GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

Identification of differentially expressed genes involved in early bolting of *Angelica sinensis* (Apiaceae)

G. Yu[1], Y.X. Ma[2], J.-A. Duan[1], B.S. Song[3] and Z.Q. He[3]

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

Corresponding author: J.-A. Duan
E-mail: dja@njutcm.edu.cn

Genet. Mol. Res. 11 (1): 494-502 (2012)
Received August 2, 2011
Accepted January 23, 2012
Published March 6, 2012
DOI http://dx.doi.org/10.4238/2012.March.6.2

ABSTRACT. *Angelica sinensis* is a highly valued medicinal herb, known as female ginseng that is widely cultivated in China. Although *A. sinensis* is in great demand due to its multiple medicinal and food applications, its early bolting rate (almost 40%) seriously affects crop quality. To better understand its flowering mechanism, cDNA-amplified RFLP analysis was employed to look for gene expression differences between flower bud and shoot apical meristem tissues. Sixty-four primer sets were used, with each primer set amplified to 60 transcript-derived fragments. Some transcript-derived fragments were expressed only in the flower bud. After cloning, sequencing and a homology search, 46 distinct sequences were obtained; 26 of these were found to have homologous sequences in databases. These included trans-caffeoyl-CoA 3-O-methyltransferase, 1-deoxy-D-xylulose 5-phosphate reductoisomerase, 15-cis-zeta-carotene isomerase, isoamylase, and calmodulin-binding protein. These genes are closely related to pollen germination and pollen tube growth, terpenoid backbone biosynthesis, and other metabolic pathways. Confirmation of differential expression

of 10 sequences was obtained by semi-quantitative RT-PCR, showing higher expression levels in flower buds.

Key words: cDNA-AFLP; Early bolting; *Angelica sinensi*s Diels; TDFs; sqRT-PCR; Female ginseng

INTRODUCTION

Angelica sinensis (Oliv.) Diels (Umbelliferae) is a world-renowned medicinal herb found in China (mainly in Gansu Province), Japan, and Korea. The radix of *A. sinensis*, called Danggui, is a popular tonic recorded as early as AD 25-225 in the book Shen Nong Ben Cao Jing during the Han Dynasty. It is recommended for the treatment of amenorrhea and dysmenorrhea, as an anti-inflammatory agent in menstrual disorders, and as a hematopoietic and tonic, among other uses (Nanjing University of Chinese Medicine, 2006; Wang et al., 2007). It is also a dietary supplement in Europe and America and a health food product for women in Asia (Yang et al., 2008). Because of this wide variety of food and medicinal applications, demand for *A. sinensis* is increasing.

In Gansu Province where the herb originated, however, its quality has been seriously affected by its high early bolting rate, restricting its further development. Flowering affects the accumulation of medicinal compounds in *A. sinensis*, reducing its effectiveness as materia medica. Scientists have, therefore, begun to focus on the ecological factors involved in early bolting as well as the nutritional status of the plant itself (Lin and Qiu, 2007; Wang et al., 2007) rather than on specific molecular mechanisms. Currently, no effective methods for the control and prevention of this characteristic exist, although the intense demand for the plant presses for an urgent solution.

Through studies of *Arabidopsis*, rice, and other model plants, researchers have found that bolting is affected by gene expression and environmental factors. Four pathways affect flowering time: the photoperiod, vernalization, autonomous, and gibberellin pathways (Yanovsky and Kay, 2003; Komeda, 2004). These pathways have not been yet studied in *A. sinensis*. Complementary DNA-amplified fragment length polymorphism (cDNA-AFLP) analysis with high reproducibility can be used to screen a large number of differentially expressed cDNAs systematically (Money et al., 1996; Habu et al., 1997; Milioni et al., 2002). In the present study, cDNA-AFLP analysis was carried out to identify the differentially expressed genes in early bolting and normal *A. sinensis*. Several A number of candidate genes that might be involved in early bolting were identified.

MATERIAL AND METHODS

Plant materials

A. sinensis, a triennial medicinal plant, is characterized by early bolting, which frequently occurs in May. Thus, since 2004, and each year after this period, flower buds and sprout-shoot apical meristems were gathered from plants in Gansu Province.

RNA preparation and cDNA synthesis

Flower buds of early bolting plants, each approximately 0.5 cm, were harvested before full blooming together with sprout-shoot apical meristems from normal plants. The samples

were frozen in liquid nitrogen and stored at -80°C. Total RNA was extracted from the plant materials using the Trizol extraction method (Invitrogen, USA). RNA integrity and quantity were determined by running 2 μL total RNA in a formamide denaturing gel. First- and second-strand cDNA synthesis was obtained using a Universal Riboclone cDNA Synthesis System (Promega, USA), according to manufacturer instructions. The resulting double-stranded cDNA was separated on agarose gel to check the sizes of the cDNA samples.

Preparation of the samples for cDNA-AFLP analysis

cDNA-AFLP was carried out according to the procedure of Habu et al. (1997) with minor modifications. Approximately 500 ng double-stranded cDNA was subjected to standard AFLP template production. The restriction enzymes used for the digestion of cDNA were *Mse*I (10 U/μL; Sangon, China) and *Eco*RI (10 U/μL; Sangon). The digested products were ligated to adapters with the following sequences: *Mse*I adapter, 5'-GACGATGAGTCCTGAG-3' and 5'-TACTC AGGACTCAT-3'; *Eco*RI adapter, 5'-CTCGTAGACTGCGTACC-3' and 5'-AATTGGTACG CAGTC-3'. The ligated products were preamplified with the corresponding preamplification primers (*Mse*I: 5'-GATGAGTCCTGAGTAA-3'; *Eco*RI: 5'-GACTGCGTACCAATTC-3'). Twenty cycles of preamplification were performed in a MyCycler PCR system (Bio-Rad, USA) as follows: 30 s at 94°C, 30 s at 56°C, and 1 min at 72°C. The final template was diluted 50-fold with sterile water in preparation for selective amplification.

cDNA-AFLP reactions and polyacrylamide gel electrophoresis of the selectively amplified products

Equal amounts of preamplified products were amplified using primers with selective nucleotides. The primers used for selective amplification were as follows: *Eco*RI+NN primer, 5'-GACTGCGTACCAATTCNN-3' and *Mse*I+NN primer, 5'-GATGAGTCCTGAGTAANN-3', in which N could correspond to A, C, G, or T. The first cycle consisted of 30 s at 94°C, 30 s at 65°C, and 1 min at 72°C. For the next 13 cycles, the annealing temperature was reduced by 0.7°C per cycle. The final 23 cycles consisted of 30 s at 94°C, 30 s at 56°C, and 1 min at 72°C. Four sets of the *Eco*RI primer and 16 sets of the *Mse*I primer were combined to give a total of 64 primer pair combinations. For high-throughput analysis of differentially expressed transcript-derived fragments (TDFs), the polymerase chain reaction (PCR) products were separated on denaturing 6% polyacrylamide gels. The differentially amplified bands were cut directly from the gels on the basis of the scanned image.

TDF isolation and reamplification

The polymorphic TDFs identified based on presence and expressed at higher levels in the flower buds were cut from the denaturing 6% polyacrylamide gel. They were initially eluted in 50 μL sterile double-distilled water at 95°C for 15 min, and then hydrated overnight at 4°C.

Two microliters of the aliquot was used for reamplification to a total volume of 20 μL with the same set of corresponding selective primers and PCR conditions as those used for selective amplification except that an annealing temperature of 56°C, 35 cycles, and a final 5-min extension were used. The PCR products were resolved on a 1.5% 1X Tris-acetate ethylenediaminetetraacetic acid-agarose gel, and each single band was isolated and eluted using a Takara gel extraction kit (Takara, Japan). The eluted TDFs were cloned into a plasmid PMD-

18T vector (Sangon) following the manufacturer protocol.

Sequence analysis

Five colonies were selected for the sequencing analysis of each TDF. Sequencing was carried out using an automated sequencer. The sequences of the TDFs were analyzed for their homology against nonredundant genes, expressed sequence tags (ESTs), or transcripts in public databases (http://www.ncbi.nlm.nih.gov/BLAST, http://www.arabidopsis.org/cgi-bin/fasta/nph-TAIRfasta.pl, and http://www.shigen.nig.ac.jp/rice/oryzabase/blast) using the BLASTN and BLASTX algorithms (Altschul et al., 1997). The homologous sequences were further analyzed in databases (http://www.genome.jp/kegg/ and http://www.geneontology.org/) to understand their function more fully.

Semi-quantitative reverse transcription PCR (sqRT-PCR)

sqRT-PCR was carried out on 2 μg total RNA using a two-step RT-PCR kit (Takara, Japan) according to manufacturer instructions. Primers were designed from the sequences of the TDFs using a Primer 3.0 web resource (http://frodo.wi.mit.edu/). Total RNA was used as the internal standard in the RT-PCR for checking the quality of the RNA template. PCR conditions were essentially the same as those described earlier for the reamplification of the TDFs. To obtain reproducibility of results, we repeated the sqRT-PCR analysis three times.

RESULTS

cDNA-AFLP fingerprinting

A total of 64 primer combinations (Table 1) were selected for cDNA-AFLP. On average, 60 clear and unambiguous bands (TDFs) were generated with each primer combination. Although most combinations of primers produced identical patterns, some TDFs were expressed at a higher level in the flower bud (Figure 1). This difference implied that minor changes in gene expression could account for the phenotypes of early bolting observed in *A. sinensis*. These differentially expressed TDFs were successfully recovered from the denaturing polyacrylamide gels for further analysis.

Table 1. Sequences of the primers used for amplified restriction fragment length polymorphism.

*Mse*I+NN	*Eco*RI+NN
GATGAGTCCTGAGTAAAA	GACTGCGTA CCAATTCCA
GATGAGTCCTGAGTAAAT	GACTGCGTA CCAATTCCT
GATGAGTCCTGAGTAAAC	GACTGCGTA CCAATTCCC
GATGAGTCCTGAGTAAAG	GACTGCGTA CCAATTCCG
GATGAGTCCTGAGTAATA	
GATGAGTCCTGAGTAATC	
GATGAGTCCTGAGTAATG	
GATGAGTCCTGAGTAATT	
GATGAGTCCTGAGTAACA	
GATGAGTCCTGAGTAACT	
GATGAGTCCTGAGTAACC	
GATGAGTCCTGAGTAACG	
GATGAGTCCTGAGTAAGA	
GATGAGTCCTGAGTAAGT	
GATGAGTCCTGAGTAAGC	
GATGAGTCCTGAGTAAGG	

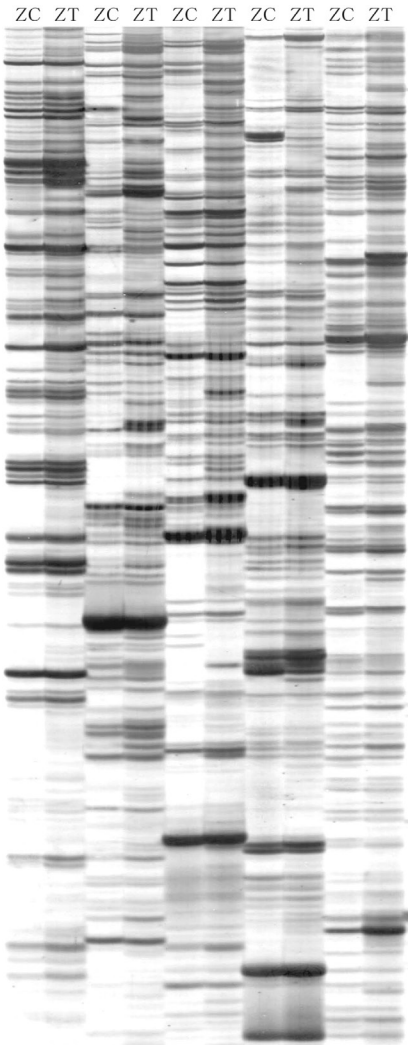

Figure 1. cDNA-AFLP autoradiogram of five of the most information primer combinations amplifying differentially expressed genes in flower bud and shoot apical meristem. The combinations of primers used were included according to the codes reported in Table 1. ZT = flower bud; ZC = shoot apical meristem.

Identification of differentially expressed cDNA clones

The recovered TDFs were reamplified with the original primer sets used for cDNA-AFLP analysis and examined on 1.5% agarose gels. The expected sizes ranged from 151 to 695 bp. After the isolated TDFs were cloned, five colonies were selected and sequenced for each TDF, generating 42 distinct nucleotide sequences. The sequences were then subjected to homology searches using BLAST against the databases, whereby 26 sequences with lower e-value were obtained (Table 2). In addition, some sequences did not result in any hits, possibly as a result of poor knowledge of the *A. sinensis* genome.

Table 2. Nucleotide homology of the transcript-derived fragments with known gene sequences in the database using BLASTN algorithm along with their expression patterns.

AFLP fragment	dbEST_Id	GenBank_ Accession No.	Homology	E-value	Length
A203-1	72420447	HS513742	Trans-caffeoyl-CoA 3-O-methyltransferase	8e-11	537
A204	72420448	HS513743	Actin-related protein ARP4	0.0035	403
A208-1	72420450	HS513745	Retrotransposon V14	1e-31	696
A210	72420452	HS513747	1-Deoxy-D-xylulose 5-phosphate reductoisomerase	7e-04	625
A211-1	72420453	HS513748	Transposable element gene; pseudogene, hypothetical protein	0.89	502
A214-1	72420455	HS513750	Isoamylase	1e-40	695
A303	72420457	HS513752	Transposable element gene	7.8e-13	359
A308-1	72420458	HS513753	Transposable element gene	4.3e-05	240
A310	72420459	HS513754	15-Cis-zeta-carotene isomerase	0.00065	422
A316-1	72420461	HS513756	*Daucus carota* subsp *sativus* clone BAC C073D12 genomic sequence	5e-11	262
A327-1	72420462	HS513757	With NAD kinase activity, the protein was also shown to bind calmodulin	2.9e-05	204
A348-1	72420463	HS513758	Transposable element gene	7e-08	151
A365-1	72420464	HS513759	Transposable element gene	2.9e-25	354
A314-1	72420466	HS513761	Galactose oxidase	0.0073	332
A352-1	72420467	HS513762	A chloroplast trans-acting factor of the psbD light-responsive promoter	0.41	288
A364-1	72420468	HS513763	*Solanum lycopersicum* strain Heinz 1706 chromosome 1 clone hba-168g4 map 1	1e-16	384
A362-1	72420469	HS513764	Transposable element gene	2.7e-24	413
A349-1	72420471	HS513766	*Daucus carota* subsp *sativus* clone BAC C061I05 genomic sequence	8e-18	335
A311-1	72420472	HS513767	Transposable element gene	3.2e-29	384
A353-1	72420473	HS513768	*Daucus carota* subsp *sativus* clone BAC C155P02 genomic sequence	1e-23	236
A309-1	72420474	HS513769	Transposable element gene	4.5e-05	442
A325-1	72420475	HS513770	Transposable element gene	3.1e-15	330
A351-1	72420476	HS513771	Calmodulin-binding protein (CAM), reverse transcriptase	2e-13	310
A241-1	72420479	HS513774	Transposable element gene	0.12	295
A242-1	72420480	HS513775	RNA-directed DNA polymerase gene	5e-06	269
238-1	72420481	HS513776	*P. trichocarpa* multidrug/pheromone exporter	3e-31	343

Confirmation of differentially expressed fragments using sqRT-PCR

Among the 26 sequences with lower e-values, 10 sequences were found with clear functions and a close relationship to bolting. Their expressions were analyzed using sqRT-PCR with the primer pairs listed in Table 3. The results showed that they were expressed at a higher level in early bolting *A. sinensis* (Figure 2).

Table 3. Sequences of the primers used for semi-quantitative reverse transcription polymerase chain reaction.

TDF clone	Forward primer (5'-3')	Reverse primer (5'-3')
A203-1	GCTCGAAGTCCGAATGAATC	ACTGCCCATGCGAAGATAAC
A210	ACGCCTTGATCTTTGCAAGT	TGTAACTGGTTTCCCCCTTG
A214-1	AAATCCAGCCAACCTCCTCT	CCAATCAAGCTGCACTGTGT
A310	GCACTTTTCAAATTGGGTCAA	GCGTGATTGGTATGTGCATC
A314-1	GTTGGCTGTTTTGGAAGCTG	ACAACCAACATGTGCACGAC
A361-1	CCCCTCCCCCTCTGTATAGT	TGGATTTGTGGCTTTGATGA
A327-1	TTTTGGCGTGATATGTTTTGT	TGGCAAAAATGTCATGAACC
A351-1	CGGATACGACGTTTTCTGAAG	ATTTTGATTCCACGCAGCTT
A242-1	TCCTTGCAAAACGAGCTTCT	TATTCAACCACCCCTTCTCC
A238-1	TGGCATCTGTGGGAATTACA	AGGTCCCGGAAGATCTGAAC

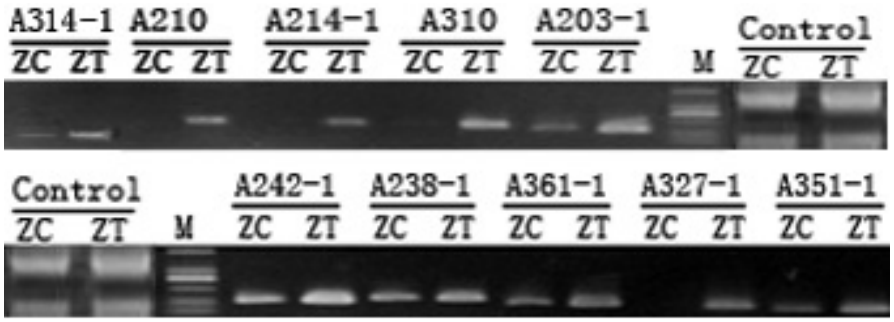

Figure 2. Confirmation by semi-quantitative RT-PCR of the higher level expression of TDFs in the flower bud of early-bolting *A. sinensis* (Oliv.) Diels. ZT = flower bud; ZC = shoot apical meristem.

DISCUSSION

cDNA-AFLP, a variation of AFLP derived from RNA fingerprint identification technology, has already become a sophisticated research tool for identifying gene expression differences. In this study, cDNA-AFLP was used to compare the gene expression profiles of flower buds and shoot apical meristems. Twenty-six sequences were obtained, and some might be involved in the control of early bolting. We focused on sequences that were identified to be trans-caffeoyl-coenzyme A (CoA) 3-O-methyltransferase, 1-deoxy-D-xylulose 5-phosphate reductoisomerase, 15-cis-zeta-carotene isomerase, isoamylase, nicotinamide adenine dinucleotide (NAD) kinase, galactose oxidase, calmodulin-binding protein (CaM), and transposable element gene. In all cases, these genes were involved in cellular pathways leading to bolting.

TDF A210, shown in Table 2, is homologous to 1-deoxy-D-xylulose 5-phosphate reductoisomerase, which is related to terpenoid backbone biosynthesis. 1-Deoxy-D-xylulose 5-phosphate reductoisomerase catalyzes the biosynthesis of isoprene, zeatin, monoterpenoid, diterpenoid, carotenoid, porphyrin, chlorophyll, N-glycan, steroids, sesquiterpenoid ubiquinone, and other terpenoid quinones. The hormone zeatin affects plant bolting; zeatin is markedly more effective than kinetin and 6-(3-methylbut-2-enyl) aminopurine in promoting frond expansion and increasing frond number of *Spirodela oligorrhiza* cultures grown under continuous illumination (Letham, 1967). The hormone gibberellin is a derivative of diterpenoid that plays a central role in modulating diverse processes throughout plant development. Gibberellins are hormones that control growth and a wide variety of plant developmental processes. The gibberellin pathway is one of the four key pathways that affect flowering (Sun and Gubler, 2004).

A242-1 is homologous to calmodulin-binding protein, which is related to CaM. CaM is a ubiquitous multifunctional calcium sensor in all eukaryotes that mediates calcium action by regulating the activity/function of many unrelated proteins. Calcium and CaM are known to play a crucial role in pollen germination and pollen tube growth (Golovkin and Reddy, 2003).

A238-1 is homologous to the ATP binding cassette (ABC) superfamily, which is a large, ubiquitous, and diverse group of proteins, most of which mediate transport across biological membranes. Sequences of more than 20 plant ABC proteins have been published, and include homologs of P-glycoprotein, multidrug resistance-associated protein and organellar transporters. Recent results indicate that the function of this protein family is not restricted

to detoxification processes. Plant ABC transporters have been demonstrated to participate in chlorophyll biosynthesis, formation of Fe/S clusters, stomatal movement, and probably ion fluxes; hence, they may play a central role in plant growth and development (Sidler et al., 1998; Møller et al., 2001).

TDF A203-1 is matched with trans-caffeoyl-CoA 3-O-methyltransferase, which is related to metabolic pathways, biosynthesis of secondary metabolites such as flavonoid phenylpropanoid, and phenylalanine metabolism. Homoeriodictyol, capsaicin, syringin, coniferin, guaiacyl lignin, and syringyl lignin are all related to trans-caffeoyl-CoA 3-O-methyltransferase. TDF A310 is homologous to 15-cis-zeta-carotene isomerase, which is related to the metabolic pathways and biosynthesis of secondary metabolites such as carotenoids. TDF A327-1 is homologous to NAD kinase 2, which encodes a protein with NAD kinase activity. A314-1 is homologous to galactose oxidase.

As summarized in Table 2, the early bolting of *A. sinensis* might be related to pollen germination and pollen tube growth, terpenoid backbone biosynthesis, metabolic pathways, biosynthesis of secondary metabolites, and transposable elements. The differentially expressed genes obtained in this study should help to elucidate the molecular basis of early bolting in *A. sinensis* and identify genes that could be targeted to prevent early bolting and increase *A. sinensis* production.

ACKNOWLEDGMENTS

We are grateful to the National Natural Science Foundation of China (Project #30801518), the National Science and Technology Pillar Program, Eleventh Five-Year Plan Period (Projects #2006BA109B05-01 and #2007BA137B02), and a Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions for support provided for this research.

REFERENCES

Altschul SF, Madden TL, Schaffer AA, Zhang J, et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25: 3389-3402.

Golovkin M and Reddy AS (2003). A calmodulin-binding protein from *Arabidopsis* has an essential role in pollen germination. *Proc. Natl. Acad. Sci. U. S. A.* 100: 10558-10563.

Habu Y, Fukada-Tanaka S, Hisatomi Y and Iida S (1997). Amplified restriction fragment length polymorphism-based mRNA fingerprinting using a single restriction enzyme that recognizes a 4-bp sequence. *Biochem. Biophys. Res. Commun.* 234: 516-521.

Komeda Y (2004). Genetic regulation of time to flower in *Arabidopsis thaliana*. *Annu. Rev. Plant Biol.* 55: 521-535.

Letham DS (1967). Regulators of cell division in plant tissues. *Planta* 74: 228-242.

Lin MH and Qiu YD (2007). Effect of root diameter on early bolting rate and yield in seedling of *Angelica sinensis* (Oliv.) Diels. *Chin. Traditional Herbal Drugs* 38: 1386-1389.

Milioni D, Sado PE, Stacey NJ, Roberts K, et al. (2002). Early gene expression associated with the commitment and differentiation of a plant tracheary element is revealed by cDNA-amplified fragment length polymorphism analysis. *Plant Cell* 14: 2813-2824.

Møller SG, Kunkel T and Chua NH (2001). A plastidic ABC protein involved in intercompartmental communication of light signaling. *Genes Dev.* 15: 90-103.

Money T, Reader S, Qu LJ, Dunford RP, et al. (1996). AFLP-based mRNA fingerprinting. *Nucleic Acids Res.* 24: 2616-2617.

Nanjing University of Chinese Medicine (2006). Dictionary of Chinese Herbal Medicines. 2nd edn. Shanghai Science and Technology Press, 1207-1213.

Sidler M, Hassa P, Hasan S, Ringli C, et al. (1998). Involvement of an ABC transporter in a developmental pathway regulating hypocotyl cell elongation in the light. *Plant Cell* 10: 1623-1636.

Sun TP and Gubler F (2004). Molecular mechanism of gibberellin signaling in plants. *Annu. Rev. Plant Biol.* 55: 197-223.

Wang YL, Liang YZ and Chen BM (2007). High-performance liquid chromatography with atmospheric pressure chemical ionization and electrospray ionization mass spectrometry for analysis of *Angelica sinensis*. *Phytochem. Anal.* 18: 265-274.

Yang XB, Zhao Y and Lv Y (2008). *In vivo* macrophage activation and physicochemical property of the different polysaccharide fractions purified from *Angelica sinensis* (Oliv.) Diels. *Carbohydr. Polym.* 71: 372-379.

Yanovsky MJ and Kay SA (2003). Living by the calendar: how plants know when to flower. *Nat. Rev. Mol. Cell Biol.* 4: 265-275.