cDNA-AFLP analysis of gene expression differences between the flower bud and sprout-shoot apical meristem of Angelica sinensis (Oliv.) Diels

Guang Yu¹, Jinao Duan¹, Hui Yan¹, Bingsheng Song² and Ziqing He²

¹Nanjing University of Chinese Medicine, Nanjing, China. ²Gansu Mingui Chinese Medicinal Materials Technology Co., Ltd., Lanzhou, China.

Abstract

Angelica sinensis (Oliv.) Diels (Umbelliferae) is a well-known medicinal plant mainly distributed in Gansu Province of China. Its local and global demand is significant because of its food and medicinal applications. However, the early bolting rate of Angelica sinensis (Oliv.) Diels reaches 20%-60%, which seriously affects its food and medicinal qualities. Thus, differences in gene expression between the flower bud and sprout-shoot apical meristem underwent analysis, by means of cDNA-amplified restriction fragment length polymorphism, to better understand the flowering mechanism. 64 primer sets, each of which amplified to 60 transcript-derived fragments (TDFs), were used. Among these TDFs, 26 were expressed specifically in the flower bud. After cloning and sequencing, 32 distinct sequences were obtained from these 26 TDFs, and 25 were found with homologous sequences in databases. Confirmation of differential expression of 13 sequences was obtained by semi-quantitative RT-PCR, their showing higher expression levels in flower buds. These homologous sequences encode transposable elements, pentatricopeptide repeat-containing proteins, DNA-binding transcription factors, zinc finger (B-box type) family proteins, NADP-dependent sorbitol 6-phosphate dehydrogenase (S6PDH), amongst others.

Key words: cDNA-AFLP, early bolting, Angelica sinensis (Oliv.) Diels, gene expression, TDFs.

Received: June 25, 2010; Accepted: February 7, 2011.
cal meristems were gathered from plants in Gansu Province. Flower buds of early-bolting plants, each around 0.5 cm in size, were harvested prior to full-bloom, together with sprout-shoot apical meristems from normal ones. The samples were frozen in liquid nitrogen and stored at -80 °C. Total RNA was extracted from the plant material, by using the Trizol extraction method (Invitrogen, USA). RNA integrity and quantity were determined by running 2 μL of total RNA in a formamide denaturing gel. First and second strand cDNA synthesis was according to manufacturer's instructions (Promega Universal Riboclone cDNA Synthesis System, USA). The resulting double-stranded cDNA was separated on agarose gel to check the sizes of cDNA samples.

cDNA-AFLP was carried out according to the procedure of Habu et al. (1997) with minor modifications. About 500 ng of double-stranded cDNA underwent standard AFLP template production. The restriction enzymes used for cDNA digestion were EcoRI10U/μL, Sangon, China) and MseI10U/μL, Sangon, China). The digested products were ligated to adapters with the following sequences: EcoRI adapter - 5’-CTCGTAGACTGCTACC-3’ and 5’-AATTGAGCTAGTCGTC-3’; MseI adapter - 5’-GACGTAGATGCTCTGA CC-3’ and 5’-TACTCAGGACTGTC-3’. The ligated products were preamplified with the corresponding preamplification primers (EcoRI: 5’-GACTGCGTACCAATTCNN-3’, MseI: 5’-GATGAGTCCTGAGTAANN-3’). Twenty preamplification cycles were undertaken in a Bio-Rad MyCycler PCR system (Bio-Rad, USA). The PCR program was carried out as follows: 30 s at 94 °C, 30 s at 56 °C, and 1 min at 72 °C. Four sets of the EcoRI primer and sequenced for each, thus giving 130 randomly selected colonies. Among these, 32 were distinct, this implying that each TDF corresponded to 1.23 distinct nucleotide sequences. A number of DNA sequences of the same length have been reported to co-migrate as a single AFLP fragment (Chen et al., 2008). As these migrate to the same position, AFLP resolution potential is low, in terms of variable base composition in DNA molecules of the same length. Thus, the 32 sequences were subjected to homology searches using BLAST against the databases, whereby 26 different sequences were obtained (Table 1), with six not resulting in any hits, this implying their possible correspondence to novel genes related to early bolting in Angelica sinensis (Oliv.) Diels.

These 26 TDFs were amplified from the polyacrylamide gels and re-amplified with the original primer sets used for cDNA-AFLP analysis. After the isolated TDFs were cloned, five colonies were selected and sequenced for each, thus giving 130 randomly selected colonies. Among these, 32 were distinct, this implying that each TDF corresponded to 1.23 distinct nucleotide sequences. A number of DNA sequences of the same length have been reported to co-migrate as a single AFLP fragment (Chen et al., 2008). As these migrate to the same position, AFLP resolution potential is low, in terms of variable base composition in DNA molecules of the same length. Thus, the 32 sequences were subjected to homology searches using BLAST against the databases, whereby 26 different sequences were obtained (Table 1), with six not resulting in any hits, this implying their possible correspondence to novel genes related to early bolting in Angelica sinensis (Oliv.) Diels.

Primer pairs were designed for 13 of the 32 distinct sequences obtained (Table 2). Differential expression of
the 13 distinct sequences was analyzed using sqRT-PCR, whereby it was shown that these sequences were expressed at a higher level in early bolting *Angelica sinensis* (Figure 2). TDFs showed different levels of homology with the genes/cDNAs of other species (with an E value ranging from 6.1 to 2e⁻¹⁷).

cDNA-AFLP, a variation of AFLP and derived from RNA fingerprint identification technology, has already become a sophisticated research tool for identifying differences in gene expression. This technique was here employed with 64 primer sets, to compare gene-expression profiles of flower buds and sprout-shoot apical meristems. We obtained 32 different sequences, some of which are possibly applicable to controlling early bolting. In this study, sequences with low E-value and definite functions were focused upon, thereby resulting in the identification of an RF2 protein (TDF A104-1), homeobox protein 25 (TDF A035-4), CMGC Ser/Thr protein kinase family (TDF A011-4), NADP-dependent sorbitol 6-phosphate dehydrogenase (TDF A021-3), ATAF-like NAC-domain transcription factor (TDF A010-1), ATTRX H1 (TDF A110-1), and senescence-associated protein (TDF A115-1), all involved in cellular pathways leading to bolting.

TDF A104-1, as shown in Table 1, is matched with the RF2 protein. The *rf2* gene is one of the two nuclear genes required for fertility restoration in male-sterile T-cytoplasm (cmsT) plants. RF2 is an aldehyde dehydrogenase, thereby inferring several mechanisms that might explain Rf2-mediated fertility restoration in cmsT maize. Aldehyde dehydrogenase, possibly involved in the detoxification of acetaldehyde produced by ethanolic fermentation during pollen development, may also play a role in energy metabolism (Cui et al., 1996).

On the other hand, TDF A035-4 showed similarity with homeobox protein 25, which encodes ZFHD2, a member of the zinc-finger homeodomain transcription factor family. The cysteine-rich zinc-binding motifs are known as RING and B-box in several unrelated proteins. Structural, biochemical, and biological studies have revealed that these motifs may mediate protein-protein interactions (Borden, 1998).

TDF A011-4, which contains a protein kinase domain, belongs to the protein kinase superfamily (CMGC Ser/Thr protein kinase family, GSK-3 subfamily). It may mediate extracellular signals that regulate transcription in differentiating cells (Yamada, 2003). The expression patterns of SHAGGY-related protein kinase genes during wild-type *Arabidopsis* inflorescence development, detected by *in situ* hybridization, have been shown to be consistent with a possible role in floral meristem patterning. SHAGGY-related protein kinase gene transcripts were detected both at the periphery of the inflorescence meristem and within the floral meristem. At later stages, their expression became localized in specific regions of developing flower-organ primordia. The plants themselves developed

---

**Figure 1** - cDNA-AFLP autoradiogram of four of the most informative-primer combinations that amplified differentially expressed genes in flower buds and sprout-shoot apical meristems. The codes for the primers used in the combinations are shown shown in Table S1. The boxes emphasize differential-expressed fragments. (ZT: flower bud, ZC: sprout-shoot apical meristem).
flowers with a higher number of perianth organs and an alteration in the apical-basal patterning of the gynoecium (Dornelas, 2000).

TDF A021-3 is matched with the NADP-dependent sorbitol 6-phosphate dehydrogenase (S6PDH) gene. Sorbitol-6-phosphate is a major photosynthetic product of several members of the Rosaceae family.

Transposons are mobile DNA molecules existing in the genomes of many organisms. The transposon superfamilies of higher plants were introduced as including LTR

<table>
<thead>
<tr>
<th>AFLP fragment</th>
<th>dbEST_Id</th>
<th>GenBank accession number</th>
<th>Similarity</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>A117-1</td>
<td>700623271</td>
<td>HO056150</td>
<td>retrotransposon protein, putative, Ty1-copia subclass</td>
<td>5e-22</td>
</tr>
<tr>
<td>A111-1</td>
<td>70062325</td>
<td>HO056148</td>
<td>senescence-associated protein</td>
<td>1.1</td>
</tr>
<tr>
<td>A112-1</td>
<td>70062323</td>
<td>HO056146</td>
<td>Reverse transcriptase (RNA-dependent DNA polymerase)</td>
<td>9.7e-06</td>
</tr>
<tr>
<td>A110-1</td>
<td>70062322</td>
<td>HO056145</td>
<td>ATTRX H1, ATTRX1</td>
<td>0.24</td>
</tr>
<tr>
<td>A107-1</td>
<td>70062320</td>
<td>HO056143</td>
<td>Ribosomal protein S8 (rps8) gene</td>
<td>3e-54</td>
</tr>
<tr>
<td>A104-1</td>
<td>70062319</td>
<td>HO056142</td>
<td>RF2 protein</td>
<td>2e-44</td>
</tr>
<tr>
<td>A103-2</td>
<td>70062318</td>
<td>HO056141</td>
<td>GLR2.6 (member of Putative ligand-gated ion channel subunit family)</td>
<td>2.9</td>
</tr>
<tr>
<td>A102-3</td>
<td>70062317</td>
<td>HO056140</td>
<td>zinc finger (C3HC4-type RING finger) family protein / ankyrin repeat family protein</td>
<td>1e-02</td>
</tr>
<tr>
<td>A040</td>
<td>70062301</td>
<td>HO056124</td>
<td>transposable element gene; CACTA-like transposase family</td>
<td>0.54</td>
</tr>
<tr>
<td>A038-1</td>
<td>70062302</td>
<td>HO056125</td>
<td>pentatricopeptide (PPR) repeat-containing protein</td>
<td>26</td>
</tr>
<tr>
<td>A037-2</td>
<td>70062303</td>
<td>HO056126</td>
<td>Putative 22 kDa kafirin cluster; Ty3-Gypsy type</td>
<td>9e-14</td>
</tr>
<tr>
<td>A035-4</td>
<td>70062305</td>
<td>HO056128</td>
<td>Homeobox protein 25</td>
<td>1.7e-09</td>
</tr>
<tr>
<td>A034-3</td>
<td>70062306</td>
<td>HO056129</td>
<td>zinc finger (B-box type) family protein; Identical to the Zinc finger protein CONSTANS-LIKE 13 (COL13)</td>
<td>6.1</td>
</tr>
<tr>
<td>A033-4</td>
<td>70062307</td>
<td>HO056130</td>
<td>transposable element gene; gypsy-like retrotransposon family</td>
<td>5.5e-13</td>
</tr>
<tr>
<td>A116-1</td>
<td>70062308</td>
<td>HO056131</td>
<td>retrotransposon protein</td>
<td>1e-13</td>
</tr>
<tr>
<td>A109-2</td>
<td>70062309</td>
<td>HO056132</td>
<td>Prunus umbellata</td>
<td>2e-17</td>
</tr>
<tr>
<td>A011-4</td>
<td>70062311</td>
<td>HO056134</td>
<td>serine/threonine kinase-like protein</td>
<td>1.2e-02</td>
</tr>
<tr>
<td>A010-1</td>
<td>70062310</td>
<td>HO056133</td>
<td>ATAF-like NAC-domain transcription factor</td>
<td>6.4e-17</td>
</tr>
<tr>
<td>A009-1</td>
<td>70062312</td>
<td>HO056135</td>
<td>putative reverse transcriptase</td>
<td>7.9e-09</td>
</tr>
<tr>
<td>A008-4</td>
<td>70062314</td>
<td>HO056137</td>
<td>transposable element gene; non-LTR retrotransposon family (LINE)</td>
<td>2.5</td>
</tr>
<tr>
<td>A007-3</td>
<td>70062313</td>
<td>HO056136</td>
<td>transposable element gene, pseudogene, hypothetical protein</td>
<td>1.1</td>
</tr>
<tr>
<td>A006-4</td>
<td>70062315</td>
<td>HO056138</td>
<td>Photosystem II protein Z [Daucus carota] protein</td>
<td>4e-16</td>
</tr>
<tr>
<td>A001-4</td>
<td>70062316</td>
<td>HO056139</td>
<td>transposable element gene; copia-like retrotransposon family</td>
<td>4.2e-09</td>
</tr>
</tbody>
</table>

Figure 2 - Confirmation, by semi-quantitative RT-PCR, of a higher level of TDF expression in the flower bud of early-bolting Angelica sinensis (Oliv.) Diels. (ZT: flower bud, ZC: sprout-shoot apical meristem).
retrotransposons, hAT, CACTA elements, Mutator and MULEs, Tc1/mariner, miniature inverted repeat transposon MITEs, and so on. TDF A117-1, A112-1, A037-2, A033-4, A116-1, A109-2, A010-1 and A001-4 are matched with the transposable-element gene. TDF A010-1 and an ATAF-like NAC-domain transcription factor are homologous. Subtractive EST analysis, and the screening of cDNA libraries derived from *Brassica napus* leaves subjected to mechanical wounding, flea beetle feeding, or cold temperatures, revealed eight genes encoding NAC-domain transcription factors. These genes were found to be differentially regulated in response to biotic and abiotic stress, such as wounding, insect feeding, *Sclerotinia sclerotiorum* infection, cold shock and dehydration (Hegedus et al., 2003).

Transposable elements may have an important effect on earlier bolting in *Angelica sinensis*. TDF A110-1 is homologous to ATTRX H1. The latter encodes a cytosolic thioredoxin that reduces the disulfide bridges in target proteins, by the reversible formation of a disulfide bridge between the two neighboring Cys residues, present at the active site. Thioredoxins have been found to regulate a variety of biological reactions in prokaryotic and eukaryotic cells.

TDF A110-1 is homologous to ATTRX H1. The latter encodes a cytosolic thioredoxin that reduces the disulfide bridges in target proteins, by the reversible formation of a disulfide bridge between the two neighboring Cys residues, present at the active site. Thioredoxins have been found to regulate a variety of biological reactions in prokaryotic and eukaryotic cells.

TDF A115-1 is homologous to senescence-associated genes (SAGs). A comparison was undertaken of the expression of several *Arabidopsis thaliana* SAGs in attached and/or detached leaves, as a possible response to age, dehydration, darkness, abscisic acid, cytokinin, and ethylene treatments. For the majority, the response to most of the treatments was similar. Detachment in darkness and ethylene were the strongest inducers of both SAGs and visible yellowing. Detachment in light, although a strong inducer of SAGs, was not of visible yellowing. The other treatments varied more in their individual effects. Responses, examined in both older and newer leaves, were generally much stronger in the former. As individual SAGs differed from the norms in various ways, this implies that their gene products play a role in overlapping, but not identical, circumstances. Some SAGs responded quickly to treatments, possibly indicating a direct response. Others responded more slowly, which may indicate an indirect response via treatment-induced senescence (Louis et al., 1998).

Early bolting in *Angelica sinensis* is related to general metabolism, transcription, signal transduction, and transposable elements (Table 1). It is a complex physiological action.

The remaining primer sets were used in analyzing the various genes expressed in early-bolting and normal *Angelica sinensis* plants. More of the TDFs involved in early bolting regulation need to be cloned, since some could be of possible use in preventing this phenomenon.

### Acknowledgments

We are grateful to the National Natural Science Foundation of China (Project No.30801518), the National Science and Technology Pillar Program, Eleventh Five-Year Plan Period (Project No. 2006BA109B05-01; Project No. 2007BA137B02) and and A Project Funded by the Priority Academic Progam Development of Jiangsu Higher Education Institutions (PAPD) for the support they provided.

### References


Internet Resources


Supplementary Material

The following online material is available for this article:
Table S1 - Sequences of the primers used for AFLP. This material is available as part of the online article from http://www.scielo.br/gmb.

Associate Editor: Dario Grattapaglia