Interaction of Polycomb-group proteins controlling flowering in *Arabidopsis*

Yindee Chanvivattana[1,*,‡], Anthony Bishopp[1,‡], Daniel Schubert[1], Christine Stock[1], Yong-Hwan Moon[2,†], Z. Renee Sung[2] and Justin Goodrich[1,§]

[1]Institute of Molecular Plant Science, School of Biology, University of Edinburgh, Mayfield Road, Edinburgh EH9 3JH, UK
[2]Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720, USA

*Present address: National Centre for Genetic Engineering and Biotechnology Thailand Science Park 113, Phahonyothin Rd., Klong 1, Klong Luang, Pathum Thani 12120 Thailand
†Present address: Department of Molecular Biology, Pusan National University, Keumjung-ku, Busan 609-735, Korea
‡These authors contributed equally to this work
§Author for correspondence (e-mail: justin.goodrich@ed.ac.uk)

Accepted 16 August 2004

Development 131, 5263-5276
Published by The Company of Biologists 2004
doi:10.1242/dev.01400

Summary

In *Arabidopsis*, the *EMBYRONIC FLOWER2* (*EMF2*), *VERNALISATION2* (*VRN2*) and *FERTILISATION INDEPENDENT ENDOSPERM2* (*FIS2*) genes encode related Polycomb-group (Pc-G) proteins. Their homologues in animals act together with other Pc-G proteins as part of a multimeric complex, Polycomb Repressive Complex 2 (PRC2), which functions as a histone methyltransferase. Despite similarities between the *fis2* mutant phenotype and those of some other plant Pc-G members, it has remained unclear how the *FIS2/EMF2/VRN2* class Pc-G genes interact with the others. We have identified a weak *emf2* allele that reveals a novel phenotype with striking similarity to that of severe mutations in another Pc-G gene, *CURLY LEAF* (*CLF*), suggesting that the two genes may act in a common pathway. Consistent with this, we demonstrate that *EMF2* and *CLF* interact genetically and that this reflects interaction of their protein products through two conserved motifs, the VEFS domain and the C5 domain. We show that the full function of *CLF* is masked by partial redundancy with a closely related gene, *SWINGER* (*SWN*), so that null *clf* mutants have a much less severe phenotype than *emf2* mutants. Analysis in yeast further indicates a potential for the CLF and SWN proteins to interact with the other VEFS domain proteins VRN2 and FIS2. The functions of individual Pc-G members may therefore be broader than single mutant phenotypes reveal. We suggest that plants have Pc-G protein complexes similar to the Polycomb Repressive Complex2 (PRC2) of animals, but the duplication and subsequent diversification of components has given rise to different complexes with partially discrete functions.

Key words: Polycomb, Flowering, VEFS domain

Introduction

A general feature of developmental patterning is that it occurs progressively so that patterns are gradually refined based on information from earlier, cruder patterns (Coen, 1999; Stern, 1968). This presents a mechanistic problem in growing embryos or organ primordia, because most patterning events are thought to involve gradients of morphogens that operate across small numbers of cells and cannot persist over larger distances (Lawrence and Struhl, 1996). A typical resolution to this problem is for the transcriptional output of early patterning events to become fixed, so that information from early events is inherited through cell division during somatic development. This often involves epigenetic changes in gene activity, i.e. changes that are heritable through mitotic (and sometimes meiotic) division but are not caused by alterations in DNA sequence (Russo et al., 1996). The advantage of epigenetic changes in a development context is that although stable they are also reversible, particularly during meiosis, so that changes that accrue during somatic development can be erased at the onset of each new generation.

In *Drosophila* and other animals, the epigenetic control of developmental patterning is mediated by members of the Polycomb group (Pc-G) and trithorax group (trx-G) of genes (for a review, see Francis and Kingston, 2001). A general feature of these genes is that they are required not for pattern initiation, but rather to ensure that the transcriptional output of early patterning events is stably inherited through somatic development. The two groups act antagonistically, so that the Pc-G genes are required for maintenance of transcriptional repression and the trx-G genes for maintenance of transcriptional activation. Recently, biochemical characterisation of their protein products has provided some mechanistic insight. The Pc-G products, which are structurally disparate from one another, are found in at least two distinct complexes, termed the Polycomb Repressive Complex 1 and 2 (PRC1 and PRC2) (Cao et al., 2002; Czermin et al., 2002; Francis et al., 2001; Kuzmichev et al., 2002; Muller et al., 2002; Saurin et al., 2001). Consistent with the epigenetic function of Pc-G proteins, the PRC2 was recently shown to modify chromatin. Thus, several groups have shown that the PRC2 has a histone methyltransferase (HMTase) activity, methylating specific residues (lysine 9 and lysine 27) on the

N-tail of histone H3 (Cao et al., 2002; Czermin et al., 2002; Kuzmichev et al., 2002; Muller et al., 2002). The precise biochemical function of the different PRC2 members is not well defined, with the exception of one member, the Enhancer of zeste [E(z)] protein, which has been shown to confer HMTase activity via a conserved motif, the SET domain. Unlike most other SET domain proteins, E(z) itself does not show HMTase activity in vitro unless associated with other members of the PRC2 complex (Czermin et al., 2002). The mK9 H3 and mK27 H3 modifications catalysed by the PRC2 are associated with the repression of transcription, although how they are interpreted and inherited through mitosis is not well understood. The Polycomb protein appears to recognise and bind mK27 H3, and may recruit other members of the PRC1 complex, which probably have roles in mediating transcriptional silencing and its propagation through mitosis (Cao et al., 2002; Czermin et al., 2002; Francis et al., 2001).

The PRC2 components are also found in plants, and were identified independently through genetic screens in *Arabidopsis* aimed at dissecting various developmental pathways. Thus, the *FERTILISATION INDEPENDENT SEED* (*FIS*) genes were mostly identified through screens for mutants that showed some aspects of seed development in the absence of fertilisation (Chaudhury et al., 1997; Grossniklaus et al., 1998; Guitton et al., 2004; Ohad et al., 1996). Currently four *FIS* genes have been identified: *MEDEA* (*MEA*), *FERTILISATION-INDEPENDENT SEED2* (*FIS2*), *FERTILISATION-INDEPENDENT ENDOSPERM* (*FIE*) and *MULTICOPY SUPPRESSOR OF IRA 1 (MSI1)*. These encode products with homology to the *Drosophila* PRC2 proteins E(z), Suppressor of zeste 12 [Su(z)12], Extra sex combs (Esc) and P55, respectively (Grossniklaus et al., 1998; Kiyosue et al., 1999; Kohler et al., 2003a; Luo et al., 1999; Ohad et al., 1999). The *FIS* genes repress expression of the MADS box gene *PHERES1* (*PHE1*) during early seed development, and presumably affect many other as yet unidentified target genes (Kohler et al., 2003b). A second group have been identified based on a common function in repressing floral homeotic gene expression. Mutants in this class are early flowering and exhibit mild homeotic transformations in flowers. The first two members identified were *CURLY LEAF* (*CLF*) and *EMBRYONIC FLOWER2* (*EMF2*), which encode proteins with homology to E(z) and Su(z)12, respectively (Goodrich et al., 1997; Yoshida et al., 2001). Recently, the *FIS* genes *MSI1* and *FIE* have also been implicated in repressing flowering homeotic genes during vegetative development. Because mutant alleles of the *FIS* genes all cause early embryo lethality when inherited maternally, this obstructed the phenotypic analysis of *fis* homozygotes during later developmental stages. However, studies of transgenic lines that confer a partial loss of *FIS* gene activity have revealed roles for *FIE* and *MSI1* beyond seed development (Hennig et al., 2003; Katz et al., 2004; Kinoshita et al., 2001). A third class of *Arabidopsis* Pc-G genes was identified on the basis of the function of the genes in the epigenetic memory of vernalisation. In *Arabidopsis*, as with many other plant species originating from temperate latitudes, flowering is accelerated if plants are first vernalised by growing for 3-6 weeks at low temperatures (4-10°C). The vernalisation response has several epigenetic features, including stability during somatic development and resetting from generation to generation. Recent studies indicate that the underlying basis for the response involves transcriptional repression of *FLC*, a gene that itself represses flowering (Michaels and Amasino, 1999; Sheldon et al., 1999). The *VERNALIZATION2* (*VRN2*) gene is required so that the cold-induced repression of *FLC* is mitotically stable during later periods of growth at warm temperatures (Gendall et al., 2001). *VRN2*, like *FIS2* and *EMF2*, encodes a protein with homology to *Drosophila* Su(z)12 (Gendall et al., 2001). The completion of the *Arabidopsis* genome sequence revealed that these comprised most of the *Arabidopsis* homologues of the core members of the PRC2. One exception was a third *E(z)* homologue, GenBank accession At4g02020, that had not been characterised genetically. In addition, there are four genes with weak similarity to *MSI1*, *MSI2-5*, with poorly defined functions (Ach et al., 1997; Hennig et al., 2003; Kenzior and Folk, 1998). Thus, unlike *Drosophila*, in which the PRC2 members are single copy genes, in *Arabidopsis* the different members are mostly small gene families. The duplicated members appear to have acquired distinct functions; thus *CLF* and *MEA* function in repressing flowering and repressing endosperm proliferation, respectively. It is not clear how this has occurred; for example, whether it simply reflects different expression patterns of *MEA* and *CLF*, or whether their protein products have also diverged in function.

In addition to the conservation of PRC2 members in plants, there is also evidence that their proteins may act together. Thus, several studies have shown that the FIE protein can interact with the E(z) homologues MEA, CLF and At4g02020 (Katz et al., 2004; Kohler et al., 2003a; Luo et al., 2000; Spillane et al., 2000; Yadegari et al., 2000). Also, compelling evidence for interaction of FIE and MSI1 proteins was recently presented (Kohler et al., 2003a). However, the role of the plant *Su(z)12* homologues has remained obscure. Despite the similarities in the phenotype of *fis2* and the other *fis* mutants, no interaction between FIS2 (or any other Su(z)12 homologue) with the other Pc-G proteins has been found.

Here, we show that the plant *E(z)* and *Su(z)12* homologues interact both genetically and physically through their protein products. We localise the interactions to motifs that are conserved between the plant and animal proteins. We show that the third *Arabidopsis E(z)* homologue functions largely redundantly with *CLF*, so that *CLF* has a more general role in control of plant development than was apparent from its single mutant phenotype. Characterisation of the misexpression phenotypes for the three *E(z)* homologues indicates that they have diverged not only in expression, but also at the protein level. We suggest that in plants an evolutionarily ancient complex (the PRC2) has been conserved, but gene duplication and divergence has given rise to several complexes with partially discrete functions.

Materials and methods

Plant materials

The *clf-2* and *clf-9* alleles arose in L*er* background and were described previously (Goodrich et al., 1997). The null *clf-50* allele (Ws background) was provided by E. Huala and harbours a deletion spanning the *CLF* locus (J.G., unpublished). The weak *emf2-10* allele arose in Ws background during a T-DNA mutagenesis experiment and was provided by M. Running. The weak *swn-1* allele (Ws background) was identified in seed pool 5887 in the University of Wisconsin *Arabidopsis* knockout collection (Krysan et al., 1999). The

swn-1 line also carried an unlinked, recessive mutation conferring late flowering. The line was backcrossed twice to the Ws progenitor and *swn-1* lines with and without the late flowering mutation were generated. The severe *swn-2* and *swn-3* alleles (Columbia background) were obtained from the SIGNAL collection of T-DNA insertion lines (Alonso et al., 2003) and correspond to accessions SALK 010213 and SALK 050195, respectively. The position of the T-DNA inserts was confirmed by PCR amplification and sequencing of genomic DNA flanking the inserts.

Yeast two-hybrid assay

Constructs for yeast two-hybrid analysis were generated using the vectors pGBT9 and pGAD424 (Clontech) that express protein fusions to the GAL4 DNA-binding domain or transcriptional-activation domain, respectively. cDNA inserts encoding plant Pc-G proteins were introduced as *Eco*RI/*Sal*I fragments. The Quik Change site-directed mutagenesis system (Stratagene) was used to introduce in-frame *Eco*RI and *Sal*I restriction sites within cDNA clones, with the exception of *EMF2* clones, which were generated by PCR amplification using mutagenic primers. PCR-generated clones were validated by sequence analysis. The methods for two-hybrid analysis were as described in the yeast protocols handbook (Clontech). The analysis was performed in yeast strain Hf7c (Feilotter et al., 1994), which carries *HIS*3 and *LacZ* reporters for reconstituted GAL4 activity, or in strain AH109 (James et al., 1996), which carries *HIS*3 and *ADE*2 reporters.

Yeast split-ubiquitin assay

Vectors were used as described in Kim et al. (Kim et al., 2002). CLF-C5 was cloned into pENTRY 3c (Invitrogen) and recombined into the bait vector using the Gateway system (Invitrogen), resulting in a CLF-C5-Cub-URA3 gene fusion. EMF2-VEFS was fused to a gene encoding the N-terminal part of ubiquitin in the vector pCGK. The plasmids were transformed into the *Saccharomyces cerevisiae* strain JD53 and interaction of the fusion proteins was monitored as ability to grow on 5-fluoroorotic acid (5-FOA) plates, containing yeast nitrogen base without amino acids (Difco) and glucose, supplemented with lysine, leucine, uracil, and 1 mg/ml 5-FOA.

In-vitro pull down assay

A similar protocol to that described in Kohler et al. (Kohler et al., 2003a) was applied. The coding region for the CLF C5 domain (amino acids 257-331) was cloned into the pGEX-4T expression vector (Amersham) as a GST-fusion, whereas the EMF2 VEFS domain (amino acids 427-631) was cloned in the pET30a expression vector (Novagen) as a HIS_6-fusion. *Escherichia coli* strain BL21 DE3 Codon-plus (Stratagene) was freshly transformed with the pGEX-CLF-C5, pGEX or pET-EMF2-VEFS plasmids and grown in LB medium at 37°C overnight. After diluting the cultures 1:100 in 250 ml LB, they were grown at 37°C (pGEX-CLF-C5 and pGEX) or 18°C (pET-EMF2-VEFS) until OD_{600}=0.7. Production of recombinant protein was induced by adding isopropyl-β-D-thiogalactopyranoside (IPTG) to 0.2 mM and after growing the cells for 3 hours at 18°C, they were harvested and resuspended in 4 ml binding buffer [BB; 20 mM Tris pH 7.5, 150 mM NaCl, 0.1% Triton X100, 1 μM ZnSO4, 1 mM Pefabloc (Roche)]. The cells were lysed by the addition of Lysozyme to 2 mg/ml and incubated for 20 minutes on ice. The solution was centrifuged (20,000 ***g***) for 10 minutes, the pellets discarded, centrifuged again and a 100 μl sample of supernatant was mixed with SDS sample buffer and frozen in liquid nitrogen (input control sample). Equal volumes of extract containing HIS_6-EMF2-VEFS were mixed with extracts containing GST-CLF-C5 or GST and 150 μl of pre-equilibrated glutathione-sepharose 4B beads (Pharmacia) and incubated with shaking for 2 hours at 4°C. The beads were washed four times with BB and then mixed with SDS sample buffer, analysed on protein blots and the HIS_6-EMF2-VEFS fusion detected with anti-HIS_6-antibodies (New England Biolabs).

Misexpression of *CLF*, *SWN* and *MEA*

Constructs for expression of *CLF*, *SWN* and *MEA* cDNAs under control of the cauliflower mosaic virus 35S promoter were assembled using the pART7 and pART27 vector systems (Gleave, 1992). A full-length SWN cDNA clone (pda05864) was obtained from the Riken Bioresource centre, Japan (Seki et al., 2002), CLF and MEA cDNA clones were isolated previously (Goodrich et al., 1997; Spillane et al., 2000). The Quik Change site-directed mutagenesis system (Stratagene) was used to engineer restriction sites within the cDNA clones that facilitated subcloning the coding sequences into pART7. The constructs were introduced into *Agrobacterium* strain GV3101 pMP90 (Koncz and Schell, 1986) and used to transform *clf-50/+* heterozygotes by floral dip transformation (Clough and Bent, 1998). At least 23 primary transformants were identified for each construct. Selected plants in the T1 and T2 generations were genotyped for presence of a transgene and for *clf-50* and *CLF+* alleles by Southern blot analysis.

In-situ hybridisation

The methods for in-situ hybridisation analysis using digoxigenin-labelled mRNA probes were described previously (Narita et al., 2004). *SWN* probes were generated from a poorly conserved 700 bp region at the 5′ end of the *SWN* coding region. *WUS* probes were generated using the clone pMH WUS 16, generously provided by R. Simon.

Scanning electron microscopy and cell size measurements

Scanning electron microscopy (SEM) was performed on a Hitachi 4700 with a Gatan Alto cryo-stage. The methods for cryo-SEM were as described previously (Jeffree and Read, 1991). For measuring cell sizes, fully expanded rosette leaves were fractured in transverse section and photographed using the cryo-SEM. The cell outlines were traced onto transparencies, scanned, and quantified using image analysis software (image tool, University of Texas, available at http://ddsdx.uthscsa.edu/dig).

Results

Similarity between *emf2-10* and *curly leaf* mutants

To identify genes acting like *CLF* to repress floral homeotic gene expression during vegetative development, we screened existing mutant collections for plants with a similar leaf curling and early flowering phenotype. We identified a single recessive mutation, designated *moe leaf*, which conferred a phenotype resembling, but more severe than, *clf* mutations. Subsequent genetic analysis (see next section) revealed that it was an unusual, weak *emf2* allele. We hereafter refer to it as the *emf2-10* allele but retain the name ‘moe leaf’ to describe the phenotype. Both *emf2-10* and *clf* mutants flowered early under both long days and short days (Table 1). However, *emf2-10* plants were significantly earlier flowering than *clf-50* plants, which carried a null *clf* allele isolated in the same genetic background as *emf2-10*. Both mutants gave small, dwarfed plants that had short slender inflorescence stems and narrow leaves that curled upwards along the leaf margin (Fig. 1A,B). *emf2-10* plants were smaller than *clf-50* plants (Fig. 1A) and, unlike *clf*, also had cotyledons that were smaller than in wild-type plants (Fig. 1C). Comparison of *emf2-10* and wild-type leaf epidermal surfaces by SEM showed that both had pavements of large cells with irregular outline (Fig. 2A,B). Unlike wild-type leaves, which had flat surfaces, the ventral (abaxial, lower) epidermis of *emf2-10* leaves was uneven and corrugated (Fig. 2B). One possibility, which is also consistent with the upward curling of the leaves, is that growth of the

Table 1. Effect of *emf2-10* and *clf-50* mutations on rosette leaf number

Genotype	Short days	Long days
Wild type (Ws)	16.6±0.3	8.0±0.1
clf-50	10.9±0.3	7.2±0.1
emf2-10	7.7±0.2	5.9±0.1

The average rosette leaf number at flowering is shown, together with the s.e.m. (n=20). Rosette leaf number is inversely correlated with flowering time, so that early flowering lines have fewer leaves. The wild type included for comparison is Ws ecotype, the progenitor background in which both the *clf-50* and *emf2-10* alleles were isolated.

ventral leaf surface was constrained by the dorsal surface during leaf development, leading to the observed corrugation. When leaves were frozen and fractured, so that they could be viewed in transverse section by SEM, *emf2-10* leaves had a similar arrangement of cell types as in wild type (Fig. 2C,D), but the cells were smaller than wild type and there were also many fewer cells in the leaf length and leaf width axes (data not shown). The *emf2-10* mutation therefore affects cell proliferation as well as cell size. Similarly, morphometric analysis has shown that *clf* mutant leaves also show reductions in both cell number and cell size (Kim et al., 1998).

The inflorescences of *emf2-10* plants produced few flowers. In wild type, the primary inflorescence produced 25-30 flowers before arresting development. In *emf2-10*, after 6-12 flowers had opened, the remaining flower buds arrested development so that the inflorescence subsequently appeared determinate. The flowers of *emf2-10* plants mostly had normal organ identity but were smaller than wild type, and petals and sepals were narrower than in wild type (Fig. 1D). The flower buds often opened later than normal, after fertilisation had occurred, so that the elongation of the developing silique (fruit) was constrained and the siliques became bent or folded over (Fig. 1E, Fig. 2E). This suggested that the sepals were impeding bud opening. Wild-type sepals have a hyaline margin, distinguishable under SEM as a region of regularly sized cells that lack the extremely elongated cells found elsewhere on the sepal (Fig. 2F). In *emf2-10* flowers the margin was less well

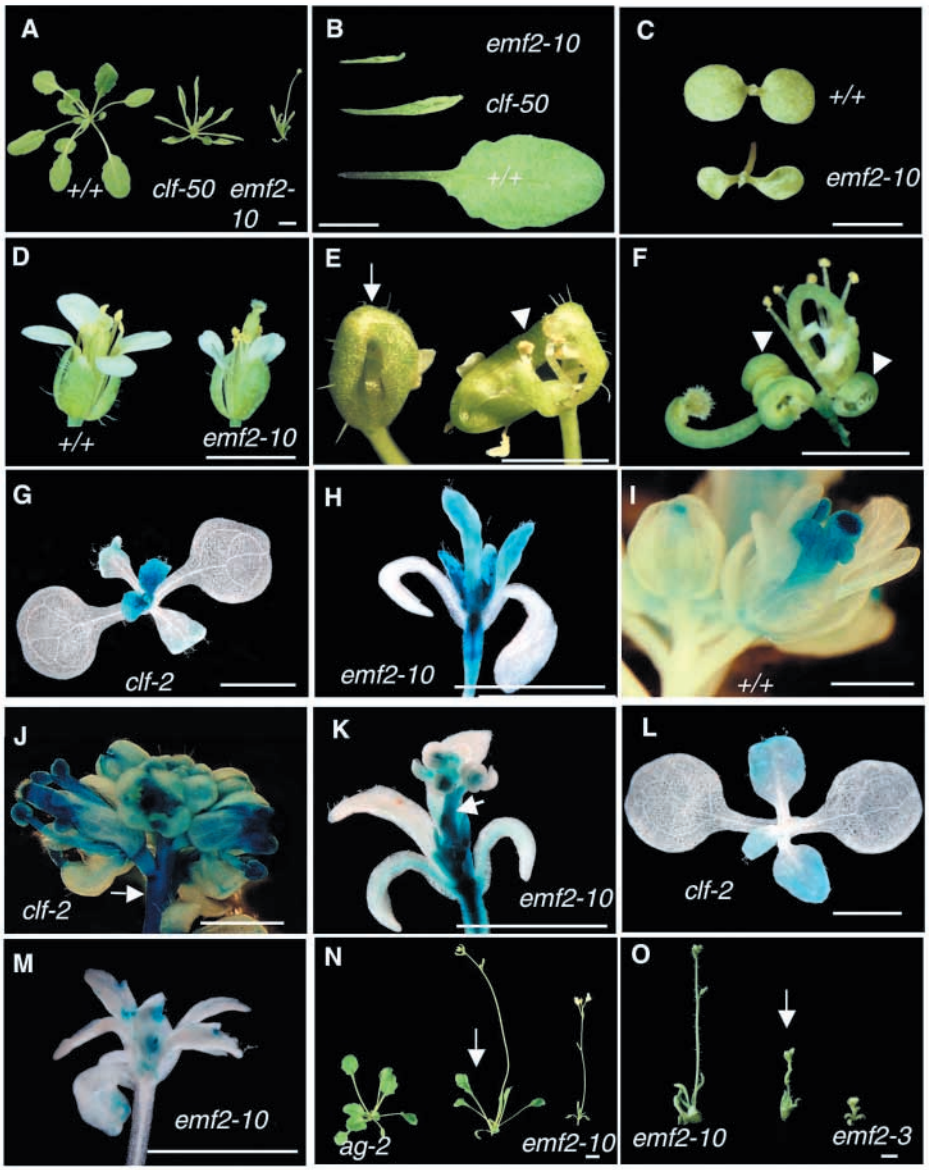

Fig. 1. The moe leaf phenotype. (A) Short-day grown plants of *clf-50*, *emf2-10* and wild-type progenitor (Ws ecotype) after 30 days. The *emf2-10* plant is slightly earlier flowering and smaller than clf-50. (B) Fifth rosette leaf of long-day grown plants. The *emf2-10* leaf is smaller and narrower than *clf-50*, but has similar upward curling of leaf margin. (C) Seedlings showing effects of *emf2-10* on cotyledon size. (D) Wild-type and *emf2-10* flowers at anthesis. (E) *emf2-10* flowers showing delayed flower opening (left) resulting in contorted siliques (right). (F) *emf2-10* flowers from apex of inflorescence. Arrowheads indicate carpelloid sepals; petals are also stamenoid in shape and have yellow anther-like sectors. Flowers appear terminal because flower buds from shoot apex (between two flowers) aborted early in development and are no longer visible. (G-M) Transgenic plants carrying *AG* or *AP3* reporter genes stained for GUS activity (blue colour). (G) *pAG-I::GUS* activity in *clf-2* seedlings. Expression was also seen in cotyledons at earlier stages. (H) *pAG-I::GUS* in *emf2-10* seedling. (I) *pAG-I::GUS* in wild-type inflorescence. (J) *pAG-I::GUS* in *clf-2* inflorescence. Arrow indicates expression in the stem. (K) *pAG-I::GUS* in *emf2-10* plant. Arrow indicates expression in inflorescence stem. (L) *pAP3::GUS* in *clf2* seedling. (M) *pAP3::GUS* in *emf2-10* seedling. (N) The *emf2-10 ag-2* double mutant (arrow) is shown between *ag-2* and *emf2-10* parent lines. Note that the double mutant is earlier flowering than *ag-2* and smaller. (O) The *emf2-10/emf2-3* heterozygote (arrow) is shown between *emf2-10* and *emf2-3* parents. Its phenotype is intermediate, both with respect to height and overall plant size. Scale bar: 5 mm in A,B; 2 mm in C-O.

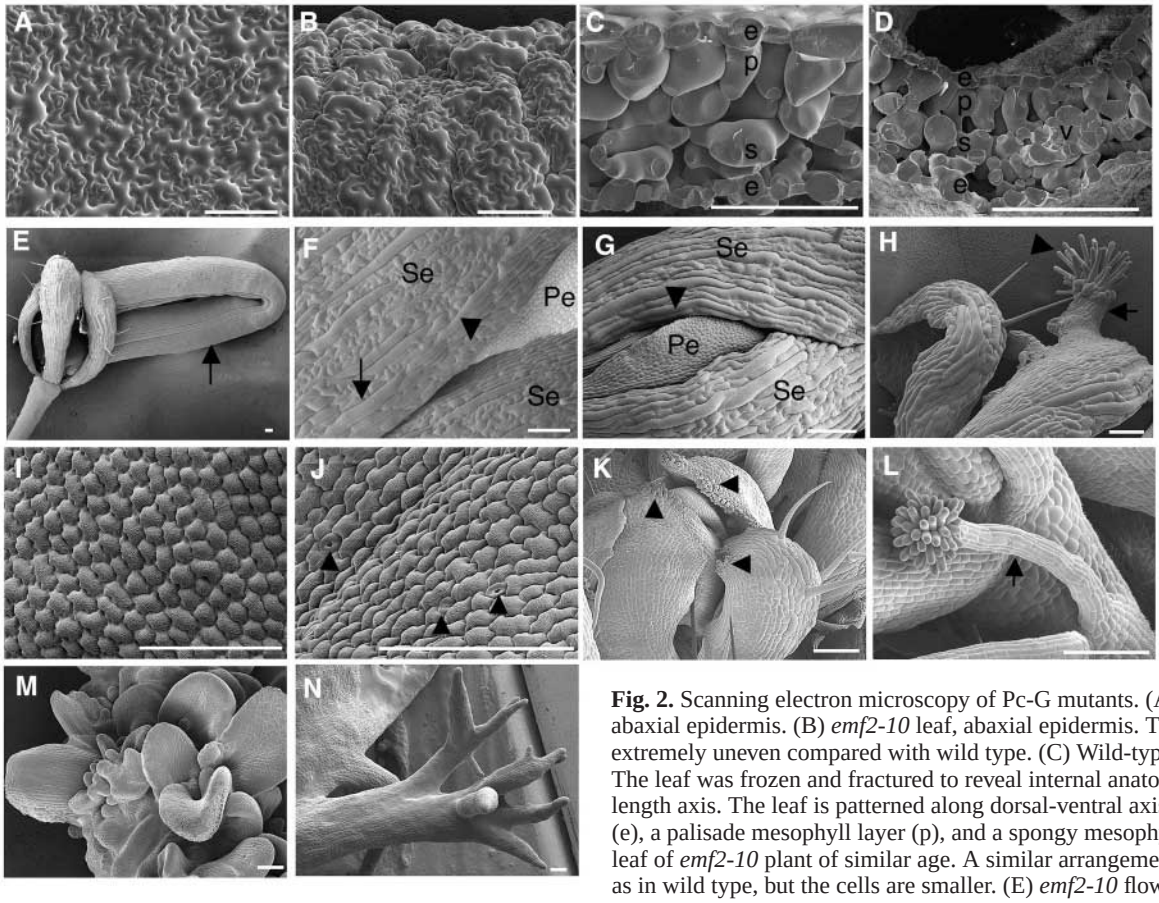

Fig. 2. Scanning electron microscopy of Pc-G mutants. (A) Wild-type leaf, abaxial epidermis. (B) *emf2-10* leaf, abaxial epidermis. The surface is extremely uneven compared with wild type. (C) Wild-type fifth rosette leaf. The leaf was frozen and fractured to reveal internal anatomy transverse to leaf length axis. The leaf is patterned along dorsal-ventral axis into epidermal cells (e), a palisade mesophyll layer (p), and a spongy mesophyll (s) (D) Fifth rosette leaf of *emf2-10* plant of similar age. A similar arrangement of cell types is seen as in wild type, but the cells are smaller. (E) *emf2-10* flower showing contorted silique. (F) Wild-type flower showing abaxial epidermi of sepals (se) and petals (pe). The sepal epidermis contains characteristic highly elongated cells (arrow); however, the margin lacks the elongated cells and has smaller, more regularly sized cells (arrowhead). (G) *emf2-10* sepal, elongated cells extend to the margin. (H) *emf2-10* flower showing carpelloid sepal. The organ has elongated cells typical of sepals but stigmatic papillae (arrowhead) and stylar cells (arrow) characteristic of carpels. (I) Wild-type petal, abaxial surface, note lack of stomates. (J) *emf2-10* petal, abaxial surface. Note presence of stomates (arrow) and cell shape, characteristic of stamen epidermis. (K) *swn-1 clf-50* inflorescence. The sepals show weak homeotic conversion to carpelloid organs. Arrowheads indicate stigmatic papillae. (L) Radialised organ (arrow) with stigmatic papillae arising from inflorescence stem of *swn-1 clf-50* double mutant in position where stipule would normally arise. (M) *swn-3 clf-50* double mutant. Organs arise with disorganised phyllotaxy. Note lack of trichomes, cells lack wall thickening and are isodiametric. (N) *swn-2 clf-50* plant showing radialised organs. Scale bar: 100 μm throughout.

defined, so that the elongated cells often extended to the margin (Fig. 2G). In addition, the sepals were more concave or boat shaped than wild type. Both features may have contributed to restricting bud opening. As with *clf* mutants, *emf2-10* flowers produced late in development showed weak homeotic transformation of sepals to carpels (Fig. 1F, Fig. 2H) and petals to stamens (Fig. 2I,J).

The similarities in phenotypes suggested that the moe leaf phenotype, like that of *clf* mutants, could be caused by misexpression of floral homeotic genes during vegetative and floral development. Previous studies have shown that the *AG* and *AP3* genes, whose expression is normally confined to flowers, are misexpressed in leaves of *clf* mutants (Finnegan et al., 1996; Goodrich et al., 1997; Serrano-Cartagena et al., 2000). We therefore used RT PCR to compare *AG* and *AP3* expression in leaves of wild-type and *emf2-10* plants. This indicated that both genes were expressed in *emf2-10* leaves (data not shown). To confirm this, we introduced reporter constructs for *AG* and *AP3* expression into the *emf2-10* background. The pAG-I::*GUS* construct contains *AG* upstream promoter sequences and intragenic sequences fused to the GUS reporter and has been shown to contain the *cis*-acting sequences necessary for repression by *CLF* (Sieburth and Meyerowitz, 1997). This construct was strongly misexpressed in seedlings of both *emf2-10* and *clf-2* mutants (Fig. 1G,H). In addition, both mutants showed misexpression in inflorescence stems (Fig. 1J,K) and occasional misexpression in the outer floral whorls. We also tested reporter constructs containing the *AG* second intron upstream of a GUS reporter gene (KB9) (Busch et al., 1999). This construct also confers the wild-type *AG* expression pattern in flowers, presumably because the second intron contains many *AG* regulatory elements (Busch et al., 1999). However, when the KB9 construct was introduced into *clf* or *emf2-10* mutant backgrounds, no expression was seen in seedlings (data not shown). This suggested that the *AG* promoter contains additional enhancers that are required for misexpression of *AG* in *clf* and *emf2-10* mutant backgrounds. An *AP3* reporter construct containing 3.7 kb of upstream

regulatory sequences (Jack et al., 1994) also showed weak expression in *emf2-10* and *clf* seedlings but not wild-type (Fig. 1L,M).

Genetic data have indicated that the phenotype of *clf* mutants is chiefly caused by ectopic *AG* expression. Thus in *clf ag* double mutant plants, in which *AG* activity is eliminated, leaf morphology is restored to near wild-type (Goodrich et al., 1997; Serrano-Cartagena et al., 2000). To test whether the moe leaf phenotype was also a result of ectopic *AG* activity, we made *emf2-10 ag-2* double mutants. Although the double mutants had larger, less curled leaves than *emf2-10* single mutants, there was less restitution of wild-type morphology than in the case of *clf ag* double mutants. Thus *emf2-10 ag-2* plants were still much smaller than wild type, their leaves retained some curling, and they flowered earlier (Fig. 1N). This indicated that although AG^+ activity contributes to the moe leaf phenotype, misexpression of other genes is also probably involved. Taken together, these results indicated that *EMF2* and *CLF* shared common functions in repressing floral homeotic gene expression, with *EMF2* required to repress a broader range of targets than *CLF*. These similarities suggested that *EMF2* might act in a common pathway with *CLF*.

The moe leaf phenotype is conferred by a weak *emf2* allele

To determine the molecular basis for the moe leaf phenotype, we employed a map-based cloning strategy and initially localised the mutation responsible to a 10 cm interval between markers nga129 and ATTED2 on the lower arm of chromosome 5. It was striking that the plant Pc-G member *EMF2* had also been located within this interval. All nine *emf2* mutant alleles previously described have much more severe phenotypes than moe leaf, producing minute plants that appear to flower soon after germination without undergoing a prior phase of vegetative development (Sung et al., 1992; Sung et al., 2003; Yang et al., 1995). Instead, a few flowers and sessile cauline leaves are produced on an inflorescence with a severely shortened bolt (Fig. 1O, Fig. 3C). However, several features made *EMF2* a promising candidate. Firstly, it was known to repress floral homeotic gene expression during vegetative development (Chen et al., 1997; Moon et al., 2003). Secondly, transgenic plants that expressed an antisense *EMF2* construct had a phenotype resembling moe leaf, which probably reflected a partial loss of *EMF2* function (Yoshida et al., 2001). To test whether the moe leaf phenotype could be caused by an unusual, weak allele of *EMF2*, we performed genetic complementation

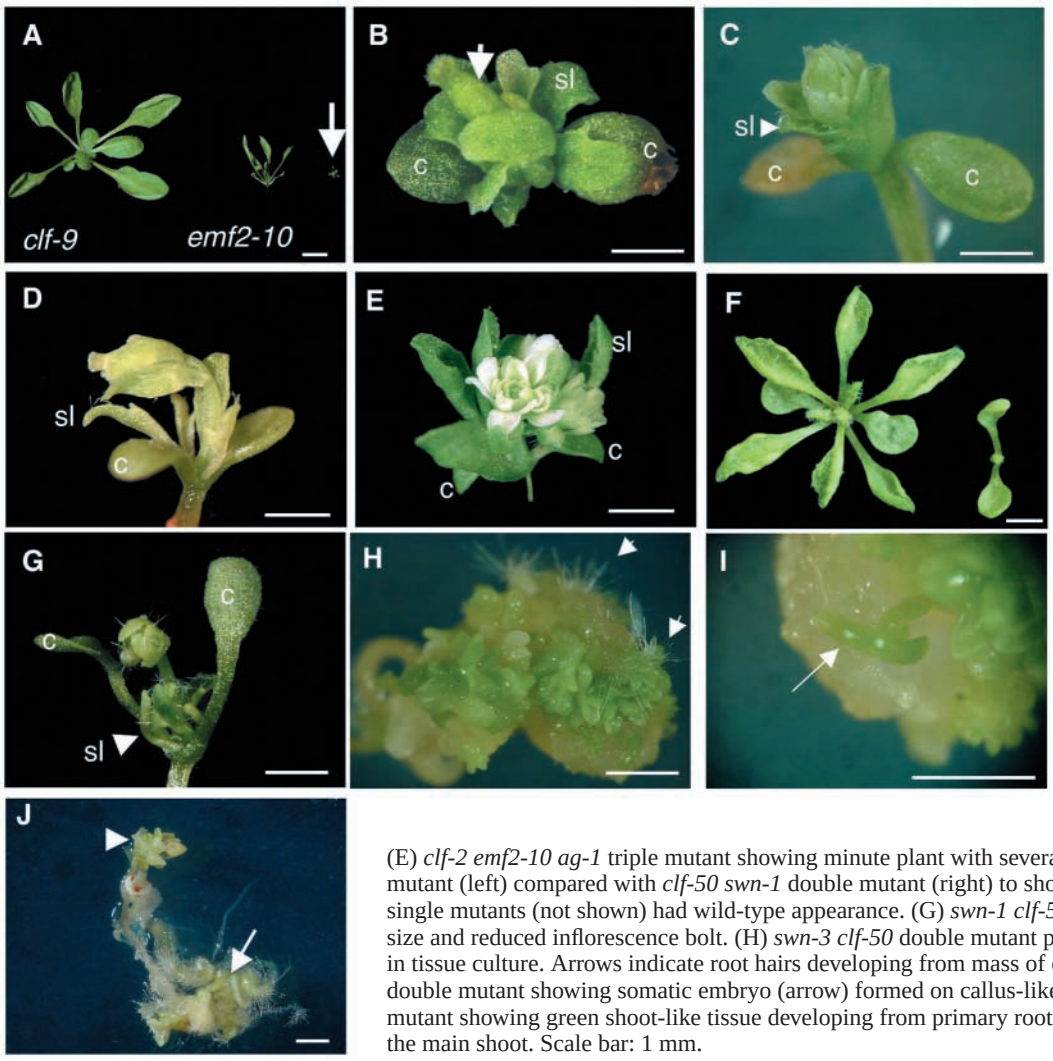

Fig. 3. Genetic interactions of Pc-G mutants. (A) Comparison of *emf2-10 clf-9* double mutant (arrow) and parental single mutants of same age. Note the minute size of the double mutant. (B) *emf2-10 clf-9* double mutant showing cotyledons (c), a few sessile leaves (sl) and a terminal flower. Tissue-culture-grown plant. (C) *emf2-3* single mutant, grown in tissue culture. Note similarity with B and D. (D) *clf-2 emf2-10* double mutant, grown in tissue culture. (E) *clf-2 emf2-10 ag-1* triple mutant showing minute plant with several *ag*– flowers. (F) *clf-50* single mutant (left) compared with *clf-50 swn-1* double mutant (right) to show enhanced phenotype. *swn-1* single mutants (not shown) had wild-type appearance. (G) *swn-1 clf-50* double mutant showing small size and reduced inflorescence bolt. (H) *swn-3 clf-50* double mutant plants after about 5 week's growth in tissue culture. Arrows indicate root hairs developing from mass of callus-like tissue. (I) *swn-3 clf-50* double mutant showing somatic embryo (arrow) formed on callus-like tissue. (J) *swn-3 clf-50* double mutant showing green shoot-like tissue developing from primary root (arrow). The arrowhead indicates the main shoot. Scale bar: 1 mm.

tests. Because *emf2* mutants are sterile, and moe leaf plants have low fertility, we crossed heterozygotes for the two mutations. The resulting F1 population of 254 plants contained 71 mutants, consistent with the two mutations being allelic (1/4 mutants expected, χ^2=1.2 P>0.1). We designated the new mutation responsible for the moe leaf phenotype as the *emf2-10* allele. The phenotype of *emf2-10/emf2-3* heterozygotes was intermediate between that of the two parental alleles, consistent with *emf2-10* being a weaker allele than *emf2-3* (Fig. 1O). To identify the lesion causing the *emf2-10* mutation, we compared the sequence of the *EMF2* locus from *emf2-10* and the wild-type progenitor. This revealed that the *emf2-10* allele carried a 17 bp deletion extending from the 3′ end of the second exon (9 bp) into the 5′ end of intron 2 (8 bp) followed by a cytosine to guanine substitution (see Fig. S1A,C in the supplementary material). Because this deletion was predicted to affect splicing of the *EMF2* pre-mRNA, we used RT-PCR to amplify *EMF2* cDNA from *emf2-10* and wild-type seedlings. Whereas a single message corresponding to the spliced *EMF2* transcript was detected in wild type cDNA, five novel transcripts were identified in *emf2-10* cDNA (see Fig. S1B in the supplementary material). Molecular cloning and sequencing of these aberrant transcripts indicated that four contained frameshift mutations likely to abolish EMF2 activity. However, one transcript was predicted to produce a variant EMF2 protein that was truncated by 17 amino acids at the N-terminus (see Fig. S1C in the supplementary material). The region deleted does not correspond to a conserved region or to a known functional domain, so the variant protein is likely to retain *EMF2*$^+$ activity. The weak emf2 phenotype may arise because only a small fraction (about 20%) of the various *emf2-10* transcripts are likely to produce a functional protein. In addition, the resulting truncation of the protein may also reduce its activity.

Genetic interaction of *EMF2* and *CLF*

The similarity in phenotypes of severe alleles of *CLF* and weak alleles of *EMF2* suggested that the two genes might act in a common genetic pathway. To test for a genetic interaction we therefore combined weak alleles of *CLF* and *EMF2* by constructing the *clf-9 emf2-10* double mutant. The weak *clf-9* allele was derived from the severe *clf-2* allele by an imprecise excision of a transposon from the *CLF* locus (P. Puangsomlee, Phd thesis, University of East Anglia, 1997) (Goodrich et al., 1997). *clf-9* plants are very similar to wild type but are slightly smaller, show earlier flowering under short days and very weak leaf curling (Fig. 3A). A synergistic interaction was observed, so that the double mutant had a much more extreme phenotype than either parent, producing extremely small plants with few, sessile leaves and very short inflorescences (Fig. 3A,B). The double mutant phenotype therefore resembled that of severe *emf2* alleles such as *emf2-3* (Fig. 3C). A similar phenotype was also observed in double mutant combinations of the null *clf-2* allele and weak *emf2-10* (Fig. 3D). In double mutant combinations of severe *emf2-3* and severe *clf* alleles, *emf2* was epistatic to *clf* (not shown). In general these observations were consistent with the two genes acting in a common pathway.

The severity of the *clf emf2* double mutant phenotype suggested that it was unlikely to result simply from misexpression of *AG*. To confirm this, we constructed *ag-1 clf-2 emf2-10* triple mutants. The triple mutants were minute plants, similar to severe *emf2* mutants in size, and had 1-3 flowers with normal petals and *ag*– phenotype (Fig. 3E). This indicated that *AG* misexpression was not responsible for the severe effects of *clf emf* mutants on overall plant size, but did account for the poor development of petals in *clf emf* mutant flowers (Fig. 3D).

Molecular interactions of CLF and EMF2

To test whether the genetic interaction of *CLF* and *EMF2* might reflect a direct interaction between their protein products, we performed yeast two-hybrid assays. We expressed full-length EMF2 protein, and a series of EMF2 truncations, as ‘prey’ fusions with the GAL4 transcriptional activation (TA) domain. We tested these fusion proteins for interaction with a ‘bait’ comprising a fusion of a truncated CLF protein (lacking the C-terminal SET domain) with the GAL4 DNA-binding domain. We did not observe an interaction between full-length EMF2 proteins with CLF in yeast. However, yeast strains expressing both CLF and a C-terminal portion of EMF2 expressed both two-hybrid reporter genes, consistent with the two proteins interacting (Fig. 4A). The C-terminal portion of EMF2 contained the VEFS domain, a motif originally defined on the basis of its conservation between plant and animal homologues of the Su(z)12 protein (Gendall et al., 2001). It was not clear why the full-length EMF2 protein, which includes the VEFS box, did not also interact with CLF.

To define the region of CLF that is required for interaction with EMF2, we tested a series of CLF truncations as baits with the EMF2 VEFS box prey construct (Fig. 4B). We thus mapped the interaction to a short 74 amino acid region of CLF that contains the C5 domain. The C5 domain contains five cysteine residues whose presence and spacing is conserved between plant, *Drosophila*, nematode and vertebrate E(z) homologues (Goodrich et al., 1997; Grossniklaus et al., 1998; Holdeman et al., 1998). No function has previously been ascribed to this domain. To verify the interaction between CLF and EMF2 in an independent system, we first used the yeast split-ubiquitin assay, which differs from the two-hybrid assay in that candidate proteins are fused to portions of the ubiquitin protein and the fusions are expressed in the cytoplasm rather than the nucleus (Johnsson and Varshavsky, 1994; Kim et al., 2002; Stagljar et al., 1998). Again, we observed an interaction of CLF with the VEFS box domain of EMF2 (Fig. 4C). Secondly, to confirm that CLF and EMF2 interact directly, we performed in-vitro binding assays (Fig. 4D). Both proteins were expressed in *E. coli*, the C5 domain of CLF as a glutathione-*S*-transferase (GST) fusion and the EMF2 VEFS domain as a HIS_6-tagged fusion. As shown in Fig. 4D, the HIS_6-EMF2 VEFS protein bound to GST–CLF C5 (lane B) but not to GST alone (lane C), suggesting a direct physical interaction between the proteins. Thus, the CLF C5 domain and EMF2 VEFS domain bound to each other in vitro as well as in yeast.

The *Arabidopsis FIS* genes *FIS2* and *MEA* encode homologues of EMF2 and CLF, respectively. Although the *FIS2* and *MEA* genes share extremely similar mutant phenotypes, suggesting that their products may interact, we were previously unable to demonstrate any interaction between the full-length proteins using the two-hybrid assay (Spillane et al., 2000). However, the observation that the interaction of EMF2 with CLF was mediated by the VEFS box domain suggested that FIS2 and MEA might also interact via the VEFS box. We therefore specifically tested the VEFS box of FIS2

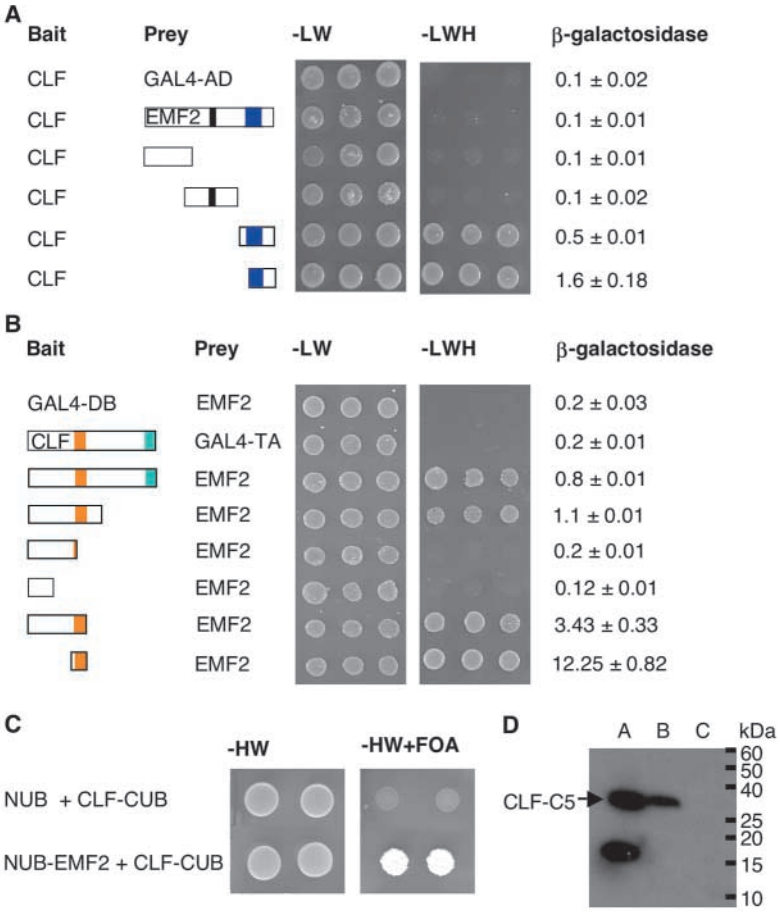

Fig. 4. Interaction of EMF2 and CLF in yeast and in vitro. Two-hybrid assays were performed in yeast strain HF7c. For nutritional assays, three independent transformants were stamped onto –LW and –LWH media. Growth on –LW selects for markers carried on the bait and prey plasmids, whereas growth on –LWH also indicates activity of the *HIS3* reporter gene. β-Galactosidase activity was quantified using the assay and units of Miller (Miller, 1972; Miller, 1992). Each value is an average from assays of three independent transformants; the standard error of the mean is also indicated. (A) The EMF2 protein is shown schematically, with the zinc finger motif indicated by the black box and the conserved VEFS domain by the blue box. The uppermost row is a control to show that the CLF bait does not have transcriptional activation activity by itself and cannot interact with an 'empty' GAL4-TA prey. The smallest region of EMF2 that was sufficient for interaction with CLF comprised residues 510-631. (B) The CLF protein is shown schematically, with the C5 domain indicated in orange and the CXC region, which precedes the SET domain, shown in turquoise. All prey fusions that contained an intact C5 domain were able to interact with the EMF2-VEFS domain (427-631). The shortest region of CLF sufficient for interaction comprised residues 257-331. (C) Split ubiquitin assay using the system of Kim et al. (Kim et al., 2002). CLF protein, lacking the C-terminal SET domain, was fused to the C-terminal half of ubiquitin (CUB) and the EMF2 VEFS domain was fused to an N-terminal portion of a modified ubiquitin (NUB). The NUB and CUB peptides are unable to interact on their own. Interaction of NUB and CUB fusions reconstitutes ubiquitin activity and results in proteolysis of a URA3 reporter. This allows growth on media containing FOA. Growth on –HW medium selects for the markers on the CUB and NUB constructs. (D) (lane A) In-vitro binding of CLF C5 domain and EMF2 VEFS domain. Bacterial extract containing His_6–EMF2 VEFS protein was tested for binding to GST–CLF C5 (lane B) or GST (lane C). Proteins that bound to GST or GST-CLF C5 were separated by SDS–PAGE, transferred to PVDF membrane, and incubated with anti-His_6 antibodies. The input lane (lane A) contains 1.5% of the volume of bacterial extract used in the binding assay. The lower band in the input lane corresponds to a His_6–EMF2 VEFS degradation product. Note that this is not bound by GST-CLF C5.

against MEA in two-hybrid assay and in this case were able to demonstrate an interaction (Fig. 5A).

Partial redundancy of *CLF* and the related *SWINGER* gene

The genetic and molecular interactions between *CLF* and *EMF2* suggested that their protein products probably acted in a common complex. However, two observations were at odds with this scenario: firstly, null *clf* alleles had much less severe phenotypes than null *emf2* alleles; secondly, the phenotype of null *clf* alleles was enhanced by *emf2* mutant alleles. One possible explanation was that the *CLF* gene showed redundancy, so that even in a null *clf* background, a similar activity was provided by other genes. This seemed a possibility because the *Arabidopsis* genome contains two other genes with strong similarity to *CLF*. The first, *MEA*, shows expression confined to the female gametophyte and early seed development and is therefore unlikely to overlap significantly with *CLF*, which is expressed predominantly during vegetative and inflorescence development (Goodrich et al., 1997; Vielle-Calzada et al., 1999). The second gene, accession At4g02020 [referred to as EZA1 by Luo et al. (Luo et al., 2000)], had not been genetically characterised. We designated this gene *SWINGER* (*SWN*), because our subsequent analysis of its protein product and mutant phenotype indicated a potential to share partners with the CLF protein (see below). Phylogenetic analysis of plant E(z) homologues (Fig. 5B) indicated that SWN and CLF belonged to distinct clades that can be clearly distinguished even in species distantly related to *Arabidopsis*, for example maize and rice. The duplication event that gave rise to *CLF* and *SWN* was therefore an ancient one within the angiosperm lineage. In addition, the *CLF* and *SWN* clades were clearly much more similar to one another than either was to *MEA*, suggesting that the function of *SWN* was more likely to resemble that of *CLF* than *MEA*.

To determine the *SWN* expression pattern, we localised its mRNA by in-situ hybridisation to sections of seedlings and inflorescences. *SWN* was expressed throughout the apical meristem and leaf primordia of 8-day-old wild-type seedlings (Fig. 6A,B). Expression was also detected in the vasculature of hypocotyls and cotyledons (Fig. 6B). In inflorescences, *SWN* was expressed throughout the inflorescence meristem and young stage 1-3 floral meristems (Fig. 6C). In older flowers, expression was weak in the sepals and stronger in the inner whorls containing developing petals, stamens and carpels (Fig. 6D,E). In stage 12 flowers, strongest expression occurred in the ovules, particularly in the funiculus and maternal tissues of the ovule (Fig. 6F). Expression was also seen in the female gametophyte, but the tissues were too poorly preserved to distinguish the different cell types within the gametopyhte (Fig. 6F). Little signal was observed when seedlings and inflorescences were

hybridised with a probe from the sense strand of the *SWN* cDNA (Fig. 6G,H), confirming that the signal was specific for the *SWN* antisense probe. As a positive control, we also hybridised seedlings with a probe for the *WUSCHEL* (*WUS*) gene and detected expression confined to the centre of the shoot meristem (Fig. 6I) as previously described (Mayer et al., 1998). The *SWN* expression pattern was therefore similar to that of *CLF* (Goodrich et al., 1997), with both genes being generally expressed during vegetative and reproductive development but with strongest expression in meristems and other regions of dividing cells.

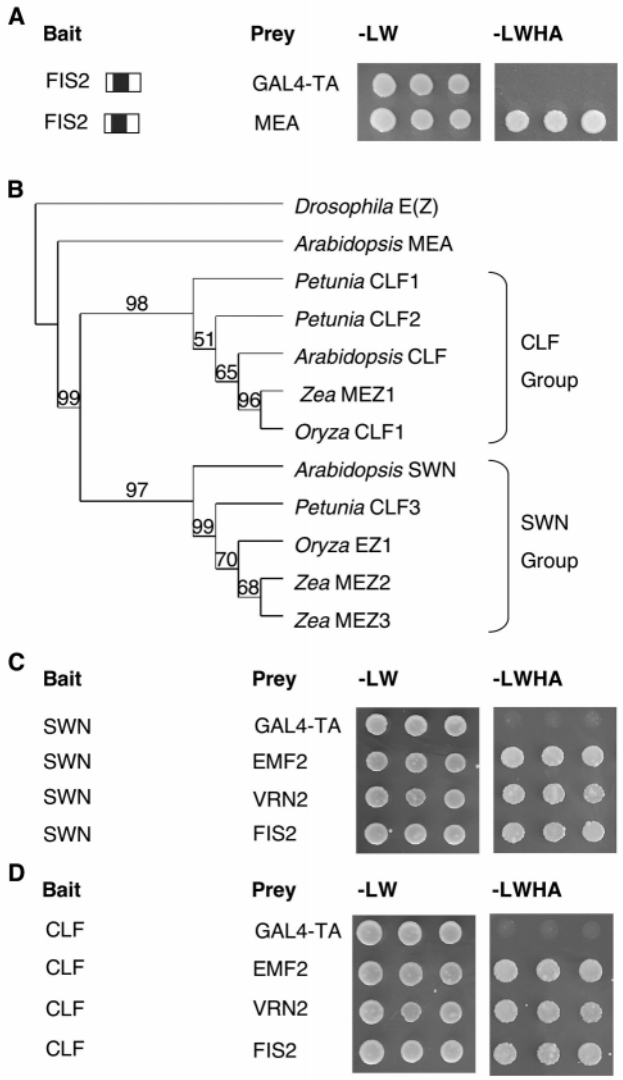

Fig. 5. Interaction between *Arabidopsis* CLF and EMF2 homologues in yeast. Two-hybrid constructs were introduced into the yeast strain AH109, which contains reporter genes that confer histidine and adenine prototrophy. The ADE3 reporter is extremely stringent. Three independent transformants were stamped onto –LW and –LWHA media. –LW selects for markers on the bait and prey constructs, while –LWHA also selects for activation of the *HIS3* and *ADE2* reporters. (A) Interaction of the VEFS domain of FIS2 (residues 466-692) with full-length MEA protein. (B) Cladogram of plant E(z) homologues. *Drosophila* E(z) is included as an outlier. Analysis was performed using the PAUP program to align the SET domains of the proteins. The bootstrap values are indicated. (C) Interaction of SWN C5 domain (252-331) with the VEFS domains of EMF2 (510-631), VRN2 (275-440) and FIS2 (394-692). (D) Interaction of CLF C5 domain (257-331) with the VEFS domains of EMF2, VRN2 and FIS2.

To test whether the SWN protein had similar properties to those of CLF, we compared their interactions in yeast two-hybrid assays. We observed an interaction between the EMF2 VEFS domain and both of the CLF or the SWN C5 domains, indicating that SWN had similar potential to interact with EMF2 to that of CLF (Fig. 5C,D). We further tested whether CLF and SWN could interact with the related *Arabidopsis* VEFS domain proteins FIS2 and VRN2. Both were able to interact with FIS2 and VRN2 in yeast, indicating a potential for one or both to function in the FIS and vernalisation response pathways (Fig. 5C,D). In addition, we found that both CLF and SWN can interact with FIE through a 110 amino acid motif at their N-termini (see Fig. S2 in the supplementary material) (Luo et al., 2000). Thus SWN and CLF showed similar interactions with both EMF2 and FIE in yeast. Taken together, the similarities in expression pattern and protein–protein interactions confirmed the potential for the *CLF* and *SWN* genes to act redundantly.

To identify the function of *SWN*, we exploited facilities for reverse genetics in *Arabidopsis* to identify a series of mutant alleles caused by T-DNA insertions within the locus. The *swn-1* allele contained an insertion 3 bp upstream of the predicted ATG start codon. This allele is unlikely to be null, as RT-PCR analysis of *swn-1* mRNA revealed chimeric transcripts that initiated within the T-DNA insertion and extended the full length of the *SWN* coding sequences (data not shown). The *swn-2* insertion is within an intron and *swn-3* within an exon, but both are upstream of the catalytic SET domain and are therefore likely to represent null alleles. All three alleles were viable as homozygotes and had no obvious phenotype that we could discern from inspection of gross plant morphology, embryo or endosperm development (data not shown). However, all three alleles strongly enhanced the *clf* mutant phenotype in *clf swn* double mutant combinations, confirming that the two genes exhibit redundancy. The *swn-1* allele gave a less severe enhancement than did *swn-2* or *swn-3*, consistent with its being a weaker allele. The *swn-1 clf-50* double mutant gave extremely small, early flowering plants with few flowers that resembled *emf2* mutants (Fig. 3F,G). SEM analysis indicated that the floral organs showed weak homeotic conversion to carpelloid structures (Fig. 2K). In addition, filamentous organs were observed in place of stipules, a phenotype that has also been observed in plants that have a partial loss of *FIE+* activity (Katz et al., 2004). Double mutants of the null *clf-50* allele with either *swn-2* or *swn-3* were more extreme, and viable plants were recovered only when seedlings were grown in sterile tissue culture. The seed germinated and produced seedlings with narrow, but relatively normal, cotyledons, hypocotyl and roots. As the plants aged they became increasingly abnormal. The cotyledons developed finger-like projections on their margins. The shoot apex did not initiate leaves, but instead developed into a disorganised mass of green tissue on which poorly differentiated organs formed (Fig. 3H). In SEM analysis of the plants, the epidermi of these organs lacked trichomes and comprised small, isodiametric cells, which did not have the surface cuticular thickening or elongated cell shape that is characteristic of epidermal surfaces

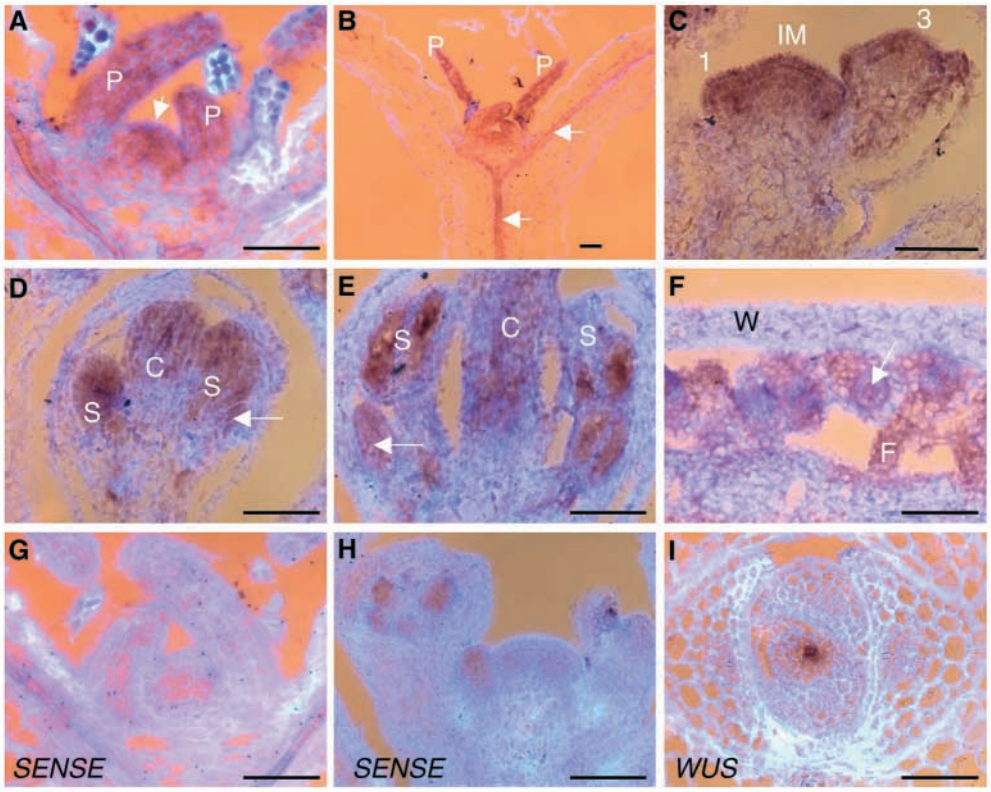

Fig. 6. Expression of *SWN*. Transcript localization by in-situ hybridisation. Tissue appears light blue, whereas the signal appears dark purple/brown. (A-F) were hybridised with an *SWN* antisense probe generated from a region at the 5′ end of the cDNA that lacked similarity with any other *Arabidopsis* gene. (A) Longitudinal section through an 8-day-old seedling showing expression in the shoot apical meristem (arrowed) and young leaf primordia (P). (B) Seedling section showing expression in vasculature (arrowed) and in older leaf primordia. (C) Longitudinal section through inflorescence showing strong expression in inflorescence meristem (IM), and throughout young stage 1 and 3 floral meristems. (D) Stage 7 flower showing expression in stamens (S), carpels (C) and emerging petal primordium (arrowed). (E) Stage 9 flower showing strong expression in emerging petal primordium (arrow), stamens (S) and carpels (C). (F) Longitudinal section through carpel of stage-12 flower showing mature ovules. Expression is low in the carpel walls (W), but strong throughout the sporophytic tissue of the ovule, particularly in the funiculus (F). Expression is also visible in the embryo sac (arrow). (G) Longitudinal section through seedling hybridised with *SWN* sense probe. (H) Longitudinal section through inflorescence hybridised with *SWN* sense probe. (I) Transverse section through 8-day-old seedling hybridised with *WUS* antisense probe. Signal is confined to the centre of the meristem, as previously described (Mayer et al., 1998).

of most of the mature floral organs (Fig. 2M,N). In addition, colourless callus-like tissue formed and eventually gave rise to somatic embryos and roots (Fig. 3H,I). Unlike the single mutants, which had normal roots, the primary root of the double mutants became opaque, swollen and eventually produced green shoot-like tissue (Fig. 3J). A similar phenotype has been observed in seedlings of rescued *fie* homozygotes (Kinoshita et al., 2001). Together, these observations suggested that weak *clf-50 swn-1* double mutants resembled *emf2* mutants, whereas the null *clf swn* doubles were more extreme and resembled plants lacking FIE^{+} activity.

Although the above data suggest that the *CLF* and *SWN* genes have very similar functions, the fact that *clf* mutants have a clear phenotype indicates that *SWN* is not identical in function to *CLF*, at least with respect to repression of *AG*. This might be due to subtle differences in level of expression between *CLF* and *SWN*, and/or changes in protein function. To clarify whether differences are solely due to changes in expression, we expressed full-length cDNA clones for each gene under control of a common promoter (the cauliflower mosaic virus 35S promoter) and introduced the two transgenes into the null *clf*-50 mutant background. Whereas the *35S::CLF* construct fully complemented the *clf*-50 mutation, the *35S:SWN* construct did not (Fig. 7). There are therefore subtle differences in function between the CLF and SWN proteins, as might be expected given the persistence of the *CLF*/*SWN* duplication within angiosperms. Expression of *35S::MEA* failed to complement the *clf*-50 mutation, indicating that the MEA protein has also diverged from CLF (Fig. 7).

Discussion

The C5 and VEFS domains mediate interaction between Pc-G proteins

The plant and animal Pc-G proteins of the E(z) class share several motifs of which the CXC and SET domains towards the C-termini of the proteins are the most highly conserved. The proteins also share a less-well-conserved region towards their N-termini, termed the C5 domain, that contains five cysteine residues with conserved spacing in the arrangement $CRRCX_2DCX_2HX_{(22\text{-}27)}CX_3CY$. The arrangement of cysteines does not correspond with any previously defined cysteine

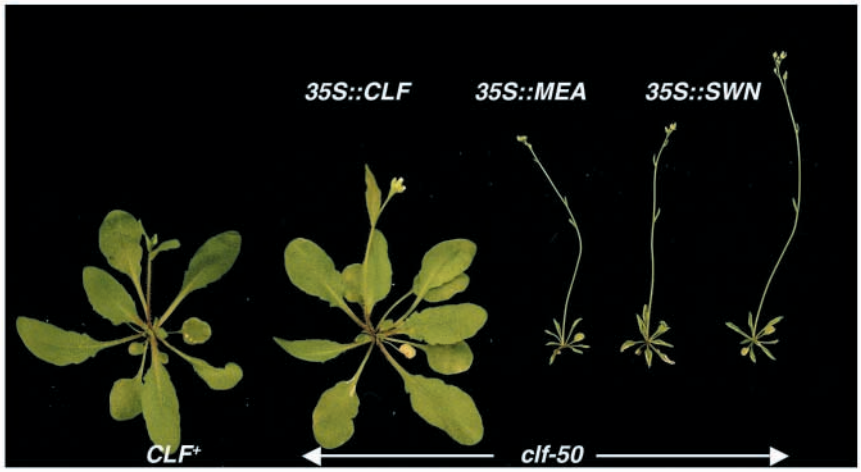

Fig. 7. Comparison of *CLF*, *MEA* and *SWN* misexpression. Transformed plants (T2 generation) containing *35S::CLF*, *35S::MEA* and *35S::SWN* transgenes in a *clf-50* mutant background. *35S::CLF* complements *clf-50* whereas the other two transgenes do not. At least 23 primary transformants were obtained for each construct. The *CLF+* progenitor and *clf-50* mutant are shown for comparison.

cluster motif such as the C2H2 zinc-finger motif involved in binding DNA. However, the region is functionally important because at least one mutant allele maps within the C5 domain: the *Drosophila E(z)*[28] allele is a mis-sense allele that swaps the fifth conserved cysteine for a tyrosine, and it gives a temperature-sensitive loss of function phenotype (Carrington and Jones, 1996). We show that this domain mediates the binding of plant E(z) homologues with the VEFS domain of plant Su(z)12 homologues. It is likely that the C5 and VEFS domains have a similar function in animals. Consistent with this, it was recently shown that mammalian Su(z)12 can interact through its VEFS domain with the mouse E(z) homologue EZH2 (Yamamoto et al., 2004). Although the region of EZH2 required for the interaction was not mapped, we note that the region expressed in this study (residues 238-746) included the C5 domain. It is also noticeable the C5 domain is less well conserved in the *Caenorhabditis elegans* E(z) homologue, MES2, than in the other animal proteins. *C. elegans*, unlike insects and vertebrates, lacks a Su(z)12 homologue, so the relatively poor conservation of the MES2 C5 domain may be because it no longer functions in this protein–protein interaction.

Conservation of the PRC2 complex between plants and animals

In animals, the core members of the PRC2 complex comprise four proteins first identified in *Drosophila* as the Esc, P55, E(z) and Su(z)12 proteins (Cao et al., 2002; Czermin et al., 2002; Kuzmichev et al., 2002; Muller et al., 2002). There is now strong evidence that structurally and functionally equivalent complexes occur in *Arabidopsis*. Thus several previous studies have shown genetic and physical interactions of FIE with MEA and CLF, and also of FIE with MSI1 (for a review, see Reyes and Grossniklaus, 2003) (Fig. 8). In particular, Kohler et al. (Kohler et al., 2003a) partially purified an FIS complex and showed that it contained FIE, MEA, MSI1 and, based on molecular weights, probably several other unidentified components. However, the role of the FIS2, VRN2 and EMF2 proteins has remained enigmatic, although the strong similarity between the *fis* mutant phenotypes suggested that FIS2 might interact with one or more of the other FIS proteins. We have now shown that EMF2 interacts physically and genetically with CLF. We extend this to show a general potential for the Su(z)12 homologues FIS2 and VRN2 to interact with the E(z) homologues MEA, CLF, and SWN, at least in yeast two-hybrid assays. Taken together, these observations strongly suggest that there are *Arabidopsis* complex(es) that are structurally equivalent to at least the core members of the animal PRC2 members. It is also likely that they have an equivalent biochemical function in mK27 H3 histone methylation. Thus, the *Arabidopsis* VRN2 protein was recently shown to be required for vernalisation-induced mK27 H3 methylation at the *FLC* gene (Bastow et al., 2004; Sung and Amasino, 2004). However, biochemical purification of the plant PRC2 complexes will be necessary to confirm that they have a direct HMTase activity.

Diversification of PRC2 function in plants

Whereas the *FIE* gene is single copy, the other *Arabidopsis* PRC2 members are represented by small gene families with three to four members. Within these families, the different

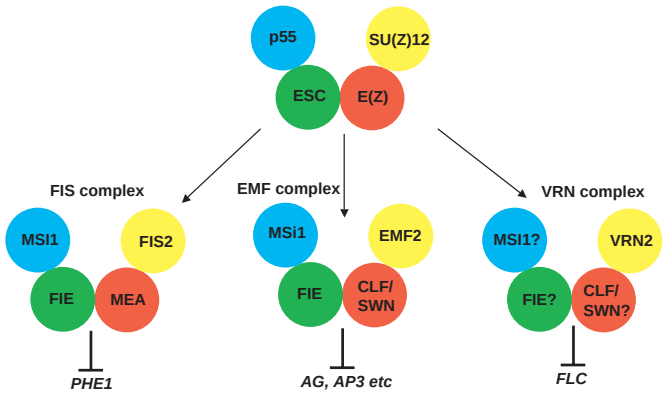

Fig. 8. *Arabidopsis* Polycomb-group protein complexes. The core components of the *Drosophila* PRC2 complex are shown at top. In *Arabidopsis*, an equivalent ancestral complex is proposed to have diversified into three similar complexes with at least partially discrete functions. The colours indicate homology; so for example, E(z) homologues are coloured red. The contacts indicate interactions; for example, FIE can interact with MEA and MSI1 but not FIS2, whereas FIS2 can interact with MEA but not with other FIS proteins. The target genes shown are not comprehensive; it is likely that all three complexes have many more targets than those shown.

members control at least three different processes: firstly, repression of endosperm proliferation during gametophyte and endosperm development (*FIS2*, *MEA*); secondly, repression of floral homeotic gene expression during embryo development and vegetative development (*EMF2*, *CLF*); thirdly, epigenetic control of vernalisation response (*VRN2*). We suggest that these reflect their participation in at least three complexes that differ in their target gene specificity (Fig. 7) (Reyes and Grossniklaus, 2003). The distinct roles of the *Arabidopsis* PRC2 members may in part reflect differences in expression patterns. For example, several studies suggest that *FIS2* and *MEA* expression is confined to female gametophyte and seed development, whereas *EMF2* and *CLF* are also expressed more broadly during vegetative development, where they act to repress genes controlling flowering time or floral development (Goodrich et al., 1997; Luo et al., 2000; Vielle-Calzada et al., 1999; Yoshida et al., 2001). However, differences in expression are not sufficient to account for the altered roles. Thus, even when *CLF*, *MEA* and *SWN* cDNAs are expressed under control of a common promoter (CaMV 35S), only the *CLF* transgene is able to complement *clf* mutants. This suggests that differences between the CLF, MEA and SWN proteins are also important. Thus, following duplication, the plant Pc-G genes appear to have diverged in protein function as well as expression.

It is unclear how these complexes might acquire specificity for different target genes, as PRC2 members appear to lack intrinsic DNA-binding specificity and the recruitment of Pc-G members to specific targets is not yet well understood either in animals or plants (Birve et al., 2001; Carrington and Jones, 1996). One possibility is that PRC2 members are recruited to targets by interaction with sequence-specific DNA-binding proteins (Wang et al., 2004). A recent alternative model is that Pc-G members could achieve sequence specificity through interactions with small RNAs (Steimer et al., 2004). Because the *FIE* gene is a single copy, and its protein product is probably common to all complexes, it is unlikely that FIE could distinguish the activity of different complexes. However, small differences between the EMF2/VRN2/FIS2 and/or MEA/CLF/SWN proteins could change their affinities for protein partners that target the complex. It is striking that the FIS2/VRN2/EMF2 class of protein is the only one of the PRC2 members that is not also conserved in *C. elegans*. This implies that this protein is not absolutely required for the biochemical activity of the complex, and might therefore play a role in specifying its targets. In addition to differences between complexes in their target gene specificities, there must also be differences between the CLF family members and the EMF2 family members in their affinity for one another. For example, if CLF has equal affinity for FIS2, EMF2 and VRN2 and the FIS2 family members have equal affinity for CLF, MEA and SWN, then proteins such as MEA and CLF should be able to cross-complement one another when misexpressed. We did not observe such differences in yeast two-hybrid assays; for example, CLF and SWN showed a similar potential to interact with each of the EMF2, VRN2 and FIS2 proteins. However, the interactions in yeast may not accurately reflect subtle differences in affinity in plants. It will be interesting to test whether swapping the C5 domains between CLF and MEA proteins, or other regions, can alter their specificity in vivo.

Partial redundancy between *CLF* and *SWN*

The *CLF* and *SWN* genes show similar expression patterns and encode closely related proteins that display identical interactions in several yeast two-hybrid assays. We tested three independent *swn* mutant alleles and all three strongly enhance the *clf* single mutant phenotype, although they are without gross morphological phenotype by themselves. This suggests that there is substantial functional redundancy between the two genes, so that the roles of *CLF* are largely masked by *SWN* activity in *clf* single mutants. For example, a role for *CLF* and *SWN* in primary root development is not apparent from either single mutant phenotypes but is revealed in null *clf swn* doubles. The partial redundancy of *CLF* and *SWN* probably explains why null *clf* mutants have much less extreme phenotypes than null *emf2* mutants, although the CLF and EMF2 proteins act together. Consistent with this, weak *swn-1 clf-50* double mutants resembled *emf2* mutants. However, null *swn clf* double mutants were more extreme than *emf2* mutants and resembled plants lacking FIE^+ activity. It is likely that the full function of *EMF2* is also masked by partial redundancy; for example, with the related *VRN2* gene with which it shares overlapping expression.

Despite overlapping functions, *CLF* and *SWN* are not completely redundant with respect to one another: firstly, *clf* mutants have a phenotype, largely caused by ectopic *AG* expression, that is not complemented by *SWN+* activity; secondly, *35S::SWN*, unlike *35S::CLF*, does not complement *clf* mutants; thirdly, phylogenetic comparisons indicate that *CLF* and *SWN* orthologues can be distinguished clearly in other plants, including monocotyledenous species such as rice and maize. This means that the *CLF/SWN* duplication is an ancient one within the angiosperm lineage. It is unlikely that the *SWN* gene would show such wide conservation if it did not have at least partially distinct functions from *CLF*. Although we did not identify gross morphological effects of null *swn* mutations, several of the phenotypes associated with other plant PRC2 members (for example, autonomous endosperm development or vernalisation response) are apparent only in specific phenotypic screens or genetic backgrounds. It is therefore likely that *swn* mutants do have a phenotype, but this was not manifest in our growth conditions or assays.

The potential for CLF and SWN to act in vernalisation response

Recently, it was shown that *VRN2* is required, after vernalisation treatments, for mK27 H3 methylation in chromatin of its target gene *FLC* (Bastow et al., 2004; Sung and Amasino, 2004). In animals, mK27 H3 methylation by the PRC2 complex requires the E(z) protein, which contains a SET domain known to have HMTase activity (Cao et al., 2002; Czermin et al., 2002; Kuzmichev et al., 2002; Muller et al., 2002). Together these observations suggest that an *E(z)* homologue will be required for the vernalisation response. Consistent with this, we show that the VRN2 protein has potential to interact, through its VEFS domain, with the C5 domain of the E(z) homologues CLF and SWN. In preliminary experiments (data not shown), we did not observe gross effects of null *clf* or *swn* mutations on the vernalisation response comparable with those of *vrn2* or other vernalisation response mutants. It is possible that *CLF* and *SWN* act redundantly with respect to the vernalisation response, so that defects will be

manifest only in double mutants. Unfortunately, the pleiotropic phenotype of *clf swn* double mutants makes it difficult to characterise their vernalisation response, at least by straightforward comparison of flowering times. One possibility will be to use chromatin immunoprecipitation (ChIP) to test whether FLC chromatin becomes enriched for CLF and/or SWN proteins following vernalisation treatments.

In summary, it is likely that the PRC2 complex is conserved between plant and animals, both structurally and also functionally in terms of its histone methylation activity. However, in plants there has been duplication of most components of the PRC2, and the duplicated members have diverged in protein function as well as in expression. This has given rise to several PRC2-like complexes in plants, with at least partially discrete functions in terms of target gene specificity. Expression of chimeric proteins that swap domains between duplicated components, such as CLF/SWN/MEA, may help identify how the changes in specificity are mediated. Despite the conservation of the PRC2 in plants, it is striking that there are no homologues of the animal Pc-G members that comprise the PRC1 complex. It is therefore possible that the mechanisms to interpret, maintain, and re-set epigenetic information conveyed by the PRC2 have evolved independently in plants. Alternatively, plants may employ similar protein motifs to those found in the animal PRC1 members, but in novel combinations.

J.G. was funded by a Royal Society University Research fellowship, Y.C. by a scholarship from the Government of Thailand, C.S. and A.B. by BBSRC PhD studentship awards, D.S. by a BBSRC postdoctoral fellowship and a fellowship from the German Academic Exchange Service and Z.R.S. by USDA99-35301-7984 and NSF IBN-0236399. We thank Chris Jeffree for help with SEM, Amelia Green for help with the yeast two-hybrid analysis, Magali Bic and Neil Haig for help with the in-vitro pull down assays, Mark Running and Elliot Meyerowitz for generously providing the *emf2-10* allele and Caroline Dean for providing vernalisation requiring backgrounds and a *VRN2* cDNA. We thank Laurent Deslandes and Imre Somssich for providing vectors for split-ubiquitin analysis and Nobamusa Yoshida for providing *EMF2* cDNA clones.

Supplementary material

Supplementary material for this article is available at
http://dev.biologists.org/cgi/content/full/131/21/5263/DC1

References

Ach, R. A., Taranto, P. and Gruissem, W. (1997). A conserved family of WD-40 proteins binds to the retinoblastoma protein in both plants and animals. *Plant Cell* **9**, 1595-1606.

Alonso, J. M., Stepanova, A. N., Leisse, T. J., Kim, C. J., Chen, H., Shinn, P., Stevenson, D. K., Zimmerman, J., Barajas, P., Cheuk, R. et al. (2003). Genome-wide insertional mutagenesis of Arabidopsis thaliana. *Science* **301**, 653-657.

Bastow, R., Mylne, J. S., Lister, C., Lippman, Z., Martienssen, R. A. and Dean, C. (2004). Vernalization requires epigenetic silencing of FLC by histone methylation. *Nature* **427**, 164-167.

Birve, A., Sengupta, A. K., Beuchle, D., Larsson, J., Kennison, J. A., Rasmuson-Lestander, A. and Muller, J. (2001). Su(z)12, a novel Drosophila Polycomb group gene that is conserved in vertebrates and plants. *Development* **128**, 3371-3379.

Busch, M. A., Bomblies, K. and Weigel, D. (1999). Activation of a floral homeotic gene in Arabidopsis. *Science* **285**, 585-587.

Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R. S. and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* **298**, 1039-1043.

Carrington, E. A. and Jones, R. S. (1996). The Drosophila Enhancer of zeste gene encodes a chromosomal protein: examination of wild-type and mutant protein distribution. *Development* **122**, 4073-4083.

Chaudhury, A. M., Ming, L., Miller, C., Craig, S., Dennis, E. S. and Peacock, W. J. (1997). Fertilization-independent seed development in Arabidopsis thaliana. *Proc. Natl. Acad. Sci. USA* **94**, 4223-4228.

Chen, L., Cheng, J. C., Castle, L. and Sung, Z. R. (1997). EMF genes regulate Arabidopsis inflorescence development. *Plant Cell* **9**, 2011-2024.

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.

Coen, E. (1999). *The Art of Genes*. Oxford: Oxford University Press.

Czermin, B., Melfi, R., McCabe, D., Seitz, V., Imhof, A. and Pirrotta, V. (2002). Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell* **111**, 185-196.

Feilotter, H. E., Hannon, G. J., Ruddell, C. J. and Beach, D. (1994). Construction of an improved host strain for two hybrid screening. *Nucleic Acids Res.* **22**, 1502-1503.

Finnegan, E. J., Peacock, W. J. and Dennis, E. S. (1996). Reduced DNA methylation in Arabidopsis thaliana results in abnormal plant development. *Proc. Natl. Acad. Sci. USA* **93**, 8449-8454.

Francis, N. J. and Kingston, R. E. (2001). Mechanisms of transcriptional memory. *Nat. Rev. Mol. Cell. Biol.* **2**, 409-421.

Francis, N. J., Saurin, A. J., Shao, Z. and Kingston, R. E. (2001). Reconstitution of a functional core polycomb repressive complex. *Mol. Cell* **8**, 545-556.

Gendall, A. R., Levy, Y. Y., Wilson, A. and Dean, C. (2001). The VERNALIZATION 2 gene mediates the epigenetic regulation of vernalization in Arabidopsis. *Cell* **107**, 525-535.

Gleave, A. P. (1992). A versatile binary vector system with a T-DNA organisational structure conducive to efficient integration of cloned DNA into the plant genome. *Plant Mol. Biol.* **20**, 1203-1207.

Goodrich, J., Puangsomlee, P., Martin, M., Long, D., Meyerowitz, E. M. and Coupland, G. (1997). A Polycomb-group gene regulates homeotic gene expression in Arabidopsis. *Nature* **386**, 44-51.

Grossniklaus, U., Vielle-Calzada, J. P., Hoeppner, M. A. and Gagliano, W. B. (1998). Maternal control of embryogenesis by MEDEA, a polycomb group gene in Arabidopsis. *Science* **280**, 446-450.

Guitton, A. E., Page, D. R., Chambrier, P., Lionnet, C., Faure, J. E., Grossniklaus, U. and Berger, F. (2004). Identification of new members of Fertilisation Independent Seed Polycomb Group pathway involved in the control of seed development in Arabidopsis thaliana. *Development* **131**, 2971-2981.

Hennig, L., Taranto, P., Walser, M., Schonrock, N. and Gruissem, W. (2003). Arabidopsis MSI1 is required for epigenetic maintenance of reproductive development. *Development* **130**, 2555-2565.

Holdeman, R., Nehrt, S. and Strome, S. (1998). MES-2, a maternal protein essential for viability of the germline in Caenorhabditis elegans, is homologous to a Drosophila Polycomb group protein. *Development* **125**, 2457-2467.

Jack, T., Fox, G. L. and Meyerowitz, E. M. (1994). Arabidopsis homeotic gene APETALA3 ectopic expression: transcriptional and posttranscriptional regulation determine floral organ identity. *Cell* **76**, 703-716.

James, P., Halladay, J. and Craig, E. A. (1996). Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics* **144**, 1425-1436.

Jeffree, C. E. and Read, N. D. R. (1991). Ambient- and low-temperature scanning electron microscopy. In *Electron Microscopy of Plant Cells* (ed. J. L. Hall and C. Hawes), pp. 313-414. London: Academic Press.

Johnsson, N. and Varshavsky, A. (1994). Split ubiquitin as a sensor of protein interactions in vivo. *Proc. Natl. Acad. Sci. USA* **91**, 10340-10344.

Katz, A., Oliva, M., Mosquna, A., Hakim, O. and Ohad, N. (2004). FIE and CURLY LEAF polycomb proteins interact in the regulation of homeobox gene expression during sporophyte development. *Plant J.* **37**, 707-719.

Kenzior, A. L. and Folk, W. R. (1998). AtMSI4 and RbAp48 WD-40 repeat proteins bind metal ions. *FEBS Lett.* **440**, 425-429.

Kim, G. T., Tsukaya, H. and Uchimiya, H. (1998). The CURLY LEAF gene controls both division and elongation of cells during the expansion of the leaf blade in Arabidopsis thaliana. *Planta* **206**, 175-183.

Kim, M. C., Panstruga, R., Elliott, C., Muller, J., Devoto, A., Yoon, H. W., Park, H. C., Cho, M. J. and Schulze-Lefert, P. (2002). Calmodulin interacts with MLO protein to regulate defence against mildew in barley. *Nature* **416**, 447-451.

Kinoshita, T., Harada, J. J., Goldberg, R. B. and Fischer, R. L. (2001). Polycomb repression of flowering during early plant development. *Proc. Natl. Acad. Sci. USA* **98**, 14156-14161.

Kiyosue, T., Ohad, N., Yadegari, R., Hannon, M., Dinneny, J., Wells, D., Katz, A., Margossian, L., Harada, J. J., Goldberg, R. B. et al. (1999). Control of fertilization-independent endosperm development by the MEDEA polycomb gene in Arabidopsis. *Proc. Natl. Acad. Sci. USA* **96**, 4186-4191.

Kohler, C., Hennig, L., Bouveret, R., Gheyselinck, J., Grossniklaus, U. and Gruissem, W. (2003a). Arabidopsis MSI1 is a component of the MEA/FIE Polycomb group complex and required for seed development. *EMBO J.* **22**, 4804-4814.

Kohler, C., Hennig, L., Spillane, C., Pien, S., Gruissem, W. and Grossniklaus, U. (2003b). The Polycomb-group protein MEDEA regulates seed development by controlling expression of the MADS-box gene PHERES1. *Genes Dev.* **17**, 1540-1553.

Koncz, C. and Schell, J. (1986). The promoter of the T-DNA gene 5 controls the tissue-specific expression of chimeric genes carried by a novel type of *Agrobacterium* binary vector. *Mol. Gen. Genet.* **204**, 383-396.

Krysan, P. J., Young, J. C. and Sussman, M. R. (1999). T-DNA as an insertional mutagen in Arabidopsis. *Plant Cell* **11**, 2283-2290.

Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P. and Reinberg, D. (2002). Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes Dev.* **16**, 2893-2905.

Lawrence, P. A. and Struhl, G. (1996). Morphogens, compartments, and pattern: lessons from Drosophila? *Cell* **85**, 951-961.

Luo, M., Bilodeau, P., Dennis, E. S., Peacock, W. J. and Chaudhury, A. (2000). Expression and parent-of-origin effects for FIS2, MEA, and FIE in the endosperm and embryo of developing Arabidopsis seeds. *Proc. Natl. Acad. Sci. USA* **97**, 10637-10642.

Luo, M., Bilodeau, P., Koltunow, A., Dennis, E. S., Peacock, W. J. and Chaudhury, A. M. (1999). Genes controlling fertilization-independent seed development in Arabidopsis thaliana. *Proc. Natl. Acad. Sci. USA* **96**, 296-301.

Mayer, K. F., Schoof, H., Haecker, A., Lenhard, M., Jurgens, G. and Laux, T. (1998). Role of WUSCHEL in regulating stem cell fate in the Arabidopsis shoot meristem. *Cell* **95**, 805-815.

Michaels, S. D. and Amasino, R. M. (1999). FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. *Plant Cell* **11**, 949-956.

Miller, J. H. (1972). *Experiments in Molecular Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Miller, J. H. (1992). In *A Short Course in Bacterial Genetics*, p. 74. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Moon, Y. H., Chen, L., Pan, R. L., Chang, H. S., Zhu, T., Maffeo, D. M. and Sung, Z. R. (2003). EMF genes maintain vegetative development by repressing the flower program in Arabidopsis. *Plant Cell* **15**, 681-693.

Muller, J., Hart, C. M., Francis, N. J., Vargas, M. L., Sengupta, A., Wild, B., Miller, E. L., O'Connor, M. B., Kingston, R. E. and Simon, J. A. (2002). Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. *Cell* **111**, 197-208.

Narita, N. N., Moore, S., Horiguchi, G., Kubo, M., Demura, T., Fukuda, H., Goodrich, J. and Tsukaya, H. (2004). Overexpression of a novel small peptide ROTUNDIFOLIA4 decreases cell proliferation and alters leaf shape in Arabidopsis thaliana. *Plant J.* **38**, 699-713.

Ohad, N., Margossian, L., Hsu, Y. C., Williams, C., Repetti, P. and Fischer, R. L. (1996). A mutation that allows endosperm development without fertilization. *Proc. Natl. Acad. Sci. USA* **93**, 5319-5324.

Ohad, N., Yadegari, R., Margossian, L., Hannon, M., Michaeli, D., Harada, J. J., Goldberg, R. B. and Fischer, R. L. (1999). Mutations in FIE, a WD polycomb group gene, allow endosperm development without fertilization. *Plant Cell* **11**, 407-416.

Reyes, J. C. and Grossniklaus, U. (2003). Diverse functions of Polycomb group proteins during plant development. *Semin. Cell Dev. Biol.* **14**, 77-84.

Russo, V. E. A., Martienssen, R. A. and Riggs, A. D. (1996). *Epigenetic Mechanisms of Gene Regulation*. Cold spring Harbor Laboratory Press, New York.

Saurin, A. J., Shao, Z., Erdjument-Bromage, H., Tempst, P. and Kingston, R. E. (2001). A Drosophila Polycomb group complex includes Zeste and dTAFII proteins. *Nature* **412**, 655-660.

Seki, M., Narusaka, M., Kamiya, A., Ishida, J., Satou, M., Sakurai, T., Nakajima, M., Enju, A., Akiyama, K., Oono, Y. et al. (2002). Functional annotation of a full-length Arabidopsis cDNA collection. *Science* **296**, 141-145.

Serrano-Cartagena, J., Candela, H., Robles, P., Ponce, M. R., Perez-Perez, J. M., Piqueras, P. and Micol, J. L. (2000). Genetic analysis of incurvata mutants reveals three independent genetic operations at work in Arabidopsis leaf morphogenesis. *Genetics* **156**, 1363-1377.

Sheldon, C. C., Burn, J. E., Perez, P. P., Metzger, J., Edwards, J. A., Peacock, W. J. and Dennis, E. S. (1999). The FLF MADS box gene: a repressor of flowering in Arabidopsis regulated by vernalization and methylation. *Plant Cell* **11**, 445-458.

Sieburth, L. E. and Meyerowitz, E. M. (1997). Molecular dissection of the AGAMOUS control region shows that cis elements for spatial regulation are located intragenically. *Plant Cell* **9**, 355-365.

Spillane, C., MacDougall, C., Stock, C., Kohler, C., Vielle-Calzada, J. P., Nunes, S. M., Grossniklaus, U. and Goodrich, J. (2000). Interaction of the Arabidopsis polycomb group proteins FIE and MEA mediates their common phenotypes. *Curr. Biol.* **10**, 1535-1538.

Stagljar, I., Korostensky, C., Johnsson, N. and te Heesen, S. (1998). A genetic system based on split-ubiquitin for the analysis of interactions between membrane proteins in vivo. *Proc. Natl. Acad. Sci. USA* **95**, 5187-5192.

Steimer, A., Schob, H. and Grossniklaus, U. (2004). Epigenetic control of plant development: new layers of complexity. *Curr. Opin. Plant Biol.* **7**, 11-19.

Stern, C. (1968). Developmental genetics of pattern. In *Genetic Mosaics and Other Essays*, pp. 130-173. Cambridge, Massachusetts: Harvard University Press.

Sung, S. and Amasino, R. M. (2004). Vernalization in Arabidopsis thaliana is mediated by the PHD finger protein VIN3. *Nature* **427**, 159-164.

Sung, Z. R., Belachev, A., Bai, S. and Betrand-Garcia, R. (1992). *EMF*, an *Arabidopsis* gene required for vegetative shoot development. *Science* **258**, 1645-1647.

Sung, Z. R., Chen, L., Moon, Y.-H. and Yoshida, N. (2003). Molecular mechanisms of shoot determinacy and flowering in *Arabidopsis*. *HortScience* **38**, 1325-1327.

Vielle-Calzada, J. P., Thomas, J., Spillane, C., Coluccio, A., Hoeppner, M. A. and Grossniklaus, U. (1999). Maintenance of genomic imprinting at the Arabidopsis medea locus requires zygotic DDM1 activity. *Genes Dev.* **13**, 2971-2982.

Wang, L., Brown, J. L., Cao, R., Zhang, Y., Kassis, J. A. and Jones, R. S. (2004). Hierarchical recruitment of polycomb group silencing complexes. *Mol. Cell* **14**, 637-646.

Yadegari, R., Kinoshita, T., Lotan, O., Cohen, G., Katz, A., Choi, Y., Katz, A., Nakashima, K., Harada, J. J., Goldberg, R. B. et al. (2000). Mutations in the FIE and MEA genes that encode interacting polycomb proteins cause parent-of-origin effects on seed development by distinct mechanisms. *Plant Cell* **12**, 2367-2382.

Yamamoto, K., Sonoda, M., Inokuchi, J., Shirasawa, S. and Sasazuki, T. (2004). Polycomb Group Suppressor of Zeste 12 Links Heterochromatin Protein 1{alpha} and Enhancer of Zeste 2. *J. Biol. Chem.* **279**, 401-406.

Yang, C. H., Chen, L. J. and Sung, Z. R. (1995). Genetic regulation of shoot development in Arabidopsis: role of the EMF genes. *Dev. Biol.* **169**, 421-435.

Yoshida, N., Yanai, Y., Chen, L., Kato, Y., Hiratsuka, J., Miwa, T., Sung, Z. R. and Takahashi, S. (2001). EMBRYONIC FLOWER2, a novel polycomb group protein homolog, mediates shoot development and flowering in Arabidopsis. *Plant Cell* **13**, 2471-2481.