The Plant Cell, Vol. 7, 1259–1269, August 1995

Diverse Roles for MADS Box Genes in Arabidopsis Development

Steven D. Rounsley, Gary S. Ditta, and Martin F. Yanofsky[1]

Department of Biology and Center for Molecular Genetics, University of California at San Diego, La Jolla, California 92093-0116

Members of the MADS box gene family play important roles in flower development from the early step of determining the identity of floral meristems to specifying the identity of floral organ primordia later in flower development. We describe here the isolation and characterization of six additional members of this family, increasing the number of reported Arabidopsis MADS box genes to 17. All 11 members reported prior to this study are expressed in flowers, and the majority of them are floral specific. RNA expression analyses of the six genes reported here indicate that two genes, *AGL11* and *AGL13* (*AGL* for *AGAMOUS*-like), are preferentially expressed in ovules, but each has a distinct expression pattern. *AGL15* is preferentially expressed in embryos, with its onset at or before the octant stage early in embryo development. *AGL12*, *AGL14*, and *AGL17* are all preferentially expressed in root tissues and therefore represent the only characterized MADS box genes expressed in roots. Phylogenetic analyses showed that the two genes expressed in ovules are closely related to previously isolated MADS box genes, whereas the four genes showing nonfloral expression are more distantly related. Data from this and previous studies indicate that in addition to their proven role in flower development, MADS box genes are likely to play roles in many other aspects of plant development.

INTRODUCTION

In recent years, tremendous progress has been made in understanding the molecular factors involved in floral development. The analysis of floral mutants has successfully elucidated some major regulatory interactions in Arabidopsis and Antirrhinum, and thus these species have become model organisms for the study of flower development (for review, see Coen and Meyerowitz, 1991; Weigel and Meyerowitz, 1994).

The development of an Arabidopsis flower occurs in two basic steps. First, the inflorescence meristem produces floral meristems from its flanks. These floral meristems continue to be produced indefinitely in a spiral fashion. Second, each floral meristem develops into a single flower consisting of four distinct organ types arranged in four concentric whorls. These whorls contain four sepals, four petals, six stamens, and two fused carpels. This development from early floral meristems to the mature flower has been divided into distinct stages based on morphology (Smyth et al., 1990). Mutations have been identified in two different classes of genes: the meristem identity genes and the floral organ identity genes (reviewed by Coen and Meyerowitz, 1991; Bowman et al., 1993). Several examples of each class of gene have been isolated and include the meristem identity genes *APETALA1* (*AP1*), *CAULIFLOWER* (*CAL*), and *LEAFY* (*LFY*) (Mandel et al., 1992; Weigel et al., 1992; Kempin et al., 1995) and the organ identity genes *AGAMOUS* (*AG*), *APETALA2* (*AP2*), *APETALA3* (*AP3*), and *PISTILLATA* (*PI*) (Yanofsky et al., 1990; Jack et al., 1992; Goto and Meyerowitz, 1993; Jofuku et al., 1994).

The gene products from the *AP1*, *CAL*, *AG*, *AP3*, and *PI* genes share a highly conserved region called the MADS domain. This domain shows similarity with domains in known transcription factors from humans (serum response factor [SRF]) and yeast (MCM1; Norman et al., 1988; Passmore et al., 1988). In both humans and yeast, several MADS box–containing genes have been cloned (Dubois et al., 1987; Yu et al., 1992). However, the size of this gene family in Arabidopsis appears to be much larger and is thought to contain at least 20 members (Ma et al., 1991). Eleven have been previously isolated, including the five floral homeotic genes mentioned previously, and all are expressed in flowers (Ma et al., 1991; Mandel et al., 1992; Kempin et al., 1995). For the five floral homeotic genes that have been isolated, the onset of expression correlates with the proposed activity of their gene products, suggesting that expression analyses of the other members of the family may provide useful clues to their function.

The large proportion of MADS box genes that have floral mutant phenotypes makes it likely that additional MADS box genes may also be involved in flower development. We have taken the approach of isolating additional MADS box genes from Arabidopsis by homology. One advantage of this approach is that it may identify important genes that would normally be

[1] To whom correspondence should be addressed at Department of Biology, 0116, University of California at San Diego, La Jolla, CA 92093-0116.

missed in classical mutant screens, particularly in the case of genes with redundant functions. For example, the closely related *AP1* and *CAL* genes have partially redundant functions involved in meristem identity such that single gene mutations in *CAL* are phenotypically wild type (Kempin et al., 1995).

In this study, we describe the isolation and initial characterization of six additional members of the Arabidopsis MADS box gene family designated *AGL11* to *AGL15*, and *AGL17* (*AGL* for *AGAMOUS*-like), bringing the total to 17 reported genes. These analyses revealed MADS box genes with specific nonfloral expression patterns. This indicates that members of this gene family are not only involved in flower development but may play wider roles in different aspects of plant development.

RESULTS

Isolation of Six New MADS Box Genes

The 56–amino acid domain encoded by the MADS box is highly conserved among the different members of this gene family. Within this highly conserved domain, there are regions that are almost invariant between the different Arabidopsis MADS domains. We used degenerate oligonucleotides encoding the peptides MGRGKV/I and VLCDAEV to amplify intra-MADS box sequences from Arabidopsis genomic DNA by using polymerase chain reaction (PCR). After sequencing, we chose those that showed significant differences from previously cloned MADS boxes for further study. To obtain additional sequence information and to confirm that these clones represented transcribed genes rather than pseudogenes, we used reverse transcription PCR to isolate a cDNA for each of these putative genes. Using oligo(dT) and oligonucleotides based on the unique genomic sequence found in the small PCR-generated genomic clones, we amplified partial cDNA clones from total RNA from different tissues. Subsequently, these were used to isolate additional cDNA clones spanning the entire coding regions of *AGL11*, *AGL12*, and *AGL13*. These cDNA clones were then used in sequence and expression analyses.

The cDNA sequences that have been isolated for these six genes show the profile of typical plant MADS box genes (Davies and Schwarz-Sommer, 1994). Each cDNA has a single long open reading frame encoding a deduced protein of between 211 and 268 amino acids, as shown in Figure 1. The similarity with other MADS box genes is most striking in the MADS box itself, and also within the K-box, a region that codes for a structurally conserved domain thought to form a coiled-coil structure such as that found in keratin (Ma et al., 1991), from which the K-box derives its name. Figures 2A and 2B show the similarity of the deduced amino acid sequences within these two domains. The C-terminal end of the deduced proteins typically shows much less similarity between members of this family, and the same is true for the genes presented here.

Restriction Fragment Length Polymorphism Mapping Analyses

The genetic map positions of *AGL11*, *AGL12*, *AGL13*, and *AGL14* were determined by restriction fragment length polymorphism (RFLP) mapping using unique sequences from the 3′ end of each cDNA as gene-specific probes. *AGL11* mapped between

```
AGL11
   MGRGKIEIKRIENSTNRQVTFCKRRNGLLKKAYELSVLCDA
   EVALIVFSTRGRLYEYANNNIRSTIERYKKACSDSTNTSTV
   QEINAAYYQQESAKLRQQIQTIQNSNRNLMGDSLSSLSVKE
   LKQVENRLEKAISRIRSKKHELLLVEIENAQKREIELDNEN
   IYLRTKVAEVERYQQHHHQMVSGSEINAIEALASRNYFAHS
   IMTAGSGSGNGGSYSDPDKKILHLG•

AGL12
   MARGKIQLKRIENPVHRQVTFCKRRTGLLKKAKELSVLCDA
   EIGVVIFSPQGKLFELATKGTMEGMIDKYMKCTGGGRGSSS
   ATFTAQEQLQPPNLDPKDEINVLKQEIEMLQKGISYMFGGG
   DGAMNLEELLLLEKHLEYWISQIRSAKMDVMLQEIQSLRNK
   EGVLKNTNKYLLDKIEENNNSILDANFAVMETNYSYPLTMP
   SEIFQF•

AGL13
   MGRGKVEVKRIENKITRQVTFSKRKSGLLKKAYELSVLCDA
   EVSLIIFSTGGKLYEFSNVGVGRTIERYYRCKDNLLDNDTL
   EDTQGLRQEVTKLKCKYESLLRTHRNLVGEDLEGMSIKELQ
   TLERQLEGALSATRKQKTQVMMEQMEELRRKERELGDINNK
   LKLETEDHDFKGFQDLLLNPVLTAGCSTDFSLQSTHQNYIS
   DCNLGYFYRLGFNNTMSKVKDLR•

AGL14
   mgrgkiEMKRIENATSRQVTFSKRRNGLLKKAFELSVLCDAE
   VALIIFSPRGKLYEFSSSSSIPKTVERYQKRIQDLGSNHKRN
   DNSQQSKDETYGLARKIEDLEISTRKMMGEGLDASSIEELQQ
   LENQLDRSLMKIRAKKYQLLREETEKLKEKERNLIAENKMLM
   EKCEMQGRGIIGRISSSSTSSELDIDDNEMEVVTDLFIGPPE
   TRHFKKFPPSN•

AGL15
   mgrgkiEIKRIENANSRQVTFSKRRSGLLKKARELSVLCDAE
   VAVIVFSKSGKLFEYSSTGMKQTLSRYGNHQSSSASKAEEDC
   AEVDILKDQLSKLQEKHLQLQGKGLNPLTFKELQSLEQQLYH
   ALITVRERKERLLTNQLEESRLKEQRAELENETLRRQVQELR
   SFLPSFTHYVPSYIKCFAIDPKNALINHDSKCSLQNTDSDTT
   LQLGLPGEAHDRRTNEGERESPSSDSVTTNTSSETAERGDQS
   SLANSPPEAKRQRFSV•

AGL17
   mgrgkiVIQKIDDSTSRQVTFSKRRKGLIKKAKELAILCDAE
   VCLIIFSNTDKLYDFASSSVKSTIERFNTAKMEEQELMNPAS
   EVKFWQREAETLRQELHSLQENYRQLTGVELNGLSVKELQNI
   ESQLEMSLRGIRMKREQILTNEIKELTRKRNLVHHENLELSR
   KVQRIHQENVELYKKAYGTSNTNGLGHHELVDQFMNPMHRLG
   CS•
```

Figure 1. Amino Acid Sequences for Six Arabidopsis MADS Box Genes.

The complete and partially deduced amino acid sequences for *AGL11* to *AGL15* and *AGL17* are shown. See Methods for GenBank accession numbers of the cDNA sequences. Lowercase letters for the first six amino acids of the *AGL14*, *AGL15*, and *AGL17* sequences indicate the translation of the degenerate oligonucleotide sequence used in the initial isolation.

A

```
AG      GRGKIEIKRIENTTNRQVTFCKRRNGLLKKAYELSVLCDAEVALIVFSSRGRLYEY
AGL1    ****************************************VI**T*******
AGL2    ***RV*L*****KI******A**********************I**N**K***F
AGL3    ****V*L*****KI******A********************I**LI**N**K***F
AGL4    ***RV*L*****KI******A**********************S*****N**K***F
AGL5    ********************************************VI**T*******
AGL6    ***RV*M*****KI******S**********************I*****K***F
AGL8    ***RVQL*****KI******S***S******H*I**************SK*K*F**
AGL9    ***RV*L*****KI******A**********************I**N**K***F
AP1     ***RVQL*****KI******S***A******H*I**********V***HK*K*F**
CAL     ***RV*L*****KI******S***T******Q*I********S*****HK*K*F**
AP3     A****Q******Q******YS******F***H**T*****R*SI*M***SNK*H**
PI      ***********AN**V***S******V***K*IT*****K****I*A*N*KMID*
AGL11   ***********S**********************************T*******
AGL12   A****QL*****PVH*********T******K*********IGVVI**PQ*K*F*L
AGL13   ****V*V*****KIT*****S**KS*****************S**I**TG*K***F
AGL14   grgki*M*****A*S*****S*********F************I**P**K***F
AGL15   grgki******ANS*****S***S******R**********V****KS*K*F**
AGL17   grgkiV*QK*DDS*S*****S***K**I***K**AI******C**I**NTDK**DF
Cons.   grgkve-krIen--nRqVTfsKRrnGLlKKAyElsvLCDAevaliiFs-rgklyey
```

B

```
AG      QQESAKLRQQIISIQNSNRQLMGETIGSMSPKELRNLEGRLERSITRIRSKKNELLFSEIDYMQKRE
AGL1    ***AS***R**RD******HIV**SL**LNF***K*******KG*S*V********VA**E******
AGL2    YR*YL**KGRYENL*RQQ*N*L**DL*PLNS***EQ**RQ*DG*LKQV**I*TQYMLDQLSDL*NK*
AGL3    Y*DYL**KSRVEIL*H*Q*H*L**ELSE*DVN**EH**RQVDA*LRQ***T*ARSMLDQLSDLKTK*
AGL4    YR*YL**KGRYENL*RQQ*N*L**DL*PLNS***EQ**RQ*DG*LKQV*CI*TQYMLDQLSDL*GK*
AGL5    ***AS***R**RD***L**HIL**SL**LNF***K***S***KG*S*V****H*M*VA**E******
AGL6    C**VT**KSKYE*LVRT**N*L**DL*E*GV***QA**RQ**AAL*AT*QR*TQVMME*MEDLR*K*
AGL8    VL*H***KARVEVLEKNK*NF***DLDSL*L***QS**HQ*DAA*KS***R**QAM*ES*SAL**KD
AGL9    ***YL**KERYDAL*RTQ*N*L**DL*PL*T***ES**RQ*DS*LKQ**ALRTQFMLDQLNDL*SKL
AP1     SM*YNR*KAK*ELLERNQ*HYL**DLQA******Q***QQ*DTALKH**TR**Q*MYES*NEL**K*
CAL     SM*YSR*KAK*ELLERNQ*HYL**ELEP**L*D*Q***QQ**TALKH***R**Q*MNESLNHL*RK*
AP3     *ETKR**LETNRNLRTQIK*RL**CLDELDIQ***R**DEM*NTFKLV*ER*FKS*GNQ*ETTK*KN
PI      SN*IDRIKKEND*L*LEL*H*K**D*Q*LNL*N*MAV*HAI*HGLDKV*DHQM*I*....ISKRRN*
AGL11   **********QT******N***DSLS*L*V***KQV*N***KA*S******H***LV**ENA****
AGL12   KD*INV*K*E*EML*KGISYMF*GGD*A*NLE**LL**KH**YW*SQ***A*MDVMLQ**QSLRNK*
AGL13   R**VT**KCKYE*LLRTH*N*V**DLEG**I***QT**RQ**GALSAT*KQ*TQVMMEQMEELRRK*
AGL14   KD*TYG*ARK*EDLEI*T*KM***GLDAS*IE**QQ**NQ*D**LMK**A**YQ**RE*TEKLKEK*
AGL15   CA*VDI*KD*LSKL*EKHL**Q*KGLNPLTF***QS**QQ*YHALITV*ER*ER**TNQLEESRLK*
AGL17   *R*AET***ELH*L*ENY***T*VELNGL*V***Q*I*SQ**M*LRG**M*REQI*TN**KELTRKR
Cons.   --e--klk---e-lq---r-l-Ge-l--l--keL--lE-qle--l--iR--k-q-m------lq-ke
```

Figure 2. Comparison of MADS and K Domains of All Arabidopsis MADS Box Genes Reported to Date.

(A) Alignment of deduced amino acid sequences for the MADS domain.

(B) Alignment of deduced amino acid sequences for the K domain.

Asterisks indicate identity with the deduced *AG* sequence. In the consensus (Cons.) sequence, uppercase letters represent residues present in all listed sequences, and lowercase letters represent residues present in more than 50% of the listed sequences. Dashes indicate residues that are more variable.

the morphological marker *brevipedicellus* (*bp*) and the RFLP marker 518 on chromosome 4 (Hauge et al., 1993). *AGL12* mapped on chromosome 1 between the morphological marker *apetala1* (*ap1*) and the RFLP marker λAt237. *AGL13* mapped between the RFLP markers λAt424 and r807.2 on chromosome 3, and *AGL14* mapped on chromosome 4 between the RFLP markers λAt456 and λAt326. We were unable to find polymorphisms for *AGL15* and *AGL17*. These data, together with those previously published (Ma et al., 1991; Heck et al., 1995; Savidge et al., 1995; M.A. Mandel and M.F. Yanofsky, manuscript in preparation), are presented in Figure 3 and show that MADS box genes are not clustered but are scattered throughout the Arabidopsis genome, although they most likely evolved from a common ancestral gene.

Phylogenetic Tree Analysis/Evolutionary Considerations

To predict the evolutionary history of these newly isolated genes within the gene family, we produced a multiple alignment of cDNA sequences of the Arabidopsis and Antirrhinum MADS

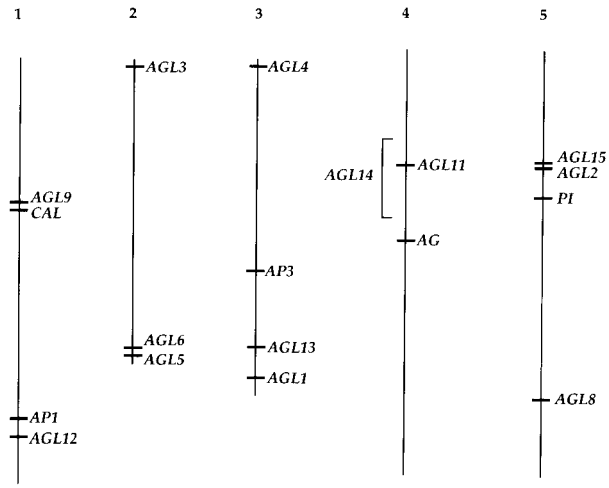

Figure 3. Map Positions of Arabidopsis MADS Box Genes.

Approximate map positions of all reported Arabidopsis MADS box genes on chromosomes 1 to 5 were determined by RFLP mapping. Relative positions have not been determined for those genes that map immediately adjacent to one another. *AGL15* was placed on this map using data from Heck et al. (1995).

box genes. This alignment was used to reconstruct a phylogenetic tree, shown in Figure 4. In a previous phylogenetic study of MADS box genes from several different plant species, we have shown that many of these genes fall into three distinct monophyletic groups (Purugganan et al., 1995). These have been referred to as the *AG* group, the *AP1/AGL9* group, and the *AP3/PI* group. *AGL11* is clearly part of the *AG* group of the tree, along with *AGL1* and *AGL5*. *AGL13* is closely related to *AGL6* within the *AP1/AGL9* group. The other four genes, *AGL12*, *AGL14*, *AGL15*, and *AGL17*, are all distinct from the previously cloned plant MADS box genes. The phylogeny showed that these four genes all diverged from an ancestral MADS box gene very early, in contrast with the genes of the *AG* group that most likely have diverged from one another relatively recently.

Expression Analyses

We performed RNA gel blot analyses using total RNA isolated from different Arabidopsis tissues for each of these genes. As shown in Figure 5, *AGL11* RNA was detected in inflorescence and developing silique tissue, whereas *AGL12*, *AGL14*, and *AGL17* transcripts were only detected in root tissue. *AGL13* transcripts were detected in inflorescence and silique tissue, with much lower levels being present in leaf and seedling tissue. *AGL15* followed a pattern similar to that of *AGL13*, although higher levels were detected in seedlings.

To begin to determine the temporal and spatial patterns of expression of some of these genes, we analyzed RNA accumulation in various tissues by in situ hybridization. Using the data from the RNA blot analyses as a starting point, we made sections from appropriate tissues of wild-type Arabidopsis (Landsberg *erecta*). We hybridized these sections with ^{35}S-labeled antisense RNA probes transcribed from the 3′ portions of the corresponding cDNAs to avoid cross-hybridization to the highly conserved MADS box.

AGL11

In situ hybridizations revealed that *AGL11* RNA accumulates only in developing ovules and associated placental tissues, as shown in Figures 6A to 6H. The onset of *AGL11* expression occurs during stage 9, at which time ovule primordia begin to arise from the placental tissue (Figures 6C and 6D). At this stage, *AGL11* RNA accumulates uniformly in the ovule primordia and the placental tissue from which they arise. At later stages of ovule development, this accumulation becomes largely restricted to the chalazal end of the ovule, the funiculus, and the outer integuments (Figures 6E and 6F). The expression at the chalazal end of the ovule is maintained beyond fertilization and into seed development (Figures 6G and 6H). No *AGL11* RNA was detected in other floral organs during earlier or later stages of flower development.

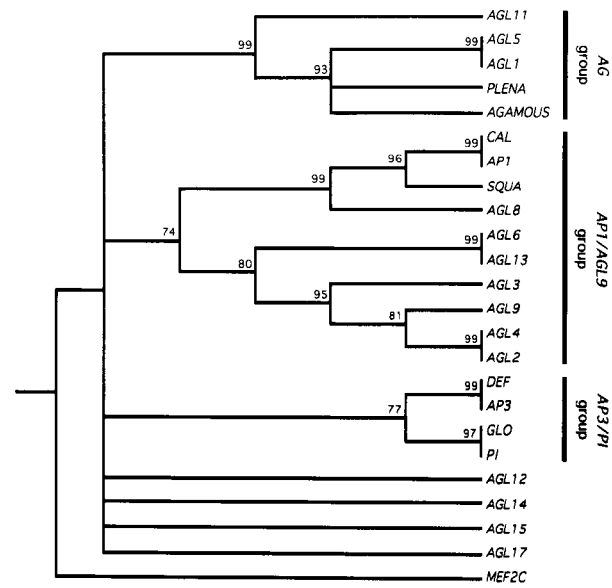

Figure 4. Phylogenetic Tree for Arabidopsis and Antirrhinum MADS Box Genes.

Shown is a neighbor–joining tree illustrating relationships between the reported MADS box genes from Arabidopsis and Antirrhinum (*PLENA*, *SQUAMOSA* [*SQUA*], *DEFICIENS* [*DEF*], and *GLOBOSA* [*GLO*]), using the mammalian MADS box gene *MEF2C* as an outgroup. The numbers show bootstrap confidence levels for each node. Nodes with less than 50% bootstrap confidence are collapsed.

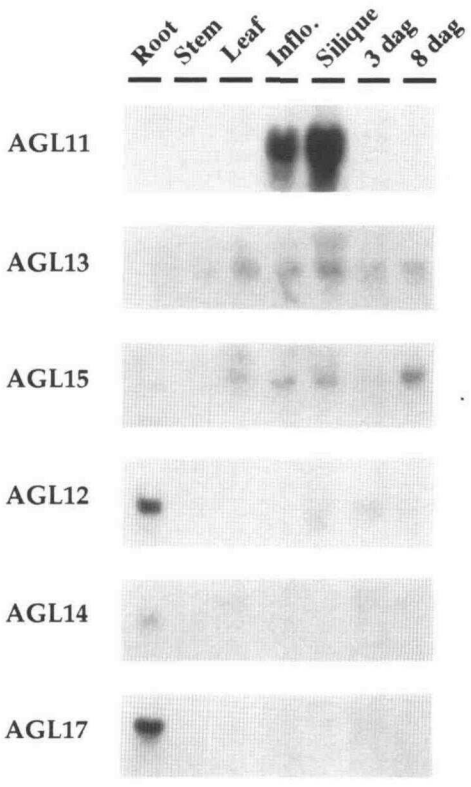

Figure 5. Expression Analyses Using RNA Gel Blot Hybridizations.

Total RNAs were used from roots of 2-week-old plate-grown plants and from stems, rosette leaves, inflorescences (Inflo.), and siliques of soil-grown plants as well as from plate-grown seedlings 3 and 8 days after germination (dag). Identical RNA gel blots were prepared for hybridization by gene-specific probes for the genes shown at left.

AGL13

The accumulation of *AGL13* RNA is restricted mainly to developing ovules, although lower levels of expression were also observed in the filaments of anthers and in the style tissue of the carpel but not in the stigmatic or transmitting tissue (Figures 6I to 6P). The onset of *AGL13* expression in ovules occurs during approximately the middle to latter part of stage 12. At this time, the integuments develop and begin to envelop the nucellus (Mansfield and Bowman, 1994). Analysis of serial sections through many carpels showed that *AGL13* appears to be expressed transiently in a small group of cells within ovules (Figures 6M to 6P). These cells appeared to stain slightly darker, perhaps indicating a difference in metabolism from other cells in the ovule. In sections through single carpels, we observed that only a subset of the ovules showed expression of *AGL13* (Figures 6I and 6J). For instance, serial sections through ovules 2 and 3 in Figures 6I and 6J show no accumulation of *AGL13* RNA, whereas ovule 1 does, suggesting that *AGL13* expression is sharply reduced at a specific stage of ovule development. After fertilization, little or no *AGL13* RNA can be detected.

AGL15

Although significant accumulation of *AGL15* RNA was not detected in flowers at any stage prior to fertilization, we observed hybridization at high levels in cells of the developing embryo, as shown in Figure 7. Early embryo tissue was present in the inflorescence RNA used for the RNA blot analysis (Figure 5). *AGL15* RNA was detected as early as the octant stage in the cells of the embryo proper but not in the suspensor cells (Figures 7A and 7B). We do not have sufficient data to determine whether expression occurs prior to this stage. As embryonic development continues, the accumulation of *AGL15* RNA remains uniform throughout the embryo (Figures 7C to 7L). Significant RNA accumulation was detected in germinating seedlings, in agreement with the RNA blot data; the seedlings showed significant expression of *AGL15* 8 days after germination (Figure 5). A lower level of RNA accumulation was also seen in sections through rosette leaves (data not shown).

DISCUSSION

We isolated six new MADS box genes from Arabidopsis, and they are designated *AGL11* to *AGL15* and *AGL17*. Each gene encodes a protein containing a MADS domain and a K-domain and thus shares the same overall organization of the previously described plant MADS box genes. This increases the number of reported Arabidopsis MADS box genes to 17. In addition to these characterized genes, our laboratory has identified seven additional MADS box genes, bringing the total number contained within the genome to at least two dozen (M.A. Mandel and M.F. Yanofsky, manuscript in preparation; S.J. Liljegren and M.F. Yanofsky, unpublished data; S.E. Gold and M.F. Yanofsky, unpublished data).

The data from RNA expression analyses for the six genes illustrate a wide range of expression profiles. Three of the genes are only expressed in root tissue, which is notable not only because no other reported MADS box genes are expressed in this tissue but also because no others are absent from floral tissues. Given the known regulatory roles of MADS box genes in flower development, it seems likely that these genes play similar roles in root development.

Transcripts of two of the genes reported here accumulate to high levels in ovules but in very different patterns. *AGL11* is expressed from the earliest visible stage of ovule development and is maintained after pollination through late seed development. The highest levels of transcript are localized toward the chalazal end of the ovule at these later stages, with lower levels throughout the funiculus and the outer integuments. The

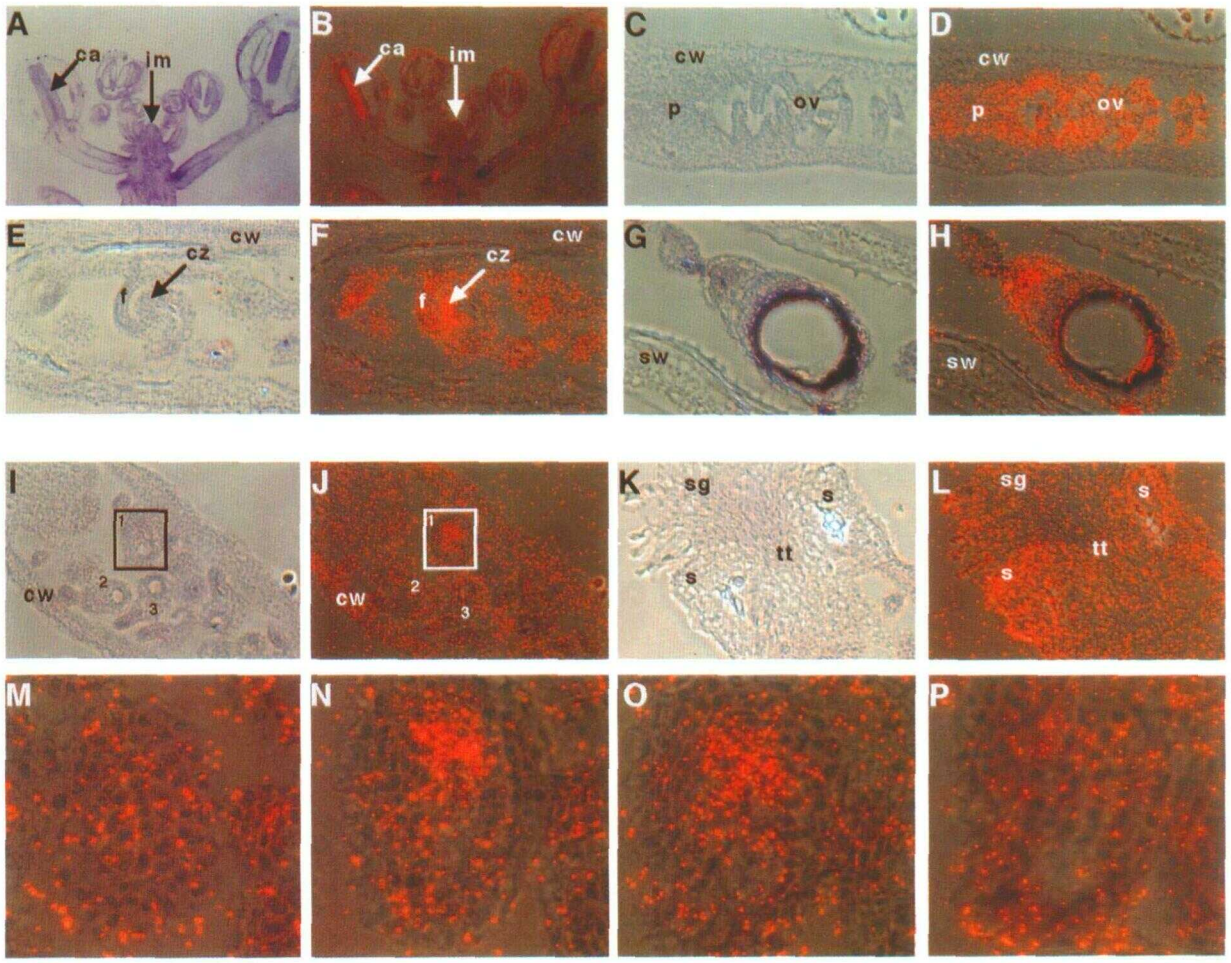

Figure 6. *AGL11* and *AGL13* Expression in Wild-Type Tissues.

In **(A)** to **(H)**, in situ hybridizations of wild-type plant tissues hybridized with the *AGL11* antisense RNA probe are shown. In **(I)** to **(P)**, the *AGL13* antisense RNA probe was used. **(A)**, **(C)**, **(E)**, **(G)**, **(I)**, and **(K)** are bright-field views of longitudinal tissue sections. **(B)**, **(D)**, **(F)**, **(H)**, **(J)**, and **(L)** to **(P)** are bright/dark-field double exposures.

(A) to **(H)** *AGL11* expression analysis. **(A)** and **(B)** show an inflorescence with flowers at several stages of development; **(C)** and **(D)** show a carpel of a stage 10 flower; **(E)** and **(F)** show a carpel of a stage 13 flower; **(G)** and **(H)** show a developing seed in a silique.

(I) to **(P)** *AGL13* expression analysis. **(I)** and **(J)** show a carpel of a stage 12 flower; **(K)** and **(L)** show a stigma and style of the same carpel; **(M)** to **(P)** are high magnifications of serial sections through the boxed ovule labeled 1 shown in **(I)** and **(J)**. **(O)** is the same section as shown in **(I)** and **(J)**. Serial sections through the ovules labeled 2 and 3 in **(I)** and **(J)** show no hybridization.

ca, carpel; cw, carpel wall; cz, chalazal region; f, funiculus; im, inflorescence meristem; ov, ovule; p, placental tissue; s, style; sg, stigmatic tissue; sw, silique wall; tt, transmitting tissue.

other gene, *AGL13*, is expressed in a more restricted manner both spatially and temporally. Its expression within the ovule is restricted to a small group of cells during a short period of time just prior to production of a visibly distinct endothelium around the embryo sac. Recent studies have identified a number of genetic loci involved in ovule development (Modrusan et al., 1994; Ray et al., 1994). It is likely that *AGL11* and *AGL13* regulate ovule-specific gene expression, perhaps through an interaction with one or more of these genes.

Although no significant accumulation of *AGL15* RNA was observed during any stage of flower development, high levels were detected shortly after fertilization. Strong and uniform *AGL15* expression was evident as early as the eight-cell stage of the embryo and remained uniform throughout development of the embryo proper. No expression was seen in the suspensor or in other maternal tissues, such as the surrounding seed coat. These data suggest a possible role for *AGL15* in regulating early events in embryogenesis. In situ data showed that

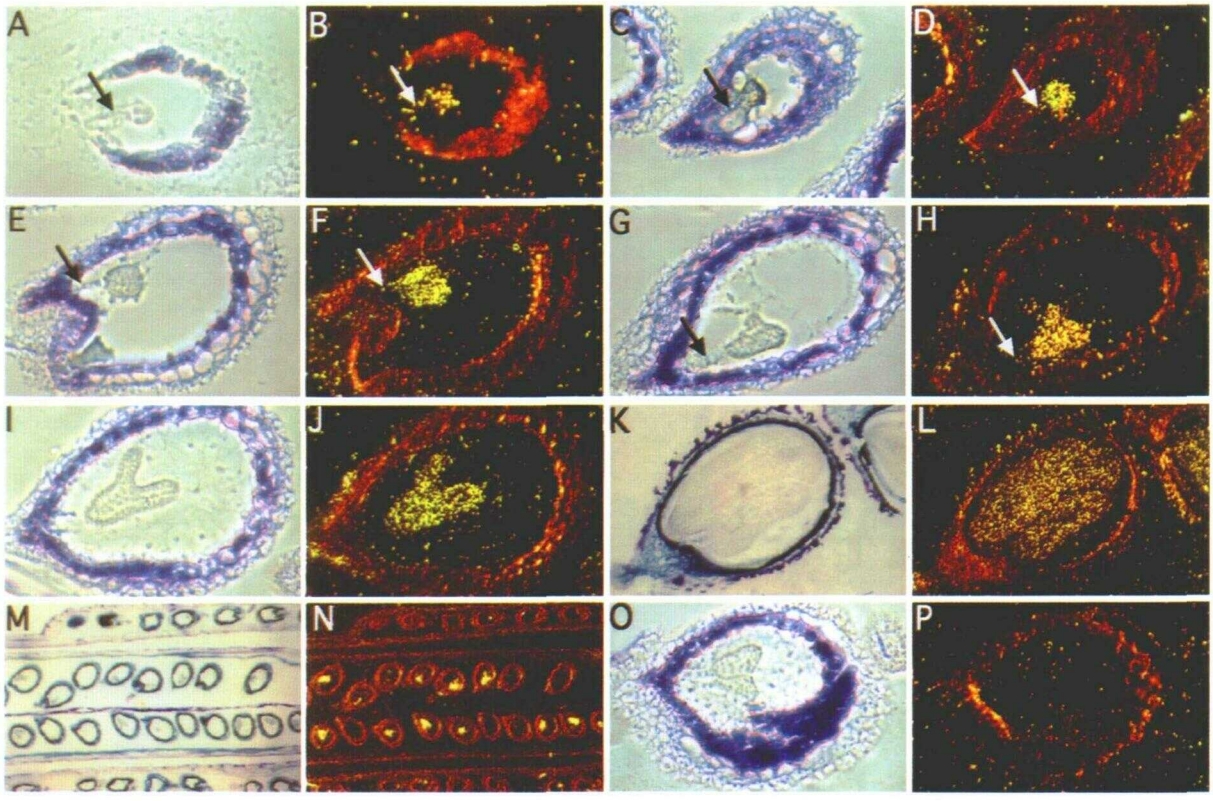

Figure 7. *AGL15* Expression in Developing Wild-Type Embryos.

In **(A)** to **(N)**, in situ hybridizations of seeds in wild-type siliques hybridized with the antisense *AGL15* RNA probe are shown. In **(O)** and **(P)**, the sense *AGL15* RNA probe was used. **(A)**, **(C)**, **(E)**, **(G)**, **(I)**, **(K)**, **(M)**, and **(O)** are bright-field views of longitudinal tissue sections. **(B)**, **(D)**, **(F)**, **(H)**, **(J)**, **(L)**, **(N)**, and **(P)** are bright/dark-field double exposures.
(A) and **(B)** An octant-stage embryo.
(C) and **(D)** An embryo at the dermatogen stage.
(E) and **(F)** A globular-stage embryo.
(G) and **(H)** A heart-stage embryo.
(I) and **(J)** An embryo at the torpedo stage.
(K) and **(L)** A mature embryo.
(M) and **(N)** A low-magnification view of a longitudinal section through adjacent siliques.
(O) and **(P)** A mature embryo.
The arrows indicate the position of the suspensor cells.

AGL15 was also expressed at a low level in rosette leaves but that no significant expression was seen after bolting had occurred (data not shown). *AGL15* may therefore be important for vegetative or nonreproductive phases of the plant life cycle, although more detailed analyses of the late expression of *AGL15* are required to investigate this possibility.

The expression patterns of the genes reported here show some correlation with the major clade groupings in the phylogenetic tree (Figure 4). The three major groups consisting of closely related genes, and a collection of more divergent genes, termed the "orphan" genes, have been identified in a more extensive survey of plant MADS box genes (Purugganan et al., 1995). The orphan genes from Arabidopsis are *AGL12*, *AGL14*, *AGL15*, and *AGL17* and are the same four genes that are shown here to have nonfloral expression patterns. These genes most likely play roles elsewhere in plant development, and the phylogenetic data support the distinct evolutionary history of these genes. In contrast, all members of the three major clades are expressed in flowers. There is good correlation between their specific floral expression patterns and the clade to which they belong. The *AGAMOUS* clade contains *AGL1*, *AGL5*, and *AGL11*, which are all expressed in tissues within the carpel (Ma et al., 1991; Savidge et al., 1995), as well as *AG*, which is necessary for carpel development (Bowman et al., 1989; Yanofsky et al., 1990). In addition, many of the genes in the *AP1/AGL9* group are expressed very early in flower development. These include the two meristem identity genes *AP1* and *CAL* (Mandel et al., 1992; Gustafson-Brown et al., 1994; Kempin et al., 1995) and

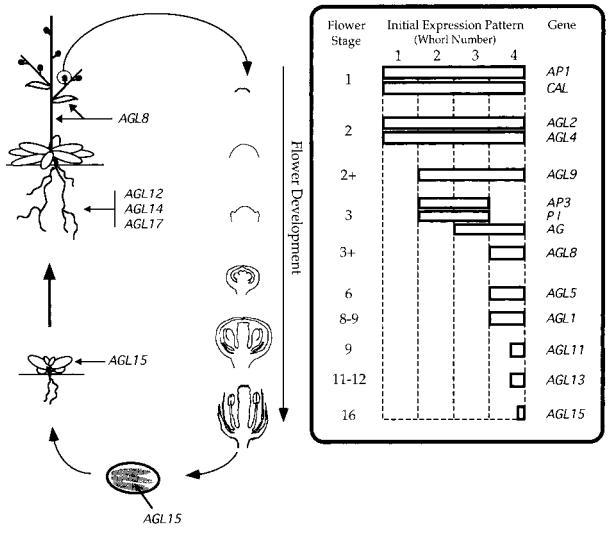

Figure 8. Spatial and Temporal RNA Expression Patterns for Arabidopsis MADS Box Genes.

Gene names indicate expression of that gene in the tissue indicated. At left, the plant's life cycle is represented, beginning with seed development at the bottom. The germinated seedling and the mature plant are also shown. The line drawings in the middle represent particular stages of flower development, corresponding to the stage number shown inside the box to the right (Smyth et al., 1990; Mansfield and Bowman, 1994). This box contains a schematic representation of MADS box gene expression during flower development. For each gene, the spatial pattern of the initial expression is shown by a horizontal bar across the four columns representing the whorls of the flower. The timing of the onset of this expression is indicated by the stage number to the left. During stage 3, *PI* is transiently expressed in the fourth whorl (Goto and Meyerowitz, 1993).

AGL2, *AGL4*, and *AGL9*, which are all expressed in young floral primordia (Flanagan and Ma, 1994; Savidge et al., 1995; M.A. Mandel and M.F. Yanofsky, manuscript in preparation). In addition, two genes in this clade that are closely related to one another, *AGL13* and *AGL6*, are both expressed at high levels in specific cell types within developing ovules (T. Araki and M.F. Yanofsky, manuscript in preparation).

The results of this and previous studies illustrate that MADS box genes play varied roles in flower development. In Arabidopsis, the analysis of floral mutant phenotypes demonstrates the involvement of MADS box genes in specifying floral meristem identity (*AP1* and *CAL*) and in determining the identity of different floral organ primordia formed by the floral meristem (*AG*, *AP3*, *PI*, and *AP1*). Expression analyses of these and other members of the family suggest that a cascade of MADS box gene activity occurs as flower development progresses (Figure 8). *AP1* and *CAL* are expressed as soon as floral meristems are visible on the flanks of the inflorescence meristem (stage 1), and this expression is uniform over the flower primordium (Gustafson-Brown et al., 1994; Kempin et al., 1995). Shortly thereafter (stage 2), *AGL2* and *AGL4* expression commences and is similarly uniform (Flanagan and Ma, 1994; Savidge et al., 1995). At this time, *AGL9* RNA also begins to accumulate in the young floral primordia (M.A. Mandel and M.F. Yanofsky, manuscript in preparation). The onset of *AGL2*, *AGL4*, and *AGL9* expression suggests a possible role for these genes in mediating between the meristem and organ identity genes. Previous studies on the *AGL9* orthologs from tomato and petunia, *TM5* and *fbp2*, respectively, support this idea. When the expression of these genes is perturbed, the resulting plants have both meristem and floral organ defects (Angenent et al., 1994; Pnueli et al., 1994). In addition, cosuppression of *fbp2* leads to inactivation of the *AG* ortholog *fbp6*, suggesting that *AGL9* may be necessary for activation of *AG*.

The cascade of MADS box gene activity continues as the organ identity genes are transcribed in whorls 2 and 3 (*AP3* and *PI*) and 3 and 4 (*AG*), respectively, during stage 3 (Yanofsky et al., 1990; Jack et al., 1992; Goto and Meyerowitz, 1993). By stage 6, the carpel-specific gene *AGL5* begins to be expressed, and this expression requires the prior activity of *AG* (Savidge et al., 1995). At progressively later stages of flower development, *AGL1*, *AGL11*, and *AGL13* RNAs begin to accumulate uniformly in developing ovules or in specific cell types within these ovules (Ma et al., 1991). Thus, the results of this and previous studies show that MADS box genes are expressed throughout all stages of flower development and may be activated in a cascade-like manner.

The expression of MADS box genes is not, however, restricted to flower development. For example, shortly after pollination of the Arabidopsis flower, *AGL15* RNA begins to accumulate in the developing embryo, and this expression is maintained throughout embryo development and seedling germination. Other MADS box genes are also expressed in nonfloral tissue. *AGL8* is expressed in stems and leaves as well as in developing flowers (M.A. Mandel and M.F. Yanofsky, manuscript in preparation), and three of the genes reported here, *AGL12*, *AGL14*, and *AGL17*, are preferentially expressed in root tissues.

In summary, these studies suggest that MADS box genes play diverse roles in plant development in addition to their well-documented roles in regulating early events in flower development. Although ultimately it will prove necessary to identify mutant alleles for each of these genes, the data presented here form the basis for future efforts designed to determine the myriad of functions carried out by members of the plant MADS box gene family.

METHODS

Polymerase Chain Reaction Amplification

Genomic DNA from the Landsberg *erecta* ecotype of *Arabidopsis thaliana* was used as the template for polymerase chain reaction (PCR) amplification of MADS boxes using the degenerate oligonucleotides MADS-1, 5′-CGGAATTCATGGGN(AC)GNGGNAA(AG)(AG)T-3′, and MADS-43, 5′-CGGGATCCIAC(TC)TCIGC(AG)TC(AG)CAIA(AG)IAC-3′,

where N represents A, G, C, or T and I represents inosine. Using DNA sequences from these PCR products, oligonucleotides were designed to amplify cDNAs containing novel MADS box sequences by reverse transcription PCR using oligo(dT) as the 3′ primer. Total RNAs from various Landsberg *erecta* tissues were used as templates for reverse transcription. We used inflorescence tissue as the RNA source to clone *AGL11*, *AGL12*, *AGL13*, and *AGL14*; leaf tissue for *AGL15* (*AGL* for *AGAMOUS*-like); and root tissue for *AGL17*. Each cDNA was independently amplified, cloned, and sequenced at least twice to control for PCR-induced mutations. Conditions for the PCR amplification with the degenerate oligonucleotides are as follows: initial denaturation at 94°C for 5 min; five cycles of 94°C for 30 sec, 37°C for 2 min, and 72°C for 1 min with 1-min ramping time between each temperature; 35 cycles of 94°C for 1 min, 55°C for 1 min, and 72°C for 90 sec. The final cycle was followed by a final extension of 5 min at 72°C. Reverse transcription PCR had the following cycles: initial denaturation for 5 min at 94°C, followed by 35 cycles at 94°C for 1 min, 55°C for 1 min, and 72°C for 90 sec, with a final extension time of 5 min at 72°C.

Cloning and Sequencing

PCR products were cloned into the pCRII vector (Invitrogen, San Diego, CA) and subcloned into the pGEM series of vectors (Promega) for sequencing. Sequencing was performed using the Sequenase version 2.0 kit (U.S. Biochemical Corp.) according to the manufacturer's protocol. DNA and predicted protein sequences were analyzed using the MacVector program (International Biotechnologies, Inc., New Haven, CT). For each of the six genes reported here, a cDNA sequence, including sequence 5′ of the MADS box where available, has been submitted to GenBank with the following accession numbers: *AGL11*, U20182; *AGL12*, U20193; *AGL13*, U20183; *AGL14*, U20184; *AGL15*, U20185; *AGL17*, U20186.

Library Screening

A probe for the partial *AGL11* cDNA was used to screen 2×10^5 plaque-forming units of a λZAPII (Stratagene) cDNA library made from poly(A)$^+$ RNA from young flowers (Weigel et al., 1992). A cDNA clone containing the entire MADS box and the likely translation start site was isolated from this screen. Genomic clones containing the entire MADS box region of *AGL12* and *AGL13* were isolated by screening 2×10^5 plaque-forming units of a genomic library made from DNA from the Columbia ecotype of Arabidopsis in λGEM-11 (Promega). This library was a kind gift from R. Davis (Stanford University, Palo Alto, CA). Sequences from these genomic clones were used to enable amplification of the entire coding region of *AGL12* and *AGL13* from RNA using reverse transcription PCR.

Restriction Fragment Length Polymorphism Mapping

Gene-specific probes were used to hybridize to parental mapping filters to identify polymorphisms between various ecotypes. (Landsberg *erecta*, Columbia, and Wassilewskija ecotypes were used.) Mapping filters were prepared with the appropriate restriction enzyme and probed with the gene-specific probe. A probe synthesized from an *AGL11* cDNA clone, pSR90, was used to score an EcoRI polymorphism between Landsberg *erecta* and Columbia ecotypes on 79 individuals. A recombinant inbred mapping population from a cross between the Landsberg *erecta* and Wassilewskija ecotypes (Reiter et al., 1992) was used to map *AGL12*, *AGL13*, and *AGL14*. A probe synthesized from an *AGL12* cDNA clone, pSR102, was used to score for a polymorphism revealed by a BamHI-BglII double digest in 27 individuals. A probe synthesized from a 567-bp EcoRI-BglII fragment of pSR109, representing the 3′ end of the *AGL13* cDNA clone, was used to score for a BamHI polymorphism in 33 individuals. An *AGL14* probe from a 388-bp HindIII fragment of pSR111, containing the 3′ end of the *AGL14* cDNA, was used to score for a BamHI-HindIII polymorphism in 25 individuals. The mapping filters were scored and the data analyzed with the Mapmaker Macintosh version 2.0 software (E.I. duPont de Nemours, Wilmington, DE) as described by Reiter et al. (1992).

Phylogenetic Analysis

Nucleotide and predicted amino acid sequences were obtained for published *Antirrhinum majus* and Arabidopsis MADS box genes, in addition to those for *AGL8* and *AGL9* (M.A. Mandel and M.F. Yanofsky, manuscript in preparation). These sequences were then used in conjunction with the sequences for *AGL11* to *AGL15* and *AGL17* to produce both amino acid and nucleotide alignments with the PILEUP program of the GCG package (Genetics Computer Group, Madison, WI). The nucleotide alignment was visually refined using the amino acid alignment as a guide. Distance calculations were performed using the Tajima-Nei algorithm of the Molecular Evolutionary Genetic Analyses software package version 1.1 (Institute of Molecular Evolutionary Genetics, Pennsylvania State University, University Park, PA). First and second codon positions were used in the distance calculations, and the tree was constructed from the results of 100 bootstrap replicates using the neighbor–joining method with the mammalian MADS box gene *MEF2C*, which was used as an outgroup.

RNA Isolation and Blot Analysis

Total RNAs from various Landsberg *erecta* plant tissues were isolated using the method of Crawford et al. (1986) and size fractionated by gel electrophoresis on formaldehyde agarose gels, transferred to Hybond-N nylon membranes (Amersham), and hybridized according to standard procedures with ^{32}P-radiolabeled probes. RNA samples were quantified using spectrophotometric methods prior to electrophoresis, and RNA gels were visualized by ethidium bromide staining prior to blotting. Probes used were gene-specific and included regions of the cDNA clones 3′ of the MADS box and 5′ of the poly(A) tail. A probe for the 3′ portion of *AGL11* was synthesized from a 321-bp HindIII fragment of pSR91, a clone consisting of the *AGL11* cDNA as an 820-bp EcoRI fragment in the pCRII vector. The 3′ probe for *AGL12* was synthesized from a 266-bp EcoRI-XbaI fragment of pSR102, a clone consisting of the *AGL12* cDNA as an 810-bp EcoRI fragment in the pCRII vector. The 3′ *AGL13* probe was synthesized from a 567-bp EcoRI-BglII fragment of pSR109, a clone containing the *AGL13* cDNA as an 850-bp EcoRI fragment in the pCRII vector. The 3′ *AGL14* probe was synthesized from a 388-bp HindIII fragment of pSR111, which is the *AGL14* cDNA inserted as a 710-bp EcoRI fragment into the pCRII vector. The *AGL15* probe was synthesized from a 651-bp HindIII fragment of pSR142, which is the *AGL15* cDNA inserted into the pCRII vector as an 890-bp EcoRI fragment. The *AGL17* probe was synthesized from the 452-bp EcoRI insert of SR192, which is a subclone of the 3′ portion of the *AGL17* cDNA.

In Situ Hybridization

Fixation of tissue and hybridization conditions were the same as previously described (Drews et al., 1991), with minor modifications. The *AGL11* ^{35}S-labeled antisense mRNA was synthesized with the SP6 RNA polymerase from an EcoRI-digested pSR101 template. The *AGL13* ^{35}S-labeled antisense mRNA was synthesized with the T7 RNA polymerase from a BglII-digested pSR176 template, and the *AGL15* ^{35}S-labeled antisense mRNA was synthesized with the SP6 RNA polymerase from an EcoRI-digested pSR150 template. Probes were used at a final concentration of 2.5 x 10^7 cpm/mL.

Image Processing

Photographic slides were scanned and digitized using a Nikon Coolscan slide scanner (Nikon Inc., Melville, NY). Brightness, contrast, and color balance were adjusted using Adobe Photoshop 3.0 (Adobe Systems Inc., Mountain View, CA). Composite figures were printed using a Tektronix Phaser IIsdx dye sublimation printer (Tektronix Inc, Wilsonville, OR).

ACKNOWLEDGMENTS

We thank Steve Strauss for providing the degenerate MADS box primers and Chuck Gasser for helpful discussions on the ovule-specific gene expression data. We also thank Greg Heck and Donna Fernandez for sharing unpublished mapping data for *AGL15* and M. Alejandra Mandel, Takashi Araki, and Beth Savidge for sharing results prior to publication. We thank Bob Schmidt, Detlef Weigel, and members of our laboratory for critical reading of the manuscript. This work was supported by a grant from the National Science Foundation (Grant No. DCB-9018749) to M.F.Y.

Received February 24, 1995; accepted April 27, 1995.

REFERENCES

Angenent, G., Franken, J., Busscher, M., Weiss, D., and van Tunen, A. (1994). Co-suppression of the petunia homeotic gene *fbp2* affects the identity of the generative meristem. Plant J. **5,** 33–44.

Bowman, J.L., Smyth, D.R., and Meyerowitz, E.M. (1989). Genes directing flower development in *Arabidopsis.* Plant Cell **1,** 37–52.

Bowman, J.L., Alvarez, J., Weigel, D., Meyerowitz, E.M., and Smyth, D.R. (1993). Control of flower development in *Arabidopsis thaliana* by *APETALA1* and interacting genes. Development **119,** 721–743.

Coen, E.S., and Meyerowitz, E.M. (1991). The war of the whorls: Genetic interactions controlling flower development. Nature **353,** 31–37.

Crawford, N.M., Campbell, W.H., and Davis, R.H. (1986). Nitrate reductase from squash: cDNA cloning and nitrate regulation. Proc. Natl. Acad. Sci. USA **83,** 8073–8076.

Davies, B., and Schwarz-Sommer, Z. (1994). Control of floral organ identity by homeotic MADS-box transcription factors. Results Probl. Cell Differ. **20,** 235–258.

Drews, G.N., Bowman, J.L., and Meyerowitz, E.M. (1991). Negative regulation of the Arabidopsis homeotic gene *AGAMOUS* by the *APETALA2* product. Cell **65,** 991–1002.

Dubois, E., Bercy, J., Descamps, F., and Messenguy, F. (1987). Characterization of two new genes essential for vegetative growth in *Saccharomyces cerevisiae*: Nucleotide sequence determination and chromosome mapping. Gene **55,** 265–275.

Flanagan, C.A., and Ma, H. (1994). Spatially and temporally regulated expression of the MADS-box gene *AGL2* in wild-type and mutant *Arabidopsis* flowers. Plant Mol. Biol. **26,** 581–595.

Goto, K., and Meyerowitz, E.M. (1993). Function and regulation of the *Arabidopsis* floral homeotic gene *PISTILLATA*. Genes Dev. **8,** 1548–1560.

Gustafson-Brown, C., Savidge, B., and Yanofsky, M. (1994). Regulation of the Arabidopsis floral homeotic gene *APETALA1.* Cell **76,** 131–143.

Hauge, B.M., Hanley, S.M., Cartinhour, S., Cherry, J.M., Goodman, H.M., Koornneef, M., Stam, P., Chang, C., Kempin, S., Medrano, L., and Meyerowitz, E.M. (1993). An integrated genetic/RFLP map of the *Arabidopsis thaliana* genome. Plant J. **3,** 745–754.

Heck, G.R., Perry, S.E., Nichols, K.W., and Fernandez, D.E. (1995). AGL15, a MADS domain protein expressed in developing embryos. Plant Cell **7,** 1271–1282.

Jack, T., Brockman, L.L., and Meyerowitz, E.M. (1992). The homeotic gene *APETALA3* of Arabidopsis thaliana encodes a MADS box and is expressed in petals and stamens. Cell **68,** 683–697.

Jofuku, K.D., den Boer, B.G.W., Van Montagu, M., and Okamuro, J.K. (1994). Control of Arabidopsis flower and seed development by the homeotic gene *APETALA2.* Plant Cell **6,** 1211–1225.

Kempin, S.A., Savidge, B., and Yanofsky, M.F. (1995). Molecular basis of the *cauliflower* phenotype in *Arabidopsis.* Science **267,** 522–525.

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.

Mandel, M.A., Gustafson-Brown, C., Savidge, B., and Yanofsky, M.F. (1992). Molecular characterization of the *Arabidopsis* floral homeotic gene *APETALA1.* Nature **360,** 273–277.

Mansfield, S.G., and Bowman, J.L. (1994). Ovules: Introduction. In *Arabidopsis*: An Atlas of Morphology and Development, J.L. Bowman, ed (New York: Springer-Verlag), pp. 299–301.

Modrusan, Z., Reiser, L., Feldmann, K.A., Fischer, R.L., and Haughn, G.W. (1994). Homeotic transformations of ovules into carpel-like structures in Arabidopsis. Plant Cell **6,** 333–349.

Norman, C., Runswick, M., Pollock, R., and Treisman, R. (1988). Isolation and properties of cDNA clones encoding SRF, a transcription factor that binds to the *c-fos* serum response element. Cell **55,** 989–1003.

Passmore, S., Maine, G.T., Elble, R., Christ, C., and Tye, B.K. (1988). A *Saccharomyces cerevisiae* protein involved in plasmid maintenance is necessary for mating in MATα cells. J. Mol. Biol. **204,** 593–606.

Pnueli, L., Hareven, D., Broday, L., Hurwitz, C., and Lifschitz, E. (1994). The *TM5* MADS box gene mediates organ differentiation in the three inner whorls of tomato flowers. Plant Cell **6,** 175–186.

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.

Ray, A., Robinson-Beers, K., Ray, S., Baker, S.C., Lang, J.D., Preuss, D., Milligan, S.B., and Gasser, C.S. (1994). *Arabidopsis* floral

homeotic gene *BELL* (*BEL1*) controls ovule development through negative regulation of *AGAMOUS* gene (*AG*). Proc. Natl. Acad. Sci. USA **91**, 5761–5765.

Reiter, R.S., Williams, J.G.K., Feldmann, K.A., Rafalski, J.A., Tingey, S.V., and Scolnik, P.A. (1992). Global and local genome mapping in *Arabidopsis thaliana* by using recombinant inbred lines and random amplified polymorphic DNAs. Proc. Natl. Acad. Sci. USA **89**, 1477–1481.

Savidge, B., Rounsley, S.D., and Yanofsky, M.F. (1995). Temporal relationships between the transcription of two Arabidopsis MADS box genes and the floral organ identity genes. Plant Cell **7**, 721–733.

Smyth, D.R., Bowman, J.L., and Meyerowitz, E.M. (1990). Early flower development in *Arabidopsis*. Plant Cell **2**, 755–767.

Weigel, D., and Meyerowitz, E.M. (1994). The ABCs of floral homeotic genes. Cell **78**, 203–209.

Weigel, D., Alvarez, J., Smyth, D., Yanofsky, M., and Meyerowitz, E. (1992). LEAFY controls floral meristem identity in Arabidopsis. Cell **69**, 843–859.

Yanofsky, M.F., Ma, H., Bowman, J.L., Drews, G.N., Feldmann, K.A., and Meyerowitz, E.M. (1990). The protein encoded by the *Arabidopsis* homeotic gene *agamous* resembles transcription factors. Nature **346**, 35–39.

Yu, Y., Breitbart, R., Smoot, L., Lee, Y., Mahdavi, V., and Nadal-Ginard, B. (1992). Human myocyte-specific enhancer factor 2 comprises a group of tissue-restricted MADS box transcription factors. Genes Dev. **6**, 1783–1798.