Maize orthologs of rice GS5 and their trans-regulator are associated with kernel development

Jie Liu, Min Deng, Huan Guo, Sharif Raihan, Jingyun Luo, Yuancheng Xu, Xiaofei Dong and Jianbing Yan*

National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan 430070, China. *Correspondence: yjianbing@mail.hzau.edu.cn

Abstract Genome information from model species such as rice can assist in the cloning of genes in a complex genome, such as maize. Here, we identified a maize ortholog of rice GS5 that contributes to kernel development in maize. The genome-wide association analysis of the expression levels of ZmGS5, and 15 of its 26 paralogs, identified a trans-regulator on chromosome 7, which was a BAK1-like gene. This gene that we named as ZmBAK1-7 could regulate the expression of ZmGS5 and three of the paralogs. Candidate-gene association analyses revealed that these five genes were associated with maize kernel development-related traits. Linkage analyses also detected that ZmGS5 and ZmBAK1-7 co-localized with mapped QTLs. A transgenic analysis of ZmGS5 in Arabidopsis thaliana L. showed a significant increase in seed weight and cell number, suggesting that ZmGS5 may have a conserved function among different plant species that affects seed development.

Keywords: Association analysis; kernel development; maize; ZmBAK1-7; ZmGS5


INTRODUCTION

Maize, a widely grown staple food, plays a critical role in supporting the growing world population. Grain size and weight are two important components of grain yield, but the genetic bases of these traits in maize are insufficiently understood.

Linkage and association mapping are two powerful and complementary methods to dissect the genetic bases of complex traits. For rice, using high-density genetic linkage maps and different types of mapping populations, researchers have resolved hundreds of QTLs and genes associated with yield traits (Xing and Zhang 2010). Map-based cloning of QTLs has identified tens of genes controlling grain size and grain yield in rice (Akiyama et al. 2000; Fan et al. 2000; Song et al. 2007; Shomura et al. 2008; Wang et al. 2008; Xue et al. 2008; Huang et al. 2009; Jiao et al. 2010; Li et al. 2010; Mao et al. 2010; Miura et al. 2010; Zhang et al. 2012), as well as many QTLs associated with yield and yield-related traits in maize. For example, the famous nested association mapping (NAM) population was used to map QTLs and elucidate the genetic architecture of complex agronomic and yield traits (Buckler et al. 2009; Kump et al. 2011; Tian et al. 2011). However, few genes have been cloned by map-based methods in maize. The genome-wide association study (GWAS) is a powerful method for gene identification in maize (Yan et al. 2011). More than one million SNPs were developed in a diverse panel using RNA sequencing strategy in maize, and 74 genes associated with oil concentration and fatty acid components were identified by GWAS (Li et al. 2015). However, using a similar strategy and the same panel, only a few loci were identified for yield and agronomic traits (Yang et al. 2014), which implies that grain yield is a more complex trait compared with fatty acids and that many genetic factors are probably involved.

Comparative genomics has proven to be a useful alternative approach to identifying the genes underlying QTLs for complex traits. Using high-resolution linkage mapping and expression analysis, Miller et al. (2007) found that cis-regulatory changes in Kit ligand (Kitlg) affect the gill and skin tissue colors of sticklebacks and, through admixture mapping, that the human KITLG genomic region has a significant effect on human skin color. The maize TEOSINTE BRANCHED 1 (TBI) gene is involved in regulating the growth of axillary buds (Clark et al. 2006; Studer et al. 2011). The rice OsTBI gene, which was identified based on its sequence similarity with maize TBI, also negatively regulates lateral branching (Takeda et al. 2003). Orthologous genes of rice kernel size-related GS3 (Fan et al. 2006) and GW2 (Song et al. 2007) were also identified in maize (Li et al. 2011a, 2011b) and wheat (Bednarek et al. 2012; Zhang et al. 2014), and they have similar roles, providing us the opportunity to functionally identify and clone orthologous genes using the well-annotated rice genome.

GS5 controls grain size and weight through the regulation of grain width and filling in rice (Li et al. 2011). Enhanced expression of GS5 competitively inhibits the interaction between OsBAK1-7 and OsMSBPI by occupying the extracellular leucine-rich repeat (LRR) domain of OsBAK1-7 (Xu et al. 2015). This inhibition could prevent OsBAK1-7 from endocytosis caused by interacting with OsMSBPI which may explain how GS5 could affect grain size. In the present study, the orthologous gene of rice GS5, ZmGS5, and 26 paralogs of ZmGS5 were identified in maize. An eQTL analysis of ZmGS5 and its paralogs found that ZmGS5 and other three GS homologous genes were regulated by a trans-regulator, ZmBAK1-7 which was a BAK1-like gene, on chromosome 7. Candidate-gene association analysis showed that ZmGS5 and
ZmBAK1-7 were associated with kernel-related traits. These two genes were also co-localized with mapped QTLs in present and previous studies. Transgenic analysis in Arabidopsis indicated that ZmGS5 could regulate kernel development, thus enhancing seed weight.

RESULTS

Identification and characterization of ZmGS5

BLAST searches using the rice GS5 protein sequence (GenBank: AE037083) as the query against the maize B73-filtered gene set version 5B.60 RefGen_v2 database identified the protein (GRMZM2G1233815) with the highest similarity to GS5 (E-Value = 0, Identity = 73.63%). This gene, which we named ZmGS5, contains 10 exons, as does GS5 in rice.

The comparative genomic analysis, using Symap software (Soderlund et al. 2011), between rice chromosome 5 and maize chromosome 3 identified syntenic fragments corresponding to rice GS5 and maize ZmGS5 regions. Both genes belonged to the peptidase family S10, which has a serine-type carboxypeptidase activity.

In addition to ZmGS5, we identified 26 other genes that were paralogs of ZmGS5 (Table S1). Most of these genes had no description, but some were predicted to have serine carboxypeptidase activity.

ZmGS5 and the other three homologous genes are regulated by a same trans-regulator

More than one million high-quality SNPs and the expression levels of 28,769 genes at 15 d after pollination were publicly available in 368 diverse maize lines (Li et al. 2013; Fu et al. 2013). SNPs and expression levels for ZmGS5 and 12 paralogs were identified in the 368 maize lines. Using these datasets, GWAS on the expression levels of these genes were performed, and four genes, including ZmGS5, were found to be regulated by the same trans-eQTL located on chromosome 7, which we named ZmBAK1-7 (GRMZM2G490551, Figure 1A) because of its BAK1-like domain. The alignments of the BAK1-like domain between ZmBAK1-7 with OsBAK1 and OsBAK1-7, which could interact with rice GS5 showed high similarity (E-value = 8e-49 and 5e-54, Identity = 34 and 36%, respectively).

Sequence analysis of the protein encoded by ZmBAK1-7 was performed through searching InterProScan and it indicated that this protein had putative leucine-rich repeat receptor-like serine/threonine-protein kinase (LRR-RLKs) activity. The HMM-HMM-based lightning-fast iterative sequence search method (Remmert et al. 2012) was used to search the Protein Data Bank, which revealed that this protein has a brassinosteroid insensitive 1 structure and a brassinosteroid insensitive 1-associated receptor structure. The secondary structure prediction using SOSUI (Hirokawa et al. 1998) showed that this protein had transmembrane helices.

The eQTL analysis demonstrated that the cis-element near this gene strongly regulates its expression level (Figure 1B). It was revealed that ZmBAK1-7 could regulate the expression of ZmGS5, as well as the other three homologous genes of GS5 (Figure 1 C-F) and that these five genes were highly co-expressed (r > 0.5, Figure 1A). These data implied that ZmGS5 and another three homologous genes might be involved in the same pathway or participate in coordinated developmental processes.

Natural variations in ZmGS5 and ZmBAK1-7 affect maize kernel-related traits

Kernel-related traits, including kernel length (KL), kernel width (KW), kernel thickness (KT) and 100-kernel weight (HKW), of 368 diverse maize inbred lines were measured under four or more environments (location and year), including Sichuan, Yunnan and Hainan Provinces in 2009, 2010, and 2011, respectively. Significant phenotypic variations were identified in the four measured traits. The smallest variation was for KL, ranging from 7.04 to 11.26 mm, and the greatest variation was for HKW, ranging from 11.11 to 29.79 g (Table 1). The broad-sense heritability for these four traits was >0.85 (Table 1).

For a candidate gene association analysis, we combined two sets of genotypic data from ZmGS5 and ZmBAK1-7. One set was the SNPs generated through RNA-Seq and the other was the polymorphisms obtained by re-sequencing the 5’ upstream regions of these two genes. We re-sequenced ZmGS5 in 155 diverse maize inbred lines (Yang et al. 2010) and another 18 polymorphic sites, including four insertions and/or deletions (InDels). Re-sequencing of ZmBAK1-7 was performed in 508 diverse maize inbred lines and generated another 43 polymorphisms, including 13 InDels. Overall, we identified 43 and 67 polymorphisms with minor allele frequencies (MAFs) ≥0.05 in ZmGS5 and ZmBAK1-7, respectively. Genotypic data for the other three GS5 homologous genes, GRMZM2G052507, GRMZM2G1233940 and GRMZM2G879749, were derived from RNA-Seq results and contained 35, 7 and 20 SNPs, respectively, with MAFs ≥0.05.

The candidate gene association analysis was performed using TASSEL software (Bradbury et al. 2007) with a mixed linear model (Yu et al. 2006) that takes population structure and kinship into consideration. ZmBAK1-7 was associated with all of the investigated kernel traits in at least two environments, and ZmGS5 was associated with KL, KW and HKW in one environment (Tables 2, S2). GRMZM2G052507, GRMZM2G879749 and GRMZM2G1233940 were associated with at least one kernel-related trait (Tables 2, S2). Most SNPs in ZmBAK1-7 and ZmGS5 were significantly associated with kernel traits in two environments. However, SNP M7c130937610 showed significant associations with KL in four environments (Table S2).

SNP M7c130937610 from ZmBAK1-7 was the most significant polymorphism in this gene and was associated with KT in Hainan Province, 2011 (n = 320, P = 3.39 × 10-3) (Figure 2A). It could lead to an amino acid change (Thr/Ser) which was predicted to locate in the leucine-rich repeat receptor-like protein kinase domain and outside the cell membrane (Figure 2B). This SNP was in strong linkage disequilibrium (LD) with three other significant SNPs (r² > 0.2, Figure 2C). Among these three SNPs, two were synonymous and the remaining one could lead to amino acid change (Ala/Ser), but it was not located in reported domains. SNP M3c61242466 from ZmGS5 was the most significant variant and associated with KL in Hainan Province, 2011 (n = 326, P = 3.90 × 10-4) (Figure 3A). This SNP and other significant SNPs were all synonymous and could not affect the amino acid sequence (Figure 3B). Notably, these SNPs were in strong LD (r² > 0.5) and located near the 3’ UTR region (Figure 3C). This suggested that the potential functional variation might be a regulatory element in the 3’ end of the gene. These two SNPs contributed 6.7% and 6.3% to the phenotypic variation of KT and KL, respectively. Comparisons of
these two significant P-values with the P-value distribution of 500 randomly chosen SNPs from 500 different genes suggested that these two significant associations were not due to false positives (Figures 2D, 3D). Considering these results, we concluded that ZmGS5 and ZmBAK1-7 were associated with kernel development.

ZmGS5 and ZmBAK1-7 are located in mapped kernel-related QTL intervals
One major QTL affecting KW and HKW was identified on chromosome 7 in a recombinant inbred lines (RIL) population derived from an inbred line BK selected from a tropical landrace with a big kernel size and the elite Chinese breeding
line Yu8701. ZmBAK1-7 was mapped near the peak region of the QTL (Figure 4A). Another QTL affecting the kernel test weight was also identified on chromosome 3 in a BC\textsubscript{F\textregistered} population derived from a teosinte accession and the elite breeding line Mo17. ZmGS5 was mapped near the peak region of the QTL (Figure 4B). Previous studies (Beavis et al. 1994; CIMMYT 1994; Doebley et al. 1994; Veldboom and Lee 1994, 1996; Ajmone-Marsan et al. 1995; Austin and Lee 1996; Melchinger et al. 1998) also identified QTLs for KW and yield in similar regions (bin 3.04 and 7.03) to where ZmGS5 and ZmBAK1-7 were located, respectively (Table S3). The QTL where ZmGS5 was located explained 8.5% of the kernel test weight variation in the Teosinte/Mo17 BC\textsubscript{F\textregistered} population, whereas the QTL where ZmBAK1-7 was located explained 10.6% and 9.8% of the KW and HKW variation, respectively, in the BK/Yu8701 RIL population. The two candidate genes were of great interest and might be the underlying genes of the two QTLs although more studies were required.

\textbf{ZmGS5 affects seed size in transgenic Arabidopsis}

The full-length cDNA sequence of ZmGS5 from maize inbred line B73 was cloned and transformed into Arabidopsis with expression vector pBinGlyRed3 consisting of CaMV\textsubscript{35S} promoter (Zhang et al. 2013). In total, 14 independent ZmGS5 transgenic lines were obtained (Figure 5A), and the 1,000-seed weights of the T\textsubscript{2} generations from 10 of 14 transgenic lines were higher than that of the wild type. In particular, the 1,000-seed weights of ZmGS5-1, ZmGS5-13 and ZmGS5-20 were significantly higher than the wild-type. To confirm the results, these positive transgenic lines, ZmGS5-1, ZmGS5-13 and ZmGS5-20, were selected for further analysis. Subsequently, the seeds of the third generation of transgenic individuals were used to measure 1,000-seed weight. The 1,000-seed weight from transgenic lines ZmGS5-1, ZmGS5-13 and ZmGS5-20 increased 6.8, 6.6 and 17.0%, respectively, over wild type (Figure 5A). The data from the T\textsubscript{3} generation showed similar trends to those of the T\textsubscript{2} generation. Semi-quantitative PCR was performed as an expression analysis of ZmGS5 in the measured materials, and a positive correlation was observed between seed size and gene expression (Figure 5B).

We also investigated the size of embryos, including cotyledons and radicals, from mature seeds (Figure 5C). The cotyledon sizes of the transgenic lines ZmGS5-1, ZmGS5-13 and ZmGS5-20 were 19.6, 24.1 and 9.9% greater, respectively, than that of wild type, and the radical sizes of the transgenic lines ZmGS5-1, ZmGS5-13 and ZmGS5-20 were 20.5, 31.0 and 9.9% greater, respectively, than that of wild type (Figure 5E). These results were in accordance with the seed weights. To understand the basis for the increased cotyledon size in transgenic plants, we analyzed the cell sizes in the central region of the cotyledons (Figure 5D). The average cell areas were significantly increased in the three transgenic lines by 11.0, 14.2 and 7.1%, respectively, over that of wild type (Figure 5F). However, the cell number in cotyledons, obtained by dividing the cotyledon size by the embryo cell area, showed no significant difference between the transgenic lines ZmGS5-1 and ZmGS5-13 and the wild type; however, ZmGS5-20 showed a non-significant decrease of 6.4% (Figure 5F). This suggested that ZmGS5 affected both seed weight and cotyledon size.

\textbf{DISCUSSION}

Candidate gene mining and association analyses based on a comparative genomic strategy are useful tools for identifying key genes affecting complex agronomic traits, especially in crops with large genomes. With the rapid development of next-generation sequencing technology, it has become cheaper and quicker to determine the reference genome sequences of many crop species. Unlike conserved marker-based comparative genomic analyses (Moore et al. 1995), the new technology allows us to compare the whole genome sequences of many crop species and identify the conserved sequences and functions of agronomically important genes (Van Bel et al. 2012). This will lead to a better understanding of the evolutionary history, domestication and improvement of crops for agronomically and economically important traits, thus enhancing plant breeding.

The method of comparative genomics has been used to identify maize orthologs of rice genes controlling grain size and weight in previous studies (Li et al. 2010a, 2010b). In this study, we also found significant associations between ZmGS5 and kernel-related traits. Three haplotypes in the promoter region of rice GSS seem to be associated with KW (Li et al. 2010a).
ZmGS5 is associated with kernel development

However, according to our results, ZmGS5 might not affect grain size in the same way as the rice gene because its relevant sites were not in the promoter region. This result was also consistent with the results of ZmGS3, ZmGW2-CHR4 and ZmGW2-CHR5 (Li et al. 2010a, 2010b). They suggest that maize and rice might have undergone different evolutionary pressures after divergence.

Seed weight and cotyledon cell number increased significantly compared with wild type in A. thaliana over-expressing ZmGS5 (Figure 5). Thus GS5 seems to be functionally conserved across different plant species. The increased seed weight, cell number and faster root growth of transgenic lines compared with wild type were consistent with the function of rice GS5 as a positive modulator upstream of cell cycle genes (Li et al. 2011). It would be interesting to determine whether GS5 is also functional in other crops, such as wheat and sorghum.

Although many genes underlying QTLs for grain size and weight have been positionally cloned in other species (Ashikari et al. 2005; Fan et al. 2006; Song et al. 2007; Shomura et al. 2008; Wang et al. 2008; Xue et al. 2008; Huang et al. 2009; Jiao et al. 2010; Mao et al. 2010; Li et al. 2010; Miura et al. 2010; Zhang et al. 2012), little is known about the pathways they involve in. Here, in addition to the association between ZmGS5 and kernel-related traits, we have also discovered using eQTL analysis that ZmBAK1-7 could regulate the expressions of ZmGS5 and other three homologous genes and all these five genes were highly co-expressed.

ZmBAK1-7 was a BAK1-like gene and contained LRR-RLKs domain. Mutations in various LRR-RLKs can affect diverse developmental processes in plants, such as the perception of the hormones brassinosteroid (Li and Chory 1997) and system in (Montoya et al. 2002; Scheer and Ryan 2002), meristem differentiation (Clark et al. 1993), endosperm and pollen development (Canales et al. 2002; Zhao et al. 2002), ovule development and early embryogenesis (Hecht et al. 2001).

Recently, it was found that rice GS5 could interact with OsBAK1-7 and its expression could be suppressed by BR. This demonstrated that GS5 involved in brassinosteroid signaling. In our study, we also found that ZmGS5 could interact with a BAK1-like gene (ZmBAK1-7). This consistence indicated that ZmGS5 might also involve in brassinosteroid signaling. The similar domain between OsBAK1-7 and ZmBAK1-7 implied that GS5 might have conserved function among different plant species. Meanwhile, we also found that ZmBAK1-7 was one of the 12 regulatory hotspots which could regulate the expressions of more than 100 genes (data not shown). When a very stringent threshold for GWAS was set ($P = 1.0 \times 10^{-10}$), ZmBAK1-7 was found to regulate expressions of 67 genes (Table S4). If the threshold was set to a less stringent level ($P = 1.8 \times 10^{-6}$), then the number of regulated genes reached $\approx 900$. Gene ontology (GO) analysis of these 67 genes showed that these genes could play role as binding factor, catalytic factor, transcription regulator and transporter and they mainly affected the process of

![Figure 2. Significant associations between natural variations of ZmBAK1-7 and kernel thickness](image-url)

(A) Candidate gene association analysis between single nucleotide polymorphisms (SNPs) of ZmBAK1-7 and kernel thickness. (B) Gene structure of ZmBAK1-7. Two significant SNPs were non-synonymous. (C) LD matrix across ZmBAK1-7. (D) Permutation test for the identified strongest associated SNP. Comparison of the strongest significant SNP ($n = 251/69, P = 2.96 \times 10^{-6}$) identified in present study with the association results of 500 randomly selected SNPs for the same trait. The results showed that the identified association is more significant than the 500 times random tests which demonstrated the identified association was not false positive but real association.
biological regulation, metabolic reaction, response to stimulus (Figure S2). Thus, ZmBAK1-7 seems to be a key regulator that interacts with ZmGS5 to affect maize kernel development. Further studies are required to reveal the function of ZmBAK1-7 that will be very helpful in understanding the maize kernel development as well as facilitating the genetic improvement of grain yield.

MATERIALS AND METHODS

Genetic materials and phenotypic data collection

An association mapping population (AM508), consisting of 508 diverse maize inbred lines, was developed for the dissection of complex quantitative traits (Yang et al. 2011). One subset of 155 lines (CAM155) (Yang et al. 2010) was used for the re-sequencing of ZmGS5, and another subset composed of 368 lines was used for RNA-Seq (Li et al. 2013). The re-sequencing of ZmBAK1-7 was done in AM508. Populations were planted during 2009, 2010 and 2011 in Sichuan, Yunnan and Hainan Provinces. Two linkage populations, BK/Yu8701 RIL and Teosinte/Mo17 BC2F7, were planted in Hainan in 2011. Kernel-related traits, such as KL, KW and KT, were measured 10 times for each line using a digital ruler, and HKW was measured three times for each line using an electronic scale. The values of each trait were then averaged to obtain the phenotypic value used in the analysis.

Identification of ZmGS5 and 26 other genes with a high similarity to GS5

Based on the protein sequence of rice GS5, we searched reference protein databases and identified the protein with the highest similarity. We also conducted comparative genomic analyses between the GS5 regions of rice and maize and the reference sequences and annotation files of the rice and maize genomes, respectively. A similar method was used to identify the paralogs of ZmGS5 based on the protein sequence of ZmGS5 and the maize B73-filtered gene set version 5b.60 RefGen_V2 database (Schnable et al. 2009).

Protein sequence analysis

First, we searched the InterProScan database to identify the domains present in these genes. This would identify, based on the amino acid sequences, families that these proteins belong...
to. Then, a faster and more sensitive method than PSI-BLAST, HMM-HMM-based lightning-fast iterative sequence search (Remmert et al. 2012), was used to search the Protein Data Bank, confirm the structure and produce a more detailed functional annotation of the protein.

Genotypic data and gene expression analysis
For association tests, we used two sets of genotypic data. One consisted of the high-quality SNPs from RNA-Seq (Li et al. 2013), and the other was the re-sequencing data for ZmGS5 and ZmBAK1-7. The primers (Table S5) used to re-sequence these two genes were designed based on the B73 reference genome sequence using Primer-BLAST in NCBI. The PCR products were sequenced, and multiple alignments of these sequences were performed using BioEdit software (Hall, 1999). Polymorphisms were generated from the alignment results using TASSEL software to extract InDels and remove SNPs with MAFs < 0.05.

Association and linkage analyses
The expression levels of ZmGS5, ZmBAK1-7 and three other paralogs of ZmGS5 were quantified by RNA-Seq (Li et al. 2013). GWAS for eQTLs and association tests between polymorphisms from candidate genes and kernel-related traits were performed using TASSEL software with a mixed linear model. The kinship matrix and population structure were estimated using 36,618 high-quality SNPs obtained from Illumina MaizeSNP50 BeadChip (Ganal et al. 2011; Li et al. 2012). For

Figure 4. ZmBAK1-7 and ZmGS5 col-located with the mapped QTL for kernel width, 100-kernel weight and kernel test weight, respectively
(A) ZmBAK1-7 fell into one of the mapped QTL intervals in BK/Yu8701 recombinant inbred line (RIL) population. The investigated traits were kernel width and 100-kernel weight in Hainan, 2011. The arrow indicated the position of ZmBAK1-7. (B) ZmGS5 fell into one of the mapped QTL intervals in Teosinte/Mo17 BC2F6 population. The investigated trait was kernel test weight in Hainan, 2011. The arrow indicated the position of ZmGS5.
GWAS, the threshold was set at $P = 1.8 \times 10^{-6}$ ($P = 1/n$, where $n$ is the number of markers used). Linkage mapping was performed using Windows QTL Cartographer Version 2.5.

**Transgenic analysis in Arabidopsis**

*Arabidopsis thaliana* used in this study was the Col-0 ecotype (provided by Dr. Yongming Zhou, Huazhong Agricultural University, Wuhan, China). Wild-type and transgenic Arabidopsis seeds were surface sterilized and sown on 0.5× MS medium containing 1% sucrose (w/v) and 0.8% agar (w/v). All plants were grown at 22 °C with a 16/8-h light/dark photoperiod.

The DNA for the genotyping analysis was extracted from the leaves of transgenic lines and wild type. For RNA extractions, 15-d-old seedlings grown on 0.5× MS were collected. Total RNA was prepared from fresh tissues at 15 d after sowing using an RNA extraction kit (BioTeke, China). For RT-PCR, the first-strand cDNA was synthesized from 4 μg total RNA using the TransScript One-Step gDNA Removal and cDNA Synthesis SuperMix kit (TransGen, China). Semi-quantitative PCR was performed for the gene expression analysis using gene-specific (15F-1 and 9R-8) and *A. thaliana* ACTIN (ATACTIN-F and ATACTIN-R) primers. The primer sequences are listed in Table S5.

The open reading frame of *ZmGS5* was amplified from the cDNA of maize inbred line B73 leaves by PCR using the gene-specific primers 15F-1M and 9R-8M (Table S5). The amplified fragment of *ZmGS5*, bearing the EcoRI and XbaI restriction sites, was inserted into the pEASY-Blunt vector (TransGen, China). Positive clones were sequenced using the primers M13F and M13R (Table S3). After sequencing, the cDNA was cloned into the expression vector pBinGlyRed3 (Zhang et al. 2013), which contained a DsRed marker for transgenic plant selection. The seed-specific glycinin promoter in pBinGlyRed3 was replaced by the CaMV35S promoter (provided by Dr. Chunyu Zhang, Huazhong Agricultural University, Wuhan, China). The final construct was introduced into *A. tumefaciens* strain GV3101 (provided by Dr. Yongming Zhou, Huazhong Agricultural University, Wuhan, China) for Arabidopsis transformations (Zhou et al. 2003).

**Figure 5. Transgenic validation of ZmGS5 in Arabidopsis**

(A) The seed weights from the third generation of ZmGS5-1/ ZmGS5-13/ ZmGS5-20 transgenic lines and wild type. The bars represented the average seed weight based on five repeats and the error bars stood for the standard errors, the significant differences were based on t-test ($P < 0.05$; **$P < 0.01$). (B) RNA expression of ZmGS5 in different transgenic lines and WT, as determined by RT-PCR for different cycles. (C) Representative embryos from seeds of ZmGS5-1/ ZmGS5-13/ ZmGS5-20 transgenic lines and WT. (D) Embryo cells in central region of cotyledons from seeds of ZmGS5-1/ ZmGS5-13/ ZmGS5-20 transgenic lines and WT. (E) Cotyledon and radial size of mature embryos from the transgenic lines and WT. The bars represent mean ± standard errors and the significant differences were based on t-test, with $n > 50$ for cotyledon size analysis, $n > 20$ for radical size analysis ($P < 0.05$; **$P < 0.01$). (F) Cell size and average cell number of mature embryos from the transgenic lines and WT. The bars represent mean ± standard errors and the significant differences were based on t-test, $n > 20$ ($P < 0.05$; **$P < 0.01$).
For cytological observations, embryos and seeds of wild-type and transgenic lines were soaked in water for 4 h and then cleared for 4 h in Hoyer’s solution containing 8:1:2 (w/v/v) hydrate/glycerol/water (Yin et al. 2012). All samples were observed and photographed using a Nikon-Eclipse 80i differential interference contrast microscope equipped with a CCD camera. The sizes of cotyledons and radicals were measured with ImageJ (http://rsbweb.nih.gov/ij/index.html). Cell size was determined using the average cell size in areas of >2,000 μm² and the cleared cell number. The cell number was calculated by dividing the cotyledon size by the embryo cell size (Cheng et al. 2013).

ACKNOWLEDGEMENTS

This research was supported by the National Natural Science Foundation of China (31222041) and the National Hi-Tech Research and Development Program of China (2012AA10A307).

REFERENCES


ZmGS5 is associated with kernel development

951

www.jipb.net

November 2015 | Volume 57 | Issue 11 | 943–953


Melchinger, AE, Utz HF, Schon CC (1998) Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. Genetics 149: 383–403


Zhang Y, Liu J, Xia X, He Z (2014) TaGS-D1, an ortholog of rice OsGS3, is associated with grain weight and grain length in common wheat. Mol Breeding 34: 1097–1107


SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher’s web-site.

Figure S1. The trends of 1,000-seed weight from the second generation of 14 transgenic lines and wild type

Figure S2. GO analysis of 67 genes whose expressions were regulated by ZmBAK1-7

Table S1. 26 paralogs of ZmGS5

Table S2. Significant associations between SNPs and kernel-related traits

Table S3. QTLs identified within bin 3.04 and 7.03 for kernel related traits in previous studies

Table S4. The genes whose expressions were regulated by ZmBAK1-7

Table S5. The primers used in this study