Plant Mol Biol Rep (2010) 28:193–203
DOI 10.1007/s11105-009-0140-1

Cloning and Characterization of *Prunus serotina AGAMOUS*, a Putative Flower Homeotic Gene

Xiaomei Liu · Joseph M. Anderson · Paula M. Pijut

Published online: 10 September 2009

Abstract Members of the *AGAMOUS* subfamily of MADS-box transcription factors play an important role in regulating the development of reproductive organs in flowering plants. To help understand the mechanism of floral development in black cherry (*Prunus serotina*), *PsAG* (a putative flower homeotic identity gene) was isolated, and its MIKC-type structure was shown to be a homolog of the *Arabidopsis thaliana AG* gene. It was a single-copy gene in black cherry. A phylogenetic tree derived from the protein sequence indicated *PsAG* to be a C-function flower homeotic gene with a high similarity to other *AG* homologs, such as those from *Prunus persica* and *Prunus mume*. *PsAG* met the criteria for *AG* subfamily gene structure with a typical MIKC structure. *In situ* hybridization showed that *PsAG* was expressed mainly in the floral meristem, such as stamen and carpel primordia during the early stage of floral development, and transcript of *PsAG* accumulated in the tissues of the ovary, stigma, style, and stamens. When the flowers matured, *PsAG* had enhanced expression in ovary, style, and stigma, with decreased expression in the stamen. *PsAG* continued to be expressed in the ovule at the late stage of flower development. The developmental patterns of expression were consistent with those of *AG* and homologs from other species. Both phylogenetic analysis and expression-pattern data suggest that *PsAG* was the black cherry homolog of *Arabidopsis AG*. An RNAi construct with a partial *PsAG* gene was constructed for black cherry transformation.

Keywords Black cherry · Carpel · Flowering · MADS-box · RNAi · Stamen

X. Liu
Dept. of Forestry and Natural Resources,
Hardwood Tree Improvement and Regeneration Center (HTIRC),
Purdue University,
715 West State Street,
West Lafayette, IN 47907, USA

J. M. Anderson
Department of Agronomy, USDA Agricultural Research Service,
Purdue University,
915 West State Street,
West Lafayette, IN 47907, USA

P. M. Pijut (✉)
Northern Research Station, HTIRC, USDA Forest Service,
715 West State Street,
West Lafayette, IN 47907, USA
e-mail: ppijut@purdue.edu
e-mail: ppijut@fs.fed.us

Introduction

Black cherry (*Prunus serotina* Ehrh.) is a valuable hardwood in the eastern USA and Canada. Demand for high-quality black cherry wood is increasing, and there is a need to establish plantations with improved black cherry trees. Genetically improved trees containing foreign genes are subject to government regulatory guidelines for field planting because of the potential for dispersal of transgenic pollen, and the environmental impact could be difficult to predict and control (Meilan et al. 2001). Most scientists agree that transgenic crops pose little risk of ecological impact since crops are highly domesticated, and the ability to hybridize with wild relatives is very low. For transgenic trees, such as black cherry, genetic containment is desirable because commercial clones have undergone little domestication, and several characteristics of black cherry make extensive, long-distance gene flow likely. Black cherry produces abundant pollen and seeds, and long-distance movement of pollen is promoted by wind dispersal

combined with tree height. To reduce the dispersion of all genes, engineering reproductive sterility will help simplify the impact and thus facilitate regulatory and public approval (Strauss et al. 1995). This will allow landowners to plant transgenic trees without concern for the effect on the ecosystem. However, information on the basics of black cherry flower and fruit development was generally lacking (Stairs and Hauck 1968).

Studies in *Arabidopsis* using floral homeotic mutants led to a simple combinatorial model (Coen and Meyerowitz 1991) that proposes three classes of genes namely A, B, and C to be expressed in adjacent, overlapping whorls of a flower. With the isolation of new MADS-box genes specifying ovule development from *Petunia hybrida*, the ABC model was extended to the ABCDE model, which includes D- and E-function genes (Honma and Goto 2001; Theissen and Saedler 2001). All of the genes involved in the model encode type II MADS-box proteins except for two *APETALA2 (AP2)* genes, which belong to another group of transcription factors. Among the floral MADS-box genes, *AGAMOUS* (*AG*) was the only C-function gene, first isolated from the model plants *Arabidopsis* (Yanofsky et al. 1990) and *Antirrhinum majus* (Coen and Meyerowitz 1991).

In *Arabidopsis*, *AG* was the only C-function gene that has been identified and extensively elucidated, while in other species such as cucumber (*Cucumis sativus*), *Petunia*, or *Antirrhinum*, multiple *AG* homologs control flower development by regulatory interactions among these homologs (Kater et al. 1998; Davies et al. 1999). In woody plants, the functional homology of the *AG* gene has been identified from a diverse number of species, such as palm (Zhang et al. 2004), spruce (Rutledge et al. 1998; Tandre et al. 1998), hazelnut (Rigola et al. 1998), birch (Lemmetyinen et al. 2004), grape (Boss et al. 2001), poplar (Brunner et al. 2000), gingko (Jager et al. 2003), rose (Kitahara and Matsumoto 2000), and apple (van der Linden et al. 2002). *In situ* hybridization and ectopic expression has been widely used to assess expression pattern of *AG* homologs. *AG* and its homologs are mainly expressed within the floral meristem, immature stamens, and immature carpel tissues (Brunner et al. 2000; van der Linden et al. 2002; Jager et al. 2003; Martin et al. 2006).

Among the floral MADS-box genes, *AG* functions sequentially in the establishment of the determinate floral meristem and later in the development of the stamens and carpels in the third and fourth whorls (Sieburth et al. 1995; Martin et al. 2006). There are several criteria for characterization of *AG* homologs. First, the protein sequence has a typical MIKC-type MADS protein structure, as with other *AG* homologs, a highly conserved MADS domain, and moderately conserved K domain (Fan et al. 1997). Second, intron 8 interrupts the codon of the last amino acid which was an ancient primary characteristic of all *AG*-like genes (Kramer et al. 2004). Third, *AG* spatial expression was restricted to two inner (the third and fourth) whorls, where the stamen and carpel are normally located (Bowman et al. 1991a, b). In *Prunus*, the *AG* gene has been cloned from peach (*Prunus persica*; Martin et al. 2006) by screening cDNA libraries, and full cDNAs of *AG* of *P. persica*, *Prunus mume*, and *Prunus dulcis* have also been obtained (GenBank).

Degenerate primers have been widely used for cloning *AG* from different tree species, such as apple (*Malus domestica*; Sung et al. 2000), sweetgum (*Liquidambar styraciflua*; Liu et al. 1999), ginkgo (*Ginkgo biloba*; Jager et al. 2003), and fern (*Ceratopteris richardii*; Hasebe et al. 1998). In apple, approximately 15 MADS-box genes have been cloned, and several of these have been characterized (Sung et al. 1999; Yao et al. 1999; Kotoda et al. 2000, 2002; van der Linden et al. 2002; Wada et al. 2002). Degenerate primers are mixtures of similar, but not identical primers. The primer design was based upon protein sequence because several different codons may code for one amino acid primer sequence corresponding to the variable bases. Using degenerate primers generally reduces the specificity of the polymerase chain reaction (PCR) amplification. The problem can be partially solved by nested PCR to improve specificity. Degenerate primers are designed by aligning gene sequences found in GenBank and then synthesizing a mixture of primers corresponding to all permutations. *AG* was a highly conserved gene in many species. It is possible to use several pairs of degenerate primers to clone MADS-box genes without screening a cDNA library, which was a traditional method to clone genes. The objective of the present research was to isolate the black cherry *AG* gene and characterize its expression pattern. We describe the methods to clone a full-length cDNA of *PsAG* by using degenerate primers and rapid amplification of cDNA ends (RACE) and characterized the *PsAG* expression pattern in black cherry flowers.

Materials and Methods

Plant Materials

Flowers of *P. serotina* from 13- to 14-year-old trees were collected on April 25, 2007, May 2, 2007, April 24, 2008, and May 5, 2008 near West Lafayette, Indiana. Tissues were immediately frozen in liquid nitrogen and stored at −80°C until used or fixed for *in situ* hybridization.

RNA Extraction, cDNA Synthesis, and Cloning

Total RNA was extracted from flower tissue according to the protocol of Salzman et al. (1999). The isolation of DNA

from black cherry leaves was performed using the DNeasy Plant Mini Kit (Qiagen, Inc., Valencia, CA). A reverse transcriptase PCR (RT-PCR)-based strategy was used for cloning. To design the degenerate primers, a total of 12 published *AG* homologous sequences were retrieved from GenBank and aligned with the Clustal W program. From these alignments, several conserved regions were identified and examined further using the OLIGO primer analysis. Three pairs of degenerate primers (Table 1: P1F, P2F, P3F, P1R, P2R, and P3R) were designed to amplify an internal fragment spanning part of the MADS-box and part of the K box.

To amplify the *PsAG* fragment, 2 μg total RNA was used for first-strand cDNA synthesis by using the ProtoScript First Strand cDNA Synthesis Kit (New England BioLabs, Inc., Ipswich, MA). The first-strand cDNA product was diluted to 50 μl, and 10 μl of the cDNA synthesis product was used for primary RT-PCR. After the internal sequence of *PsAG* was obtained, gene-specific primers (5′ RACE outer primer, 5′ RACE inner primer, 3′ RACE outer primer, 3′ RACE inner primer) were designed to clone the cDNA ends (Table 1). Primers were designed to give PCR products with a 250 to 350 bp overlap in order to provide convenience in subsequent confirmation and other manipulations. Total RNA isolated from immature flowers (tissue collected on April 24, 2008) was used for cloning.

For amplification of MADS-box sequences, oligo-dT and a degenerate primer, P2F, 5′-TN AAR AAR GCN TAY GAR YT-3′, was chosen for primary PCR. A 50 μl PCR reaction was prepared containing 5 μl 10× PCR buffer (Invitrogen, Carlsbad, CA), 2 mM $MgCl_2$, 0.2 mM each of dNTP, 0.5 mM each of primer (P2F and oligo-dT), and 1 U of Taq DNA polymerase (Invitrogen). The cycling program consisted of an initial denaturation at 94°C for 4 min, 20 cycles of 94°C for 1 min, 48°C for 40 s, 72°C for 1 min, and 25 additional cycles with fresh Taq polymerase 94°C, 50°C, 72°C, 1 min each, followed by a final extension at 72°C for 7 min. The PCR products (2 μl) were used to provide templates for nested PCR with primers P3F (5′-TCN GTN YTN TGY GAY GCN GA-3′) and P2R (5′-TT YTG CAT RTA YTC DAT YTC NGC RA-3′) corresponding to the conserved amino acid sequences SVLCDAE and FAEIEYMQK, respectively. The thermocycler program was: 3 min at 94°C, 40 cycles of 45 s at 94°C, 40 s at 55°C, 1 min at 72°C, and a final extension step of 7 min at 72°C. PCR product was separated on 1.5% agarose gel. Half-nested PCR was performed with primer (P3F, P1R) using the same thermocycler program as described previously. Several products from P3F–P2R or P3F–P1R with the expected size between 300 and 400 bp were purified with the QIAquick Gel Extraction Kit (Qiagen, Inc.) and then cloned into the pGEM-T Easy Vector (Promega, Corp., Madison, WI).

Transformation of *Escherichia coli* by Electroporation

E. coli DH5α competent cells were made according to the protocol of Dower et al. (1988). The ligation product was used for electroporation. *E. coli* cells were cultured on Luria–Bertani media with appropriate antibiotics (for pGEM-T Easy; 50 mg l^{-1} ampicillin). Colony PCR was conducted by using T7 and SP6 primers. Plasmid DNA was isolated with the QIAprep Spin Miniprep Kit (Qiagen, Inc.), and insert was confirmed by PCR with T7 and SP6 primers. Twenty clones were sequenced at the Purdue University Genomic Center (West Lafayette, IN) using the BigDye terminator sequencing with T7 and SP6 primers. The 3′ RACE and 5′ RACE were conducted following manufacturer's procedure using FirstChoice RLM-RACE Kit (Applied Biosystems, Inc., Foster City, CA). The annealing temperature was 60°C. PCR products were purified and sequenced as described previously.

Table 1 Degenerate primers and RACE primers used to amplify the *PsAG* gene

Name	Primer sequence 5′–3′	Corresponding amino acid sequence	T_m
P1F	GN ATH GAR AAY ACN ACN AA	RIENTTN	46.7
P2F	TN AAR AAR GCN TAY GAR YT	LKKAYEL	47.7
P3F	TCN GTN YTN TGY GAY GCN GA	SVLCDAE	57.7
P3R	TC RTT YTC NGC DAT YTT NG	RAKIAENE	49.9
P2R	TT YTG CAT RTA YTC DAT YTC NGC RA	FAEIEYMQK	55.5
P1R	AR YTC RTT YTT YTT NGA NCK DA	IRSKKNEL	50.8
5′ RACE outer R	CAGGAGCTCATTCTTCTTGGA	SKKNELL	54.3
5′ RACE inner R	GTTCTTCAGGTCCTTCATCTTCA	MKDLKN	54.5
3′ RACE outer F	GCCGAGGTTGCTCTCATAGT	AEVALIV	54
3′ RACE inner F	TTGAGAGGTACAAGAAGGCAT	YKKACAE	53.1
SP6	CATACGATTTAGGTGACACTATAG		52
T7	TAATACGACTCACTATAGGG		48

Determination of *PsAG* Intron Positions and Sizes

Five sets of primers (I1F/I1R, I2F/I2R, I3F/I3R, I4567F/I4567R, and I8F/I8R) (Table 2) were designed to clone eight introns in *PsAG* using the corresponding sequence of the *AG* gene of *Arabidopsis thaliana* (Yanofsky et al. 1990) and *PtAG*1 in poplar (Brunner et al. 2000). Primers were designed from the coding sequence, each primer set bracketed a putative intron site, and the genomic structure of the *PtAG*1 gene of poplar was used as a model for looking for introns in *PsAG*. Genomic DNA was used as the template for PCR amplification. PCR conditions were as described previously, but annealing temperature was adjusted according to each primer pair T_m value. The PCR products were gel purified, quantified, and sequenced as described previously.

Phylogenetic Analysis and Amino Acid Alignment

The *AG* homologs were identified by Blastp searches at NCBI (http://www.ncbi.nlm.nih.gov/blast). Translation of nucleotide sequence to protein and alignments were conducted using the ExPASy translate tool (http://ca.expasy.org/tools/dna.html). Alignment of conceptual amino acid sequences was conducted using the BioEdit program. A phylogenetic tree was constructed using the neighbor-joining methods in PAUP 4b. BETA 10 (Swofford 2003); amino acids from 0 to 49 and from 234 to 296 were excluded because of poor alignment. The tree was bootstrapped with 1,000 interactions (node cutoff value of 50%), and *A. thaliana* sequence was used as an outgroup to root the tree.

Southern Blot Analysis

Genomic DNA from black cherry was isolated, and 10 μg genomic DNA was digested with four restriction enzymes (*Bam* HI, *Xba* I, *Xho* I, and *Kpn* I) at 37°C in a water bath overnight; the resulting genomic DNA fragments were separated on 0.8% agarose gels. After electrophoresis, the agarose gel was soaked in 0.25 M HCl for 10 min and denaturing solution (1.5 M NaCl, 0.5 M NaOH) for 30 min; then the DNA fragments from the gel were transferred to a nylon membrane by the alkaline transfer method. The transferred DNA was immobilized by UV irradiation (Stratagene UV crosslinker, 120 mJ). The templates used as probes were prepared by PCR amplification from *PsAG* cDNA clones with gene-specific primers in nonconserved regions (Table 1). Primer pair 5′-TTGAGAGGTACA AGAAGGCAT-3′ and 3′ RACE inner reverse primer were used to amplify 600 bp downstream of the MADS-domain region. The 20 ng DNA in 45 μl TE buffer was labeled by 5 μl [α-^{32}P] dCTP by the Rediprime II random prime labeling system. Prehybridization was performed for at least 1 h in the hybridization solution (Sigma, St. Louis, MO, Cat # H703) at 65°C with gentle shaking; after adding the denatured probe, the hybridization was carried out overnight under the same conditions. After hybridization, the membranes were washed twice (30 min each, using the same conditions described above) with a solution containing 2× saline sodium citrate (SSC) and 0.1% sodium dodecyl sulfate (SDS), then once by 0.2× SSC and 0.1% SDS at 60°C for 30 min. After washing, the membrane was wrapped in plastic wrap and used to expose to autoradiography film.

Table 2 Primers used to amplify introns

Primer name	Primer sequence 5′–3′
Intron 1F	CTG CGT TTG CTG GCT TTG ATG
Intron 1R	TTG CAG AAG GTG ACT TGA CG
Intron 2F	TCA GTG CTA TGC GAC GCC GA
Intron 2R	AGC TTC GGA AAC AGA TCC AG
Intron 3F	TTG AGA GGT ACA AGA AGG CAT
Intron 3R	CTT CTT GGA TCT GAT TCT GC
Intron 4567F	GTA CTA CCA ACA AGA AGC TGC
Intron 4567R	TGG TGA TTG GGT TGT AAT GC
Intron 8F	GTG GAT GCA TTA CAA CCC AAT C
Intron 8R	CAT TGC TGC AGG CAC ATA TT

Tissue Embedding, *In Situ* Hybridization, and Microscopy

Shoot segments that contained flowers at various developmental stages were collected and soaked in 4% paraformaldehyde fixation solution immediately after removal from trees. Protocols for tissue fixation, embedding, and *in situ* hybridization were according to Jackson (1991) and Shu et al. (1999). Flower materials were fixed for 24 h at 4°C in fixation buffer (4% paraformaldehyde, 0.1 M phosphate buffer, pH 7), dehydrated through a graded series of ethanol and tert-butyl alcohol, embedded in Paraplast® (Monoject, St Louis, MO), and sectioned to 6 μm thick with a microtome. The sections were mounted onto ProbeOn Plus Slides (Fisher Scientific, Pittsburgh, PA) on a 42°C slide warmer. The labeled RNA probe was a fragment that corresponded to the *PsAG* cDNA and lacked the MADS-box, but contained part of the I region and the remaining 3′ terminal sequence, which was also used for Southern blot analyses to test its specificity. A 600 bp *PsAG* cDNA was subcloned into pGEM-T vector, and SP6 and T7 primer were used to amplify a suitable template. The RNA probes were synthesized using the DIG RNA Labeling Kit (Roche Applied Science, Indianapolis, IN), and the T7 and SP6 RNA polymerases were used to synthesize sense RNA and antisense RNA. Paraplast was removed from tissue sections with xylene, rehydrated

through an ethanol series, and then pretreated with proteinase K (10 ngμl^{-1}) in Tris–HCl, pH 8.0, at 37°C for 30 min. Digestion was stopped by washing with 0.2× phosphate-buffered saline (PBS)/glycine and then rinsed twice with PBS for 2 min each. Tissues were acetylated by acetylation buffer (0.1 M triethanolamine, 0.5% acetic anhydride, pH 8.0) for 10 min. After dehydrating in ethanol baths, hybridization was performed at 45°C overnight with 1 ngμl^{-1} of the digoxigenin-labeled RNA probe in hybridization solution. RNA probe was prepared in 50% deionized formamide and mixed with prehybridization buffer as 1:4 mix, (preparation of 1,000 µl prehybridization solution: 500 µl 50% deionized formamide, 125 µl 10× *in situ* salts, 250 µl 50% dextran sulfate, 50 µl 20 mg ml^{-1} tRNA, 10 µl 100× Denhart's, 65 µl ddH_2O; preparation of 10× *in situ* salts: 100 mM Tris–HCl, pH 6.8, 3 M NaCl, 50 mM Na_2HPO_4, 50 mM NaH_2PO_4, 50 mM EDTA in diethylpyrocarbonate-treated ddH_2O) . After hybridization, slides were washed in 2× SSC at 50°C for 30 min and twice in 1× NTE (Tris–HCl 1 mM, NaCl 0.5 M, EDTA 1 mM, pH 8.0) at 37°C for 5 min each. An RNase A digestion (100 µg ml^{-1} in 1× NTE) was carried out for 30 min at 37°C and stopped by washing with 1× NTE at 37°C. Final washes were conducted in 2× SSC, 1× SSC, and 0.5× SSC for 30 min each at 42°C prior to rinsing in PBS. For immunological signal detection, samples were incubated in blocking reagent (Roche; 10% (*w*/*v*) in maleic acid buffer: 0.1 M maleic acid, 0.15 M NaCl, pH 7.5) for 30 min and then for 30 min in blocking reagent containing antidigoxigenin alkaline phosphatase-conjugated Fab fragment antibody (Roche) diluted at 1:5,000. Samples were washed twice for 15 min in washing buffer (0.1 M maleic acid, 0.15 M NaCl, pH 7.5, 0.3% Tween 20). Tissues were equilibrated in detection buffer (100 mM Tris–HCl, pH 9.5, 100 mM NaCl) for 5 min prior to incubating in the same buffer supplemented with 0.2 mM nitroblue tetrazolium and 0.2 mM bromochloroindolyl phosphate substrates (Roche Applied Science) for color visualization. Development times varied between 16 h to 2 days, depending on the signal strength and the pattern of expression. Photomicrography was conducted using an Olympus Vanox-S microscope (Olympus, Tokyo, Japan) equipped with a Sport RT CCD camera (Diagnostic Instruments, Sterling Heights, MI), and images were captured with a Live Spot image analysis system.

Results

Isolation of *PsAG* cDNA by Degenerate PCR

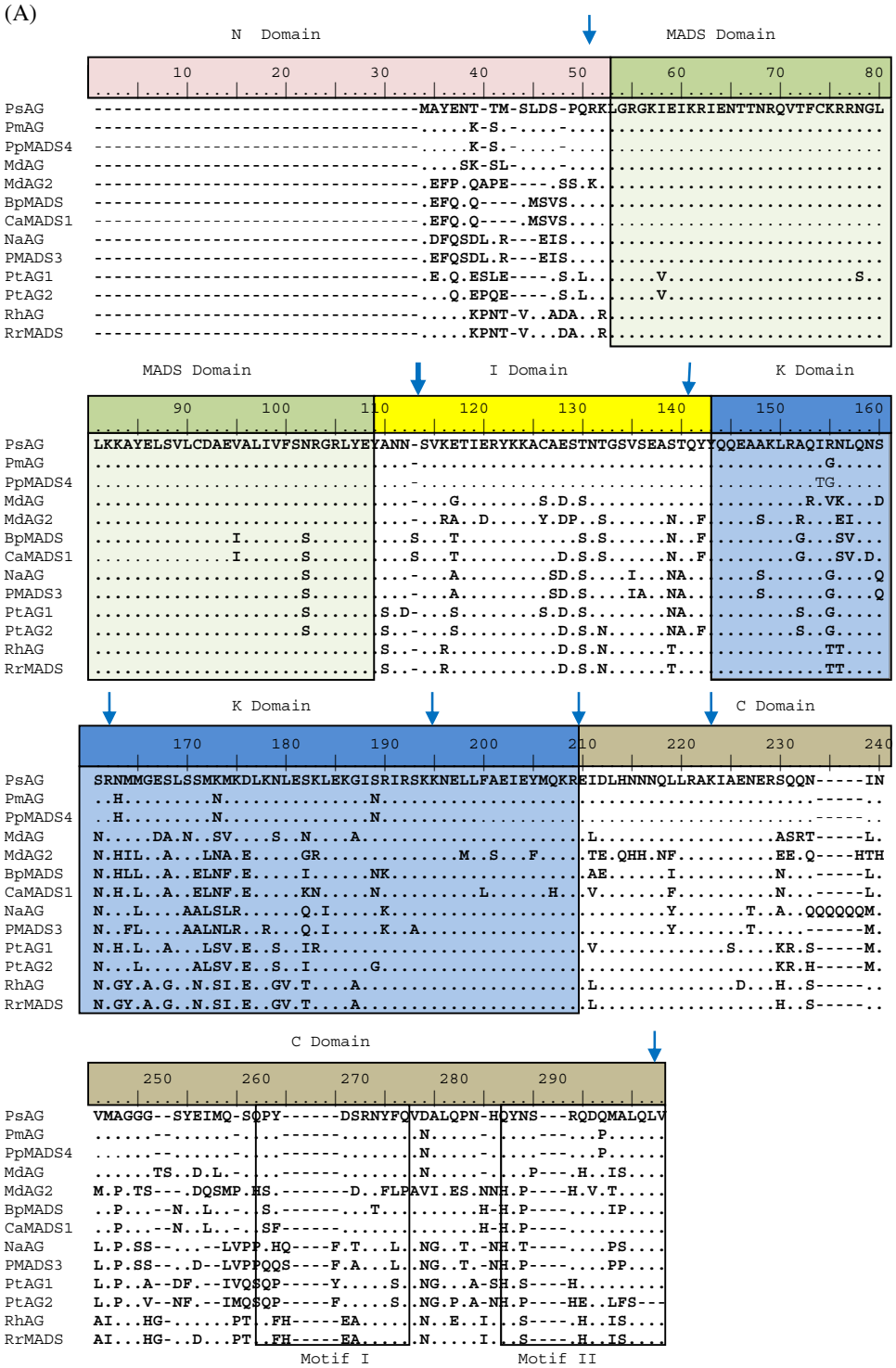
RT-PCR methods enabled the use of flower buds (April collection) as cloning materials. At this stage, stamen and carpel primordia were just appearing, and the *AG* homologous gene was expected to be expressed at a high level, versus most other MADS-box genes. Degenerate primers were designed to be specific to *AG* or its homologs and to discriminate against *AG*-like sequences. The nested PCR approach was used to amplify *PsAG*. The primary PCR program consisted of 20 cycles with an annealing temperature of 48°C to maximize the chance of annealing, and another 25 cycles with the annealing temperature increased to 50°C and 55°C in nested PCR in order to maintain the high specificity of PCR. A total of 20 clones from two independent PCR reactions were sequenced. Among these, BLAST similarity searches identified four clones which were identical and homologous to the corresponding region of *AG*, whereas the others were novel as no similar sequences were found in GenBank. The longest clone was 329 bp. No *AG*-like or other MADS-box sequences were cloned. The cDNA ends were identical.

In the 3′ RACE, a 750 bp product was amplified by nested PCR. In the 5′ RACE, a 650 bp product was amplified by nested PCR. The full-length sequence of the *PsAG* cDNA was obtained by assembling internal *AG* and RACE results. Partial 3′ UTR and 5′ UTR were obtained (GenBank accession number EU938540.1).

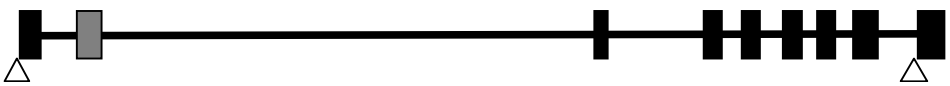
The 729 bp predicted open reading frame of *PsAG* encoded a putative protein of 243 amino acid residues. A homology search in GenBank showed high score match between *PsAG* and other *AG* homologs. A global alignment of *PsAG* and other *AG* amino acid sequences revealed that the black cherry *PsAG* protein was most similar to *PmAG* (*P. mume*) and *PpMADS4*, a functional homolog from *P. persica*. These shared 96.7% and 96.2% identity, respectively. *PsAG* was also 84.7% identical with *MdAG*, an *AG* homolog from apple (*M. domestica*). All of the functional domains defined in *AG* were present in *PsAG*. Within the MADS domain, the black cherry sequence and many other *AG* sequences had only one or two conservative amino acid substitutions (Fig. 1a). Similarity was not limited to the MADS domain, but extended over other regions as well. The least conserved region among *AG* homologs has been in the C-terminus. A comparison in this region revealed several short amino acid motifs that were conserved among *PsAG* and *AG* homologs (Fig. 1a). In summary, our sequence analyses indicated that this black cherry cDNA was the cognate homolog of *AG*.

A comparison of the deduced protein sequence revealed all of the features that were characteristic of *AG* homologs. It had a typical MIKC structure with an N-terminal extension preceding the MADS domain (Kramer et al. 2004; Fig. 1a). This was also seen in other *AG* homologs of woody species. The *PsAG* MADS domain was highly conserved with those of the rose family *AG*, such as *PmAG*, *PpMADS4*, and *MdAG* with sequence similarities of up to

(A)



(B) *PsAG* gene structure



◀ **Fig. 1** **a** The sequence of *AGAMOUS* homologs from a range of plant species. Included are *PmAG* (ABU41518; *P. mume*), *PpMADS4* (AAU29513; *P. persica*), *PsAG* (EU938540.1; *P. serotina*), *RhAG* (AAD00025; *Rosa hybrid*), *RrMADS* (BAA90744; *Rosa rugosa*), *MdAG* (CAC80858; *M. domestica*), *BpMADS* (CAB95649; *Betula pendula*), *CaMADS1* (AAD03486; *Corylus avellana*), *MdAG2* (AQ03090; *M. domestica*), *NaAG* (Q 43585; *Nicotiana tabacum*), *PMADS3* (Q40885; *P. hybrida*), *PtAG1* (AF052570; *Populus trichocarpa*), and *PtAG2* (AF052571; *P. trichocarpa*). The domains were indicated by *different shades*, and conserved amino acids were *highlighted*. Motifs I and II within the C domain were indicated. Introns positions were indicated by *arrows*. **b** *PsAG* gene structure. Exons are depicted in *gray* and introns by *lines*. The position of translational start and stop codons were indicated with *triangles*

100% (Fig. 1a; Table 3). The sequences of the other domains were more variable, but showed that black cherry *PsAG* was more similar to *PmAG* and *PpMADS4* than *PtAG1* (Table 3). The C domain of *PsAG* possessed the *AG* motif I and *AG* motif II sequences (Fig. 1a), which were particularly conserved in lineage that includes the clade with *AG* and *MdMADS15* (Kramer et al. 2004). The sequence data indicated that *PsAG* was structurally very similar to the homologs of *A. thaliana AG* and in particular those from other woody species such as *P. persica* and *P. mume*.

Relationship Among Members of the *AG* Subfamily

Different species of *AG* homolog was compared at the protein level (Table 3). Previous phylogenetic analysis using MIK region revealed that most plants MADS-box genes were organized into monophyletic clades; these clusters generally corresponded to groups of genes that shared related functions (Purugganan et al. 1995). Based on these studies, we selected a representative subset of dicot MADS-box genes to show that *PsAG* belongs to members of *AG* clade. A phylogenetic tree (Fig. 2) estimated by maximum parsimony had a similar topology (data not shown), though a few differences were apparent. *PsAG* was most closely related with *PmAG* and *PpMADS4* with a 94% bootstrap support.

Table 3 Comparison between *PsAG* and related genes in the MADS and K domains and full-length amino acid sequence

AG homolog	MADS (%)	K domain (%)	Full-length identity (%)
PmAG	100	96	96.7
PpMADS4	100	95	96.2
MdAG	100	83	84.7
PtAG1	94.5	79	76.3
AG	98.1	67	70.3

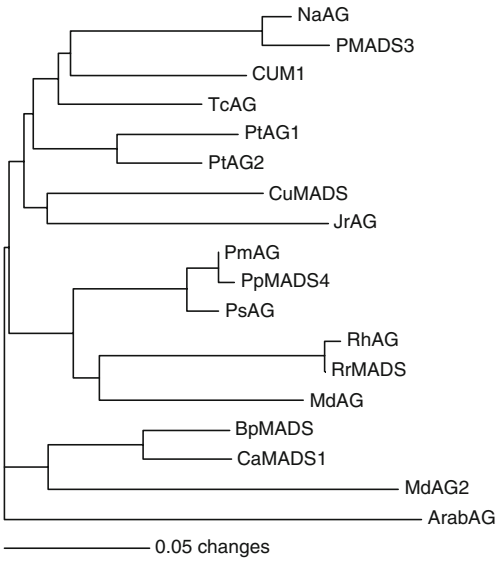


Fig. 2 Phylogenetic analysis of MADS-box genes related to *P. serotina PsAG*. The neighbor-joining tree was generated by Clustal W. The *numbers next to each node* give bootstrap values from 1,000 replicates. GenBank accession numbers of amino acid sequences used: *ArabAG* (X53579; *A. thaliana*), *PmAG* (ABU41518; *P. mume*), *PpMADS4* (AAU29513; *P. persica*), *PsAG* (EU938540.1; *P. serotina*), *RhAG* (AAD00025; *R. hybrid*), *RrMADS* (BAA90744; *R. rugosa*), *MdAG* (CAC80858; *M. domestica*), *CUM1* (AAC08528; *C. sativus*), *BpMADS* (CAB95649; *B. pendula*), *CaMADS1* (AAD03486; *C. avellana*), *MaAG2* (AQ03090; *M. domestica*), *NaAG* (Q 43585; *N. tabacum*), *PMADS3* (Q40885; *P. hybrida*), *PtAG1* (AF052570; *P. trichocarpa*), *PtAG2* (AF052571; *P. trichocarpa*), *CuMADS* (BAF 34911; *Citrus unshiu*), *TcAG* (ABA 39727; *Theobroma cacao*), *JrAG* (CAC 38764; and *Juglans regia*)

Genomic Organization of the *PsAG* Gene

Because the *AG* gene belongs to the dicot C-group, many investigators have reported that the *AG* gene had a particular intron–exon structure with eight introns at conserved positions (Yanofsky et al. 1990; Rutledge et al. 1998; Tandre et al. 1998; Brunner et al. 2000; Kramer et al. 2004). To understand the evolution of genomic organization in the genes of *AG* family, we determined the intron positions of the *PsAG* gene. The intron–exon pattern of the *PsAG* gene was shown in Fig. 1b. Intron 8 was 136 bp in size, which was exactly the same size as the peach intron 8, and larger than intron 8 from *A. thaliana AG*, which was 109 bp. Furthermore, the DNA sequence of the genomic clones revealed the presence of intron 8 that interrupts the last codon of the protein (Fig. 1b). Intron 8 was found in all members of the *AG* subfamily (Kramer et al. 2004). We failed to clone intron 2 by using primer I2F/I2R, but it was possible that the size of intron 2 was too large for PCR by using general Taq DNA polymerase. In *AG* homologous

genes, intron 2 was 5 to 7 kb long (Brunner et al. 2000). All other introns were retrieved in *PsAG* at the same position as in *PtAG1*. Intron sizes were moderate, being generally larger in *PsAG* than in *AG*, but shorter than *PtAG1*. Similar to other known MADS-box genes, exons in the *PsAG* gene did not correspond exactly to protein domains (Shore and Sharrocks 1995). For example, the K domain spans three exons, while the MADS domain only exists in one exon.

Southern blot analysis using a black cherry gene-specific probe derived from the *PsAG* cDNA 3′ end demonstrated the presence of only a single *PsAG* gene in the genome of black cherry (Fig. 3). The restriction fragment pattern was different for different enzymes. It indicated that *Kpn* I cut genomic DNA sequence of *PsAG* once because it resulted in two hybridization bands when genomic DNA was digested with this enzyme. Sequencing data confirmed this, and it is at 871 bp at 3′ UTR.

In Situ Hybridization

The expression patterns of *PsAG* in various floral tissues at different stages were shown in Fig. 4. Longitudinal sections were hybridized with an antisense or sense probes. In the early stages of flower development, *PsAG* transcripts were present in the corolla and ovary primordia (Fig. 4b). *PsAG* mRNA could be seen in the immature corolla and immature ovary (Fig. 4d). At the middle stage of flower development, the *PsAG* gene signal was apparent in mature pollen, anthers, filament, but not in the corolla (Fig. 4f). Once the ovaries developed in the late spring, *PsAG* mRNA could be seen to accumulate in the ovary (Fig. 4g), and expression decreased in the stamen, but was enhanced in the ovule relative to the earlier floral development stage. During the late stage of flower development, *PsAG* mRNA continued to accumulate at high levels in the ovules (Fig. 4h). *PsAG* transcripts were not detected in the flower using sense RNA probes (Fig. 4a, c, e). In the late floral development stage, *AG* RNA was present in integuments, becoming restricted to the endothelium (cell layer surrounding the embryo sac) in mature ovules (Fig. 4i). This pattern was expected for C-function genes such as *A. thaliana AG*, and this matches the observations made by *in situ* hybridization from other studies (Bowman et al. 1991a, b; Coen and Meyerowitz 1991).

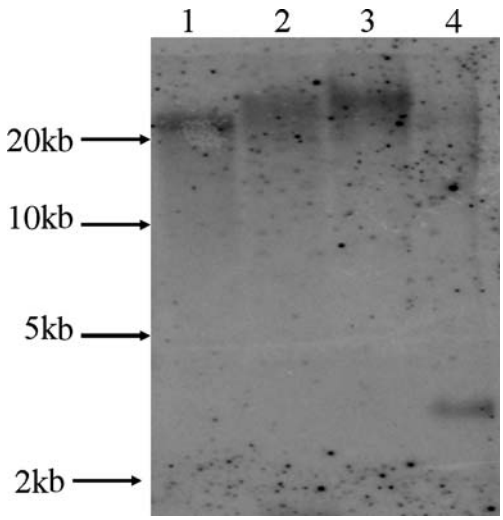


Fig. 3 Southern blot analysis of the *PsAG* gene. Genomic DNA was digested in *lane 1* by *Bam* HI, *lane 2 Xba* I, *lane 3 Xho* I, and *lane 4 Kpn* I; a single band was observed in *lanes 1–3* suggesting that *PsAG* was a single-copy gene. Two hybridization bands were observed in *lane 4* indicating that *Kpn* I cut within the genomic sequence of *PsAG* once

Discussion

Phylogenetic analysis with various angiosperm genes have revealed that MADS genes contain highly conserved motifs (Kramer et al. 1998; Vandenbussche et al. 2003; Kramer et al. 2004). Kramer et al. (2004) reported that the *AG* gene family had two monophyletic groups, namely the C-class genes and the D-class genes. Studies revealed that all plant *AG* homologs had the characteristic MIKC (MADS, intervening, keratin-like C-terminal) domain structure that was characteristic of the type II or MIKC-type MADS-box proteins (Kramer et al. 2004). Over 100 MADS-box genes in *A. thaliana* (Martinez-Castilla and Alvarez-Buylla 2003; Parenicová et al. 2003) have been isolated. A large number of MADS-containing genes were expected to be present in black cherry. A comparison of the deduced protein sequence indicated that *PsAG* was an *AG* homolog (Fig. 1a). First, it had a MIKC structure with an N-terminal extension preceding the MADS domain. Second, intron 8 fell within the last codon of the protein (Fig. 1b), which happens in all members of the *AG* subfamily (Kramer et al. 2004). Third, *PsAG* had *AG* motifs I and II (two diagnostic amino acid motifs) in its C- terminal region, as defined by Kramer et al. (2004). Fourth, the expression pattern was restricted to the two inner whorls, where stamens and carpels were situated.

The intron 2 of *A. thaliana* was about 3 kb and contains functionally important regulatory sequences for *AG* expression. In *PsAG*, intron 2 was about 5 to7 kb, much larger than that of *A. thaliana*. Intron 2 of *PsAG* might contain more complex regulatory sequences than that seen in *AG*. Previous studies indicated that intron 2 was important for its normal expression pattern (Deyholos and Sieburth 2000). The 3′ sequence of intron 2 was required for *AG* carpel-specific expression, while sequence within the 5′ part of intron 2 was required for stamen expression (Sieburth and Meyerowitz 1997; Deyholos and Sieburth 2000). The

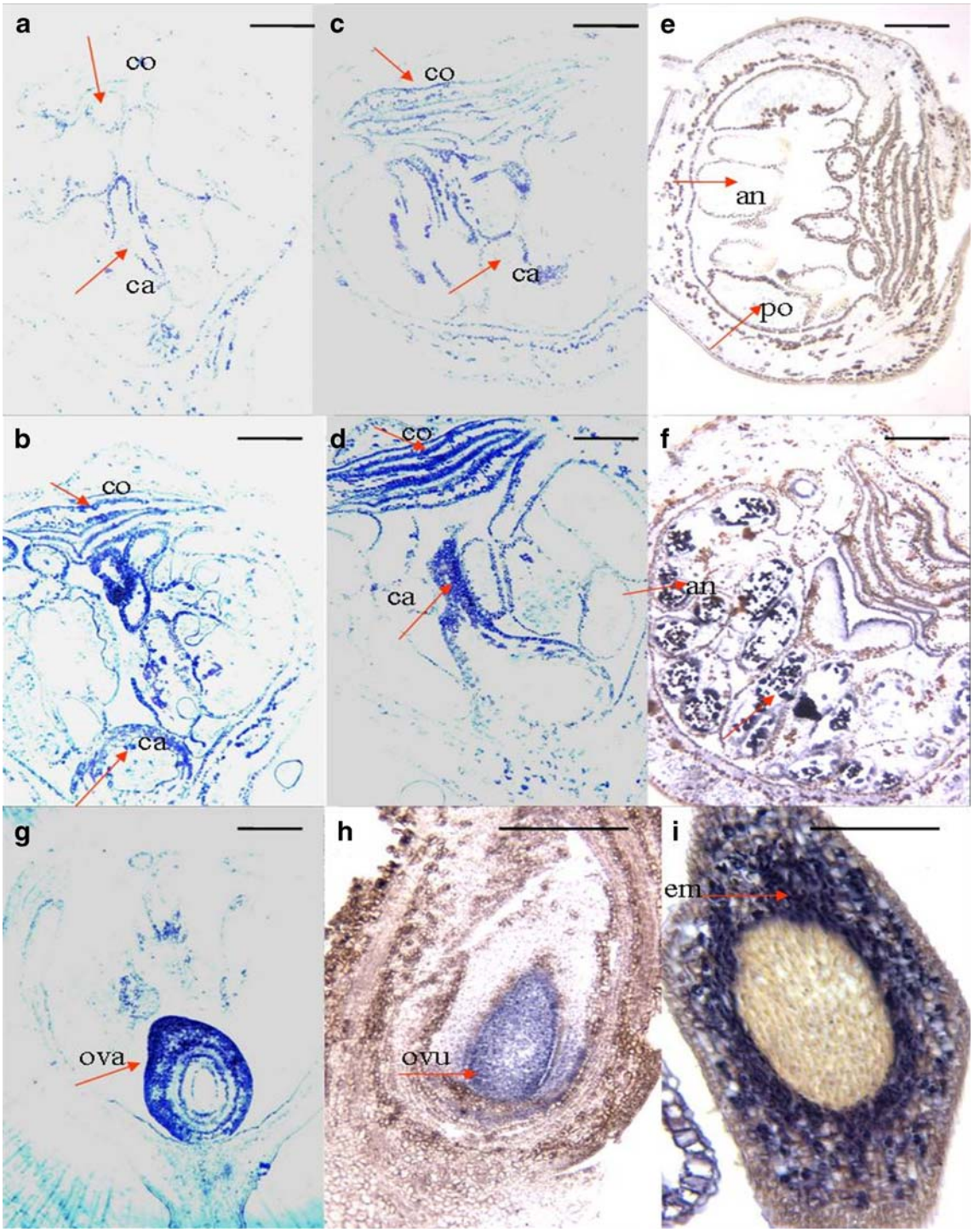


Fig. 4 *In situ* hybridization. Signals were in *blue* or *purple* and indicated by *arrows*. **a**–**b**, **c**–**d**, and **e**–**f** were pairs; sections came from the same flower sample. **a**, **c**, and **e** Controls, hybridized with sense probe. **b** *PsAG* showed expression in corolla and ovary primordia; **d** expression in the corolla, immature ovary, and stamen; **f** *PsAG* mRNA could be seen in the pollen, anthers, but not in the corolla; **g** the late stage of flower development, *PsAG* transcript expression decreased in the stamen but was enhanced in the ovary; **h** *PsAG* expressed in the ovules and disappeared in the stamen; **i** *PsAG* RNA was present in the endothelium in mature ovules. Scale bar=200 μm, *co* corolla, immature petals; *ca* carpel; *ova* ovary; *ovu* ovule; *sa* stamen; *sti* stigma; *st* style; *po* pollen; *an* anther

level of *AG* expression was determined by cis-regulatory elements localized in intron 2 (Hong et al. 2003). Fusing *A. thaliana AG* enhancer elements from intron 2 with 35 S promoter and DT or Barnase imparted transgenic *A. thaliana* with 68% and 89% ablation of stamens and carpels, and the resulting plants were completely sterile (Liu and Liu 2008). In tobacco, a 1,853-bp nucleotide sequence from the 3′ end of the second *AG* intron was fused with *AG* and Barnase (*AG*-I-35 S::Barnase) and resulted in 98% sterility in transgenic tobacco (Wang et al. 2008). The regulatory regions in woody species have not yet been fully elucidated. However, it would be valuable information for engineering reproductive sterility.

We characterized the expression pattern of an *AG* gene from *P. serotina*. The pattern of *PsAG* expression was typical of C-function floral genes (Bowman et al. 1991a, b). The expression pattern of the *PsAG* was analyzed by RT-PCR and *in situ* hybridization. Using total RNA isolated from stems, leaves, and flower buds, the RT-PCR revealed that *PsAG* was expressed specifically in flowers (data not shown). No transcript was detected in vegetative tissues such as stems and leaves. However, *in situ* hybridization detected low levels of expression in the corolla at the early stage of flower development, which disappeared when the flowers matured. This was similar with what has been observed in poplar and apple, where *AG* had low expression in vegetative tissues such as leaves (Brunner et al. 2000; van der Linden et al. 2002). *In situ* hybridization revealed that *PsAG* transcripts accumulated in floral primordia and developing flowers in the inner two whorls. The high level of expression of *PsAG* detected in the carpel throughout flower development from early stage until flowers matured supported that *PsAG* was a C-function gene. This result was consistent with the expression pattern in *PpAG* (Martin et al. 2006) and *MdMADS4* (Sung et al. 2000).

It was reported that in the Evergrowing peach mutant, a mutation in one of the MADS-box genes affects the ability to enter dormancy and survive in southern Mexico (Bielenberg et al. 2004). This also demonstrated that controlling important qualities can be achieved by MADS-box transcription factors. In our study, an RNAi construct was generated with the *PsAG* gene, without its MADS-box domain, into a binary vector. In *A. thaliana*, 80% of transformants were silenced using a similar type of construct (Wesley et al. 2001). It is anticipated that expression of this construct will lead to the silencing of the *PsAG* and impart sterility in black cherry. This RNAi construct is currently being used in studies for transformation of black cherry.

Acknowledgments The authors gratefully acknowledge Dr. Jin-Rong Xu (Purdue University) for his advice, technical guidance, and support with this research, and Drs. Brian Miki and Michael J. Zanis for their constructive review and suggestions for the improvement of this manuscript. Mention of a trademark, proprietary product, or vendor does not constitute a guarantee or warranty of the product by the US Department of Agriculture and does not imply its approval to the exclusion of other products or vendors that also may be suitable.

References

Bielenberg DG, Wang Y, Fan S, Reighard GL, Scorza R, Abbott AG (2004) A deletion affecting several gene candidates is present in the Evergrowing peach mutant. J Hered 95:436–444

Boss PK, Vivier M, Matsumoto S, Dry IB, Thomas MR (2001) A cDNA from grapevine (*Vitis vinifera* L.) shows homology to *AGAMOUS* and *SHATTERPROOF* is not only expressed in flowers but also throughout berry development. Plant Mol Biol 45:541–553

Bowman JL, Drews GN, Meyerowitz EM (1991a) Expression of the *Arabidopsis* floral homeotic gene *AGAMOUS* is restricted to specific cell types late in flower development. Plant Cell 3:749–758

Bowman JL, Smyth DR, Meyerowitz EM (1991b) Genetic interactions among floral homeotic genes of *Arabidopsis*. Development 112:1–20

Brunner AM, Rottmann WH, Sheppard LA, Krutovskii K, DiFazio SP, Leonardi S, Strauss SH (2000) Structure and expression of duplicate *AGAMOUS* orthologues in poplar. Plant Mol Biol 44:619–634

Coen ES, Meyerowitz EM (1991) The war of the whorls: genetic interactions controlling flower development. Nature 353:31–37

Davies B, Motte P, Keck E, Saedler H, Sommer H, Schwarz-Sommer Z (1999) *PLENA* and *FARINELLI*: redundancy and regulatory interactions between two *Antirrhinum* MADS-box factors controlling flower development. EMBO J 18:4023–4034

Deyholos MK, Sieburth LE (2000) Separable whorl-specific expression and negative regulation by enhancer elements within the *AGAMOUS* second intron. Plant Cell 12:1799–1810

Dower WJ, Miller JF, Ragsdale CW (1988) High efficiency transformation of *E. coli* by high voltage electroporation. Nucl Acids Res 16:6127

Fan HY, Hu Y, Tudor M, Ma H (1997) Specific interactions between the K domains of *AG* and *AGLs*, members of the MADS domain family of DNA binding proteins. Plant J 12:999–1010

Hasebe M, We CK, Kat M, Banks JA (1998) Characterization of MADS homeotic genes in the fern *Ceratopteris richardii*. Proc Natl Acad Sci 95:6222–6227

Hong RL, Hamaguchi L, Busch MA, Weigel D (2003) Regulatory elements of the floral homeotic gene *AGAMOUS* identified by phylogenetic footprinting and shadowing. Plant Cell 15:1296–1309

Honma T, Goto K (2001) Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. Nature 409:525–529

Jackson D (1991) *In-situ* hybridization in plants. In: Gurr SJ, McPherson MJ, Bowles DJ (eds) Molecular plant pathology: a practical approach. Oxford University Press, Oxford, pp 163–174

Jager M, Hassanin A, Manuel M, Le Guyader H, Deutsch J (2003) MADS-box genes in *Ginkgo biloba* and the evolution of the *AGAMOUS* family. Mol Biol Evol 20:842–854

Kater MM, Colombo L, Franken J, Busscher M, Masiero S, Van Lookeren Campagne MM, Angenent GC (1998) Multiple *AGAMOUS* homologs from cucumber and petunia differ in their ability to induce reproductive organ fate. Plant Cell 10:171–182

Kitahara K, Matsumoto S (2000) Rose MADS-box genes '*MASAKO C1* and *D1*' homologous to class C floral identity genes. Plant Sci 151:121–134

Kotoda N, Wada M, Komori S, Kidou S, Abe K, Masuda T, Soejima J (2000) Expression pattern of homologues of floral meristem identity genes *LFY* and *AP1* during flower development in apple. J Amer Soc Hort Sci 125:398–403

Kotoda N, Wada M, Kusaba S, Kano-Murakami Y, Masuda T, Soejima J (2002) Overexpression of *MdMADS5*, an *APETALA1*-like gene of apple, causes early flowering in transgenic *Arabidopsis*. Plant Sci 162:679–687

Kramer EM, Dorit RL, Irish VF (1998) Molecular evolution of gene controlling petal and stamen development: duplicate and divergence within the *APETALA3* and *PISTILLATA* MADS box gene lineages. Genetics 149:765–783

Kramer EM, Jaramillo M, Di Stilio VS (2004) Patterns of gene duplication and functional evolution during the diversification of the *AGAMOUS* subfamily of MADS box genes in angiosperms. Genetics 166:1011–1023

Lemmetyinen J, Hassinen M, Elo A, Porali I, Keinonen K, Makela H, Sopanen T (2004) Functional characterization of *SEPALLATA3* and *AGAMOUS* orthologues in silver birch. Physiol Plant 121:149–162

Liu ZR, Liu ZC (2008) The second intron of *AGAMOUS* drives carpel and stamen-specific expression sufficient to induce complete sterility in *Arabidopsis*. Plant Cell Rep 27:855–863

Liu JY, Huang YH, Ding B, Tauer CG (1999) cDNA cloning and expression of a sweetgum gene that shows homology with *Arabidopsis AGAMOUS*. Plant Sci 142:73–82

Martin T, Hu M, Labbé H, McHugh S, Svircev A, Miki B (2006) PpAG1, a homolog of *AGAMOUS*, expressed in developing peach flowers and fruit. Can J Bot 84:767–776

Martinez-Castilla LP, Alvarez-Buylla ER (2003) Adaptive evolution in the *Arabidopsis* MADS-box gene family inferred from its complete resolved phylogeny. Proc Natl Acad Sci USA 100:13407–13412

Meilan R, Brunner AM, Skinner JS, Strauss SH (2001) Modification of flowering in transgenic trees. In: Morohoshi N, Komamine A (eds) Molecular Breeding of Woody Plants, pp. 247–256

Parenicová L, de Floter S, Kieffer M, Horner DS, Favalli C, Busscher J, Cook HE, Ingram RM, Kater MM, Davies B, Angenent GC, Colombo L (2003) Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in *Arabidopsis*: new openings to the MADS world. Plant Cell 15:1538–1551

Purugganan MD, Rounsley SD, Schmidt RJ, Yanofsky MF (1995) Molecular evolution of flower development: diversification of the plant MADS-box regulatory gene family. Genetics 140: 345–356

Rigola D, Pè ME, Fabrizio C, Me G, Sari-Gorla M (1998) *CaMADS1*, a MADS box gene expressed in the carpel of hazelnut. Plant Mol Biol 38:1147–1160

Rutledge R, Regan S, Nicolas O, Fobert P, Cote C, Bosnich W, Kauffeldt C, Sunohara G, Seguin A, Stewart D (1998) Characterization of an *AGAMOUS* homologue from the conifer black spruce (*Picea mariana*) that produces floral homeotic conversions when expressed in *Arabidopsis*. Plant J 15:625–634

Salzman RA, Fujita T, Zhu-Salzman K, Hasegawa PM, Bressan RA (1999) An improved RNA isolation method for plant tissues containing high levels of phenolic compounds or carbohydrates. Plant Mol Biol Rep 17:11–17

Shore P, Sharrocks AD (1995) The MADS-box family of transcription factors. Eur J Biochem 229:1–13

Shu GP, Baum DA, Mets LJ (1999) Detection of gene expression patterns in various plant tissues using non-radioactive mRNA *in situ* hybridization. The World Wide Web J Biology (http://epress.com/w3jbio/vol4/shu/paper.htm)

Sieburth LE, Meyerowitz EM (1997) Molecular dissection of the *AGAMOUS* control region shows that cis elements for spatial regulation are located intragenically. Plant Cell 9(3):355–365

Sieburth LE, Running MP, Meyerowitz EM (1995) Genetic separation of third and fourth whorl functions of *AGAMOUS*. Plant Cell 7:1249–1258

Stairs G, Hauck W (1968) Reproductive cytology of black cherry (*Prunus serotina* Ehrh.) Proc 15th NE For Tree Improvement Conf, Morgantown, WV, pp. 42-53

Strauss SH, Rottmann WH, Brunner AM, Sheppard LA (1995) Genetic engineering of reproductive sterility in forest trees. Molec Breed 1:5–26

Sung SK, Yu GH, An G (1999) Characterization of *MdMADS2*, a member of the *SQUAMOSA* subfamily of genes, in apple. Plant Physiol 120:969–978

Sung SK, Yu GH, Nam J, Jeong DH, An G (2000) Developmentally regulated expression of two MADS-box genes, *MdMADS3* and *MdMADS4*, in the morphogenesis of flower buds and fruits in apple. Planta 210:519–528

Swofford DL (2003) *PAUP**: phylogenetic analysis using parsimony (* and other methods). Version 4.0b10. Sinauer Associates, Sunderland

Tandre K, Svenson M, Svensson ME, Engstrom P (1998) Conservation of gene structure and activity in the regulation of reproductive organ development of conifers and angiosperms. Plant J 15:615–623

Theissen G, Saedler H (2001) Plant biology: floral quartets. Nature 409:469–471

Vandenbussche M, Theissen G, Van de Peer Y, Gerats T (2003) Structural diversification and neo-functionalization during floral MADS-box gene evolution by C-terminal frameshift mutation. Nucl Acids Res 31:4401–4409

van der Linden CG, Vosman B, Smulders MJM (2002) Cloning and characterization of four apple MADS box genes isolated from vegetative tissues. J Exp Bot 53:1025–1036

Wada M, Cao Q, Kotoda N, Soejima J, Masuda T (2002) Apple has two orthologues of *FLORICAULA/LEAFY* involved in flowering. Plant Mol Biol 49:567–577

Wang HZ, Hu B, Chen GP, Shi NN, Zhao Y, Yin QC, Liu JJ (2008) Application of *Arabidopsis AGAMOUS* second intron for the engineered ablation of flower development in transgenic tobacco. Plant Cell Rep 27:251–259

Wesley SV, Helliwell CA, Smith NA, Wang MB, Rouse DT, Liu Q, Gooding PS, Singh SP, Abbott D, Stoutjesdijk PA, Robinson SP, Gleave AP, Green AG, Waterhouse PM (2001) Construct design for efficient, effective and high-throughput gene silencing in plants. Plant J 27:581–590

Yanofsky MF, Ma H, Bowman JL, Drews GN, Feldmann KA, Meyerowitz EM (1990) The protein encoded by the *Arabidopsis* homeotic gene *AGAMOUS* resembles transcription factors. Nature 346:35–39

Yao JL, Dong YH, Kvarnheden A, Morris B (1999) Seven MADS-box genes in apple are expressed in different parts of the fruit. J Amer Soc Hort Sci 124:8–13

Zhang P, Tan HTW, Pwee KH, Kumar PP (2004) Conservation of class C function of floral organ development during 300 million years of evolution from gymnosperms to angiosperms. Plant J 37:566–577