The protein encoded by the *Arabidopsis* homeotic gene *agamous* resembles transcription factors

Martin F. Yanofsky*, Hong Ma*, John L. Bowman, Gary N. Drews, Kenneth A. Feldmann† & Elliot M. Meyerowitz‡

Division of Biology 156-29, California Institute of Technology, Pasadena, California 91125, USA
† E. I. DuPont de Nemours and Co., Wilmington, Delaware 19880, USA

Mutations in the homeotic gene *agamous* of the plant *Arabidopsis* cause the transformation of the floral sex organs. Cloning and sequence analysis of *agamous* suggest that it encodes a protein with a high degree of sequence similarity to the DNA-binding region of transcription factors from yeast and humans and to the product of a homeotic gene from *Antirrhinum*. The *agamous* gene therefore probably encodes a transcription factor that regulates genes determining stamen and carpel development in wild-type flowers.

ARABIDOPSIS THALIANA is widely used for molecular and genetic studies of many developmental processes in plants[1,2]. We have focused on the molecular basis of pattern formation in developing *Arabidopsis* flowers. Flower development remains one of the most complex processes unique to plants, and little is known about the molecular mechanisms involved. Flowers develop from a few undifferentiated cells into a structure that has several types of organs occupying precisely defined positions. *Arabidopsis* has typical mustard flowers (Fig. 1*a*) consisting of four concentric whorls of organs[3]. The first or outermost whorl consists of four green sepals. Interior to and alternating with the sepals are four white petals which make up the second whorl. The third whorl consists of six stamens, four long and two short, each of which has a filament capped by a pollen-bearing anther. The fourth and innermost whorl consists of two fused carpels capped by a short style and stigmatic papillae.

Many mutations affecting floral morphogenesis have been identified in *Arabidopsis*. Mutations in at least four genes result in homeotic transformations of floral organs[1,3-7]. In general, these homeotic mutations affect the development of adjacent whorls of floral organs. For example, the *ap2-1* mutation results in the conversion of the sepals into leaf-like organs, and in the conversion of petals into stamen-like organs[3]. Mutations in the *AP3* (*ap3-1*) or *PI* (*pi-2*) genes result in the conversion of petals into sepals and stamens into carpels. Mutations in the *agamous* (*AG*) gene result in the overall phenotype of a flower within a flower and the absence of stamens and carpels. The six stamens have been replaced in *ag* mutant flowers by six petals, and the carpels have been replaced by a new flower. Thus, *ag* flowers consist of ten petals (four normal and six in the place where stamens are normally found) inside of the four sepals, and inside the petals are again four sepals and ten petals (Fig. 1*c*). This pattern repeats itself and the resulting flower can consist of more than 70 organs. Other species of plants with similar double flower phenotypes were recognized as long ago as 2,000 years[8]. The first published report of *Arabidopsis* flowers with an *ag* mutant phenotype was more than a century ago[9], and another *Arabidopsis* mutant having similar flowers has been described by Conrad[10]. The extensively characterized[3] mutant allele, *ag-1*, was isolated after ethylmethane sulphonate (EMS) mutagenesis and was first described by Koornneef *et al.*[11]. The *AG* locus has been mapped to chromosome 4 (ref. 11).

Here we describe the molecular cloning and characterization of the *AG* gene, which was facilitated by a T-DNA insertion mutation[12]. The deduced AG protein product is similar to transcription factors from humans (SRF) and yeast (MCM1, ARG80), and to the product, DEF A, of a recently isolated homeotic gene from the snapdragon *Antirrhinum majus*.

Insertion mutant of *agamous*

Recently, a method has been developed to tag genes by insertion of the T-DNA from *Agrobacterium tumefaciens*[13]. One of the

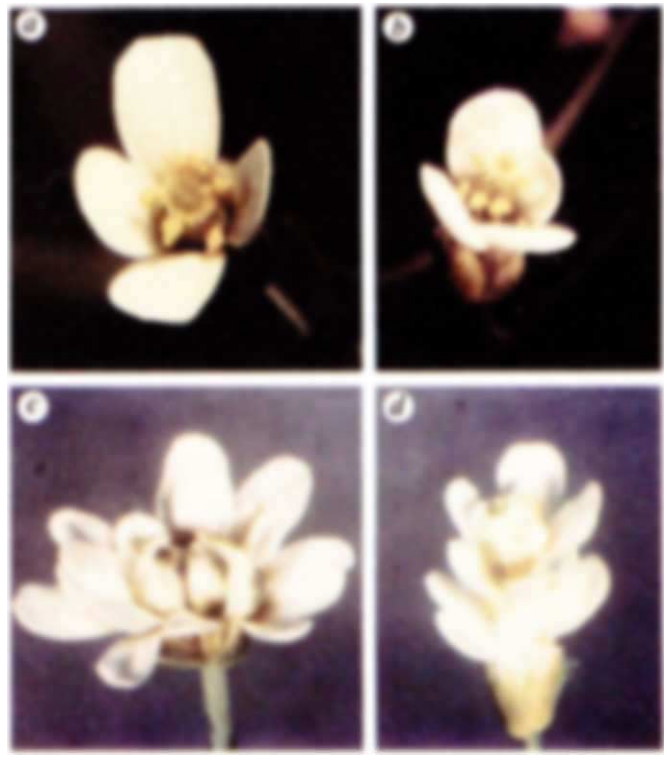
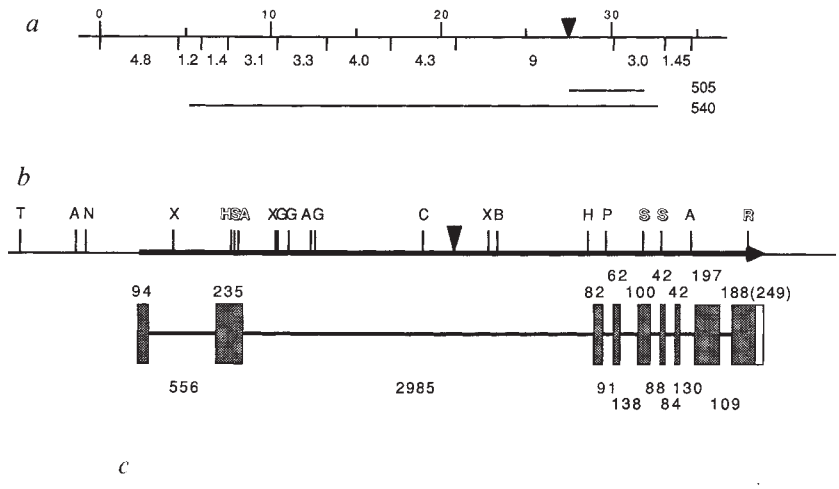


FIG. 1 Photographs of *Arabidopsis* flowers. *a*, Wild-type; *b*, *ag-2* plants transformed with pCIT540; *c*, *ag-1*; *d*, *ag-2*. Complementation of the *ag-2* mutation was tested by introducing the cosmid pCIT540 into the *Agrobacterium* strain ASE (ref. 26). This strain was used to infect leaf pieces of homozygous *ag-2* mutant plants by standard methods[27]. A regenerated plant that displayed wild-type flower morphology is shown here.

* Present addresses: Department of Biology C-016, University of California at San Diego, La Jolla, California 92093, USA (M.F.Y.) and Cold Spring Harbor Laboratory, P.O. Box 100, Cold Spring Harbor, New York 11724, USA (H.M.).
‡ To whom correspondence should be addressed.

mutants isolated by this approach (Fig. 1*d*) displays a phenotype that closely resembles that of *ag-1* mutant plants (Fig. 1*c*). Like *ag-1*, this insertional mutant lacks stamens and carpels and has the overall phenotype of a flower within a flower. The only noticeable difference between the insertional mutant and *ag-1* mutant plants is that the pedicels of inner flowers in the former are elongated, resulting in a greater separation of internal flowers. The elongated pedicel is not due to the insertional mutation, but is the result of different genetic backgrounds (M.F.Y., H.M., J.L.B., G.N.D., K.A.F. and E.M.M., unpublished results). To demonstrate whether the insertional mutation is an allele of *AG*, plants heterozygous for *ag-1* were crossed to plants heterozygous for the insertional mutation (homozygous mutant plants could not be used because they are sterile). As the *ag-1* and insertional mutants are both recessive to wild-type, and as a quarter of the resulting F1 progeny produced *ag* mutant flowers, this confirms that the two mutations are allelic. We have named the insertional allele *ag-2*.

Genetic analyses indicate that the kanamycin resistance marker encoded by the T-DNA co-segregates with the mutant flower phenotype (data not shown). This, together with the demonstration that the insertion mutation is allelic to *ag-1*, indicates that the T-DNA can be used as a molecular probe to isolate the *AG* gene. Plasmid rescue was used to isolate sequences flanking the insertion site and one of the resulting clones, pCIT505, contained plant sequences and was analysed further (Fig. 2*a*). To determine whether these sequences were derived from the *AG* genomic region, the genetic map position of the



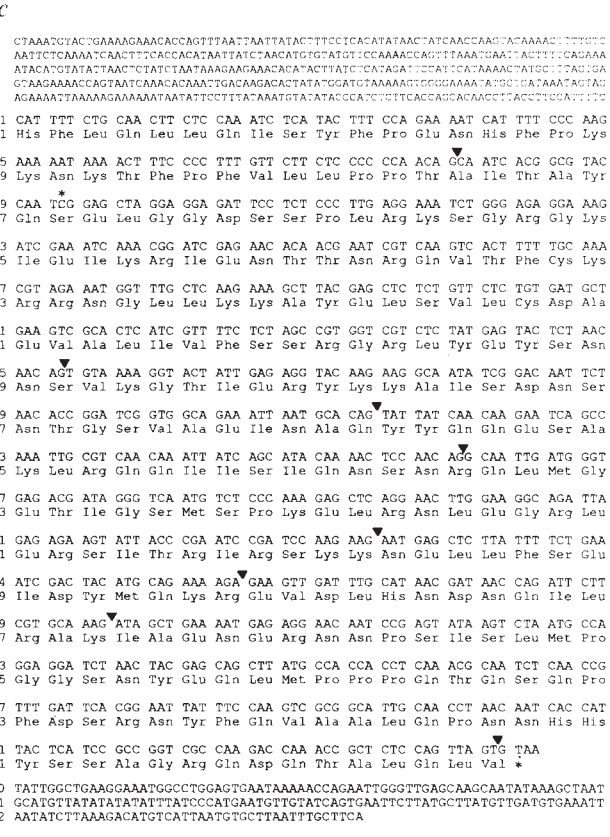


FIG. 2 *a*, Map of the *agamous* region. *Eco*RI restriction sites are shown for a 35-kb region from the Columbia ecotype flanking the *AG* gene. Size of fragments is in kb. ▼, site of T-DNA insertion in the *ag-2* mutant. Below the restriction map, horizontal lines indicate the region contained in the plasmid-rescued clone pCIT505, and the cosmid pCIT540. *b*, Structure and restriction map of *AG* gene. The diagram below the restriction map is of introns (lines) and exons (shaded boxes); size in nucleotides is given. One cDNA extended an additional 61 nucleotides (open box). Restriction sites include: A, *Sca*I; B, *Bam*HI; C, *Nco*I; H, *Hin*dIII; N, *Nde*I; S, *Sac*I; R, *Eco*RI; T, *Pst*I; X, *Xba*I. Restriction sites in outline indicate sites in exon sequences. ▼, site of the 35-kb T-DNA insertion in *ag-2* allele. Arrow, direction of transcription. *c*, Nucleotide sequence of *AG* cDNAs and the deduced amino-acid sequence for the open reading frame. Nucleotide 1 corresponds to the first nucleotide in the longest isolated cDNA and nucleotide 1,043 is the site of polyadenylation. The region preceding nucleotide 1 was derived from genomic sequence. ▼, position of intron sequences as deduced from a comparison of cDNA and genomic sequence data. *, position of the translational initiation codon for each of the *B. napus* genes.

METHODS. Cosmids were isolated from genomic libraries of nuclear DNA prepared from the Columbia and Landsberg *erecta* ecotypes of *Arabidopsis*. This DNA was partially digested with *Sau*3A, size fractionated by sucrose gradient centrifugation, ligated to the vector pCIT30 and packaged. Plasmid pCIT30 was derived from the plasmid pMON200 (ref. 26) by replacing the kanamycin resistance gene with a fragment that encodes hygromycin resistance. Plasmid rescue was achieved by digesting DNA from homozygous *ag-2* mutant plants with *Sal*I, followed by ligation and transformation into *Escherichia coli*. The digest with *Sal*I produces a fragment containing an *E. coli* origin of replication and the ampicillin resistance (Amp^r) gene encoded by the T-DNA, as well as plant sequences flanking the insertion[28]. Plasmid pCIT505 was placed on the RFLP map of Chang *et al.*[14] using an *Eco*RI RFLP and found to reside at about the same position as *ag-1*. To map directly pCIT505 relative to *ag-1* recombinants derived from a cross between Landsberg *erecta* (Ler) and Niederzenz (Nd-O) ecotypes that included the *ag-1* mutation were analysed. Plants heterozygous for the *ag-1* mutation (*ag-1* (Ler)/+(Ler)) were crossed to plants that were wild-type for *AG in the* Nd-O background (+(Nd-O)/+(Nd-O)). The resulting F1 progeny that were heterozygous for the *ag-1* mutation (*ag-1* (Ler)/+(Nd-O)) were allowed to self-fertilize and DNA was isolated from the F2 plants that displayed the *ag* phenotype (*ag-1/ag-1*). DNA blots were prepared and hybridized with probes revealing *Eco*RI and *Bgl*II RFLPs. The cDNA library was prepared from 2 μg poly(A)$^+$RNA isolated from immature floral buds[3,20] before stage 12 using the Stratagene kit. After size-selection of the cDNA (>400 bp) *Eco*RI linkers were added before ligating to the vector λgt10 and packaging using Gigapack gold (Stratagene). Many cDNA clones (3×10^6) were obtained, a portion of which were screened with a probe derived from the region flanking the T-DNA insertion site in *ag-2* mutant plants. Twenty seven of the longest independently isolated cDNAs were subcloned and fully or partially sequenced. DNA sequence analysis of one of these clones, pCIT559, demonstrated that it contained the fourth (138 base) intron. All of the other cDNAs seemed to represent the same fully processed RNA. Sequence analysis was performed on double-stranded plasmids using the USB Sequenase Kit. Both strands of cDNAs and the corresponding genomic region were sequenced.

sequences in pCIT505 was determined by restriction fragment length polymorphism (RFLP) analysis[14]. No crossovers were observed between a pCIT505-linked polymorphism (Fig. 2*a*) and the *ag-1* mutation in 118 meiotic products scored, indicating that this cloned segment maps less than two centimorgans from the *AG* gene. Taken together, these results suggest that the T-DNA is inserted in or near the *AG* gene and that we have recovered DNA sequences from this region. The pCIT505 plasmid was then used as a probe to screen a cosmid library of wild-type *Arabidopsis* sequences and several clones were isolated and characterized. One of these cosmids, pCIT540 (Fig. 2*a*), when introduced into the genome of homozygous *ag-2* mutant plants, complemented the mutation and the resulting flowers appeared to be of wild type, with stamens and carpels (Fig. 1*b*). As the T-DNA insertion has been stably maintained through eight generations, and the introduced sequences restore a wild-type flower phenotype to this stable strain, it seems that this region indeed contains the wild-type *AG* gene.

Nucleotide sequence

Restriction fragments spanning the T-DNA insertion site were used to probe a complementary DNA library prepared from RNA isolated from wild-type *Arabidopsis* flowers. Only one class of cDNA clones spanned the T-DNA insertion site. The nucleotide sequence of several overlapping cDNA clones and the corresponding genomic region was determined (Fig. 2*c*). These results indicate that this gene contains at least nine exons ranging in size from 42 to 235 bases, and eight introns ranging in size from 84 to 2,985 bases (Fig. 2*b*). Sequence analysis of the T-DNA-plant junction fragment indicates that the foreign DNA in the *ag-2* allele is inserted in the 2,985-base intron (Fig. 2*b*).

To confirm that these sequences represent the *AG* gene, genomic fragments were isolated and sequenced from the EMS-generated mutant *ag-1*. The *ag-1* mutation was induced in the same genetic background, Landsberg *erecta*, from which we have isolated and characterized the wild-type sequences. This analysis revealed a single nucleotide change that is presumably responsible for the *ag-1* mutant phenotype. This nucleotide substitution changes the AG dinucleotide to AA at the acceptor site of the fourth (138-base) intron shown in Fig. 2*b*.

The deduced protein product contains at least 285 amino-acid residues, although from analysis of the sequences it seems that none of the cDNA clones are of full length. As the deduced open reading frame continues to the 5′ end of the longest cDNA, and a translational initiation codon has not been identified, the full-length AG protein could be slightly longer. But many attempts with primer extension, anchored polymerase chain reaction and the screening of two different cDNA libraries have been unsuccessful at obtaining longer clones. Nearly all of the isolated cDNAs end within a region of a few nucleotides, suggesting that the RNA may have a secondary structure that prevents extension beyond this region. As the longest cDNAs

```
                      10         20         30         40         50
                      |          |          |          |          |
AG          52 RGKIEIKRIENTTNRQVTFCKRRNGLLKKAYELSVLCDAEVALIVFSSRGRLYEYS
DEF A        3 RGKIQIKRIENQTNRQVTYSKRRNGLFKKAHELSVLCDAKVSIIMISSTQKLHEYI
SRF        144 RVKIKMEFIDNKLRRYTTFSKRKTGIMKKAYELSTLTGTQVLLLVASETGHVYTFA
MCM1        18 RRKIEIKFIENKTRRHVTFFKRKHGIMKKAFELSVLTGTQVLLLVVSETGLVYTFS
ARG80       80 RRKQPIRYIENKTRRHVTFSKRRHGIMKKAYELSVLTGANILLLILANSGLVYTFT

AG/DEF A      RGKI IKRIEN TNRQVT  KRRNGL KKA ELSVLCDA V  IV SS   L EY   41/56
AG/MCM1       R KIEIK IEN T R VTF KR  G  KKA ELSVL    V L V S  G  Y  S  32/56
AG/SRF        R KI   I N   R  TF KR  G  KKAYELS L    V L V S  G  Y      25/56
```

FIG. 3 Amino-acid sequence comparison (single-letter code). An alignment is shown for the deduced amino-acid sequences for the gene products from *Arabidopsis thaliana* (AG), *Antirrhinum majus* (DEF A), humans (SRF) and yeast (MCM1, ARG80). Identical residues between AG and each of DEF A, MCM1 and SRF are shown below, as are the fractions of identical amino acids shared by these proteins. Numbers to the left of each sequence indicate the position of the first amino acid shown for each protein.

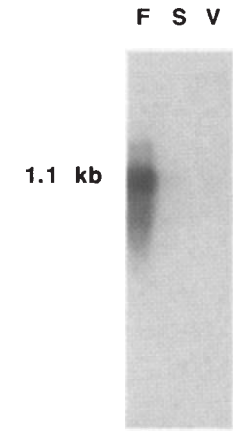


FIG. 4 RNA blot. An autoradiogram is shown for an RNA blot hybridized with ^{32}P-labelled *AG* cDNA. RNA was from immature floral buds (lane F), floral stems (lane S) and from vegetative tissue (lane V).
METHODS. Poly(A)$^+$RNA was isolated from *Arabidopsis thaliana* Landsberg (*er*) as previously described[29] and 2 μg each RNA type was size-fractionated on a 1.5% formaldehyde agarose gel, transferred to Hybond N nylon membrane (Amersham), and hybridized with a radiolabelled gel-isolated DNA fragment from the *AG* cDNA, pCIT565. Hybridization was in 50% formamide, 5 × SSC buffer, 10 × Denhardt's buffer, 100 μg ml^{-1} single-stranded salmon sperm DNA and 10% Dextran sulphate. The filter was washed three times in 2 × SSC, 1% SDS at room temperature and then four times in 0.2 × SSC, 1% SDS at 58 °C. The flower RNA was prepared from a collection of young floral buds before stage 12 (before the flowers open)[3,20]. Vegetative RNA was obtained by growing plants in a liquid culture for ~18 days before collection; the resulting tissue consisted of roots, stems and leaves. Floral stem RNA was prepared from floral stems with cauline leaves. The vegetative and floral stem RNA preparations also contain a small amount of floral RNA as these tissues contain immature stage-3 flowers[20]. To prove that RNA was present in the floral stem and vegetative lanes, this RNA blot was also hybridized with a cDNA specific for the *Arabidopsis* G-protein gene *GPA1* (ref. 30); this clone hybridized more strongly to the stem and vegetative RNA (lanes S and V) than to flower RNA.

are of 1,043 base pairs (bp) and the polyadenylated messenger RNA is close to 1.1 kilobases (kb) in length (see below), these cDNAs are of nearly full length. We have recently isolated the two presumed *AG* homologues from the allotetraploid mustard *Brassica napus* and each of these genes contains a translational initiation codon corresponding to the TCG codon at nucleotides 112–114 of the *AG* cDNA sequence (M.F.Y., H.M. and E.M.M., unpublished results). This indicates that all the evolutionarily conserved sequences of the *Arabidopsis* gene have been identified.

The deduced *AG* amino-acid sequence was used to search the database of protein sequences using the program of Lipman and Pearson[15]. The AG protein shares significant sequence similarity to a class of transcription factors from humans[16] and yeast[17,18], and to the product of a recently isolated *Antirrhinum* homeotic gene[19] (see Fig. 3).

RNA analysis

To study the pattern of *AG* gene expression, poly(A)$^+$ RNA was isolated from several different tissue types and hybridized with an *AG* cDNA. An RNA from flowers (length ~1.1 kb) hybridized with the *AG* cDNA probe (Fig. 4, lane F). After a much longer exposure, a similarly sized transcript was detected in RNA from inflorescence stems and whole plants before the appearance of the primary inflorescence stem (Fig. 4, lanes S and V), although the intensity of this signal was about 1% of that in flowers. The plant tissue used for these RNA isolations did contain a small number of early developing flowers[20], and

it is possible that this low level of signal is the result of RNA in these developing flowers. We estimate the level of *AG* RNA to be $\sim 10^{-4}$ in poly(A)$^+$ RNA of flowering apices, on the basis of both reconstruction experiments and the frequency of cDNA isolation.

To determine if *AG* expression is organ-specific, *in situ* hybridizations were performed. Sections of wild-type *Arabidopsis* flowers were prepared and probed with ^{35}S-labelled RNA synthesized from the *AG* cDNA template (Fig. 5). These results indicate that *AG* is expressed in stamens and carpels and not (or at much lower levels) in sepals or petals.

Discussion

We have isolated and characterized a flower-specific *Arabidopsis* homeotic gene, *AG*. The insertion mutant allele, *ag-2*, and the EMS-induced mutant allele, *ag-1*, both display alterations in the genomic region from which the presumed *AG* RNA is transcribed. The changes in the mutant alleles seem to interfere with normal transcription or splicing.

The deduced sequence of the AG protein shows striking similarity to the sequences of transcription factors from both humans (SRF, ref. 16) and yeast (encoded by *MCM1*, ref. 17, and *ARG80*, ref. 18), and to the deduced amino-acid sequence of a recently isolated homeotic gene (*def A*) from the flowering plant *Antirrhinum majus*[19]. In humans, the serum response factor (SRF) is thought to be required for the serum-inducible transcriptional activation of genes such as c-*fos*, a nuclear proto-oncogene of mammals[16]. In yeast, the product of the *MCM1* gene (GRM/PRTF) is a transcriptional regulator of mating-type-specific genes[21]. The region shared by these proteins spans ~56 amino-acid residues and, allowing for conservative substitutions, *AG* is ~80% similar in this region to all of these proteins (Fig. 3). This same region of SRF encodes the DNA-binding and dimerization domains[16,22]. On the basis of this sequence similarity and the *ag* phenotype we propose that the *AG* protein is a transcription factor involved in regulating genes that determine stamen and carpel development in wild-type flowers.

Like those of *Arabidopsis*, flowers of *Antirrhinum* consist of four concentric whorls of organs (sepals, petals, stamens and carpels). Mutations in the *Antirrhinum def A* homeotic gene cause the conversion of petals into sepals and stamens into carpels. This phenotype does not resemble that of *ag*, but closely resembles the phenotype of the *ap3* and *pi* mutants of *Arabidopsis*[3], suggesting that the *Antirrhinum def A* gene could be the homologue of either the *AP3* or *PI* gene of *Arabidopsis*. The *AG* gene product displays a high degree of similarity to DEF A in the DNA-binding region (Fig. 3), but not elsewhere. As the phenotypes of *Arabidopsis ag* and *Antirrhinum def A* mutants are very different, it seems likely that these genes function in completely different regulatory steps in flower development. The AG and DEF A proteins are more similar to each other than either is to the SRF, MCM1 and ARG80 proteins (Fig. 3).

We have recently isolated a family of other genes from *Arabidopsis*, some of which are preferentially expressed in flowers, that all share extensive sequence similarity to the DNA-binding regions of the SRF–MCM1 class of transcription factors (H.M., M.F.Y. and E.M.M., unpublished results). This indicates that there is a family of related genes that may act to control many of the steps in organ development in *Arabidopsis* flowers. This situation might be analogous to animal development where specific DNA-binding motifs have been adopted for different regulatory circuits. For example, *Drosophila* homeotic selector genes share the homeobox motif, the *Drosophila* gap genes all encode zinc-finger-type proteins, and genes involved in neurogenesis share the helix–loop–helix motif[23,24]. The MCM1[21] and SRF[25] gene products interact with other factors in regulating gene expression, and it is therefore possible that both AG and DEF A also interact with other factors to exert their regulatory control; some of these factors may be encoded by other homeotic genes known to exist in *Arabidopsis* and *Antirrhinum*.

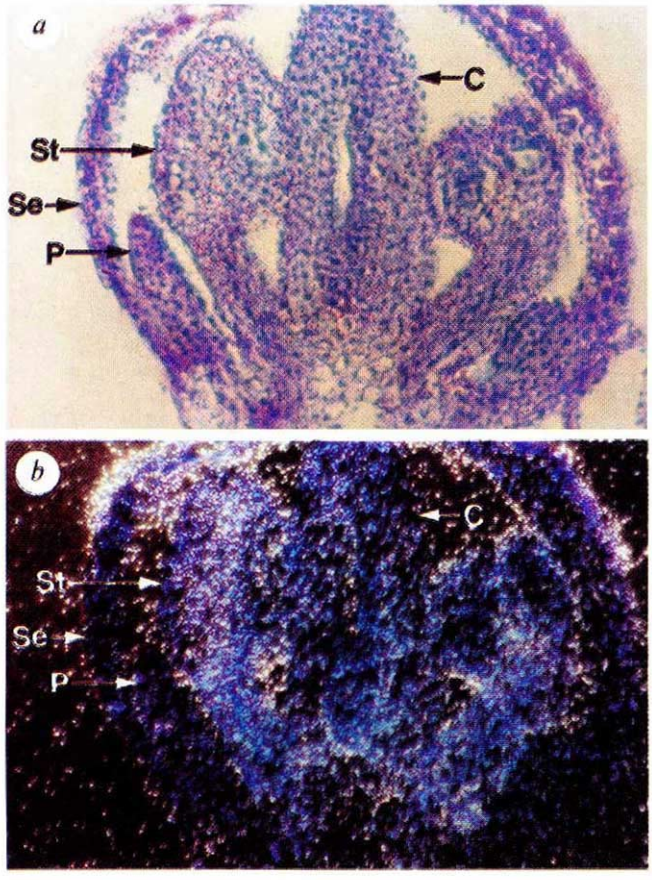


FIG. 5 *In situ* hybridization of *AG*. To determine the pattern of *AG* expression, ^{35}S-labelled antisense RNA was synthesized from an *AG* cDNA template (pCIT565), and hybridized to 8 μm longitudinal sections of stage-9 *Arabidopsis* flowers[3,20] as previously described[31]. *a*, Bright field and *b*, dark field. Arrows indicate the position of floral organs including Se (sepals), P (petals), St (stamens) and C (carpels). Hybridization of antisense RNA to the *AG* mRNA is localized to the positions of stamens and carpels and not to sepals and petals, indicating that it is in these organs that the *AG* gene is preferentially expressed. The signal over the upper portion of the sepals is observed both with sense and antisense RNA probes and therefore does not represent specific hybridization.

In addition, we used RNA blot analysis to identify the 1.1-kb *AG* mRNA and found it to be preferentially expressed in flowers. *In situ* hybridization experiments show that this RNA is localized to the stamens and carpels of wild-type flowers, the same organs that are lacking in *ag* mutants. It will be interesting to determine if the pattern of *AG* gene expression is altered in any of the other *Arabidopsis* homeotic flower mutants. □

Received 24 April; accepted 1 June 1990.

1. Meyerowitz, E. M. *Rev. Genet.* **21**, 93-111 (1987).
2. Meyerowitz, E. M. *Cell* **56**, 263-269 (1989).
3. Bowman, J. L., Smyth, D. R. & Meyerowitz, E. M. *The Plant Cell* **1**, 37-52 (1989).
4. Koornneef, M. in *Genetic Maps 1987* 742-745 (Cold Spring Harbor Laboratory, New York, 1987).
5. Pruitt, R. E., Chang, C., Pang, P. P. & Meyerowitz, E. M. in *45th Symposium of the Society for Developmental Biology* (ed. Loomis, W.) 327-338 (Liss, New York, 1987).
6. Komaki, M. K., Okada, K., Nishino, E. & Shimura, Y. *Development* **104**, 195-203 (1988).
7. Kunst, L., Klenz, J. E., Martinez-Zapater, J. & Haughn, G. W. *The Plant Cell* **1**, 1131-1135 (1989).
8. Meyerowitz, E. M., Smyth, D. R. & Bowman, J. L. *Development* **106**, 209-217 (1989).
9. Braun, A. *Gesell. Naturforsch. Freunde z. Berlin* p. 75 (1873).
10. Conrad, D. *Biol. Zentralbl.* **90**, 137-144 (1971).
11. Koornneef, M., de Bruine, J. H. & Goettsch, P. *Arabidopsis Inf. Serv.* **17**, 11-18 (1980).
12. Feldmann, K. A., Marks, M. D., Christianson, M. L. & Quatrano, R. S. *Science* **243**, 1351-1354 (1989).
13. Feldmann, K. A. & Marks, M. D. *Molec. Gen. Genet.* **208**, 1-9 (1987).
14. Chang, C., Bowman, J. L., DeJohn, A. W., Lander, E. S. & Meyerowitz, E. M. *Proc. natn. Acad. Sci. U.S.A.* **85**, 6856-6860 (1988).
15. Lipman, D. J. & Pearson, W. R. *Science* **227**, 1435-1441 (1985).
16. Norman, C., Runswick, M., Pollock, R. & Triesman, R. *Cell* **55**, 989-1003 (1988).
17. Passmore, S., Maine, G. T., Elble, R., Christ, C. & Tye, B. K. *J. molec. Biol* **204**, 593-606 (1988).
18. Dubois, E., Bercy, J. & Messenguy, F. *Molec. Gen. Genet.* **207**, 142-148 (1987).
19. Sommer, H. *et al. EMBO J.* **9**, 605-613 (1990).
20. Smyth, D. R., Bowman, J. L. & Meyerowitz, E. M. *Plant Cell* **2** (in the press).
21. Herskowitz, I. *Nature* **342**, 749-757 (1990).
22. Hayes, T. E., Kitchen, A. M. & Cochran, B. H. *Proc. natn. Acad. Sci. U.S.A.* **84**, 1272-1276 (1987).
23. Murre, C. *et al.* Cell **58**, 537-544 (1989).
24. Ingham, P. W. *Nature* **335**, 25-34 (1988).
25. Schröter, H., Mueller, C. G. F., Meese, K. & Nordheim, A. *EMBO J.* **9**, 1123-1130 (1990).
26. Fraley, R. T. *et al. Bio/Technology* **3**, 629-635 (1985).
27. Lloyd, A. M. *et al. Science* **234**, 464-466 (1986).
28. Velten, J. & Schell, J. *Nucleic Acids Res.* **13**, 6981-6998 (1985).
29. Crawford, N. M., Campbell, W. H. & Davis, R. W. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8073-8076 (1986).
30. Ma, H., Yanofsky, M. F. & Meyerowitz, E. M. *Proc. natn. Acad. Sci. U.S.A.* **87**, 3821-3825 (1990).
31. Barker, S. J., Harada, J. J. & Goldberg, R. B. *Proc. natn. Acad. Sci. U.S.A.* **85**, 458-462 (1988).

ACKNOWLEDGEMENTS. We thank S. Kempin, L. Medrano, Y. Hu for technical assistance, L. Huang for help in cDNA library construction, and T. Stearns for help with the protein sequence database search. We also thank U. Vijayraghavan for helpful discussions, and A. van der Bliek, C. Chang, S. Kempin and D. Weigel for critical reading of the manuscript. K.A.F. thanks M. L. Christianson for being allowed to experiment on the transformation of intact tissues, and for numerous insightful discussions, while a postdoctoral fellow in Christianson's lab. This work was supported by the NSF and the Lucille P. Markey Charitable Trust (E.M.M.). M.F.Y., H.M. and G.N.D. were supported by the NSF (plant biology), the Helen Hay Whitney Foundation and the NIH, respectively.

Detectability of γ-rays from clumps of dark matter

George Lake

Department of Astronomy, FM-20, University of Washington, Seattle, Washington 98195, USA

IF the dark matter in our Galaxy is made up of weakly interacting massive particles (WIMPs) with masses of the order of several GeV (for example, photinos or Higgsinos), γ-rays produced by their annihilation would in principle be observable[1,2]. But the expected flux[3] from a smoothly distributed dark matter halo[4,5] is much smaller than the observed diffuse background[6], and although narrow lines might be produced, their intensity would be much too low to see with the Gamma Ray Observatory (GRO)[3,7]. A complementary approach is to consider unique spatial signatures. Numerical simulations of galaxy formation[8] show that even in the central bulge of the Galaxy, the mean density of the dark matter could be equal to that of the stars. If this were so, GRO could see the Galactic Centre as a source of annihilating dark matter[1,3]. Other lumps formed as part of the hierarchical formation of the Galaxy could also produce sources that would be recognized by the shape of their continuum spectrum[2,3] and a line feature in sufficiently bright sources[3,7]. Even Geminga[9,10], the second strongest source of γ-rays at energies greater than 50 MeV, could be annihilating dark matter.

Lumps clearly formed during the hierarchical formation of the Galaxy. Clusters of galaxies are large-scale versions of this process, several per cent of their mass is contained in galaxian lumps with masses of $\sim 10^{-4} M_{\rm cluster}$ and density contrasts of $\sim 10^2$. As shown below the density contrast needs to be higher for the lumps to survive in the Galaxy, so I assume that a lower percentage of the mass is in such lumps. These lumps must either be the remnant cores of larger structures or have formed from non-gaussian initial conditions[11,12].

The dark matter in the halo has a local density[4,5] $\rho_{\rm h,0} \approx 10^{-2}\, M_\odot\, {\rm pc}^{-3}$. The distance to the nearest lump is $D_{\rm near} \approx (M_{\rm cl}/\eta\rho_{\rm h,0})^{1/3}$, where η is the fraction of the dark matter in lumps with cluster masses, $M_{\rm cl}$, of $M_6 \times 10^6\, M_\odot$. The flux of γ-rays with energies >50 MeV observed from this lump is

$$F = 5.8\times 10^{-11}\ {\rm cm}^{-2}\ {\rm s}^{-1} (\sigma V)_{26} \left(\frac{X}{7}\right) m_{\lambda,5}^{-2} M_6^{-2/3} \times \left(\frac{\eta\rho_{\rm h,0}}{10^{-2}\, M_\odot\, {\rm pc}^{-3}}\right)^{2/3} \int_0^\infty \rho_{\rm cl}^2 r^2\, {\rm d}r \qquad (1)$$

where $(\sigma V)_{26}$ is the annihilation cross-section in units of $10^{-26}\ {\rm cm}^3\ {\rm s}^{-1}$, $m_{\lambda,5}$ is the WIMP mass in units of 5 GeV and the integral over the density in the lump, $\rho_{\rm cl}$ has units $M_\odot^2\ {\rm pc}^{-3}$. X, the average number of γ-rays with energies >50 MeV per annihilation, is fixed by e^+-e^- collider data[13] and is 7.1 $m_{\lambda,5}^{1/2}$. For the particular case of photinos[7] with masses >5 GeV

$$(\sigma V)_{26} \approx 0.1 (\Omega_\lambda h_{100}^2)^{-1} m_{\lambda,5}^{-2} \qquad (2)$$

where h_{100} is the Hubble constant, H, in units of $100\ {\rm km\ s^{-1}\ Mpc^{-1}}$, and Ω_λ is the fraction of the closure density in WIMPs.

I adopt a density profile appropriate for lumps that have undergone mutual collisions and tidal shocking by the disk of the Galaxy[14]

$$\rho_{\rm cl} = \frac{\rho_0}{(1 + r^2/r_{\rm c}^2)^2}; \qquad \int_0^\infty \rho_{\rm cl}^2 r^2\, {\rm d}r = 2.3 \overline{\rho_{1/2}} M_{\rm tot} \qquad (3)$$

where $M_{\rm tot}$ is the total mass, $\overline{\rho_{1/2}}$ is the mean density inside the half-mass radius, and $r_{\rm c}$ is a core radius.

The density $\overline{\rho_{1/2}}$ constrains the redshift of dissipationless formation for the lumps[15] and determines their ability to survive in the Galaxy. When galaxies form without dissipation, they reach maximum expansion and turn around when their density is a factor of 5.5 greater than the critical density $\rho_{\rm crit} = 3H^2/8\pi G$, where G is the gravitational constant. If pressure is unimportant, the protogalaxy collapses by a factor of two to become bound and virialized. Hence, the relationship between the redshift of turn-around, $z_{\rm turn}$ and $\overline{\rho_{1/2}}$ is

$$(1 + z_{\rm turn}) = 43 h_{100}^{-2/3} \left(\frac{\overline{\rho_{1/2}}}{1 M_\odot\, {\rm pc}^{-3}}\right)^{1/3}$$

The density $\overline{\rho_{1/2}}$ also determines whether the lumps survive mutual collisions and tidal shocking by the galactic disk. Using