Cell, Vol. 89, 737–745, May 30, 1997, Copyright ©1997 by Cell Press

FCA, a Gene Controlling Flowering Time in Arabidopsis, Encodes a Protein Containing RNA-Binding Domains

Richard Macknight,*# Ian Bancroft,*# Tania Page,*‡ Clare Lister,* Renate Schmidt,*§ Karina Love,* Lore Westphal,*‖ George Murphy,* Sarah Sherson,† Christopher Cobbett,† and Caroline Dean*

*Department of Molecular Genetics
John Innes Centre
Norwich Research Park
Colney, Norwich NR4 7UH
United Kingdom
†Department of Genetics
University of Melbourne, Parkville
Melbourne, Victoria 3052
Australia

Summary

A strong promoter of the transition to flowering in Arabidopsis is encoded by *FCA*. *FCA* has been cloned and shown to encode a protein containing two RNA-binding domains and a WW protein interaction domain. This suggests that FCA functions in the posttranscriptional regulation of transcripts involved in the flowering process. The *FCA* transcript is alternatively spliced with only one form encoding the entire FCA protein. Plants carrying the *FCA* gene fused to the strong constitutive 35S promoter flowered earlier, and the ratio and abundance of the different *FCA* transcripts were altered. Thus, *FCA* appears to be a component of a posttranscriptional cascade involved in the control of flowering time.

Introduction

The transition to flowering is a major developmental switch in the life cycle of plants. The apical meristem changes from vegetative to reproductive development to form flowers and seed. In most plant species, this transition is strongly influenced by the environmental cues, photoperiod, and temperature. A long, cold temperature treatment (i.e., a winter season) induces or accelerates flowering. This is known as vernalization, and, for temperate species, it ensures that flowering occurs in spring when conditions are favorable for growth and seed development.

In recent years, there has been a concerted effort to understand the mechanisms controlling flowering time and floral development utilizing a molecular genetic approach in Arabidopsis. Loci influencing flowering time and vernalization requirement have been mapped in natural ecotypes (Burn et al., 1993; Lee et al., 1993; Clarke and Dean, 1994), and mutations that confer either late or early flowering have been identified (Koornneef et al., 1991; Hicks et al., 1996). Late flowering mutants derived from the early flowering parent Landsberg *erecta* (L*er*) have been well characterized (Martinez-Zapater and Somerville, 1990; Koornneef et al., 1991). They affect multiple aspects of both vegetative and reproductive development (Martinez-Zapater et al., 1995) and have been classified into two groups based on their response to day length and vernalization. The mutant phenotype of the first group of late flowering mutants, *fe*, *ft*, *fd*, *fwa*, *co*, and *gi*, is relatively unaffected by both vernalization and short days (the noninductive photoperiod in Arabidopsis), and so their gene products are likely to play a role in the response to environmental cues. *CO* has been cloned and encodes a protein containing two putative zinc finger domains reminiscent of GATA-1 transcription factors (Putterill et al., 1995). The level of *CO* expression is very low throughout the developmental stages analyzed but is slightly higher in long days relative to short days. The flowering time of the second group of late flowering mutants *fca*, *fve*, *ld*, *fy*, and *fpa* is delayed in short days and accelerated to varying degrees by vernalization. *LD* has been cloned and shown to encode a nuclear localized, glutamine-rich protein containing a possible homeodomain (Lee et al., 1994; Aukerman and Amasino, 1996). *LD* expression is throughout the plant and is unaffected by photoperiod. As this second group of mutants are recessive but are able to respond to environmental cues, they are considered to be disrupted in an autonomous promotion pathway (Martinez-Zapater and Somerville, 1990; Amasino, 1996). The acceleration of their flowering time by vernalization is thought to be due to a parallel vernalization-dependent promotion pathway. The balance of the autonomous and vernalization-dependent promotion pathways may explain the different flowering times and vernalization responses of the different Arabidopsis ecotypes. The latter pathway is likely to play a significant role in controlling the flowering time of the majority of Arabidopsis ecotypes that flower late and whose flowering time is significantly accelerated by vernalization. The former may have evolved to bypass the vernalization requirement for early flowering, enabling Arabidopsis ecotypes to adapt to more temperate climes.

We have investigated the role of *FCA* in the control of flowering time. A number of *fca* alleles, induced with either EMS (*fca-1*, *fca-5*, *fca-6*), fast-neutron (*fca-2*), or X-ray mutagenesis (*fca-3*, *fca-4*) are available (Koornneef et al., 1991). *fca-1* is one of the latest flowering of all of the L*er* mutants; however, it is the most responsive to vernalization, with the mutants flowering at the same time as wild-type after saturating vernalization (Koornneef et al., 1991). We report here on the isolation of *FCA*, show that it encodes a protein that is likely to function as a posttranscriptional regulator, and provide evidence that *FCA* itself is posttranscriptionally regulated.

‡Present address: Horticulture Research International, Wellsbourne, Warwick CV35 9EF, United Kingdom.
§Present address: Max-Delbrück Laboratory, Carl-von-Linne Weg 10, D50829, Köln, Germany.
‖Present address: Institut für Angewandte Genetik, Universität Hannover, Herrenhäuser Str 2, 3000 Hannover 21, Germany.
#R. M. and I. B. are joint first authors.

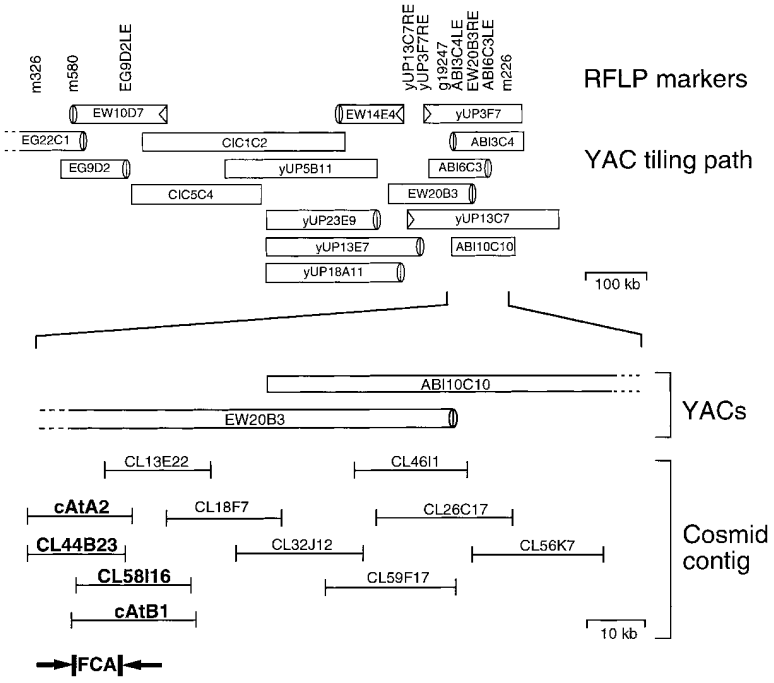

Figure 1. YAC and Cosmid Contig Covering *FCA*

The upper part of the figure shows the RFLP markers used to position the *FCA* locus within a YAC contig generated from Schmidt et al. (1996) with relevant ABI YAC clones integrated. The length of the rectangles represent the insert size of the YAC clones. Circles at one end of the YAC clones represent left end-fragments, and triangles represent right-end fragments. Not all of EG22C1 is shown. The lower part of the figure shows the enlargement of part of the YAC clones EW20B3 and ABI 10C10 and the cosmid contig. Those cosmids complementing the *fca* mutant phenotype are shown in bold. The overlapping DNA within the cosmids that contains *FCA* is ~10 kb.

Results

Positional Cloning of *FCA*

The *fca* mutation was mapped to the interval on chromosome 4 between g19247 and m226, covered by two overlapping YAC clones EW20B3 and ABI10C10 (Figure 1). These YAC clones were hybridized to a cosmid library of L*er* genomic DNA in an Agrobacterium binary vector. Eleven cosmid clones were introduced into the Arabidopsis *fca-1* mutant, and four cosmids, cAtA2, CL44B23, CL58I16, and cAtB1, complemented the *fca* mutation.

FCA Encodes a Protein Containing Two RNA-Recognition Motifs and a WW Protein Interaction Domain

The overlapping region of the complementing cosmids was 10 kb. The complete genomic sequence of this genomic region was determined. It was also hybridized to 1 × 10^6 phage clones of the PRL-2 cDNA library (Newman et al., 1994). This identified two cDNA clones with insert sizes ~1650 bp and 1300 bp. The first of these (cDNAI) hybridized to a ~3 kb transcript in wild-type and *fca-1*, *fca-2*, *fca-3*, *fca-5*, and *fca-6* and a ~2 kb transcript in *fca-4*. The other clone (cDNAII) hybridized to the same sized transcript (1.3 kb) in all alleles examined. The smaller transcript in *fca-4* suggested that the 3 kb mRNA was a transcript from *FCA*. RT–PCR experiments were performed using primers designed from within cDNAI and cDNAII and the genomic interval between them. Alternatively spliced forms were detected (PCRi and PCRii, Figure 2), and it became clear that the two cDNA clones represented alternatively spliced products, both from *FCA*.

The genomic sequence of *FCA* (Columbia allele 1–9763, including all of the promoter region contained in the complementing cosmid CL44B23) is available through the EMBL database (accession no. Z82992). The transcript start site of *FCA* was determined using RNAse protection experiments. The major start site was at nucleotide 1119. The first methionine residue in-frame with the rest of the protein started at nucleotide 1532, giving a predicted 5′ untranslated leader of 414 bp. The structure of *FCA* is shown in Figure 2. The gene contains 21 exons and covers 8.1 kb, and it encodes a predicted

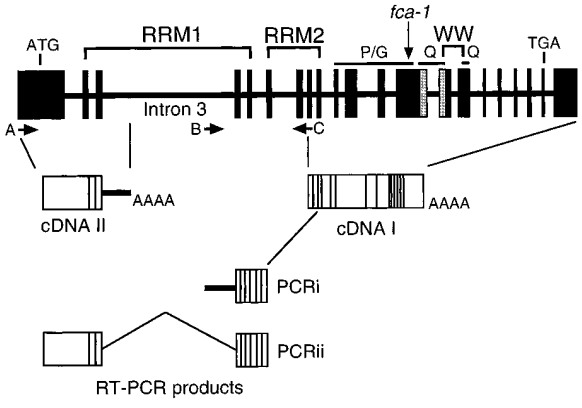

Figure 2. Schematic Representation of the *FCA* Gene and Location of RRM and WW Domains

The differently sized exons are illustrated as differently sized closed boxes, with the intron sequences shown as lines. The exons encoding the two RRMs, the proline/glycine-rich regions (P/G), glutamine-rich regions (Q), and the WW domain are indicated. Also shown are the extent and positions of cDNAI, cDNAII, and the two RT–PCR products isolated during the characterization of the transcripts: PCRi (containing intron 3 sequences), isolated using primers B and C, and PCRii (lacking intron 3 sequences), isolated using primers A and C. The position of the mutation in exon 13 in *fca-1*, which converts a glutamine codon (CAA) into a stop codon (TAA), is indicated. The hatched regions of exons 13 and 14 are removed in transcript δ.

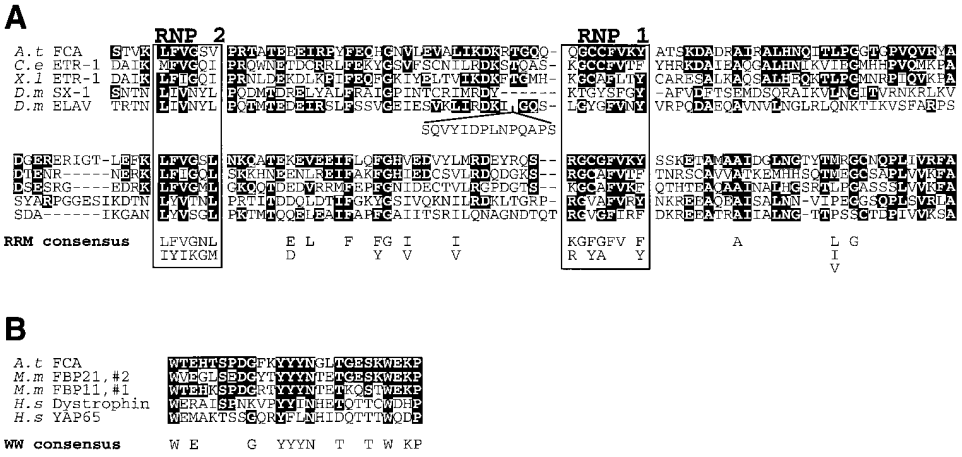

Figure 3. Alignment of RRM and WW Domains

(A) The amino acid sequence encompassing the two RRMs in the FCA protein aligned to RRMs from C. elegans ETR-1, X. laevis ETR-1, and D. melanogaster SX-l and ELAV. The general consensus RRM sequence is shown underneath (Burd and Dreyfuss, 1994). Amino acid identity between the RRMs of the selected proteins is indicated by closed boxes. These RRMs show much greater homology to each other than to RRMs in general. The spacing between the RRMs is also conserved. ELAV contains 14 extra amino acids within the first RRM 5′ to RNP1. This is shown as an insert on the line below to conserve the maximum alignment.

(B) The 26 amino acid region identified as the WW domain in FCA aligned to WW domains from mouse FBP21 (the second of two in this protein) and FBP11 (the first of two in this protein), human dystrophin, and human YAP65 (Yes-associated protein). The WW consensus is shown underneath (Chan et al., 1996).

protein of 747 amino acids. Comparison of FCA with sequences in the databases revealed significant homology to a class of proteins bearing RNA recognition motifs (RRMs). RRMs comprise a loosely conserved region of approximately 80 amino acids, within which lie two highly conserved sequences called RNP1 and RNP2. The FCA protein contains two RRMs (between amino acid residues 118–199 and 209–289) that show most similarity to RRMs present in a subfamily of RNA-binding proteins encoded by *ELAV*-like genes (Figure 3A) (Good, 1995). *ELAV* is a Drosophila gene required for the proper development and maintenance of the nervous system (Robinow et al., 1988), and related neural-specific genes have been found in humans and other vertebrates (Good, 1995). The closest similarity over the RRMs is to ETR-1 (ELAV-type ribonucleoprotein-1) from C. elegans (P. Good, unpublished data; accession number U53931) and X. laevis (Richter et al., 1988; Knecht et al., 1995). Most of the *ELAV*-like genes encode proteins consisting of two consecutive RRMs within the amino-terminal half of the protein, connected by a tether region to a third RRM at the carboxyl terminus. FCA, like the Drosophila SEX-LETHAL protein (Bell et al., 1988), only shares homology with the first two RRMs of the ELAV family.

The FCA protein also contains a recently identified domain called WW or WWP (because of two conserved tryptophan [W] and proline [P] residues; Figure 3B) (Bork and Sudol, 1994; Chen and Sudol, 1995) between amino acids 597 and 622. The WW domain has been found in a diverse group of proteins involved in cell signaling or regulation, and it is considered to play a role in mediating protein:protein interactions. The WW domain in FCA shows greatest similarity to WW domains present in a group of proteins that were recently identified due to their ability to bind in vitro to a proline-rich region of the mouse formin protein (formin binding proteins, or FBPs [Chan et al., 1996]). Formin, which is involved in limb and kidney development, also interacts with proteins containing an SH3 domain via this proline-rich region. The strong homology in the WW domains between FBPs and FCA suggests that FCA will bind to a proline-rich sequence similar to that found in formin.

RNA-binding proteins often have auxiliary domains rich in particular amino acids such as proline, glycine, glutamine, or arginine/serine dipeptide (Burd and Dreyfuss, 1994). The FCA protein contains a region (from amino acid residue 290 to 490) that has a relatively high proline (19%) and glycine (11%) content. On either side of the WW domain, the protein is glutamine rich (25% of the amino acids residues between 491 and 596 are glutamine, and 49% between 623 and 672 are glutamine).

RNA-Binding Ability of FCA

Since the cellular target of FCA is not known, we used a semispecific assay to determine whether FCA is a functional RNA-binding protein. The assay, which involves binding to different homopolymeric RNAs in vitro, has been extensively used to study the RNA-binding properties of hnRNP and other proteins (Swanson and Dreyfuss, 1988; Dreyfuss et al., 1993). FCA bound only to poly(G) and poly(U) sequences and not to poly(A), poly(C), ssDNA, or dsDNA at moderate salt concentrations (150 mM KCl, Figure 4). At higher salt concentrations (1.2 M KCl), binding to both poly(G) and poly(U) ribohomopolymers was still maintained. Thus, we conclude that FCA binds RNA in vitro with a preference for U- and G-rich sequences. We are now using SELEX experiments (Tuerk and Gold, 1990) to analyze the sequence requirements for FCA binding.

The *FCA* Transcript Is Alternatively Spliced

Alternative splicing of the *FCA* transcript had been detected during the analysis of cDNA clones and identification of the *FCA* ORF. To fully characterize the different

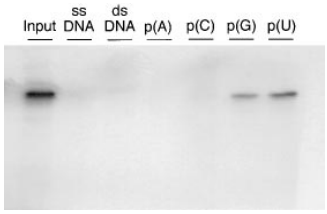

Figure 4. Binding of FCA to Ribohomopolymers

In vitro translated FCA protein was tested for its ability to bind various nucleic acids, as indicated above the lanes. An amount corresponding to 10% of the input of each assay was loaded as a size standard.

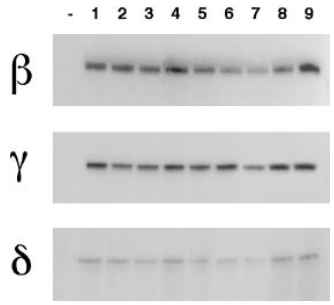

Figure 6. RNAse Protection Assays to Analyze the Relative Abundance of Transcripts β, γ, and δ in Different Tissues, at Different Developmental Stages, and after Growth in Different Conditions

RNAse protection assays, using probes specific for either transcript β, γ, or δ performed on 50 μg RNA isolated from: lane (−), negative control (yeast tRNA); lane 1, plants with 4 leaves grown under a short day photoperiod; lanes 2–5, plants grown under a long day photoperiod with 2, 4, 6, and 8 leaves, respectively (flower buds were visible on plants with 8 leaves); lane 6, roots; lane 7, top 2 cm of inflorescences; lane 8, fully expanded leaves from long day grown plants; and lane 9, plants vernalized for 4 weeks then grown in long day photoperiod until 4 leaves had emerged.

transcripts formed, a combination of RT–PCR and RNAse protection experiments were undertaken. Schematic representations of the transcripts found are shown in Figure 5. The major transcript, termed transcript β, accounts for ~55% of the *FCA* message in RNA isolated from young seedlings and was represented by the initial cDNA clone, cDNAII. It is ~1.3 kb and is formed from cleavage and polyadenylation within intron 3, introns 1 and 2 having been excised. A protein produced from transcript β would consist of the N terminus of the protein plus 45 amino acids of the 82 amino acid long first RRM domain, terminating just at the start of the RNP1 submotif. Transcript γ, the second most abundant transcript, is 3 kb. It is the only transcript that encodes the putative full-length FCA protein. It contains the 21 exons (as shown in Figure 2) and represents ~35% of *FCA* transcripts. A third *FCA* transcript, termed transcript δ, has all introns excised but is alternatively spliced at intron 13. The 5′ and the 3′ splice junctions are shifted either 75 or 81 bp 5′ and 76 or 82 bp 3′. The absolute junctions could not be established because of a 6 bp repeated sequence, but neither incorporates canonical splice sites. The 5′ and 3′ dinucleotides of the larger intron 13 are either /CA, AG/ or /CT, AC/. The extension of the intron 5′ and 3′ results in a deletion of part of exon 13 and incorporation of an in-frame TAA codon immediately at the start of exon 14. Thus, a protein produced from transcript δ would contain both RRMs but would lack the WW domain and the 63 amino acids just upstream of it. Approximately 10% of *FCA* transcripts are estimated to be of the δ form. A fourth transcript is 5 kb, and 20 of the 21 introns have been excised but intron 3 is retained. It is termed transcript α and is present at low levels (<1% of *FCA* transcripts).

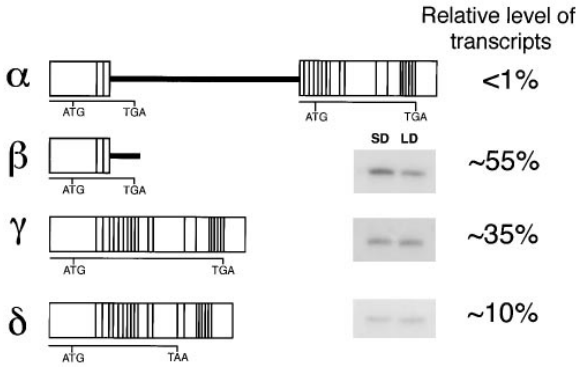

Figure 5. Schematic Illustration and Abundance of *FCA* Transcripts

The exon/intron composition of transcripts α, β, γ, and δ is shown. Where intron sequences are included, these are shown as thick horizontal lines. The thin lines beneath each transcript represent possible ORFs. RNAse protection assays, using probes specific for either transcripts β or γ, performed on RNA isolated from aerial parts of plants grown under short day (SD) and long day (LD) photoperiods, are shown. The relative abundance of the transcripts, as assayed by RNAse protection assays and RT–PCR, is shown at the right-hand side of the figure.

The Mutation in *fca-1* Would Yield a Protein Similar to That Encoded by Transcript δ

The mutation in the strong EMS-induced allele, *fca-1*, was determined. It was found to be a point mutation, at position 6861 within exon 13, that converted a C residue to a T and introduced an in-frame stop codon, CAA to TAA. An FCA polypeptide in *fca-1* would be 468 amino acids long and would contain both RRMs but would lack the glutamine-rich region at the C-terminal end of the protein and the WW domain (Figure 2). This protein would be 65 amino acids shorter than that encoded by transcript δ. That *fca-1* is one of the strongest alleles phenotypically suggests that this partial protein is nonfunctional and that the *fca-1* allele is a null; however, this remains to be confirmed.

Relative Abundance of the Major *FCA* Transcripts in Different Tissues, at Different Developmental Stages, and after Growth in Different Conditions

The abundance of *FCA* transcripts was found to be low, and so RNAse protection assays were used to quantitate expression. The levels of transcripts β, γ, and δ and the ratios relative to each other were constant in all of the

Table 1. Flowering Time of Wild-Type L*er*, *fca-4* and *35S–FCA* Transgenic Lines

Line	LD	ESD	SD
35S–FCA-15 (in *fca-4*)	5.8 ± 0.2	5.5 ± 0.3	16.1 ± 0.3
35S–FCA-7 (in L*er*)	6.0 ± 0.2	4.4 ± 0.2	14.8 ± 0.2
L*er*	6.7 ± 0.3	6.1 ± 0.3	17.2 ± 0.2
fca-4	17.0 ± 1.0	16.7 ± 0.5	ND

Rosette leaf number at flowering time for wild-type L*er*, *fca-4* and two independent, homozygous, single copy, transgenic lines containing the *35S–FCA* fusion. The plants were grown in three different environmental conditions LD (16 hr light), extended SD (10 hr high light + 6 hr low irradiance light) and SD (10 hr light). Numbers are means ± S.E. for 40 plants for ESD and SD conditions and 10 plants for LD conditions. (ND), not determined.

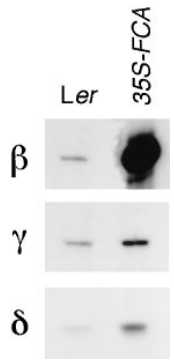

Figure 7. RNAse Protection Assay Showing the Levels of *FCA* Transcripts in Wild-Type and *35S–FCA* Plants

RNAse protection assays performed on 50 μg of total RNA isolated from young seedlings of wild-type L*er* and a transgenic line *35S–FCA-15* (carrying the *35S–FCA* fusion) Three different probes specific for *FCA* transcripts β, γ, and δ were used.

developmental stages analyzed (Figure 6). They were also present at the same levels in RNA isolated from leaves, roots, inflorescences, and from plants grown in SD or subjected to a vernalization treatment. Transcript α was barely detectable (<1% of *FCA* transcript) in any of the samples analyzed. RNA was also isolated from old leaf tissue, and transcript γ and β levels were assayed. Both were reduced (per μg RNA) as compared to RNA isolated from the relatively young material.

Introduction of a *35S–FCA* Gene Fusion Results in Accumulation of Transcript β

To investigate the effects of increasing transcription of *FCA*, plants containing the entire *FCA* gene under the control of the strong constitutive 35S promoter (from the cauliflower mosaic virus) were generated. Two homozygous independent, single insert, transgenic lines, one in an L*er* background (*35S–FCA-7*) and one in *fca-4* (*35S–FCA-15*), were generated. They were grown alongside a L*er* control, and rosette leaf number at flowering was counted. In three different environmental conditions, LD (16 hr light), extended SD (10 high light + 6 hr low irradiance light), and SD (10 hr light), both transformants flowered slightly but reproducibly earlier than the nontransgenic control (Table 1). The presence of the *35S–FCA* transgene not only complemented the *fca-4* mutant phenotype but accelerated flowering in both lines compared to wild-type plants. The levels of the different *FCA* transcripts were assayed in the plants carrying the *35S–FCA* transgene. Transcript β accumulated to very high levels in both lines, 150 times higher than wild-type (Figure 7 shows *35S–FCA-15* data). The levels of transcript γ and δ also increased but to a much lower extent, 5- and 10-fold higher than wild-type in *35S–FCA-15* and *35S–FCA-7*, respectively. This experiment showed that the ratio of abundance of the different *FCA* transcripts can change significantly, and the large accumulation of transcript β in the *35S–FCA* lines indicates that the processing of intron 3 is regulated. It also shows that an increase in the level of *FCA* transcripts is correlated with an acceleration of flowering time.

Overexpression of Transcript β Does Not Restore Early Flowering

From our previous analyses, we had concluded that transcript γ was the only transcript that could encode a functional FCA protein. The small acceleration in flowering time of the *35S–FCA* transgenic lines was therefore likely to have resulted from the increased levels of transcript γ. The higher levels of γ in *35S–FCA-7*, the earlier flowering of the two lines, supports this idea. To test whether the very significant elevation of transcript β levels affected flowering time at all, plants were generated that expressed only transcript β from the 35S promoter. The flowering time of ten transformants in an *fca-1* background was analyzed in greenhouse conditions. All of the plants from each transgenic line flowered at the same time as the *fca-1* parent, showing that high levels of the protein produced by transcript β did not even partially complement the late flowering phenotype of the *fca-1* mutant.

Discussion

The identification of the *FCA* gene product as an RNA-binding protein and the alternative processing of the *FCA* transcript has demonstrated that posttranscriptional regulation is an important mechanism in the control of flowering time in Arabidopsis.

RNA-binding proteins function in a variety of processes including splicing, nuclear export, polyadenylation, RNA stability, and translation. The two RRMs in the FCA protein share most homology with the first two RRMs present within the C. elegans ETR-1 protein (41% amino acid identity). The homology and spacing between RRM1 and RRM2 is also highly conserved between FCA and the two Drosophila proteins ELAV and SX-1, which play critical roles in establishing developmental pathways. ELAV is important in neuron differentiation, and ectopic expression of ELAV in imaginal discs is sufficient to mediate neuron-specific alternative splicing of *NEUROGLIAN* transcripts (Koushika et al., 1996). SX-1 protein regulates the somatic sex determination pathway and acts to cause alternative splicing of *SX-1* pre-mRNA and the pre-mRNA of a downstream gene *TRA*, yielding active SX-1 and TRA protein only in female flies (Bell et al., 1988). Given the homology between the RRMs of ELAV, SX-1, and FCA, it is possible that FCA

also functions in alternative splicing of transcripts. However, other ELAV-related RRM-containing proteins do not appear to act through regulation of splicing, e.g., HEL-N1 binds the 3′ UTR of growth regulator mRNAs, targeting them to polysomes (Gao and Keene, 1996). We are determining the cellular location of the FCA protein and which transcripts it binds to establish its role in the posttranscriptional regulation of transcripts involved in the flowering process.

The presence of the WW protein–protein interaction domain suggests that FCA also interacts with at least one other protein, probably with a proline-rich ligand. That the *fca-1* mutation causes premature termination of the protein just upstream of this domain suggests that it is essential for *FCA* function. WW-containing proteins are widely distributed in nature (Sudol et al., 1995). They are thought to function like SH2, SH3, and PH domains, bringing proteins together intracellularly in regulatory pathways (Andre and Springael, 1994; Bork and Sudol, 1994). They have been found in a diverse set of proteins including formin binding proteins YAP65, a protein that associates in vitro with the SH3 domain of the Yes protooncogene product; dystrophin, a cytoskeletal protein from the Duchenne muscular dystrophy locus; and a human putative GTPase-activating protein. Identification of proteins that bind to the FCA WW domain will help to determine the function of this domain in the FCA protein.

Our current data suggests that transcript γ is the only transcript that encodes a functional FCA protein. The other *FCA* transcripts, however, could produce different protein isoforms that may have distinct functions and be expressed preferentially at different stages or in different places in the plant, as has been found for the Drosophila TRA-2 protein (Mattox et al., 1996). Transcript β would produce a protein carrying the N terminus of the protein and the first part of the first RRM. Overexpression of this form did not change the flowering time of the *fca-1* mutant, but these plants were not tested for any other parameter. Transcript δ could produce a protein containing the RRMs but lacking the WW domain. We obtained no evidence for differential accumulation of these different transcripts based on using RNAse protection experiments. However, the RNA used in these experiments was isolated from whole seedlings or organs, and any subtle differences in particular cells would not have been detected using this technique. We are investigating this further through analysis of lines carrying *FCA* fused at different positions to a β-glucuronidase marker.

The *FCA* transcript is alternatively spliced, and four different mRNAs have been detected. This suggested that *FCA* was a component of a posttranscriptional regulatory cascade with the alternative processing limiting FCA production. Increasing the transcription of the *FCA* gene by fusing it to the strong constitutive viral 35S promoter resulted in a large increase in transcript β levels, a smaller increase in transcript γ and δ levels, and an acceleration in flowering time. These data support a model where the splicing of *FCA* transcripts is regulated, with a major control point being the processing of intron 3. That transcript β is the most abundant transcript in wild-type cells and that it accumulates to such high levels in the *35S–FCA* lines suggests that the correct splicing of intron 3 requires a specific factor(s) that is limiting. This factor(s) may be present at a low level in all cells or it may be present in specific cells or at a specific developmental stage resulting in limited production of transcript γ. The default pathway would then be polyadenylation and the production of transcript β. This would account for the large increase in transcript β in plants ectopically expressing *FCA* from the 35S promoter. An alternative model that would account for high transcript β levels would be one where FCA negatively regulated the production of transcript γ. The result of increasing FCA levels would be a feedback regulation preventing more transcript γ formation and increased levels of transcript β.

Of the differential RNA processing events so far described in plants, alternative splicing/polyadenylation is relatively frequent (Simpson and Filipowicz, 1996). In the absence of a 5′ splice site, polyadenylation within an AU-rich intron is the default pathway (Niwa et al., 1992; Luehrsen and Walbot, 1994). However, the 5′ splice site at intron 3 (AAG:GUAUGU) conforms to the Arabidopsis 5′ splice consensus sequence (Brown, 1996) at all but one position. This may suggest that the 5′ splice site of intron 3 is masked until a certain factor is present. Regulation at intron 3 may not be the only point of posttranscriptional regulation of *FCA*. The alternative splicing at intron 13 occurs at a relatively high frequency. Not only are both 5′ and 3′ splice sites changed, but noncanonical sites are used. This may indicate a novel splicing pathway utilized to add a second level of control in the formation of transcript γ, and the significance of this is under study.

The flowering time of the *35S–FCA* lines was slightly but reproducibly earlier than wild-type. This is likely to have been caused by the increased levels of transcript γ. Whether increased levels of transcript γ in the absence of the large increase in transcript β would have caused even earlier flowering is under investigation. We have shown that transcript β overexpression does not restore early flowering to *fca-1*, but we do not yet know whether it would delay flowering in wild-type plants. Transgenic lines (in both wild-type and *fca-1* backgrounds) carrying either an intronless *FCA* gene or a 35S–γ fusion will further establish the significance of the alternative processing of *FCA* transcripts in the control of flowering time.

The *fca* mutation affects multiple phases of development, suggesting that the *FCA* gene product is required throughout development rather than just at the point of the vegetative to reproductive transition. The similar level of transcript γ expression in different organs and at different developmental stages is consistent with this. Furner et al. (1996) have addressed whether *FCA* function is required throughout the plant using X-irradiation to generate null *fca* sectors. Their data suggested that *FCA* was not required in the inner two (L2 and L3) meristematic layers to produce a phenotypically normal plant. They concluded that *FCA* or downstream gene products involve diffusible signals. It will be interesting to generate sectors expressing *FCA* in an otherwise *fca* background to further investigate the non–cell autonomy and when *FCA* expression is required.

Localization of the FCA protein, identification of the RNA and protein ligands, and *FCA* expression analysis in different mutant backgrounds should help elucidate the mechanism of FCA action and the interactions of the gene products in the floral promotion pathways. We can also begin to investigate whether *FCA* is involved in the production of promotive flowering signals or in affecting competence of the meristem to respond to these signals, or both.

Experimental Procedures

Genetic Mapping of the *fca* Mutation

To map the *FCA* locus relative to molecular markers, 160 late flowering F2 individuals (homozygous recessive class) from a cross between the late flowering mutant *fca-1* (in a L*er* background, NASC stock number NW52) and the polymorphic early flowering ecotype Columbia (Col NASC stock number N933) were analyzed with RFLP markers m210, m326, m580, m226, g10086, g4546, g4108 (Chang et al., 1988; Nam et al., 1989), pCITd23 (from E. Meyerowitz, Caltech, Pasadena, CA), and g19247 (from B. Hauge and H. Goodman, MGH, Boston, MA). Plant genomic DNA preparation and Southern blot analysis was performed as described by Dean et al. (1992). Preliminary mapping positioned the *FCA* locus between the markers m326 and m226. Other RFLP markers used were: the left end-fragment (LE) of EG9D2; right end-fragment (RE) of yUP13C7; RE of yUP3F7; RE of EW20B3; LE of ABI3C4; LE of ABI6C3; and two cosmids g19247 and cAtA2.

Additional plants carrying crossovers closely linked to *FCA* were selected from the F2 generation of a cross between *fca-1* and *ara1*, a semidominant mutation conferring arabinose sensitivity (in Col) that maps ~0.2 cM proximal to the *FCA* locus (Dolezal and Cobbett, 1991). Individuals that were arabinose resistant and flowered early were selected, and their progeny were checked to confirm that they were homozygous for the wild-type *ARA1* allele and heterozygous for the *fca* mutation. Analysis of these plants mapped *FCA* distal to the polymorphism identified by g19247. The interval between g19247 and m226 was covered by two overlapping YAC clones EW20B3 and ABI10C10 (Figure 1).

Analysis of YAC Clones

The four Columbia YAC libraries used and the generation of YAC end-probes are described in Schmidt et al. (1996). The ABI library was obtained from E. Grill and C. Somerville (Michigan State University). Protocols for YAC screening, yeast DNA preparations, and Southern blot analysis of YAC clones are described in Schmidt and Dean (1995).

Screening the Cosmid Library

The YAC clones EW20B3 and ABI10C10 were gel purified and hybridized to filters carrying 25500 cosmid clones each carrying inserts of 15–20 kb of L*er* genomic DNA in a cosmid vector pCLD 04541 carrying Agrobacterium LB, RB sequences, and a 35S–NPTII fusion (C. L. and C. D., unpublished data). The library was screened by gridding offsets from 16 microtitre plates onto LB-tet (10 μg/ml) plates and then taking colony lifts onto Hybond N filters. Positively hybridizing colonies were analyzed by hybridizing each clone to Southern blots carrying all the cosmid clones digested with HindIII, EcoRI, and BamHI. This generated a restriction map for the insert of each cosmid and indicated which clones carried overlapping inserts. The two cosmid clones, cAtA2 and cAtB1, were isolated from a cosmid library carrying Columbia genomic DNA (Olszewski et al., 1988). The result of this analysis was a cosmid contig covering the 100 kb interval containing the *FCA* locus, shown in Figure 1.

Transformation of Arabidopsis

For the complementation experiments, eleven cosmid clones were mobilized into Agrobacterium tumefaciens C58C1, and the T-DNA introduced into Arabidopsis *fca-1* mutant plants were marked with a *cer4* mutation (conferring a waxless phenotype) using the root explant transformation protocol (Valvekens et al., 1988). Multiple independent transformants carrying each cosmid clone were generated. Seeds were collected from self-fertilized kanamycin-resistant individuals (all of which were confirmed to carry the *cer4* marker) and analyzed with respect to their kanamycin segregation and flowering time. Multiple transformants (3 out of 4, 7 out of 8, 6 out of 11, and 4 out of 4, respectively) carrying four cosmids, cAtA2, CL44B23, CL58I16, and cAtB1, complemented the *fca* mutation. Three-quarters of the progeny flowered early with a similar flowering time to kanamycin-resistant L*er* transformants carrying the binary vector alone, planted alongside. One-quarter flowered late with more than 25 leaves (similar to the *fca-1* plants grown alongside). Cosegregation of kanamycin resistance and early flowering was confirmed in two transformants for each of the four complementing cosmid clones. Multiple transformants carrying the other seven cosmids did not yield any plants segregating for early flowering.

Transgenic lines, in either L*er* or *fca-1* backgrounds, carrying the *35S–FCA* and *35S*–β fusions were generated using the vacuum infiltration transformation procedure (Bechtold et al., 1993).

Isolation of cDNA by RT–PCR

Total RNA was isolated from whole seedlings at the 2–3 leaf stage grown under a 16 hr photoperiod according to Dean et al. (1985). First-strand cDNA synthesis and PCR were performed according to standard procedures. The following pairs of primers were used to isolate cDNA extending 5′ to the cDNAI clone; primers B (5′-ATTGAG ATTCTTACATACTG-3′; positioned 1669 bp into intron 3) and C (5′-TCTTTGGCTCAGCAAACCG-3′; 43 bp into exon 9) were used to isolate a 777 bp cDNA fragment extending to within intron 3; and primers A (5′-CGGATCCTTCATCATCTTCGATACTCG-3′; 25 bp into exon 1) and C were used to isolate a 1240 bp product extending to within exon 1. For identification of the mutation in *fca-1*, *FCA* cDNA was isolated in two overlapping fragments from *fca-1* and wild-type Landsberg *erecta* using primer pairs A and C, and D (5′-AAACAAGC AAGCCACTG-3′; 34 bp into exon 6) and E (5′-AGGCCATTGTTTGGC AGCTC-3′; 225 bp into exon 21).

RNAse Protection Assays

The following fragments were cloned into Bluescript KSII (Stratagene) to generate antisense RNA transcripts. Transcript β was analyzed using a 345 bp DraI fragment (bp 49 to bp 394 within intron 3); transcript γ was analyzed using a 166 bp TaqI fragment (from 43 bp into exon 14 to 32 bp into exon 15); and transcript δ was analyzed using a 212 bp NheI/TaqI fragment (from 310 bp into exon 13 to 32 bp into exon 15). The transcription start site was determined using a 284 bp AluI fragment (143 bp before exon 1 to 141 bp into exon 1). The plasmids were linearized and in vitro transcribed with either T3 or T7 RNA polymerase (Promega). RNA transcripts, labeled with [α-^{32}P]UTP, were purified on 6% polyacrylamide–8 M urea gels. RNAse protection assays were performed using Hyspeed RPA kit (Ambion Inc.) according to the manufacturer's instructions.

Ribohomopolymer Binding Assay

The FCA protein was generated by in vitro transcription/translation using the TNT T7 coupled rabbit reticulocyle lysate system (Promega). RNA binding was performed using 2 μl of the ^{35}S-labeled translation product diluted in 500 μl of KHN buffer (150 mM KCl; 20 mM HEPES [pH 7.9], and 0.01% NP–40). Twenty microliters of poly(rA), poly(rC), poly(rG), and poly(rU) Sepharose beads, calf thymus ssDNA, and dsDNA cellulose beads (Sigma) was added and rotated for 10 min at room temperature. The beads were centrifuged and washed three times with 1 ml of KHN buffer. Proteins were eluted in SDS loading buffer and analyzed by SDS-PAGE.

Construction of the *35S–FCA* and 35S–β Fusions

The *35S–FCA* fusion was produced by subcloning the cauliflower mosaic virus 35S promoter from the vector pJJ3411 (Jones et al., 1992) as an EcoRI/XhoI fragment into the EcoRI (within cloning cassette)/SalI (bp 1470) sites of a pBluescript KSII vector containing the 9763 bp *FCA* gene. The 35S–β construct, in which only transcript β is expressed from the 35S promoter, was produced by digesting the plasmid containing the *35S–FCA* fusion with SpeI [which cuts 72 bp before the start of exon 4 and 87 bp before poly(A) site within

exon 21] and religating the vector. Both constructs were subcloned as EcoRI/XhoI fragments into the binary vector pSLJ1714 (Jones et al., 1992).

Acknowledgments

Correspondence regarding this paper should be addressed to C. D. (e-mail: caroline.dean@bbsrc.ac.uk). The authors thank Prof. Maarten Koornneef for the *fca* alleles and encouragement at the start of this project. They also acknowledge the technical help of Rob Ewing during the early stages of the project and Drs. Gordon Simpson, Paul Dijkwel, Jonathan Jones, George Coupland, and Nick Harberd for comments on the manuscript. This work was funded through the BBSRC CSG grant to the JIC, BBSRC Plant Molecular Biology programmes I and II, and EC Bridge grant (BIOT CT 90-0207).

Received November 21, 1996; revised April 1, 1997.

References

Amasino, R.M. (1996). Control of flowering time in plants. Curr. Opin. Genet. Dev. *6*, 480–487.

Andre, B., and Springael, J.-Y. (1994). WWP, a new amino acid motif present in single or multiple copies in various proteins including dystrphin and the SH3-binding Yes-associated protein YAP65. Biochem. Biophys. Res. Commun. *205*, 1201–1205.

Aukerman, M.J., and Amasino, R.M. (1996). Molecular genetic analysis of flowering time in *Arabidopsis*. Semin. Cell Dev. Biol. *7*, 427–433.

Bechtold, N., Ellis, J., and Pelletier, G. (1993). *In planta Agrobacterium* mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. CR Acad. Sci. (Paris) *316*, 1194–1199.

Bell, L.R., Maine, E.M., Schedl, P., and Cline, T.W. (1988). *Sex-lethal*, a Drosophila sex determination switch gene, exhibits sex-specific RNA splicing and sequence similarity to RNA binding proteins. Cell *55*, 1037–1046.

Bork, P., and Sudol, M. (1994). The WW domain: a signalling site in dystrophin? Trends Biochem. Sci. *19*, 531–533.

Brown, J.W.S. (1996). *Arabidopsis* intron mutations and pre-mRNA splicing. Plant J. *10*, 771–780.

Burd, C.G., and Dreyfuss, G. (1994). Conserved structures and diversity of functions of RNA-binding proteins. Science *265*, 615–621.

Burn, J.E., Smyth, D.R., Peacock, W.J., and Dennis, E.S. (1993). Genes conferring late flowering in *Arabidopsis thaliana*. Genetica *90*, 145–157.

Chan, D.C., Bedford, M.T., and Leder, P. (1996). Formin binding proteins bear WWP/WW domains that bind proline-rich peptides and functionally resemble SH3 domains. EMBO J. *15*, 1045–1054.

Chang, C., Bowman, J.L., DeJohn, A.W., Lander, E.S., and Meyerowitz, E.M. (1988). Restriction fragment length polymorphism linkage map for *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA *85*, 6856–6860.

Chen, H.I., and Sudol, M. (1995). The WW domain of Yes-associated protein binds a proline-rich ligand that differs from the consensus established for Src homology 3–binding modules. Proc. Natl. Acad. Sci. USA *92*, 7819–7823.

Clarke, J.H., and Dean, C. (1994). Mapping *FRI*, a locus controlling flowering time and vernalization response in *Arabidopsis thaliana*. Mol. Gen. Genet. *242*, 81–89.

Dean, C., van-den-Elzen, P., Tamaki, S., Dunsmuir, P., and Bedbrook, J. (1985). Differential expression of the eight genes of the petunia ribulose bisphosphate carboxylase small subunit multi-gene family. EMBO J. *4*, 3055–3061.

Dean, C., Sjodin, C., Page, T., Jones, J., and Lister, C. (1992). Behaviour of the maize transposable element *Ac* in *Arabidopsis thaliana*. Plant J. *2*, 69–81.

Dolezal, O., and Cobbett, C.S. (1991). Arabinose kinase-deficient mutant of *Arabidopsis thaliana*. Plant Physiol. *96*, 1255–1260.

Dreyfuss, G., Matunis, M.J., Pinol-Roma, S., and Burd, C.G. (1993). hnRNP proteins and the biogenesis of mRNA. Annu. Rev. Biochem. *62*, 289–321.

Furner, I.J., Ainscough, F.X., Pumfrey, J.A., and Petty, L.M. (1996). Clonal analysis of the late flowering *fca* mutant of Arabidopsis thaliana: cell fate and cell autonomy. Development *122*, 1041–1050.

Gao, F.-B., and Keene, J.D. (1996). Hel-N1/Hel-N2 proteins are bound to poly(A)+ mRNA in granular RNP structures and are implicated in neuronal differentiation. J. Cell Sci. *109*, 579–589.

Good, P.J. (1995). A conserved family of *elav*-like genes in vertebrates. Proc. Natl. Acad. Sci. USA *92*, 4557–4561.

Hicks, K.A., Sundas, A., and Meeks-Wagner, D.R. (1996). *Arabidopsis* early-flowering mutants reveal multiple levels of regulation in the vegetative-to-floral transition. Semin. Cell Dev. Biol. *7*, 409–418.

Jones, J.D.G., Shlumukov, L., Carland, F., English, J., Scofield, S.R., Bishop, G.J., and Harrison, K. (1992). Effective vectors for transformation, expression of heterologous genes, and assaying transposon excision in transgenic plants. Transgenic Res. *1*, 285–297.

Knecht, A.K., Good, P.J., Dawid, I.B., and Harland, R.M. (1995). Dorsal-ventral patterning and differentiation of noggin-induced neural tissue in the absence of mesoderm. Development *121*, 1927–1936.

Koornneef, M., Hanhart, C.J., and Van der Veen, J.H. (1991). A genetic and physiological analysis of late flowering mutants in *Arabidopsis thaliana*. Mol. Gen. Genet. *299*, 57–66.

Koushika, S.P., Lisbin, M.J., and White, K. (1996). ELAV, a Drosophila neuron-specific protein, mediates the generation of an alternatively spliced neural protein isoform. Curr. Biol. *6*, 1634–1641.

Lee, I., Bleecker, A., and Amasino, R. (1993). Analysis of naturally occurring late flowering in *Arabidopsis thaliana*. Mol. Gen. Genet. *237*, 171–176.

Lee, I., Aukerman, M.J., Gore, S.L., Lohman, K.N., Michaels, S.D., Weaver, L.M., John, M.C., Feldmann, K.A., and Amasino, R.M. (1994). Isolation of *LUMINIDEPENDENS*: a gene involved in the control of flowering time in Arabidopsis. Plant Cell *6*, 75–83.

Luehrsen, K.R., and Walbot, V. (1994). Intron creation and polyadenylation in maize are directed by AU-rich RNA. Genes Dev. *8*, 1117–1130.

Martinez-Zapater, J.M., and Somerville, C.R. (1990). Effects of light quality and vernalization on late-flowering mutants of *Arabidopsis thaliana*. Plant Physiol. *92*, 770–776.

Martinez-Zapater, J.M., Jarillo, J.A., Cruz-Alvarez, M., Roldán, M., and Salinas, J. (1995). *Arabidopsis late-flowering fve* mutants are affected in both vegetative and reproductive development. Plant J. *7*, 543–551.

Mattox, W., McGuffin, M.E., and Baker, B.S. (1996). A negative feedback mechanism revealed by functional analysis of the alternative isoforms of the Drosophila splicing regulator *transformer-2*. Genetics *143*, 303–314.

Nam, H.-G., Giraudat, J., den Boer, B., Moonan, F., Loos, W, Hauge, B.M., and Goodman, H.M. (1989). Restriction fragment length polymorphism linkage map of *Arabidopsis thaliana*. Plant Cell *1*, 699–705.

Newman, T., de Bruijn, F., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M., Retzel, E., and Somerville, C. (1994). Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous cDNA clones. Plant Physiol. *106*, 1241–1255.

Niwa, M., MacDonald, C.C., and Berget, S.M. (1992). Are vertebrate exons scanned during splice-site selection? Nature *360*, 277–280.

Olszewski, N.E., Martin, F.B., and Ausubel, F.M. (1988). Specialized binary vectors for plant transformation: expression of the *Arabidopsis thaliana* AHAS gene in *Nicotiana tobacum*. Nucl. Acids Res. *16*, 10765–10782.

Putterill, J., Robson, F., Lee, K., Simon, R., and Coupland, G. (1995). The *CONSTANS* gene of Arabidopsis promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. Cell *80*, 847–857.

Richter, K., Grunz, H., and Dawid, I.B. (1988). Gene expression in the embryonic nervous system of *Xenopus laevis*. Proc. Natl. Acad. Sci. USA *85*, 8086–8090.

Robinow, S., Campos, A.R., Yao, K.-M., and White, K. (1988). The *elav* gene product of *Drosophila,* required in neurons, has three RNP consensus motifs. Science *242*, 1570–1572.

Schmidt, R., and Dean, C. (1995). Hybridization analysis of YAC clones. Meth. Mol. Cell. Biol. *5*, 309–318.

Schmidt, R., West, J., Cnops, G., Love, K., Balestrazzi, A., and Dean, C. (1996). Detailed description of four YAC contigs representing 17Mb of chromosome 4 of *Arabidopsis thaliana* ecotype Columbia. Plant J. *9*, 755–765.

Simpson, G.G., and Filipowicz, W. (1996). Splicing of precursors to mRNA in higher plants: mechanism, regulation, and sub-nuclear organisation of the spliceosomal machinery. Plant Mol. Biol. *32*, 1–41.

Sudol, M., Bork, P., Einbond, A., Kastury, K., Druck, T., Negrini, M., Huebner, K., and Lehman, D. (1995). Characterization of the mammalian YAP (Yes-associated protein) gene and its role in defining a novel protein module, the WW domain. J. Biol. Chem. *270*, 14733–14741.

Swanson, M., and Dreyfuss, G. (1988). Classification and purification of proteins of heterogeneous nuclear ribonucleoprotein particles by RNA-binding specificities. Mol. Cell. Biol. *8*, 2237–2241.

Tuerk, C., and Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science *249*, 505–510.

Valvekens, D., Van Montagu, M., and Van Lisjebettens, M. (1988). *Agrobacterium tumefaciens* –mediated transformation of *Arabidopsis thaliana* root explants by using kanamycin selection. Proc. Natl. Acad. Sci. USA *85*, 5536–5540.

EMBL Accession Numbers

The accession numbers for the nucleotide sequences of transcripts α, β, γ, and δ are Z82993, Z82991, Z82989, and Z82990, respectively. The accession number for the genomic sequence of *FCA* is Z82992.