Gene 678 (2018) 23–32

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

Molecular characterization of the *TaWTG1* in bread wheat (*Triticum aestivum* L.)

Han Zhang[a,1], Jian Ma[a,*,1], Jiajun Liu[a], Yang Mu[a], Huaping Tang[a], Yaxi Liu[a], Guangdeng Chen[b], Qiantao Jiang[a], Guoyue Chen[a], Yuming Wei[a], Youliang Zheng[a], Xiujin Lan[a,*]

[a] Triticeae Research Institute, Sichuan Agricultural University, Chengdu, Sichuan 611130, China
[b] Institute of Ecological and Environmental Sciences, Sichuan Agricultural University, Wenjiang, Chengdu 611130, China

ARTICLE INFO

Keywords:
Expression pattern
Structure
Transcription factor
Stress
Comparative analysis

ABSTRACT

OsWTG1 (LOC_Os08g42540.1) functions as an important factor determining grain size and shape in rice. Our understanding on its ortholog in wheat, *TaWTG1*, is limited. Here, we identified and mapped *TaWTG1* in wheat, characterized its gene and protein structures, predicted transcription factor binding sites of its promoter, and the expression patterns was also analysed bases on real-time quantitative PCR and public available microarray data. The *WTG1* orthologs in barley (*HvWTG1*), rice (*OsWTG1*), *Aegilops tauschii* (*AtWTG1*), *Triticum urartu* (*TuWTG1*), *Triticum turgidum* (*TtWTG1*) and *Brachypodium distachyon* (*BdWTG1*) were also identified for comparative analyses. *TaWTG1* was mapped onto the short arms of group 7 chromosomes (7AS, 7BS, and 7DS). Multiple alignments indicated that *WTG1* possesses eight exons and seven introns in all of the orthologs, except for the orthologs on 7A of wild emmer and on 7D of *A. tauschii* (seven exons and six introns). An exon–intron junction composed of intron 2 to intron 3 and exon 2 to exon 4 was highly conserved. The protein of WTG1 exists a conserved domain (Peptidase_C65). *WTG1* was mainly expressed in wheat roots, spikes and grains, in barley caryopsis and roots, and in rice anthers. Drought and heat stresses significantly regulated the expression of *TaWTG1* in wheat. In barley, *WTG1* was significantly down-regulated under Fusarium at late stage. In addition, significant correlations between the expression patterns of predicted transcription factors and *WTG1* were also detected. Overall, the results presented here broaden our knowledge on *WTG1* and will be helpful for its manipulation aiming at dissecting its function in plants.

1. Introduction

Common wheat (*Triticum aestivum* L., 2n = 6x = 42, AABBDD) is a major staple crop consumed by the world population. Improving wheat yield is one of the most important targets of breeding work. However, the current annual rate of increase in wheat productivity is considerably lower than that in the food demand by human and even seems to be stagnating in some regions (Reynolds et al., 2012). Therefore, it is essential to identify and introduce genes or alleles associated with high yield in wheat breeding programs.

Wheat yield is a complex quantitative trait which could be intensively influenced by interacting genetic and environmental factors. Wheat grain size and shape are key agronomic traits determining grain yield and milling performance (Gegas et al., 2010). Quantitative trait loci (QTL) mapping provides a useful tool for dissecting genetic components, and it has been extensively used to study yield-related traits in wheat. To date, numerous QTLs correlated with grain size and weight have been identified in wheat (Brinton et al., 2017; Cui et al., 2016; Su et al., 2016; Li et al., 2015; Wu et al., 2015). However, as the huge size and highly repetitive nature of the wheat genome and the lack of complete gene annotations, it is still hard to directly isolate yield-related genes by map-based cloning.

In comparison to wheat, quite a few QTLs controlling grain size and shape have been isolated and characterized because of the complete genome sequences in rice, such as *TGW3/TW3/GL3.3* (Ying et al., 2018; Hu et al., 2018; Xia et al., 2018), *GS9* (Zhao et al., 2018), *GSN1* (Guo et al., 2018), *qLGY3* (Liu et al., 2018) and *GS5* (Xu et al., 2015). It is widely believed that the linear order of genes among different grass

Abbreviations: QTL, Quantitative trait loci; TFBS, transcription factor binding sites; CDS, coding sequences; ESTs, Expression sequence tags; RPKM, Reads per kilobase per million reads; tpm, transcripts per million

* Corresponding authors.
E-mail addresses: jianma@sicau.edu.cn (J. Ma), lanxiujin@163.com (X. Lan).
[1] Contributed equally to this paper.

https://doi.org/10.1016/j.gene.2018.08.010
Received 24 April 2018; Received in revised form 15 July 2018; Accepted 3 August 2018
Available online 04 August 2018
0378-1119/

including rice, wheat, barley, *Brachypodium*, and sorghum is well conserved (Choi et al., 2004; Delseny, 2004). In recent years, the wheat genomics landscape has changed radically and increasing public resources including complete genome sequences and high-quality gene models are becoming available (Clavijo et al., 2017; Uauy, 2017). Comparative genomics provide a possibility for homology-based cloning of yield-related genes in wheat. For instance, *TaGS5*, the orthologous gene of *OsGS5* in rice, was isolated and found to be associated with grain weight in wheat and exhibited a large potential application in wheat high-yield breeding programs (Wang et al., 2016; Wang et al., 2015; Ma et al., 2015).

WIDE AND THICK GRAIN 1 (*WTG1*), plays a role in determining grain size and shape in rice and encodes an otubain-like protease with deubiquitination activity in rice (Huang et al., 2017). To date, *WTG1* in wheat has not been well reported. In this study, we identified and chromosomally mapped *WTG1* gene in wheat, characterized its nucleotide sequence and protein structures, predicted its *cis*-elements and transcription factor binding sites (TFBS), also analysed its expression by real-time PCR and employing available transcriptome data from public databases.

2. Materials and methods

2.1. Collection of sequence data

Pseudomolecules and coding sequences (CDS) of the wheat cultivar 'Chinese Spring' (International Wheat Genome Sequencing Consortium Reference Sequence v1.0) were downloaded from http://wheat-urgi.versailles.inra.fr/ hosted by Unité de Recherche Génomique Info (URGI). Assembled whole genome sequences of *B. distachyon* v2.0 and rice (*Oryza sativa*) were obtained from Phytozome v10 and v9.0, respectively (http://www.phytozome.org/). Pseudomolecules and CDS of *A. tauschii* (AetV4.0) were downloaded from https://www.ncbi.nlm.nih.gov/assembly/GCA_002575655.1/#/def_asm_Primary_Assembly (Luo et al., 2017). Whole genome shotgun sequences of *T. urartu* were downloaded from the National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/) in 2013. Assembled whole genome and CDS data of *T. turgidum* were obtained from https://www.dropbox.com/sh/3dm05grokhl0nbv/AAC3wvlYmAher8fY0srX3gX9a?dl=0%22 (Avni et al., 2017). The assembled barley genome sequences and CDS were downloaded from http://webblast.ipk-gatersleben.de/barley_ibsc/downloads/ (Mascher et al., 2017). CDS data for *T. urartu* was downloaded from GIGA_DB (http://gigadb.org/) in 2013. The *B. distachyon* v2.1genome assembly was obtained from Phytozome v10, and the rice genome was obtained from Phytozome v9.0.

2.2. Identification of WTG1 genes and chromosomal locations of WTG1 genes

We blasted the CDS regions of the reported *WTG1* (*LOC_Os08g42540.1*) (Huang et al., 2017) of rice against the CDS database for wheat, barley, *A. tauschii, T. turgidum, T. urartu*, and *B. distachyon*. Full-length genomic sequences was determined using the BLAST++ BLASTN algorithm with an E value cut-off of 10^{-5} (Ma et al., 2013b) based on the above genome sequences. The 1000 bp of 5′ flanking sequence of *WTG1* was isolated as the promoter sequences. The collected gDNA regions for *WTG1* sequences from each species were blasted against the physical maps (E value cut-off of 10^{-5}; Ma et al., 2013b) to identify their locations.

2.3. Isolation of the WTG1 gene in wheat

To validate the retrieved *WTG1* orthologous sequences in wheat obtained by BLAST, we further isolated the orthologs from wheat cultivar 'Chinese Spring'. The seedling leaves at 20 days after germination were collected for total DNA extraction from leaves. Thirteen sets of primers (Table 1 and Fig. 1a) were designed to isolate the full-length of *TaWTG1* sequence based on the genomic sequences of wheat cultivar 'Chinese Spring', *T. urartu*, *A. tauschii*, and *T. turgidum* (Table S1). Twenty microliters PCR reaction mixtures contained 10 μL mix (which from Vazyme Biotech Co., Ltd), 100 ng of genomic DNA, 0.5 μM of each primer and 7.5 μL ddH_2O. The PCR cycling parameters were 94 °C for 5 min for pre-denaturation, followed by 35 cycles of 94 °C for 45 s, 53–63 °C as given in Table 1 for 40 s, 72 °C for 5 min, and a final extension at 72 °C for 15 min. The annealing temperature of each primer is listed in Table 1. Amplification products were separated on 1.5% agarose gels, and target fragments were purified using QIAquick Extraction Kit (QIAGEN, Hilden, Germany). The purified PCR products were sent to Tsingke Biotech (Beijing) Co., Ltd. for gene cloning and sequencing. At least three independent clones for each amplicon were sequenced in both directions.

2.4. Nucleotide and protein structures of WTG1

Intron-exon structures of *WTG1* in different species were determined using GSDS 2.0 (Hu et al., 2015). Sequence alignments were carried out using DNAman 6.0 (Lynnon Biosoft, San Ramon, CA, USA) to assess the similarity between sequences in diverse species including each intron and exon. Information of genomic and coding sequences is listed in Table S1. The amino acid sequences were translated by MEGA6.0 (Tamura et al., 2013) according to CDS and were uploaded into SMART (Letunic et al., 2015) for protein motifs analysis.

2.5. Sequence alignments and phylogenetic analysis

Multiple alignments of coding sequences for different species were performed by MEGA 6.0; and further improved by visual examination and manual adjustment. The phylogenetic tree was generated using CDS of *H. vulgare* (*HvWTG1*), *A. tauschii* (*AtWTG1*), *T. urartu* (*TuWTG1*), *T. turgidum* (*TtWTG1-7* and *TtWTG1-7B*), *T. aestivum* (*TaWTG1-7AS*, *TaWTG1-7BS*, and *TaWTG1-7DS*), *B. distachyon* (*BdWTG1*) and *O. sativa* (*OsWTG1*). A dendrogram for the evolutionary relationship of the *WTG1* genes among different species was generated using MEGA 6.0 with the neighbour-joining method (Tamura et al., 2013).

Table 1
Details of primers used in this study.

Primer	Sequence (5′-3′)	Annealing temperature (°C)	Description
A-F2	CGATTGATCCAGCAGTATGT	59 °C	Isolation
AB-R2	ACTGGAACTCTGCTGCTAAG	59 °C	Isolation
ABD-F3	CTTAGCAGCAGAGTTCCAG	60 °C	Isolation, RT-qPCR
A-R3	CTGAACCACGGTCGAATTTGTC	60 °C	Isolation
ABD-F4	ATGTTCATTGAGGAGCTGC	60 °C	Isolation
A-R4	AGCATGCTCTGATTACGC	60 °C	Isolation
B-R1	AAACCCTGGCATAGACTAGGC	63 °C	Isolation
B-F2	CTATTGGTCCAGCAGTATGC	59 °C	Isolation
B-R3	CTGAGCCACGGTCGAATTTGTC	60 °C	Isolation
B-R4	CTAGAGCACACTTACAACAAGG	60 °C	Isolation
D-F1	CAGCGTCACGAAGTACATCATG	59 °C	Isolation
D-R1	GCACTGTGTCTACTGATGTC	59 °C	Isolation
D-F2	GACATCAGTAGACACAGTGC	53 °C	Isolation
D-R2	CTCCTCGTGTTCGCCTTA	53 °C	Isolation
D-R3	CTGAGTCACGGTCGAATTTGTC	60 °C	Isolation
D-R4	GAGGCACCATACCTAATG	60 °C	Isolation
D-F5	TATGCAACACTGGGTGGTGG	56 °C	Isolation
D-R5	CGAATTTGAGGGCAGTGAC	56 °C	Isolation
ABD-R1	GGAGAACATAAAGCTTCG	60 °C	RT-qPCR
β-actin-F1	ACCTTCAGTTGCCCAGCAAT	60 °C	RT-qPCR
β-actin-R1	CAGAGTCGAGCACAATACCAGTTG	60 °C	RT-qPCR