Molecular cloning, characterization and expression of *WAG-2* alternative splicing transcripts in developing spikes of *Aegilops tauschii*

SHUHONG WEI

Key Laboratory of Southwest China Wildlife Resources Conservation (Ministry of Education), China West Normal University, Nanchong, Sichuan 637009, People’s Republic of China

Abstract

*WAG-2* is a C-class MADS-box gene, which is orthologous to *AGAMOUS (AG)* in *Arabidopsis*. The AG group C-class MADS-box genes are involved in stamen and pistil identity. In this study, two *WAG-2* transcripts, namely, *WAG-2f* and *WAG-2g*, were isolated and characterized from *Aegilops tauschii*. The open reading frames of *WAG-2f* and *WAG-2g* were 825 and 822 bp, respectively, encoding 275 and 274 amino acid residues. BLAST searches of partial *WAG-2* genomic sequence against the draft sequence of *Ae. tauschii* genome database revealed the complex structure of the *WAG-2* gene, which consisted of seven exons and six introns. The *WAG-2f* and *WAG-2g* cDNAs were two alternative splicing transcripts. The alternative splicing events were produced by an alternative 5′ splice site. The expression level of *WAG-2f* transcript, which was extremely weak in young spikes of floret primordium formation stage, increased as the spikes developed. The highest expression was observed in the spikes at the anther separation stage. Low expression levels of *WAG-2f* were also detected at the tetrad stage. The *WAG-2g* transcript was expressed at all four stages of spike development but at a relatively low level. The expression pattern of the two transcripts was distinctly different during floral development, thereby suggesting a functional divergence.

Introduction

Flower development has been the subject of intensive genetic and molecular studies over the past few decades (Hama et al. 2004). Studies on *Arabidopsis* and *Antirrhinum* have provided the ABC model to understand the floral organ development in higher plants (Davies and Schwarz-Sommer 1994; Weigel and Meyerowitz 1994; Theissen and Saedler 1999; Theissen 2001). In the ABC model, floral organ identity is defined by three classes of homeotic genes, namely, A, B and C. A-class genes specify sepals in the first whorl, A-class genes and B-class genes specify petals in the second whorl, B-class genes and C-class genes specify stamens in the third whorl and C-class genes specify carpels in the fourth whorl. All these genes belong to the MADS-box gene family (Riechmann and Meyerowitz 1997), except for the A-class gene *AP2*. The ABC model is later extended to the ABCD model on account of D function genes, which specify ovules (Angenent and Colombo 1996). The ABCD model is now extended to the ABCDE model, in which an E function is involved in the specification of petal, stamen, carpel identity and possible ovule identity (Theissen 2001). Studies on MADS-box genes in monocot grasses indicated that the ABC model can essentially be extended to rice (Kang et al. 1998; Kyozuka et al. 2000; Kyozuka and Shimamoto 2002) and maize (Ambrose et al. 2000).

As explained by the ABC model, C-class MADS-box genes are required for stamen and pistil identity. Wheat *AGAMOUS (AG)* orthologue *WAG-2* is a C-class gene. Previous studies showed that *WAG-2* transcript is preferentially detected in fertile and sterile pistils, as well as in the ovule primordium. The transformed stamens also show intensive accumulation of *WAG-2* transcripts (Mizumoto et al. 2009). These results indicated that the *WAG-2* is mainly correlated with pistil development rather than pistil differentiation.

Common wheat (*Triticum aestivum*) is an allohexaploid with genome AABBDD, and each genome originates from three diploid species (Feldman 2001). Therefore, the hexaploid wheat genome generally contains triplicated homeologous genes derived from the ancestral diploid species. Goat grass, *Aegilops tauschii* (*2n = 2x = 14, DD*) is considered the D-genome donor of bread wheat. In the present study, *WAG-2* transcripts were isolated and characterized from *Ae. tauschii*. The expression patterns

Keywords. *WAG-2* gene; alternative splicing; expression pattern; functional divergence; *Aegilops tauschii*. 
of WAG-2 transcripts at different developmental stages of young spikes were examined. This study aimed to investigate the sequence features of WAG-2 gene in Ae. tauschii, determine its expression pattern from spikes at various developmental stages and present evidence for alternative splicing of transcripts from WAG-2 in developing spikes. Knowledge on WAG-2 alternative splicing transcripts in Ae. tauschii will facilitate functional research on WAG-2 gene in hexaploid wheat.

Materials and methods

Plant materials

One accession of Ae. tauschii (2n = 2x = 14, DD) was used in this study. Fresh young spikes at pistil and stamen primordial differentiation stages (spikes ~10 mm) were used for total RNA isolation. Young spikes at various developmental stages were selected for real-time PCR analysis.

RNA isolation

Total RNAs were extracted using the EZgene Plant RNA Isolation kit (Biomiga, Shanghai, China). The total RNAs were synthesized into first-strand cDNA by using the manufacturer’s instructions (Clontech, Palo Alto, USA).

WAG-2 cDNA cloning

Rapid amplification of cDNA ends (RACE) method was employed to amplify 3′UTR and 5′UTR sequences, whereas 3′RACE fragment and 5′RACE fragment were spliced with the intermediate fragment. The primers (WAG-2F and WAG-2R) were designed to clone WAG-2 partial coding sequence (CDS). The 5′RACE inner primer (R240-1), 5′RACE outer primer (R240-2), 3′RACE inner primer (112-1) and 3′RACE outer primer (112-2) were designed to clone 5′UTR and 3′UTR after the partial CDS was obtained. WAG-2 full-length cDNA sequence was spliced with fragment overlapping areas and submitted to GenBank. On the basis of WAG-2 mRNA sequence, the primers (WAG-2P1 and WAG-2P2) were designed to clone a WAG-2 complete cDNA. All the primers were designed through Oligo 7.0 and synthesized by Shanghai Sangon Biological Engineering Technology and Services (table 1).

Sequencing reactions and analysis

PCR amplification was conducted in a thermocycler (My-Cycle, Biorad, San Diego, USA). The PCR cycling included pre-denaturation at 94°C for 5 min followed by 35 cycles of 94°C for 30 s, 60°C for 30 s and 72°C for 1 min, and then a final extension at 72°C for 10 min. Target DNA bands were recovered and purified from the gels with Qiaquick Gel Extraction kit (Qiagen, Shanghai, China). Purified PCR products were cloned in a pMD20T vector following the manufacturer’s instructions (Takara, Dalian, China). Transformants were plated on LB agar containing ampicillin. Clones with inserts were identified by blue/white colony selection. Positive clones were then screened and sequenced by Taihe Biotechnology (Beijing, China).

WAG-2 gene open reading frame (ORF) was predicted by DNA star; thus, the amino acid sequence was deduced. The sequence alignments were conducted using ClustalW.

Real-time PCR

Total RNAs were isolated from spikes at various developmental stages. cDNA was synthesized using PrimeScript Perfect real-time RT reagent kit (TaKaRa, Dalian, China). The primers WAPF and WABF were designed with Primer Express 2.0 software to amplify short fragments of WAG-2. TaqMan probe Tzs were labelled with FAM at the 5′end and with TAMR at 3′end. The primers and probe sequences are listed in table 1.

All the samples were analysed in triplicate, and the fold change in the number of RNA transcripts was calculated through 2−ΔΔCt method with the wheat housekeeping genes ubiquitin (DQ086482) and actin (AB181911) acting as internal controls.

Results

Molecular cloning of full-length WAG-2 cDNAs

WAG-2 partial CDS was successfully cloned using WAG-2F and WAG-2R. RACE method was used to amplify 3′UTR and 5′UTR sequences to obtain WAG-2 full-length cDNA sequence. The lengths of partial CDS, 3′RACE product, and 5′RACE product were 425 or 422, 459 and 514 bp, respectively.

Nucleotide sequencing of a subset of the cloned PCR fragments revealed two distinct but very similar WAG-2 gene sequences from 1234 and 1231 bp (accession numbers KT188778 and KT188779, respectively). The corresponding gene sequences were named WAG-2f and WAG-2g, to keep with the previously identified WAG-2 gene sequences from Chinese Spring (accession numbers AB465688). WAG-2f

---

**Table 1. Primers and probes for WAG-2 cloning and expression from Ae. tauschii.**

<table>
<thead>
<tr>
<th>Primer or probe</th>
<th>Sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>WAG-2F</td>
<td>5′-CATTTGAGAGGTCAAAAGGCT-3′</td>
</tr>
<tr>
<td>WAG-2R</td>
<td>5′-GGGAAGTTTCTGGAAATCGAAG-3′</td>
</tr>
<tr>
<td>R240-1</td>
<td>5′-TCCAGCTGGCCTTCTA-3′</td>
</tr>
<tr>
<td>R240-2</td>
<td>5′-GGCTTGAGGACCTCTGCTTG-3′</td>
</tr>
<tr>
<td>112-1</td>
<td>5′-CCCTAGGACGCAAGGTCTGAGTAAG-3′</td>
</tr>
<tr>
<td>112-2</td>
<td>5′-CCCTAGGACGCAAGGTCTGAGTAAG-3′</td>
</tr>
<tr>
<td>WAG-2P1</td>
<td>5′-GCACTCCTTCTTTTCTCTTGATC-3′</td>
</tr>
<tr>
<td>WAG-2P2</td>
<td>5′-TACCAAGCAAATTTAGAATTGGTACC-3′</td>
</tr>
<tr>
<td>WAPF</td>
<td>5′-AGATCCAACTCCTGCAAGG-3′</td>
</tr>
<tr>
<td>WAPR</td>
<td>5′-TGCTATGCTTTTCTGGAC-3′</td>
</tr>
<tr>
<td>TZ1</td>
<td>5′-TCAGTGTGCAAAACTCAACAGTA-3′</td>
</tr>
<tr>
<td>TZ2</td>
<td>5′-TCAGTGTGCAAAACTCAACAC-3′</td>
</tr>
</tbody>
</table>

---

Shuhong Wei
Analysis of WAG-2 alternative splicing transcripts in Ae. tauschii

Figure 1. Protein sequence alignment of WAG-2f and WAG-2g in Ae. tauschii. The positions of the MADS-box and K-domain are indicated by red and blue boxes, respectively. The change of serine in K-domain is indicated by an arrow.

Figure 2. Schematic diagrams of alternatively spliced transcripts structure of WAG-2 gene from Ae. tauschii. Exons are indicated by gray rectangles and numbered E1–E7, while introns are indicated by black thin lines and numbered I1–I6. Black rectangles represent constitutively spliced sequences. The dashed lines represent splicing option.

and WAG-2g ORF lengths were 825 and 822 bp, respectively, encoding 275 and 274 amino acid residues.

A comparison between the corresponding gene sequences and Arabidopsis AG identified the characteristic protein domains of MADS-box protein. They all contained the characteristic modular structure of a highly conserved MADS box (M), weakly conserved intervening (I), keratin-like (K) and C-terminal domains (C). The similarity between WAG-2f and WAG-2g in M, I and C domains was typically striking, with nearly 100% identity in these regions to one another. However, a distinct sequence divergence that was 3 bp insertion/deletion (+/−) mutation, namely, 5′-AGT-3′, was identified in K domain, thereby resulting in the change of Serine in146 site (figure 1).

Structure of WAG-2 genomic sequences in Ae. tauschii

Basic local alignment search tool (BLAST) search of the partial WAG-2 genomic sequence (GenBank accession number JF330187) against the genome sequence of Ae. tauschii genome database (https://urgi.versailles.inra.fr/blast/blast.php) revealed that the partial WAG-2 sequence of Ae. tauschii showed 100% homology to the TGAC_WGS_tauschii_v1_contig_109481. The results from the DNA sequence comparison and assembly showed that the whole genomic sequence of WAG-2 in Ae. tauschii was 7315 bp. Alignment with the corresponding cDNA sequences revealed the complex structure of WAG-2 gene, which consisted of seven exons and six introns (figure 2).

Alternative splicing of WAG-2 genes

A total of 20 cDNAs from young spikes (~1 mm in length) through RT-PCR of WAG-2 were cloned and sequenced to assess the possibility that the two amplicons observed in the WAG-2 RACE reaction were produced by alternative splicing. The sequence alignments of the 20 cDNAs with one another and with the full-length WAG-2 cDNA of Ae. tauschii revealed that the transcript population can be split into two distinct groups. Among the cDNA sequences, 15 were identical and the length of ORF was 825 bp. The remaining six cDNAs contained the regions of identity to one another, and the length of ORF was 822 bp. The variant in the first group was WAG-2f, and the isolate represented by the second cDNA group was WAG-2g. WAG-2f and WAG-2g cDNAs were two alternative splicing transcripts. Alternative splicing events were produced by the alternative 5′ splicing
Expression analysis of two alternative splicing transcripts: WAG-2f and WAG-2g

Real-time PCR was conducted on RNA templates at four stages of spike development to determine the expression patterns of the two alternative splicing genes. As shown in figure 3, the expression patterns of the two WAG-2 genes were distinctly different from each other. The expression level of WAG-2f was extremely weak in young spikes at floret primordium formation stage; however, it increased as the spikes developed. The highest expression was observed in spikes at the anther separation stage. The low expression level of WAG-2g transcript was also detected at the tetrad stage. The WAG-2g transcript was expressed at all four stages of spike development, but at relatively low level. The expression patterns of the two genes were different during floral development. This finding suggested a functional divergence.

Discussion

Alternative splicing is a process through which multiple mRNAs are generated from a single pre-mRNA, thereby resulting in functionally distinct proteins. Recent genomewide analyses of alternative splicing indicated that alternative splicing in high eukaryotes is an important mechanism that generates proteomic complexity and regulates gene expression (Macknight et al. 2002; Cheng et al. 2003; Lightfoot et al. 2008).

Extensive studies demonstrated that alternative splicing is most prevalent in animals and plants. In humans, up to 95% pre-mRNAs from multieoxon genes undergo alternative splicing (Pan et al. 2008; Wang et al. 2008), whereas this rate is >40% in plants (Filichkin et al. 2010). In Arabidopsis, ~21.8% of the genes show alternative splicing events. In rice, ~21.2% of the expressed genes are alternatively spliced (Wang and Brendel 2006). Five Rose MASAKO C transcripts belonging to AG subclass are related to alternative splicing events (Kitahara and Matsumoto 2000). Splicing variants are also detected in at least 20 of ~100 Arabidopsis MADS-box transcripts, including AG (Cheng et al. 2003; Parenicova et al. 2003; Wang and Brendel 2006). Although many alternatively spliced MADS-box genes appear in the literature and database compilations, none of them were obtained from Ae. tauschii and wheat.

In the present study, 20 transcripts from Ae. tauschii ~1 mm whole spikes were cloned. All the transcripts can be assigned to one of the two groups based on the sequence similarity with one another. The transcripts are identical to one another in coding sequences within each group; however, they differ in the presence or absence of certain three bases of AGT at 5′ of exon 4. Moreover, Ae. tauschii is a self-pollinated plant, and its genotype within a population is homozygous rather than heterozygous. These results suggest that WAG-2f and WAG-2g are two alternative splicing transcripts.

A critical question is whether the alternative transcripts perform a biological function in floral organ development. If the alternatively spliced transcripts are translated, the modified proteins may demonstrate novel properties, thereby changing their roles or allowing them to participate in new molecular pathways. Analyses of the expression patterns of the two transcripts WAG-2f and WAG-2g at the four stages of spike development suggested a role for Ae. tauschii WAG-2 cDNAs in spike development. From floret primordium formation to tetrad stage, WAG-2f plays a vital role as spikes develop, and the highest expression is observed at the anther separation stage. Nevertheless, the expression of WAG-2g transcript is extremely weak at the spikes. The young spikes measuring <3 mm in length are in the floret primordium formation, whereas those 3–10 mm in length are in the floret differentiation stage when the stamen and pistil primordia are initiating. The floral organs begin to mature in the spikes at the anther separation stage to the tetrad stage (10 mm in length). Therefore, the results suggest that the WAG-2f transcript is involved in the development of floral organs rather than in their differentiation. The two transcripts WAG-2f and WAG-2g share different functions in spikes during floral organ development. However, whether the alternatively spliced transcript WAG-2g plays important roles in other tissues, such as vegetative organs, needs further verification.

Nonetheless, the amplification of transcripts from whole spikes provides limited information, similar to the data on complex tissues or organs. Other alternative splicing transcripts for Ae. tauschii WAG-2 gene may exist when several different tissues or organs, i.e. stamen and pistil in different development processes are accessed. Thus, this topic needs further research.
Analysis of WAG-2 alternative splicing transcripts in *Ae. tauschii*

**Acknowledgements**

This work was financially supported by the National Natural Science Foundation of China (31501304) and the Doctor Startup Foundation Project of China, West Normal University (13E002).

**References**


Received 15 October 2015, in revised form 8 December 2015; accepted 14 December 2015

Unedited version published online: 17 December 2015

Final version published online: 19 July 2016

Journal of Genetics, Vol. 95, No. 3, September 2016