

Extended expression of B-class MADS-box genes in the paleoherb *Asarum caudigerum*

Yin-He Zhao · Michael Möller · Jun-Bo Yang · Ting-Song Liu ·
Jin-Feng Zhao · Li-Na Dong · Jin-Peng Zhang · Cheng-Yun Li ·
Guo-Ying Wang · De-Zhu Li

Received: 9 June 2009 / Accepted: 23 October 2009 / Published online: 11 November 2009
© Springer-Verlag 2009

Abstract *Asarum caudigerum* (Aristolochiaceae) is a paleoherb species that is important for research in origin and evolution of angiosperm flowers due to its basal position in the angiosperm phylogeny. In this study, a subtracted floral cDNA library from floral buds of *A. caudigerum* was constructed and cDNA arrays by suppression subtractive hybridization were generated. cDNAs of floral buds at different stages before flower opening and of leaves at the seedling stage were used. The macroarray analyses

of expression profiles of isolated floral genes showed that 157 genes out of the 612 unique ESTs tested revealed higher transcript abundance in the floral buds and uppermost leaves. Among them, 78 genes were determined to be differentially expressed in the perianth, 62 in the stamens, and 100 genes in the carpels. Quantitative real-time PCR of selected genes validated the macroarray results. Remarkably, *APETALA3* (*AP3*) B-class genes isolated from *A. caudigerum* were upregulated in the perianth, stamens and carpels, implying that the expression domain of B-class genes in this basal angiosperm was broader than those in their eudicot counterparts.

Electronic supplementary material The online version of this article (doi:10.1007/s00425-009-1048-6) contains supplementary material, which is available to authorized users.

Y.-H. Zhao · J.-B. Yang · L.-N. Dong · D.-Z. Li (✉)
Key Laboratory of Biodiversity and Biogeography,
Kunming Institute of Botany, Chinese Academy of Sciences,
650204 Kunming, China
e-mail: dzl@mail.kib.ac.cn

Y.-H. Zhao
e-mail: yhzha0808@163.com

Y.-H. Zhao · C.-Y. Li
College of Agronomy and Key Laboratory for Plant
Pathology of Yunnan Province, Yunnan Agricultural University,
650201 Kunming, China

M. Möller
Royal Botanic Garden Edinburgh, 20A Inverleith Row,
Edinburgh EH3 5LR, Scotland, UK

T.-S. Liu · J.-F. Zhao
State Key Laboratory of Agrobiotechnology,
China Agricultural University, 100094 Beijing, China

J.-P. Zhang · G.-Y. Wang (✉)
Institute of Crop Sciences,
Chinese Academy of Agricultural Sciences,
Zhongguancun Nandajie 12, 100081 Beijing, China
e-mail: gywang@cau.edu.cn

Keywords B-class MADS-box genes ·
Floral gene expression · cDNA macroarray ·
Phylogenetic analysis

Abbreviations

<i>A. caudigerum</i>	<i>Asarum caudigerum</i>
<i>AP3</i>	<i>APETALA3</i>
<i>PI</i>	<i>PISTILLATA</i>
<i>SEP</i>	<i>SEPALLATA</i>
<i>TM6</i>	<i>TOMATO MADS-BOX GENE 6</i>
<i>AcAP3-1</i>	<i>Asarum caudigerum APETALA3-1</i>

Introduction

Typical angiosperm flowers comprise concentric arrangements of four types of organs, arranged from outward in: sepals, petals, stamens, and the inner carpels (Meyerowitz et al. 1989). All of these whorls are interpreted as modified leaves, as Goethe proposed 200 years ago (Pelaz et al.

2001; Ditta et al. 2004). Flowers develop under the control of homeotic genes, many of them from the MADS-box gene family. They encode transcription factors that play crucial roles in the development of floral primordia and the establishment of floral organ identities (Weigel 1998; Theissen et al. 2000; Theissen 2001; Becker and Theissen 2003).

The genetic regulation of floral organ formation in typical eudicot flowers, such as *Arabidopsis thaliana* or *Antirrhinum majus*, was initially described with the ABC model for the specification of floral organ identities (Bowman et al. 1991; Coen and Meyerowitz 1991). Since then, new classes of genes have been identified, such as D-function genes thought to be involved in ovule identity (Angenent et al. 1995). The four class E *SEPALLATA* genes encode proteins that are apparently required in floral organ identity determination (Pelaz et al. 2000, 2001; Ferrario et al. 2003; Ditta et al. 2004). The products of the A, B, C and E genes act in a combinatorial manner to achieve three types of activity. In the outermost whorl, A and E genes function to control sepal identity. In the second whorl, A, B, and E genes together specify petal identity. In the third whorl, B, C, and E genes act in concert to direct stamen identity, while in the innermost whorl, C and E genes determine carpel identity (Bowman et al. 1991; Coen and Meyerowitz 1991; Honma and Goto 2001; Pelaz et al. 2000, 2001; Theissen 2001; Ditta et al. 2004). The functions of A and C genes are considered mutually antagonistic and the expression of one represses that of the other (Bowman et al. 1991; Coen and Meyerowitz 1991; Drews et al. 1991; Egea-Cortines et al. 1999; Pelaz et al. 2000; Theissen 2001).

In higher eudicots, B-class gene products are represented by homologs of the *Arabidopsis thaliana* genes *APETALA3* (*AP3*) and *PISTILLATA* (*PI*), which control petal and stamen identity in the second and third whorls, respectively (Jack et al. 1992; Goto and Meyerowitz 1994). In addition to petal and stamen expression, however, *AP3* and *PI* transcripts are detected in carpel tissue in basal angiosperms such as *Amborella* and *Nuphar* (Kim et al. 2005). Other angiosperms and gymnosperm further possess a sister clade of B genes, termed Bsister (Bs) genes, and expression studies revealed that these genes are predominantly expressed in female reproductive organs (including carpels and ovules) (Becker et al. 2002).

Asarum caudigerum Hance is a paleoherb belonging to family Aristolochiaceae of the magnoliid order Piperales, which is phylogenetically near the base among angiosperms (Kramer and Irish 2000; Angiosperm Phylogeny Group 2003). For another basal angiosperm genus, *Amborella* (Amborellales) actually the most basal lineage of the extant angiosperms (Mathews and Donoghue 1999; Qiu

et al. 1999; Soltis et al. 1999; Angiosperm Phylogeny Group 2003), Buzgo et al. (2004) and Soltis et al. (2007a) hypothesized that there is a gradual transition in the expression of B-class genes throughout the floral parts in basal angiosperms. They proposed a “fading borders” model of floral gene expression in these basal angiosperms, and Kim et al. (2005) suggested that their B-class genes expression is broader than those of their counterparts in eudicots and monocots.

In the present study, we investigated this hypothesis by studying MADS-box genes in a little-studied basal angiosperm species, *A. caudigerum*. We took a genome-wide screening approach to study its MADS box gene homologs. We analyzed the expression profiles of all genes expressed during floral bud development by macroarray. We identified MADS-box homologs in *A. caudigerum* by phylogenetic analysis, and determined how its B-class genes are expressed through gene expression pattern analyses, qRT-PCR and RNA in situ hybridization. We further showed that the B-class genes isolated from this basal angiosperm have a broader expression pattern than those of their counterparts in higher eudicots.

Materials and methods

Plant materials, total RNA extraction, DNA sequencing, data analysis, and data deposition

Plants of *Asarum caudigerum* Hance were introduced from the wild in Qiubei county of Yunnan province, SW China, and cultivated in the Botanical Garden of the Kunming Institute of Botany (Chinese Academy of Sciences, Kunming, Yunnan, China). Seedling leaves and floral organs of buds at several stages before anthesis were collected and transferred directly to liquid nitrogen. Total RNA was isolated from the frozen leaves and floral buds using Trizol (Shanghai Huashun Company, Shanghai, China) according to the manufacturer’s protocol. To identify floral organ-specific genes and to elucidate their expression patterns, 612 unique EST genes were used in the present study: 567 unique ESTs were previously published (Zhao et al. 2006; GenBank EST database DV038159–DV038720 and DV075851–DV075856), and 45 unique genes from 400 newly sequenced clones of the previously prepared cDNA library described in Zhao et al. (2006) were obtained and deposited in the GenBank EST database (EE127746 and EE605642–EE605686). Protein similarity searches were performed against the NCBI database using the BLASTx program to assign putative functions to these ESTs. In the present study, *e*-values less than $1e^{-5}$ with more than 100 nucleotides of the ESTs were considered significant.

cDNA macroarray preparation

A cDNA macroarray was prepared according to a previously published method (Ji et al. 2003), with minor modifications. The PCR product from each unique EST was transferred from a 384-well plate to a nylon membrane (Amersham Biosciences, Arlington Heights, IL, USA) using the Biomek 2000 Laboratory Automation Workstation (Beckman Coulter, Fullerton, CA, USA). The PCR products were reamplified by PCR (Perkin-Elmer GeneAmp PCR System 9600) using nested PCR primers 1 and 2R provided in the PCR-Select cDNA Subtraction kit (Clontech, Palo Alto, CA, USA). These primers were complementary to sequences flanking both sides of the cDNA insert. Thermo-cycling conditions were as follows: one step at 94°C for 3 min, followed by 28 cycles of 95°C for 10 s, 68°C for 3 min, and 72°C for 7 min.

Each clone was blotted in quadruplicates with the spots 1.125 mm in diameter and 1.25 mm apart. After air-drying, the membranes were denatured in 0.6 M NaOH for 5 min, neutralized in 0.5 M Tris-HCl (pH 7.5) for 5 min and then rinsed in distilled water for 3 min. The blotted cDNA samples were cross-linked to membranes using a low-energy UV source and were baked for 2 h at 80°C. SARS virus genes, distilled water, and the PCR reaction solution were also transferred onto the membrane as negative controls.

Macroarray hybridization, washing, and radioactive scanning

Total RNAs prepared from leaves at the seedling stage, uppermost leaves (upper leaves which close to floral buds), perianth, stamens, and carpels were reverse transcribed and used as probes for expression profile analysis. The reverse transcription reaction was performed in 20 µl volumes set up as follows: 1 µl oligo(dT)₁₈ (10 mM/µl), 5 µg total RNA and distilled water (up to 8 µl). The mix was heated to 65°C for 5 min, then quickly chilled on ice, and the contents collected by centrifugation at 10,000g for 20 s. Then, 4 µl of 5×first-strand buffer, 2 µl of 0.1 M DTT, 1 µl of 10 mM dNTP mix (10 mM each dATP, dTTP and dGTP), 1 µl RNasin (40 U µl⁻¹), 3 µl [³²P]-dCTP (370 000 Bq µl⁻¹) and 1 µl (200 U) of SuperScriptTM II polymerase (Invitrogen, Carlsbad, CA, USA) was added, the solutions mixed by gentle vortexing and incubated at 42°C for 1 h. The probes were denatured at 100°C for 5 min, and then chilled for 5 min on ice before hybridization. Membranes were pre-hybridized in 20 ml Church solution (1% BSA, 1 mM EDTA, 0.25 M Na₂HPO₄-NaH₂PO₄, 7% SDS) at 65°C for 5 h. The denatured probes were then added to the Church solution and hybridization was carried out overnight at 65°C. After hybridization, the membranes

were washed at 65°C in 2 × SSC, 0.5% SDS for 10 min, in 1 × SSC, 0.5% SDS for 10 min, then in 0.5 × SSC, 0.5% SDS for 10 min, and finally in 0.1 × SSC, 0.1% SDS for 10 min. They were then exposed to storage phosphor screens (Amersham) for 3 days. Images were acquired by scanning the membranes with a Typhoon 9210 scanner (Amersham).

Data analysis

Data were analyzed using the GPC Visualgrid software (<http://www.gpc-biotech.com>). The radioactive intensity of each spot was quantified as volume values and the local background levels subtracted, resulting in subtracted volume values, designated sVOL. The mean of all spot intensities in each membrane was used as the internal control, the subtracted volume value of which was designated sRef. All images were normalized by dividing the sVOL of each spot by the sRef value within the same image, resulting in a normalized volume value (nVOL) for each spot. The nVOL values were comparable between all images. The ratios of the signal intensities for each EST in the uppermost leaves, perianth, stamens, and carpels, to those of the leaves at the seedling stage were calculated as measures of the changes in the differential expression of the genes represented by the cDNA spots on the macroarrays. A twofold expression cutoff was applied to make the analysis more stringent, i.e., spots with a ratio equal to or more than two were judged significantly upregulated (Jia et al. 2006) and loaded into the program Hierarchical Clustering Explorer 3.0 (HCE3.0) for analysis (<http://www.cs.umd.edu/hcil/hce/hce3.html>).

Real-time quantitative PCR analysis

Total RNA was reverse-transcribed with oligo-dT and reverse transcriptase (Promega, Madison, WI, USA) following the supplier's protocols. To examine the expression of the three putative MADS-box transcription factors (DV038650, DV038420, DV038434) and one gibberellin-regulated gene (DV038510), quantitative real-time PCR (qRT-PCR) was carried out with an ABI Prism 7900 HT sequence detection system (Applied Biosystems, Foster City, CA USA). The *A. caudigerum* 28S rRNA (DV038186) gene was used as internal control for normalization of the template cDNA. The DV038420 sense primer was DV038420F (5'-GTCCATGGCGGGGAGTTTCTCTCTTTTC-3') and the anti-sense primer DV038420R (5'-GAGCAGCCATTCCAAGGAGTGAG-3'); the DV038434 sense primer was DV038434F (5'-AGCCATTCCAAGGAGTGAG-3'), and the anti-sense primer DV038434R (5'-GCATACTTACATTCCAGGTCTTC-3'); the DV038510 sense primer was DV038510F (5'-ATTTCGATTTTCGTA

TCCAAC-3'), and the anti-sense primer DV038510R (5'-TGTTATGCAAGGCGCAGATCGTG-3'); the DV038650 sense primer was DV038650F (5'-AAGACCAAGGAA GGAGGAC-3'), and the anti-sense primer DV038650R (5'-CGGCCGCGACCACGCTAATC-3'); the DV038186 sense primer was DV038186F (5'-AGGAGACTGCGTTG ATGTG-3'), and the anti-sense primer DV038186R (5'-CAAATCCAAACGAAAGGGACAAT-3'). qRT-PCR was performed in 1× SYBR Green I PCR Master Mix (Applied Biosystems), containing 200 nM of each primer and 1 µl 1:10 diluted cDNA. The PCR was performed under thermal cycling conditions as follows: 1 cycle at 50°C for 2 min, 1 cycle at 95°C for 10 min, 40 cycles at 95°C for 30 s, at 60°C for 30 s, and at 72°C for 18 s. Following amplification, melting curve analyses were performed at 95°C for 10 min followed by 40 cycles of 95°C for 30 s, 60°C for 30 s, and 72°C for 20 s. The data collected during each extension phase were analyzed initially using SDS2.1 (Applied Biosystems). The 28S rRNA gene was used as an internal calibrator to standardize the RNA content of the different tissues. Measurement of the leaves at the seedling stage was used as a sample calibration control. The abundance of DV038650, DV038420, DV038434 and DV038510 transcripts was calculated using the relative $2^{-\Delta\Delta CT}$ analytical method (Livak and Schmittgen 2001). The mean of triplicates of the same RNA sample was used as final result of gene expression, and the standard deviation of the three reactions calculated.

Scanning electron microscopy (SEM)

Floral primordia at different stages of development were dissected in the greenhouse, fixed in FAA (5% formaldehyde, 5% acetic acid, in 70% alcohol) for 13 h, and were then dehydrated through a series of alcohol solutions ranging from 70 to 100%. The materials were further dissected under a stereomicroscope, and the alcohol replaced by isopentanol acetate before the samples were dried in a Hitachi HCP-2 CO₂ Critical Point Dryer (CPD). The dried material was mounted on stubs and coated with gold–palladium. Observations were made using a Hitachi KY Amray-100B SEM at 25KV and 0.5–1 mm working distance.

Database searches and isolation of a putative B-class genes

Using the previously described EST DV038650 (Zhao et al. 2006), we performed RACE using the 5'-RACE cDNA Amplification Kit (Clontech) according to the manufacturer's protocol. We subsequently performed database searches in GenBank for the resulting mRNA and predicted protein sequences. This enabled the elaboration

of additional cDNA sequence, deposited in GenBank as EU368583. The complete coding region of EU368583 was amplified by PCR using the specific sense primer EU368583F (5'-GATCCATGGGCTGCGCGACGTCCAA-3') and anti-sense primer EU368583R (5'-GTAGGTGACC ACTTTGTTATGCAAGGCGCAGATCG-3'). The conditions for amplification were 94°C for 3 min followed by 30 cycles at 94°C for 30 s, at 54°C for 30 s and at 72°C for 1 min, plus a final extension at 72°C for 7 min. PCR products were purified and cloned into pGEM T-easy vectors (Promega) according to the manufacturer's protocol, and finally the clones were sequenced.

Construction of phylogenetic trees

The putative MADS-box amino acid EU368583 isolated from *A. caudigerum* was aligned with those closely related to B-class genes and MADS-box genes from GenBank using Clustal X (Thompson et al. 1997) followed by manual adjustment where necessary. Alignments were created from the relatively conserved MIK amino acids (the nucleotides were too variable to be aligned unambiguously) while the C domain was excluded. A Neighbor-Joining (NJ) tree was constructed using the pairwise deletion option in MEGA4 (Tamura et al. 2007). Genetic distances were estimated under the Poisson correction model.

RNA in situ hybridization

Floral buds were fixed in FAA, and were dehydrated through a standard ethanol series. The buds were transferred to liquid nitrogen, shock-frozen, and were stored at –80°C. Frozen tissues were later embedded with an embedding Optimum Cutting Temperature (OCT) compound ("Tissue-Tek"; Miles Laboratories Inc., Elkhart, IN, USA), sectioned at 10 µm thickness and mounted on glass slides. RNA in situ hybridization was carried out according to Yang et al. (2005) with minor modifications. Immunodetection with anti-DIG antibodies conjugated with alkaline phosphatase was carried out using Dig Northern Starter Kit (Roche, Mannheim, Germany) according to the manufacturer's protocol. A 326-bp antisense probe was prepared using the DV038650R (5'-AGATTTAGACCGTA GAGT-3') and the DV038650 T7 primer containing the T7 promoter (5'-TAATACGACTCACTATAGGG-3'). The sense probe was created using primer DV038650F (5'-GG TATAGCGTATAATGAGAC-3') and the T3 promoter primer (5'-AATTAACCCTCACTAAAGGG-3'). In situ hybridizations were performed at 46°C and subsequent washes were carried out at 37°C. Images were captured with an Olympus microscope (Olympus Company, Guangzhou, Guangdong, China).

Results

Expression patterns of floral genes in *A. caudigerum*

Total RNA from seedling leaves, uppermost leaves, perianth tissue, stamens, and carpels was used to synthesize a set of probes for the macroarray experiments (Fig. 1). In total, 157 genes out of the 612 unique EST genes tested showed an upregulated expression (Table S1). Of these, 43, 78, 62, and 100 were upregulated in the uppermost leaves, perianth, stamens and carpels, respectively. The expression of 29, 38 and 32 upregulated genes overlapped in the uppermost leaves and perianth, perianth and stamens, and stamens and carpels, respectively (Fig. 2a–c and Table S1). The gene expression pattern in the perianth was very similar to that in the uppermost leaves (Fig. 1a–d). However, we did not detect transcripts encoding putative MADS-box transcription factors in the uppermost leaves (Table 1 and Table S1). Of the 78 upregulated perianth transcripts (Table S1), seven putative MADS-box transcription factors (DV038213, DV038420, DV038528, DV038434, DV038183, DV038296, DV038162), including

three putative APETALA3-like proteins, two AGL6-like proteins, and two similar to putative MADS542 protein were found (Table 1). Sixty-two genes were upregulated in the stamens (Table S1). Of the stamen-specific transcripts, three genes encoded MADS-box transcription factors (DV038577, DV038682, DV038420), including one putative MADS box transcription factor, AP3-like, and one putative MADS542 protein (Table 1). Carpels showed more upregulated genes than the other organs, probably due to their complex tissue types, with 100 genes showing significant upregulation. They included three transcripts encoding putative MADS-box transcription factors (Table S1), two encoding putative AP3-like proteins (DV038213, DV038650), and one encoding a putative MADS542 protein (DV038528) (Table 1).

Validation of gene expression by qRT-PCR

The expression of three putative MADS-box transcription factors (DV038650, DV038420, DV038434) and one gibberellin-regulated protein (DV038510) was further validated by qRT-PCR (Fig. 3). Our qRT-PCR results showed

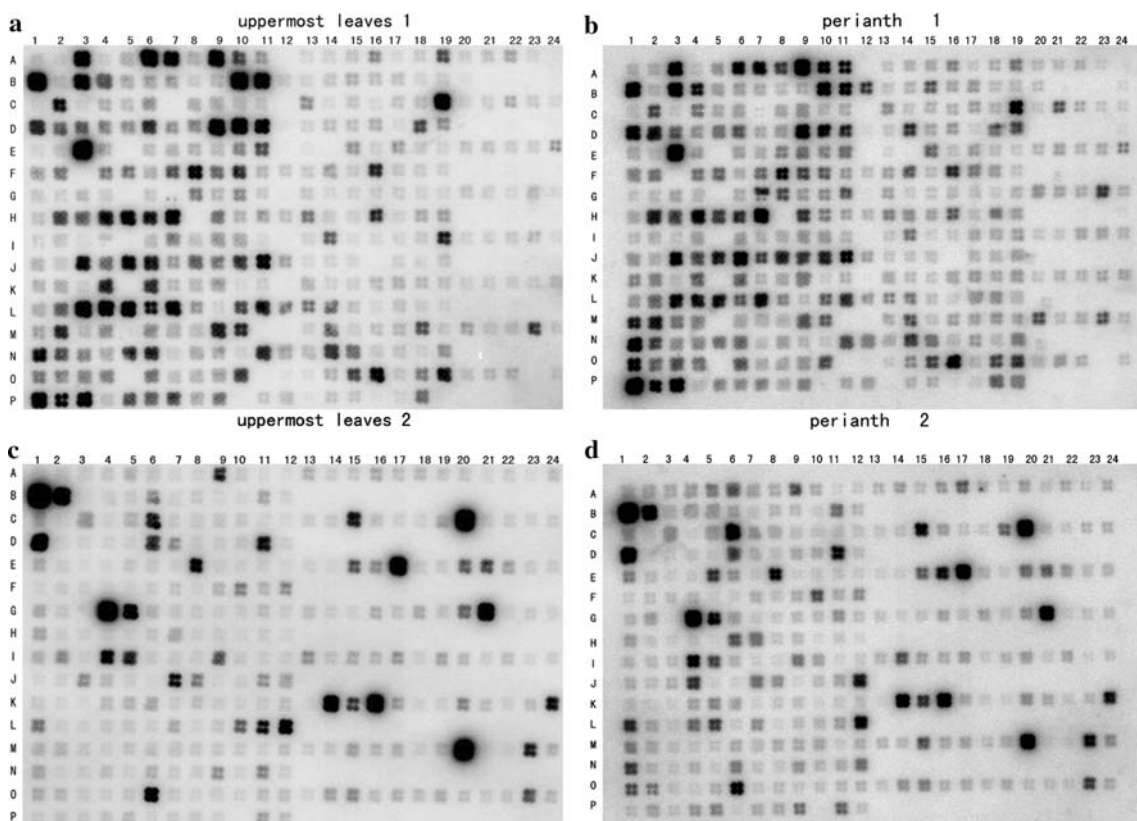
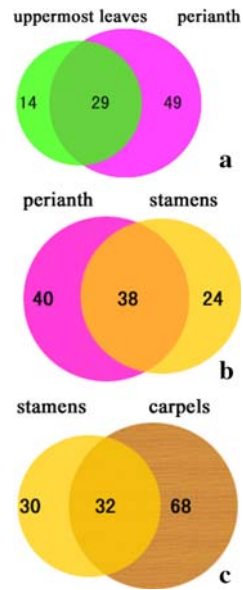


Fig. 1 Examples of cDNA macroarray hybridization of uppermost leaves and perianth transcripts. **a, b** The same cDNA macroarray probed with [³²P]-labeled probes made from the total RNA from the uppermost leaves and the perianth, respectively. Negative controls included distilled water (20–23 of B, D, F, H, J, L, N, P), PCR

reaction solution (B24, D24, F24, H24, J24, L24, N24, P24), and SARS virus genes (A12, C12, E12, G12, I12, K12, M12, O12) on the two membranes. **c, d** Another cDNA macroarray probed with [³²P]-labeled probes made from the total RNA of the uppermost leaves and perianth, respectively

Fig. 2 Comparison of the distribution of upregulated floral organ genes showing more than twofold increase in transcript abundance between different whorls and uppermost leaves.

a Genes that were expressed in both the uppermost leaves and the perianth. **b** Genes that were upregulated in both the perianth and stamens. **c** Genes that were upregulated in both the stamens and carpels



the upregulation of DV038650 in the carpels, DV038434 in the perianth, DV038420 in the perianth and stamens, and DV038510 in all the floral organs tested. This pattern confirmed the expression pattern of these genes observed in the macroarray experiments.

Early floral organ development

In *A. caudigerum*, floral primordia arise under transverse protuberances in axillary bract primordia. The whorl of the perianth appears first and then the first whorl of the androecium, quickly followed by the carpel primordia and then the second androecium whorl (Fig. S1). *A. caudigerum* usually has six carpels and two whorls of six stamens, although some buds have only five carpels and five stamens per androecium whorl. The second androecium whorl

continues to develop when the first whorl has already matured. Therefore, the series of floral organ development progresses from the first whorl to the third, then fourth, and lastly to the second whorl.

Identification of gene EU368583 and comparison of its amino acid sequence

We used RACE to obtain the full-length coding sequence of EU368583 and confirmed its homology with the previously published sequence DV038650 (Zhao et al. 2006). EU368583 cDNA had an open reading frame (ORF) of 678 bp, encoding 225 amino acids. The predicted protein showed the highest sequence similarity (94% identity) to the *A. europaeum* MADS box transcription factor *AP3-1*. Alignment with 31 putative B-class homologs and 7 putative Bsister (Bs) homologs showed that the putative B-class *AP3*-like gene isolated from *A. caudigerum* shared the PI and paleo*AP3* motifs of other *AP3* B-class genes (Fig. 4). The paleo*AP3* motif was not present in PI genes, and was not well conserved in the Bs genes (Fig. 4).

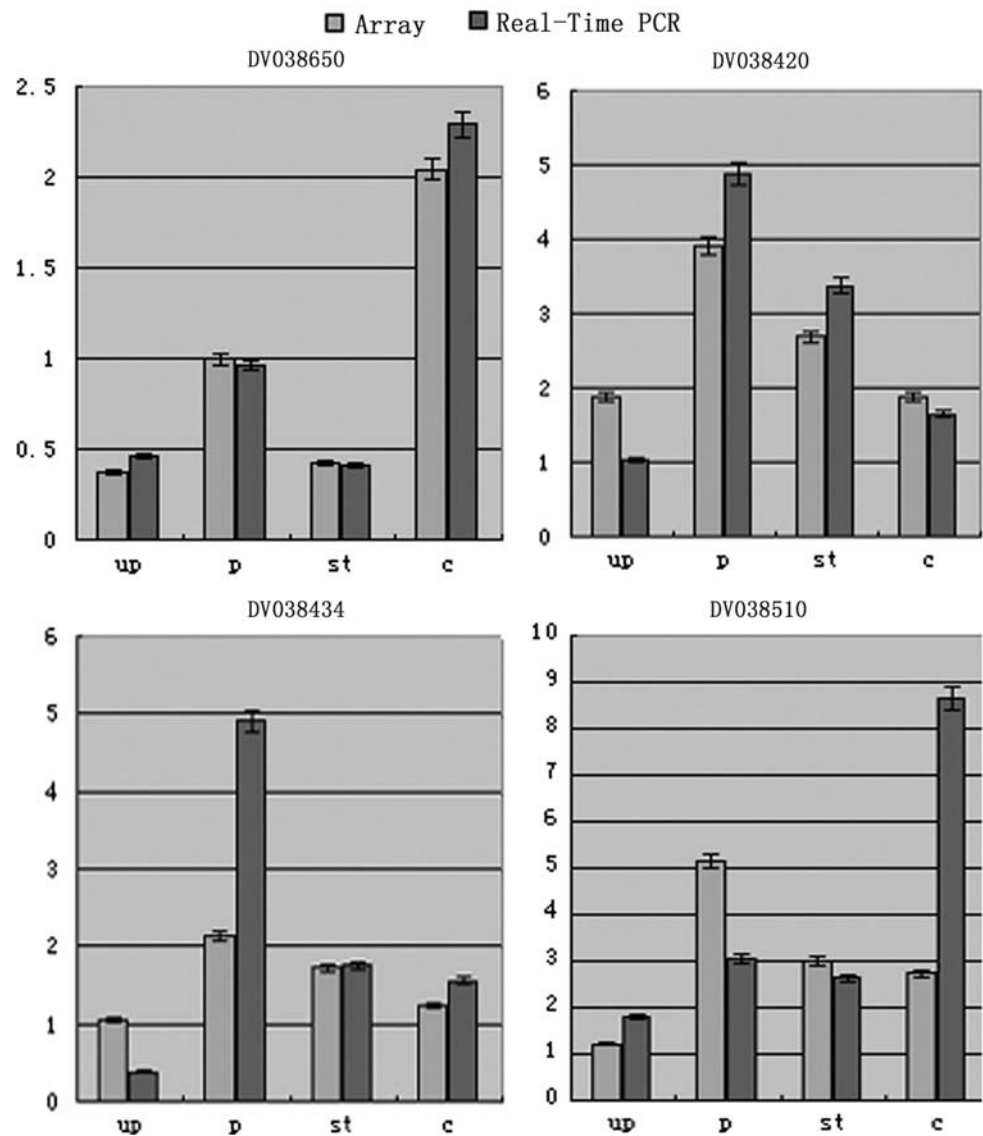
Homology testing and phylogeny reconstruction

Phylogenetic analysis was performed to further clarify the homology and relationships of the full-length putative *AP3*-like MADS-box transcription factor gene of *A. caudigerum* isolated in this study. The Neighbor-Joining analysis of the amino acid alignment yielded a high bootstrap support for the clade of *AP3*-like MADS-box transcription factor genes from *Asarum* (bootstrap, BS = 100%), including EU368583 (the full ORF length of DV038650) isolated from *A. caudigerum*. The latter formed a clade with *AP3-1* of *A. europaeum* (BS = 94%) (Fig. 5), and thus was not a member of PI or Bs gene

Table 1 Upregulated MADS-box genes in floral parts of *A. caudigerum*

Accession number	Putative gene function	Gene symbol	Uppermost	A & B class	B & C class	B & C class
			leaves	Perianth	Stamens	Carpels
DV038162	Putative APETALA3-like protein		1.02	2.02	1.29	1.94
DV038183	MADS-box transcription factor Pe.am.AGL6.2		1.88	3.91	1.69	1.89
DV038434	AGL6-like protein		1.40	2.13	1.72	1.68
DV038296	Putative APETALA3-like protein		1.76	2.07	1.74	1.37
DV038420	Similar to putative MADS542 protein		1.72	3.76	2.62	1.78
DV038528	Similar to putative MADS542 protein		1.15	2.16	0.92	3.34
DV038213	APETALA3-like protein		1.60	2.15	1.60	2.15
DV038650	Putative AP3-like of the MADS-box transcription factor	<i>AcAP3-1</i>	0.57	0.99	0.42	2.19
DV038577	Putative MADS box transcription factor AP3-like		0.76	1.31	2.15	1.31
DV038682	MADS-box transcription factor		1.10	1.56	2.19	1.17

Fig. 3 Validation of macroarray results by qRT-PCR for selected genes (DV038650, DV038420, DV038434, DV038510). *A. caudigerum* 28S ribosomal RNA (28S rRNA) was used as internal control for normalization of the template cDNA. Each PCR was repeated three times and the error bars represent standard deviation (SD). *Up* uppermost leaves, *p* perianth, *st* stamen, *c* carpel, gray bars represent Array; black bars represent real-time PCR



families. In addition, the *AP3-like* B-class genes from *Asarum* formed a cluster with *AP3-1* of *Saruma henryi* (BS = 77%).

In situ hybridization studies of the putative *AP3*-like homolog in *A. caudigerum*

To obtain information on the spatial expression pattern of the putative *AP3* homolog isolated from *A. caudigerum*, RNA in situ hybridization was performed. The putative MADS-box protein DV038650, a partial fragment of EU368583 was upregulated in carpels at two different stages of development: mid-development (Fig. 6c, e, f) or late-development (Fig. 6a, b, d). In longitudinal sections, the signal was especially strong at the adaxial base of the cupules, where ovules would later develop, but was very weak in the perianth, and absent from the stamens

(Fig. 6a–c, e). In transverse cross sections, a strong signal was also detected in the female reproductive structures (Fig. 6d, f). The use of sense probes for DV038650 as a control resulted in non-specific signals (Fig. 6g, h).

Discussion

We have performed a macroarray analysis of gene expression in leaves and flowers at different stages of development in the paleoherb *A. caudigerum*, with special focus on MADS-box genes. Because genome information for *A. caudigerum* is not available, we employed a large-scale screening approach to identify the target genes. In our study, it was noticeable that a considerable number of upregulated transcripts constituted heretofore-uncharacterized genes, or at least hypothetical proteins with unknown

Fig. 4 Alignment of the conserved C-terminal domain of predicted proteins from B and Bs lineage homologs. The PI and paleoAP3 motifs (Kramer et al. 1998; Jaramillo and Kramer 2004) are shaded

	PI Motif-derived	paleoAP3 Motif	
PI	<i>Asarum speciosum</i> PI	FAFCVQ-PMQPNLHQKK	
	<i>Asarum splendens</i> PI	FAFCVQ-PMQPNLHQKK	
	<i>Asarum europaeum</i> PI	FAFCVQ-PMQPNLHQKK	
	<i>Saruma henryi</i> PI	FAFCVQ-PMQPNLHQKK	
	<i>Drimys winteri</i> PI	LSFCIQ-PIQPNLQQNK	
AP3	<i>Asarum caudigerum</i> EU368583	FAFRIQ-PCQPNLHNAG---	YGIHDLRLA
	<i>Asarum speciosum</i> AP3	FAVRIQ-PCQPNLHNAG---	YGIHDLRLA
	<i>Saruma henryi</i> AP3	FAFRVQ-PCQPNLHDAG---	YGTHDLRLA
	<i>Asarum europaeum</i> AP3-1	FAFRIQ-PCQPNLHNAG---	YGIHDLRLA
	<i>Asarum splendens</i> AP3	FAFRIQ-PCQPNLHNAG---	YGIHDLRLA
	<i>Asarum speciosum</i> AP3	FAVRIQ-PCQPNLHNAG---	YGIHDLRLA
	<i>Dendrobium crumenatum</i> AP3	YSYRTQ-PSQPNLQGMG--	YGPHDLRLA
	<i>Dendrobium crumenatum</i> AP3	YSYRTQ-PSQPNLQGMG--	YGPHDLRLA
	<i>Magnolia praecocissima</i> MpMADS7	-----LHDTG---	FGIHDRLA
	<i>Magnolia grandiflora</i> AP3	-----LHDTG---	FGIHDRLA
	<i>Liriodendron tulipifera</i> LtAP3	-----LHNTG---	FGIHDRLA
	<i>Michelia figo</i> MfAP3	-----LHDTG---	FGIHDRLA
	<i>Drimys winteri</i> AP3-2	FAFRMQ-PYQPNLHDPS---	YGIHDLHLA
	<i>Drimys winteri</i> AP3-4	FAFHVQ-PCQPNLQDNG---	YGCHDLRLA
	<i>Drimys winteri</i> AP3-1	FAFRVQ-PYQPNLHDPS---	YGIHDRIRS
	<i>Drimys winteri</i> AP3-3	FAFRVQ-PFQPNLHDPS---	YGIHDLHLA
	<i>Lindera erythrocarpa</i> AP3-1	FAFCIQ-PCQPNLQDTR---	YGSHNLRLA
	<i>Lindera erythrocarpa</i> AP3-2	FAFCIQ-PCHPNLQETR---	YGSHDLRLS
	<i>Calycanthus floridus</i> AP3-2	FALCVQ-PCHPNLQDAR---	SGHDLRLA
	<i>Alpinia hainanensis</i> AP3	YSYKVV-PNQPQLQAIRSY-	ASHHDLRLA
	<i>Hordeum vulgare</i> AP3	YAFRVV-PSQPNLHGMA-Y-	GGSHDLRHL
	<i>Sanguinaria canadensis</i> AP3	FAFRLQ-PSQPNLHDG----	YSSNDLRLA
<i>Dicentra eximia</i> DeAP3-1	FAFRLQ-PSQPNLHDGGG--	YGSHDLRLA	
<i>Meliosma dilleniifol</i> AP3-3	FAFRLQ-PSQPNLQD-GV--	YGLQDLRLA	
<i>Meliosma dilleniifol</i> AP3-2	FAFRLQ-PTQPNLQDSGA--	YGLQDLRLA	
<i>Meliosma dilleniifol</i> AP3-3	FAFRLQ-PSQPNLQDGGGA--	YGMRDRLA	
Bsister	<i>Zea mays</i> ZMM17	LTFRQL-PTQPNLQDPGCSS	AAGHGLQLW
	<i>Asarum europaeum</i> AeAP3-2	LSFRLQ-PAQPNLQDANLLP---	HDLQL-
	<i>Drimys winteri</i> Bs	LSFRLQ-PTQSNLQDAN---	-----LQL-
	<i>Gnetum gnemon</i> GGM13	LGFRQL-PTQPNLQESSIVP---	PVLQL-
	<i>Aquilegia alpina</i> Bs	LSFRLQ-PTQPNLQERGLQY---	PDLQL-
	<i>Petunia</i> PBP24	ASYRLQ-PSHPNLQDSHVHG---	PSYD--
<i>Arabidopsis thaliana</i> TT16	ATYNLQ-LAQPNLQNDPTAQ---	ND----	

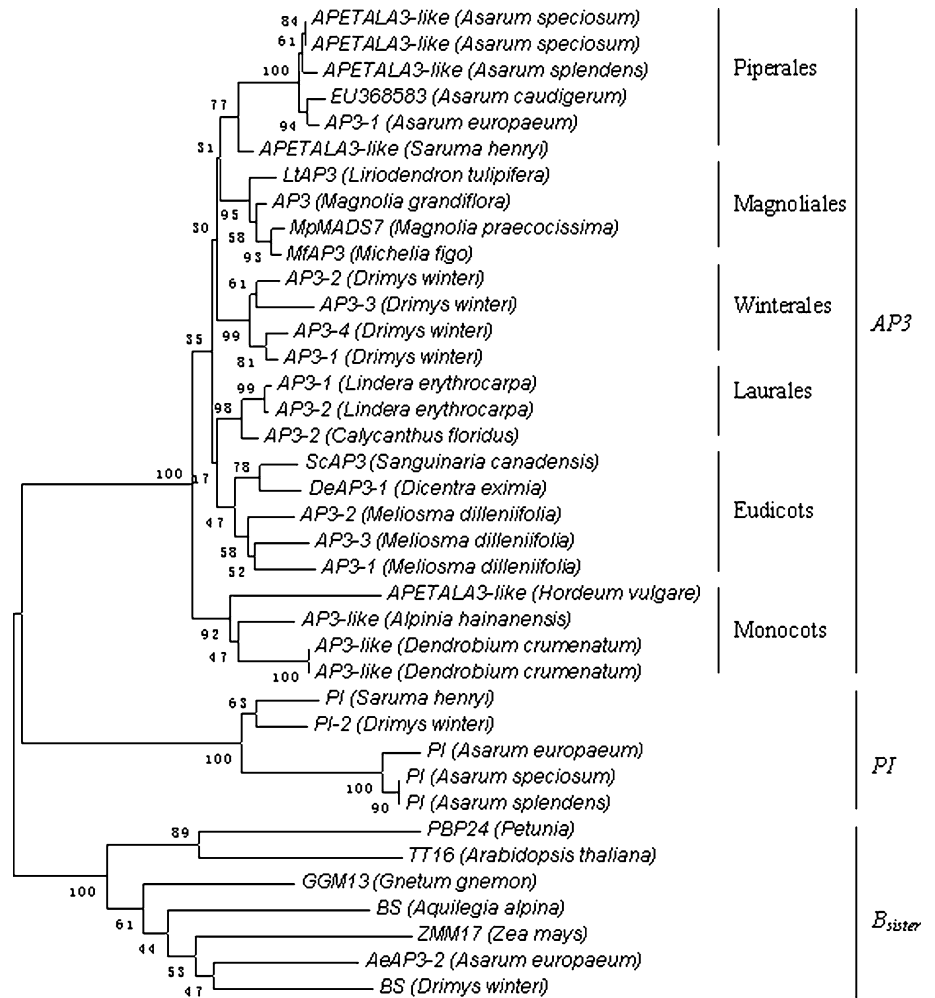
gene function. In addition, many transcription factors were differentially upregulated in the flower whorls.

Basal angiosperms are very interesting to the study of the origin, diversification, and evolution of angiosperms. Using Neighbor-Joining analysis, we demonstrated that the putative MADS-box transcription factor *AcAP3-1* isolated from *A. caudigerum* is a B-class gene based on the presence of conserved PI motif-derived and the paleoAP3 motif (Fig. 4) and the positioning in the phylogeny (Fig. 5). It did not fall anywhere near *AeAP3-2* and the other Bsister genes (Fig. 5). Although *AcAP3-1* most closely resembled *AeAP3-1*, the former showed carpel expression that had not been detected in the latter (Kramer and Irish 2000) (Fig. 6). The Bsister gene *AeAP3-2*, however, is expressed in female reproductive structures (Kramer and Irish 2000). Recent comprehensive expression studies in other plants groups revealed

that Bsister genes are mainly transcribed in female reproductive organs (Becker et al. 2002; de Folter et al. 2006). Why *AP3-1* from *A. europaeum* should behave differently from its *A. caudigerum* homolog is not yet understood.

In the typical ABC model based on *A. thaliana*, the B-class genes, *AP3* and *PI* control the specification of petals in conjunction with A-class genes and stamens in conjunction with C-class genes in the second and third whorls, respectively (Bowman et al. 1989; Coen and Meyerowitz 1991; Ma and dePamphilis 2000). The borders of B-function gene expression are also not so clear-cut in *Arabidopsis*: *AP3* and *PI*, the major *Arabidopsis* B-function genes, are not entirely restricted to the second and third whorl of the flowers; instead, *AP3* is expressed in parts of the first whorl, and *PI* is expressed in parts of the fourth whorl at early stage (Jack et al. 1992; Goto and

Fig. 5 Neighbor-Joining analysis of the 38 B-class and Bs putative MADS-box transcription factors, using 117 sites and 500 bootstrap iterations. Representative B-class and Bs lineage genes from GenBank were included in the analysis. EU368583 was isolated from *A. caudigerum* based on amino acid of MIK domain. The NJ tree was constructed under Poisson correction model



Meyerowitz 1994; Chen et al. 2000). It is believed that physical interactions of the two proteins are required for protein stability and in turn maintenance of their expression, so that in older flowers their expression is limited to whorls two and three (Goto and Meyerowitz 1994; Riechmann et al. 1996). It is possible that such a stabilization mechanism will not be present in basal angiosperms. Based on our cDNA microarray data, the expression domain of B-class AP3-like genes in *A. caudigerum* was found to be broader than for their counterparts in eudicots (Fig. 7), as *AcAP3-1* expression was found in carpels, and other AP3-like genes were also expressed in the perianth and stamens (Fig. 7; Table 1). Comparative studies of B-class AP3 gene homologs conducted in the family Aristolochiaceae reported AP3 gene expression for the genera *Saruma* and *Aristolochia* (Jaramillo and Kramer 2004). Specifically, B-class AP3 genes were found to be expressed in the second, third and fourth whorls of *Saruma henryi*, as well as the third and fourth whorls of *Aristolochia manshuriensis* (Jaramillo and Kramer 2004). This is consistent with the view of a broader expression pattern of B-class genes in basal angiosperms (Buzgo et al. 2004; Kim et al. 2005).

Based on our SEM observations, we found that the early floral organ development initiated in the sequence (1) first whorl, (2) third whorl, (3) fourth whorl, and (4) second whorl. Therefore, the early floral organ development of *A. caudigerum* does not progress from the outer whorl inward, but instead the development of the second whorl is delayed. It is consistent with the hypothesis of Buzgo et al. (2004) who suggested a gradual transition in the expression of floral B-class genes in *Amborella*. Our cDNA microarray data, RNA in situ hybridization, and qRT-PCR results indicated that *AcAP3-1/DV038650* is expressed in carpels, and particularly strongly in the endothelium of ovules. This gene's unusual expression may explain the retarded development of the second-whorl organs.

Kramer and Irish (1999, 2000) proposed that the ABC model was not rigidly fixed during the earliest stages of angiosperm evolution. Irish (2003) cautioned that the applicability of this model outside of the core eudicots would require testing. Indeed, several lines of evidence suggest patterns of gradual transitions among flowers that prompt a “fading borders” view of gene expression. This evidence includes the morphological transition between

Fig. 6 In situ expression of DV038650 in developing flower buds of *A. caudigerum*. **a–f** In situ hybridization with a DV038650 anti-sense probe. **g, h** In-situ hybridization with a DV038650 sense probe as negative control. **a–c, e, h** Longitudinal sections; **d, f, g** cross sections of floral buds. *p* perianth, *st* stamen, *c* carpel, *ov* ovule. The DV038650 expression can be detected in the ovule and stigma (arrowhead). Scale bars 500 μ m

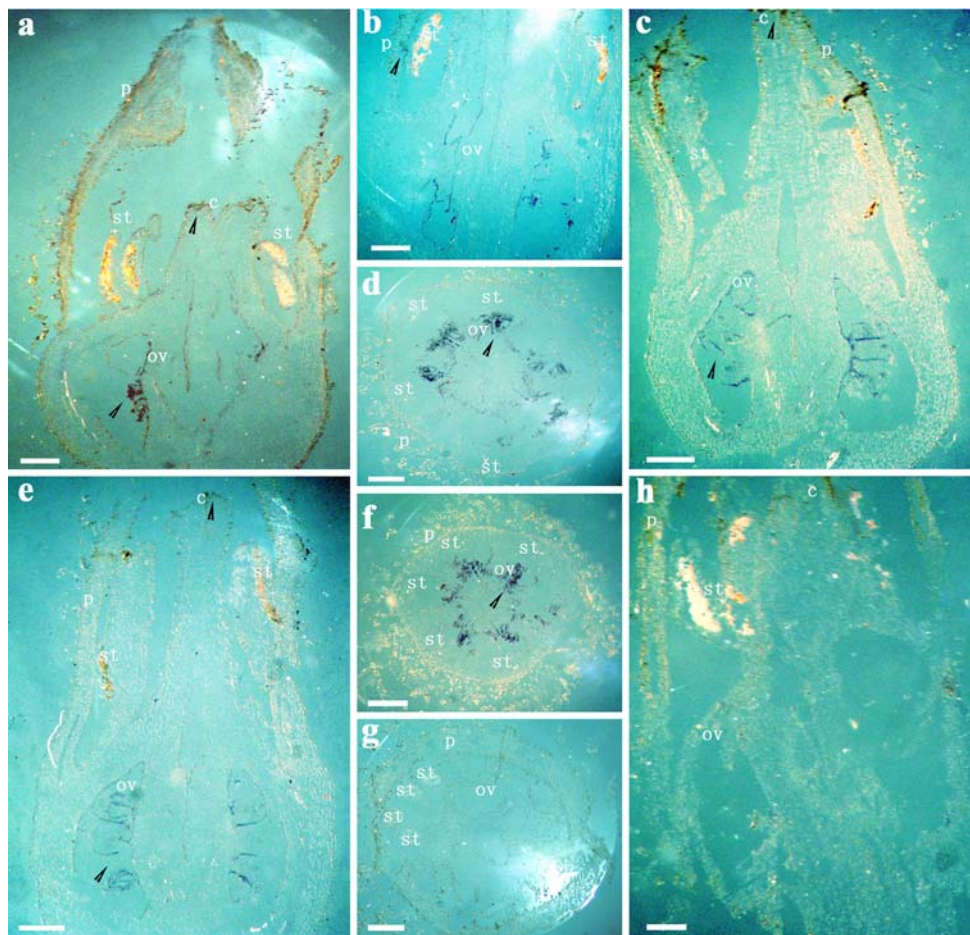
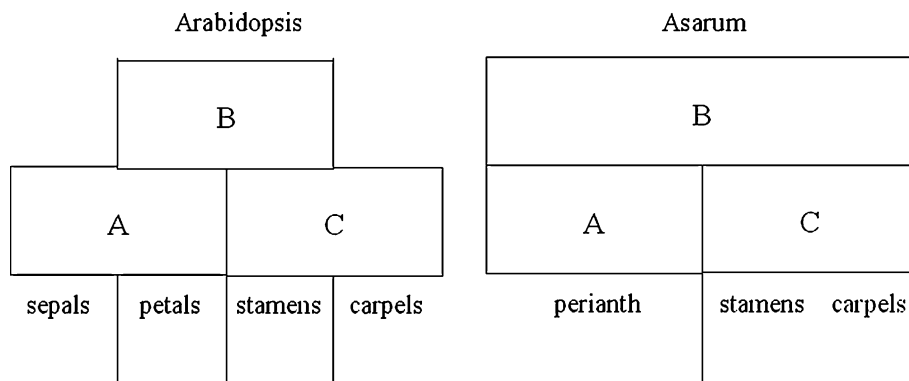


Fig. 7 Models of expression domains of MADS-box genes in floral organs in *Arabidopsis thaliana* and *A. caudigerum*, illustrating the hypothesized extended domain of B-class genes in *A. caudigerum*



organ identities from perianth to carpels, the results from genome-scale phylogenomics research, and the implied gradual shift in the expression of the B-class genes across the flower in the basal angiosperms (Kramer et al. 2003; Buzgo et al. 2004; Kim et al. 2005; Soltis et al. 2007a, b). However, there could be antagonistic interactions between the A-class gene products, which are partially responsible for perianth identities, and the C-class genes, which specify stamen and carpel identities in the higher eudicots (Coen and Meyerowitz 1991; Drews et al. 1991; Theissen 2001). In *Arabidopsis thaliana*, *APETALA2* (*AP2*) of the A-class

genes promotes the B gene expression domain by antagonizing *AGAMOUS* (*AG*) (Zhao et al. 2007). Winter et al. (2002a, b) pointed out that an ancestral B protein from the gymnosperm *Gnetum gnemon* binds DNA in a sequence-specific manner as a homodimer, whereas in *Antirrhinum majus* the DEF-like protein binds to DNA only as a heterodimeric complex with the GLO-like protein. Less strict requirements for heterodimerization or autoregulatory upregulation may facilitate spatial shifts in the B-function in taxa outside of the eudicots, as detected by the broader gene expression of B-class genes among flowers of

A. caudigerum, *Amborella* and other basal angiosperm lineages. Therefore, we support the contention that the ABC model was not established at the outset of angiosperm evolution, but instead occurred in a simpler form in the basal angiosperms, and gradually evolved into the canonical ABC model in the more derived higher eudicot lineages.

Acknowledgments We would like to thank Li-Jia Qu, Long Hong (National Laboratory for Protein Engineering and Plant Genetic Engineering, Peking–Yale Joint Research Center for Plant Molecular Genetics and Agrobiotechnology, College of Life Sciences, Peking University, Beijing, China) for helpful preparation and scanning of the cDNA macroarray. We are also very grateful to Ying Zhang and Jun-Jie Fu (State Key Laboratory of Agrobiotechnology, China Agricultural University, Beijing 100094, China) for help with macroarray data analysis and Bing Su (Kunming Institute of Zoology, the Chinese Academy of Sciences) and Yue-Qiu He (Yunnan Agricultural University) for help with *Asarum caudigerum* in situ hybridization experiments. We are also very grateful to Zheng Meng (Institute of Botany, the Chinese Academy of Sciences), Li-Zhi Gao, Zhen-Hua Guo, Hong Wang (Kunming Institute of Botany, the Chinese Academy of Sciences) and Zachary Larson-Rabin (University of Wisconsin-Madison) for helpful discussion and comments on the manuscript. This research was supported by the Ministry of Science and Technology of China (Project 2004DKA30430 to De-Zhu Li), the National Natural Science Foundation of China (NSFC 30860130 to Yin-He Zhao). The Royal Botanic Garden Edinburgh is supported by the Scottish Government Rural and Environment Research and Analysis Directorate (RERAD).

Conflict of interest statement None.

References

- Angenent GC, Franken J, Busscher M, van Dijken A, van Went JL, Dons HJM, Van Tunen AJ (1995) A novel class of MADS box genes is involved in ovule development in petunia. *Plant Cell* 7:1569–1582
- Angiosperm Phylogeny Group (2003) An update of angiosperm phylogeny group classification for the orders and families of flowering plants: APG II. *Bot J Linnean Soc* 141:399–436
- Becker A, Theissen G (2003) The major clades of MADS-box genes and their role in the development and evolution of flowering plants. *Mol Phylogenet Evol* 29:464–489
- Becker A, Kaufmann K, Freialdenhoven A, Vincent C, Li MA, Saedler H, Theissen G (2002) A novel MADS-box gene subfamily with a sister-group relationship to class B floral homeotic genes. *Mol Gen Genom* 266:942–950
- Bowman JL, Smyth DR, Meyerowitz EM (1989) Genes directing flower development in *Arabidopsis*. *Plant Cell* 1:37–52
- Bowman JL, Drews GN, Meyerowitz EM (1991) Genetic interactions among floral homeotic genes of *Arabidopsis*. *Development* 112:1–20
- Buzgo M, Soltis SP, Soltis DE (2004) Floral developmental morphology of *Amborella trichopoda* (Amborellaceae). *Int J Plant Sci* 165:925–947
- Chen XM, Riechmann JL, Jia DX, Meyerowitz E (2000) Minimal regions in the *Arabidopsis PISTILLATA* promoter responsive to the *APETALA3/PISTILLATA* feedback control do not contain a CArG box. *Sex Plant Reprod* 13:85–94
- Coen ES, Meyerowitz EM (1991) The war of whorls: genetic interactions controlling flower development. *Nature* 353:31–37
- de Folter S, Shchennikova AV, Franken J, Busscher M, Baskar R, Grossniklaus U, Angenent GC, Immink RGH (2006) A B-sister MADS-box gene involved in ovule and seed development in petunia and *Arabidopsis*. *Plant J* 47:934–946
- Ditta G, Pinyopich A, Robles P, Pelaz S, Yanofsky FM (2004) The *SEP4* gene of *Arabidopsis thaliana* functions in floral organ and meristem identity. *Curr Biol* 14:1935–1940
- Drews GN, Bowman JL, Meyerowitz EM (1991) Negative regulation of the *Arabidopsis* homeotic gene *AGAMOUS* by the *APETALA2* product. *Cell* 65:991–1002
- Egea-Cortines M, Saedler H, Sommer H (1999) Ternary complex formation between the MADS-box proteins *SQUAMOSA*, *DEFICIENS* and *GLOBOSA* is involved in the control of floral architecture in *Antirrhinum majus*. *EMBO J* 18:5370–5379
- Ferrario S, Immink GH, Shchennikova A, Busscher-Lange J, Angenent GC (2003) The MADS box gene *FBP2* is required for *SEPALLATA* function in *Petunia*. *Plant Cell* 15:914–925
- Goto K, Meyerowitz EM (1994) Function and regulation of the *Arabidopsis* floral homeotic gene *PISTILLATA*. *Gene Dev* 8:1548–1560
- Honma T, Goto KJ (2001) Complexes of MADS-box proteins are sufficient to convert leaves into floral organ. *Nature* 409:525–529
- Irish VF (2003) The evolution of floral homeotic gene function. *BioEssays* 25:637–646
- Jack T, Brockman LL, Meyerowitz EM (1992) The homeotic gene *APETALA3* of *Arabidopsis thaliana* encodes a MADS box and is expressed in petals and stamens. *Cell* 68:683–697
- Jaramillo MA, Kramer EM (2004) *APETALA3* and *PISTILLATA* homologs exhibit novel expression patterns in the unique perianth of *Aristolochia* (Aristolochiaceae). *Evol Dev* 6:449–458
- Ji SJ, Lu YC, Feng JX, Wei G, Li J, Shi YH, Fu Q, Liu D, Luo JC, Zhu YX (2003) Isolation and analyses of genes preferentially expressed during early cotton fiber development by subtractive PCR and cDNA array. *Nucleic Acids Res* 31:2534–2543
- Jia JP, Fu JJ, Zheng J, Zhou X, Huai JL, Wang JH, Wang M, Zhang Y, Chen XP, Zhang JP, Zhao JF, Su Z, Lv YP, Wang GY (2006) Annotation and expression profile analysis of 2073 full-length cDNAs from stress-induced maize (*Zea mays* L.) seedlings. *Plant J* 48:710–727
- Kim ST, Koh J, Yoo MJ, Kong HZ, Hu Y, Ma H, Soltis PS, Soltis DE (2005) Expression of floral MADS-box genes in basal angiosperms: implications for the evolution of floral regulators. *Plant J* 43:724–744
- Kramer EM, Irish VF (1999) Evolution of genetic mechanisms controlling petal development. *Nature* 399:144–148
- Kramer EM, Irish VF (2000) Evolution of the petal and stamen developmental programs: evidence from comparative studies of the lower eudicots and basal angiosperms. *Int J Plant Sci* 161:29–40
- Kramer EM, Dorit RL, Irish VF (1998) Molecular evolution of genes controlling petal and stamen development: duplication and divergence within the *APETALA3* and *PISTILLATA* MADS-box gene lineages. *Genetics* 149:765–783
- Kramer EM, Di Stilio VS, Schluter PM (2003) Complex patterns of gene duplication in the *APETALA3* and *PISTILLATA* lineages of the Ranunculaceae. *Int J Plant Sci* 164:1–11
- Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* 25:402–408
- Ma H, dePamphilis CW (2000) The ABCs of floral evolution. *Cell* 101:5–8
- Mathews S, Donoghue MJ (1999) The root of angiosperm phylogeny inferred from duplicate phytochrome genes. *Science* 286:947–950

- Meyerowitz EM, Smyth DR, Bowman JL (1989) Abnormal flowers and pattern formation in floral development. *Development* 106:209–217
- Pelaz S, Ditta G, Baumann E, Wisman E, Yanofsky MF (2000) B and C floral organ identity functions require *SEPALLATA* MADS-box genes. *Nature* 405:200–202
- Pelaz S, Tapia-Lopez R, Alvarez-Buylla ER, Yanofsky MF (2001) Conversion of leaves into petals in *Arabidopsis*. *Curr Biol* 11:182–184
- Qiu YL, Lee JH, Bernasconi-Quadroni F, Soltis DE, Soltis PS, Zanis M, Zimmer EA, Chen ZD, Savolainen V, Chase MW (1999) The earliest angiosperms: evidence from mitochondrial, plastid and nuclear genomes. *Nature* 402:404–407
- Riechmann JL, Krizek BA, Meyerowitz EM (1996) Dimerization specificity of *Arabidopsis* MADS domain homeotic proteins, *APETALA1*, *APETALA3*, *PISTILLATA*, and *AGAMOUS*. *Proc Natl Acad Sci USA* 93:4793–4798
- Soltis PS, Soltis DE, Chase MW (1999) Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. *Nature* 402:402–404
- Soltis DE, Ma H, Frohlich MW, Soltis PS, Albert VA, Oppenheimer DG, Altman NS, dePamphilis C, Leebens-Mack J (2007a) The floral genome: an evolutionary history of gene duplication and shifting patterns of gene expression. *Trends Plant Sci* 12:358–367
- Soltis DE, Chanderbali A, Kim S, Buzgo M, Soltis PS (2007b) The ABC model and its applicability to basal angiosperms. *Ann Bot* 100:155–163
- Tamura K, Dudley J, Nei M, Kumar S (2007) *MEGA4*: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596–1599
- Theissen G (2001) Development of floral organ identity: stories from the MADS house. *Curr Opin Plant Biol* 4:75–85
- Theissen G, Becker A, Rosa AD, Kanno A, Kim JT, Munster T, Winter KW, Saedler H (2000) A short history of MADS-box genes in plants. *Plant Mol Biol* 42:115–149
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25:4876–4882
- Weigel D (1998) From floral induction to floral shape. *Curr Opin Plant Biol* 1:55–59
- Winter KU, Weiser C, Kaufmann K, Bohne A, Kirchner C, Kanno A, Saedler H, Theissen G (2002a) Evolution of class B floral homeotic proteins: Obligate heterodimerization originated from homodimerization. *Mol Biol Evol* 19:587–596
- Winter KU, Saedler H, Theissen G (2002b) On the origin of class B floral homeotic genes: Functional substitution and dominant inhibition in *Arabidopsis* by expression of an orthologue from the gymnosperm *Gnetum*. *Plant J* 31:457–475
- Yang T, Li J, Wang HX, Zeng Y (2005) A geraniol-synthase gene from *Cinnamomum tenuipilum*. *Photochemistry* 66:285–293
- Zhao YH, Wang GY, Zhang JP, Yang JP, Peng S, Gao LM, Li CY, Hu JY, Li DZ, Gao LZ (2006) Expressed sequence tags (ESTs) and phylogenetic analysis of floral genes from a paleoherb species, *Asarum caudigerum*. *Ann Bot* 89:157–163
- Zhao L, Kim YJ, Dinh TT, Chen XM (2007) miR172 regulates stem cell fate and defines the inner boundary of *APETALA3* and *PISTILLATA* expression domain in *Arabidopsis* floral meristem. *Plant J* 51:840–849