The Plant Journal (2001) **28**(6), 619–631

Functional importance of conserved domains in the flowering-time gene CONSTANS demonstrated by analysis of mutant alleles and transgenic plants

Frances Robson[1,†], M. Manuela R. Costa[1], Shelley R. Hepworth[1], Igor Vizir[1], Manuel Piñeiro[1], Paul H. Reeves[1], Joanna Putterill[1,†], and George Coupland[1,2,*]

[1]John Innes Centre, Colney Lane, Norwich NR4 7UH, UK and, [2]Max-Planck-Institut für Züchtungsforschung, Carl-von-Linné-Weg 10, D-50829 Köln, Germany

Received 11 July 2000; accepted 6 August 2001.
*For correspondence (fax +49 2215062207; e-mail coupland@mpiz-koeln.mpg.de).
[†]Present address: School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland, New Zealand.

Summary

CONSTANS **promotes flowering of** ***Arabidopsis*** **in response to long-day conditions. We show that** ***CONSTANS*** **is a member of an** ***Arabidopsis*** **gene family that comprises 16 other members. The CO-Like proteins encoded by these genes contain two segments of homology: a zinc finger containing region near their amino terminus and a CCT (CO, CO-Like, TOC1) domain near their carboxy terminus. Analysis of seven classical** ***co*** **mutant alleles demonstrated that the mutations all occur within either the zinc finger region or the CCT domain, confirming that the two regions of homology are important for CO function. The zinc fingers are most similar to those of B-boxes, which act as protein–protein interaction domains in several transcription factors described in animals. Segments of CO protein containing the CCT domain localize GFP to the nucleus, but one mutation that affects the CCT domain delays flowering without affecting the nuclear localization function, suggesting that this domain has additional functions. All eight** ***co*** **alleles, including one recovered by pollen irradiation in which DNA encoding both B-boxes is deleted, are shown to be semidominant. This dominance appears to be largely due to a reduction in** ***CO*** **dosage in the heterozygous plants. However, some alleles may also actively delay flowering, because overexpression from the CaMV 35S promoter of the** ***co-3*** **allele, that has a mutation in the second B-box, delayed flowering of wild-type plants. The significance of these observations for the role of** ***CO*** **in the control of flowering time is discussed.**

Keywords: Arabidopsis, flowering, photoperiod, B-box zinc finger.

Introduction

The *CONSTANS* (*CO*) gene was originally identified because of the late-flowering phenotype of *co* mutant plants (Koornneef *et al*., 1991; Redei, 1962). The phenotype of the mutant suggested that CO protein promotes the transition from vegetative growth to flowering, and this was supported by the demonstration that plants carrying extra copies of *CO* (Putterill *et al*., 1995) or overexpressing *CO* from the 35S promoter (Onouchi *et al*., 2000) flowered earlier than wild-type. The *CO* gene was cloned (Putterill *et al*., 1995), and the predicted protein product contains two regions of 43 amino acids towards the amino terminus of the protein that are closely related in sequence. Each of these regions contains an arrangement of cysteine residues similar to that present in the zinc fingers of GATA transcription factors, but little direct homology to these proteins was detected (Putterill *et al*., 1995). The construction of a translational fusion of CO to the ligand binding domain of the rat glucocorticoid receptor (CO:GR; Simon *et al*. 1996) provided further evidence that CO acts to influence transcription. Introduction of CO:GR into *co* mutant plants did not correct the mutant phenotype until the plants were treated with the steroid dexamethasone (dex). A similar fusion of the plant transcription factor LEAFY to the GR domain was retained in the cytoplasm until dex was added, suggesting that the GR domain operates in plants as it does in animals by sequestering the

fusion protein in the cytoplasm (Wagner *et al.*, 1999). Furthermore, induction of CO:GR with dex is associated with the rapid transcription of likely target genes, such as *SOC1* that encodes a MADS box transcription factor (Samach *et al.*, 2000). Taken together these data support the idea that CO acts in the nucleus to promote flowering by altering the transcription of downstream target genes.

Despite the considerable genetic and molecular data available on genes that interact with CO to regulate flowering (Kardailsky *et al.*, 1999; Kobayashi *et al.*, 1999; Onouchi *et al.*, 2000; Samach *et al.*, 2000; Soppe *et al.*, 2000), little is known of the function of the CO protein nor of the roles of the different domains of the protein. Here several molecular-genetic approaches are used to address the function of CO. The availability of the genomic sequence of *Arabidopsis* (Arabidopsis Genome Initiative, 2000) allowed us to identify an extensive family of related proteins and thereby to recognize domains conserved between these proteins that may be functionally important. The functional significance of these homologies is addressed by analysing the sequence of eight mutant alleles and determining the positions of the mutations within the CO protein.

The blocks of homology identified in the sequence comparisons and the positions of the mutations suggest that CO protein has a modular structure with two zinc fingers near the amino terminus and a domain of unknown function near the C-terminus. Transcription factors frequently have a modular structure in which for example the DNA-binding domain is separable from the transcriptional activation or repression domains (Hope and Struhl, 1986; Keegan *et al.*, 1986). Mutant proteins in which one domain is inactivated but the other is intact can act as dominant negative forms that repress the function of the wild-type protein and this has been used to investigate the function of plant transcription factors (e.g. Mizukami and Ma, 1997; Unger *et al.*, 1993). This approach was used to address the function of CO by making transgenic wild-type plants overexpressing mutant proteins in which one or other of the protein domains are affected.

The intracellular location of plant proteins can be determined by constructing translational fusions with green fluorescent protein (GFP; Haseloff *et al.* 1997; Grebenok *et al.*, 1997). This approach was also used to determine whether CO is present in the nucleus, and whether mutations that affect one of the protein domains prevent nuclear localization.

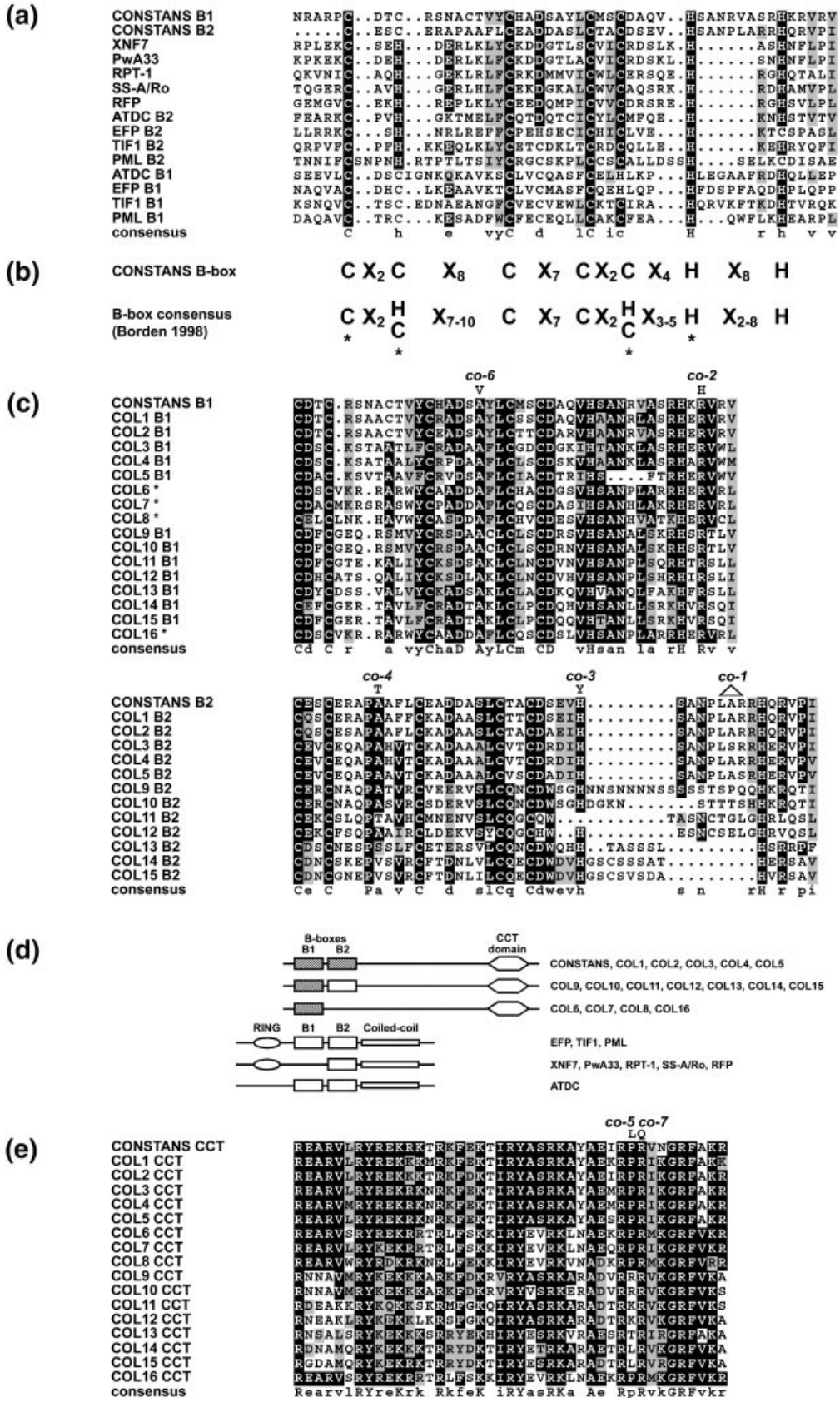
The results of these approaches are used to propose models of how CO acts to regulate flowering time.

Results

CONSTANS contains two B-boxes that are altered in five mutant alleles

Analysis of the CO protein using the SMART program (Schultz *et al.*, 1998, 2000) identified strong similarity between the proposed zinc fingers of CO and those of the B-box (Figure 1a and b). The B-box is a class of zinc finger, usually of the type C-X^2-H-X^7-C-X^7-C-X^2-C-X^5-H-X^2-H, that was identified in a variety of animal proteins including several transcription factors (XNF7, RPT-1, EFP), ribonucleoproteins (SS-A/Ro, PwA33) and proto-oncogene products (RFP, PML, TIF1) (reviewed in Borden, 1998; Reddy *et al.*, 1992). The CO protein contains two B-box motifs that show 46% identity and 86% similarity with each other (Figure 1c). There are seven conserved residues that could act as metal-binding residues within the B-box motif, and all of these are conserved in both CO B-boxes (Figure 1b). Four of these residues were shown to bind zinc in the B-box structure of the Xenopus protein XNF-7 (Borden *et al.*, 1995), and these four residues are conserved in both of the CO B-boxes (Figure 1b).

Figure 1. Comparison of CO with CO-Like proteins and with B-box containing proteins of animals.
(a) Comparison of the B-boxes of CO with several B-box proteins of animals. The animal B-box proteins are XNF7 (Miller *et al.*, 1989), PwA33 (Bellini *et al.*, 1993), RPT-1 (Pataraca *et al.*, 1988), SS-A/Ro (Chan *et al.*, 1991), RFP (Takahashi *et al.*, 1988), ATDC (Leonhardt *et al.*, 1994), EFP (Inoue *et al.*, 1993), TIF1 (Miki *et al.*, 1991) and PML (Goddard *et al.*, 1991). B1 and B2 are the most N-terminal or C-terminal B-box, respectively. (b) The consensus spacing of C and H residues in animal B-box proteins compared with the spacing between these residues in the CO B-boxes. X represents any amino acid. The asterisks indicate those residues predicted to bind zinc in the NMR structure of the XNF7 B-box (Borden *et al.*, 1995). (c) Alignment of the first and second B-boxes of CO with those of the CO-Like proteins. The predicted amino acid substitutions in *co-2*, *co-3*, *co-4* and *co-6*, as well as the predicted deletion in *co-1* are indicated. Accession numbers are CO (embl X94937; At5g15840; Putterill *et al.* 1995), COL1(embl Y1005; At5g15850; Putterill *et al.*, 1997), COL2(gb L81120; At3g02380; Ledger *et al.* 1996), COL3(gb AC006585; At2g24790), COL4(gb AF069716; At5g24930), COL5(dgb AB018118; At5g57660), COL6(gb AC011915; At1g68520), COL7(gb AC016662; At1g73870), COL8(gb AC016041; At1g49130), COL9(gb AC009176; At3g07650), COL10(dgb AB023039; At5g48520), COL11(emb Z97338; At4g15250), COL12(dgb AP000739; At3g21880), COL13(gb AC005309; At2g47890), COL14(gb AC002332; At2g33500), COL15(gbAC069471; At1g28050) and COL16(gbAC079281; At1g25440). (d) Schematic representation of the sequence motifs found within the CO and COL proteins, or within the animal B-box proteins. The CO and COL proteins contain one or two B-boxes near their N-terminus, and a C-terminal domain recently named the CCT domain (Strayer *et al.*, 2000). The second B-box of COL9, 10, 11, 12, 13, 14 and 15 is relatively dissimilar to that of CO and is shown as an open rectangle. Moreover, the second B-Boxes of COL9, COL10, COL11, and COL12 do not exactly match the B-box consensus and therefore may not be active B-Box domains. The animal proteins carry the illustrated arrangement of RING fingers, B-boxes and coiled-coil domains. (e) Alignment of the carboxy-terminal CCT domains of CO and COL1-16. The predicted amino acid substitutions in *co-5* and *co-7* are illustrated.


(a)
CONSTANS B1
CONSTANS B2
XNF7
PwA33
RPT-1
SS-A/Ro
RFP
ATDC B2
EFP B2
TIF1 B2
PML B2
ATDC B1
EFP B1
TIF1 B1
PML B1
consensus
(b)
CONSTANS B-box
C X2 C X8 C X7 C X2 C X4 H X8 H
B-box consensus
(Borden 1998)
C X2 H C X7-10 C X7 C X2 H C X3-5 H X2-8 H
(c)
co-6
co-2
CONSTANS B1
COL1 B1
COL2 B1
COL3 B1
COL4 B1
COL5 B1
COL6 *
COL7 *
COL8 *
COL9 B1
COL10 B1
COL11 B1
COL12 B1
COL13 B1
COL14 B1
COL15 B1
COL16 *
consensus
co-4
co-3
co-1
CONSTANS B2
COL1 B2
COL2 B2
COL3 B2
COL4 B2
COL5 B2
COL9 B2
COL10 B2
COL11 B2
COL12 B2
COL13 B2
COL14 B2
COL15 B2
consensus
(d)
B-boxes
B1
B2
CCT domain
CONSTANS, COL1, COL2, COL3, COL4, COL5
COL9, COL10, COL11, COL12, COL13, COL14, COL15
COL6, COL7, COL8, COL16
RING
B1
B2
Coiled-coil
EFP, TIF1, PML
XNF7, PwA33, RPT-1, SS-A/Ro, RFP
ATDC
(e)
co-5 co-7
CONSTANS CCT
COL1 CCT
COL2 CCT
COL3 CCT
COL4 CCT
COL5 CCT
COL6 CCT
COL7 CCT
COL8 CCT
COL9 CCT
COL10 CCT
COL11 CCT
COL12 CCT
COL13 CCT
COL14 CCT
COL15 CCT
COL16 CCT
consensus

Two *CO-LIKE* (*COL*) genes were previously described in *Arabidopsis* (Figure 1c; Ledger *et al.*, 1996, 2001; Putterill *et al.*, 1997). A BLAST search of the *Arabidopsis* genome sequence was performed to determine whether there were other members of this gene family. By analysing the *Arabidopsis* genome sequence (Arabidopsis Genome Initiative, 2000) a total of 16 *COL* genes were identified that contained one or two B-boxes at the amino terminus of the predicted protein (Figure 1c), and another highly conserved domain at the carboxy terminus (see below). Five COL proteins have two B-box motifs closely related to those of CO, seven contain a second B-box less closely related to those of CO and four have only one B box motif.

The *CO* gene was amplified by PCR from each of the seven mutants and the resulting fragments sequenced to identify the mutations. Five of the seven *co* alleles contain mutations that affect residues in the B-boxes (Figure 1c), suggesting that these are important for CO protein function. The *co-6* mutation causes substitution of an alanine for a valine in the first B-box. The *co-2* mutation converts an arginine to a histidine towards the carboxy-terminal end of the first B-box. The *co-1*, *co-3* and *co-4* mutations affect the second B-box. The *co-3* mutation affects a histidine residue that, based on the analysis of the XNF-7 B-box structure, is likely to be required for zinc binding (Figure 1a, b and c; Borden *et al.*, 1995).

CONSTANS contains a carboxy-terminal domain that is conserved among related proteins and is functionally important

The two remaining mutant alleles, *co-5* and *co-7*, do not affect the B-box structures. These mutations affect adjacent proline and arginine residues in a highly conserved basic domain of approximately 43 amino acids near the C-terminus of the protein (Figure 1e). This novel domain was previously proposed to contain a nuclear localization sequence (Robert *et al.*, 1998), and is also found in all 16 COL proteins (Figure 1e). Homology to this domain was also recently described in proteins that do not contain B boxes (Kurup *et al.*, 2000; Makino *et al.*, 2000; Strayer *et al.*, 2000).

The carboxy-terminal domain of CONSTANS is sufficient to localize GFP to the nucleus, but the co-7 *allele does not affect nuclear localization*

To test whether CO protein is localized to the nucleus, and whether this is conferred by the carboxy-terminal domain translational fusions were constructed between GFP and CO or CO derivatives.

A translational fusion of GFP and CO (GFP:CO) was constructed in plasmid pAVA121 (von Arnim *et al.*, 1998; Experimental procedures). GFP was fused to the N-terminus of the CO protein and expressed from a double 35S promoter (Experimental procedures). A transient expression assay in onion bulb epidermal cells was used to assess the cellular location of the fusion protein. Cells bombarded with the control plasmid pAVA121 showed GFP localization both in the cytoplasm and in the nucleus (Figure 2a), as previously shown (Grebenok *et al.*, 1997; Haseloff *et al.*, 1997). However, GFP:CO fusion protein was located in the nucleus and was not detected in the cytoplasm (Figure 2b). CO can therefore localize GFP to the nucleus of these cells, suggesting that CO is a nuclear protein. To confirm that GFP:CO retained biological activity, the 35S::GFP:CO fusion was introduced into *co-2* mutants. The transgenic plants were early flowering and showed a similar phenotype to 35S::CO (Onouchi *et al.*, 2000). The cellular location of the fusion protein was analysed in root, hypocotyl and cotyledon cells and shown to be located in the nucleus (Figure 2g, h and i).

To determine whether the conserved domain at the C-terminus of the CO protein is responsible for nuclear localization, a series of CO deletion derivatives were fused to GFP. The plasmid containing GFP fused to the region of CO between amino acids 304 and the C-terminal amino acid 373 (GFP:CtermCO; Experimental procedures) was bombarded into onion cells. The GFP:CtermCO fusion protein was observed only in the nucleus (Figure 2c), suggesting that this C-terminal sequence is sufficient to target GFP to the nucleus. Some proteins contain multiple NLSs (Varagona *et al.*, 1992), and therefore to test whether another NLS was present in the N-terminal portion of CO, a truncated CO-protein lacking the C-terminal residues 302–373 was produced (GFP:NtermCO). This fusion protein was tested for intracellular localization as described earlier. Fluorescence was observed both in the cytoplasm and nucleus (Figure 2d), demonstrating that this portion of CO does not localize GFP exclusively to the nucleus. Taken together these experiments suggest that the only region of CO containing an NLS is between amino acids 304 and 373.

The mutant alleles *co-5* and *co-7* contain mutations in the C-terminal region of the protein that was demonstrated above to be sufficient for nuclear localization. The C-terminal regions (between amino acids 304 and 373) from these mutant proteins were fused to GFP to determine whether the mutations affected nuclear localization of the protein. Cells bombarded with *35S::GFP:co-5* DNA showed GFP localization in the cytoplasm and in the nucleus (Figure 2e), suggesting that the *co-5* mutation affects the subcellular localization of the CO protein. The distribution of GFP:*co-5* was similar to that of the GFP control but the level of expression was lower. However, GFP:*co-7* showed GFP localization only in the nucleus (Figure 2f), suggesting that this mutation does not affect nuclear import of the CO protein.

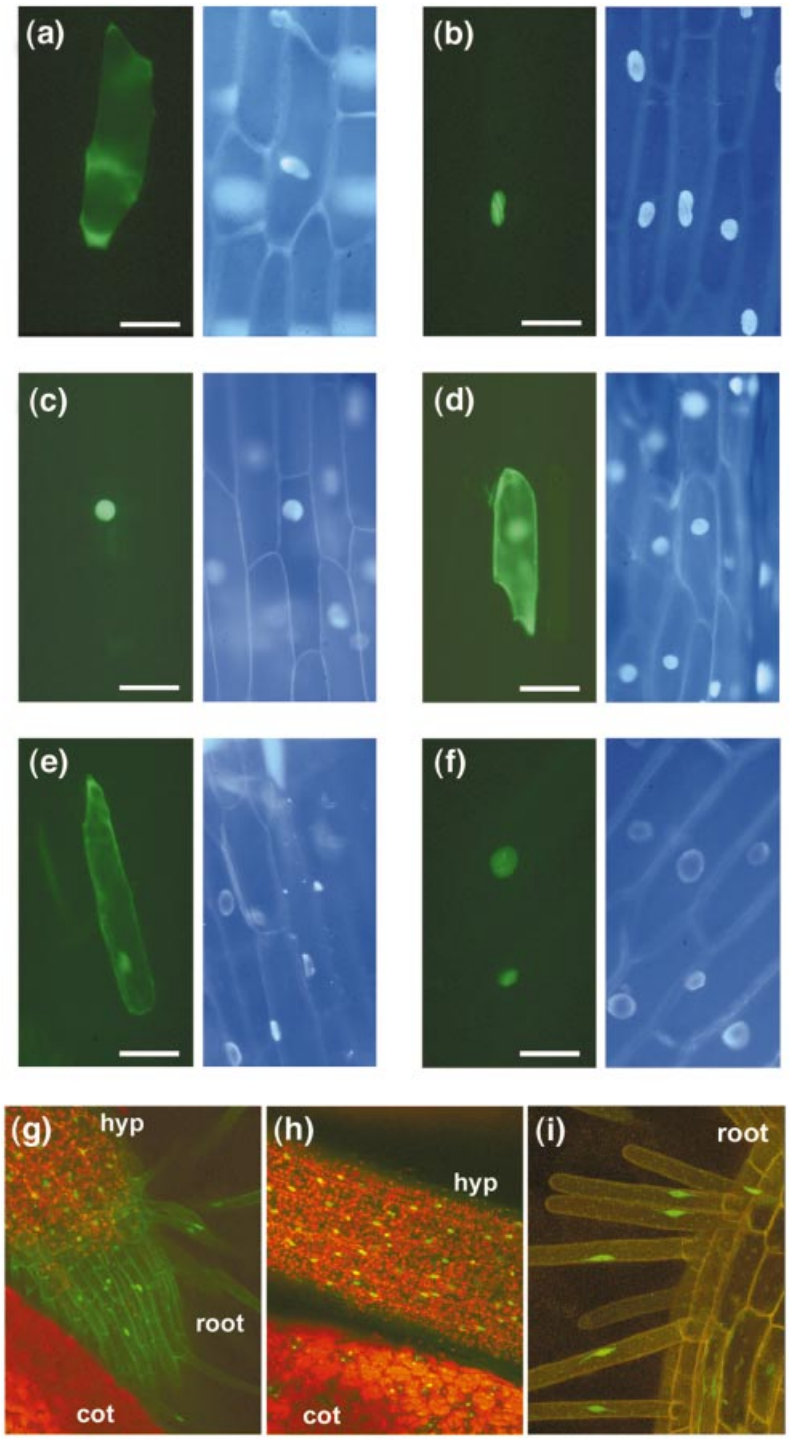


Figure 2. Subcellular localization of CO:GFP fusion protein.
(a)–(f) Images of onion epidermal cells stained with DAPI. In each case the tissue was viewed using epifluorescence optics with blue excitation to detect GFP (left) and UV-excitation to detect nuclei. Cells bombarded with 35S::GFP. Bars = 100 μm. (a) Cells bombarded with 35S::GFP. (b) Cells bombarded with 35S::GFP:CO. (c) Cells bombarded with GFP:CtermCO. (d) Cells bombarded with GFP:NtermCO. (e) Cells bombarded with GFP:*co-5*. (f) Cells bombarded with GFP:*co-7*. All samples were stained with DAPI and viewed under epifluorescence optics with blue (left) and UV (right) excitation. (g)–(i) Images of 11-day-old-transgenic *Arabidopsis* plants carrying 35S::CO:GFP. (g) Junction of hypocotyl and root. Merged images of green and red channels. GFP fluorescence detected in the green channel and chlorophyll autofluorescence in the red channel. (h) Hypocotyl tissue. Merged images of green and red channels as for G. (i) Root hairs.

Isolation and characterization of co-8*, a likely null allele*

Based on their sequences none of the seven classical *co* mutant alleles were predicted to certainly abolish CO function. All seven caused in-frame changes, including *co-1* which was induced with X-rays. This raised the possibility that complete loss of function may generate pleiotropic effects, and that the seven classical alleles were all hypomorphs identified by screening phenotypically for late-flowering plants. Therefore, pollen irradiation was used to identify loss-of function alleles without making the assumption that these would only cause late flowering. An

appropriate population of plants made with γ-irradiated pollen was described previously (Vizir *et al.*, 1994). Pollen from Landsberg *erecta* plants was irradiated and used to cross-fertilize plants homozygous for the genetically linked mutations *lu co-1 ms1 ttg*. Twenty-six late-flowering M1 plants were identified, and these potentially carried novel *co* mutant alleles derived from irradiation. These plants were self-fertilized, and six of them gave rise to M2 progeny that were all late flowering, although only one of these lines was fertile. The mutation in this late-flowering line was preliminarily called *co-8*, and was tested further at the genetic and molecular levels.

Southern analysis of *co-8* using probes derived from the *CO* genomic region demonstrated that a deletion had removed 1.3 kb from the 5′ end of the *CO* gene. The deletion included the sequence encoding the two B-boxes and approximately 1 kb of the 5′ non-coding region (Figure 3a). However, the pattern of hybridizing fragments detected in the Southern analysis could not be explained by a simple deletion. For example, probes that flanked the deletion on either side did not hybridize to the same *Eco*RI or *Hin*dIII fragments, although they would have been predicted to do so had a simple deletion occurred. This suggested that an insertion or an inversion had probably occurred at the deletion breakpoints.

To analyse the structure of *co-8* more carefully both junctions between *CO* and the presumed rearrangement were isolated by using Inverse Polymerase Chain Reaction (IPCR) and by constructing a cosmid library from DNA extracted from *co-8* mutants (Robson, 1998; Experimental procedures). The exact locations of the breakpoints were defined by DNA sequencing using primers (co52 and co17 in Experimental procedures; Figure 3a) designed to the *CO* gene to sequence into the rearrangement on either side. This identified the breakpoints within *CO* to be 1025 bp upstream of the ATG and 304 bp downstream of the ATG. Adjacent to the breakpoints at both ends of the rearrangement was novel DNA not associated with the *CO* gene in wild-type plants. Both of these unknown sequences were used in a BLAST search against the database and both were identical to sequences within P1 clone MPI10 that contains DNA from *Arabidopsis* chromosome 5. The two junction sequences however, are not directly adjacent in the sequence of MPI10 but 1208 bp apart. On the wild-type chromosome, the *CO* gene (on BACF14F8) is approximately 17 Mb from the DNA within clone MPI10 according to the physical maps of chromosome 5 (TAIR). The *co-8* allele was probably therefore derived from an irradiation-induced event in which two deletions occurred approximately 17 Mb apart. One of these was within the *CO* gene, and the other within the DNA cloned in BAC MPI10. These deletions were then repaired such that the intervening 17 Mb segment of chromosome 5 was inverted (Figure 3a).

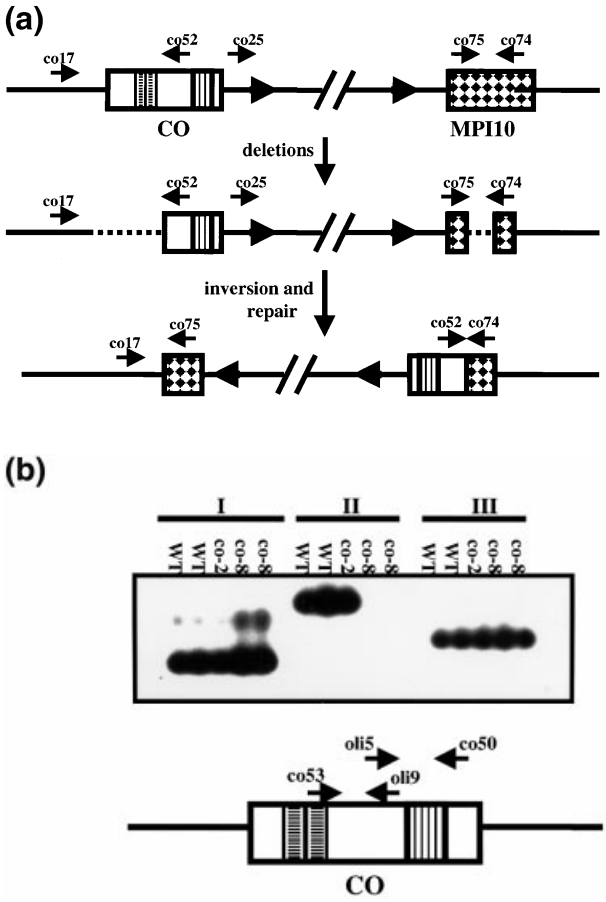


Figure 3. Structure of the *co-8* allele.
(a) Proposed derivation and final structure of the *co-8* allele. The first diagram illustrates the structure of the wild-type chromosome. The *CO* gene is illustrated as a rectangle containing the B-boxes (marked with internal horizontal lines) and the CCT domain (marked with internal vertical lines). DNA within BAC MPI10 is marked with black squares. CO and MPI10 are approximately 17 Mb apart, and the orientation of this intervening DNA is denoted by arrowheads. The positions of primers used to analyse the structure of the mutant allele are illustrated, their use is described in the text and their sequences appear in the Experimental procedures. The second diagram shows the location of two deletions that are proposed to have occurred in the generation of the *co-8* allele, and their positions relative to the primer sequences. The third diagram illustrates the proposed final structure of the allele. The intervening DNA is inverted such that the MPI10 sequence containing co75 is adjacent to CO sequence containing co17 and MPI10 sequence containing co74 is adjacent to MPI10 sequence containing co74. (b) Analysis of *CO* RNA present in *co-8* mutants by RT–PCR. Experiment I was performed using primers oli5 and co50 (Experimental procedures). These detected the 3′ end of the wild-type *CO* mRNA extending from the single intron over the region encoding the CCT domain. Experiment II was performed using primers oli9 and co53 (Experimental procedures). These detected the 5′ end of the wild-type CO mRNA extending from the single intron to the position at which co53 anneals within the first exon, as shown in the diagram. Experiment III shows the detection of the *APETALA2* mRNA as a control for the abundance of cDNA used in each of the previous RT–PCR experiments.

In the *co-8* allele 1025 bp of the 5′ untranslated region is deleted and the truncated *CO* ORF is fused to DNA from BAC MPI10 that is not normally associated with the *CO* gene. Reverse transcriptase–PCR was used to test whether

CO mRNA was present in *co-8* plants (Figure 3b). No transcript was amplified from cDNA made from *co-8* mutants when one of the primers (co53) used for the PCR was designed to anneal to DNA removed by the deletion in *co-8*, although cDNA made from wild-type plants produced a transcript of the expected size. However, PCR primers (co50 and oli5) that annealed to part of the CO ORF that is retained in *co-8*, amplified a product of the expected size from cDNA made from both *co-8* mutants and wild-type plants. Therefore, the remaining portion of the CO ORF is still transcribed in the *co-8* rearrangement.

Although a novel transcript is detected in the *co-8* mutant, the deletion of both B-boxes suggests that the mutant will lack any CO activity.

All eight constans *alleles are semidominant*

The *co-2*, *co-3* and *co-4* alleles were previously reported to be semidominant (Koornneef *et al.*, 1991). Based on the phenotype of homozygous *co-2* mutant plants carrying transgenic copies of the wild-type gene, Putterill *et al.* (1995) proposed that the semidominance of *co-2* was due to haploinsufficiency rather than to *co-2* encoding an altered product that delayed flowering.

To determine whether all seven classical alleles and the new *co-8* allele all cause semidominance, they were each independently crossed to Landsberg *erecta*. The F1 progeny were sown in long-day conditions and their flowering time compared to that of wild type and homozygous mutant controls. All of the F1 plants showed an intermediate flowering time phenotype (Table 1), indicating that all eight alleles are semidominant. This was confirmed in the F2 generation, in which approximately 50% of plants showed intermediate flowering times (data not shown).

Transgenic wild-type plants over-expressing the co-3 protein are late flowering

The *CO* gene contains two functional domains based on homology searches and analysis of mutant alleles. Both of these domains may facilitate interactions between CO and other proteins (see Discussion). This suggested that mutant forms of *CO* in which one domain is altered but the other is intact might sequester interacting proteins into inactive complexes, and thereby lead to a late-flowering phenotype. Such a dominant negative function has been proposed recently to explain the effect of mutant forms of B-box proteins (Peng *et al.*, 2000). To test the effectiveness of this approach, the *co-3* allele, that carries a mutation in the second B-box of CO, and the *co-7* allele, that carries a mutation in the carboxy-terminal domain, were each expressed from the CaMV 35S promoter (Experimental procedures). These transgenes were then introduced into wild-type Landsberg *erecta* plants.

Table 1. Flowering time of wild type and the *constans* homozygous and heterozygous mutants, measured as the total number of leaves produced before the onset of flowering. Data from 20 individuals for each genotype ± SE

Genotype	Homozygote (leaf number)	Heterozygote (leaf number)	Mutagen
WT	7.5 ± 0.2	–	–
co-4	11.3 ± 0.4	9.6 ± 0.2	EMS
co-5	14.4 ± 0.4	11.5 ± 0.4	EMS
co-2	19.4 ± 0.9	13.3 ± 0.3	EMS
co-8	22.5 ± 0.3	16.2 ± 0.7	γ – ray
co-6	23.7 ± 0.5	11.9 ± 0.3	EMS
co-1	25.6 ± 0.6	13.5 ± 0.3	X – ray
co-7	28.5 ± 0.8	10.2 ± 0.2	EMS
co-3	29.1 ± 1.2	15.2 ± 0.6	EMS

Approximately 150 kanamycin-resistant T1 plants were identified after infiltration of Landsberg *erecta* plants with *Agrobacterium* cells carrying the *35S::co-3* construct. Around 20 of these T1 plants appeared to flower at least slightly later than wild-type plants. The T1 plants were self-fertilized and individuals homozygous for the T-DNA were identified in five of the late-flowering lines. T3 progeny of these five lines were then scored for flowering time under long-day conditions (Figure 4). All five lines flowered significantly later then wild-type plants, although none were as late flowering as the *co-2* or *co-3* mutant. Northern blots demonstrated that *co-3* mRNA was over-expressed in the late-flowering *35S::Co-3* lines, and therefore that the late-flowering phenotype was not due to cosuppression causing a reduction in expression of the *co-3* and *CO* mRNAs. This indicates that over-expression of the *co-3* allele, that carries a mutation in one of the B-boxes, can delay flowering of wild-type plants.

A similar experiment was performed with the *35S::co-7* construct and a total of 29 kanamycin-resistant T1 plants were identified. However, none of these flowered later than wild-type plants. These were self-fertilized and in the T2 generation their flowering time was tested under long and short days. None of the transformants flowered late under long days, suggesting that the *35S::co-7* construct did not generate a dominant negative phenotype. Some of the *35S::co-7* transformants flowered earlier than wild-type under short days, indicating that the protein encoded by the *co-7* allele may retain some residual CO activity.

Discussion

CO is a member of a novel family of B-box containing proteins

The CO zinc finger regions are most similar to those of B-box proteins. The seven potential zinc-binding residues

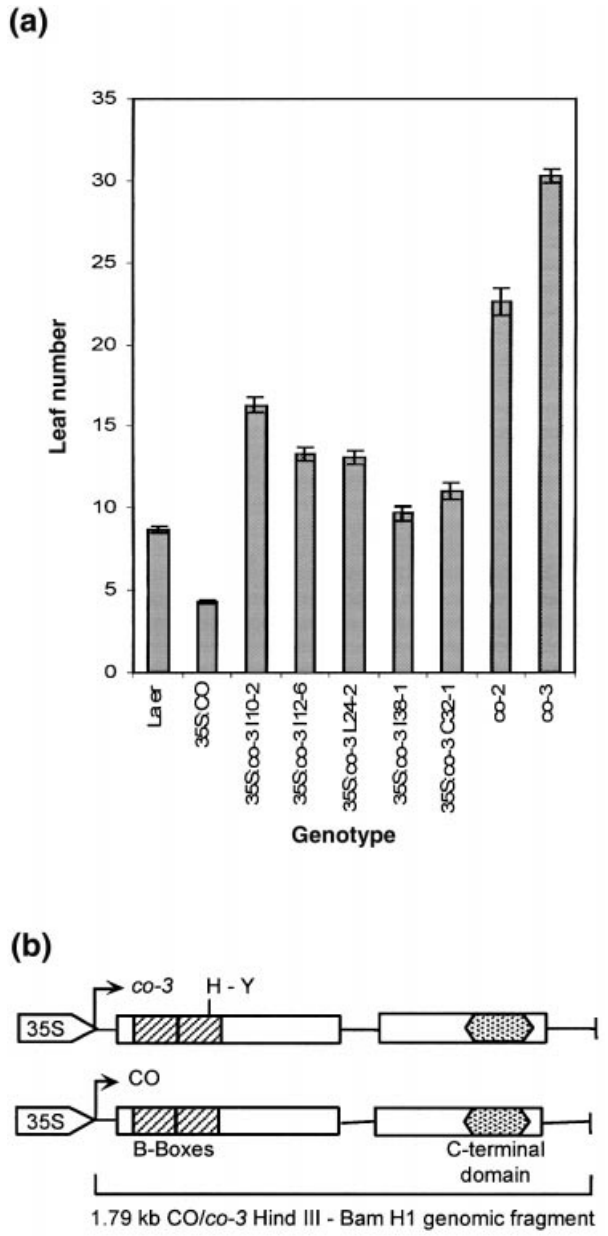


Figure 4. The effects of overexpression of the *co-3* allele. (a) Flowering times of Landsberg *erecta, co-2, co-3* and *35S::CO* plants compared to the flowering times of five independent *35S::co-3* transformants. The total numbers of leaves formed prior to the onset of flowering are shown. Error bars represent the SE for 20 individulas for each genotype growing under long day conditions. (b) A schematic illustration of the *35S::CO* and *35S::co-3* constructs.

within the B-box consensus sequence are conserved in the CO and most of the COL B-boxes. Furthermore, the *co-3* allele, which causes the most severe delay in flowering time, alters a histidine in the second B-box that corresponds to a histidine shown to bind zinc in the solution structure of the B-box of the Xenopus protein XNF-7 (Borden *et al.*, 1995). Although the zinc fingers of CO were originally compared with those of GATA transcription factors (Putterill *et al.*, 1995), based on the spacing of four of the cysteine residues within the CO fingers, the similarity to the more recently described B-boxes is much stronger. Putterill *et al.* also pointed out the lack of direct homology with GATA transcription factors (Putterill *et al.*, 1995).

The function of B-box proteins in plants has not previously been discussed. In animal proteins, the B-box domain is usually part of a tripartite motif comprising a zinc-binding RING finger and a B-box domain followed closely (5–8 amino acids) by a predicted α-helical coiled-coil domain (RBCC family; Figure 1). The spacing between the three elements is highly conserved suggesting that the relative position of the domains is of functional importance. Proteins in a subfamily of this group, defined by the gene for ataxia–telangiectasia group D (ATDC) (Leonhardt *et al.*, 1994), have one or two B-boxes and a coiled coil domain but no RING finger. Another variation in the arrangement of RBCC domains is found in the protein kinase C-interacting protein (RBCK1) which has two coiled coil domains followed by a RING finger, a B-box and a B-box-like domain. This is the only published example of a protein that does not contain the coiled-coil domain after the B-box motif (Tokunaga *et al.*, 1998).

CO and the other COL proteins are unusual in containing one or two B-box domains with no coiled-coil domain or RING finger. There are several plant proteins containing RING fingers (COP1, Deng *et al.*, 1992; A-RZF, Zou and Taylor, 1997; PRT1 Potuschak *et al.*, 1998) and in COP1 this is followed by a coiled-coil domain (Deng *et al.*, 1992). However none of these plant RING finger proteins contain B-boxes.

The RBCC motif is believed to mediate protein–protein interactions (Borden, 1998; El-Husseini *et al.*, 2000; Peng *et al.*, 2000; Tsuzuki *et al.*, 2000). However, the two B-boxes in CO may not function in a similar way to those in the RBCC motif. For example, the RBCC domain of the transcriptional corepressor KAP-1 appears to function as an integrated structural unit in which the RING finger, the B-box and the coiled-coil region are all required for interaction with the transcriptional repression module KRAB (Peng *et al.*, 2000). However, in other cases the B-boxes appear to function autonomously. For example, the transcription factor GATA-2 interacts specifically with the B-box region of promyelocytic leukaemia protein (Tsuzuki *et al.*, 2000).

The role of the conserved carboxy-terminal domain of CO

The carboxy-terminal region of CO was sufficient to direct GFP to the nucleus, suggesting that nuclear localization is one function of this region. Such a function for this region

was originally proposed based on the similarity of a portion of it to the consensus sequence for an NLS (Robert *et al.*, 1998). More recently, a related region in the TOC1 protein was shown to direct TOC1 to the nucleus in transient expression assays, and was termed the CCT (CO, COL, TOC1) domain (Kurup *et al.*, 2000; Makino *et al.*, 2000; Strayer *et al.*, 2000). Nevertheless the carboxy-terminal domain of TOC1 shows only 51% identity to that of CO whereas the least closely related COL protein, COL14, shows 60.5% identity. The experiments described here establish that the CCT domain in CO shares the nuclear localization function with the related domain in TOC1, and the early flowering of the 35S::CO:GFP plants confirmed that CO:GFP located in the nucleus retains biological activity.

However, in addition to nuclear localization, the CCT domain probably has other functions. This was originally suggested because the conserved region is 43 amino acids long, which is a longer stretch of contiguous homology than is shown by nuclear localization sequences (Raikhel, 1992). The demonstration that the *co-7* allele has a severe effect on flowering time but does not affect the nuclear localization function of this domain further suggests that the domain has an additional role in CO activity.

A CCT domain is present in at least 18 proteins, including TOC1, that do not contain B-boxes (Strayer *et al.*, 2000 and data not shown). There are also a further 13 proteins that contain one or two B-boxes but do not have the CCT domain. These include the salt-tolerance protein STO (Lippuner *et al.*, 1996). The existence of proteins containing only one of these domains, either B-box or CCT-domain, suggests that these domains act independently of one another. This is supported by the observation of Kurup *et al.* (2000) who showed that the CCT domains of CO and TOC1 (also called ABI3 Interacting Protein 1) interact in yeast cells with the *Arabidopsis* transcription factor ABI3. This interaction was reduced approximately two-fold by both the *co-5* and *co-7* mutations. Therefore, the carboxy-terminal region probably has a role in protein–protein interaction as well as in nuclear localization.

The dominance of the co *mutations*

Three *co* mutant alleles were previously shown to be semidominant with the heterozygotes showing a phenotype intermediate between the homozygous mutants and wild-type (Koornneef *et al.*, 1991; Redei, 1962). Putterill *et al.* (1995) proposed that this was likely to be caused by haploinsufficiency, in which the heterozygotes did not produce enough CO protein to promote early flowering, rather than the mutant allele encoding an altered gain of function protein. This was proposed because transgenic mutants homozygous for the *co-2* allele and carrying wild-type *CO* as a transgene flowered as early as wild-type plants. We have now shown that all eight mutant alleles are semidominant. The novel *co-8* allele, which we isolated, may be a null allele because the DNA encoding the translational start site and both B-boxes is deleted, although the remaining portion of the *co-8* mRNA may still be translated to produce a truncated protein. This truncated protein would carry the CCT domain (Figure 3) and may actively delay flowering, as was shown for overexpression in wild-type plants of the *co-3* allele, which also carries an intact CCT domain and impaired B-box domain. In contrast, overexpression in wild-type plants of the *co-7* allele, which encodes intact B-boxes and an impaired CCT domain, did not delay flowering, although this allele was semidominant when tested in heterozygous plants (Table 1). Therefore, the observation that *co-8* and *co-7* alleles are semidominant is consistent with the proposal that the dominance of *co* mutations is caused by haploinsufficiency.

Nevertheless, at least for some alleles the semidominance may be caused by a combination of haploinsufficiency and the mutant allele encoding an altered product that actively delays flowering. The late-flowering phenotype of Landsberg *erecta* plants carrying the 35S::*co-3* transgene clearly indicates that at least when overexpressed this allele can actively delay flowering. The *co-3* mutation affects a histidine residue that is predicted to be essential for zinc-binding within the second B-box. The active delay in flowering time caused by overexpression of this protein may be a consequence of the co-3 protein sequestering wild-type CO protein or proteins required for CO function into inactive complexes. The sequestration may occur by proteins binding to the first B-box of co-3 or to the CCT domain, neither of which is affected by the co-3 mutation.

Implications for the roles of CO and CO-Like proteins in regulating flowering time

CO is a nuclear protein (Figure 2) that acts to promote flowering time by rapidly inducing the expression of downstream flowering-time genes such as *SOC1* and *FT* (Samach *et al.*, 2000). The zinc fingers of CO are required for CO function and are most similar to B-box motifs, which are predicted to mediate protein–protein interaction rather than DNA binding. This suggests that to activate transcription of downstream genes, CO may be recruited to promoters by DNA binding proteins. Such a role for B-box proteins has been described in animals. For example, the transcription factor GATA-2 recruits the B-box protein promyelocytic leukemia (PML) to DNA, and PML enhances the ability of GATA-2 to activate transcription (Tsuzuki *et al.*, 2000). Similarly, the Krüppel associated box (KRAB) that acts as a transcriptional repression module must interact with the RBCC protein KAP-1 in

order to cause gene silencing (Peng *et al.*, 2000). The KAP-1 protein is recruited to DNA by zinc-finger DNA binding proteins that carry the KRAB domain. These examples may describe a paradigm for CO function, and suggest that it may interact with specific DNA binding proteins that enable its recruitment to DNA. The observation that the carboxy-terminal regions of CO and TOC1 will interact with the DNA-binding protein ABI3 (Kurup *et al.*, 2000), suggests that ABI3 or transcription factors of the same class may be responsible for the recruitment of CO and COL proteins to DNA.

The evolution of the family of 16 COL proteins that contain the B-boxes and the carboxy-terminal domain was recently discussed (Lagercrantz and Axelsson, 2000), but their function is unknown. Overexpression of *COL1* shortened the period length of circadian clock regulation, but did not affect flowering time (Ledger *et al.*, 2001). In some cases the B-boxes are closely related in sequence to those of CO (Figure 1), however, so far there is no evidence that they regulate flowering time, and they may interact with transcription factors that do not associate with CO, and thereby regulate a different set of target genes. Closely related RBCC proteins were previously shown to interact with specific protein partners (Cainarca *et al.*, 1999).

Further understanding of the function of the CO and COL family is likely to come from identifying interacting proteins, some of which may recruit the B-box proteins to specific sets of target genes.

Experimental procedures

Plant material and growth conditions

Seeds from Landsberg *erecta* – Ler-0 (NW20), *tt4–1* (N85) and EMS mutants *co-2* (N175), *co-3* (N176), *co-4* (N177), *co-5* (N178), *co-6* (N179) and *co-7* (N180) were obtained from M. Koornneef. These mutants are all in a Landsberg *erecta* background. Seeds from Redei's X-ray mutant *co-1* (N3122) were also provided by M. Koornneef. This mutant is in Landsberg – La-0 (N1298). *co-1* is also available in Landsberg *erecta* – *co-1 er-1* (N3135). Seeds from γ-irradiated *lu-1 co-1 ms1–1 ttg-1* (N240) were provided by I. Vizir.

In the summer plants in the glasshouses were grown in natural daylight. In the winter supplementary light was provided so that the minimum day length was 16 h. Flowering time was measured under defined conditions by growing plants in Sanyo Gallenkamp as described by Putterill *et al.* (1995) and Robson (1998).

DNA and RNA extraction

Plant genomic DNA was extracted as described by Dean *et al.* (1992). To make the *co-8* cosmid library the DNA was further purified on a caesium chloride gradient prior to digestion (Sambrook *et al.*, 1989). RNA for analysis of *co-8* by RT–PCR was extracted as described by Putterill *et al.* (1995).

Cloning and sequencing of the co mutant alleles

DNA was extracted from seedlings as described above. A pair of primers designed to amplify the *CO* gene had previously been designed (Putterill *et al.*, 1995); co41 (5′-GGTCCCAACGA-AGAAGTGC-3′) and co42 (5′-CAGGGAGGCGTGAAAGTGT-3′). These were used to amplify a 1.95-kb fragment from wild-type and co mutants *co-1* to *co-7*, in duplicate PCR reactions. The PCR products were blunt-ended using T4 DNA polymerase and cloned into the *Eco* RV site of pBluescript (SK +).

Library construction and screening

DNA from the *co-8* mutant was extracted and purified as described above. The library was constructed as described in Schaffer *et al.* (1998) by ligating plant DNA partially digested with *Sau* 3 A into the *Bam* HI site of cosmid vector c04541 (Jones *et al.*, 1992). The recombinant cosmids were packaged in Gigapack II Gold packaging extracts (Stratagene, La Jolla, CA, USA) according to the manufacturer's instructions and plated using *Escherichia coli* XL1 Blue MR.

Analysis of expression by RT–PCR

RNA was extracted from duplicate samples of tissue from 10-day-old-seedlings as described above and cDNA prepared as described by Putterill *et al.* (1995). Primer pairs used were as follows: to amplify the 5′ end of the CO transcript primers co53 (5′-ACGCCATCAGCGAGTTCC-3′) and oli9 (5′-AAATGTATGC-GTTATGGTTAATGG-3′) were used. To amplify the 3′ end of the CO transcript primers co50 (5′-CTCCTCGGCTTCGATTTCTC-3′) and oli5 (5′-CATTAACCATAAC-GCATACATTTC-3′) were used. Oli5 and Oli 9 were designed to anneal to the exon sequence either side of the single CO intron to prevent the amplification of contaminating DNA (Simon *et al.*, 1996). The position of the intron is marked in the primer sequence by a hyphen. To amplify the *APETALA2* cDNA as a control, primers AP2 Oli3 (5′-CTCAATGCCG-AGTCATCAGG-3′) and AP2 Oli4 (5′-CATG-AGAGGAGGTTGGAAGC-3′) were used. The resulting PCR products were fractionated on an agarose gel, Southern blotted onto Hybond N+ (Amersham, Little Chalfont, UK) according to the manufacturer's instructions and probed with the *CO* cDNA.

Primers used to analyse the co-8 *rearrangement*

The following primers were used to identify and characterize the *co-8* rearrangement:

co25 (5′-TACTGTTGTGCAAATGG-3′)and co52 (5′-GGAACAGCC-ACGAAGCAAC-3′) were used in the IPCR experiment to amplify the DNA adjacent to the deletion in *co-8*.

co17 (5′-ATGGATCATGTGGACTAG-3′)anneals in the CO promoter and was used to first identify the inversion in *co-8*.

co74 (5′-GATGGGCTACGTATGCGGC-3′)and co75 (5′-GGACTA-AGCATATACGACACATCTC-3′) were designed to anneal to DNA brought adjacent to each side of the deletion in *co-8* by the inversion. In wild-type they anneal to DNA within P1 clone MPI10.

Construction of transformation vectors

p35S::co3 was constructed by first isolating the CaMV 35S promoter as a 350 bp *Cla* I – *Hin*d III fragment from pJIT62 (Guerineau *et al.*, 1992) and cloning it into these sites in

pBluescript. A 1.7 kb *Hin*d III fragment containing the *co-3* genomic region, including the native *CO* polyadenylation sequences, was isolated from pco-3. *Hin*d III cuts in this clone in the polylinker near the 3′ end of the gene and also in the plant DNA 70 bp upstream of the ATG. This was cloned into the *Hin*d III site adjacent to the 35S promoter and orientated correctly by restriction mapping. *35S:co3* was moved as a *Cla* I – *Bam* HI fragment into the binary vector pSLJ1711 (Jones *et al.*, 1992). p35S::CO was constructed in essentially the same way, as described by Onouchi *et al.* (2000).

Transformation of Arabidopsis

Landsberg *erecta* plants were transformed with the *35S::co-3* construct by floral dipping (Bechtold *et al.*, 1993; Clough and Bent, 1998). The *Agrobacterium* strain used was C58C1 pGV101 pMP90. Kanamycin-resistant transformants (T1 generation) were selected on 1/2 × Murashige and Skooge (MS) agar. Flowering time was measured in the T3 generation using lines homozygous for the T-DNA from several independent transformants.

Bombardment of onion bulb epidermal cells

From the inner layer of onion bulb, a peel of epidermis was taken and placed inside up on top of a 50 μl drop of liquid MS on a plate containing solid MS medium (Varagona *et al.*, 1992). The medium contained per litre, 4.3 g MS, 1 mg thiamine, 10 mg myo-inositol, 180 mg KH2PO4 and 30 g sucrose, the pH was adjusted to 5.7 with KOH. After autoclaving, 2.5 mg of amphotericin (in DMSO) was added to the medium. The onion epidermal layers were prepared just before bombardment. 20 μg of plasmid GFP(S65T):CO was precipitated onto gold particles and bombardment performed as described by McCabe and Christou (1993). After bombardment, onion cell layers were incubated at 20 °C for 5 h in complete darkness. To visualize the distribution of cellular DNA the onion peels were immersed in a solution of 0.1% (v/v) DAPI (Sigma-Aldrich, Dorset, UK) for 5 min Subsequently, they were mounted in water and examined by epifluorescence microscopy (Nikon E-800, Nikon, Melville, NY, USA).

Plasmid construction

The GFP-vector pAVA121 was provided by Dr A.G. von Arnim (von Arnim *et al.*, 1998). This plasmid is based on the expression cassette of pRTL2 (Restrepo *et al.*, 1990) that contains a double 35S promoter from CaMV, the translational leader sequence from tobacco etch virus (TEV), and the 35S polyadenylation signal from CaMV. The *GFP* cDNA is a modified version of *mGFP4* (Haseloff *et al.*, 1997) (GFP(S65T), in which the serine 65 residue is substituted by a threonine, resulting in increased absorbance of blue light and reduced absorbance of UV light (Heim *et al.*, 1995). The *CO* cDNA was inserted in frame in the *Bgl*II restriction site of the C-terminus of GFP(S65T). The region corresponding to the CO C-terminus (Met304-Phe373) was amplified by PCR using the primers 5′TERCO (5′-CAA CT**C GGA TCC** ATG GAG AGA GAA GCC-3′) and 3′TERCO (5′- AAT C**AG ATC T**TT CTT TTT GCC ACA GGA G-3′). The 5′TERCO primer introduces a *Bam*HI site before the first codon of the sequence (methionine 304) and the 3′TERCO primer introduces a *Bgl*II restriction site after the CO coding sequence. The PCR fragment was digested with *Bam*HI and *Bgl*II and cloned into the *Bgl*II restriction site of the vector pAVA121 resulting in an in-frame translational fusion at the C-terminus of GFP. This fusion was called GFP(S65T):Cterm. The region corresponding to the first 303 amino acid residues was amplified by PCR using primers OLIGO2 (5′-TGA **GGA TCC** ATG TTG AAA CAA GAG AGT A-3′) and C′-STOP (5′-CT G**AG ATC T**CA ACT GAG TTG TGT TAC T G-3′). Oligo2 maintains the *Bam*HI restriction site before the start codon of the *CO* gene and the reverse primer transforms the proline 302 codon (CCA) into a stop codon (TGA) as well as introducing a *Bgl*II restriction site after the stop codon. The amplified fragment was inserted at the 3′ end of the GFP gene, as described previously for the CO C-terminus. This fusion was called (GFP(S65T):Nterm. The primers 5′TERCO and 3′TERCO were also used to amplify the DNA encoding the C-terminal region (Met304-Phe373) of CO from the mutants *co-5* and *co-7*. The PCR fragments were digested with *Bam*HI and BglII and cloned into the *Bgl*II restriction site of the vector pAVA121 (creating fusion proteins GFP(S65T):*Co-5* and GFP(S65T):*Co-7*). All the PCR fragments were sequenced to check for PCR errors.

Acknowledgements

This work was funded by grants from the BBSRC and EC to G.C. S.R.H. was supported by an EMBO long-term fellowship, M.M.R.C. by a Ph.D. studentship (PRAXIS XXI/BD/3781/94) from Fundação para a Ciência e a Tecnologia, Portugal, and P.H.R. by the EC through the REGIA project.

References

Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, **408**, 796–815.

von Arnim, A.G., Deng, X.-W. and Stacey, M.G. (1998) Cloning vectors for the expression of green fluorescent protein fusion proteins in transgenic plants. *Gene*, **221**, 35–43.

Bechtold, N., Ellis, J. and Pelletier, G. (1993) *In Planta Agrobacterium* mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. *C.R. Acad. Sci. Paris, Sciences de la vie/Life Sci.* **316**, 1194–1199.

Bellini, M., La Croix, J.-C. and Gall, J.G. (1993) A putative zinc binding protein on lampbrush chromosome loops. *EMBO J.* **12**, 107–114.

Borden, K.L.B. (1998) RING fingers and B-boxes: zinc-binding protein–protein interaction domains. *Biochem. Cell Biol.* **76**, 351–358.

Borden, K.L., Lally, J.M., Martin, S.R., O'Reilly, N.J., Etkin, L.D. and Freemont, P.S. (1995) Novel topology of a zinc-binding domain from a protein involved in regulating early Xenopus development. *EMBO J.* **14**, 5947–5956.

Cainarca, S., Messali, S., Ballabio, A. and Meroni, G. (1999) Functional characterization of the Opitz syndrome gene product (midin): evidence for homodimerization and association with microtubules throughout the cell cycle. *Human Mol. Genet.* **8**, 1387–1396.

Chan, E.K.L., Hamel, J.C., Buyon, J.P. and Tan, E.M. (1991) Molecular definition and sequence motifs of the 52kD component of human ss-A/Ro auto-antigen. *J. Clin. Invest.* **87**, 68–76.

Clough, S.J. and Bent, A.F. (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735–743.

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

Deng, X.-W., Matsui, M., Wei, N., Wagner, D., Chu, A.M., Feldmann, K.A. and Quail, P.H. (1992) COP1, an Arabidopsis regulatory gene, encodes a protein with both a zinc-binding motif and a G-beta homologous domain. *Cell*, **71**, 791–801.

El-Husseini, A.E.D., Kwasnicka, D., Yamada, T., Hirohashi, S. and Vincent, S.R. (2000) BERP, a novel ring finger protein, binds to alpha-actinin-4. *Biochem. Biophys. Res. Comm.* **267**, 906–911.

Goddard, A.D., Borrow, J., Freemont, P.S. and Solomon, E. (1991) Characterization of a novel zinc finger gene disruption by the t (15; 17) in acute promyelocytic leukemia. *Science*, **254**, 1371–1374.

Grebenok, R.J., Pierson, E., Lambert, G.M., Gong, F.C., Afonso, C.L., HaldemanCahill, R., Carrington, J.C. and Galbraith, D.W. (1997) Green-fluorescent protein fusions for efficient characterization of nuclear targeting. *Plant J.* **11**, 573–586.

Guerineau, F., Brooks, L. and Mullineaux, P. (1992) Effect of deletions in the cauliflower mosaic virus polyadenylation sequence on the choice of the polyadenylation sites in tobacco protoplasts. *Mol. Gen. Genet.* **226**, 141–144.

Haseloff, J., Siemering, K.R., Prasher, D.C. and Hodge, S. (1997) Removal of a cryptic intron and subcellular localization of green fluorescent protein are required to mark transgenic Arabidopsis plants brightly. *Proc. Natl Acad. Sci. USA*, **94**, 2122–2127.

Heim, R., Prasher, D.C. and Tsien, R.Y. (1994) Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proc. Natl Acad. Sci. USA*, **91**, 12501–12504.

Hope, I.A. and Struhl, K. (1986) Functional dissection of a eukaryotic transcritional activator protein, GCN4 of yeast. *Cell*, **46**, 885–894.

Inoue, S., Orimo, A., Hosoi, T., Kondo, S., Toyoshima, H., Kondo, T., Ikegami, A., Ouchi, Y., Orimo, H. and Muramatsu, M. (1993) Genomic binding site cloning reveals an eastrogen-responsive gene that encodes a RING finger protein. *Proc. Natl Acad. Sci. USA*, **90**, 11117–11121.

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

Kardailsky, I., Shukla, V.K., Ahn, J.H., Dagenais, N., Christensen, S.K., Nguyen, J.T., Chory, J., Harrison, M.J. and Weigel, D. (1999) Activation tagging of the floral inducer FT. *Science*, **286**, 1962–1965.

Keegan, L., Gill, G. and Ptashne, M. (1986) Separation of DNA binding from the transcrition-activating function of a eukaryotic regulatory protein. *Science*, **231**, 699–704.

Kobayashi, Y., Kaya, H., Goto, K., Iwabuchi, M. and Araki, T. (1999) A pair of related genes with antagonistic roles in mediating flowering signals. *Science*, **286**, 1960–1962.

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. General Genet.* **229**, 57–66.

Kurup, S., Jones, H.D. and Holdsworth, M.J. (2000) Interactions of the developmental regulator ABI3 with proteins identified from developing Arabidopsis seeds. *Plant J.* **21**, 143–155.

Lagercrantz, U. and Axelsson, T. (2000) Rapid evolution of the family of CONSTANS LIKE genes in plants. *Mol Biol. Evol.* **17**, 1499–1507.

Ledger, S., Dare, A. and Putterill, J. (1996) COL2 (Accession, L81119 and L81120) is a homologue of the Arabidopsis flowering time gene CONSTANS (PGR 96–081). *Plant Physiol.* **112**, 862.

Ledger, S., Strayer, C., Ashton, F., Kay, S.A. and Putterill, J. (2001) Analysis of the function of two circadian-regulated CONSTANS-LIKE genes. *Plant J.* **26**, 15–22.

Leonhardt, E.A., Kapp, L.N., Young, B.R. and Murnane, J.P. (1994) Nucleotide-sequence analysis of a candidate gene for Ataxia-Telangiectasia Group-D (ATDC). *Genomics*, **19**, 130–136.

Lippuner, V., Cyert, M.S. and Gasser, C.S. (1996) Two classes of plant cDNA clones differentially complement yeast calcineurin mutants and increase salt tolerance of wild-type yeast. *J. Biol. Chem.* **271**, 12859–12866.

Makino, S., Kiba, T., Imamura, A., Hanaki, N., Nakamura, A., Suzuki, T., Taniguchi, M., Ueguchi, C., Sugiyama, T. and Mizuno, T. (2000) Genes encoding pseudo-response regulators: Insight into His-to-Asp phosphorelay and circadian rhythm in *Arabidopsis thaliana*. *Plant Cell Physiol.* **41**, 791–803.

McCabe, D. and Christou, P. (1993) Direct DNA transfer using electric discharge particle acceleration (ACCELLTM technology). *Plant Cell Tissue Organ Cult.* **33**, 227–236.

Miki, T., Fleming, T.P., Crescenzi, M., Molloy, C., Blam, S., Reynolds, S. and Aaronson, S. (1991) Development of a highly efficient cDNA cloning system: application to oncogene isolation. *Proc. Natl Acad. Sci. USA*, **88**, 5167–5171.

Miller, M., Kloc, M., Reddy, B.A., Eastman, E., Dreyer, C. and Etkin, L.D. (1989) x1gv7: a maternal geneproduct localised in nuclei of the central nervous system in *Xenopus laevis*. *Genes Dev.* **3**, 572–583.

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

Onouchi, H., Igeno, M.I., Perilleux, C., Graves, K. and Coupland, G. (2000) Mutagenesis of plants over-expressing CONSTANS demonstrates novel interactions among Arabidopsis flowering-time genes. *Plant Cell*, **12**, 885–900.

Patarca, R., Schwartz, J., Singh, R.P., Kong, Q.T., Murphy, E., Anderson, Y., Sheng, F.Y.W., Singh, P., Johnson, K.A., Guathagia, S.M., Durfee, T., Blattner, F. and Cantor, H. (1988) rpt-1, an intracellular protein from helper/inducer T cells that regulates gene expression of interleukin-2 receptor and human immunodeficiency virus type 1. *Proc. Natl Acad. Sci. USA*, **85**, 2733–2737.

Peng, H.Z., Begg, G.E., Schultz, D.C., Friedman, J.R., Jenson, D.E., Speicher, D.W. and Rauscher, F.J. (2000) Reconstitution of the KRAB-KAP-1 repressor complex: a model system for defining the molecular anatomy of RING-B box-coiled-coil domain-mediated protein–protein interactions. *J. Mol. Biol.* **295**, 1139–1162.

Potuschak, T., Stary, S., Schloegelhofer, P., Beker, F., Nejinskaia, V. and Bachmair, A. (1998) PRT1 of *Arabidopsis thaliana* encodes a component of the plant N-end rule pathway. *Proc. Natl Acad. Sci. USA*, **95**, 7904–7908.

Putterill, J.J., Ledger, S.E., Lee, K., Robson, F., Murphy, G. and Coupland, G. (1997) The flowering time gene CONSTANS and homologue CONSTANS LIKE 1 exist as a tandem repeat on chromosome 5 of Arabidopsis. *Plant Physiol.* **114**, 396.

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.

Raikhel, N. (1992) Nuclear targeting in plants. *Plant Physiol.* **100**, 1627–1632.

Reddy, B.A., Etkin, L.D. and Freemont, P.S. (1992) A novel zinc finger coiled-coil domain in a family of nuclear proteins. *Trends Biochem. Sci.* **17**, 344–345.

Redei, G.P. (1962) Supervital mutants of Arabidopsis. *Genetics*, **47**, 443–460.

Restrepro, M.A., Freed, D.D. and Carrington, J.C. (1990) Nuclear transport of plant potyviral proteins. *Plant Cell,* **2**, 987–998.

Robert, L.S., Robson, F., Sharpe, A., Lydiate, D. and Coupland, G. (1998) Conserved structure and function of the Arabidopsis flowering time gene CONSTANS in *Brassica napus*. *Plant Mol Biol.* **37**, 763–772.

Robson, F. (1998) Characterization of CONSTANS an Arabidopsis gene that promotes flowering. PhD Thesis, Milton, Keynes, UK: The Open University/John Innes Centre.

Samach, A., Onouchi, H., Gold, S.E., Ditta, G.S., Schwarz-Sommer, Zs, Yanofsky, M.F. and Coupland, G. (2000) Distinct roles of CONSTANS target genes in reproductive development of Arabidopsis. *Science*, **288**, 1613–1616.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Schaffer, R., Ramsay, N., Samach, A., Corden, S., Putterill, J., Carre, I. and Coupland, G. (1998) The late elongated hypocotyl mutation in Arabidopisis disrupts circadian rhythms and the photoperiodic control of flowering. *Cell*, **93**, 1219–1229.

Schultz, J., Copley, R.R., Doerks, T., Ponting, C.P. and Bork, P. (2000). SMART: a web-based tool for the study of genetically mobile domains. *Nucl. Acids Res.* **28**, 231–234.

Schultz, J., Milpetz, F., Bork, P. and Ponting, C.P. (1998) SMART, a simple modular architecure research tool: identification of signalling domains. *Proc. Natl Acad. Sci. USA*, **95**, 5857–5864.

Simon, R., Igeno, I.M. and Coupland, G. (1996) Activation of floral meristem identity genes in Arabidopsis. *Nature*, **384**, 59–62.

Soppe, W.J.J., Jacobsen, S.E., Alonso-Blanco, C., Jackson, J.P., Kakutani, T., Koornneef, M. and Peeters, A.J.M. (2000) The late flowering phenotype of fwa mutants is caused by gain-of-function epigenetic alleles of a homeodomain gene. *Mol. Cell*, **6**, 791–802.

Strayer, C., Oyama, T., Schultz, T.F., Raman, R., Somers, D.E., Mas, P., Panda, S., Kreps, J.A. and Kay, S.A. (2000) Cloning of the Arabidopsis clock gene TOC1, an autoregulatory response regulator homolog. *Science*, **289**, 768–771.

Takahashi, M., Inaguma, Y., Hiai, H. and Hirose, F. (1988) Developmentally regulated expression of a human 'Finger'-containing gene encoded by the 5′ half of the ret transforming gene. *Mol. Cell. Biol.* **8**, 1853–1856.

Tokunaga, C., Kuroda, S., Tatematsu, K., Nakagawa, N., Ono, Y. and Kikkawa, U. (1998) Molecular cloning and characterization of a novel protein kinase C-interacting protein with structural motifs related to RBCC family proteins. *Biochem. Biophys. Res. Commun.* **244**, 353–359.

Tsuzuki, S., Towatari, M., Saito, H. and Enver, T. (2000) Potentiation of GATA–2 activity through interactions with the Promyelocytic Leukemia Protein (PML) and the t (15; 17) - generated PML-Retinoic Acid receptor ± oncoprotein. *Mol. Cell. Biol.* **20**, 6276–6286.

Unger, E., Parsons, R.L., Schmidt, R.J., Bowen, B. and Roth, B.A. (1993) Dominant-negative mutants of opaque2 suppress transactivation of a 22-kd zein promoter by opaque2 in maize endosperm cells. *Plant Cell*, **5**, 831–841.

Varagona, M.J., Schmidt, R.J. and Raikhel, N.V. (1992) Nuclear localization signal (s) required for nuclear targeting of the maize regulatory protein Opaque-2. *Plant Cell*, **4**, 1213–1227.

Vizir, I.Y., Anderson, M.L., Wilson, Z.A. and Mulligan, B.J. (1994) Isolation of deficiencies in the arabidopsis genome by gamma-irradiation of pollen. *Genetics*, **137**, 1111–1119.

Wagner, D., Sablowski, R.W.M. and Meyerowitz, E.M. (1999) Transcriptional activation of APETALA1 by LEAFY. *Science*, **285**, 582–584.

Zou, J.T. and Taylor, D.C. (1997) Cloning and molecular characterization of an *Arabidopsis thaliana* RING zinc finger expressed preferentially during seed development. *Gene*, **196**, 291–295.