Plant Molecular Biology 44: 619-634, 2000. © 2000 Kluwer Academic Publishers. Printed in the Netherlands.

Structure and expression of duplicate AGAMOUS orthologues in poplar

Amy M. Brunner, William H. Rottmann[1], Lorraine A. Sheppard[2], Konstantin Krutovskii, Stephen P. DiFazio, Stefano Leonardi[3] and Steven H. Strauss*

*Department of Forest Science, Oregon State University, Corvallis, OR 97331, USA (*author for correspondence; e-mail: strauss@fsl.orst.edu); present addresses: 1 Westvaco Forest Science and Technology, PO. Box 1950, Summerville, SC 29484, USA; 21nstitute of Forest Genetics, USDA Forest Service c/o Department of Environmental Horticulture, One Shields Ave., University of California, Davis, CA, 95616, USA; 3Department of Environmental Science, University of Parma, Parco Area delle Scienze 33a, 43100 Parma, Italy*

Received 9 November 1999; accepted in revised from 24 July 2000

Key words: AGAMOUS, cottonwoods, dioecy, floral development, MADS-box, Populus

Abstract

To investigate the homeotic systems underlying floral development in a dioecious tree, and to provide tools for the manipulation of floral development, we have isolated two Populus *trichocarpa* genes, *PTAGI* and *PTAG2,* homologous to the *Arabidopsis* floral homeotic gene *AGAMOUS (AG). PTAGI* and *PTAG2* are located on separate linkage groups, but their non-coding regions are highly similar, consistent with a phylogenetically recent duplication. Intron/exon structure is conserved in relation to *AG* and the *Antirrhinum AG* orthologue, *PLENA (PLE),* and low-stringency Southern analysis demonstrated the absence of additional genes in the poplar genome with significant *PTAG1/2* homology. *PTAGI* and *PTAG2* exhibit an *AG-like* floral expression pattern, and phylogenetic analysis of the *AG* subfamily strongly supports evolutionary orthology to C-class organ identity genes. The high degree of similarity shared by *PTAGI* and *PTAG2* in both sequence (89% amino acid identity) and expression indicates that they are unlikely to be functionally associated with specification of tree gender. Unexpectedly, *PTAG* transcripts were consistently detected in vegetative tissues.

Abbreviations: AG, AGAMOUS; MIK region (region encompassing the MADS-box, intervening region and Kbox); MYA, million years ago; *PLE, PLENA; PTAG, Populus trichocarpa AGAMOUS;* SSR, simple sequence repeat; UTR, untranslated region

Introduction

Studies in a variety of plants indicate that the genetic mechanisms determining floral organ identity are generally conserved among angiosperms. These mechanisms were first elucidated in *Arabidopsis* and *Antirrhinum,* and formulated as the ABC model of flower development (Coen and Meyerowitz, 1991). According to this model, three classes of homeotic genes (A, B, and C) act alone and in combination to specify the four types of floral organs found in the typical angiosperm flower. Most floral homeotic genes encode proteins belonging to a family of transcription factors named after its highly conserved DNA binding and dimerization region, the MADS domain (reviewed in Riechmann and Meyerowitz, 1997). Most plant members of the MADS-box family consist of four domains, including a second conserved domain, the K-box, which is named for its similarity to the coiled coil region of keratin. The two additional domains are called the intervening region (1-region), which lies between the MADS domain and K-box and is necessary for dimerization, and the highly variable C-terminal region.

We are studying floral regulatory genes in poplar (genus *Populus*, including aspens and cottonwoods).

Poplars display characteristics commonly found in temperate trees, such as inflorescences adapted for wind pollination, an extended juvenile phase, and a prolonged period of floral development interrupted by dormancy. The species we study is *Populus trichocarpa* Torn and Gray (black cottonwood), which is native to the Pacific Northwest and widely used in poplar breeding programs in many parts of the world. Populus is widely acknowledged as a model system for genetic and molecular analyses of longlived, woody plants because of its fast growth, small genome (5 x 108 by haploid), ease of controlled breeding and vegetative propagation, and facile transformation/regeneration systems (reviewed in Bradshaw, 1998).

The reproductive development of poplars is similar to that of willows (*Salix* spp.), the other genus in the taxonomically distinct family Salicaceae (Eckenwalder, 1996). Flowering typically begins at 5 to 15 years of age in the wild. Floral development is initiated in late spring, nearly a year before anthesis occurs (Boer and Strauss, 1994). Flowers are borne on pendulous racemes (i.e., catkins) that appear before flushing of vegetative buds. After anthesis, the inflorescences for the following year are initiated in the axils of leaves. The inflorescences develop as axillary buds and differentiate acropetally. Floral meristems arise spirally in the axils of bracts, with their differentiation continuing through autumn. After dormancy, sporogenesis and gametogenesis occur, and the inflorescences rapidly elongate and emerge from the bud scales.

Poplars are dioecious and their flowers are highly simplified, consisting only of an outer whorl organ called the perianth cup and an inner whorl of either stamens or carpels. The homologies of the perianth cup with sepals and petals are uncertain. Poplar's reduced flowers have traditionally been considered as highly evolved (Eckenwalder, 1996), and recent phylogenetic studies of flowering plants based on DNA sequences (Soltis *et al.,* 1999) indicate that Populus is of relatively recent origin and probably derived from an ancestor with a typical four-whorled dicot flower. The Salicaceae are in the order Malipighiales, which includes diverse families, most of which have flowers with all four organ types. In agreement with earlier classifications (reviewed in Eckenwalder, 1996), molecular analyses showed that genera of the tropical family Flacourtiaceae were most closely related to the Salicaceae.

Unlike most dioecious and monoecious species, which selectively arrest development of stamen or carpels, poplars do not visibly initiate primordia of the alternate sex (e.g., Boes and Strauss, 1994; Kaul, 1995). However, rare occurrences of hermaphroditic flowers, as wel I as flowers of the opposite gender, have been reported for most poplar species (e.g., Stettler, 1971).

The mechanism of sex determination in the Salicaceae is unknown. Though gender is genetically controlled, neither sex chromosomes nor ratios of autosome to sex-determining loci appear to be involved (McLetchie and Tuskan, 1994, Alstro in -Rapaport *et al.,* 1997). Analyses of sex ratio and of a DNA marker associated with femaleness in Salix *viminalis* suggest that gender is controlled by multiple loci interacting in an epistatic manner (Al strom-Rapaport *et al.,* 1997, 1998). Environmental factors also appear to affect gender determination (reviewed in McLetchie and Tuskan, 1994).

In the hermaphroditic flowers of *Arabidopsis* and *Antirrhinum,* B and C class genes act together to specify stamens in whorl 3, while C activity alone specifies carpels in whorl 4 and determinacy of the floral meristem. In *Arabidopsis, AG* specifies C function; the corresponding homologue in *Antirrhinuan is PLE* (Yanofsky *et al.,* 1990; Bradley *et* al., 1993). In poplars and other plants in which sex determination occurs very early in floral development, the expression patterns of at least some of the genes specifying reproductive organ identity are likely to differ between developing female and male flowers. For example, in sorrel, a dioecious annual plant, a putative C-function gene exhibits a sex-specific expression pattern, while class B gene homologues do not (Ainsworth *et al.,* 1995). Moreover, some species contain two *AG* homologues that in some cases exhibit diversification in function and/or expression. In maize, two closely related genes that have distinct but partially redundant activities may specify C-function (Mena *et al.,* 1996). Loss-of-function mutations in the maize *AG* homologue ZAGl cause an indeterminate pattern of growth, but stamens develop normally and fourthwhorl organs have carpelloid features. ZAGI RNA accumulates more in carpels than in stamens, whereas a second AG homologue, *ZMM2,* exhibits the reverse pattern.

To begin to understand the specification of reproductive organs in poplars, we have isolated cDNA and Genomic clones for two AG homologues from *P. trichocarpa.* Through analysis of sequence, expression, and phylogeny of the AG subfamily, we show that these genes were produced by an evolutionarily re

cent duplication of the poplar *AG* orthologue. Both *PTAGI* and *PTAG2* are expressed in the inner whorl of developing floral meristems from female and male trees, consistent with a function in specifying reproductive organ identity and floral determinacy. Unlike most other close *AG* homologues, they also show considerable vegetative expression.

Materials and methods

Plant materials and nucleic acid extraction

Tissues were gathered from wild *P. trichocarpa* trees growing in the vicinity of Corvallis, OR. Immature inflorescence tissue was collected in late May to early June when floral meristems had already initiated and floral organ primordia had begun to form in some of the meristems. Mature inflorescences were collected in late February as they were emerging from the buds. After removing the bud scales, entire inflorescences were collected, and tissue samples from different trees were kept separate. Post-dormancy vegetative buds were collected after inflorescences had emerged. Newly initiated vegetative buds, leaves and stems were collected in early June from basal, non-flowering branches or root sprouts from mature trees. Genomic DNA was isolated from vegetative buds using a modified CTAB method (Wagner *et al.,* 1987). Total RNA was extracted according to the method of Hughes and Galau (1988; also see http://www.fsl.orst.edu/tgerc/rnaextra.htm).

Isolation of cDNA and genomic clones

Immature and mature female floral cDNA libraries were prepared using the λZAP cloning kit (Stratagene). To construct the genomic library, DNA from a single male tree was partially digested with Sau3AI, ligated into a,GEM12 (Promega) at partially filledin *XhoI* sites, and packaged with GigaPack Gold 11 (Stratagene). To avoid cross-hybridization with other MADS-box genes, all probes lacked most or all of the MADS-box and were labeled with a random primer labeling kit (Boehringer Mannheim). A 737 by *HindIII-EcoRI fragment* of the *AG* cDNA (Yanofsky *et al.,* 1990) was isolated from pCIT565 (provided by E. Meyerowitz, California Institute of Technology) and used to screen 6 x 105 genomic clones. Hybridization was performed at 65 °C in 5 x SSPE, 5 x Denhardt's solution, 0.2%o SDS, 100μg/ml denatured salmon sperm DNA. Filters were washed at 42 °C in 2x SSPE, 0.2% SDS and in 0.2x SSPE, 0.2% SDS.

A *PTAGI* cDNA clone was obtained by probing 1.3 x 105 clones of an immature female floral cDNA library with the *HindIII-Xbal* fragment of pCIT565. To isolate a *PTAG2* cDNA clone, genespecific primers 5'-CGACAGCACATGAATTTG-3' and 5'-TTACACTAACTGAAGAGCTGG-3' were designed based on the *PTAG2* genomic sequence. A 147 by fragment corresponding to the 3' end of the *PTAG2* coding region was obtained by PCR using a mature female floral cDNA library as the template. The fragment was then used to probe 9 x 105 clones of a mature female floral library at high stringency. Hybridization conditions were as described above, but washes were performed at 65 °C.

Fragments of the genomic clones and the *PTAGI* cDNA clone were subcloned into pBluescript 11 KS (Stratagene) to facilitate sequencing. Sequencing was done at the Oregon State University Central Services Laboratory using an Applied Biosystems model 377 sequencer.

Genomic DNA blots

DNA samples (5 μg) from a male and a female tree were digested with restriction enzymes and analyzed using standard procedures. Three different probes were used and all were labeled with a random primer kit (Boehringer Mannheim). One probe was an 800 by *PstI-XhoI* fragment from the *PTAGI* cDNA clone, which lacked the MADS-box region, but cross-hybridized with *PTAG2.* A 274 by *PTAGI* speciflc probe, corresponding to the last 147 by of the coding region and part of the 3'-UTR, was synthesized using 5'-CGACAGAGCATGAATTTG-3' and 5'-CCAGACAAATATGATTTACG-3'. The *PTAG2* probe was the 147 by PCR fragment used to isolate the cDNA clone (described above). Hybridization was performed at 65 °C in 5 x SSPE, 5 x Denhardt's solution, 0.2% SDS, 0.05 x Blotto, 100 μg/ml denatured salmon sperm DNA, and final washes were done at 65 °C in 0.1 x SSPE, 0.5% SDS or at a reduced stringency at 42 °C in 0.2x SSPE, 0.2% SDS.

Mapping of PTAGI *and* PTAG2

The *PTAGI* and *PTAG2* loci were placed on a genetic linkage map created from a three-generation pedigree resulting from an interspecific cross between *P. trichocarpa* and *P. deltoides* (Bradshaw, 1994). We analyzed segregation of *PTAG1* and *PTAG2* alleles among

89 *FZ* progeny from this pedigree. Allelic polymorphisms resulted from AT/TA simple sequence repeats (SSRs) located in the second intron of PTAGI and the 5'-flanking region of PTAG2. We designed primers for the regions flanking the SSRs and determined optimal PCR conditions for each locus. PTAGI primers were 5'-CTTGTAATTAAGAGCAAGCCA-3' and 5'ATGTTAAACTACCTCAAACATATCC-3'; PTAG2 primers were 5'-CGAATATAGTGGATGGTTATTG3' and 5'-CGAATCTGAGTAGGAGAGATG-3'. PCR reactions were carried out in a 20 μl volume containing a total of 12 ng of genomic DNA, 0.25 μM of each oligonucleotide primer (Operon), 1 unit *Taq* DNA polymerase (Life Technologies), 200 μM of each dNTP, 20 mM Tris-HCl pH 8.4, 50 mM KCI, 3 MM MgC12 (PTAGI) or 4 mM $MgC1_2$ (PTAG2). Amplification conditions were as follows: one cycle of 94 °C for 4 min, 40 cycles of 94 °C for 15 s, 50 °C (PTAGI) or 56 °C (PTAG2) for 15 s, and 72 °C for 1 min; and one cycle of 72 °C for 10 min. PCR products were detected using a 5' fluorescent label and the ABI Prism 377 DNA Sequencer. Loci were mapped with MAPMAKER 3.0 (Lincoln *et al.,* 1992) as described in Bradshaw *et al.* (1994).

Sequence analyses

Pairwise sequence alignments were performed using the ALIGN program of the FASTA version 2.0 sequence analysis package (Pearson, 1990). Gaps were not counted in determining percentage of identity. Maximum likelihood estimates of synonymous and non-synonymous nucleotide substitution rates and the corresponding estimates of Nei and Gojobori (1986) were calculated using the SYNDIST program (Muse, 1996). Multiple alignments of protein sequences for phylogenetic analyses were done using Clustal X (Thompson *et al.,* 1994) and refined visually. All sequences apart from PTAGI and PTAG2 were obtained from GenBank.

The neighbor-joining tree method (Saitou and Nei, 1987) of the MEGA computer program (Kumar *et al.,* 1993) was used for phylogenetic analysis using amino acid sequences and p-distance matrices with Poisson correction. Consensus trees and estimates of statistical confidence were inferred from 1000 bootstrapped data sets. Tajima and Nei's (1984) genetic distance *(D)* was used as a measure of nucleotide divergence between sequences. *D* was calculated for a 420-480 by region encompassing the MADS box, the I region, and K box (MIK region), and using only the 1st and 2nd codon positions to minimize the saturation effect of the mostly synonymous substitutions at the 3d codon position between distantly related sequences. We applied this evolutionary distance to calibrate nucleotide substitution rates (λ) within the AG subfamily, and to estimate time of divergence (T) between the PTAGI and PTAG2 genes. Nucleotide substitution rate was inferred from the linear relationship $D = 2\lambda T$ using estimated distance *(D)* and assumed time of divergence (T) between rice and maize (Crepet and Feldman, 1991), monocots and dicots (Wolfe *et al.,* 1989), and gymnosperms and angiosperms (Savard *et al.,* 1994). For better calibration we also used a linear regression analysis and SYSTAT v.8.0 (1998) to estimate λ and its variance (Snedecor and Cochran, 1989; see also Hillis *et al.,* 1996).

Expression analyses

Total RNA samples (20 μg) were electrophoresed in formaldehyde agarose gels and blotted onto Hybond (Amersham). Short PCR fragments corresponding to portions of the 3'-untranslated regions (UTRs) were synthesized for use as probes to eliminate any cross-reaction between PTAGI and PTAG2. A 115 by PTAGI fragment was obtained using the primers 5'-CCTGGGTTTCCATTGAGC3'and 5'-GGATAGTTAATACATAGAGGAAGAG-3', and a 118 by PTAG2 fragment was generated with the primers 5'-GTACCTACTATTTCACTGAGCG-3' and 5'-AAAGCAATACATGGAGGAAGAG-3'. Fragments were subcloned into pBluescript 11 KS (Stratagene), linearized with EcoRI, and ^{32}P-labeled antisense RNA probes were synthesized using T7 polymerase and Ambion's Maxiscript kit. Blots were hybridized at 65 °C in 50% formamide, 5 x SSC, 5 x Denhardt's solution, 0.5% SDS, 100 μg/ml denatured salmon sperm DNA and final washes were at 65 °C in 0.1 x SSC, 0.1 % SDS. The specificity of the probes was verified under identical conditions with *in vitro* transcribed PTAG1 and PTAG2 RNAs (data not shown). To correct for variation in RNA loading onto gels, blots were stripped and re-probed with a *SstI-XbaI* fragment of the *P. deltoides* 18S rDNA isolated from pPD5 (D'Ovidio *et al.,* 1991). Blots were exposed to Kodak BioMax Film, and quantitation of RNA bands was done using a Molecular Dynamics phosphorimager and ImageQuaNT software version 4.2.

PTAGI transcripts from floral and vegetative tissues were analyzed using the Superscript One-Step re

verse transcriptase-PCR (RT-PCR) system (Life Technologies). cDNAs were synthesized with 200 ng of total RNA according to the manufacturer's protocol. Primers 5'-AAGATCCTCACTTTCTCTACAC-3' and 5'-CCCAGACAAATATGATTTAC-3' were used to synthesize and amplify a product containing all intron-exon junctions.

In situ *hybridization studies*

Tissue samples were fixed, embedded, sectioned, and hybridized following a modified version of the procedure described by Kelly *et al.* (1995). Antisense and sense ^{35}S probes were generated from a plasmid consisting of a 800 by *PstI-XhoI* fragment from the PTAGI cDNA clone inserted in pBluescript II KS (Stratagene), and were not alkaline-hydrolyzed. These probes lacked the MADS box and included part of the I region and the remaining 3' sequences of the cDNA. Short, gene-specific probes derived from the PTAGI and PTAG2 3'-UTRs (described above for northern blots) were also used (data not shown). Probes were applied to 10 ltm sections at a final concentration of 2 x 107 cpm/ml (ca. 30-40 ng/ml). Slides were photographed with a Contax camera mounted on a Zeiss Axioskop microscope equipped with a dark-field illuminator. Photographic slides were scanned, digitized, and adjusted for contrast, brightness, and color balance with Adobe PhotoShop version 5.0.

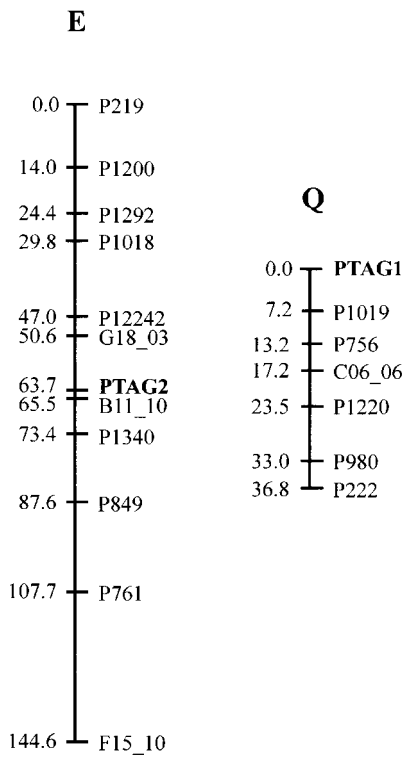

Figure 1. Map positions of PTAGI and PTAG2. Using SSR markers, the loci were placed on a linkage map derived from segregation of DNA markers in an interspecific hybrid poplar pedigree (Bradshaw et al., 1994). Cumulative distances (cM) are shown on the left of each linkage group.

Results

Isolation and mapping of PTAGI *and* PTAG2 *cDNAs and genes*

A genomic library prepared with DNA isolated from a single male tree was screened with a fragment of the AG cDNA that lacked most of the MADS-box region. Based on restriction mapping, two classes of genomic clones were identified. Sequencing of the genomic clones revealed that the two classes corresponded to two closely related genes. With a fragment of the AG cDNA 3' to the MADS box as a probe, a near-full-length PTAGI cDNA clone was isolated from an immature female floral library. Based on PTAG2 genomic sequence, a partial PTAG2 cDNA clone was isolated from a mature female floral library by PCR, and this sequence was then used to isolate a near-full-length cDNA clone.

The cDNA sequences are identical to the corresponding genomic sequences except for two nucleotide changes in the 5'-UTR and one synonymous base transition in the I region of PTAG2. PTAGI encodes a 241 amino acid protein, while PTAG2 encodes a slightly shorter (238 amino acid) protein, apparently due to a frameshift caused by a 4 by deletion near the C-terminal end of the coding sequence. A total of 11485 by of the PTAGI gene locus contained in a single genomic clone, and 10,007 by of the PTAG2 locus contained in two overlapping clones, were sequenced and are available as GenBank accession numbers AF052570 and AF052571, respectively. Both PTAGI and PTAG2 contain 8 introns at conserved positions; all introns have canonical donor (GT) and acceptor (AG) sites.

The PTAGI and PTAG2 cDNAs show high sequence identity in both the coding and untranslated regions (Table 1). Several of the coding domains share greater nucleotide identity than amino acid identity. Ratios of non-synonymous substitution rate to synonymous substitution rate revealed that the MADS box is under the greatest level of constraint, while the C-

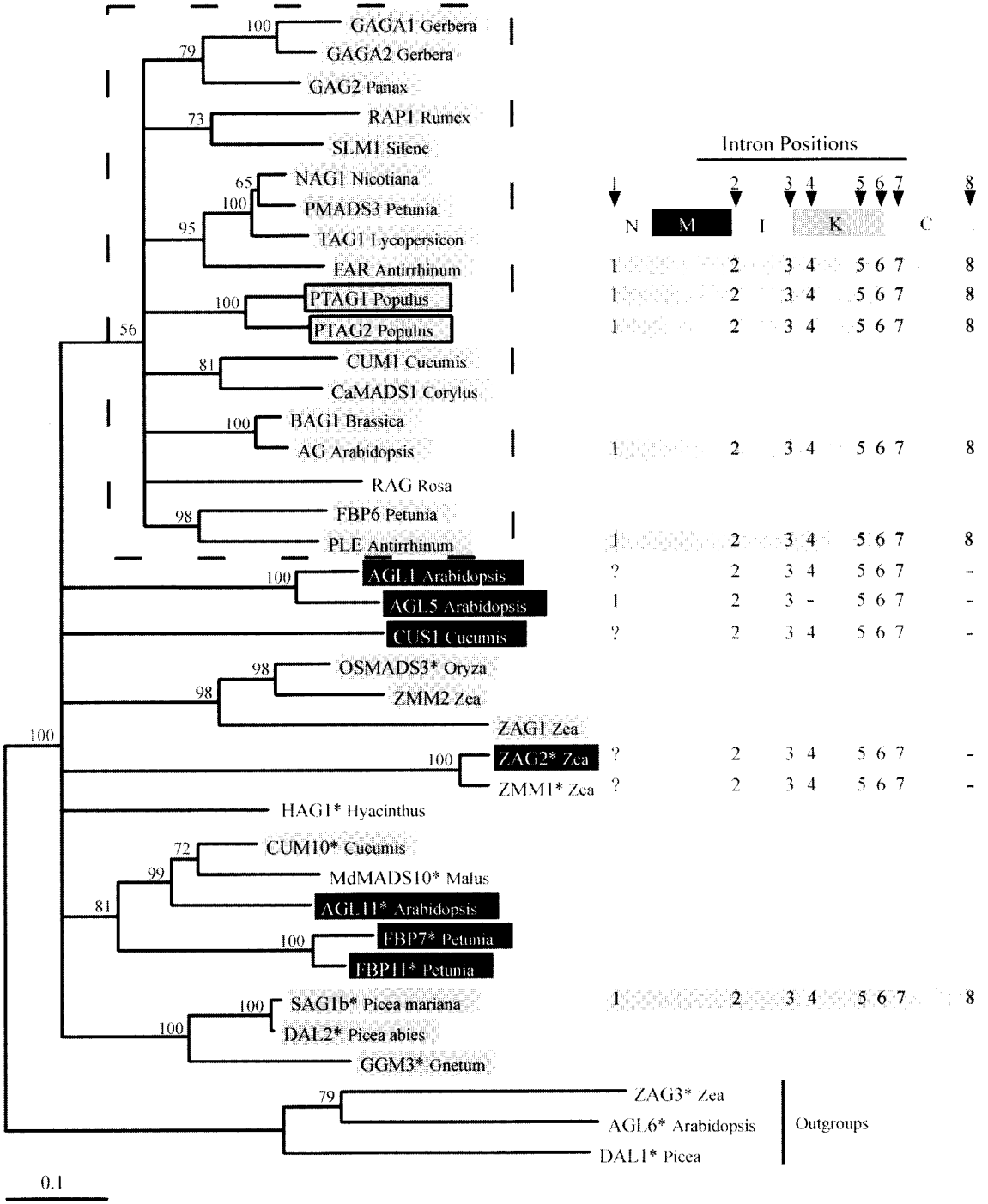

Figure 2. Phylogeny and gene structure of the AG subfamily based on neighbor-joining analysis. Predicted amino acid sequences of the entire coding region were used on sequences obtained from GenBank. Bootstrap values are shown at nodes; nodes with values less than 50% are collapsed. The dotted line encases the genes referred to as the dicot C group (see text). Only genus names are indicated except where sequences from more than one species within a genus were used. Asterisks indicate proteins that lack an N-terminal domain. Genes expressed in both female and male tissues are in gray boxes, while those expressed only in female tissues are in black boxes. Expression results for most of the genes are summarized in Irish and Kramer (1998). Additional references are: Kater *et al.,* 1998 *(CUM/, /*0); Filipecki *et al.,* 1997 (CUSl); Kim *et al.,* 1998 (CAG2); Rigola *et al.,* 1998 *(CaMADSl); Yu et al.,* 1999 (GAGA/, 2); Winter *et al.,* 1999 *(CCM3);* Davies *et al.,* 1999 *(FAR).* For the gene structures, N/M/I/K/C refer to the domains described in the text, - indicates intron absence and ? indicates that intron presence/absence is unclear. Ambiguity is due to the usual location of intron 1 in the 5'-UTR, and that some of the isolated cDNA clones did not extend tar enough into the 5'-UTR to conclusively verify its absence. Gray shading indicates the presence of all eight introns. See text for discussion concerning the positions of intron 1 and intron 8.

7hhle I. Comparison of PIAGI and PTAG2 by domain.

Region (length)	Nucleotide identity (%)	Amino acid identity (%)[a]	K_a/K_S[b]
N-terminal (48 bp)	87.5	81.3 (81.3)	0.295
MADS-box (168 bp)	95.8	98.2 (100)	0.050
I region (99 bp)	89.9	87.9 (90.9)	0.176
K-box (201 bp)	92.0	92.5 (98.5)	0.118
C-terminal (207/198 bp)	88.9	78.8 (87.9)	0.344
Entire coding (723/714 bp)	91.5	88.7 (93.7)	0.196
5′-UTR[c] (195/200 bp)	88.1		
3′-UTR (284/299 bp)	87.2		

[a]Percent amino acid similarity is shown in parenthesis.
[b]Maximum likelihood estimates of non-synonymous substitution rate (K_a)/synonymous substitution rate (K_S) (Muse, 1996). The corresponding estimates of Nei and Gojobori (1986) were not significantly different.
[c]Start corresponds to the first nucleotide of the *PTAG1* cDNA; the first 102 bp of the *PTAG2* 5′-UTR comes from the genomic sequence.

terminal region is evolving at the fastest rate, with a ratio seven times greater than the MADS-box ratio. Comparison of the genomic sequences further demonstrates the striking similarity between the two genes. Homologous introns share significant sequence identity (71-82%), with most of the differences due to insertions/deletions. The 5'-flanking regions also exhibit strong evidence of homology, particularly within 500 by of the start of the cDNAs (79% sequence identity). Sequence similarity decreases markedly in more distal upstream sequences and appears largely insignificant, although short (ca. 10 bp) motifs that possibly have some regulatory function are shared by the genes in this region.

Intron 2 is large (ca. 3-5 kb) in *PTAGI, PTAG2, AG, PLE,* and *FARINELLI (FAR),* an *Antirrhinum* gene closely related to *PLE* (Yanofsky *et al.,* 1990; Bradley *et al.,* 1993; Davies *et al.,* 1999). Studies by Sieburth and Meyerowitz (1997) have shown that sequences in this intron are essential for normal *AG* expression. Although differing by almost 1 kb in length, the large second introns of *PTAGI* and *PTAG2* contain several stretches of around 100 to 400 by that are about 85% identical. Sequence comparisons with the *AG* intron 2 (GenBank accession number AL021711) revealed numerous short stretches (ca. 10 bp) of sequence identity and some longer stretches of 30-200 by with identity ranging from about 62% to 85%, suggesting that these regions may be of regulatory significance. The second intron of both *PTAGI* and *PTAG2* contains four copies of the core consensus sequence (CCANTG) of the *LEAFY* binding site identified in the *APETALAI* promoter and the *AG* intron 2 (Busch *et al.,* 1999). None of these sites, however, are located in the larger regions of sequence similarity with *AG,* and only one site is located within a region clearly shared by the *PTAGI* and *PTAG2* introns.

Southern analysis using gene-specific probes demonstrated that single copies of *PTAGI* and *PTAG2* are present in both female and male genomes (data not shown). When Southern blots were probed at high or low stringency (65 °C vs. 42 °C washes) with a fragment of the *PTAGI* cDNA that included the I region, K-box and C-terminal region, only *PTAGI* and *PTAG2* gene fragments were detected, indicating that additional closely related *AG* homologues are not present in the poplar genome (data not shown). Analyses of SSR markers derived from the *PTAGI* and *PTAG2* genomic sequences in a mapping pedigree (Bradshaw *et al.,* 1994) clearly localized (LOD scores > 10) the genes in different linkage groups (Figure 1). The genetic map generated using this pedigree consists of 19 linkage groups, which is equal to the haploid number of chromosomes present in the poplar genome.

Relationships among members of the AG subfamily

Previous phylogenetic analyses of the MIK region revealed that most plant MADS-box genes are organized into monophyletic clades; these clusters generally correspond to groups of genes that share related functions (e.g., Purugganan, 1997). In order to study the relationships within the AG subfamily, we performed phylogenetic analyses using most members of this group that were available in GenBank (partial

sequences and some orthologous sequences from different species in the same genus were not included). By limiting analysis to this group, entire amino acid sequences could be aligned reasonably well. Because entire protein sequences cannot be unambiguously aligned with sequences from other subfamilies, we used three sequences that are orthologous to each other as outgroups: a dicot, a monocot and a gymnosperm member of the API/AGL9 clade (Purugganan, 1997). Which clade is closest to the AG clade is unresolved. Similar results (data not shown) were obtained using fern MADS domain sequences as outgroups, which do not group with the known subfamilies from seed plants (Munster *et al.,* 1997).

Figure 2 depicts a phylogenetic tree derived from the neighbor-joining procedure. Phylogenetic trees estimated by maximum parsimony and a weighted genetic distance matrix produced very similar topologies (data not shown). In all trees, PTAGI and PTAG2 cluster with all but one of the dicot genes expressed in both female and male reproductive primordia (boxed area of Figure 2). Because this group includes all known C-function organ identity genes in dicots and the remaining genes exhibit AG-like floral expression patterns, we refer to it as the dicot C group. The only reported bisexually expressed dicot gene not in the dicot C group is CUMIO (Kater *et al.,* 1998), which forms a well-supported group (81 % bootstrap support) that includes three dicot proteins expressed in ovules (Angenent *et al.,* 1995; Colombo *et al.,* 1995; Rounsley *et al.,* 1995). The female-specific *Arabidopsis* genes AGL1 and AGL5 (Ma *et al.,* 1991) and Cucumis gene CUSI (Filipecki *et al.,* 1997) are most closely related to the dicot C group in all trees; however, bootstrap support is less than 50%. Although only six monocot genes are represented, their separation into groups appears to be correlated with their expression pattern. The conifer and gnetophyte genes *DAL2, S*AGIb and GGM3 (Tandre *et al.,* 1995; Rutledge *et al.,* 1998; Winter *et al.,* 1999) form a group separate (100% bootstrap) from all other members of the AG subfamily in all trees.

Complete or nearly complete gene structures have been reported for eleven members of the AG subfamily (Figure 2). Although intron number varies, with two exceptions intron position is highly conserved. While intron 1 is within the predicted translation of AG, it is located in the 5'-UTR of *PLE, FAR, AGL5*, *SAG1b*, *PTAG1* and *PTAG2* (Yanofsky *et al.,* 1990; Bradley *et al.,* 1993; Savidge *et al.,* 1995; Rutledge *et al.,* 1998; Davies *et al.,* 1999). Unlike the other genes, SAG] b does not encode an N-terminal domain. In contrast, AG encodes an N-terminal domain that is longer than the corresponding domain of most other AG subfamily members (Figure 2) and is also unusual in that it does not contain an initiating methionine (Yanofsky *et al.,* 1990). Although the location of intron 1 in a highly variable region causes its location to vary somewhat, its position is essentially homologous in all seven genes.

The position of intron 8 in PTAG2 is unique: it is located in the 3'-UTR rather than in the last amino acid codon. However, alignment of the PTAGI and PTAG2 cDNAs (data not shown) clearly shows that this is due to a frameshift introducing a slightly premature stop codon rather than a shift in intron position. All eight introns are present in the five genes belonging to the dicot C group and the conifer gene SAGI b. In contrast, the three genes most closely related to the dicot C group have somewhat different structures. CUSI, AGL1 and AGL5 lack intron 8 and AGL5 also lacks intron 4 (Ma *et al.,* 1991; Savidge *et al.,* 1995; Filipecki *et al.,* 1997).

The PTAGI and PTAG2 proteins are 76% and 72% identical to AG respectively and share a similar level of identity with other members of the dicot C group. Figure 3 is an alignment of PTAGI and PTAG2 with other members of the dicot C group. Even in the most variable C-terminal domain blocks of amino acids are conserved among all members of this group, particularly near the N-terminal end of this domain. Riechmann and Meyerowitz (1998) have noted that the region predicted to form coiled-coil structures is not limited to the K-box domain, but extends into the C-terminal domain. In addition to the characteristic conservation of hydrophobic residues in this region, charged residues are also highly conserved.

Tajima and Nei's (1984) genetic distance (D) based on the 1st and 2nd codon positions of the MIK region between PTAGI and PTAG2 was low (0.0363 ± 0.0110). The observed frequencies of the four nucleotides (A, T, C, G) deviated substantially from equality for the MIK region in PTAGI (30.8%, 24.6%,, 19.2%, 25.4%, respectively) and PTAG2 (31.2%, 25.2%, 17.7%, 25.9%), as well as in all other AG subfamily sequences on average (32.9 ± 0.4%, 21.7 ± 0.5%, 20.5 ± 0.5%, 24.8 ± 0.4%). The transition/transversion ratio was close to unity (1.063 for PTAGI vs. PTAG2, and 1.075 ± 0.009 for all other AG group sequences on average), justifying the use of Tajima and Nei's (1984) genetic distance. Based on this distance and assumed times of divergence we

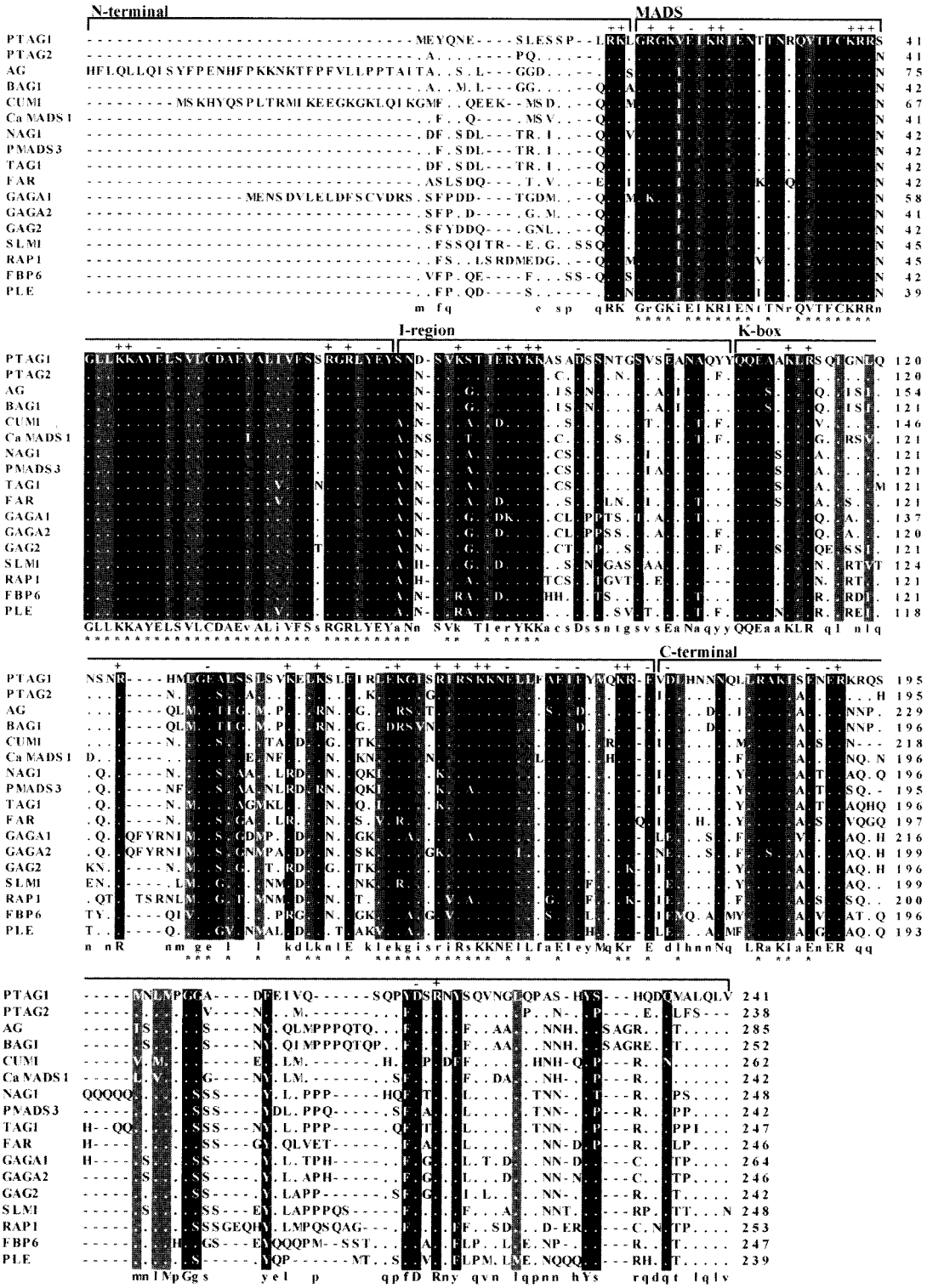

Figure 3. Alignment of PTAG I with other members of the dicot C-group of the AG subfamily. Amino acids identical to PTAGI residues are indicated by dots; dashes indicate gaps. Residues that are identical or similar in all sequences are shaded black or gray (hydrophobic amino acids). Groups of amino acids considered to be similar were L, I, V, M (hydrophobic), D, E (acidic), N, Q (amide), F, Y, W (aromatic), H, K, R (hydrophilic, basic), and P, A, G, S, T (small, neutral or weakly hydrophobic). Positively (+) and negatively (-) charged residues conserved in all sequences are indicated above the alignment. A consensus sequence is displayed at the bottom with residues identical in all listed sequences shown in upper case, and residues present in more than 50% of the listed sequences shown in lower case. Asterisks denote residues conserved among all member of the AG subfamily shown in Figure 2.

Table 2. Estimation of divergence time for *PTAG1* and *PTAG2*[a].

Pairwise comparison	N[b]	T, MYA	D, mean ± SE	λ, mean ± SE ($\times 10^{-10}$)		
				$\lambda = D/2T$	linear regression of D on T	Purugganan (1997)
1. Rice vs. maize (OsMADS3 vs. *ZAG1*, *ZAG2*, *ZMM1* and *ZMM2*)	4	65[c]	0.1667 ± 0.0252	13.89 ± 3.56	12.82 ± 2.85	12.12 ± 1.91 (1)[b]
2. Monocot vs. dicot (*OsMADS3*, *ZAG1*, *ZAG2*, *ZMM1*, *ZMM2* and *HAG1* vs. 26 dicot sequences)	156	200[d]	0.2303 ± 0.0314	5.76 ± 0.08	5.76 ± 0.08	6.35 ± 0.76 (26)
3. Gymnosperm vs. angiosperm (*SAG1b*, *DAL2* and *GGM3* vs. 32 angiosperm sequences)	96	285[e]	0.2614 ± 0.0344	4.59 ± 0.06	4.59 ± 0.06	4.56 ± 0.54 (18)
1 + 2	160	65/200	irrelevant	5.96 ± 0.15	5.78 ± 0.09	
1 + 2 + 3	256	65/200 285	irrelevant	5.43 ± 0.10	5.12 ± 0.07	7.68 ± 3.95 (45) 5.60 ± 1.02[f] (56)

[a]Nucleotide substitution rates (λ) per nucleotide site per year and its standard error (SE) are given for *AG* orthologues based on the MIK region. Divergence estimates were calibrated using estimated Tajima and Nei's (1984) measure of nucleotide divergence (genetic distance D) at the 1st and 2nd codon positions and assumed times of divergence (T) from the literature.
[b]Number of pairs.
[c]Crepet and Feldman, 1991.
[d]Wolfe *et al.*, 1989.
[e]Savard *et al.*, 1994.
[f]Estimated for MIK-region in AG subfamily sequences, and for AG and AP1/AGL9 groups together, respectively (Purugganan, 1997, Table 2)

computed nucleotide substitution rates (λ) per nucleotide site per year and its standard error (SE) for AG orthologues (Table 2). Our estimates are similar to those obtained in an earlier study (Purugganan, 1997), but we believe that our estimates of λ and its standard error are more accurate due to the larger number of AG subfamily sequences used in our pairwise comparison (256 vs. 45, respectively, Table 2). Using λ = 5.12 x 10-10 and the variance V_{T-D} estimated in the regression analysis we calculated the divergence time and its standard deviation for PTAGI and PTAG2 genes to be 35.5 ± 18.3 MYA. If we exclude the gymnosperm sequences the estimate is λ = 5.78 x 10-10, giving a divergence time of 31.4 ± 15.7 MYA for PTAGI and PTAG2.

PTAG 1 *and* PTAG2 *expression*

Gene-specific probes corresponding to portions of the PTAG1 and PTAG2 3'-UTRs were used to probe gel blots of RNA isolated from developing and mature female and male inflorescences and vegetative tissues (Figure 4). Both probes gave a very similar pattern of expression; transcripts were detected in immature and mature flowers from both male and female trees. Surprisingly, weaker bands were also detected in all vegetative tissues tested. Vegetative buds, stems and leaves from the current year's growth contained the lowest levels of both transcripts. Compared to these vegetative tissues, transcript levels were approximately two-fold greater in vegetative buds that had undergone dormancy. For both PTAG1 and PTAG2, transcripts from vegetative tissues were shorter than the floral transcripts, and this size difference was consistently observed across separate RNA gel-blot experiments. To investigate whether this size difference is due to alternate splicing, PTAGI transcripts were analyzed via RT-PCR with gene-specific primers encompassing all intron-exon junctions. Products of the expected size (940 bp) were synthesized from both vegetative and floral tissues (data not shown), demonstrating that normal splicing occurs in vegetative tissues.

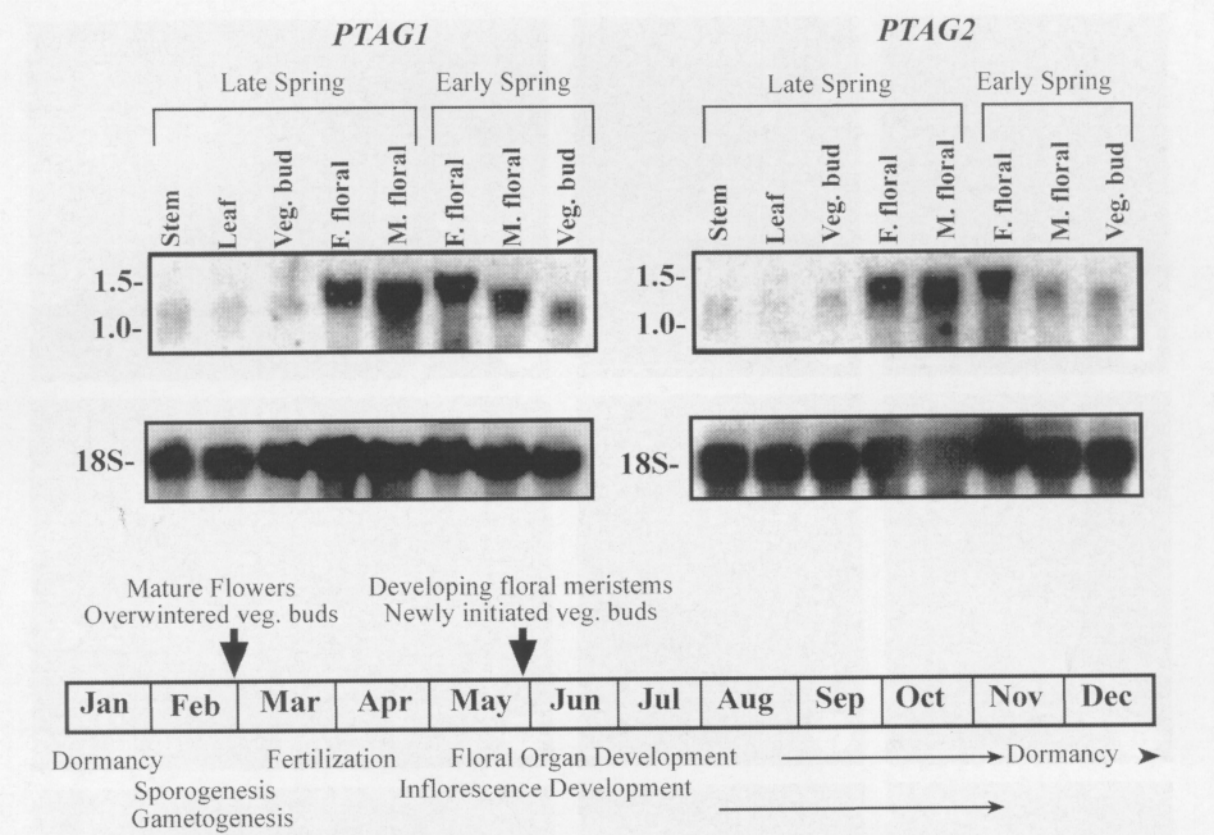

Figure 4. Expression of PTAGI and PTAG2 in floral and vegetative tissues. Each lane contains 20 pg of total RNA. Molecular weight markers (kb) are indicated at the left. RNA was prepared from female (F) and male (M) inflorescences at both immature (late spring) and mature (early spring) developmental stages. Blots were stripped and re-probed with an 18S rDNA fragment. A typical annual flowering cycle for P trichocarpa in the vicinity of Corvallis, OR, is depicted at the bottom, with arrows indicating when tissues were collected.

Poplar catkins differentiate acropetally; floral meristems arise in the axils of bracts (Figure 5a). In *P. trichocarpa* male flowers, 30 to 50 stamen primordia arise centrifugally. Three to four carpel primordia form in female flowers; thirty to fifty anatropous ovules develop from a parietal placenta. The expression pattern of PTAG1/2 in various floral tissues is shown in Figure 5. Sections were hybridized to antisense or sense probes corresponding to a fragment of the PTAG1 cDNA that lacked the MADS box. The antisense probe cross-reacts with PTAG2; however, studies using the very short, gene-specific probes (described above) produced hybridization patterns that were indistinguishable from the patterns produced using the longer probe, though they did produce weaker signals (data not shown).

PTAG expression was first detected in the central whorl of male and female meristems when the perianth cup primordia had barely begun to emerge, but before stamen and carpel primordia were visible (Figure 5a, b, f, h). Expression continued in developing stamens and carpels (Figure 5c, d, j). Transcripts were not detected in the outer whorl where the perianth cup forms at any developmental stage (Figure 5c, f). As the stamen primordia differentiated, PTAG expression appeared to decrease and was confined to the filament, connective, and anther walls; expression levels were near the detection limit in mature male flowers, though a weak signal appeared to be associated with the connective and filament (data not shown). In mature female flowers, PTAG transcripts were evident in the placenta, but were not detected in the stigma (Figure 5j, 1). PTAG expression was seen in the ovule integuments, but not in the nucellus or embryo sac; a band of expression was detected where the funiculus merges with the ovule (Figure 5k, 1).

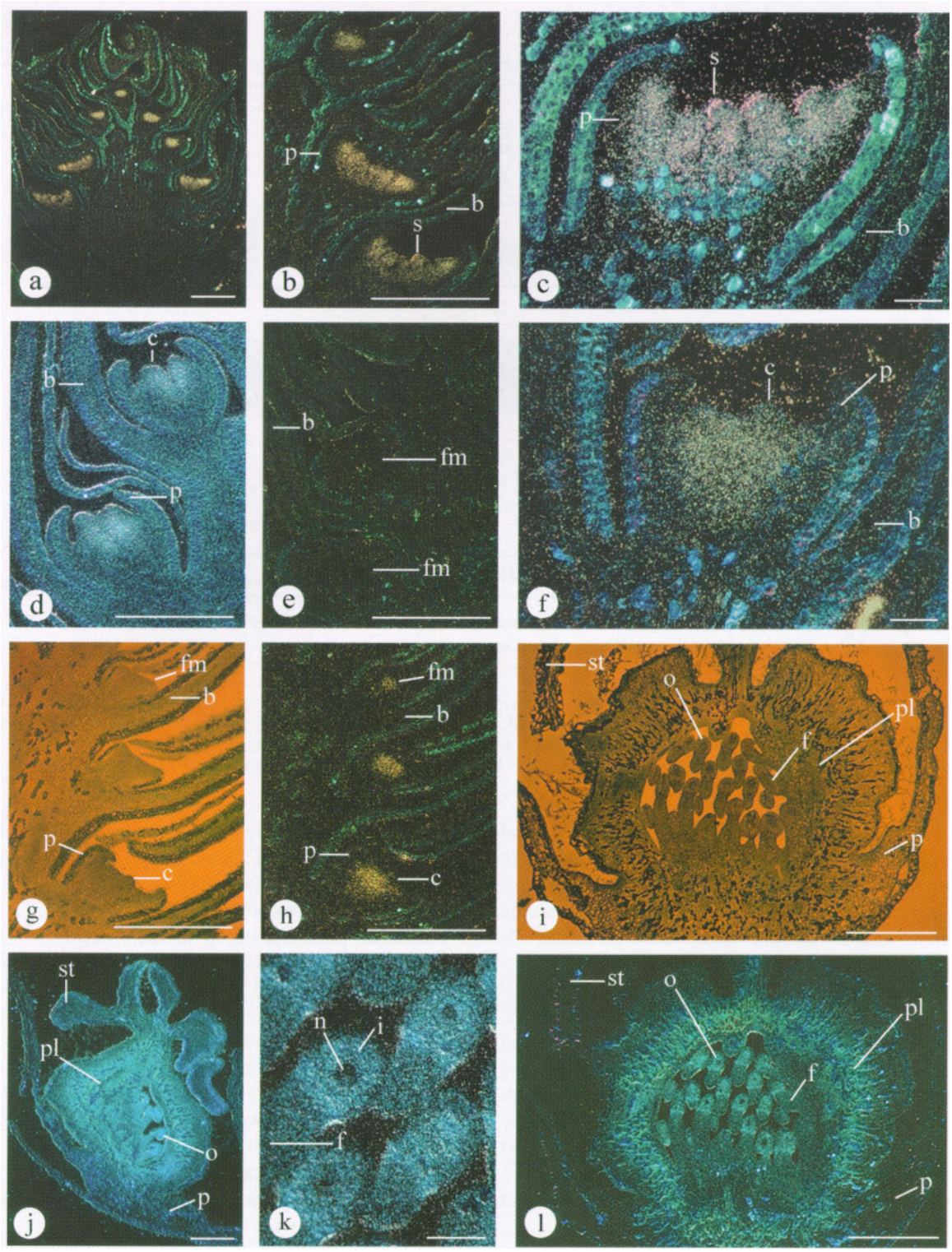

Figure 5. Expression of P7AC: in developing male and female flowers. a. Immature male inflorescence. b. Developing male floral meristems; stamen primordia have not initiated in the top meristem, but arc forming in the lower two. c. Male flower with stamen primordia. d. Female flowers with developing carpets. e. Negative control (sense probe), developing male flowers, f. Female flower with carpet primordia just forming, g. Bright-field photo of developing female floral meristems; carpet primordia have just initiated in the bottom flower. h. Dark-field photo of g. i and 1. Bright-field and dark-field photos of mature female flower, showing ovules. j. Mature female flower. k. Ovules. b, bract; c, carpet; f, funiculus; fm, floral meristem; i, integument; n, nucellus; o, ovule; p, perianth cup; pl, placenta; s, stamen, st, stigma. Bars = 500 µm, except in panels c, f, and k where bars = 100 µm.

Discussion

We have shown that the poplar genome contains two closely related members of the AG clade that appear to be the result of a relatively recent duplication of an AG orthologue. The coding regions of PTAGI and PTAG2 share high nucleotide identity with each other (92%), and intronic and near upstream sequences also exhibit strong homology. Based on substitution rates calculated within the AG clade (Table 2), we estimated the time of the PTAG duplication to be 35.5 ± 18.3 MYA. This suggests that the duplication occurred during evolution of Populus or Salicaceae as the earliest fossil evidence is leaves of Populus section Abaso from 58 MYA (Eckenwalder, 1996). The earliest fossil evidence for the other five Populus sections and for Salix is later, ranging from about 18 to 40 MYA. Interestingly, black cottonwood (sect. Tacamahaca, first fossils found from ca. 30 MYA) also contains two closely related API homologues that have a level of sequence identity and an estimated divergence time that are very similar to the corresponding values for the PTAG gene pair (A. Brunner, unpublished data). It will be of interest to study whether other sections of Populus or Salix also possess these duplications.

It is unclear whether the PTAG duplication is part of a larger genome duplication. There are some indications that Populus is a derived polyploid (Masterson, 1994). All poplar species have the same number of chromosomes (2n = 38), and a significant portion (11 %) of the loci in a genetic map of a *P. trichocarpa* x *P. deltoides* hybrid are duplicated (Bradshaw et al., 1994). The only duplicated restriction fragment length polymorphism (RFLP) marker near to either *PTAG1* or PTAG2 is marker P12242, located 16.7 cM from the *PTAG2* locus (Figure 1). However, its homologous locus is found on a separate linkage group (i.e., neither the same linkage group as *PTAG1* nor *PTAG2*). Therefore, *PTAG1* and *PTAG2* do not appear to be part of a larger duplication event; however, higher-density RFLP maps are needed to conclusively resolve this.

Our phylogenetic analyses using the entire coding region discerned some larger clusters within the AG clade that show a general correspondence with spatial expression pattern (Figure 2). *PTAG1* and *PTAG2* clearly belong to the dicot C group, which includes all but one of the reported bisexually expressed dicot genes. In addition to AG and *PLE,* other members of the dicot C group have also been shown to cause homeotic organ transformations indicative of C function (e.g., Kater et al., 1999), further supporting that this group corresponds to C-function genes. However, two exceptions indicate that evolutionary variances in function and regulation have also occurred within the dicot C group. Not surprisingly, these diversifications are associated with paralogous duplications.

Although ectopic expression of the Antirrhinum gene *FAR* in tobacco resulted in partial C-function organ identity transformations, *FAR* is incapable of specifying reproductive organ identity in *ple* mutants (Davies et al., 1999). Conversely, *PLE* can specify organ identity in far mutants, although pollen development is often disrupted. These observations may be at least partially explained by differences in regulation because *FAR* expression is reduced in *ple* mutants. In petunia, ectopic expression of *FBP6* affected petal size and shape but did not alter organ identity (Kater et al., 1998). In contrast, ectopic expression of *PMADS3* did result in partial C-function organ transformations (Tsuchimoto et al., 1993). Interestingly, the phylogenetic analyses (Figure 2) suggest that the petunia and Antirrhinum paralogues may be the result of the same duplication that occurred during the evolution of the Asteridae. *FAR* forms a strongly supported (95% bootstrap support) group with *PMADS3,* NAGI and TAG1, while *PLE* and *FBP6* form a strongly supported (98% bootstrap support) pair.

Substantial functional or regulatory diversification appears unlikely for the poplar C-function paralogues. The PTAGs appear to be the result of a duplication within the Populus lineage, and their structure and expression are highly similar. The PTAG proteins are more similar (89% identity) than the paralogues from petunia (70%) and Antirrhinum (69%). The different domains of these genes are diverging at rates (Table 1) similar to the rates reported for the AG subfamily (Purugganan et al., 1995), suggesting that both genes are under similar forms of functional selection. Moreover, Southern analyses indicated that the poplar genome does not contain additional related C-function genes. Identification of functional differences, should they exist, are likely to require ectopic expression and gene-specific suppression studies in transgenic poplar.

In addition to the phylogenetic analyses, the floral expression patterns of *PTAG1* and *PTAG2* are also consistent with C-class organ identity function. Their nearly identical expression patterns in both female and male trees do not suggest a diversification in function correlated with tree gender. Like AG, *PLE* and other genes shown to specify C-functions (e.g., Yanofsky *et al.,* 1990; Bradley *et al.,* 1993), *PTAG1* and *PTAG2* are not expressed in the whorl giving rise to perianth

structure, but are expressed in the whorls giving rise to stamen and carpets both before these primordia begin to form and in the developing stamen and carpels. Furthermore, their expression pattern in mature female flowers is similar to the patterns reported for *AG* (Bowman *et* al., 1991): *PTAG1/2* transcripts are detected in the placenta, funiculus, and ovule integuments, but not in the nucellus or embryo sac. In contrast to *AG*, *PTAG* expression was not seen in the stigma.

Intron-exon structure provides insight into evolutionary relationships of the AG subfamily. In a previous analysis, Doyle (1994) proposed that an ancestral gene with 6 exons (numbers 2-7 in Figure 2) provided the most parsimonious explanation for the origin of introns in the plant MADS-box family. This requires both intron loss and intron gain to explain the known gene structures. Within the AG subfamily, complete or nearly complete gene structures for eleven genes have now been reported. The five genes of the dicot C-group, *AG*, *PLE, FAR, PTAG1* and *PTAG2*, all share the same gene structure with 8 introns (Figure 2: Yanofsky *et al.,* 1990; Bradley *et al.,* 1993; Davies *el al.,* 1999). All other plant MADS-box genes for which gene structures have been reported have fewer than 8 introns, and both intron 1 and intron 8 appear to be specific to the *AG* subfamily.

The female-specific dicot genes, *CUS1*, *AGL1* and *AGL5*, group most closely to the dicot C group and all lack intron 8 (Ma *et* al., 1991; Filipecki *et al.,* 1997). The maize genes, ZAG2 and *ZMMI,* which are well separated from the dicot C group (Figure 2), also lack intron 8 (Theissen *et al.,* 1995). In contrast, SAGIb from *Picea* contains all eight introns (Rutledge *et al.,* 1998). The presence of all eight introns in both bisexually expressed angiosperm and gymnosperm genes suggests that the addition of intron 1 and intron 8 may have occurred early in the evolution of the AG subfamily and preceded the addition of the N-terminal domain. However, additional gene structures and AG subfamily sequences are required for a clear understanding of the AG subfamily history. Nonetheless, the conservation of intron structure between the PTAGs and other C-class genes provides further evidence of their orthology to genes specifying C-function.

Unlike AG, *PLE and* most C-class orthologues, we detected PTAG transcripts in vegetative tissues, though at lower levels than in floral tissues (Figure 4). Studies in *Arabidopsis indicated* that the level of AG expression is important for determining organ identity (e.g., Mizukami and Ma, 1995). Thus, the level of PTAG vegetative expression in poplar may be insuf

ficient for function. Alternatively, vegetative tissues in these trees may simply be incompetent to respond to PTAG expression. Although constitutive expression of AG in transgenic *Arabidopsis causes* early flowering, these plants exhibit a degree of vegetative growth, suggesting that other factors are necessary to make cells competent to respond to AG (Mizukami and Ma, 1997). Finally, the short size of the vegetative PTAG transcripts indicates that they are differentially transcribed or processed. RT-PCR experiments indicated that this size difference was not due to alternate splicing (data not shown); however, differences in the site of transcription initiation or poly(A) addition could account for the small size difference (150-200 bp) and affect translational efficiency, mRNA stability, and thus PTAG function.

We observed that PTAG112 expression levels in vegetative buds change with season and/or developmental state. Vegetative buds that had over-wintered were collected in early spring, the time when inflorescences initiated in the previous year were nearing anthesis. Expression levels in these buds were at least two times the levels observed in newly initiated vegetative buds collected in late spring. The buds that had undergone dormancy would soon flush, giving rise to shoots with inflorescence buds as well as vegetative buds. Although it is unlikely that inflorescence meristems had initiated in the over-wintered buds at the time of collection, the time that the progenitor cells are committed to form inflorescence or vegetative meristems is unknown. Certainly, floral meristems were not yet present. In contrast, the newly initiated vegetative buds would not give rise to shoots with inflorescences for at least one year; they would either remain as buds through the next dormant season or give rise to sylleptic branches during the current growing season.

Several genes are known to regulate AG expression. For example, *LEAFY* is an important activator of AG in floral meristems (Busch *et al.,* 1999), while *CURLY LEAF is* necessary to prevent AG expression in leaves during the vegetative phase (Goodrich *et al.,* 1997). Given the complexity of AG regulation during both vegetative and reproductive phases, it would not be surprising to find expression differences among disparate species. PTAG vegetative expression may be the result of less stringent repression controls in poplars, or a result of derepression that occurs as a tree approaches reproductive competence and flowering. In addition to reproductive competence acquired in the juvenile-to-adult transition, adult poplars only initiate inflorescences at a particular time during the annual

growing cycle. Higher *PTAG1/2* expression levels in the post-dormancy vegetative buds of mature trees may be correlated with a localized or seasonal derepression associated with floral induction, as these buds could soon give rise to shoots with inflorescences. Further analysis of vegetative expression at the cellular level, at various times during a year, by comparison of juvenile and mature trees, and by studies of ectopic expression in transgenic trees would help to determine if vegetative expression of *PTAG 1/2* has any functional significance, and if changes in the level of vegetative expression can serve as a molecular marker for maturation or floral induction.

Acknowledgements

We thank Toby Bradshaw for providing the F2 progeny DNA and for placing the PTAG loci on the poplar linkage snap. We also thank Victoria Hawkins, Sheila Seguin and Shuping Cheng for technical assistance; Jace Carson and Virginia Gewin for help in preparing the figures; and Elliot Meyerowitz for providing the AGAMOUS cDNA clone. We acknowledge several sources for research support, including: industrial members of the Tree Genetic Engineering Research Cooperative at Oregon State University (Alberta Pacific, Boise Cascade, Electric Power Research Institute, Fort James, Georgia Pacific/The Timber Company, International Paper, MacMillan Bloedel/Pacifica Poplar, Monsanto, Potlatch, Shell, Union Catnp, Westvaco, and Weyerhaeuser), the U.S. Department of Energy Biofuels Feedstock Program, and the Conrad Wessela Scholarship Fund.

References

Ainsworth. C., Crossley, S., Buchanan-Wollaston, V., Thangavelu, M. and Parker, J., 1995. Male and female flowers of the dioecious plant sorrel show different patterns of MADS-box gene expression. Plant Cell 7: 1583-1598.

Alslrom-Rapaport, C., Lascoux, M. and Gullberg, U., 1997. Sex determination and sex ratio in the dioecious shrub. Salix riminalis L. Theor. Appl. *Genet. 94*: 493-497.

Alstrom-Rapaport, C.. Lascoux. M., Wang, Y.C.. Roberts, G. and Tuskan, G.A., 1998. Identification of a RAPD marker linked to sex determination in the basket willow (Sali.c rimiucdis L.). J. Hered. 89: 44-49.

Angenent, G.C., Franken, J., Busscher, M., van Dijken. A., van Went, JJ.L., Dons, H.J.M. and van Tunen. A.J., 1995. A novel class of MADS box genes is involved in ovule development in petunia. Plant Cell 7: 1569-1582.

Boes, T.K. and Strauss, S.H., 1994. Floral phenology and morphology of black cottonwood, Populus trichocarpa (Salicaceae). Am. J. Bot. 81:5 62-567.

Bowman, JL, Drews, G.N. and Meyerowitz, E.M., 1991. Expression of the Arabidopsis floral homeotic *gene AGAMOUS* is restricted to specific cell types late in flower development. Plant Cell 3: 749-758.

Bradley, D., Carpenter, R., Sommer, H., Hartley, N. and Coen, E., 1993. Complementary floral homeotic phenotypes result from opposite orientations of a transposon at the Plena locus of Antirrhinum. Cell 72: 85-95.

Bradshaw. H.D. Jr.. 1998. Case history in genetics of long-lived plants: molecular approaches to domestication of a fast-growing forest tree: Poptdus. In: A.H. Paterson (Ed.) Molecular Dissection of Complex Traits, CRC Press, Boca Raton, FL, pp. 219228.

Bradshaw, H.D. Jr., Villar, M., Watson. B.D., Otto, K.G., Stewart, S. and *Stettler, R*.F., 1994. Molecular genetics of growth and development in Populus III. A genetic linkage map of a hybrid poplar composed of RFLP, STS, and RAPD markers. Theor. Appl. Genet. 89: 167-178.

Busch, M.A., Bomblies, K. and Weigel, D., 1999. Activation of a floral homeotic *gene in* Arabidopsis. Science 285: 585-587.

Colombo, L., Franken, J., Koetjc, E., van Went, J., Dons, H.J.M., Angenent, G.C. and van Tunen, A.J., 1995. The petunia MADS box gene *FBPII* determines ovule identity. Plant Cell 7: 18591868.

Crepet, W.L. and Feldman, G.D., 1991. The earliest remains of grasses in the fossil record. Am. J. Boi. 78: 1010-1014.

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

D'Ovidio. R., Mugnozza, G.S. and Tansarella, O.A., 1991. rDNA cloning and rapid hybrid identification in the Populus spp. (Salicaceae). Plant Syst. Evol. 177: 165-174.

Doyle, J.J., 1994. Evolution of a plant homeotic multigene family: toward connecting molecular systematics and molecular developmental *genetics. Syst*. Biol. 43: 307-328.

Eckenwalder, J.E., 1996. Systematics and evolution of Populus. In: R.F. Stealer, H.D. Bradshaw, Jr., P.E. Heilman and T.M. Hinckley (Eds.) Biology of Populus and its Implications for Management and Conservation, Part 1, Chapter I , NRC Research Press, National Research Council of Canada, Ottawa, pp. 7-32.

Filipecki, M.K., Sommer, H. and Malepszy, S., 1997. The MADSbox gene CUSI is expressed during cucumber somatic embryoQenesis. Plant Sci. 125: 63-74.

Goodrich, J., Puangsomlee, P., Martin, M., Long, D., Meyerowitz, E.M. and Coupland, G., 1997. A polycomb-group gene regulates homeotic gene expression in Arabidopsis. Nature 386: 44-51.

Hillis, D.M., Mablc, B.K. and Moritz, C., 1996. Applications of Molecular Systematics. In: D.M. Hillis, C. Moritz, and B.K. Mable (Eds.) Molecular Systematics, 2nd ed., Sinauer Associates, Sunderland. MA, pp. 515-_543.

Hughes. D.W. and Galau, G., 1988. Preparation of RNA from cotton leaves and pollen. Plant Mol. Biol. Rep. 6: 253-257.

Irish, V.F. and Kramer, E.M., 1998. Genetic and molecular analysis of angiosperm flower development. In: J.A. Callow (Ed.) Advances in Botanical Research. Vol. 28, Academic Press, San Diego, CA, pp. 197-230.

Kater, M.M., Colombo. L., Franken, J.. Busscher. M.. Masiero, S., van Lookeren Campagne, M.M. and Angenent. G.C., 1998. Multiple AGAMOUS homologs from cucumber and petunia dif

fer in their ability to induce reproductive organ fate. Plant Cell 10:171-182.

Kaul, R.B., 1995. Reproductive structure and organogenesis in a cottonwood, Populus *deltoides* (Salicaceae). Int. J. Plant Sci. 156: 172-180.

Kelly, AJ., Bonnlander. M. and Meeks-Wagner, D.R.. 1995. *NFL,* the tobacco homolog of *FLORICAULA and LEAFY,* is transcriptionally expressed in both vegetative and floral meristems. Plant Cell 7: 225-234.

Kim. Y-S., Lee, H.-S., Lee, M.-H., Yoo, O.-J. and Liu, J.-R., 1998. A MADS box gene homologous to AG is expressed in seedlings as well as in flowers of ginseng. Plant Cell Physiol. 39: 836-845.

Kumar, S, Tamura, K. and Nei, M., 1993. MEGA: Molecular Evolutionary Genetic Analysis, version 1.0. Pennsylvania State University, University Park, PA

Lincoln, S., Daly, M. and E. Lander, 1992. Constructing genetic maps with MAPMAKER/EXP 3.0. Whitehead Institute Technical Report, 3rd ed.

Ma, H., Yanofsky, M.F. and Meyerowitz, E.M., 1991. *AGL1-AGL6,* an *Arabidopsis* gene family with similarity to floral homeotic and transcription factor genes. Genes Dev. 5: 484-495.

Masterson, J., 1994. Stomata] size in fossil plants: evidence for polyploidy in majority of angiosperms. Science 264: 421-424.

McLetchie, D.N. and G.A. Tuskan, 1994. Gender determination in Populus. Norw. J Agric. Sci. Suppl. 18: 57-66

Mena, M., Ambrose, B.A., Meely, R.B., Briggs, S.P., Yanofsky, M.F. and Schmidt, R.J., 1996. Diversification of C-function activity in maize flower development. Science 274: 1537-1540.

Mizukami, Y. and Ma, H., 1995. Separation ofAG function in floral meristem determinacy from that in reproductive organ identity by expressing antisense AG RNA. Plant Mol. Biol. 28: 767-784.

Mizukami, Y. and Ma, H., 1997. Determination of *Arabidopsis* floral meristem identity by AGAMOUS. Plant Cell 9: 393-408.

Munster, T., Pahnke, J., Di Rosa, A., Kim, J.T., Martin, W., Saedler, H. and Theissen, G., 1997. Floral homeotic genes were recruited from homologous MADS-box genes preexisting in the common ancestor of ferns and seed plants. Proc. Natl. Acad. Sci. USA 94: 2415-2420.

Muse, SX, 1996. Estimating synonymous and nonsynonymous substitution rates. Mol. Biol. Evol. 13: 105-114.

Nei, M. and Gojobori, T., 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. 3: 418-426.

Purugganan, M.D., 1997. The MADS-box floral homeotic gene lineages predate the origin of seed plants: phylogenetic and molecular clock estimates. J. Mol. Evol. 45: 392-396.

Purugganan, M.D., Rounsley, S.D., Schmidt, R.J. and Yanofsky, M.F., 1995. Molecular evolution of flower development: diversification of the plant MADS-box regulatory gene family. Genetics 140: 345-356.

Riechmann, J.L. and Meyerowitz, E.M., 1998. MADS domain proteins in plant development. Biol. Chem. 373: 1079-1 101.

Rigola, D., Pe, M.E., Fabrizio, C., Me, G. and Sari-Gorla, M., 1998. *CaMADSI,* a MADS box gene expressed in the carpel of hazelnut. Plant Mol. Biol. 38: 1147-I 160.

Rounsley, S.D., Ditta, G.S. and Yanofsky, M.F., 1995. Diverse roles for MADS box genes in *Arabidopsis development*. Plant Cell 7: 12_59-1269.

Rutledge, R., Regan, S., Nicolas, O., Fobert, P., Cote, C., Bosnich, W., Kauffeldt, C., Sunohara, G., Seguin, A. and 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.

Saitou, N. and Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4: 406-425

Savard, L., Li, P., Strauss, S.H., Chase, M.W., Michaud, M. and Bousquet, J., 1994. Chloroplast and nuclear gene sequences indicate Late Pennsylvanian time for the last common ancestor of extant seed plants. Proc. Nail. Acad. Sci. USA 91: 5163-5167.

Savidge, B., Rounsley, S.D. and Yanofsky, M.F., 1995. Temporal relationship between the transcription of two *Arabidopsis* MADS box genes and the floral organ identity genes. Plant Cell 7: 721733.

Sieburth, L.E. and Meyerowitz, E.M., 1997. Molecular dissection of the AGAMOUS control region shows that cis elements for spatial regulation are located intragenically. Plant Cell 9: 355-365.

Snedecor, G.W. and Cochran, W.G., 1989. Statistical Methods, 8th ed. Iowa State University Press, Ames, IA.

Soltis, P.S., Soltis, D.E. and Chase, M.W., 1999. Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. Nature 402: 402-403.

SYSTAT, 1998. SYSTAT for Windows: Statistics, Version 8.0. SPSS.

Tajima, F. and Nei, M., 1984. Estimation of evolutionary distance between nucleotide sequences. Mol. Evol. Biol. 1: 269-285.

Tandre, K., Svenson, M., Svenson, M.E. and Engstrom, P., 1998. Conservation of gene structure and activity in the regulations of reproductive organ development of conifers and angiosperms. Plant J. 15: 615-623.

Theissen, G., Strater, T., Fischer, A. and Saedler, H., 1995. Structural characterization, chromosomal localization and phylogenetic evaluation of two pairs of AGAMOUS-like MADS-box genes from maize. Gene 156: 155-166.

Thompson, J.D., Higgins, D.G. and Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl. Acids Res. 22: 46734680.

Tsuchimoto, S., van der Krol, A.R. and Chua, N.-H., 1993. Ectopic expression of *PMADS3* in transgenic petunia phenocopies the petunia blind mutant. Plant Cell 5: 843-853.

Wolfe, K.H., Gouy, M., Yang, Y.W., Sharp, P.M. and Li, W.H., 1989. Date of the monocot-dicot divergence estimated from chloroplast DNA sequence data. Proc. Natl. Acad. Sci. USA 86: 6201-6205.

Winter, K.-U., Becker, A., Munster, T, Kim, J.T., Saedler, H. and Theissen, G., 1999. MADS-box genes reveal that gnetophytes are more closely related to conifers than to flowering plants. Proc. Nail. Acad. Sci. USA 96: 7342-7347.

Yu, D., Kotilainen, M., Pollanen, E., Mehto, M., Elomaa, P., Helariutta, Y., Albert, VA. and Teeri, T.H., 1999. Organ identity genes and modified patterns of flower development in *Gerbera hybridia* (Asteraceae) Plant J. 17: 51-62.