGAAAGGGCTGGTGCTATGGA	The 2nd time PCR for CAPS	418
	Reverse: GTAACGCTTGATAAACATAGGTAAT		
TaGW2-4	Forward: GCAGAACAATCGCTCCAACA	*TaGW2* real-time PCR	
	Reverse: GCCAAATCGCTTCCATAACC		
TaGW2-5	Forward: CACTGGAATGGTCAAGGCTG	Internal control for *TaGW2* real-time PCR	
	Reverse: CTCCATGTCATCCCAGTTG		
	PR-31: CCTTCCTCCTCCCTCCTATTCTG		
	PR-469: GATAAACATAGGTAATGCTTTCGTA		
	PR-804: CGCTTCCCTCGTCACTGG		
	PR-1260: CCAGAGGTAACGTTTTTCATGACT		
	PR-1772: GGGCTTTACAAATGACACCAACA		

[a] The sizes of PCR fragments were relative to the *TaGW2-6A* sequence

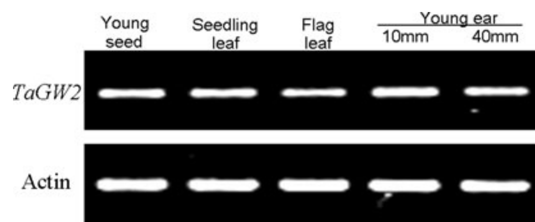

Fig. 1 Expression of *TaGW2* analyzed by RT-PCR in different organs in wheat, *actin* was used as a control

−173 and the start transcription site at −141 from the ATG initiation codon (Fig. 4b).

A series of primers to amplify the promoter region in wheat was designed according to the BAC sequences. Among them, primer set *TaGW2*-3 (Table 2) covering about 1.2 kb of upstream sequences and the first exon (198 bp) of *TaGW2* were selected. Sequence analysis further proved three different sequences in the PCR products, indicating that primer set TaGW2-3 simultaneously amplified the three orthologous *TaGW2* promoter sequences. Compared with the *TaGW2* CDS, more nucleotide substitutions and insertions/deletions existed in the promoter regions in the three orthologous sequences. This formed a basis for designing genome-specific primers for chromosomally locating the *TaGW2* genes in wheat.

Chromosomal locations of *TaGW2*

Based on sequence differences in the promoter regions, the genome-specific primer pairs, Hap-6A-P1, TaGW2-6B and TaGW2-6D (Table 2) were designed for chromosome location using the CS nulli-tetrasomic lines. *TaGW2* were located on the homoeologous group 6 chromosomes. No PCR product was detected using primer set Hap-6A-P1 to amplify genomic DNA from CS ditelosomic 6AL, indicating that *TaGW2-6A* was located on the short arm of homologous group 6 (Fig. 2).

The *TaGW2* upstream sequence amplified from BAC345F12BAC was identical to that from chromosome 6A, demonstrating that it was from chromosome 6A. Primers based on variation in the promoter and first exon regions clearly discriminated the three orthologous *TaGW2*

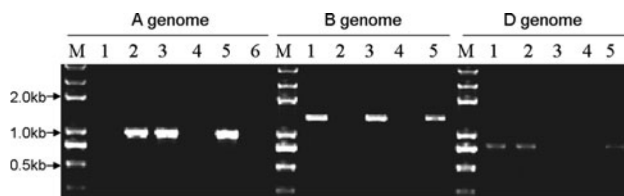

Fig. 2 PCR amplification of CS homoeologous group 6 nullisomic-tetrasomic lines and ditelosomic line 6AL with the genome-specific primer sets Hap-6A-P1, TaGW2-6B and TaGW2-6D. *M* DNA ladder, *1* CS nullisomic 6A-tetrasomic 6B (N6A-T6D), *2* N6B-T6D, *3* N6D-T6B, *4* H_2O, *5* CS, *6* CS ditelosomic 6AL

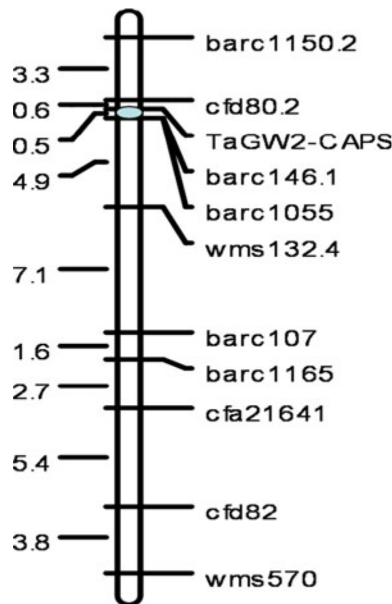

Fig. 3 Fine mapping of *TaGW2-6A* based on a recombinant inbred line (RIL) population (Xiaoyan 54 × Jing 411)

(Fig. 2). Furthermore, using the recombinant inbred line (RIL) population (Xiaoyan 54 × Jing 411) the 6A gene was mapped to a position about 0.6 cM from marker *cfd*80.2 which is very close to the centromere (Fig. 3).

Identification of sequence diversity in *TaGW2*

In rice, a 1-bp deletion in exon 4 led to a null mutation of *GW2* and resulted in increased grain width and weight (Song et al. 2007). However, no similar mutation was detected in wheat. There were, however, some nucleotide substitutions between the three orthologous sequences of *TaGW2,* the sequences for the same chromosomes were completely conserved between varieties with variable grain widths. These results implied that the grain width (weight) variation was not attributable to sequence mutations in the CDS of *TaGW2*, and that the mechanism affecting grain width and weight possibly differed from that in rice.

Comparison of the upstream 1.2 kb sequences revealed that genotypes with wide grains were −593 (A) and −739 (G), whereas those with slim grains were −593 (G) and −739 (A) (Fig. 4b; Fig. S2). The two SNPs formed a typical haplotype in the promoter region of *TaGW2-6A*, and, accordingly, the two alleles of *TaGW2-6A* were designated as Hap-6A-A and Hap-6A-G, respectively. To check for other alleles in the promoter region of *TaGW2-6A*, a genome-specific primer set Hap-6A-P1 covering the SNP region was used for amplifying DNA from 227 accessions in the mini core collection (Table S1). Sequencing of the PCR products revealed only the above two *TaGW2-6A* alleles.

Development of CAPS marker for *TaGW2-6A*

The nucleotide diversities at *TaGW2-6A* produced a restriction enzyme *TaqI* recognition site (TCGA) in wide-grain genotypes (Hap-6A-A) at SNP-593-A, but not in (Hap-6A-G) SNP-593G (TCGG) in the slim-grain genotypes (Fig. 4b, c). This SNP provided an opportunity to develop a cleaved amplified polymorphism sequence (CAPS) marker to distinguish the *TaGW2-6A* alleles.

In order to discriminate the orthologous wheat sequences, genomic-specific primer set Hap-6A-P1 was used firstly to amplify a 949-bp fragment of *TaGW2-6A* from all cultivars. However, more than five *Taq*I recognition sites within the amplified fragment made it unsuitable for direct digestion by *Taq*I. Taking into account the need for obtaining high-quality DNA for digestion and to avoid interference from unrelated digested fragment, primer pair Hap-6A-P2 was designed for a second PCR (Table 2). After the 418bp product from the second PCR was digested by *TaqI*, a length polymorphism 167- vs 218-bp was generated in wide and slim varieties, respectively, which could be easily distinguished on agarose gels (Fig. 4c, d).

Association of *TaGW2-6A* haplotypes with grain traits

The Chinese common wheat core collection consists of landraces and modern variety subpopulations (Hao et al. 2008; Zhang et al. 2002). The traits TGW, GL, GW and GT were all significantly higher in modern varieties than in landraces, indicating their likely selection during modern wheat breeding (after 1949) (Table 3). Grain trait data for a total of 265 accessions were used for association analysis (Table 4). There were significant differences ($P < 0.01$) in GW, TGW and GL/GW, but not in GL between Hap-6A-A and Hap-6A-G accessions. An effect of *TaGW2-6A* on GT was detected in 2002 ($P < 0.01$), but not in 2006 ($P = 0.726$). This suggested that the main contributions of *TaGW2-6A* were to GW and TGW. Hap-6A-A accessions

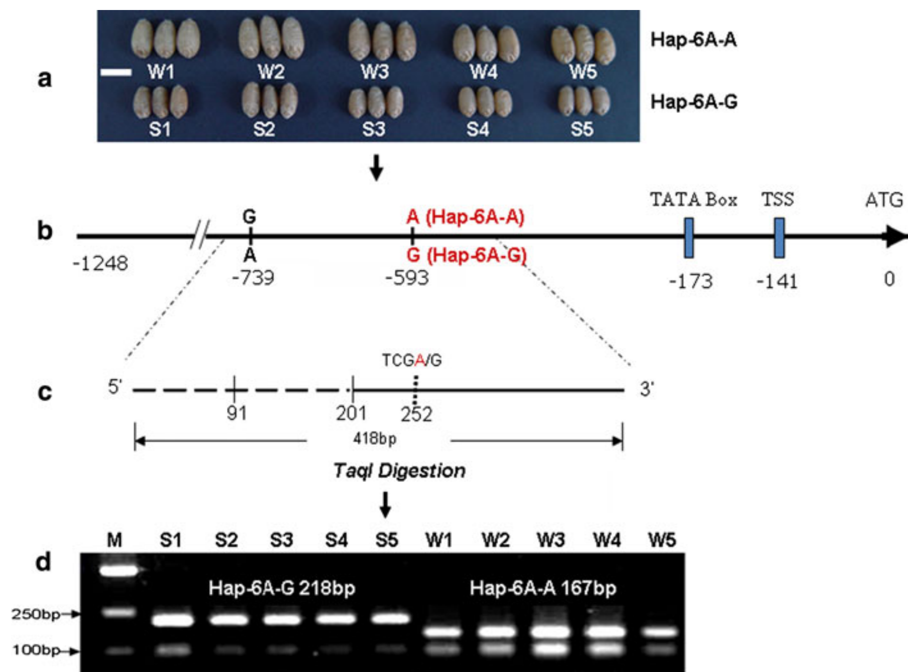

Fig. 4 Functional marker development in the promoter region of *TaGW2-6A* between varieties with different grain widths and weights. **a** Varieties with different grain widths and weights were used for identification of sequence differences in *TaGW2*. Detail information of these varieties is listed in the Table 1. *Bar* 5 mm. **b** The positions of Hap-6A-A and Hap-6A-G SNPs in the promoter region of *TaGW2-6A*. The predicted core promoter element (*TATA Box*), transcription start site (*TSS*) and the start code (*ATG*) are indicated. **c** SNP252 (A/G) in the second PCR CAPS product permitted generation of different *TaqI* restriction fragments, 167 and 218 bp in the varieties with Hap-6A-A and Hap-6A-G, respectively. Other *TaqI* restriction sites in the second PCR product of CAPS are indicated (91 and 201). **d** Validation of CAPS in varieties with Hap-6A-A and Hap-6A-G on 2% agarose gel. *M* marker, *S1–S5* and *W1–W5* are varieties with contrasting grain sizes listed in Table 1

Table 3 Comparison of grain traits between landraces and modern varieties (mean ± SD)

Subpopulation	No. of varieties	TGW (g)	GL (mm)	GW (mm)	GT (mm)
Landraces	151	32.96 ± 5.70	6.38 ± 0.49	3.05 ± 0.16	2.77 ± 0.17
Modern varieties	114	41.37 ± 5.97	6.74 ± 0.47	3.30 ± 0.18	2.93 ± 0.15
P values (*t* test)		<0.001	<0.001	<0.001	<0.001

Means of grain traits were based on of 2 years of data

possessed higher mean grain widths and weights than Hap-6A-G accessions.

Association analysis of the landrace subpopulation indicated no significant differences in GL and TGW between Hap-6A-G and Hap-6A-A accessions. Although differences in GW, GT and GL/GW ratio reached significance levels ($P = 0.05$), the effects were not the same in different years. However, differences between Hap-6A-A and Hap-6A-G in the subset of modern varieties were significant for GW ($P < 0.001$ in both years) and TGW ($P < 0.01$ in 2002, $P < 0.05$ in 2006). TGW differences between Hap-6A-A and Hap-6A-G in the modern varieties were 3.2 and 3.0 g in 2002 and 2006, respectively. GT differences were significant in 2002 ($P < 0.01$), but not in 2006. There was no GL difference between the Hap-6A-A and Hap-6A-G genotypes in either year.

Collectively, the results demonstrated that *TaGW2*, like *OsGW2* in rice, was involved in grain development, mainly affecting GW and TGW. Because Hap-6A-A had a significantly positive effect on grain size, it was considered a potentially superior allele for the improvement of grain yield in wheat.

The relationship between *TaGW2* expression level and grain width

At 10 days post-flowering (dpf), the primer set TaGW2-4 was used to analyze expression of *TaGW2* in immature seeds of 14 varieties by real-time qRT-PCR. The expression level of *TaGW2* was negatively correlated with grain width, consistent with results for rice (Song et al. 2007). Moreover, the average expression level of *TaGW2* in

Table 4 Association analysis of grain traits between Hap-6A-A and Hap-6A-G genotypes over 2 years

Trait/genotype	02LY				06LY			
	Hap-6A-A[a]	Hap-6A-G[b]	*F* value	*P*[c]	Hap-6A-A[a]	Hap-6A-G[b]	*F* value	*P*[c]
Overall								
Seed length (mm)	6.598 ± 0.500	6.510 ± 0.567	1.810	0.180	6.530 ± 0.520	6.500 ± 0.557	0.223	0.637
Seed width (mm)	3.191 ± 0.215	3.044 ± 0.226	29.160	0.000***	3.239 ± 0.208	3.129 ± 0.180	21.049	0.000***
Seed thickness (mm)	2.863 ± 0.191	2.791 ± 0.202	8.942	0.003**	2.857 ± 0.185	2.849 ± 0.173	0.119	0.730
SL/SW ratio	2.072 ± 0.145	2.144 ± 0.186	12.547	0.000***	2.019 ± 0.152	2.080 ± 0.181	8.781	0.003**
1,000 grain weight (g)	38.080 ± 7.098	34.604 ± 7.641	14.725	0.000***	38.151 ± 7.214	35.408 ± 6.849	10.066	0.002**
Landraces								
Seed length (mm)	6.373 ± 0.452	6.411 ± 0.565	0.193	0.661	6.294 ± 0.446	6.403 ± 0.553	1.683	0.196
Seed width (mm)	3.035 ± 0.162	2.966 ± 0.187	5.550	0.020*	3.108 ± 0.177	3.091 ± 0.150	0.407	0.525
Seed thickness (mm)	2.764 ± 0.179	2.759 ± 0.207	0.019	0.890	2.748 ± 0.148	2.809 ± 0.169	5.295	0.023*
SL/SW ratio	2.104 ± 0.157	2.166 ± 0.192	4.514	0.035*	2.030 ± 0.161	2.075 ± 0.186	2.410	0.123
1,000 grain weight (g)	33.255 ± 5.518	32.379 ± 7.150	0.663	0.417	32.980 ± 4.976	33.301 ± 5.709	0.129	0.720
Modern varieties								
Seed length (mm)	6.798 ± 0.456	6.712 ± 0.521	0.847	0.359	6.739 ± 0.493	6.693 ± 0.519	0.221	0.639
Seed width (mm)	3.329 ± 0.153	3.204 ± 0.216	12.890	0.000***	3.355 ± 0.158	3.208 ± 0.212	17.967	0.000***
Seed thickness (mm)	2.952 ± 0.156	2.856 ± 0.177	9.111	0.003**	2.954 ± 0.159	2.932 ± 0.153	0.508	0.478
SL/SW ratio	2.044 ± 0.128	2.100 ± 0.166	4.086	0.046*	2.010 ± 0.145	2.092 ± 0.172	7.320	0.008**
1,000 grain weight (g)	42.361 ± 5.406	39.158 ± 6.572	7.970	0.006**	42.738 ± 5.617	39.719 ± 7.024	6.386	0.013*

02LY: Luoyang (2002), 06LY: Luoyang (2006)

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

[a] Hap-6A-A: landrace, $N = 63$; modern variety, $N = 71$

[b] Hap-6A-G: landrace, $N = 88$; modern variety, $N = 43$

[c] P values calculated by the F statistics

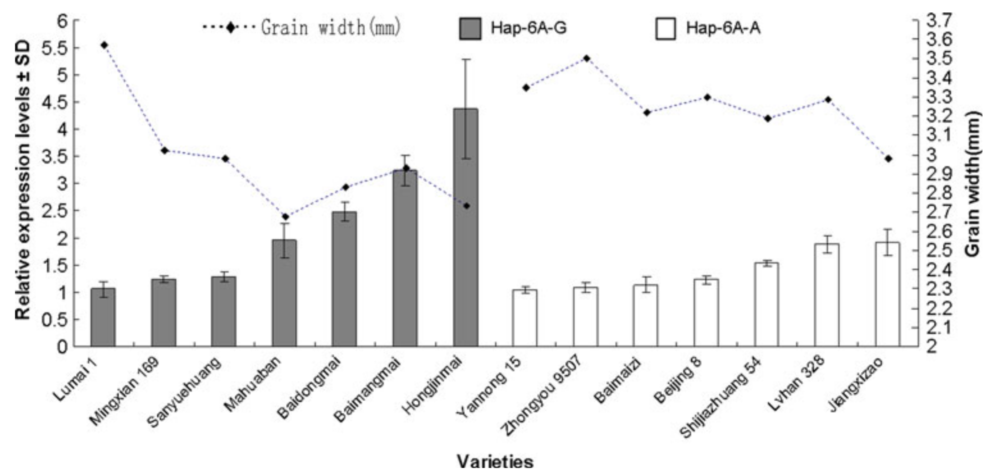

Fig. 5 The relationship between relative expression level of *TaGW2* and grain width in immature seeds at ten dpf. For variety information, see Supplement Table 1. The actin gene was used as the endogenous control

varieties with Hap-6A-G was higher than that in Hap-6A-A (Fig. 5). These results further supported the association of Hap-6A-A with higher grain width and weight.

Association of *TaGW2-6A* haplotypes with heading and maturing date

Strong associations of *TaGW2-6A* haplotypes with heading date and maturity date were found in both Chinese landraces and modern varieties; Hap-6A-A genotypes were earlier in heading and maturity (Fig. 6). Among landraces, the heading date differences between the two haplotypes were 6.6 and 4.5 days in the two growing seasons and for maturity date the corresponding differences were 4.0 and 5.9 days, respectively. Among improved varieties, the differences between the two haplotypes were 3.6 and 3.3 days in 2002 and 2006 for heading, and 2.8 and 2.0 days for maturity, respectively (Fig. 6; Table S3). Thus

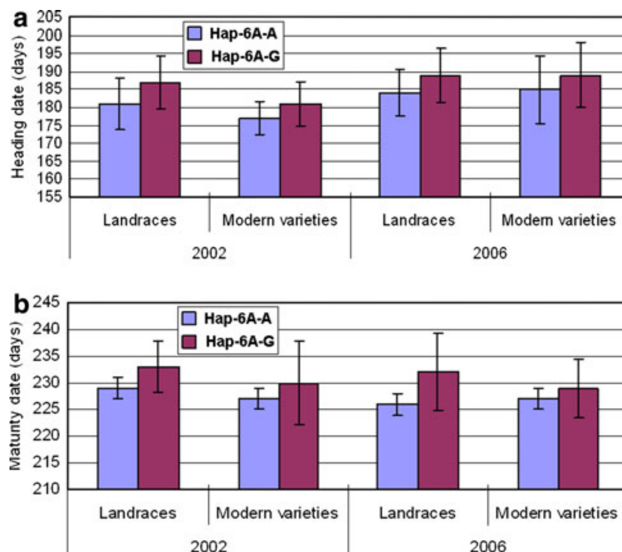

Fig. 6 Differences in heading (**a**) and maturity (**b**) between Hap-6A-A and Hap-6A-G in landraces and modern varieties over 2 years

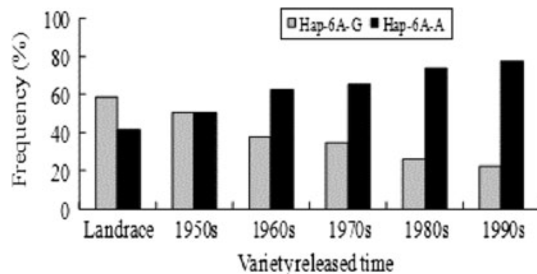

Fig. 7 Frequencies of *TaGW2-6A* Hap-6A-A and Hap-6A-G in Chinese wheat varieties released since the 1950s. Ten varieties released before 1950s were accounted in the subset of 1950s

despite overlapping ranges in heading and maturity in the 2 years, Hap-6A-A varieties were on average consistently earlier than Hap-6A-G varieties. We concluded that *TaGW2-6A* Hap-6A-A is associated with larger grain size and earlier heading and maturity, features that contribute to overall yield and the system of within-year multiple cropping.

Past selection of *TaGW2-6A* in wheat breeding

In landraces, *TaGW2-6A* Hap-6A-G is the predominant allele (58.28%), while the Hap-6A-A is the most frequent one in the modern varieties. The relative proportions of the two haplotypes among varieties released in 10-year groupings from the 1950s to 1990s are shown in Fig. 7. From 50.0% in the 1950s when varieties were either landraces or selections from landraces the frequency of Hap-6A-A consistently increased to current levels of 77.42%. This strongly indicates that *TaGW2-6A* Hap-6A-A was positively selected and that it is beneficial for grain-yield improvement.

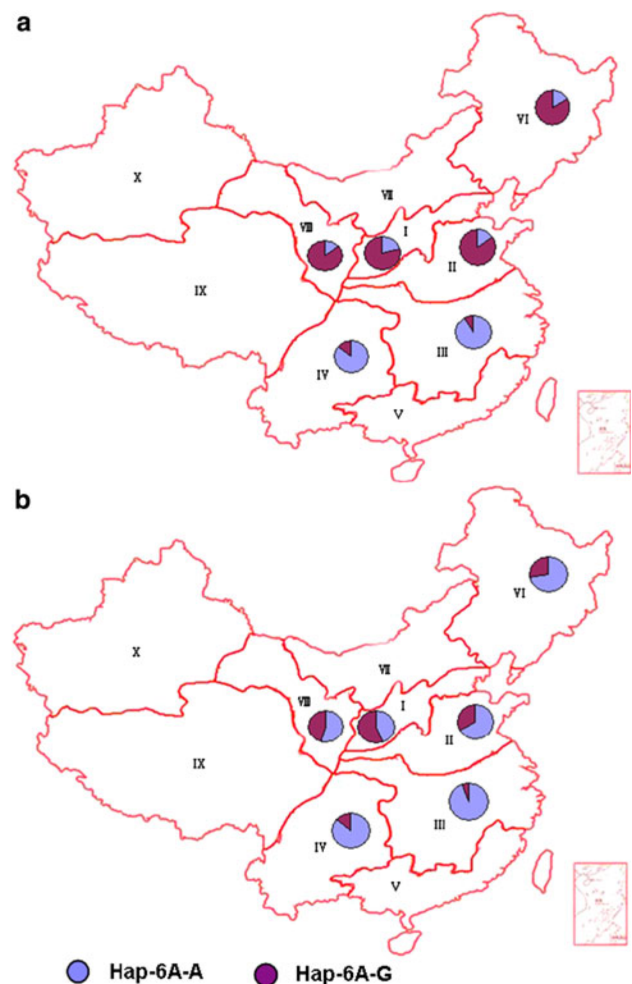

Fig. 8 Haplotype distributions at *TaGW2-6A* in Chinese landraces (**a**) and modern varieties (**b**) in the most important six agro-ecological production zones. *I* Northern winter wheat region, *II* Yellow and Huai River valley winter wheat region, *III* low and middle Yangtze River valley winter wheat region, *IV* southwestern winter wheat region, *V* southern winter wheat region, *VI* northeastern spring wheat region, *VII* northern spring wheat region, *VIII* northwestern spring wheat region, *IX* Qinghai-Tibet spring-winter wheat region, *X* Xinjiang winter-spring wheat region

Geographic distribution of *TaGW2-6A* haplotypes

The Chinese wheat production area is divided into ten ecological production zones based on environment, variety type and growing season (Zhang et al. 2002; Zhuang 2003). Because zones I, II, III, IV, VI and VIII account for more than 85% of the national wheat area and production, most of the wheat haplotyped varieties came from these regions. Among landraces, Hap-6A-A was very frequent in the neighboring autumn-planted spring wheat regions, III (91.30%) and IV (85.7%), whereas Hap-6A-G was more frequent (78.95–85.71%) in zones I, II, VI and VIII, the winter, facultative and spring-sown wheat zones. This indicated that the preferred Hap-6A-A was initially

selected and used in lower latitude areas of China (Fig. 8a; Table S4).

Compared with landraces, the frequencies of modern varieties with Hap-6A-A were increased in zones I, II, VI and VIII and were maintained at high levels in III (94.44%) and IV (86.21%). Hap-6A-A has increased to be predominant (54.17–94.44%) in all major wheat production zones in China, except the Northern Winter Wheat Zone I (44.19%) (Fig. 8b; Table S5; Fig. S1). This confirmed that the superior allele of *TaGW2-6A* can be used in different regions, and its contributions to grain width and weight are not limited by environmental conditions such as light and temperature.

Among European varieties released during 1899–1999, Hap-6A-G was more frequent, and Hap-6A-A was mainly restricted to varieties released in the former Yugoslavia, Italy, Bulgaria, Hungary and Portugal. The Hap-6A-A distribution showed a strong geographic bias, being mainly present in southern European varieties and at extremely low frequencies in northern Europe (Fig. 9).

Discussion

A feasible way to isolate yield-related genes in wheat

Since the 1960s the greatest contribution to yield increases in wheat came from the introduction and use of the reduced height genes *Rht1* and *Rht2* (Hedden 2003); however, no similar increases have occurred in recent years. In order to meet the future increasing demand for wheat, an important strategy will be the application of molecular tools to identify important yield-related genes to enhance breeding efficiency. Common wheat is an allohexaploid with a large genome (16,000 Mb), a high proportion (80%) of repetitive DNA, and few currently available genome sequences (Gupta et al. 2008). Whereas it is very difficult to isolate genes in wheat by map-based methods, comparative genetics has demonstrated that orthologous genes across genomes and species have parallel functions in regulating phenotypes, thus providing a feasible and effective way to isolate genes in wheat (Devos 2005). Many QTLs or genes associated with grain yield have been identified in model crops; for example, *Gn1a* regulating grain number, *GS3, GW2, qSW5* and *GIF1* associated with grain weight, and *Tb1* and *MOC1* controlling tiller number in rice (Ashikari et al. 2005; Fan et al. 2006; Li et al. 2003; Song et al. 2007; Takeda et al. 2003; Wan et al. 2008; Wang et al. 2008), the species that provides an important genetic resource to identify grain-yield related genes in wheat.

In the present study, we cloned an ortholog of rice *GW2* in wheat, and showed that the expression pattern and the function of the predicted protein was the same as in rice (Song et al. 2007). Although the mutation sites in *GW2* between the two species were not the same, the phenotypes of respective mutants affected grain width and weight in both cases. The effect of *TaGW2-6A* Hap-6A-A in wheat was similar to a loss-of-function mutation *OsGW2* in rice, leading to increased grain width and weight. Li et al. (2010) recently isolated the *GS3* gene in maize and showed that it might be involved in maize kernel development, in a similar way to rice *GS3*. Therefore, based on yield-related genes identified in the model crop, it is feasible to isolate grain-yield candidate genes in wheat and then to identify functional or superior alleles for use in breeding.

Effect and the putative mechanism of the *TaGW2-6A* gene in determining grain size

Association analysis of modern varieties indicated that the *TaGW2-6A* Hap-6A-A was a superior allele for grain size. The effect of *TaGW2-6A* in improving grain weight was mainly to increase grain width, with little effect on grain length and grain thickness. Mean grain width and weight of varieties with Hap-6A-A were significantly higher than those of varieties with Hap-6A-G. Nevertheless, some varieties with Hap-6A-G also had high grain width and weight (Table S1). This could be explained by other genes (QTLs) associated with grain development. Compared with modern varieties, the function of *TaGW2-6A* could not be detected in the landraces. The main reason for this was that the effect of *TaGW2-6A* was likely hidden by other genes associated with grain size, due to the higher overall genetic diversity in landraces (Hao et al. 2008). Interestingly, our result is somewhat like the *GS3* gene in rice (Takano-Kai et al. 2009) where there were no phenotypic differences between varieties carrying different *GS3* alleles in an *O. rufipogon* (rice wild progenitor) subpopulation, but there was a significant effect on grain length in *O. sativa*.

In rice, *GW2* negatively regulates grain width and weight, and 1 bp deletion in the coding region resulted in loss-of-function of the *GW2* allele, leading to increased grain width and weight. Transgenic rice plants with reduced *GW2* expression level had increased grain width and weight (Song et al. 2007). We found no sequence differences in the coding region of *TaGW2* among ten varieties with variable grain widths, but two SNPs in the promoter region affected both grain width and grain weight. The qRT-PCR results also implied that the effect of *TaGW2* on grain size was due to the level of gene expression. Thus for this gene, the two important crops shared a common mechanism affecting grain size (Fig. 5).

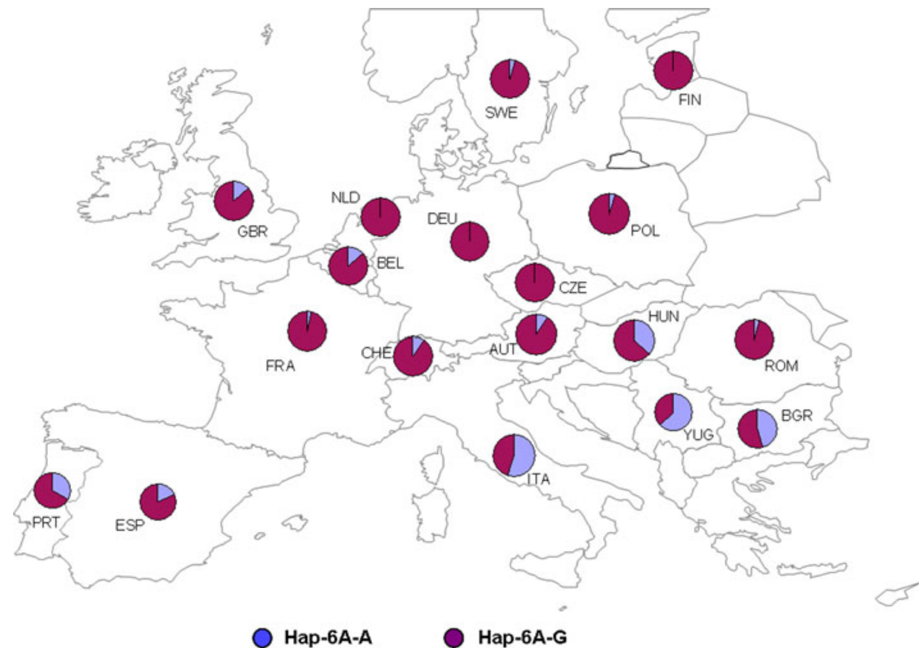

Fig. 9 Haplotype distribution at *TaGW2-6A* in European varieties. All abbreviations of country name are listed in the supplementary Table 2

A functional marker for grain width and weight in wheat

Functional or perfect markers derived from polymorphic sites within genes causally involved in phenotypic trait variation (Andersen and Lübberstedt 2003; Bagge et al. 2007) are ideal for marker-assisted breeding. However, compared with diploid crops, development of gene-derived (functional) markers is more complex in wheat because of the allohexaploid nature (Bagge et al. 2007). For most genes, there are at least three orthologs on the homoeologous chromosomes, and since their sequences and functions are very similar it is extremely difficult to characterize them separately. Since the breeding behavior of wheat is similar to a diploid species (i.e., disomic inheritance) markers must be found that uniquely identify the orthologs as well as the particular alleles at the individual loci. We designed a genome-specific primer set to differentiate the orthologous sequences. Then, based on the sequence difference between the alleles of *TaGW2-6A* in different varieties, a CAPS marker was developed to distinguish them. It is co-dominant and can be easily implemented in the laboratory. Its potential value for selection of grain width and weight, and hence TKW, was validated by association analysis. The Hap-6A-A allele identified with the CAPS marker not only led to higher grain width and weight, but also to earlier heading and maturity. This allele apparently had been positively selected in Chinese wheat breeding (Fig. 6). Now it can be used as a functional marker in wheat breeding programs aimed at improving grain size.

TaGW2-6A was mapped about 0.6 cM from marker *Xcfd80.2*, which is close to the 6A centromere. This gene therefore might be related to the major yield QTL repeatedly mapped to the chromosome 6A pericentromeric region (Huang et al. 2004, 2006; Snape et al. 2007; Sun et al. 2009). The *TaGW2-6A* Hap-6A-A allele not only increases TKW, but also advances maturity by about 3 days, a useful trait in multiple cropping situations that require rapid turnover from one crop to the next. It also seems that this allele could make a contribution to yields in northern Europe where it currently occurs at a very low frequency. Finally, this work demonstrates the value of comparative genomics in isolating important yield-related genes in wheat, a crop with a huge genome size.

Acknowledgments The authors are grateful to Prof. Yiping Tong, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, for fine mapping of *TaGW2-6A* gene. We also gratefully acknowledge help from Prof. Robert A McIntosh, University of Sydney, with English editing. This research was supported by the Chinese Ministry of Science and Technology (2010CB125900), National Natural Science Foundation of China (30900898), funding from Ministry of Agriculture (2008ZX08009) and Core Research Budget of the Non-profit Governmental Research Institution.

References

Andersen JR, Lübberstedt T (2003) Functional markers in plants. Trends Plant Sci 8:554–560

Ashikari M, Sakakibara H, Lin S, Yamamoto T, Takashi T, Nishimura A, Angeles ER, Qian Q, Kitano H, Matsuoka M (2005) Cytokinin oxidase regulates rice grain production. Science 309:741–745

Bagge M, Xia XC, Lübberstedt T (2007) Functional markers in wheat. Curr Opin Plant Biol 10:211–216

Botwright TL, Condon AG, Rebetzke GJ, Richards RA (2002) Field evaluation of early vigour for genetic improvement of grain yield in wheat. Aust J Agric Res 53:1137–1146

Breseghello F, Sorrells ME (2006) Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. Genetics 172:1165–1177

Campbell KJ, Bergman CJ, Gualberto DG, Anderson JA, Giroux MJ, Hareland G, Fulcher RG, Sorrells ME, Finney PL (1999) Quantitative trait loci associated with kernel traits in a soft × hard wheat cross. Crop Sci 39:1184–1195

Campbell BT, Baenziger PS, Gill KS, Eskridge KM, Budak H, Erayman M, Dweikat I, Yen Y (2003) Identification of QTLs and environmental interactions associated with agronomic traits on chromosome 3A of wheat. Crop Sci 43:1493–1505

Choi HK, Mun JH, Kim DJ, Zhu HY, Baek JM, Mudge JM, Roe B, Ellis N, Doyle J, Kiss GB, Youn ND, Cook DR (2004) Estimating genome conservation between crop and model legume species. Proc Natl Acad Sci USA 101:15289–15294

Devos KM (2005) Updating the 'crop circle'. Curr Opin Plant Biol 8:155–162

Dholakia BB, Ammiraju JSS, Singh H, Lagu MD, Röder MS, Rao VS, Dhaliwal HS, Ranjekar PK, Gupta VS (2003) Molecular marker analysis of kernel size and shape in bread wheat. Plant Breed 122:392–395

Fan CC, Xing YZ, Mao HL, Lu TT, Han B, Xu CG, Li XH, Zhang QF (2006) *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 112:1164–1171

Fulton TM, Van der Hoeven R, Eannetta NT, Tanksley SD (2002) Identification, analysis, and utilization of conserved ortholog set markers for comparative genomics in higher plants. Plant Cell 14:1457–1467

Gale MD, Devos KM (1998) Plant comparative genetics after 10 years. Science 282:656–659

Guo ZA, Song YX, Zhou RH, Ren ZL, Jia JZ (2010) Discovery, evaluation and distribution of haplotypes of the wheat *Ppd-D1* gene. New Phytol 185:841–851

Gupta PK, Mir RR, Mohan A, Kumar J (2008) Wheat genomics: present status and future prospects. Int J Plant Genomics 2008:1–36

Hao CY, Wang LF, Zhang XY, You GX, Dong YS, Jia JZ, Liu X, Shang XW, Liu SC, Cao YS (2006) Genetic diversity in Chinese modern wheat varieties revealed by microsatellite markers. Sci China: Ser C Life Sci 3:218–226

Hao CY, Dong YC, Wang LY, You GX, Zhang HN, Ge HM, Jia JZ, Zhang XY (2008) Genetic diversity and construction of core collection in Chinese wheat genetic resources. Chin Sci Bull 53:1518–1526

He XY, Zhang YL, He ZH, Wu YP, Xiao YG, Ma CX, Xia XC (2008) Characterization of phytoene synthase 1 gene (*Psy1*) located on common wheat chromosome 7A and development of a functional marker. Theor Appl Genet 116:213–221

Hedden P (2003) The genes of the Green Revolution. Trends Genet 19:5–9

Huang XQ, Kempf H, Ganal MW, Röder MS (2004) Advanced backcross QTL analysis in progenies derived from a cross between a German elite winter wheat variety and a synthetic wheat (*Triticum aestivum* L.). Theor Appl Genet 109:933–943

Huang XQ, Cloutier S, Lycar L, Radovanovic N, Humphreys DG, Noll JS, Somers DJ, Brown PD (2006) Molecular detection of QTLs for agronomic and quality traits in a doubled haploid population derived from two Canadian wheats (*Triticum aestivum* L.). Theor Appl Genet 113:753–766

Kato K, Miura H, Sawada S (2000) Mapping QTLs controlling grain yield and its components on chromosome 5A of wheat. Theor Appl Genet 101:1114–1121

Li XY, Qian Q, Fu ZM, Wang YH, Xiong GS, Zeng DL, Wang XQ, Liu XF, Teng S, Hiroshi F, Yuan M, Luo D, Han B, Li JY (2003) Control of tillering in rice. Nature 422:618–621

Li Q, Yang XH, Bai GH, Warburton ML, Mahuku G, Gore M, Dai JR, Li JS, Yan JB (2010) Cloning and characterization of a putative *GS3* ortholog involved in maize kernel development. Theor Appl Genet 120:753–763

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

Matsuoka M, Ashikari M (2007) A quantitative trait locus regulating rice grain width. Nat Genet 39:583–584

Moore G, Devos KM, Wang Z, Gale MD (1995) Cereal genome evolution: grasses, line up and form a circle. Curr Biol 5:737–739

Paterson AH (1995) Molecular dissection of quantitative traits: progress and prospects. Genome Res 5:321–333

Peng JR, Richards DE, Hartley NM, Murphy GP, Devos KM, Flintham JE, Beales JB, Fish LJ, Worland AJ, Pelica F, Sudhakar D, Christou P, Snape JW, Gale MD, Harberd NP (1999) 'Green revolution' genes encode mutant gibberellin response modulators. Nature 400:256–261

Peng JH, Ronin YF, Fahima T, Röder M, Li YC, Nevo E, Korol A (2003) Domestication quantitative trait loci in *Triticum dicoccoides*, the progenitor of wheat. Proc Natl Acad Sci USA 100:2489–2494

Rajaram S (2005) Role of conventional plant breeding and biotechnology in future wheat production. Turk J Agric For 29:105–111

Röder MS, Huang XQ, Börner A (2008) Fine mapping of the region on wheat chromosome 7D controlling grain weight. Funct Integr Genomics 8:79–86

Sharp PJ, Kreis M, Shewry PR, Gale MD (1988) Location of β-amylase sequences in wheat and its relatives. Theor Appl Genet 75:286–290

Shomura A, Izawa T, Ebana K, Ebitani T, Kanegae H, Konishi S, Yano M (2008) Deletion in a gene associated with grain size increased yields during rice domestication. Nat Genet 40:1023–1028

Snape JW, Foulkes MJ, Simmonds J, Leverington M, Fish LJ, Wang YK, Ciavarrella M (2007) Dissecting gene × environmental effects on wheat yields via QTL and physiological analysis. Euphytica 154:401–408

Song XJ, Huang W, Shi M, Zhu MZ, Lin HX (2007) A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. Nat Genet 39:623–630

Sorrells ME, La Rota M, Bermudez-Kandianis CE, Greene RA, Kantety R, Munkvold JD (2003) Comparative DNA sequence analysis of wheat and rice genomes. Genome Res 13:1818–1827

Sun XY, Wu K, Zhao Y, Kong FM, Han GZ, Jiang HM, Huang XJ, Li RJ, Wang HG, Li SS (2009) QTL analysis of kernel shape and weight using recombinant inbred lines in wheat. Euphytica 165:615–624

Takano-Kai N, Jiang H, Kubo T, Sweeney M, Matsumoto T, Kanamori H, Padhukasahasram B, Bustamante C, Yoshimura A, Doi K, McCouch S (2009) Evolutionary history of *GS3*, a gene conferring grain length in rice. Genetics 182:1323–1334

Takeda T, Suwa Y, Suzuki M, Kitano M, Ueguchi-Tanaka M, Ashikari M, Matsuoka M, Ueguchi C (2003) The *OsTB1* gene negatively regulates lateral branching in rice. Plant J 33:513–520

Wan XY, Weng JF, Zhai HQ, Wang JK, Lei CL, Liu XL, Guo T, Jiang L, Su N, Wan JM (2008) Quantitative trait loci (QTL) analysis for rice grain width and fine mapping of an identified QTL allele gw-5 in a recombination hotspot region on chromosome 5. Genetics 179:2239–2252

Wang ET, Wang JJ, Zhu XD, Hao W, Wang LY, Li Q, Zhang LX, He W, Lu BR, Lin HX, Ma H, Zhang GQ, He ZH (2008) Control of

rice grain-filling and yield by a gene with a potential signature of domestication. Nat Genet 40:1370–1374

Weng JF, Gu SH, Wan XY, Gao H, Guo T, Su N, Lei CL, Zhang X, Cheng ZJ, Guo XP, Wang JL, Jiang L, Zhai HQ, Wan JM (2008) Isolation and initial characterization of GW5, a major QTL associated with rice grain width and weight. Cell Res 18:1199–1209

Zhang XY, Li CW, Wang LF, Wang HM, You GX, Dong YS (2002) 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 106:112–117

Zhuang QS (2003) Chinese wheat improvement and pedigree analysis. Agricultural Press, Beijing (in Chinese)