Investigating the roles of *Arabidopsis* Polycomb-group genes in regulating flowering time and during plant development by (I) challenging silencing and (II) developing approaches to dissect Pc-G action

Kate M. Creasey

Supervisor Dr. Justin Goodrich

Thesis submitted for the Degree of Doctor of Philosophy

University of Edinburgh

2009

II. Declaration

I declare that this is my own work and contributions made by others are clearly cited.

III. Abstract

Polycomb-group (Pc-G) proteins regulate homeotic gene silencing associated with the repressive covalent histone modification, trimethylation of histone H3 lysine 27 (H3K27me3). Pc-G mediated silencing is believed to remodel chromatin, rendering target genes inaccessible to transcription factors. Pc-G mediated silencing might result in irreversible changes in chromatin structure, however, there has been little analysis addressing whether Pc-G mediated silencing is reversible. In this work we focused on *CURLY LEAF* (*CLF*), the first Pc-G homologue discovered in *Arabidopsis.* CLF mediated repression of the floral homeotic gene *AGAMOUS* (*AG)* was challenged during early and late leaf development. *AG* was activated by the late leaf promoter, revealing that Pc-G mediated silencing can be overcome in old leaves in the presence of CLF. *AG* was also activated in young leaf primordia, yet did not persist in older leaves, revealing that transient activation of a Pc-G target is not epigenetically stable. To address the mechanism of Pc-G action within an endogenous environment, the histone dynamics at the *APETALA1* (*AP1*) locus were characterized by Chromatin Immunoprecipitation. Unexpectedly, we found that the activation of *AP1* in leaves did not require the removal of H3K27me3, questioning whether H3K27me3 is sufficient to silence.

The roles of *CLF* in leaf and flower development are masked due to partial redundancy with *SWINGER* (*SWN*). *clf- swn-* mutants form a callus-like mass on sterile-tissue culture with no distinguishable plant organs. The role of *CLF* in regulating flowering time in natural populations of *A. thaliana* was investigated by complementing *clf-* mutants with *CLF* alleles from two accessions. We found that natural variation in *CLF* did not affect flowering time. To dissect the roles of *CLF* and *SWN* in late leaf and flower development, two approaches were developed for targeted expression. Firstly, *CLF* was introduced into the *LhG4/ pOp* transactivation system to provide *CLF* during early plant development. For mosaic analysis, *CLF* was introduced into the *CRE lox* recombination system in order to create *clf-* sectors surrounded by *CLF+ SWN+* and *CLF+ swn-* cells.

IV. Acknowledgements

I would like to thank my supervisor Dr. Justin Goodrich for giving me the opportunity to carry out my PhD in his group. I have gained so much through this experience and feel very fortunate to have been given the freedom to pursue my own ideas as well as his guidance in all aspects of research.

I would like to take this opportunity to thank my parents, Mr. Kenneth W. Creasey and Mrs. Susan J. Creasey for encouraging my interest in science by giving me the right books and allowing me to keep pets when I was younger. I am eternally grateful for being given the independence to make my own choices, but always having their unfailing support and confidence in me.

I would like to say a great thank you to my colleagues; Dr. Richard Milne for his advice and amazing ability to keep up with my thoughts and follow the random tangents. To Dr. Eve-Marie Josse for all her great insight and practical advice in the lab, and Miss. Olga Linnik for her positive attitude and laughter. I cannot say thank you enough to them for giving me their time to discuss everything and anything, they are all truly great friends.

To acknowledge the appreciation and gratitude for my partner Jan would never touch upon how thankful I am to share my life with someone who understands what all of this means to me. Thank you for being there for me, supporting and encouraging me to see the best in any situation.

V. Abbreviations

bp	base pair(s)
cDNA	complementary DNA
CIAP	calf intestinal alkaline phosphotase
CSPD	Disodium3-(4-methoxyspiro1,2-dioxetane-3,2'-(5-chloro)tricycle [3.3.1.1] decan}-4-yl) phenol phosphate
DNA	deoxyribonucleic acid
dNTPs	deoxyribonucleotide triphosphate
dsDNA	double stranded DNA
EDTA	ethylenediaminetetraacteic acid (disodium salt)
g	gravity
kb	kilobase
M	Molar
min	minute(s)
ml	millilitre
mM	millimolar
nM	nanomolar
mRNA	messenger RNA
pM	picomolar
RNA	ribonucleic acid
RNAase	ribonuclease
sec	second(s)
SDS	sodium dodecyl sulphate
TE	Tris-EDTA buffer
Tween	polyethylenesorbitan monolaurate
µl	microlitre
UV	ultraviolet
X-Gal	5-bromo-4-chloro-3-indolyl-β-galactopyranoside

VI. Contents

1. Introduction

1.i. Cellular memory

One of the main questions in biology is how a single cell, a zygote, develops into a multicellular organism, not only giving rise to specific cell types with distinct functions, but maintaining established gene expression patterns. Developmental patterning of multicellular organisms occurs progressively, based on crude gene expression patterns that are further refined during the early stages of development (Coen 1999). Transcription factors involved in establishing gene expression patterns are often transiently expressed, yet developmental patterns are maintained over many cell generations. Therefore, cell fate must be maintained through cell division. Cell fate may be maintained cell intrinsically (autonomously), or extrinsically, which is flexible (based on neighbouring cell to cell signalling). The mechanisms maintaining cellular memory convert an extrinsically determined cell fate to an intrinsic one.

Regulation of gene expression during development is vital for both ordered cell differentiation and proliferation and maintenance of pluripotent stem cell states. During plant embryogenesis, distinct gene expression patterns are established and cell differentiation occurs forming distinctive tissue types and organs; roots, cotyledons and shoot and root apical meristems. Patterning of reproductive organs; sepals, petals, stamens and carpels occurs post-embryonically. Plant cell fate, in contrast to animal cell fate, is largely extrinsically maintained via neighbouring cell to cell signalling. Cell to cell signalling is essential for coordinated development in both plants and animals. In *Drosophila*, early patterning events are established by the diffusion gradient of a morphogen. During embryogenesis the *Bicoid* protein gradient is established along the anterior-posterior axis and ultimately results in the formation of an anterior-posterior pattern of different combinations of homeotic genes. These gene expression patterns set up distinct parasegment identities in which cell fate is independent of neighbouring segments (Carroll et al. 1988). In plants, most experiments indicate that cell fate is relatively labile and readily

reprogrammed following external signals. *In vivo* laser ablation studies of the root apical meristem demonstrated that despite the stereotypic root cell lineage of *Arabidopsis*, cell fate depends on positional information (van den Berg et al. 1995). Studies using marked clones of cells in tobacco leaf development, illustrated that at fairly late stages of leaf development, epidermal cell fate can be re-specified according to position (Stewart & Burk 1970). In contrast to the poor capacity for de-differentiation of animal cells, plant cells retain a high degree of totipotency and an ability to regenerate from de-differentiated callus cells following the application of the appropriate hormones in tissue-culture (Steeves & Sussex 2002). This argues against the intrinsic maintenance of cell fate in plants. Although, the mechanisms ensuring cellular memory in plants may differ than that in animals, one may still argue that cellular memory in plants is intrinsically maintained. Late germline specification and lack of cell migration in plants (in animals, cell migration promotes the need for autonomous cell fate), highlights the need for plant cell fate to be flexible and easily re-specified following developmental cues. The flexibility of plant cell fate, may, in part, be associated with the mechanisms that maintain cellular memory.

1.ii. Epigenetic Marks

Epigenetics was first used to describe the developmental events leading from fertilization to the mature organism (Waddington 1947). Since then, its definition has evolved and it now encompasses a wide range of phenomena, that are collectively defined as 'mitotically and / or meiotically heritable changes in gene function that cannot be explained by changes in DNA' (Russo, Martienssen, & Riggs 1996). The packaging of DNA and associated proteins (nucleosomes and chromatin) within the nucleus has effects on gene expression. Chromatin was generally considered to be a passive structure, classed as either euchromatic or heterochromatic. Nuclear staining during interphase revealed patterns of DNA that were either in de-condensed regions (euchromatic) or highly condensed regions (heterochromatic). Heterochromatin can be further defined as, constitutive (permanently silenced gene expression states, rarely expressed), or facultative (repressed gene expression in some cells during a specific developmental stage). It is now evident that

chromatin is dynamic and pliable, being continually remodelled by various protein complexes. Chromatin remodelling has been defined as 'any event that alters the nuclease sensitivity of a region of chromatin' (Aalfs & Kingston 2000). Chromatin modifications and remodelling maintain gene expression patterns and global gene silencing (Li 2002). Unlike changes in the DNA sequence, chromatin modifications are readily reversible but they can be persistent through mitosis and even meiosis. Interactions between DNA and chromatin modifiers contribute to nuclear architecture and gene expression. However, how chromatin architecture and its biological properties are maintained over many cell generations is not fully understood. Chromatin remodelling has been characterised as epigenetic due to the mitotic heritability and reversibility of such modifications which are not encoded in the DNA.

Epigenetic phenomena include; DNA methylation, covalent histone modifications and RNA interference. Epigenetic modifications regulate gene expression through blocking the accessibility of the gene transcription domain. Epigenetic modifications are developmentally advantageous as they do not result from changes in the DNA sequence, yet are stable throughout organogenesis and reversible during the production of gametes. Therefore, changes in gene expression that occur during somatic development can subsequently be erased before the start of the following generation. Gene expression patterns must not only be inherited through cell division, but the plasticity of plant development requires that the cell fate is always open to change. The stability of the genetic code limits such flexibility, but epigenetic maintenance of gene expression patterns via chromatin state is both stable and relatively flexible.

1.iii. Genetic identification of Pc-G genes in *Drosophila*

Polycomb-group (Pc-G) and trithorax-group (trxG) proteins mediate covalent histone modifications, promoting chromatin remodelling and regulating gene repression and activation, respectively. In plants and animals, Pc-G and trxG genes are important regulators of homeotic gene expression. Homeotic genes encode a set of evolutionary conserved proteins that specify the anterior-posterior axis during early development. In plants, floral

homeotic genes encode MADS-box proteins that specify floral cell identity autonomously (Coen & Meyerowitz 1991). The expression patterns of homeotic genes are established early during development and maintained by Pc-G and trxG proteins.

Pc-G genes were first identified from genetic studies in *D. melanogaster* whilst characterizing the mutants *extra sex combs* (*esc*) and *polycomb* (*pc*) (Lewis 1978). The common phenotype of these mutants is extra sex combs on the legs of males caused by ectopic expression of *SEX COMBS REDUCED (SCR)* in posterior segments. *SCR* as well as other homeotic genes, for example *Ultrabithorax* (*UBX*), of the Bithorax (BX-C) and Antennapedia (ANT-C) complexes (that establish parasegment identity along the anterior-posterior axis) are ectopically expressed in *Pc-G* mutants (Lehmann & Nusslein-Volhard 1987). Transcription factors that establish initial homeotic gene expression patterns during embryogenesis are transiently expressed. The gap gene *HUNCHBACK (HB)* initially specifies *UBX* expression, yet following the loss of HB after gastrulation, *UBX* expression is maintained. In *enhancer of zeste* (*e(z)*) mutants, regulation of *UBX* expression is not maintained following embryogenesis, becoming ectopically expressed in the anterior body parts (Jones & Gelbart 1993). Genetic and molecular analysis revealed that homeotic gene expression is initially established in *Pc-G* mutants, but expression patterns are not maintained. Thus, Pc-G proteins do not establish homeotic gene expression, but regulate established gene expression patterns.

The *Pc-G ESC*, *E(z)* and *SUPPRESSOR OF ZESTE* (*Su(z)12*) loss of function mutants share common phenotypes associated with the mis-regulation of homeotic genes such as *UBX*, however, further molecular analysis revealed that they have distinctive functions during development. *ESC* is expressed during embryogenesis and regulates spatial expression of BX-C and ANT-C complexes. *esc-* loss of function mutants show severe homeotic transformations, the thoracic and abdominal segments develop into the eight abdominal segment, and are lethal (Struhl & Brower 1982). The maternal *ESC+* contribution during oogenesis rescues *esc-/ esc-* progeny resulting in only mild phenotypes (extra sex combs).

Although *e(z)*- loss of function mutants display similar phenotypes to *esc*- mutants, unlike *ESC*, *E(z)* is required continuously to regulate homeotic gene expression throughout development (Jones & Gelbart 1990). *Su(z)12* suppresses position-effect variegation (PEV) in which chromosomal rearrangements result in an euchromatic gene being silenced due to being positioned near heterochromatin.

Molecular isolation of *ESC*, *Su(z)12* and E*(z)* revealed that they encode diverse array of proteins with distinctive functions. *ESC* encodes a nuclear-localised protein containing seven WD-40 repeat motifs and is homologous to EMBRYONIC ECTODERM DEFICIENT (EED) in mice and FERTILIZATION-INDEPENDENT ENDOSPERM (FIE) in *Arabidopsis* (Ohad et al. 1999). *Su(z)12* encodes a protein containing a VEFS domain and a C_2H_2 zinc finger and is homologous to VERNALIZATION2 (VRN2), EMBRYONIC FLOWER2 (EMF2) and FERTILIZATION-INDEPENDENT SEED2 (FIS2) in *Arabidopsis* (Gendall et al. 2001; Luo et al. 1999; Yoshida et al. 2001). *E(z)* encodes a SET domain-containing protein with adjacent cysteine-rich domains (known as the PRE-SET and POST-SET domains) and is homologous to EZH1 and EZH2 in mice, CURLY LEAF (CLF), SWINGER (SWN) and MEDEA (MEA) in *Arabidopsis* (Goodrich et al. 1997; Chanvivattana et al. 2002; Grossniklaus et al. 1998).

1.iv. Biochemical purification of Pc-G protein complexes

Through biochemical purification, Pc-G proteins were discovered to form distinct multimeric complexes; Polycomb Repressive Complex1 (PRC1) and PRC2 (Figure 1.1). The PRC1 and PRC2 complexes function interdependently, despite their distinct biochemical roles. PRC1 contains Pc, POLYHOMEOTIC (PH), POSTERIOR SEX COMBS (PSC) and dRING1 (Saurin et al. 2001; Shao et al. 1999) and functions in nucleosome remodeling and ubiquitination of H2A resulting in stable long-term gene silencing (Schwartz & Pirrotta 2008). PRC1 inhibits nucleosome remodelling by inhibiting the action of SWI/SNF on nucleosomal arrays *in vitro* (Shao et al. 1999). Pc recruits PRC1 complexes to *UBX* downstream of Polycomb-response elements (PREs) in the promoter region of *UBX* (Francis & Kingston 2001).

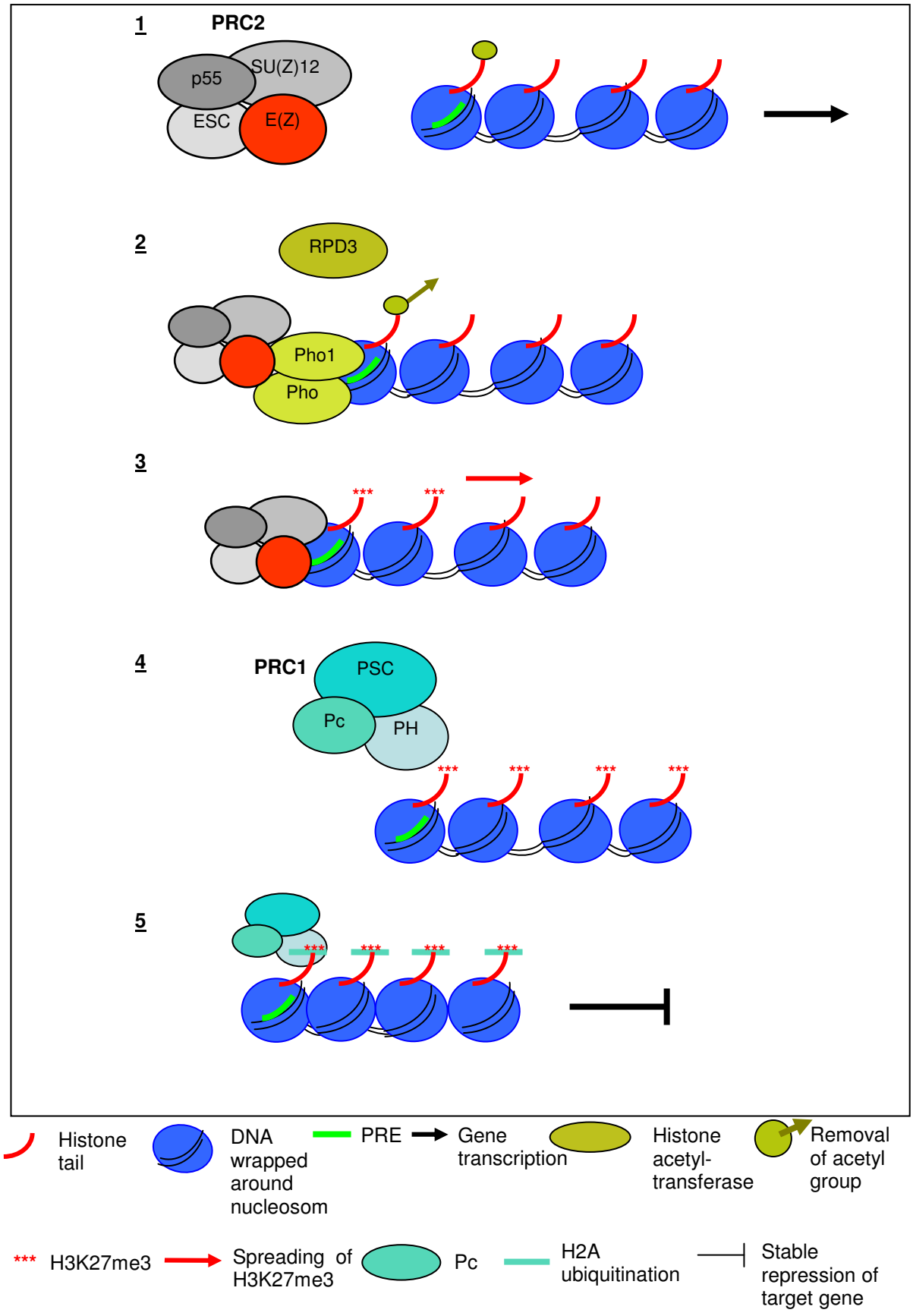

Figure 1.1 Model of *Drosophila* Pc-G mediated silencing

Figure 1.1 Model of *Drosophila* Pc-G mediated silencing

(1) The PRC2 complex contains ENHANCER OF ZESTE (E(z)), SUPPRESSOR OF ZESTE (Su(z)), EXTRA SEX COMBS (ESC) and P55. DNA wrapped around histones as observed under electron microscopy, the 11 nm fibre.

(2) The PRC2 complex is recruited to target genes via PhoRC (Pho and Pho1) that directly binds to DNA sequences present in Polycomb Response Elements (PREs). The removal of the acetyl group is mediated by histone deacetylase RPD3.

(3) The PRC2 complex trimethylates H3K27 and spreading of H3K27me3 occurs.

(4) The PRC1 complex PRC1 contains POLYCOMB (Pc), POLYHOMEOTIC (PH) and POSTERIOR SEX COMBS (PSC) and is recruited to H3K27me3 established by the PRC2 complex.

(5) The PRC1 complex then promotes stable long term gene silencing accompanied by nucleosome remodelling and ubiquitination of H2A.

PREs are 100 to 130 bp *cis*-acting sequence elements that are recognized by the PRC2 complex (reviewed in (Schuettengruber et al. 2007)) (Figure 1.1). The PRC2 complex contains E(Z), SU(Z), ESC and P55 (Ringrose & Paro 2004). The SET-domain containing Pc-G proteins encode HMTases specifically target lysine residues on core histone tails resulting in covalent histone modifications that affect gene activity by remodelling chromatin (discussed further in section 1.iv).

1.v. Histones and histone modifications

Nucleosomes are the fundamental repeating units of chromatin and contain a core particle that has 145 – 147 bp of DNA wrapped around in 1.67 turns of a flat, left-handed superhelix (Figure 1.2). At low ionic concentrations, the characteristic 11 nm 'beads-on-a-string' appearance can be viewed by electron microscopy (Olins & Olins 1974). The core particle contains an octamer of the core histones, consisting of two histone H2A-H2B dimers and a H3-H4 tetramer (Figure 1.2) (Luger et al. 1997). The N-terminal of core histones have lysine rich tails that are subjected to post-translational covalent modifications, including; acetylation, methylation, phosphorylation, ubiquitination, glycosolation and sumoylation. Covalent histone modifications constitute a 'histone code' that is recognized and interpreted by chromatin-associated proteins that establish higher order chromatin structure, regulating gene transcription (Strahl & Allis 2000). Histone tails are well conserved between species, presumably because of their role in chromatin remodelling and higher order structure. Further investigation of the histone code has revealed that the position, nature and order of histone modification are important in defining chromatin state and, hence, gene expression. Euchromatic regions are associated with 'open' chromatin conformation and gene transcription, characterised by active modifications, for example histone H3 lysine 9 acetylation (H3K9ac) and H3 K4 trimethylation (H3K4me3). Heterochromatic regions are associated with repressed gene transcription states, characterised by repressive modifications, loss of acetylation and H3K27me3. Histone deacetylation and methylation have been associated with the establishment and regulation of Pc-G silencing and are the main histone modifications to be discussed in this section.

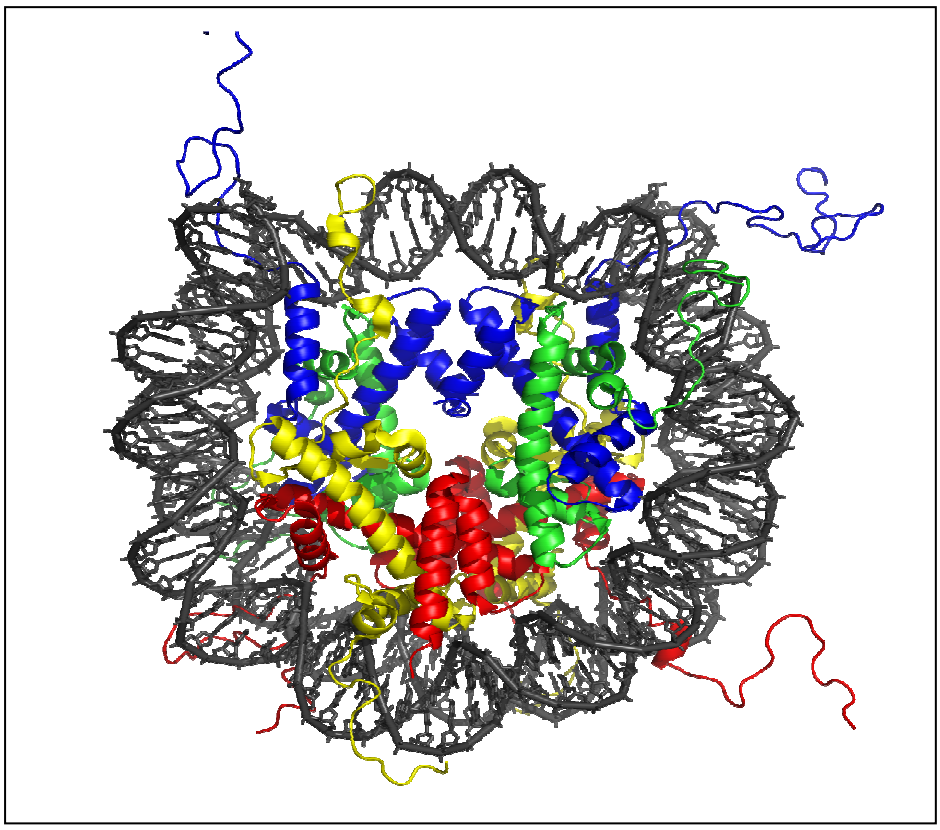

Figure 1.2 Crystal structure of the nucleosome core particle

Nucleosomes, the fundamental repeating units of chromatin, contain a core particle that has 145 to 147 bp of DNA wrapped around in 1.67 turns of a left-handed superhelix. The core particle contains an octamer of the core histones, consisting of two histone H2A (yellow) - H2B (red) dimers and a H3 (blue) - H4 (green) tetramer. From Luger et al., (1997).

Histones in euchromatic regions are hyperacetylated compared to histones in heterochromatic regions. Histone acetylation is less stable than other histone modifications. Acetyl groups can be added to the tails of all the core histones, mediated by histone acetyltransferases (HATs) and are removed by histone deacetylases (HDACs). Deacetylation of H3K9 is necessary for initiation of Pc-G mediated silencing (Figure 1.3). Histone methylation is mediated by histone methyltransferases (HMTases) and lysine residues can be mono- (me1), di- (me2) and tri- (me3) methylated. Histone methylation has been found to be associated with both gene transcription and gene repression (Francis et al. 2001). Trimethylation of H3K4 is associated with the action of trxG proteins, promoting gene transcription, whereas, K3K27me3 is associated with the action of Pc-G proteins, promoting gene silencing. Bernstein et al., (2006) showed H3K4me3 and H3K27me3 was associated with genes that encode developmentally important transcription factors in mouse embryonic stem cells. The enrichment for both H3K4me3 and H3K27me3 was characterised as a 'bivalent' chromatin structure. In mice, bivalent chromatin states are believed to mark silent genes poised for activation. Bivalency might form the basis of the mechanism that allows rapid activation of repressed genes. The involvement of EED was illustrated by Azuara et al., (2006). *eed*- mutant embryonic stem cells express differentiation-specific genes that are otherwise repressed. Boyer et al., (2006) and Lee et al., (2006) identified large numbers of Pc-G binding sites associated with genes that are activated during differentiation. These groups both demonstrated that the loss of bivalent marks correlated with stem cell differentiation. Yet, how H3K27me3 is reset, and how H3K4me3 is established (gene activation) remains unclear. If H3K27me3 is a default mark at most target genes, this raises the questions does activation require the removal of H3K27me3, and is H3K27me3 enrichment sufficient to silence?

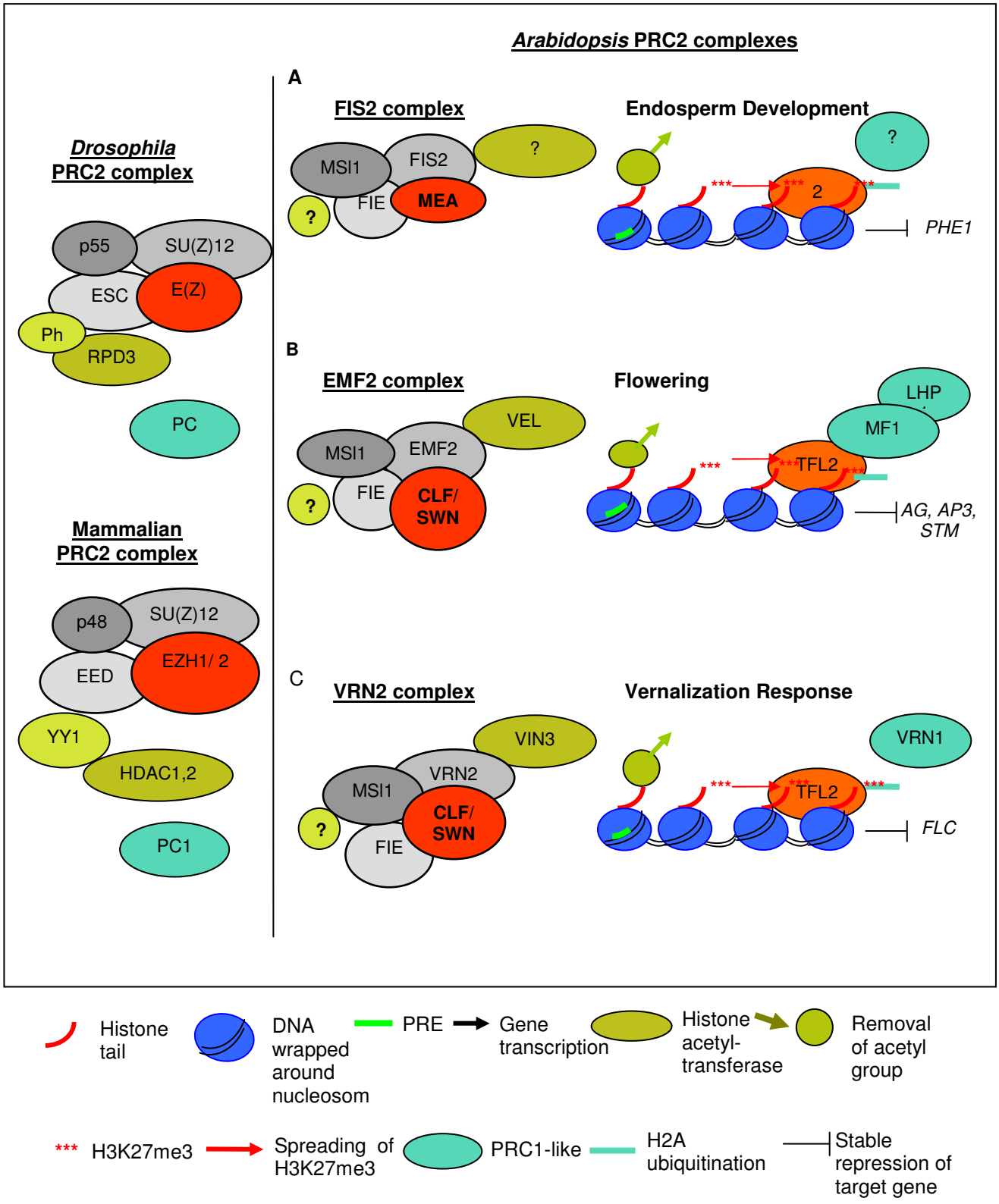

Figure 1.3 PRC2 homologues in *A. thaliana* silence target genes via Pc-G mediated histone methylation

Figure 1.3 PRC2 homologues in *A. thaliana* silence target genes via Pc-G mediated histone methylation

Core proteins of the PRC2 complex are functionally conserved in mammals, flies and plants (colours indicate homology). In *A. thaliana* Pc-G proteins form specific PRC2 complexes that regulate different target genes during different stages of development; (A) the FERTILISATION INDEPENDENT SEED2- (FIS2) complex represses targets, such as *PHE1* during endosperm development. (B) The EMBRYONIC FLOWER2- (EMF2) complex silences floral homeotic genes during vegetative development. (C) The VERNALIZATION2- (VRN2) complex represses *FLC.* This is associated with the vernalization response. The PRC2 is recruited to primary sequences by an unknown signal, associated with the action of histone deacetylases, VERNALIZATION INSENSITIVE3 (VIN3) and VERNALIZATION5 (VEL). Local tri-methylation of histone H3 lysine 27 (H3K27me3) is catalysed by SET-domain proteins CLF/ SWN and MEA. H3K27me3 accumulates at target gene loci and spreads, which is associated with long-term stable silencing. Other proteins are recruited to H3K27me3, potential PRC1-like homologues, that mediate long term silencing. VRN1 may have a more specific function, where as EMF1 appears to be associated with many target genes. The chromodomain protein, LIKE HETEROCHROMATIN PROTEIN 1 (LHP1) is anchored to target loci, via H3K27me3 and could promote long term silencing.

1.vi. Identification of plant Pc-G complexes and roles during development

As previously mentioned, Pc-G proteins regulate gene repression established during embryogenesis. In animals, Pc-G members form two complexes, the PRC1 and the PRC2 complexes. Based on sequence homology, no PRC1 homologue has been identified in plants. Structural analysis revealed that the PRC2 complexes are conserved in plants where, unlike in animals, they form small gene families in which the duplicated members play distinct roles during plant development (Figure 1.3). Evidence for functional conservation of PRC2 in plants is weaker as *in vitro* biochemical histone methyltransferase activity has yet to be demonstrated, however, it is inferred from Pc-G loss of function mutant analysis.

Seed development results from a double fertilisation event leading to the formation of the diploid embryo (from the fertilized egg cell) and the triploid endosperm (from the fertilized central cell). *fis-5* mutants exhibit a maternal-induced seed abortion phenotype and endosperm development in the absence of fertilisation (Figure 1.4) (Chaudhury et al. 1997; Grossniklaus et al. 2001). Therefore, before fertilisation, the egg and central cell divisions are actively repressed by the FIS complex. Identified *FIS* genes include: *MEA* (Grossniklaus et al. 1998), *FERTILISATION INDEPENDENT ENDOSPERM (FIE)* (Ohad et al. 1999), *FERTILISATION INDEPENDENT SEED DEVELOPMENT2 (FIS2)* (Luo et al. 1999) and *MULTI COPY SUPPRESSOR OF IRA1 (MSI1)* (Guitton et al. 2004). Analysis of *FIS* gene expression by in situ hybridisation demonstrated that all *FIS* genes are expressed in the female gametophyte prior to fertilisation and exclusively expressed from the maternal allele in the embryo and endosperm (Grossniklaus et al. 2001). Genomic imprinting is a process in which either the maternal or paternal alleles are silenced in a parent-of-origin fashion. Although imprinting has evolved independently in animals and plants, there appears to be some functional homology. In female mice, the *FIE* homologue, *EED*, controls imprinted inactivation of the paternal X chromosome in the placenta (Wang et al. 2001).

The VRN2 complex is involved in the vernalization response (Bastow et al. 2004; Chandler et al. 2001; Gendall et al. 2001). Vernalization promotes flowering after prolonged periods (4

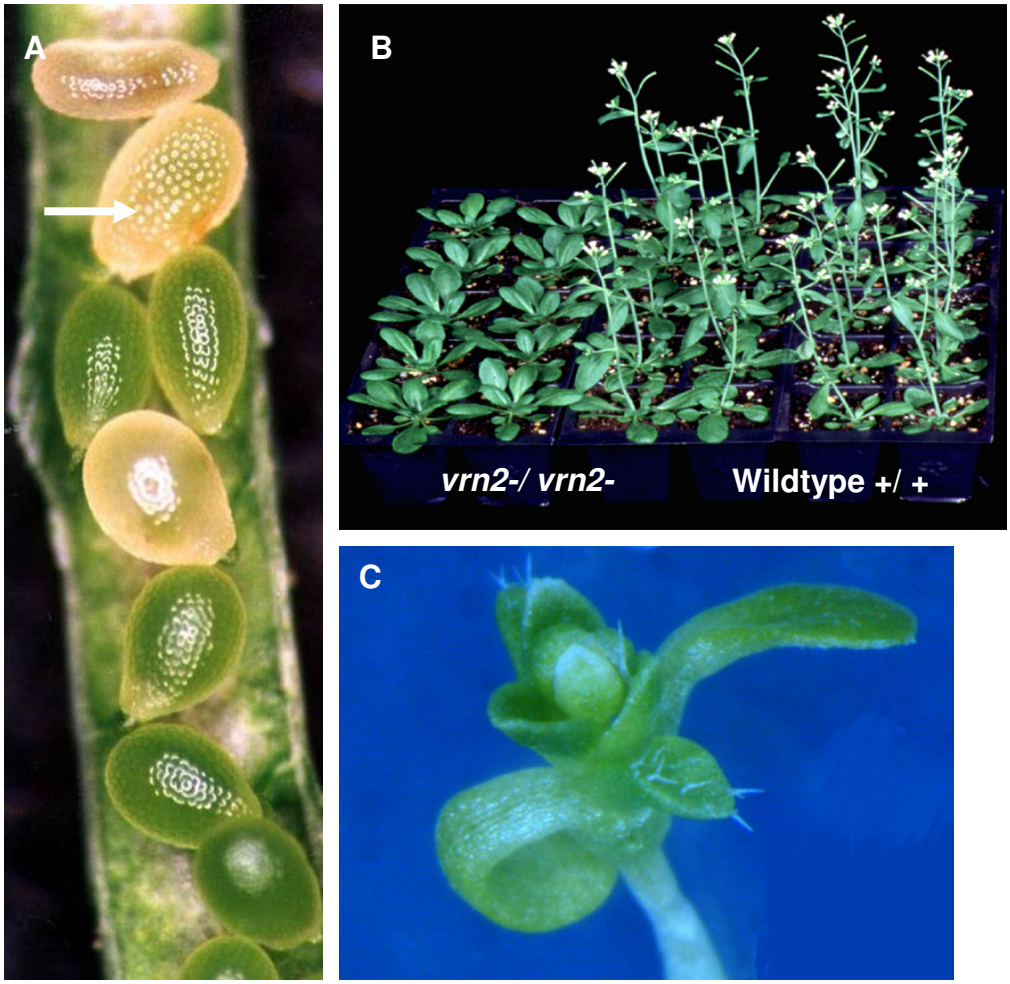

Figure 1.4 Pc-G loss of function mutant phenotypes reveal their discrete roles during *Arabidopsis* development

(A) The silique from a *mea-/+* female and a +/+ male, illustrating the maternal effect of *mea* ovules on seed development, *mea-* loss of function mutants shown by arrow that are embryo lethal. (B) *vrn2-* loss of function mutants do not respond to vernalization, wildtype *A. thaliana* respond to vernalization treatment and flower earlier. (C) *emf2-3* mutant plants are extremely small and produce small sessile leaves and several flowers. From Goodrich & Tweedie (2002).

to 12 weeks) at low temperatures (1 to 7℃). By imposing a requirement for prolonged cold, it prevents seed that germinates in late summer from flowering in autumn or winter (Sung & Amasino 2004). The involvement of Pc-G proteins in vernalization was discovered by genetic screens for *A. thaliana* mutants that did not respond to vernalization (Figure 1.4) (Chandler et al. 1996). The *vernalization2- (vrn2-)* mutant has decreased *FLC* expression following vernalization treatment, however, *FLC* repression is not maintained once plants return to warmer conditions, resulting in late flowering. The VRN2 complex contains VRN2, CLF/ SWN, FIE, MSI1 and FIE. Vernalization is further discussed in section 1.vi.

The EMF2 complex is involved in flowering and floral organ identity (Jack et al. 1994; Weigel et al. 1992). The EMF2 complex contains EMF2, CLF/ SWN, MSI1 and FIE. *emf2-* mutants are very small plants and show extremely early flowering (Figure 1.4), due to the ectopic expression of *AP1* and *AG* (Chen et al. 1997).

Floral organ identity is determined by homeotic genes that are expressed in the floral meristem during flower development. The initial activation of homeotic genes is mediated by *LEAFY*, *AP1* (Weigel & Meyerowitz 1993) *UNUSUAL FLORAL ORGANS* (*UFO*) and *WUSCHEL* (*WUS*). Combinations of homeotic genes from the classes A, B and C contribute to floral organ identity (Coen & Meyerowitz 1991). Floral organs are arranged in whorls around the floral meristem. Class A genes, *AP1* and *AP2* are both involved in floral meristem identity and expressed in sepals and in petals (whorl 1). Class B genes, *AP3*, *PISTILLATA* (*PI*) are expressed with class A in petals (whorl 2) and class C in stamens (whorl 3). Class C gene, *AG* is expressed in stamens and carpels (whorl 4) (Krizek & Fletcher 2005) (Figure 1.5). The ectopic expression of floral homeotic genes in *emf2-* and *clf-* implies that one of the main functions of Pc-G proteins is to repress floral identity during vegetative development. Regulating floral development is an important role for *CLF* and *SWN* as the transition from vegetative to reproductive development is critical for angiosperms to ensure reproductive success.

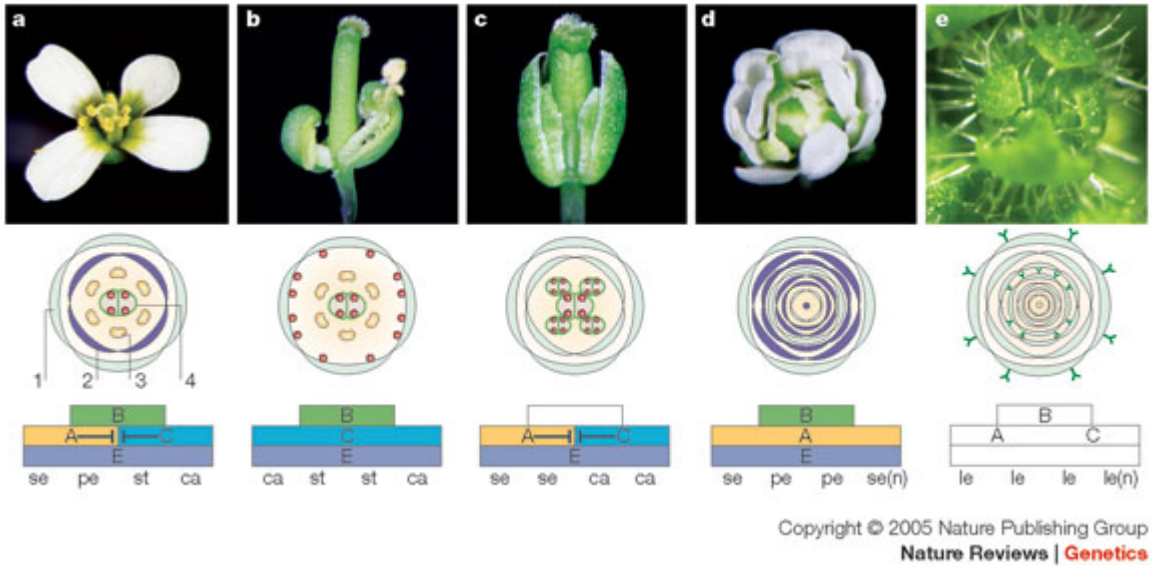

Figure 1.5 Roles of floral homeotic genes in floral identity

(A) Wildtype flower of *A. thaliana* with four sepals, 4 petals, 6 stamens and 2 fused carpels. Combinations of homeotic genes from the classes A, B and C contribute to whorl identity. Class A genes, *AP1* and *AP2* are both involved in floral meristem identity and expressed in sepals (whorl 1) and petals (whorl 2). Class B genes, *AP3* and *PISTILLATA* (*PI*) are expressed with class A in petals and class C in stamens (whorl 3). Class C gene, *AG*, is expressed in stamens and carpels (whorl 4). (B) Class A, *apetala-2*, mutant flowers do not have sepals or petals, but carpels and stamens. (C) Class B, *pistallata*, mutant flowers do not have petals or stamens, but sepals and carpels. (D) Class C, *agamous*, mutant do to not have stamens or carpels, but sepals and petals. (E) *sepallata1- sep2- sep3- sep4-* quadruple mutant lacks class E activity, resulting in the loss of floral determinacy and floral identity. From Krizek and Fletcher 2005.

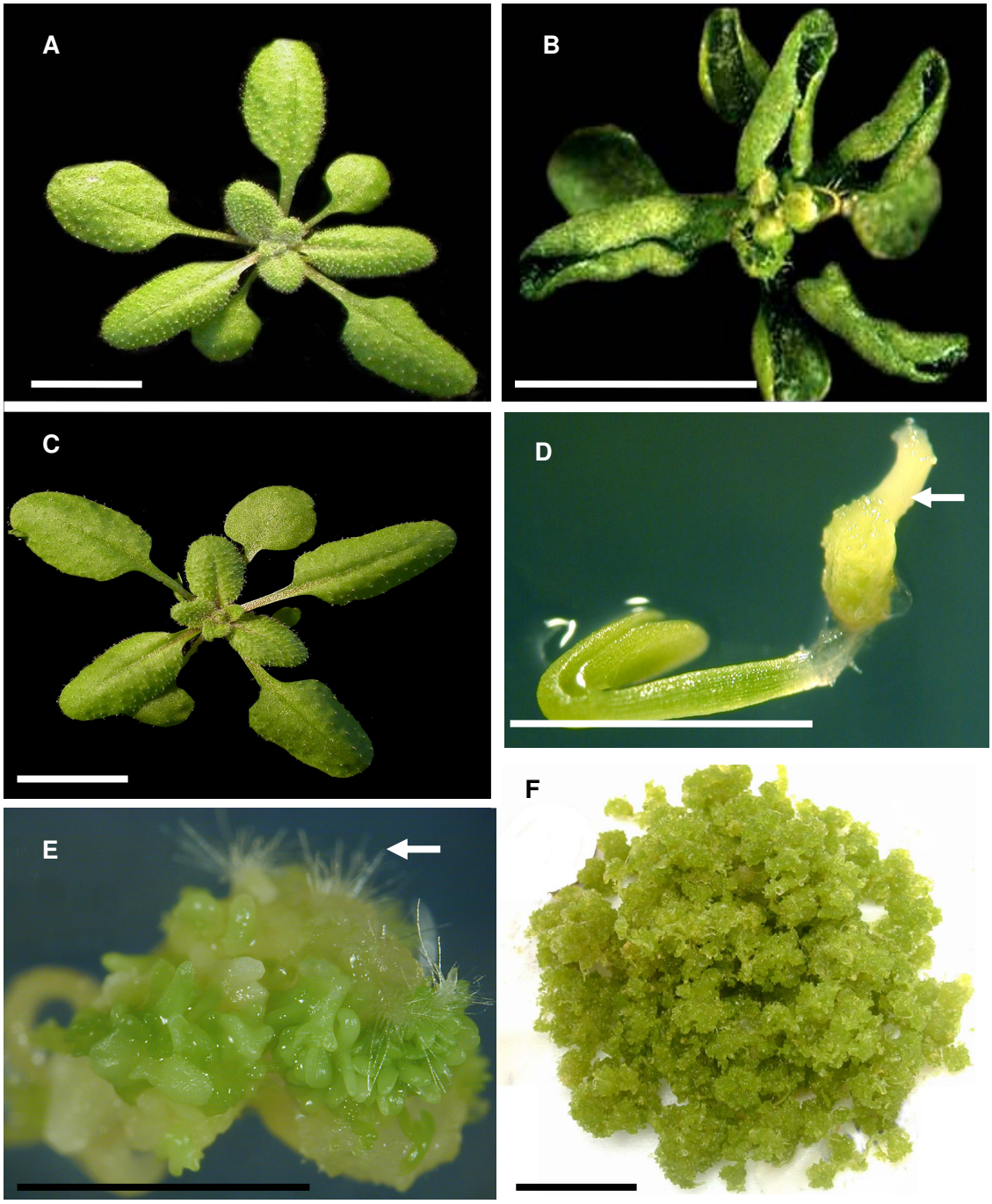

Figure 1.6 *clf-* and *swn-* loss of function mutant phenotypes

Figure 1.6 *clf-* and *swn-* loss of function mutant phenotypes

(A) Compared to wildtype Columbia-0 *A. thaliana,* (B) *clf-2* loss of function mutants are dwarfed with curled leaves. (C) *swn-* loss of function mutants do not show any gross morphological phenotype. *clf-50 swn-3* germinated on sterile tissue culture develop cotyledons, hypocotyl and primary roots. (D) The arrow refers to root to shoot transformation, a characteristic feature of *clf- swn-* loss of function mutants. (E) *clf-50 swn-3* grown on sterile tissue culture for 5 weeks showing root hairs (arrow) and somatic embryos. (F) *clf-50 swn-3* forms a callus like mass that can grow indefinitely on sterile tissue culture. Scale bar 1 cm. From Chanvivattana et al., (2002).

Molecular characterisation of Pc-G mutant phenotypes suggests that floral homeotic gene expression is repressed during embryogenesis and vegetative development by Pc-G proteins FIE, EMF2 and CLF. Pc-G loss of function mutants are early flowering, yet the role of Pc-G proteins in regulating flowering time has not been fully addressed.

clf- mutants are dwarfed with curled leaves (Figure 1.6) and are early flowering with partial homeotic conversion of sepals and petals to stamens and carpels (Goodrich et al. 1997). In *clf-* mutants, ectopic expression of *AG* occurs in leaves and flowers, resulting in curled leaves and partial homeotic transformation (Figure 1.6). *CLF* is partially redundant with a related gene *SWN* (Chanvivattana et al. 2002). The *swn-* mutant displays no gross morphological phenotype (Figure 1.6). Yet, the severity of the *clf- swn-* double mutants implies that the roles of *CLF* and *SWN* during development are greater than originally observed in single mutants (Figure 1.6). *clf- swn-* double mutants develop cotyledons, hypocotyl and primary roots during embryonic development (Figure 1.6) however, post-embryonic development arrests, forming a callus-like mass with no true plant-like structures (Figure 1.6). The mis-expression of *FUSCA3* (*FUS3*) in *clf- swn-* double mutants may underlie the transformation to embryonic callus-like tissue. *clf- swn-* double mutations are lethal on soil, presumably due to defective root development, but can grow indefinitely in sterile tissue-culture (Figure 1.6).

Flowering in *A. thaliana* is promoted by the photoperiod and vernalization pathways in response to long days (LD) and prolonged cold temperature, respectively, and the autonomous and gibberellic acid-dependent pathways that are controlled by internal cues (Figure 1.7 shows the main flowering time pathways in *A. thaliana*). Under LD, *CONSTANS (CO)* is expressed during the light phase, activating *FLOWERING LOCUS T* (*FT*) which is antagonistic to the action of *FLOWERING LOCUS C* (*FLC*) (An et al. 2004). FLC binds FT, repressing activity and delaying flowering. FT interacts with FD and up-regulates *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1 (SOC1/ AGL20)* and *LFY* that

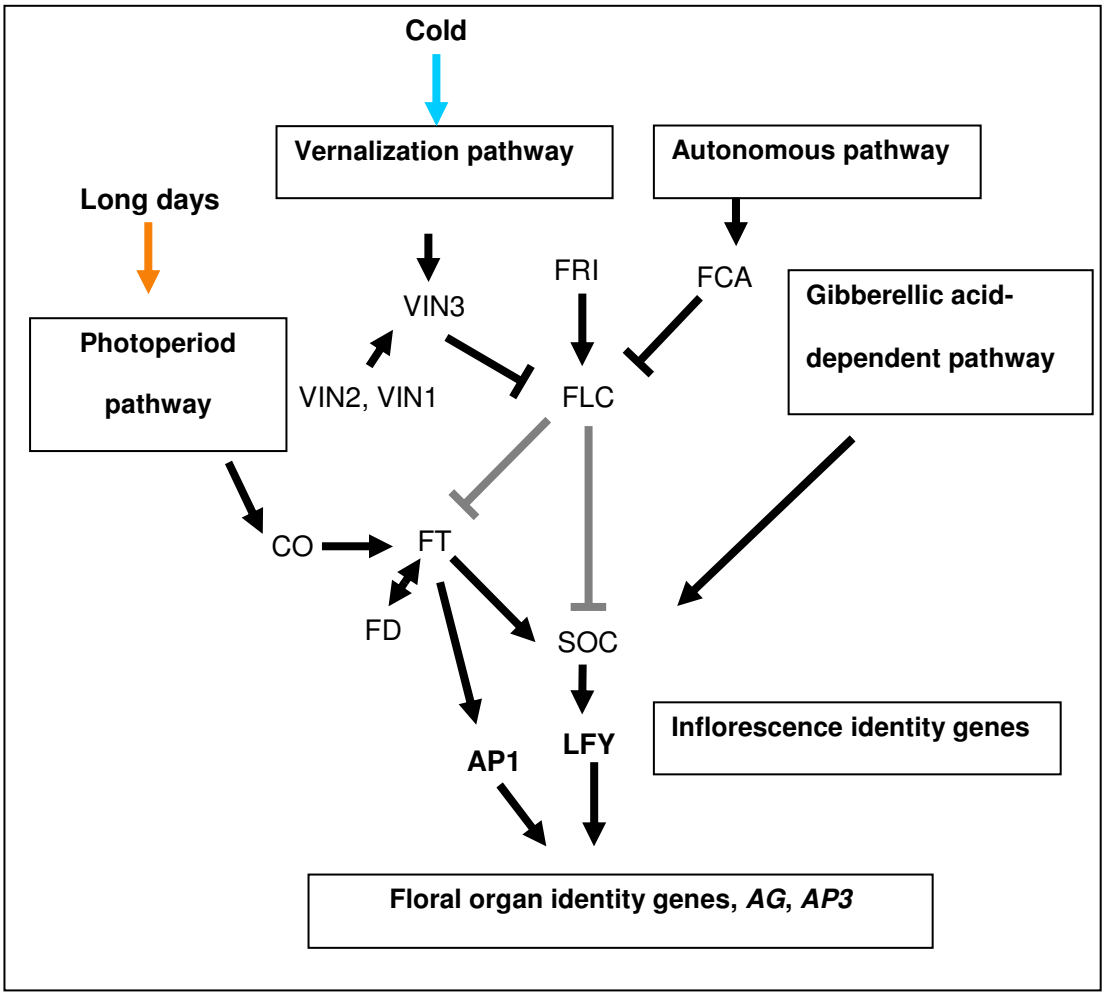

Figure 1.7 Flowering time pathways in *A. thaliana*

The four main flowering time pathways in *A. thaliana* are; the photoperiod pathway, the vernalization pathway, the gibberellic acid-dependent pathway and the autonomous pathway. The gibberellic acid-dependent and autonomous pathways are controlled by internal cues. The photoperiod pathway responds to day length. Under LD, CO accumulates in the light and is able to promote the transcript of FT, promoting flowering. The vernalization pathway responds to periods of prolonged cool temperature that leads to stable silencing of *FLC*.

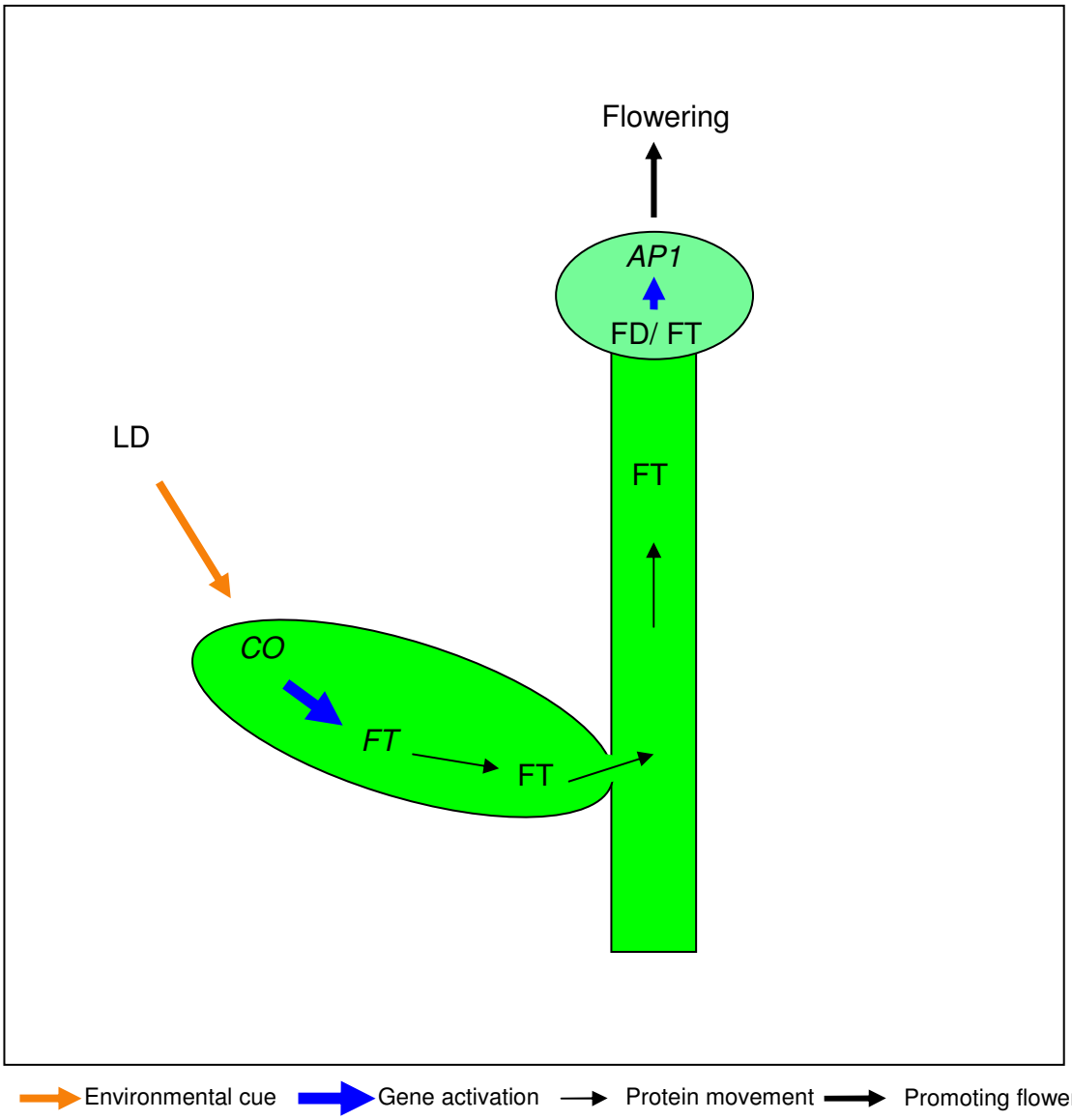

Figure 1.8 FT protein moves from the leaves to the apex and interacts with FD promoting *AP1* expression

In LD, *CO* expression during the light period induces *FT* expression in the leaves. FT then moves from the leaves to the shoot apical meristem where it interacts with FD promoting *AP1* expression, converting the vegetative meristem into an inflorescence meristem, resulting in flowering. From Wigge et al., (2005).

reprogram the shoot apex to form an inflorescence meristem. Flowering is then initiated and the flower primordium expresses *AP1*, *AG* and *AP3* (Figure 1.8). There is considerable variation in *A. thaliana* flowering time; many accessions are late flowering due to active *FRIGIDA* (*FRI*) alleles (Johanson et al. 2000). A latitudinal cline for early flowering alleles due to selection on polymorphisms in the *FRI* locus, causing loss of function *fri*- accessions, has been discovered (Stinchcombe et al. 2004). *FRI* represses flowering by promoting high levels of *FLC* expression. *FLC* encodes a MADS-box class transcription factor that represses flowering during vegetative development. *FLC* expression represses transcription of *FT* and *SOC1* (Michaels & Amasino 2001) that are involved in combining regulatory signals to initiating flowering.

1.vii. Mechanism of Pc-G action

Although the mechanistic role of Pc-G genes has been widely studied independently in both plants and animals, there have been few experiments addressing the question whether Pc-G mediated silencing in plants has similar mitotic stability and autonomy of action compared to animals. Chromatin remodelling by Pc-G complexes is thought to result in chromatin compaction (Schwartz & Pirrotta 2008). In flies, PRC1 is necessary to confer stable long-term silencing by binding H3K27me3 via chromodomain protein, Pc (Figure 1.1). As plants do not appear to have PRC1 homologues, this raises the question whether Pc-G mediated silencing in plants is as stable in animals. Plants may have developed other mechanisms that confer a PRC1-like function, for example, *EMF1* has been suggested to be a functional homologue of *PSC* (Calonje et al. 2008), which in animals is a part of the PRC1 complex. In plants, unlike animals, most Pc-G target genes are enriched for H3K27me3, covering transcribed and promoter regions gene specifically, rather than covering large genomic regions, as is the case in animals (Schubert et al. 2006; Zhang et al. 2007). PRC2 complexes have been isolated in various forward genetic screens (Figure 1.3) showing distinct gene specificity. In flies, *cis*-acting sequences, PREs, recruit the PRC2 complex, ensuring target gene recognition; however, no such elements have been identified in plants. This may not be all that surprising, when considering that the PREs could be recognized by

all Pc-G complexes. This could lead to competition between Pc-G proteins for binding sites as well as limiting accessibility. There may also be dosage effects mediated by different Pc-G complexes that have higher or lower affinity for the PRE at the target sequence. How Pc-G proteins are initially recruited to target sequences in plants is still unknown. One possible mechanism could be that Pc-G proteins are recruited to target genes via signals provided during patterning events. HMTase activity could induce chromatin remodelling that further attracts Pc-G complexes following cell division, providing faithful propagation and maintenance of established gene silencing. The current view of the mechanism of Pc-G silencing in plants is illustrated in Figure 1.3. Plant Pc-G proteins form distinct PRC2 complexes that are recruited to target genes by an unknown mechanism. VERNALIZATION INSENSITIVE3 (VIN3) and VERNALIZATION5 (VEL) are required for histone deacetylation and might recruit PRC2 complexes by protein interactions (Greb et al. 2007). H3K27me3 is catalyzed by SET-domain containing Pc-G proteins CLF, SWN and MEA, that appears to spread over target genes.

A key feature of epigenetic silencing is that although stable and heritable during cell division, epigenetic modifications are reversible. The definitive example of epigenetic phenomena in plants is the vernalization response (Figure 1.9). The vernalization response demonstrates epigenetic features; the transient signal, cold, is remembered long after the initial exposure has gone and the progeny of vernalized plants require vernalization. The vernalization response functions mainly by reducing *FLC* expression that represses flowering. *FRI* represses flowering by promoting high levels of *FLC* expression and so confers a vernalization requirement. *VRN2* represses the expression of *FLC* following prolonged cold treatment (Gendall et al. 2001) and appears to provide the epigenetic basis of the vernalization response by remodelling chromatin in the first intron of *FLC*, leading to mitotically stable *FLC* silencing.

The vernalization response illustrates the mitotic stability of Pc-G mediated silencing providing cellular memory long after the stimulus has gone, but also retaining the ability to be

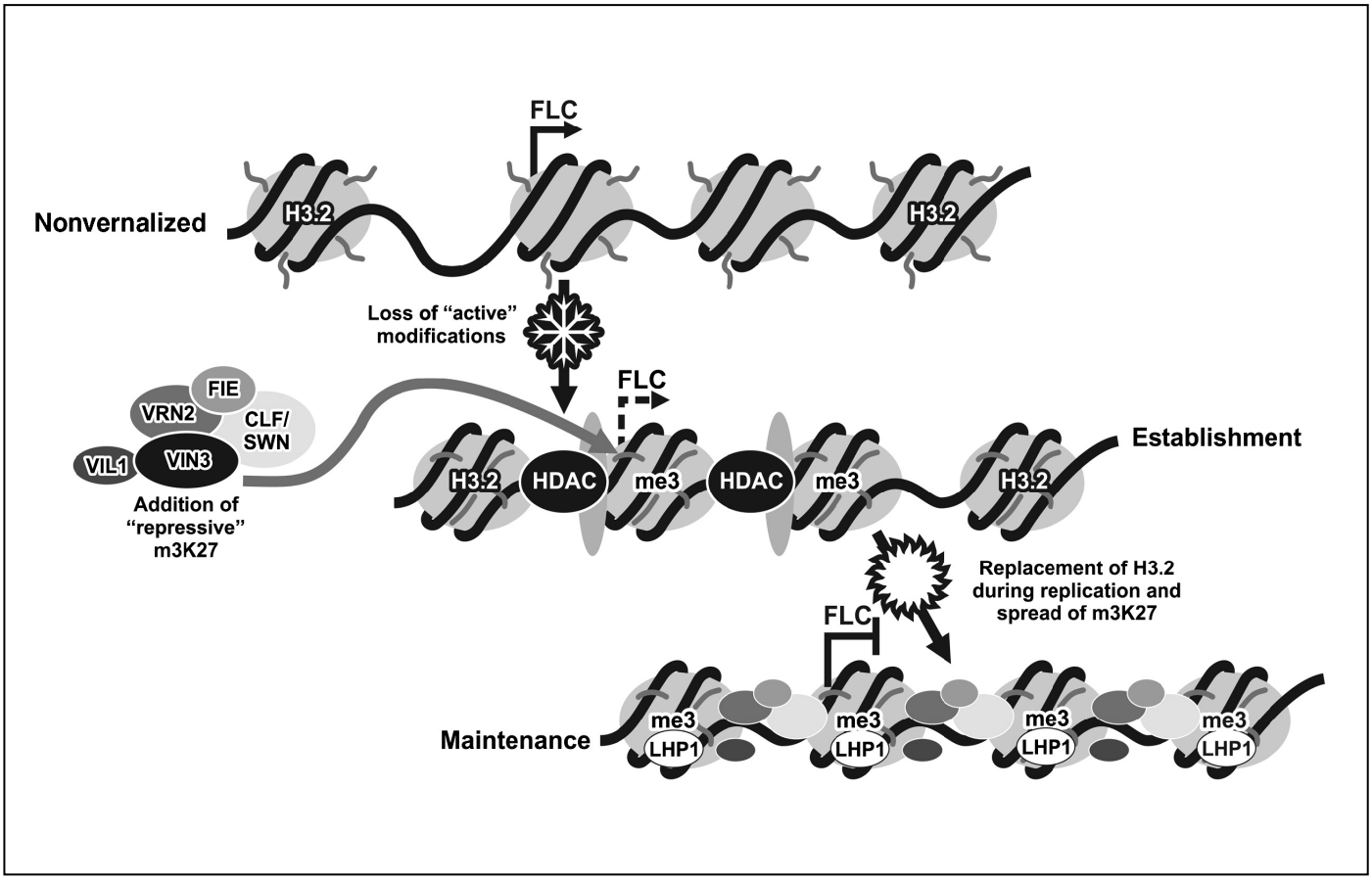

Figure 1.9 Model for the chromatin modifications at *FLC* that occur during vernalization

In nonvernalized plants *FLC* chromatin is marked by histone acetylation, the active histone mark, H3K4me3 and the variant H3.2 is enriched at the *FLC* promoter and coding regions. Upon cold induction, VIN3 represses *FLC* expression via deacetylation of histones and dimethylation of H3K4me3 to H3K4me2. The VRN2 PRC2 complex promotes the repressive mark, H3K27me3. Following vernalization treatment, when conditions become warmer, PRC2 and LHP1 are associated with *FLC* chromatin and during DNA replication, H3.2 is replaced by H3.1, which acquires K27me3. From Finnegan and Dennis (2008).

reset, as the progeny of vernalized plants also require vernalization. Upon vernalization, the initiation of H3K27me3 at *FLC* and its propagation was illustrated by Finnegan *et al* (Finnegan & Dennis 2007). During vernalization, H3K27me3 was found in a small region of *FLC* that then spread over the entire gene once the plant was transferred to warmer conditions. The spreading of H3K27me3 was also reported by Schubert *et al* (Schubert et al. 2006), revealing that H3K27me3 spread over the transgenic sequences (non-plant) of the *AG* reporter construct. This implies that once initiated, H3K27me3 can propagate without Pc-G proteins binding directly to that sequence. *CLF* has been shown to localise specifically at promoter sequences, although the localisation of *CLF* does not always coincide with H3K27me3 enrichment (Clarenz and Goodrich unpublished). This could indicate the HMTase activity of another Pc-G protein, or that *CLF* does not always need to remain at the target genes for H3K27me3 to remain. In *clf-* mutants, silencing of target genes is lost, therefore, *CLF* must be required for regulating target gene silencing. This has been confirmed by studies using *clf-* plants carrying a conditional steroid-dependent *CLF* transgene whereby *CLF+* activity is dependent on steroid application; *clf-* mutants are complemented by steroid application. In the absence of steroid, the *clf-* phenotype is observed with ectopic expression of *AG* (Schubert et al. 2006). Surprisingly, ectopic expression of *AG* in leaves was stable, restoring *CLF* activity only silenced *AG* if the steroid was supplied early in leaf development (Schubert et al. 2006). This suggests that *CLF* activity is required continuously during leaf development to maintain transcriptional silencing of *AG*, even after mitosis is completed. Due to the uneven uptake during seed formation and persistence of steroid in the plant, it is unclear exactly when *CLF* is inactivated. Additionally, the compensatory effects of *SWN* activity may have masked the full effects of the single *clf-* mutant. As *CLF* is lost during mitosis (Schubert et al. 2006), it may not constitute a heritable mark transmitted through mitosis. However, this poses the problem of how *CLF* re-establishes itself to target sequences following mitosis.

During floral transition, the shoot apical meristem undergoes stable changes, converting from a vegetative meristem to an inflorescence meristem. *AG* is stably silenced during

vegetative development and is enriched with H3K27me3 (Schubert et al. 2006). *AG* is activated upon floral transition and enriched with H3K4me3. Saleh *et al.,* (2007) showed that *AG* was enriched during vegetative development for both H3K4me3 and H3K27me3, yet *AG* is stably silenced during vegetative development. This suggests that histone modifications are not mutually exclusive. It is apparent that H3K27me3 is abundant throughout the genome (Zhang et al. 2007). It is not clear how H3K27me3 is propagated through mitosis, whether maintained by the recruitment of Pc-G proteins after cell division. How Pc-G mediated repression is maintained during plant development, two important questions are raised; how stable is Pc-G mediated silencing? And does activation of a Pc-G target gene require the removal of H3K27me3?

1.viii. Resetting Pc-G mediated silencing

A defining feature of epigenetic modifications is that they can be reversed. Presumably, genes regulating embryonic development need to be silenced following germination and the transition from vegetative development to flowering must involve resetting of Pc-G mediated silencing of floral homeotic genes. Reprogramming of gene expression states in animals occurs in gametogenesis or early embryogenesis. In plants, key regulatory developmental genes (such as *FLC*) are active or silenced (such as *AG* and *AP1*) for many cell generations during vegetative development. Following floral induction, Pc-G mediated silencing of *AG* and *AP1* is reset. The vernalization-induced repression of *FLC* is reset between generations. During vegetative development *FLC* is stably expressed and is enriched for H3K4me3. Following vernalization treatment, *FLC* is rapidly silenced, accompanied by the loss of H3K4me3 and enrichment for H3K27me3 and H3K9me2 (Bastow et al. 2004). Finnegan *et al.,* (2007) showed that there was a quantitative relationship between the length of vernalization and the amount of H3K27me3 and H3K4me3 enrichment. This suggests that there may be a correlation between levels of enrichment and gene transcription.

In the progeny of vernalized plants, *FLC* expression is restored. Sheldon *et al.,* (2008) used a reporter transgene for *FLC* to establish when Pc-G silencing of *FLC* was reset during

reproduction. They showed that the paternal and maternal derived *FLC* reporter transgene differed in the timing of their expression. The paternal copy of the *FLC* reporter transgene was active during early gametogenesis and in the single-celled zygote. The maternal copy of the *FLC* reporter transgene was not expressed until the early multicellular embryo stage of seed development (Sheldon et al. 2008).

1.ix. Main aims to be addressed in this thesis

It is apparent that Pc-G mediated silencing is an important mechanism in maintaining cellular memory during development. Microarray analysis reveals that many flowering time genes are up-regulated in *clf-* mutants (Thorpe and Goodrich unpublished). The role of *CLF* in flowering is highlighted by the early flowering phenotype of *clf-* mutants, yet it is difficult to address due to the compensatory effects of *SWN* and the effects of floral homeotic gene mis-expression. Therefore, it is of interest to investigate how genes within the flowering time pathway are regulated by *CLF*. The aim of Chapter 3 is to investigate the role of natural variation in *CLF* in regulating flowering time in natural populations of *A. thaliana*.

Floral homeotic genes are silenced by Pc-G proteins during vegetative development and expression is restricted to specialized cell types upon floral transition. *CLF* regulates the floral homeotic gene, *AG* (Goodrich et al. 1997). *CLF* mediated repression of *AG* in leaves might result in irreversible changes in chromatin structure that render *AG* inaccessible to transcription factors. However, there has been little analysis to address whether Pc-G mediated silencing is reversible. The aim of Chapter 4 is to address the stability of Pc-G mediated silencing of *AG*, by challenging *AG* silencing, to determine whether Pc-G mediating silencing can be reversed.

The repressive histone modification associated with Pc-G silencing, H3K27me3, is abundant throughout the genome (Zhang et al. 2007). Studies suggest that H3K27me3 is dispersed over promoter sequences and spreads over target loci (Schubert et al. 2006; Sheldon et al. 2008; Zhang et al. 2007). How Pc-G mediated silencing is reset, whether by a passive

mechanism or an active removal of modifier followed by histone modifications, is still unknown. The aim of Chapter 5 is to ascertain the histone dynamics at the *AP1* locus to address whether activation of a Pc-G target requires the removal of H3K27me3.

Microarray analysis, ChIP and ChIP on chip (probing tiling arrays with ChIP DNA) revealed that *CLF* and *SWN* have many target genes (Zhang et al. 2007). Further understanding of the biological relevance of these candidate target genes and identifying new targets is difficult due to the downstream effects of target gene mis-expression. The roles of *CLF* in late leaf and flower development remain unknown due to the compensatory effects of the partially redundant Pc-G gene, *SWN*. The severity of *clf- swn-* mutants limits the analysis of the roles Pc-G proteins, as the callus-like tissue does not develop leaves or distinguishable plant-like organs. The main aim of Chapter 6 is to develop two approaches to gain insight into the developmental roles of *CLF* and *SWN* in late leaf and flower development.

2. Materials and Methods

2.i. Plant materials

2.i.i. Plant lines used in this thesis

clf-28 has a T-DNA insertion in the fourth exon of *CLF*. The T-DNA insertion is in a Columbia-0 (wildtype) background and was obtained from the Salk Institute Genome Analysis Laboratory (SIGnAL) SALK_139371. *clf-28* has a mild leaf curling phenotype when grown on soil and confers kanamycin resistance. *clf-2* was derived from transposon mutagenesis, and the mutation is in a Landsberg *erecta* background. The Ds insertion is positioned at the junction between the first and second exon, previously described in (Goodrich et al. 1997). *clf-2* is a null allele and has a strong leaf curling phenotype when grown on soil. *swn-7* has a T-DNA insertion in the eighth exon of *SWN*. s*wn-7* mutation is in a Columbia-0 background and was obtained from SIGnAL, SALK_109121. This is a null allele and has no gross morphological phenotype. The T-DNA insertion confers kanamycin resistance. The Columbia-0 and Edinburgh-0 *A. thaliana* accessions were provided by Prof. Magnus Nordborg (University of Southern California, USA) for flowering time analysis associated with Chapter 3. The *LhG4* driver lines utilized in Chapters 4 and 6 were provided by Prof. Yuval Eshed (Weizmann Institute of Science, Israel). The *35S::LhGR* driver line was provided by Dr. Ian Moore (University of Oxford, UK) and is in a Columbia-0 background. *35S::FD* is an over-expression line of the bZIP transcription factor *FD* that was provided by Dr. Philip Wigge (John Innes Centre, UK). It is a single homozygous line conferring kanamycin resistance. *35S::FD AP1::GUS* confers both kanamycin and BASTA® (phosphinothricin-tripeptide (PTT) herbicide resistance, also provided by Dr. Philip Wigge.

2.i.ii. Sterilizing *A. thaliana* seeds

When germinating seeds in sterile tissue-culture, seeds were either sterilized by ethanol wash or by chlorine gas. In the ethanol treatment, the seeds were placed in microcentrifuge tubes together with 70 % ethanol + 0.05 % triton. They were left rolling for 15 min, the

solution was then replaced with 95 % ethanol. They were then left rolling for 5 min. The seeds were then dispensed onto sterile filter paper to dry and then scattered sparsely onto Murashige and Skoog basal salt mixture (MS) agar plates (1.65 g ammonium nitrate, 6.2 mg borac acid, 332.2 mg calcium chloride anhydrous, 0.025 mg cobalt chloride, 0.025 mg cupric sulfate, 27.8 mg ferrous sulfate, 180.7 mg magnesium sulfate, 0.25 mg sodium salt, 0.83 mg potassium iodide, 1.9 g potassium nitrate, 170.0 mg potassium phosphate monobasic and 8.6 mg zinc sulfate). Larger quantities of seed were sterilized by chlorine gas overnight. In this case, the seeds were placed in an air-tight bell jar (under a fume hood) containing a beaker with 100 ml of sodium hypochlorite mixed with 3 ml HCl. Seeds were sterilized overnight and following sterilization the bell jar was opened and left in the fume hood for a further 20 min in order to remove trace amounts of chlorine gas.

2.i.iii. Growing *A. thaliana* on sterile MS agar plates and on soil

All plants were grown either under long day growth conditions (LD) (16 hrs light and 8 hrs dark) or short day growth conditions (SD) comprising of 8 hrs of light and 16 hrs dark. *A. thaliana* seeds were grown on sterile MS agar plates in order to select for primary transformants or for T-DNA insertion lines, or both. Seeds were dispersed onto agar plates sparsely. Plates were sealed with micropore tape and seeds were stratified at 4 ℃ for 2 to 3 days in darkness, in order to overcome dormancy and synchronise germination. The plates contained the appropriate antibiotic concentration for selection and were grown under LD or SD at 23℃. Hygromycin sensitive plants did not grow roots, and so resistant plants were detected by root growth. Kanamycin sensitive plants had white cotyledons and resistant plants had green cotyledons, allowing detection of the plants carrying the binary vector. Seedlings were transferred from plates onto soil after 4 to 7 days. Seeds and seedlings were grown on soil made up of a 2: 1: 1 ratio of Levington F2 compost, sand and fine grit.

2.i.iv. Gene transfer into *A. thaliana* by floral dip transformation

To introduce plasmid DNA into *A. thaliana* by floral dip transformation, 60 to 100 plants were grown in SD and the first inflorescence bolts were cut. Plants were then transferred to LD and when many inflorescence shoots had emerged plasmids were introduced into the plants

by floral dip transformation. 1 L cultures of recombinant *A. tumefaciens* were grown to stationary phase. The cultures were then centrifuged at 3500 x g for 20 min and the supernatant discarded. The pellets were resuspended in 500 ml 0.5 x MS salts with 0.2 % silwet and 5 % sucrose. Inflorescences were submerged into the *A. tumefaciens* suspension culture for 10 sec ensuring that there was no air bubbles trapped between the flowers. The trays were covered overnight with plastic bags and taken off the following morning. The floral dip was repeated after 7 days. Following 2 to 3 weeks, siliques dried and seeds were collected in bulk and the T1 (first progeny transformed) were selected using the relevant antibiotic or herbicide resistance marker gene conferred by the binary vector.

2.i.v. Crossing of *A. thaliana*

Young flowers of female plants were emasculated, removing the sepals, petals and anthers, just before the anthers had matured. Mature stamens from the paternal plant were brushed against the carpels of the female plants until the stigmatic papillae were yellow with pollen. This was repeated the following day. The siliques were left to grow and then dry, and the F1 seeds from the crosses were collected.

2.i.vi. Small scale plant DNA extraction for PCR

Small leaves, flowers and young seedlings were collected into microcentrifuge tubes and immediately frozen in liquid nitrogen. The plant material was ground into a powder and 200 µl of Extraction buffer (100 mM Tris.HCl pH 8.0, 1.4 M NaCl, 20 mM EDTA, 2 % CTAB, 0.2 % β-mercaptothanol) was added to the samples, which were then incubated at 65 °C for 10 min. Chloroform extraction and ethanol precipitation followed. The DNA was dissolved in Tris-hydrochloride (10 M Tris.HCl pH 7.4) that buffers the solution and ethylene diamine tetraacetic acid (10 mM EDTA) which chelates the Mg^{2+} ions, therefore, protecting the DNA from degradation (as DNase enzymes require Mg^{2+} as cofactor).

2.i.vii. Large scale plant DNA extraction

Plants were grown under standard LD and 500 mg of young leaf tissue was collected and frozen immediately in liquid nitrogen. Samples were ground using a pestle and mortar (also

frozen with liquid nitrogen) and 500 µl of extraction buffer (50 mM EDTA, 100 mM NaCl, 100 mM Tris.HCl pH 8.0, 1 % SDS) was added. A phenol/chloroform extraction was carried out and the DNA was precipitated with 0.1 volumes of 3 M NaOAc (pH 5.2) and 0.7 volume isopropanol followed by washing in 70 % ethanol. The DNA was dissolved in 50 µl of TE.

2.i.viii. Southern Blot Analysis of genomic DNA

Genomic DNA was extracted from *A. thaliana* as described in 2.ii.viii. 15 µl of DNA (about 2 µg) was digested in a reaction consisting of; 3 µl of enzyme buffer, 1.5 µl of enzyme, 0.25 µl of spermidine and 5.25 H_2O. The digestions were incubated overnight at 37 ℃. The samples were loaded onto a 0.7 % agarose gel and run at 0.35 v/ cm for 16 hrs. Digested fragments were separated according to size. The gel was trimmed and immersed twice in 0.25 HCl for 15 min to partially depurinate the DNA (to aid transfer of fragments more than 10 Kb). The gel was then immersed twice in a denaturation solution (1.5 M NaCl, 0.5 M NaOH) for 15 min, and then twice in a neutralization solution (1.5 M NaCl, 0.5 M Tris.HCl pH 7.2, 1 mM EDTA) for 15 min. The DNA was then capillary transferred to a Hybond-N Nylon (Amersham) membrane overnight. The membrane was then rinsed in 2 x SSC and cross-linked to the membrane by UV transilluminator at 40 $mJcm^2$ followed by baking at 68 ℃ for 1 hr. The membrane was then blocked in 70 ml pre-hybridization buffer (5 x SSC, 0.1 % N-lauryl sarcosine, 0.02 % SDS and 1 % w/v milk powder) at 45 ℃ for 2 hrs to limit nonspecific binding.

To hybridize the DNA probe to the Southern blot, the probe was denatured (boiled for 3 min) and hybridized to the membrane in 50 ml of hybridization buffer. The probe was then removed and stored at -20℃ (for re-use) and the membrane was washed twice for 5 min at room temperature in 2 x SSC and 0.1 % SDS and washed twice again in stringent wash made up of 0.5 x SSC and 0.1 % SDS at 68 ℃ for 30 min. To detect the Digoxigenin labeled probe, the membrane was blocked in 50 ml maleic acid buffer (10 % w/v blocking reagent, 100 mM maleic acid, 150 mM NaCl pH 7) for 1 hr at room temperature. This prevented nonspecific antibody binding. The membrane was then incubated with 2 µl of anti-Digoxigenin antibody (Roche) linked to alkaline phosphatase (anti-GIG AP) in 20 ml of

maleic acid blocker for 30 min at room temperature. Following this, the membrane was washed twice with 100 ml 1 x maleic acid buffer (100 mM maleic acid pH 7.5, 150 mM NaCl, 0.3 % tween) for 15 min at room temperature. Then the membrane was equlibriated in detection buffer for 3 min and wetted with chemiluminescent substrate, chloro 3-(40methoxy spiro [1,2-dioxetane-3-2'-tricuclo[3.3.1.1.]-decan]-4-yl) phenyl phosphate (CSPD) (Roche) sealed and exposed to film for 1 – 6 hrs, depending on the signal strength.

2.ii. DNA manipulation and analysis

2.ii.i. DNA sequence analysis

For cloning, alignments, sequence analysis, predicted restriction digestions and constructing plasmid maps I used Vector NTI®. The software was freely available to download from www.invitrogen.com/site/us/en/home/LINNEA-Online-Guides/LINNEA-Communities/Vector-NTI-Community/Sequence-analysis-and-data-management-software-for-PCs.reg.uk.html.

2.ii.ii. Online resources used for working with *A. thaliana*

For nucleotide sequence analysis including BLAST and sequence viewer within the *A. thaliana* genome, I used the *Arabidopsis* Information Resource website found at www.arabidopsis.org. In order to establish how a gene of interest was expressed under abiotic and biotic stresses as well as expression at different developmental stages I used Genevestigator, an online *A. thaliana* microarray database and analysis toolbox found at iii.genevestigator.ethz.ch/at.

2.ii.iii. Preparing *Escherichia coli* competent cells

The *E. coli* strain used was DH5α with the genotype F- endA1 hsdR17 supE44 thi-1 recA1 gyrA (Nal-r) relA1 delta (lacIZYA-argF) U169 deoR (phi80dlac[delta]lacZM15. Competent *E. coli* cells were made by inoculating colonies in 250 ml Super Optimal Broth (SOB medium), shaking at 16 ℃ for approximately 48 hours until the OD_{600} of the cultures reached 0.6. The culture was then placed on ice for 10 min and centrifuged at 5200 x g at 4 ℃ for 10 min. The supernatant was removed and the pellet resuspended in 80 ml ice cold Tris-borate buffer (TB) and incubated on ice for 10 min, followed by centrifugation. The pellet was resuspended

in 20 ml ice cold TB with 7 % DMSO and incubated on ice for 10 min. Cells were then gently aliquoted, placed in liquid nitrogen and stored at -80 ℃.

2.ii.iv. Transforming *E. coli* with plasmid DNA

The bacterial strain DH5α was thawed on ice, using 100 µl per transformation. 1 – 50 ng of plasmid DNA was added. For ligations and mutagenesis reactions larger quantities of DNA was used in transformations. The competent cell and DNA mix was incubated on ice for 30 min, followed by a 45 second heat-shock at 42 ℃ (in a water bath), and then incubated again 2 min on ice. 0.5 ml of Luria-Bertani (LB) media was added to the cells which were then incubated at 37 ℃ for 45 min shaking. At this stage antibiotics were not added to the LB medium to allow the expression of the antibiotic resistance genes. The cells were then centrifuged at 5200 x g for 30 sec and resuspended in 100 µl of LB medium. In the case of re-transforming plasmid, for which a lot of transformants are expected, 10 µl of resuspended cells was diluted in 100 µl of LB medium and plated on LB agar plates with the appropriate antibiotic concentration. For ligation reactions, the whole volume was plated out on LB agar plates. Kanamycin was used at a final concentration of 50 µm/ ml, Ampicillin and Spectinomycin were used at a final concentration of 100 µm/ ml for both LB agar and LB liquid medium. Following overnight incubation, colonies were picked individually and 5 – 100 ml cultures were inoculated (depending on the plasmid preparation required). Cultures were grown overnight (at least 16 hrs) at 37 ℃ shaking.

2.ii.v. Plasmid DNA extraction from *E.coli*

To extract plasmid DNA from *E. coli*, 1.5 ml of culture was centrifuged at 14 000 x g 1 min, the supernatant was removed and the bacterial pellet was resuspended in 150 µl resuspension buffer (5 M Tris.HCl pH 8.0, 10 mM EDTA pH 8.0, 100 µg/ ml RNAase) and vortexed. 150 µl of lysis solution (0.2 NaOH, 1 % SDS) was added and gently mixed by inversion. The lysis solution bursts the cells, allowing the bacterial DNA to be recovered. 200 µl of ice cold neutralization buffer (3 M potassium acetate, 11.5 % glacial acetic acid) was added and gently mixed. The tubes were then incubated on ice for 5 min and the bacterial lysate was centrifuged at 14000 x g for 5 min. The supernatant was removed and transferred

to a new microcentrifuge tube and the plasmid DNA was precipitated by ethanol precipitation. Twice the volume of ethanol was added at room temperature, the solution was vortexed and left for 2 minutes before centrifuging for 5 min in a microcentrifuge at 14 000 x g. The supernatant was removed and 70 % ethanol was added (mixed gently by inversion) and the plasmid DNA was recovered by centrifugation at 14000 x g for 2 min. The supernatant was removed by gentle aspiration and the tube was left open to allow the remaining ethanol to evaporate for about 10 min. The plasmid DNA was dissolved in 50 µl of TE (pH 8.0) containing 20 µg/ ml DNAase free RNAase (pancreatic RNAase) and stored at -20 ℃. For large scale *E. coli* plasmid purifications (50 ml cultures) Qiagen midi plasmid preparation kits were used following the manufacturer's instructions. The handbook can be downloaded from www1.qiagen.com/literature/handbooks/literature.

2.ii.vi. Endonuclease Restriction digestion of DNA

Restriction digestions were performed in 10 µl volumes consisting of 5.5 µl of distilled water, 1 µl of 10 x buffer supplied with the enzyme, 1 µl of BSA (1 mg/ml), 50 ng of DNA (in TE 0.1 – 1 µg of DNA) and 0.5 µl of restriction enzyme (2 – 10 units). Reactions were incubated for 1 – 3 hrs at 37 ℃ (unless otherwise stated by the manufacturer's instructions). For larger quantities of DNA (in the case of preparing DNA for Southern blot analysis), the volumes were increased to 50 µl and the reactions were incubated overnight. Double digests were performed under the guidelines provided by the manufacturer New England Biolabs (NEB) (www.neb.com/nebecomm/DoubleDigestCalculator.asp). If unable to find a compatible buffer for 100 % enzyme activity, digestions were performed sequentially, whereby following the first digestion the enzyme was heat inactivated and the DNA was ethanol precipitated.

2.ii.vii. Agarose gel electrophoresis and gel extraction

To resolve different fragment sizes (ranging from 100 bp – 10 Kb) by gel electrophoresis, the % agarose (w/ v) dissolved in 0.5 % TBE was varied. For larger fragments of less than 10 Kb 1 % final gel volume was used (5 g in 500 ml 0.5 x TBE) and smaller fragments (less than 100 bp) were ran in 2 - 5 % final gel volume. The agarose was dissolved in 0.5 x TBE (4 mM Tris.borate, 1 mM EDTA) buffer and stored at 60 ℃. Ethidium bromide solution (0.5 µg/ ml

final concentration of gel) was added to molten agarose, and it was poured into casts and allowed to solidify (approximately 30 min at room temperature) and then placed into a gel electrophoresis tank with 0.5 x TBE buffer. 10x loading buffer (0.25 % bromophenol blue, 0.25 % xylene cyanol FF, 15 % ficoll type 400) was added to samples to a final concentration of 1 x loading buffer. The voltage applied to the electrophoresis was decided based on the DNA fragments to be resolved. For larger DNA fragments ($\geq$ 3 Kb) 100 V was applied for 1 hr. Lower voltages were applied to fragments $\leq$3 Kb to allow greater separation by gel electrophoresis, 10 V over 3 days. The gel was eventually placed on an ultraviolet (UV) light source, where ethidium bromide (which intercalates into double stranded DNA and fluoresces upon activation with UV light) was visualized. QIAquick Gel Extraction Kit was used to extract DNA after the gel had run for sufficient time to resolve products. The protocol was followed as per manufacturer's instructions.

2.ii.viii. Converting 5' -overhang to a blunt ended terminus

When performing cloning with 2 restriction enzymes, digestion products leave cohesive overhangs that can subsequently annealed and be covalently joined through ligation reactions. However, when unable to use the same enzyme for vector and insert DNA, blunt-end cloning was necessary. 5' overhangs were converted into blunt ended termini by filling with dNTPs using Klenow DNA polymerase (NEB). Following digestion and purification the DNA was dissolved in 20 µl of 1 x Klenow buffer (50 mM Tris.HCl pH 7.2, 10 mM $MgSO_4$, 100 µM DTT, 40 µM of dNTPs (all) and 0.1 mg/ml of BSA). 1 unit of Klenow (DNA polymerase I large-fragment) was added per µg of DNA. The reaction was incubated at room temperature for 10 min, and was terminated by incubating at 75 ℃ for 10 min, denaturing the Klenow DNA polymerase (NEB).

2.ii.ix. Dephosphorylating binary vectors

If vectors are cut with a single restriction enzyme, linearized vector backbones can re-ligate without incorporating insert DNA. In order to prevent re-ligation, 5' phosphate groups were removed by alkaline phosphatase treatment. Following restriction digestion, the linearized vector DNA was purified by ethanol precipitation and resuspended in 50 µl of 1 x calf

intestinal alkaline phosphatase (CIAP), buffer (50 mM Tri-HCL pH 9.3, 1 mM $MgCl_2$, 100 µM $ZnCl_2$, 1mM spermadine). 0.01 units of CIAP enzyme per pmol of ends to be dephosphorylated was added, and the reaction was incubated at 37 ℃ for 30 min. A further 0.01 units of CIAP/ pmol of ends was added and incubated for 30 min. The reaction was stopped by the addition of 5 µl of 0.5 M EGTA and incubating at 65 ℃ for 20 min. To calculate the pmol of 5' ends of linear double stranded DNA, the amount µg of DNA is divided by the length of DNA in Kb and multiplied by 3.04.

2.ii.x. Ligation Reactions

The concentrations of vector and insert DNA were quantified using a Nanodrop (Nanodrop1000 Thermo) to measure optical absorbance (OD_{260}) of small (1 to 2 µl) volumes of DNA. For larger vectors ($\geq$ 3 Kb) and small insert DNA fragments ($\leq$ 1 Kb) the molar ratio of vector to insert DNA was 1:3. For smaller vectors ($\leq$ 3 Kb) and larger insert DNA fragments ($\geq$ 6 Kb) the molar ratio of vector to insert DNA was 2:1. In cases where vector and insert DNA were of similar sizes the molar ratio used was 1:2. Ligations were performed in 5 to 20 µl reactions (made up to with distilled water) depending on the size and therefore concentration of total DNA. With larger insert DNA fragments, the insert DNA was first incubated at 30 ℃ and then made up to 20 µl ligation reactions and incubated at room temperature for 10 min. With smaller insert DNA, ligations were made up to 10 µl ligation reaction incubated at 16 ℃ overnight. Reactions consisted of 2 µl of 10 x ligation buffer, 1 unit of T4 DNA ligase and 6- 16 µl H_2O.

2.ii.xi. Polymerase Chain Reaction (PCR)

PCRs were carried out for (i) Amplifying from genomic DNA, (ii) Genotyping T-DNA insertion lines, (iii) Genotyping background colonies and (iv) Labeling DNA probes with Digoxigenin (Roche). The following recipe was used in typical PCR reactions made up on ice; 1 µl DNA (50 ng), 2 µl 10 x polymerase buffer, 1 µl forward oligonucleotide primer, 1 µl reverse oligonucleotide primer, 0.4 µl dNTP mix (10 mM), 0.2 µl Taq polymerase (5 u/ µl) and 14.4 µl H_2O. Oligonucleotide primers were designed using Primer3 (fokker.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) and Perlprimer was used to design primers spanning exon-

intron junctions in mRNA sequence for quantitative PCR, downloaded from perlprimer.sourceforge.net.

(i) To amplify from *A. thaliana* genomic DNA the PCR and cycling conditions were:

1. Initial denaturation 98℃ for 30 sec
2. Denaturation 98℃ for 10 sec
3. Annealing 60 ℃ for 30 sec (45 ℃ - 65 ℃)
4. Extension 72 ℃ for 4 min 30 sec (15 – 30 sec / Kb)

Repeated 2 – 4 x 19 cycles

5. Final extension 72 ℃ for 10 min
6. 4 ℃

The reaction conditions for amplifying genomic DNA were: 100 ng DNA, 1 µl of both forward and reverse oligonucleotide primers, 0.5 µl dNTP mix, 4 µl of 5 x Phusion HF buffer (NEB) and 0.3 µl Phusion high-fidelity DNA polymerase (NEB). Reactions were made up to 20 µl with H_2O. Reactions were performed following manufacturer's instructions.

(ii) The cycling conditions used for PCR based genotyping for T-DNA lines were:

1. Initial denaturation 95℃ for 5 min
2. Denaturation 95℃ for 20 sec
3. Annealing 57 ℃ for 20 sec
4. Extension 72 ℃ for 80 sec

Repeated 2 – 4 x 34 cycles

5. Final extension 72 ℃ for 10 min
6. 4 ℃

(iii) To quickly establish successful *E. coli* transformation, PCR was carried out to genotype background colonies. Single colonies (grown overnight on LB agar with the appropriate antibiotic) were picked and inoculated into 15 µl H_2O and then onto an agar plate. The inoculated tube was boiled for 5 min, followed by vortexing and centrifugation for 7 sec at

5200 x g. 2- 3 µl of the boiled colony preparation was used for template in the following PCR reaction; 2.5 µl 10 x buffer, 0.5 dNTPs (mix), 1 µl of both forward and reverse oligonucleotide primers and 0.5 µl Taq polymerase, made up to 15 µl H_2O. Cycling conditions were:

1. Initial denaturation 94℃ for 2 min
2. Denaturation 94℃ for 30 sec
3. Annealing 60 ℃ for 30 sec
4. Extension 72 ℃ for 90 sec

Repeated 2 – 4 x 34 cycles

5. Final extension 72 ℃ for 10 min
6. 4 ℃

(iv) DNA probes for Southern blots were labeled with Digoxigenin (Roche). 50 pg of template plasmid DNA was amplified in a 50 µl reaction made up of 5 µl PCR buffer (with $MgCl_2$), 5 µl PCR Digoxigenin (Roche) probe synthesis mix, 5 µl of both forward and reverse oligonucleotides, 0.75 µl enzyme mix and H_2O. Cycling conditions were:

1. Initial denaturation 95℃ for 2 min
2. Denaturation 95℃ for 30 sec
3. Annealing 60 ℃ for 30 sec
4. Extension 72 ℃ for 40 sec

Repeated 2 – 4 x 10 cycles

5. Denaturation 95℃ for 30 sec
6. Annealing 60 ℃ for 30 sec
7. Extension 72 ℃ for 40 sec (+ 20 sec for fragments $\geq$ 3 Kb)

Repeated 5 – 7 x 20 cycles

8. Final extension 72 ℃ for 10 min
9. 4 ℃

Both Digoxigenin PCR and non-Digoxigenin PCRs were loaded onto an agarose gel, to determine if the Digoxigenin had been incorporated, as this fragment migrated more slowly

compared to the non-incorporated product. All oligonucleotide primers used in this thesis are listed in table 2.1.

2.ii.xii. Site-directed mutagenesis by PCR

To introduce new polymorphisms into known DNA sequences, two complementary oligonucleotides were designed containing the desired mutation flanked by unmodified nucleotide sequence. These primers were designed using PrimerX available from www.bioinformatics.org/primerx/cgi-bin/DNA_1.cgi. Site-directed mutagenesis reactions were performed using *Pfu Turbo* polymerase (Stratagene). Reactions consisted of 5 µl 10 x *Pfu Turbo* buffer (Stratagene), 10 ng double stranded DNA template, 1.25 µl of each oligonucleotide (at 100 ng / µl), 1 µl of 10 mM dNTP mix and 1 µl of *Pfu Turbo* polymerase (Stratagene). Reactions were made up to 50 µl dH_2O. The PCR cycling conditions used for site-directed mutagenesis were:

1. Initial denaturation 95°C for 30 sec
2. Denaturation 95°C for 30 sec
3. Annealing 65 °C for 1 min
4. Extension 68 °C for 15 min

Repeated 2 – 4 x 34 cycles

5. Final extension 68 °C for 5 min
6. 4 °C

After the program finished, 1 µl of *Dpn* I restriction enzyme was added directly to the reaction followed by 3 hr incubation at 37 °C. *Dpn* I recognizes 6-methyladenosine with the recognition sequence GATC (Gm6ATCb). Therefore, the original template that does not contain the mutation is digested by *Dpn*I because the original sequence was derived from dam^+ bacteria that have Gm6ATC, whereas the newly synthesized product has un-methylated GATC. Competent *E. coli* were then transformed with the digested PCR product as described in section 2.ii.iii.

Oligonucleotide	Forward	Reverse
CLF_gen	TCCACCAATATTCTGTGAATGA	AGAATTTGAAAATGAGACGGCTAC
Col-0tgtg	GAAGGTAAAACTGGCACTGGCACTT CATCAGATGGTGCAGG	CCTGCACCATCTGATGAAGTGCCA GTGCCAGTTTTACCTTC
AG_PST1	TCTGGCAAAATAATGTAAAATG	ACAGAGTGAAAAAGAGAGGGAGA
LhG4	ATGGCGAATGGGAAATTGTA	CCGAAAAGTGCCACCTGAC
AP1_H3K4	TCATGGACCCGACATTAGTACG	CTTGTTCTCTATCCTCTTCAATT
AP1_H3K27	GGGATCAAAACCCTAGTTCAAA	CCCATAAGGGAAAACTCTTCG
AP1_RT-PCR	ACATGTTACCTGTTTAGAAAGC	GGCTAAGATTGAGCTTTTGGA
AG5'	CACAAAAGAAAAGGGAATAGAGCTG	TAAGGACACCCCCAAATTGA
AG3'	TATTACCCGAATCCGATCCA	CTTAATCGGTGCAATGCAAA
Actin	CGTTTCGCTTTCCTTAGTGTTAGCT	TCCAAATTTCCTGAGGTGCTTGTA
GAL4	TTTCACTGGAGTTGTCCCAAT	AACACATGAGCGAAACCCTA
pOperator	GCCAAGTCGGCCTCTAATAC	AGGCGTCTCGCATATCTCAT
LFY_H3K4	TTGCTCGAGTGGTCATTTT	CCCTCTAAACCACCAAGTCG
LFY_H3K27	TTGCTCGAGTGGTCATTTTT	CCCTCTAAACCACCAAGTCG
pOPCLF_gen	GCATATGTCGAGGAATTCGAG	TTGCGAAGATCGATCCACTA
svp-41	GACCCACTAGTTATCAGCTCAG	AAGTTATGCCTCTCTAGGTT
svp-41_WT	GACCCACTAGTTATCAGCTCAG	AAGTTATGCCTCTCTAGGAC
pHSCRE_vect	ATGGGCTTTTGCAAAGAAGA	AATCGCGAACATCTTCAGGT
pCB1_vect	TCTGTGGAGCCCGTTCTTAC	CTATTACGGTGGCCAACAGC
pBnUASPTn	GTGACAGCCCTCCGAAGC	TGCTCAACACATGAGCGAAA
CLFcDNA	CCTCCAACATTTCAGATTTCG	CAATGTGTTCTTCCGTGTGG

Table 2.1 Oligonucleotide sequences

2.ii.xiii. Sequencing DNA

Sequences were generated by the University of Edinburgh Sequencing service. Samples to be sequenced were in 5 µl reactions consisting of 5 µM primer and 300 ng of DNA. Chromatogram files were opened and analyzed with Chromas. This program is freely available to download from www.technelysium.com.au/chromas14x.html.

2.ii.xiv. Preparing *Agrobacterium tumefaciens* competent cells

The *A. tumefaciens* strain GV3101 conferring Rifampicin (100 µg/ ml) carries pMP90 Ti-plasmid. To prepare competent *A. tumefaciens* 2 ml of Bacto Yeast extract (YEP) liquid media was inoculated with a single *A. tumefaciens* colony and incubated at 28 ℃, shaking for 16 hrs. This was then used to inoculate a 50 ml YEP culture which was grown at 28 ℃ shaking until the culture was at the optimum absorbance (0.5 OD_{260}). The culture was then centrifuged at 3500 x g for 5 min, the supernatant was discarded and the pellet washed in 10 ml 0.15 M NaCl and resuspended in 1 ml of ice cold 20 mM $CaCl_2$. Cells were dispensed into 200 µl aliquots, frozen in liquid nitrogen and stored at -80 ℃.

2.ii.xv. Transforming *A. tumefaciens* with Plasmid DNA

Competent *A. tumefaciens* cells were thawed gently on ice (2- 3 hrs). 1 µg of DNA was added to 200 µl of cells and incubated on ice for 30 min. The cells were cold shocked for 1 min in liquid nitrogen and thawed for 1 min at 37 ℃ in a water bath. 1 ml of YEP medium was added and the cells were then incubated for 2 – 4 hrs at 28 ℃ shaking. The culture was plated out on YEP agar (plus the appropriate antibiotics, concentrations identical to those used for *E. coli*) and incubated for 3 days at 28 ℃.

2.ii.xvi. Plasmid DNA extraction from *A. tumefaciens*

To ensure the transformation had been successful, plasmid DNA was extracted from *A. tumefaciens.* 5 ml YEP liquid medium were inoculated with single colonies of *A. tumefaciens* (with antibiotic selection marker) and grown overnight at 28 ℃, shaking. 1.5 ml of culture was centrifuged at 12000 x g for 3 min and the supernatant discarded. The bacterial pellet

was resuspended in 100 µl of resuspension buffer (5 M Tris.HCl pH 8.0, 10 mM EDTA pH 8.0, 100 µg/ ml RNAase, 10 % lysosome) and incubated at room temperature for 30 min. 200 µl of lysis buffer (0.2 M NaOH, 1 % SDS) was added, and the tube was inverted several times and incubated on ice for 5 min. 150 µl ice-cold neutralization buffer was added (3 M potassium acetate, 11.5 % glacial acetic acid), vortexed for 10 sec and incubated on ice for a further 5 min. Bacterial chromosomal DNA and denatured proteins were precipitated by alkaline lysis. The supernatant was transferred into a fresh tube with 350 µl isopropanol for precipitation, washed in 70 % ethanol and dissolved in 10 µl TE. Due to the low yield and impurity of the DNA extracted from *A. tumefaciens*, the plasmid DNA could not be analyzed directly. 3 µl of plasmid DNA was then reintroduced into *E. coli* by transformation (as described in section 2.ii.iii) to confirm successful transformation of *A. tumefaciens* with plasmid DNA.

2.ii.xvii. Chromatin Immunoprecipitation (ChIP)

To characterize histone modifications at particular loci, ChIP was carried out using the protocol from Prof. R. Martienssen Lab (CSHL, USA) modified by Dr. R. Bastow *et al* (John Innes Centre, UK). Seedlings were harvested and placed onto miracloth, tied, submerged into ice-cold H_2O and rinsed 3 times. The samples were then compressed; removing as much H_2O as possible and 1 % formaldehyde was added to just cover the samples. The samples were lightly weighted down and placed in a vacuum for 15 min at 25 PSI to crosslink. Glycine was added to a final concentration of 0.125 M (2.5 ml of 2 M glycine in 37 ml H_2O). Seedlings were then rinsed 3 times with 40 ml H_2O. 1- 2 g of material was ground in liquid nitrogen to a fine powder and stored on liquid nitrogen, or at -80 ℃.

2.ii.xviii. Antibodies

The antibody used to characterize H3K4me3 enrichment in this ChIP was ordered from Upstate. The antibody used to characterize H3K27me3 enrichment in ChIP was ordered from Abcam.

2.ii.xix. Preparing Beads

The 25 µl of Dynabeads (Roche) per sample were washed 3 times in ChIP dilution buffer (1.1 % Triton X-100, 1.2 mM EDTA, 16.7 M Tris-HCl pH 8, 167 mM NaCl 0.1 mM PEFA Block (Roche) and phenylmethylsulfonyl (PMSF), 1 ml water + 1 complete mini protease inhibitor cocktail tablet (Roche)). 1- 2 µg per bead of antibody was added to the beads, to a total volume of 100 µl and were incubated rotating at 4 ℃ for 3 hrs (including no antibody controls).

2.ii.xx. Preparing Chromatin

30 ml of extraction buffer 1 (0.4 M sucrose, 10 mM Tris-HCl pH 8, 10 mM MgCl, 5 mM BME (2-mercaptoethanol), 0.1 mM of protease inhibitor PEFA block (Roche) and phenylmethylsulfonyl (PMSF), 2 complete protease inhibitor cocktail tablets (Roche), 100 ml +1 mM EDTA/ 100 ml) was added to the samples in a 50 ml falcon tube and left on ice for 5 min. The solution was filtered through miracloth twice and centrifuged for 20 min at 12 000 x g at 4 ℃. The pellet was resuspended in 1 ml of extraction buffer 2 (0.25 M Sucrose, 10 mM Tris-HCl pH 8, 10 mM $MgCl_2$, 1 % Triton X-100 5 mM BME, 0.1 mM PEFA block PMSF 1 ml water + 1 complete mini protease inhibitor cocktail tablet (Roche), 1 mM EDTA), transferred to a 1.5 ml eppendorf and centrifuged at 12000 g for 10 min at 4 ℃. The pellet was gently added to 300 µl of extraction buffer 3 (1.7 M sucrose, 10 mM Tris-HCl pH 8, 0.15 % triton X-100, 2 mM $MgCl_2$, 5 mM BME, 0.1 mM PEFA Block PMSF, 1 ml water + 1 complete mini protease inhibitor cocktail tablet (Roche), 1 mM EDTA) and spun for 1 hr at 16000 g at 4 ℃. The supernatant was removed and resuspended in 200 µl of nuclei lysis buffer (50 mM Tris-HCl ph 8, 10 mM EDTA, 1 % SDS, 1 ml water + 1 complete mini protease inhibitor cocktail tablet (Roche), 0.1 mM PEFA Block PMSF) by pipetting and vortexing. 2 µl were kept back for a pre-sonication sample. The solution (ensuring that there were no air bubbles) was sonicated for 10 sec 4 times and then put on ice for 1 min in between sonication treatment. The solution was centrifuged for 5 min at 4 ℃ to pellet debris. The supernatant was transferred to a new eppendorf tube. 10 µl from each sample was taken for total DNA control and stored at -20 ℃. 1 µl of sonicated chromatin was run on a gel along side 1 µl of pre-

sonicated sample to check the efficiency of sonication. The sonicated samples compared to the unsonicated samples, showing smears the sheered DNA fragments were 100 bp to 1000 bp (Figure 2.1).

The supernatant was transferred to a 10 ml falcon and made up to 3 ml with ChIP dilution buffer (1.1 % triton X-100, 1.2 mM EDTA, 16.7 M Tris-HCl pH 8, 167 mM NaCl 0.1 mM PEFA Block PMSF, 1 ml water + 1 complete mini protease inhibitor cocktail tablet (Roche), to dilute the 1 % SDS to 0.1 % SDS. The antibody coated beads were washed 3 times in ChIP dilution buffer. 750 µl of sonicated chromatin (for 4 samples) or 1 ml (for 3 samples) was added to the antibody coated beads and incubated rotating at 4 ℃ overnight.

Beads were washed 2 times in the following sequence:

1. Low salt wash buffer; 150 mM NaCl, 0.1 % SDS, 1 % TritonX-100, 2 mM EDTA, 20 mM Tris-HCL (Ph 8.1) for 5 min twice.
2. High salt wash buffer; 500 mM NaCl, 0.1 % SDS, 1 % TritonX-100, 2 mM EDTA, 20 mM Tris-HCL (pH 8.1) for 5 min twice.
3. LiCl wash buffer; 0.25 LiCl, 1 % NP40, 1 % sodium deoxycholate, 1 mM EDTA, 10 mM Tris-HCL (pH 8.1) for 5 min twice
4. TE buffer; 10 m/m Tris-HCL (pH 8.1), 1 mM EDTA once.

Immune complexes were eluted by adding 250 µl elution buffer (1 % SDS, 0.1 M $NaHCO_3$) to the pelleted beads and mixed briefly by vortex followed by incubation at 65 ℃ for 15 min with gentle agitation. The supernatant was added to the magnetized beads and transferred to a new eppendorf and elution of the beads was repeated. Eventually, both eluates were combined. 20 µl of 5 M NaCl was added to eluate and reverse crosslinked at 65 ℃ for at least 6 hrs overnight. The total DNA controls were made up to 500 µl with Elution Buffer and 20 µl of 5M NaCl and placed with the samples at 65 ℃ overnight.

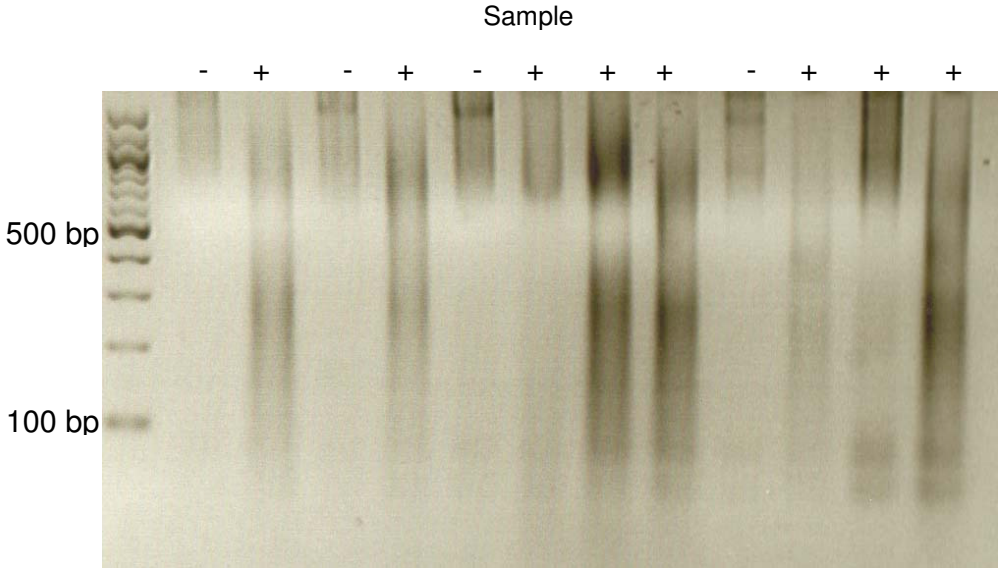

Figure 2.1 DNA is sheared into smaller fragments (≤ 500 bp) following sonication treatment

ChIP samples were sonicated for 10 seconds sonication followed by 1 minute on ice between sonication, repeated 4 times. Sheared DNA fragments were separated by agarose gel electrophoresis before sonication (-) and after sonication (+) treatment. Following sonication treatment, DNA is sheared to smaller fragments between 100 bp 10 1000 bp.

10 µl of 0.5M EDTA, 20 µl Tris-HCl 1M (pH 6.5) and 2 µl of 10 mg/ ml proteinase K were added to the samples followed by 1 hr incubation at 45 ℃ to reverse cross-link DNA to histones and other chromatin components. DNA was recovered by phenol chloroform extraction. An equal volume of phenol/chloroform (550 µl) was added and vortexed, followed by a 5 min centrifugation. The aqueous phase was transferred to a new eppendorf, 2 volumes of ethanol and 2 µl of glycogen were added, and left at -20 ℃ for 2 hrs. The samples were centrifuged at 4 ℃ fro 30 min and pellets were washed with 70 % ethanol and left to air dry. Pellets were resuspended in 50 µl of H_2O.

The input and IPs were diluted 1/ 100, 1/ 400, 1/ 1200 and 1/ 6400 in H_2O and 5 µl per dilution was used in 25 µl PCR reactions. ChIP DNA was stored at -80 ℃ and dilutions for PCR reactions were stored at -20 ℃. The PCR cycling conditions used for ChIP were:

1. Initial denaturation 95℃ for 2 min
2. Denaturation 95℃ for 1 min
3. Annealing 60 ℃ for 1 min
4. Extension 72 ℃ for 20 sec

Repeated 2 – 4 x 34 cycles

5. Final extension 72 ℃ for 5 min
6. 4 ℃

2.iii. RNA and Gene expression analysis

2.iii.i. Histochemical staining of plants carrying GUS reporter genes

β-glucuronidase activity in transgenic plants carrying GUS reporter genes was analyzed by histochemical analysis. Plant tissue was incubated in 100 m/M $NaPO_4$ pH7.2, 2.5 mM 5-bromo-4-chloro-3-indolyl-β-D-glucuronide (X-Gluc, Direct), 0.5 mM $K_3Fe(CN)_6$, 0.5 mM $K_4Fe(CN)_6$, 10 mM Na_2EDTA, 1 % Triton X-100 and H_2O. Plant tissue was vacuum-infiltrated followed by incubation at 37 ℃ overnight. After staining, chlorophyll was cleared from samples by dehydration in an ethanol series.

For precise tissue and cell-specific localization of β-glucuronidase activity, plant tissue was mildly fixed in 0.05 M $NaHPO_4$ pH 7.2, 0.3 % formaldehyde, 0.5 mM $K_3Fe(CN)_6$, 0.5 mM $K_4Fe(CN)_6$, 1 mM EDTA, 1 % Triton X-100 and H_2O for 60 min (a light vacuum was applied). Samples were rinsed 3 times in 0.05 M $NaHPO_4$ pH 7.2, 0.5 mM $K_3Fe(CN)_6$, 0.5 mM $K_4Fe(CN)_6$, 1 mM EDTA and H_2O. Following the final rinse, the solution was removed by aspiration and the samples were stained with 0.05 M $NaHPO_4$ pH 7.2, 0.5 mM $K_3Fe(CN)_6$, 0.5 mM $K_4Fe(CN)_6$, 1 mM EDTA, 1 % Triton X-100, 2 mM X-Gluc and H_2O. The plant tissue was just submerged and incubated at various times at 37 ℃ (generally 10 hrs to overnight). Tissue was then immediately fixed in 5 % formaldehyde in PBS overnight. The samples were rinsed 3 times with PBS and chlorophyll was removed by ethanol series. Samples were mounted in 20 % ethanol and 20 % glycerol and photographs were taken either using a Nikon dissecting microscope and a Nikon differential interference contrast (DIC) microscope.

2.iii.ii. RNA analysis

To extract RNA from plants in order to quantify gene activity, 50 mg of plant tissue was collected and immediately frozen in liquid nitrogen. The material was ground and 1 ml Trizol (Invitrogen) was added. The samples were then centrifuged for 2 min at 12000 x g and the supernatant removed to a fresh tube. 250 µl of chloroform were added and samples were vortexed for 15 sec followed by a centrifugation at 12000 x g for 15 min at 4 ℃. The aqueous phase was removed to a fresh tube, 250 µl of isopropanol, 500 µl of 0.8 M Sodium citrate and 1.2 M NaCl were added, and the samples were incubated for 10 min on ice. Samples were centrifuged at 12000 x g for 10 min at 4 ℃. The RNA pellets were ethanol precipitated and resuspended in 50 µl of RNAase free water (left over night at 4 ℃). The RNA was DNase treated as per manufacturer instructions (Promega). The RNA was quantified by Nanodrop at OD_{260} (Nanodrop1000 Thermo) and used for cDNA synthesis at a concentration of 40 µg/ µl.

For cDNA synthesis, 0.5 µg of RNA was reverse transcribed using ImProm-II Reverse Transcription Kit (Promega). 1 µl of random primers and H_2O to 7 µl was heated to 85 ℃ for 2 min and then cooled to 4 ℃. The following was then added to the reaction; 4 µl of ImProm-

II 5 x RT buffer, 4.8 µl $MgCl_2$, 1 µl dNTPs, 0.5 µl RNase Out (Invitrogen), 1 µl ImProm-II RT and 1.7 µl H_2O. The RT conditions were:

1. 25 ℃ 5 min
2. 42 ℃ 45 min
3. 70 ℃ 15 min
4. 4 ℃

Semi-quantitative PCR was performed to quantify gene expression. The following PCR reaction mixture was made up; 5 µl cDNA, 1.25 µl of both forward and reverse oligonucleotides primers, 0.5 µl dNTPs, 2.5 µl $MgCl_2$, 2.5 µl 5x Crimson *Taq* reaction buffer (NEB), 0.3 µl Crimson *Taq* DNA polymerase (NEB) and 11.7 µl H_2O. The semi-quantitative PCR conditions were:

1. Initial denaturation 95℃ for 30 sec
2. Denaturation 95℃ for 20 sec
3. Annealing 58 ℃ for 30 sec
4. Extension 68 ℃ for 10 sec

Repeated 2 – 4 x 19 cycles

5. Final extension 72 ℃ for 5 min
6. 4 ℃

3. Investigating the role of *CURLY LEAF* in regulating flowering time in natural populations of *Arabidopsis thaliana*

3.i. Introduction

A role for *CLF* in flowering is revealed by the early flowering phenotype of *clf*- mutants, yet difficult to address due to the compensatory effects of *SWN* and the down stream consequences of floral homeotic gene mis-expression. Microarray analysis reveals that many flowering time genes are up-regulated in *clf*- mutants (Thorpe and Goodrich unpublished). Flowering time pathways have been widely investigated, but it is not clear how *CLF* fits into the pathways. There are many accessions of *A. thaliana* that occur over wide range of latitudes and environmental conditions across the globe that vary in flowering time. This Chapter describes how I have screened some of the accessions which very in flowering time for correlation between time to flower and natural variation in *CLF*.

Polymorphisms in *CLF* between accessions of *A. thaliana* may be indicative of changes in *CLF* function which have been selected for in different environments. Polymorphisms that are driven by directional selection typically reach a higher frequency within the population much faster than those under genetic drift. This results in a larger region of identity around the polymorphism than expected given the frequency within the population, as there is less time for recombination further breaking up of the ancestral haplotype. Genome-wide association mapping or linkage disequilibrium mapping can be used to identify alleles that may have been under strong directional selection in *A. thaliana* (Nordborg et al. 2002). By using pairwise haplotype scoring (PHS), a system based on the estimation of the shared length around a polymorphism from a pairwise comparison between individual haplotypes, Toomajian *et al* (Toomajian et al. 2006) identified alleles likely to have been under directional selection. A peak was discovered on chromosome 2 at 9.9 Mb, where amongst other genes,

CLF is located (Table 3.1). Furthermore, an in-frame threonine-glycine (ThrGly) insertion in CLF showed characteristics of selection, with a high PHS score of 2.26 (Toomajian et al. 2006). This result appears to be similar to that reported in alleles of *FRI*, in which directional selection on loss of function *fri*- early flowering alleles has occurred over a wide range of geographical locations (Toomajian et al. 2006). The ThrGly insertion in CLF was found in 27 out of the 96 accessions of *A. thaliana* sampled (Table 3.1). The polymorphism occurs over a wide range of geographical locations, although it appears to be centred in North-West Europe. It is common in the UK, France and USA, but not present in Northern Sweden, Central Asia or the Iberian Peninsula.

The aim of this chapter is to investigate the role of *CLF* in regulating flowering time in natural populations of *A. thaliana*. It is important to establish the functional significance of the ThrGly insertion in *CLF* and whether this polymorphism, in particular, effects flowering time. By complementing *clf*- plants with different alleles of *CLF* we can address whether *CLF* plays a role in regulating flowering time.

3.ii. The threonine glycine insertion is in a non-conserved region of CLF

There are distinctive regions that are highly conserved amongst Pc-G proteins; the SET domain, the cysteine rich CXC domain (POST-SET) and the conserved five cysteines (C5) (PRE-SET) (Figure 3.1 A). The ThrGly insertion occurs outside of these conserved regions Figure 3.1), yet this region may be significant. In order to investigate if the region of the insertion was highly conserved amongst other plant *CLF*-like genes, we aligned the *A. thaliana CLF* to *Petunia* and *Antirrhinum CLF*-like genes. If there was sequence homology in this region, it would provide evidence that the ThrGly insertion was functionally significant. Unfortunately, we discovered that the ThrGly insertion is in a non-conserved region of *CLF* with little or no sequence homology found within this region (Figure 3.1 C). Although the ThrGly insertion is in a region that is not well conserved within eudicots, it was discovered to be highly significant in association with a flowering time effect. At first it seems unlikely that the ThrGly insertion in *CLF* would regulate flowering time.

Chr	Position*	Allele	Freq	Sample	Left end	Right end	PHS	Location	Gene ID	Description
1	3607837-75	39 bp	6	96	1042028	5923561	4.76	3'UTR	At1g10840	TIF3H1; eukaryotic translation initiation factor 3 subunit 3
1	13925316	T	6	96	3229453	23821158	5.78	Repetitive element	At1g36770	gypsy-like retrotransposon family
1	21080288	A	74	95	20738912	21141319	0.64	Syn	At1g56290	CwfJ-like family protein
1	21669059	T	61	84	21563945	21669588	0.6	Syn	At1g58350	ZW18; expressed protein
1	28190589	C	71	96	27959121	28408814	0.74	Syn	At1g75080	BZR1; brassinosteroid signalling positive regulator
2	5086023	G	6	93	2171791	7433513	4.63	Repetitive element	At2g12490	copia-like retrotransposon family
2	5250540	A	5	62	3393477	7352838	4.51	Repetitive element	At2g12770	non-LTR retrotransposon family
2	7073005	G	3	95	172277	11850986	6.57	Pseudogene	At2g16310	pseudogene, RER1-related protein
2	9256270-1	2 bp	6	94	6654153	11595138	3.99	Intron	At2g21620	RD2; universal stress protein (USP) family protein
2	9361318	T	10	91	8659687	10050365	4.87	Nonsyn	At2g21950	SKP1 interacting partner 6 (SKIP6)
2	9819642	C	3	75	9550714	11850986	5.95	3'UTR	At2g23050	phototropic-responsive NPH3 family protein
2	**9964707**	**ACTGGC**	**27**	**96**	**9550714**	**10400703**	**2.26**	**In-frame insertion**	**At2g23380**	**CLF; curly leaf protein; leaf morphogenesis**
2	10134400	C	19	96	9550714	11178939	2.95	Nonsyn	At2g23790	expressed protein
2	10566612	C	14	96	9428891	11850986	3.54	Nonsyn	At2g24762	expressed protein
2	11457225	C	12	96	10900127	12187491	3.18	Nonsyn	At2g26850	F-box family protein; similar to SKP1 interacting partner 2
2	11456894-7	4 bp	67	95	11028504	11457499	0.91	Intron	At2g26850	F-box family protein; similar to SKP1 interacting partner 2
2	11564935	A	20	96	10566416	12014448	3.23	Syn	At2g27070	ARR13; two-component responsive regulator family protein
3	11350490	A	3	62	7590422	14673136	5.86	Intergenic	At3g29515	non-LTR retrotransposon family: pseudogene
4	269962	16 bp	9	95	142440	749157	5.67	Frameshift deletion	At4g00650	FRI; FRIGIDA protein
4	428792	T	6	95	268715	2241604	4.93	Intron	At4g00990	transcription factor jumonji (jmjC) domain-containing protein
4	641613	G	12	96	541028	1322536	4.96	Intergenic	At4g01500	DNA-binding protein, putative;similar to ARV1-like family
4	1055208	T	75	95	747636	1055626	1.29	Syn	At4g02400	U3 ribonucleoprotein (Utp) family protein
4	1145098	C	10	94	747636	1322536	3.22	Intron	At4g02600	seven transmembrane MLO family protein
4	2441355	A	11	95	1055093	5078001	5.53	Intergenic	At4g04800	methionine sulfoxide reductase domain-containing protein
4	9578374	C	3	80	9376858	13358833	5.17	Intron	At4g17020	transcription factor-related
4	14732311	G	4	94	12978424	17152239	4.93	Intron	At4g30120	HMA3; ATPase E1-E2 type family protein
4	15727497	C	31	96	15368323	16025877	1.85	Syn	At4g32610	mitochondrial glycoprotein family protein
5	2244619	G	20	95	1841010	3289479	2.34	Nonsyn	At5g07200	YAP169; gibberellin 20-oxidase
5	2296407	A	44	90	2244524	2799420	1.28	Nonsyn	At5g07290	RNA recognition motif (RRM)-containing protein
5	2799315	7 bp	62	96	2591893	2900468	0.79	Intron	At5g08630	DDT domain-containing protein
5	9723136	G	11	90	9722715	9926079	2.96	Intron	At5g27540	GTP-binding protein-related
5	10909786	C	5	71	7004132	15768566	5.23	Pseudogene	At5g28892	pseudogene, similar to putative helicase

Table 3.1 Whole genome-wide association mapping identified candidate alleles under directional selection

Alleles with high pairwise haplotype scores (PHS) were identified by Toomajian *et al* (Toomajian et al 2006). An in-frame threonine-glycine (ThrGly) insertion in *CURLY LEAF* was found in 27 out of 96 *A. thaliana* accessions sampled. This data suggests that *CLF* may be under directional selection. Data provided by our collaborator Magnus Nordborg[1].

[1]University of Southern California, USA.

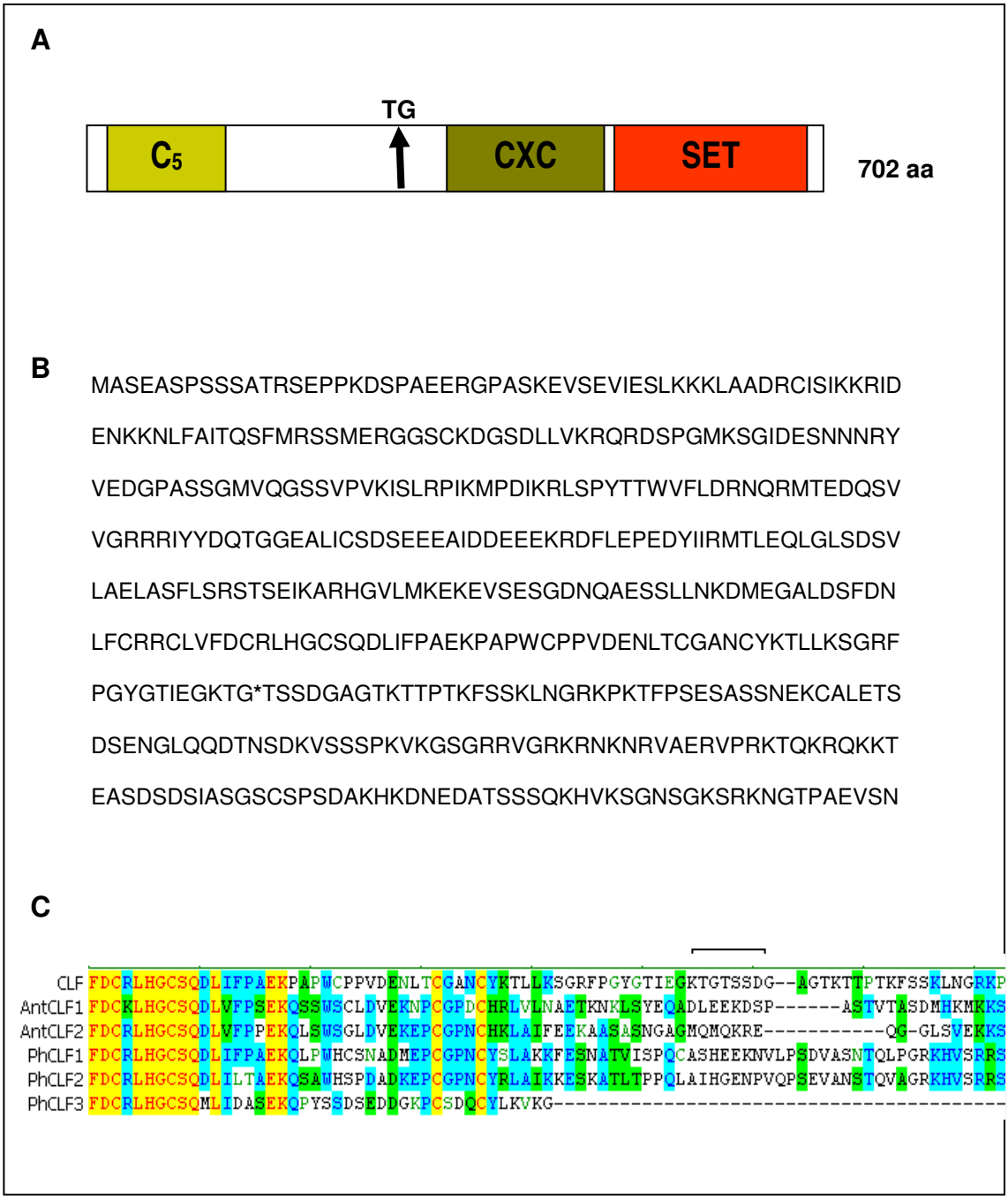

Figure 3.1 The threonine-glycine polymorphism in *CLF* is found outside functional domains and is poorly conserved in *Antirrhinum* and *Petunia*

(A) The primary structure of CLF showing the conserved domains found in Pc-G proteins, and location of the in-frame threonine-glycine insertion (marked by arrow) which is found outside the conserved C_5, CXC and SET domains. (B) The position of the in-frame ThrGly insertion within the full length CLF sequence (marked by *). (C) The alignment of the *Petunia* and *Antirrhinum CLF*-like genes to *CLF* revealed that the threonine-glycine insertion is located in a non-conserved region of *CLF* (for the full alignment, refer to the Appendix).

We decided to test this by complementing *clf*- null mutants using alleles of *CLF* with and without the polymorphism. As there are several polymorphisms in *CLF* between accessions, it is important to address whether it is specifically the ThrGly insertion which affects flowering time.

The five *A. thaliana* accessions that contained the ThrGly insertion in CLF were compared to five accessions that do not, in order to investigate whether differences in flowering time were observed under our growth conditions. 24 plants per *A. thaliana* accession were grown in SD and LD, flowering time was measured by the number of days to 1 cm bolt. Figure 3.2 shows the SD to LD flowering time ratio from the flowering time results. With 4 weeks vernalization treatment, the accessions with the polymorphism in *CLF* show a reduced SD to LD flowering time ratio (in gray). This implies that plants with the polymorphism in *CLF* are less sensitive to day length changes. Light and day length are an important developmental stimuli that regulate key developmental decisions, including flowering (Krizek & Fletcher 2005). Plants that are less susceptible to photoperiod may therefore have a selective advantage in varying environmental conditions.

The flowering time phenotype associated with a non-synonymous change in *CLF* raises the following question; could natural variation in *CLF* regulate flowering time? It would be surprising if polymorphisms within *CLF* were beneficial under natural conditions because *CLF* regulates several target genes, including the floral homeotic gene *AG* and *AP3*. Moreover, *clf*- mutants are generally dwarfed, sub-viable plants with low fertility. It is possible, however, that the effects of a specific polymorphism could be less pleiotropic, disrupting a specific function or interaction, so that most target genes are left unaffected. If the ThrGly insertion had been inserted in one of the well conserved functional domains, it is probable that CLF function would have been more severely affected. In this situation, the ThrGly insertion is in a non-conserved region of CLF, away from the histone methyltransferase domain, and one can envisage that it could change the protein conformation slightly leading to slight changes to a specific interaction.

A. thaliana Accession	SD:LD flowering time ratio
Columbia-0	2.57
Estonia-0	2.60
RMX-A180	2.25
Var-2-1	2.60
Bayreuth-0	2.44
Edinburgh-0	1.54
Omo-2-3	1.70
Wietze-0	1.68
Kindalville-0	1.63
Gowttingen-0	1.82

Figure 3.2 SD to LD flowering time ratios of *A. thaliana* accessions

Flowering time was measured by the number of days between germination and the production of a 1 cm bolt. 24 plants per *A. thaliana* accession were vernalized for 4 weeks and grown in either LD (16 hrs light/ 8 hrs dark) or SD (8 hrs light/ 16 hrs dark) at 22 ℃. The SD to LD flowering time ratios show that accessions that carry the ThrGly polymorphism in *CLF* (gray) have a reduced SD to LD flowering time ratio compared to those without (white).

There are several known examples of alleles that have specific effects, leaving other known targets unaffected. Prof. Richard Amasino identified a *clf* allele, *clf-59*, that has a mutation in the CXC region of CLF (pers. comm.). *clf-59* is a very specific dominant gain of function allele that alters the interaction of *CLF* with *FLC*. The *clf-59* allele does not affect *AP3* or *AG* expression, and consequently does not have the characteristic leaf curling phenotype associated with *clf-* mutants. *fca- clf-59* mutants are early flowering (*fca-* mutants are late flowering). Similarly, a novel gain of function allele of *CRY2*, which encodes the blue-light photoreceptor cryptochrome-2, was discovered by quantitative trait loci (QTL) analysis. A single amino-acid substitution results in the reduction of light-induced down regulation of CRY2 under SD, leading to early flowering (El-Assal et al. 2001). Allelic variation in *PHYTOCHROME C* (*PHYC*), was reported to be mediating natural variation in flowering time (Balasubramanian et al. 2006). *PHYC* alleles are highly polymorphic with varying levels of activity in different accessions of *A. thaliana*. These examples support the preliminary findings that polymorphisms in *CLF* may act as part of a mechanism regulating flowering time in the wild.

3.iii. The Columbia-0 and Edinburgh-0 *CLF* alleles complement null *clf-28* mutants

To address whether *CLF* regulates flowering time in natural populations of *A. thaliana*, *CLF* alleles were amplified and cloned from two *A. thaliana* accessions with and without the ThrGly insertion. The Columbia-0 *A. thaliana* accession does not contain the ThrGly polymorphism in *CLF*, whereas the Edinburgh-0 accession contains the ThrGly insertion. The *CLF* alleles were compared in a null isogenic background lacking *CLF+* activity, *clf-28*. Apart from the ThrGly polymorphism there are several other differences between the Columbia-0 and Edinburgh-0 CLF sequences. In order to address whether the ThrGly insertion alone is responsible for the reduced SD to LD flowering time phenotype, the Columbia-0 *CLF* allele was mutagenized to incorporate the ThrGly insertion and used to complement the null allele *clf-28*. *clf-28* contains a T-DNA insertion in the fourth exon of *CLF* within Columbia-0 background (refer to Materials and Methods).

The 7.5 Kb genomic fragment that spans the *CLF* locus, is sufficient to complement *clf* mutants (Schubert et al. 2006), and therefore contains all the necessary regulatory sequences for *CLF* function, was amplified from Edinburgh-0 and Columbia-0. The amplification products were sequenced, confirming that the PCR amplification did not introduce any mutations. The genomic fragments were then cloned into the *pSC-B1* plasmid and cloned into the binary vector *pCAMBIA1390* (Figure 3.3) to create the clones *pCAMBIA::CLF_Col-0*, *pCAMBIA::CLF_Edi-0* and *pCAMBIA::CLF_Col-tgtg*. *clf-28* mutants have low fertility with poor floral dip transformation efficiency, so large numbers of seed need to be screened to find primary transformants (T1). Therefore, the transgenes were introduced into *clf-28/CLF+* heterozygotes (that have a wildtype phenotype) by floral dip transformation. Primary transformants were selected on hygromycin, the selectable marker carried by *pCAMBIA1390* binary vector. The T2 progeny of selfed primary transformants were grown on sterile media containing hygromycin to identity single locus insertion lines segregating at 3:1 for resistance. Figure 3.4 A shows *clf-28* plant complemented by *pCAMBIA::CLF_Col-0* and Figure 3.4 B shows the sibling of *clf-28* homozygous plants (100 % kanamycin resistant). Figure 3.4 C shows the complementation of *clf-28* homozygous plants by *pCAMBIA::CLF_Edi-0* and Figure 3.4 D shows the sibling *clf-28* homozygous plant. The complementation of *clf-28* by our transgenic constructs confirmed that *CLF+* function was restored by the transgenic *CLF*.

3.iv. *pCAMBIA::CLF_Col-0, pCAMBIA::CLF_Edi-0* and *pCAMBIA::CLF_Col-tgtg* do not affect flowering time

For each construct, we identified 4 independent transgenic lines that were homozygous for *CLF* transgene and homozygous for *clf-28*. These lines were vernalized for 4 weeks and then grown in either SD or LD. We determined flowering times of the parental lines, Edinburgh-0 and Columbia-0, and the transgenic lines in each condition. Under LD the Edinburgh-0 accession, which contains the ThrGly insertion, flowered on average 12 $\pm$ 2 days later than

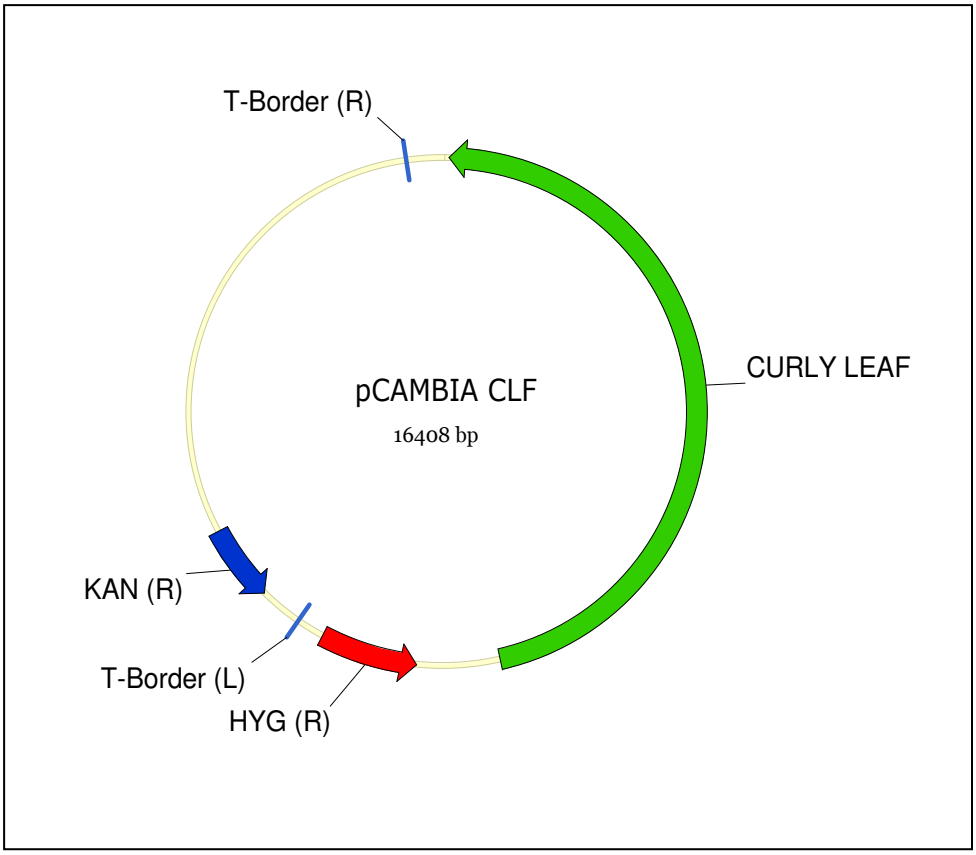

Figure 3.3 Construct design of *pCAMBIA::CLF*

The 7.5 Kb genomic fragment including the *CLF* gene was amplified from Columbia-0 and Edinburgh-0 by PCR, subcloned into *pSC-B* and cloned into the binary vector *pCAMBIA1390*, conferring hygromycin resistance. The resulting two clones were named *pCAMBIA::CLF_Col-0* and *pCAMBIA::CLF_Edi-0*. *pCAMBIA::CLF_Col-tgtg* was created by mutagenising *CLF* amplified from Columbia-0 and cloning into *pCAMBIA1390*. The 3 clones were individually transformed into *A. tumefaciens* and were introduced into *A. thaliana* by floral dip transformation.

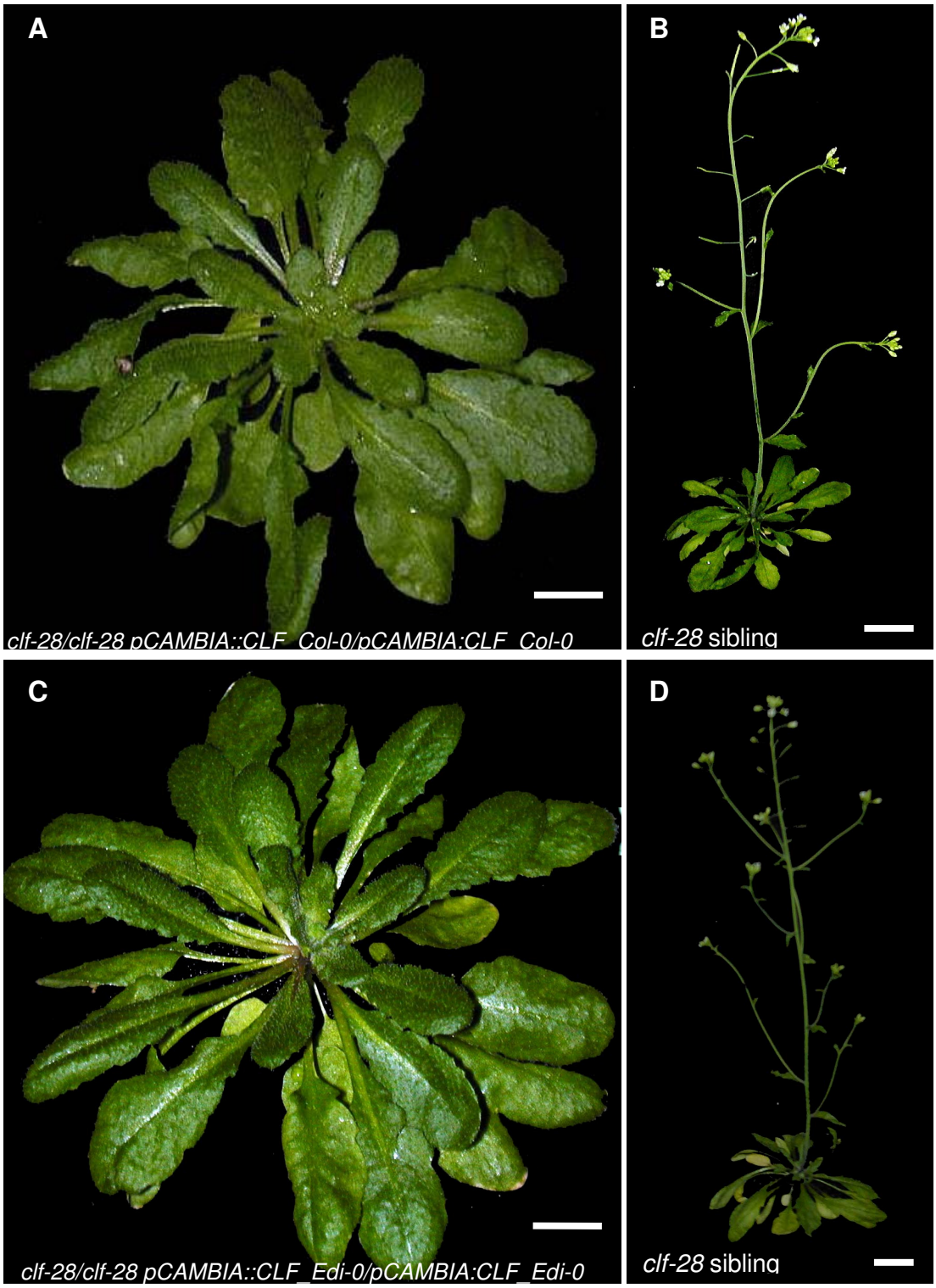

Figure 3.4 Complementation of *clf-28* by transgenic *CLF*

56 day old sibling plants grown in SD showing 3:1 segregation for *clf-* phenotype. Independent transgenic lines that confer 100% kanamycin resistance, homozygous for *clf-28,* and 75% hygromycin resistant for the *CLF* transgene. (A) Complementation of *clf-28* by *pCAMBIA::CLF_Col-0*. (B) Sibling *clf-28* homozygote that does not contain the *CLF* transgene. (C) Complementation of *clf-28* by *pCAMBIA::CLF_Edi-0*. (D) Sibling *clf-28* homozygote that does not contain the transgene. Scale bar 1 cm.

Columbia-0 (Figure 3.5). The flowering times of the 4 independent transgenic lines for both *pCAMBIA::CLF_Col-0* and *pCAMBIA::CLF_Edi-0* (Figures 3.5 and 3.6) did not differ from the flowering time of the parent line Columbia-0. No flowering time phenotype was observed of the 3 independent transgenic lines carrying *pCAMBIA::CLF_Col-tgtg*. Figure 3.6 shows the flowering time of the transgenic plants in SD. Edinburgh-0 flowered on average 10 $\pm$ 1.5 days earlier than the parental Columbia-0 flowering time (Figure 3.6).

Edinburgh-0 showed a reduced SD to LD flowering time ratio whilst Columbia-0 did not (Figure 3.7). The transgenic lines carrying the ThrGly insertion did not show reduced SD to LD flowering time ratios under the standard growth conditions observed in this experiment (Figure 3.7). This indicates that the ThrGly insertion is not responsible for the flowering time phenotype observed. The flowering times of *pCAMBIA::CLF_Col-tgtg* transgenic plants illustrated that the ThrGly polymorphism does not affect this part of CLF function in a Columbia-0 background. The flowering times of *pCAMBIA::CLF_Edi-0* under LD and SD in a Columbia-0 background is the same as *pCAMBIA::CLF_Col-0* indicating that Columbia-0 is not sensitive to slight changes in *CLF*.

3.v. Conclusion

Accessions of *A. thaliana* that contained the ThrGly polymorphism in *CLF* showed a reduced SD to LD flowering time ratio, indicating that the accessions are less sensitive to changes in day length. The role of *CLF* in regulating this variation in flowering time in natural populations of *A. thaliana* was investigated by comparing *CLF* alleles in a null isogenic background, *clf-28*. The Columbia-0 accession was chosen as representative of *CLF* alleles which do not contain the ThrGly polymorphism. The Edinburgh-0 accession was chosen as representative of *CLF* alleles which contain the ThrGly polymorphism. The two alleles were transformed into a null *clf-28* background and flowering times of independent transgenic lines were scored in both SD and LD. The flowering time ratios were calculated and expressed as the ratio of the number of days to flower in SD to LD. *pCAMBIA::CLF_Edi-0* in a Columbia-0 background flowered under SD and LD the same as *pCAMBIA::CLF_Col-0.*

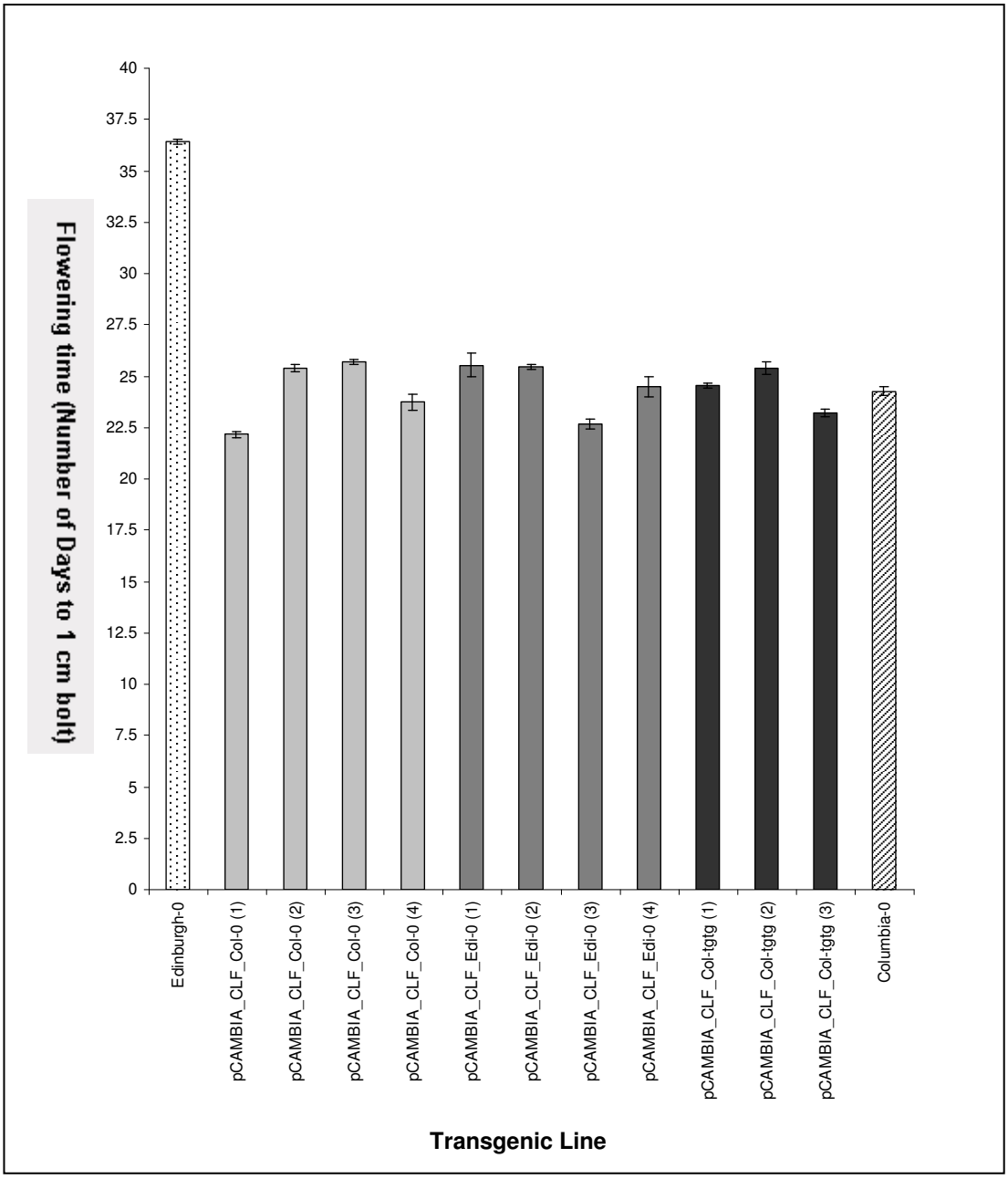

Figure 3.5 Flowering time of selected lines under LD

Flowering time in LD was scored for 4 independent transgenic lines of *pCAMBIA::CLF_Col-0,* and *pCAMBIA::CLF_Edi-0*, and 3 independent lines of *pCAMBIA::CLF_Col-0tgtg*. Flowering times of parental lines, Columbia-0 and Edinburgh-0 were scored along side. Flowering time was measured by the number of days between germination and the production of a 1 cm bolt. Measurements were taken from 20 plants per line after 4 weeks vernalization treatment. Growth conditions were 16 hrs light/ 8 hrs dark at 22 ℃. $\pm$ Standard error of the mean (SEM) bars were calculated by dividing the standard deviation (SD) by the square root of the number (n) of plants 20.

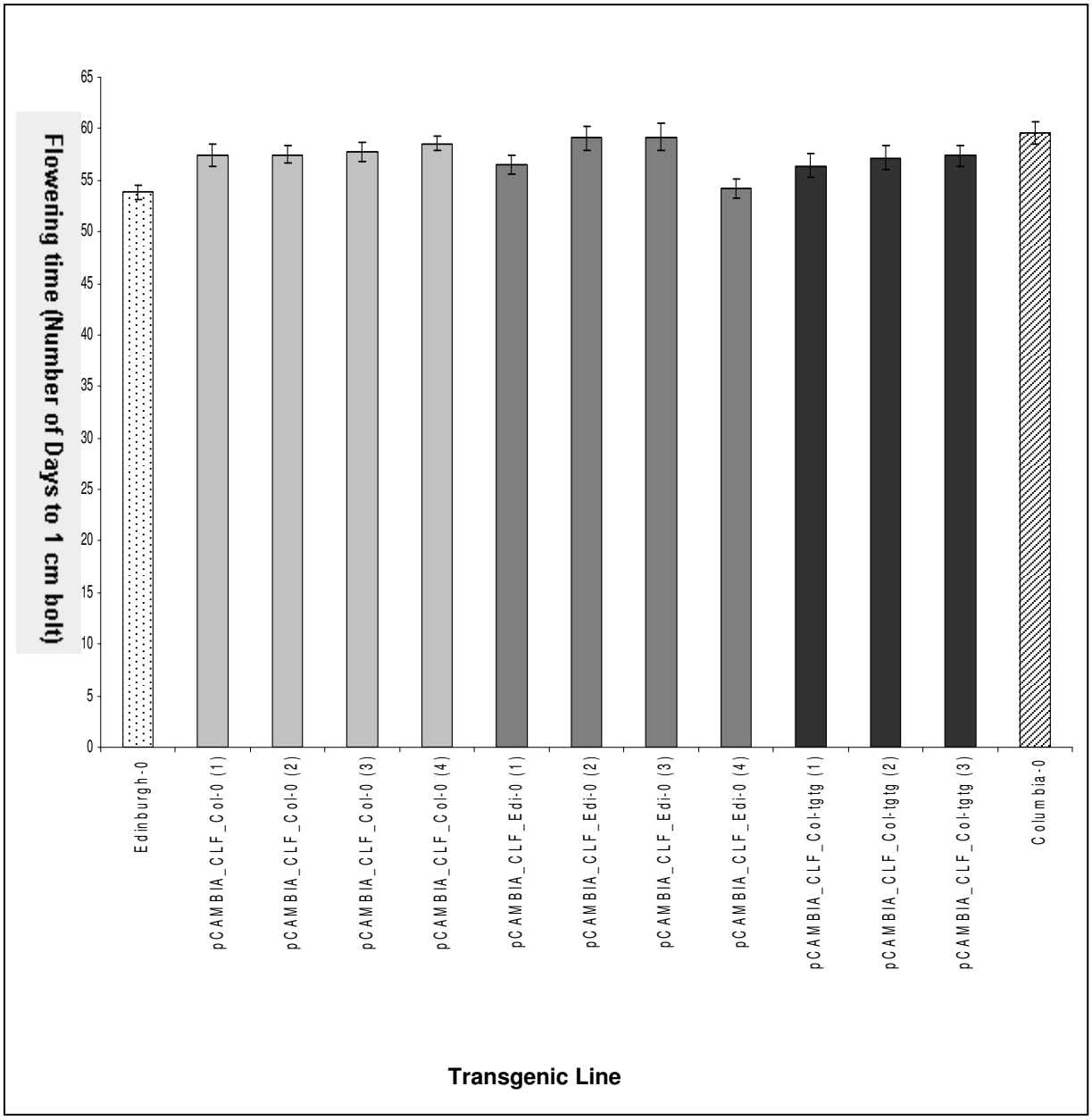

Figure 3.6 Flowering time of selected lines under SD

Flowering time in SD was scored for 4 independent transgenic lines of *pCAMBIA::CLF_Col-0,* and *pCAMBIA::CLF_Edi-0*, and 3 independent lines of *pCAMBIA::CLF_Col-0tgtg*. Flowering times of parental lines, Columbia-0 and Edinburgh-0 were scored along side. Flowering time was measured by the number of days between germination and the production of a 1 cm bolt. Measurements were taken from 20 plants per line after 4 weeks vernalization treatment. Growth conditions were 8 hrs light/ 16 hrs dark at 22 ℃. ± Standard error of the mean (SEM) bars were calculated by dividing the standard deviation (SD) by the square root of the number (n) of plants 20.

***A. thaliana* Accession**	**SD:LD flowering time ratio**
Edinburgh-0	1.47
pCAMBIA::CLF_Col-0 (1)	2.59
pCAMBIA::CLF_Col-0 (2)	2.26
pCAMBIA::CLF_Col-0 (3)	2.25
pCAMBIA::CLF_Col-0 (4)	2.47
pCAMBIA::CLF_Edi-0 (1)	2.24
pCAMBIA::CLF_Edi-0 (2)	2.60
pCAMBIA::CLF_Edi-0 (3)	2.33
pCAMBIA::CLF_Edi-0 (4)	2.42
pCAMBIA::CLF_Col-0tgtg (1)	2.30
pCAMBIA::CLF_Col-0tgtg (2)	2.35
pCAMBIA::CLF_Col-0tgtg (3)	2.47
Columbia-0	2.46

Figure 3.7 SD to LD flowering time ratios of selected lines

The SD to LD flowering time ratios of independent transgenic lines carrying *pCAMBIA::CLF_Col-0*, *pCAMBIA::CLF_Edi-0* and *pCAMBIA::CLF_Col-0tgtg.* The Columbia-0 and Edinburgh-0 parental lines show the originally observed flowering time ratio phenotypes. The SD to LD flowering time ratios of the *CLF* transgenic lines did not differ to the Columbia-0 parental line.

This indicates that Columbia-0 is not sensitive to slight changes in *CLF*. *pCAMBIA::CLF_Col-tgtg* flowered in a similar number of days as the parental Columbia-0 line confirming that the ThrGly polymorphism did not affect CLF function in a Columbia-0 background. These results show that the ThrGly insertion itself is not responsible for the reduced SD to LD flowering time ratio observed. Under standard growth conditions it may be concluded that the ThrGly polymorphism in *CLF* does not play a role in regulating flowering time in natural populations of *A. thaliana*. The reduced SD to LD flowering time ratio observed in accessions of *A. thaliana* with *CLF* alleles that contain the ThrGly polymorphism may be caused through the interactions of other closely linked genes. The ThrGly polymorphism in *CLF* could hinder or alter CLF activity, which could then be complemented by *SWN*. It is reasonable to assume that natural variation in *CLF* highlights potential roles for *SWN* in the regulation of floral transition. The reduced SD to LD flowering time ratio could be accentuating the effects of variation in functional Pc-G protein activity.

4. Challenging Pc-G mediated silencing of *AGAMOUS*

4.i. Introduction

Floral homeotic genes are silenced by Pc-G proteins during vegetative development and expression is restricted to specialized cell types upon floral transition. The curled leaf phenotype is one of the main characteristics of *clf*- mutants, and is caused by the ectopic expression of the floral homeotic gene, *AG* (Goodrich et al. 1997). By using a GUS::reporter construct for *AG,* we can see that *AG* expression is confined to whorls 3 and 4 of flowers, where it confers carpel and stamen identity (Figure 4.1). *CLF* regulation of *AG* expression (Goodrich et al. 1997) is associated with covalent histone modifications that are mitotically heritable. ChIP analysis reveals that *AG* is enriched for H3K27me3 (a repressive mark associated with Pc-G silencing) (Schubert et al. 2006). In plants, H3K27me3 marks are associated with specific genes, precisely covering single loci but seldom extending into neighbouring genes (Zhang et al. 2007). H3K27me3 marks at the *AG* locus are necessary for maintaining *AG* silencing (Schubert et al. 2006).

The mechanisms through which *CLF* maintains the established expression pattern of *AG* during vegetative development and the first and second stages of floral meristem development are unknown. It is not clear how histone modifications are propagated through mitosis and maintained during cell division. If Pc-G mediated silencing were to be labile early in leaf development, associated with cell proliferation, target gene repression may be less stable in young developing leaves. In older leaves, the mitotically heritable state of *AG* repression may become more fixed in cells that are no longer proliferating. *AG* is not expressed during vegetative development and is confined to stamens and carpels of flowers indicating that Pc-G mediated silencing has to be reset in particular cell types.

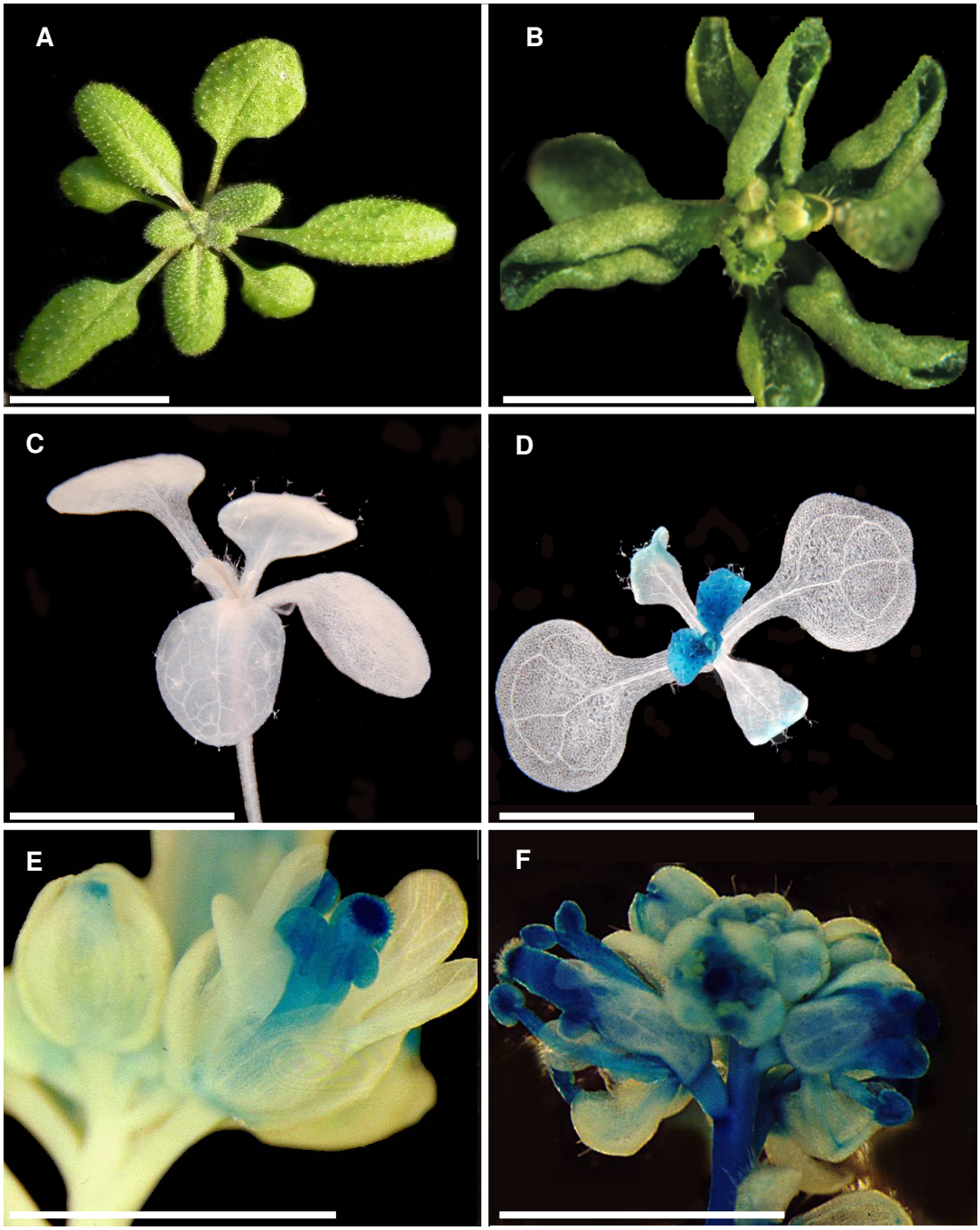

Figure 4.1 *AGAMOUS* expression in wildtype and in *clf*- mutants

(A) The wildtype *A. thaliana*. (B) *clf*- mutants are dwarfed plants with curled leaves. (C) An *AG* promoter::GUS reporter construct was introduced into wildtype and *clf*- mutants and β-glucuronidase activity was analyzed. *AG* is not expressed in leaves of wildtype plants. (D) Ectopic expression of *AG* occurs in leaves of *clf*- mutants. (E) *AG* expression is restricted to carpels and stamens in wildtype plants. (F) However, in *clf*- mutants, the ectopic expression of *AG* occurs throughout the whole plant. Scale bar 1 cm. From Goodrich et al. 1997 and Chanvivattana et al., (2002).

Whether Pc-G mediated silencing is maintained and reset is a passive mechanism or an active removal of histone modifications, is still unknown. *CLF* mediated repression of *AG* in leaves might result in irreversible changes in chromatin structure that in turn render *AG* inaccessible to transcription factors. However, there has been little analysis to address whether Pc-G mediated silencing is reversible. This chapter describes how I have targeted the *LhG4* trans-factor to *AG* to test whether Pc-G mediating silencing can be reversed during early and late leaf development.

4.ii. Targeting the *LhG4* trans-activator to *AG* to challenge Pc-G mediated silencing

The aims of this chapter are to challenge Pc-G mediated silencing and to ascertain the stability of Pc-G mediated silencing during plant development by introducing *AG* into the *LhG4/ pOp* bipartite transactivation system (Moore et al. 1998) (Figure 4.2). Driver lines express the *LhG4* chimeric transcription factor which is comprised of a modified LacDB DNA binding domain (a mutant that has increased affinity to the *Lac Operator (Lac Op)*) translationally fused to the transcriptional-activation domain-II from *GAL4* trans-factor of *Saccharomyces cerevisiae*. There are a range of driver lines that express the *LhG4* trans-factor under a variety of tissue-specific promoters. The *LhG4* driver lines induce the expression of constructs carrying *Lac Op* binding sites upstream of a reporter gene.

In order to test whether Pc-G mediated silencing of *pOp:AG-I::GUS* is reversible and stable in older leaves, the late leaf *LhG4* promoter, *p650::LhG4*, was selected as a driver. Overcoming Pc-G silencing may differ during development, so to discover whether Pc-G silencing of *pOp:AG-I::GUS* is reversible early in development, the early leaf *LhG4* promoter, *pAINTEGUMENTA::LhG4* (*pANT::LhG4)* was also selected. Using the *pANT::LhG4* promoter line, the stability of transient activation of *pOp:AG-I::GUS* would also be tested.

4.iii. *pOp:AG-I::GUS* faithfully mimics endogenous *AG* expression

The *pAG-I::GUS* reporter transgene contains the necessary *cis*-acting sequences recognised by *CLF* (Sieburth & Meyerowitz 1997) and therefore, necessary for repression.

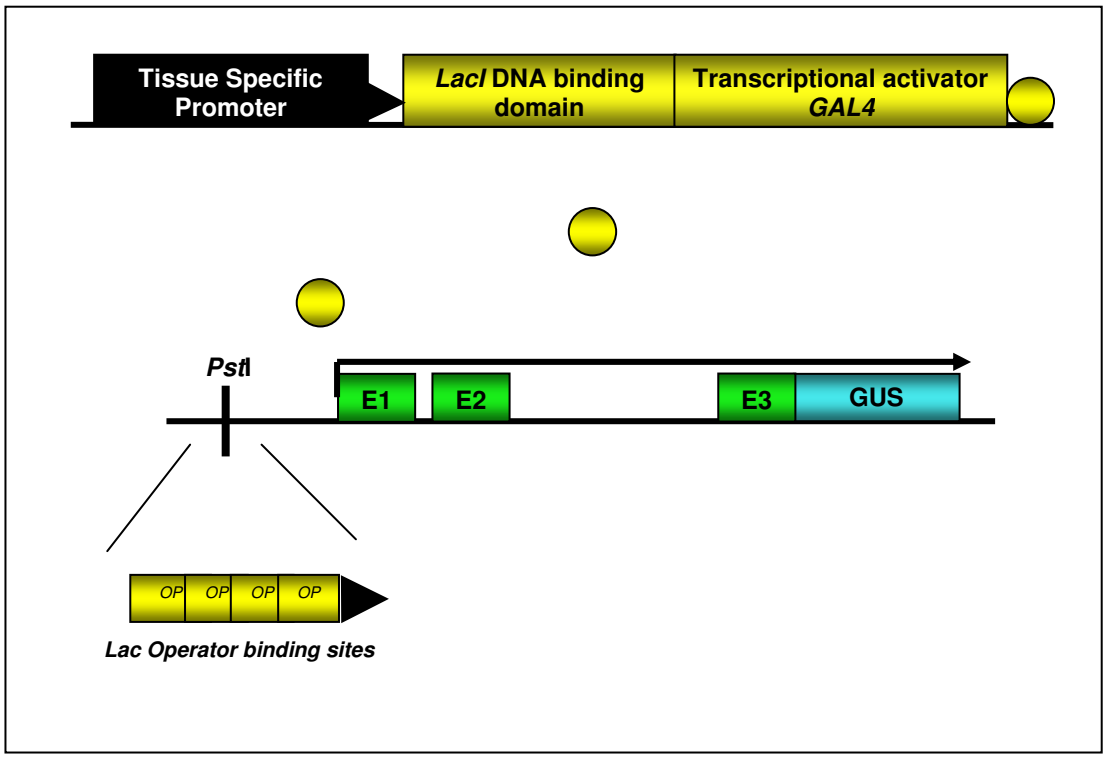

Figure 4.2 Challenging *AG* silencing using the *LhG4/ pOp* transactivation system

The *LhG4/ pOp* transactivation system is a two component system. The driver line is comprised of a tissue specific promoter that is upstream of a *LacI* DNA binding domain that is translationally fused to the transcriptional-activation domain-II from *GAL4 Saccharomyces cerevisiae*. *Lac Op* sites were cloned upstream of the *pAG-I::GUS* construct (provided by Sieburth et al 1997). *GAL4* is expressed under the tissue specific promoter and once crossed to plants containing *pOp:AG-I::GUS* reporter construct, *AG* and GUS are expressed under the control of the promoter of the driver.

pAG-I::GUS has been previously shown to faithfully mimic the expression pattern of native *AG* gene and is regulated by *CLF* (Schubert et al. 2006). The *pOp:AG-I::GUS* reporter transgene contains 5 x *Lac Op* sequences that were inserted into a *Pst*I site upstream of *pAG-I::GUS* (Figure 4.2). To confirm that *pOp:AG-I::GUS* mimics endogenous *AG* expression, *pOp:AG-I::GUS* was introduced into wildtype plants (Landsberg e*recta*) by floral dip transformation. β-glucuronidase activity was analyzed in the progeny of primary transformants, that had been selected for *pOp:AG-I::GUS* that confers BASTA® (PPT) herbicide resistance. Figure 4.3 A shows that *pOp:AG-I::GUS* is not expressed in young leaves in Landsberg *erecta* (wildtype) background. Figure 4.3 C is magnified from Figure 4.3 A (marked by *) showing *pOp:AG-I::GUS* is not expressed in the vasculature of young and older leaves. Figure 4.3 E shows β-glucuronidase activity in whorls 3 and 4 of flowers. These results confirm that *pOp:AG-I::GUS* expression in wildtype plants is restricted to carpels and stamens in flowers.

To ensure *pOp:AG-I::GUS* was repressed by *CLF* as endogenous *AG* (Schubert et al. 2006), *pOp:AG-I::GUS* was introduced into *clf-2/+* heterozygotes by floral dip transformation, as *clf-2* homozygotes have low fertility with poor transformation efficiency. In the progeny of the transformed lines, *clf-2* primary transformants (T1) were identified by phenotype. β-glucuronidase activity was analyzed in 14 day old *clf-2* mutants carrying the *pOp:AG-I::GUS* reporter transgene. Figure 4.3 B shows β-glucuronidase activity in *clf-2* showing ectopic expression of *pOp:AG-I::GUS* in leaves and throughout the whole plant. Figure 4.3 D is magnified from Figure 4.3 B (marked by *) showing ectopic expression of *pOp:AG-I::GUS* in young leaves, expressed in the vasculature of leaves. Figure 4.3 F shows β-glucuronidase activity of 10 day old *clf-2* seedlings. Taken together, these observations lead us to the conclusion that *pOp:AG-I::GUS* faithfully mimics the endogenous expression pattern of *AG.* The insertion of the 5 x *Lac Op* sites has not disrupted *AG* regulation.

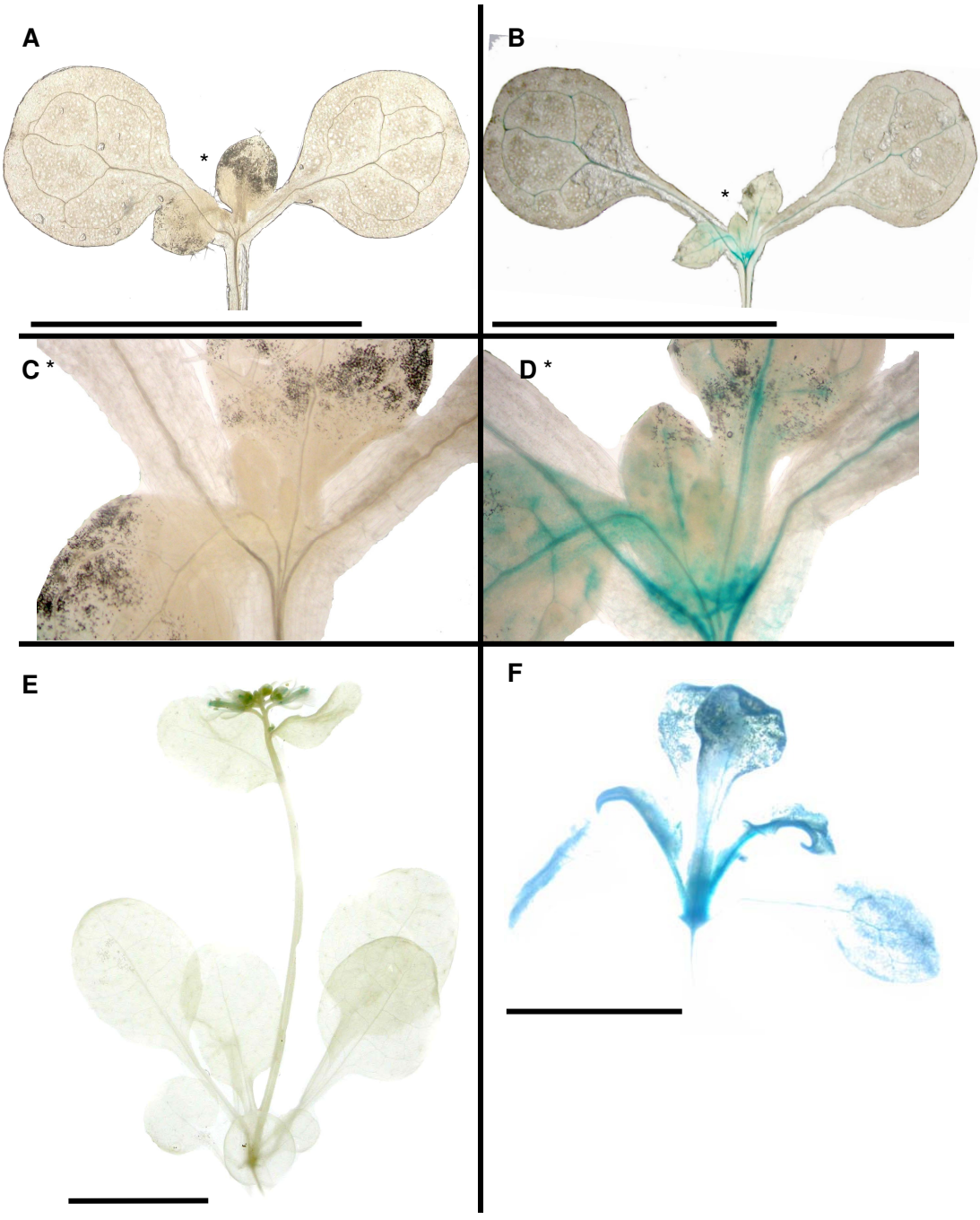

Figure 4.3 *pOp:AG-I::GUS* faithfully mimics endogenous *AG* expression

Figure 4.3 *pOp:AG-I::GUS* faithfully mimics endogenous *AG* expression

(A) *pOp:AG-I::GUS* was introduced into wildtype plants (Landsberg *erecta*) by floral dip transformation and primary transformants were selected for resistance to BASTA® (PTT) herbicide. β-glucuronidase activity was analyzed in 7 day old T2 *pOp:AG-I::GUS* plants, confirming that *pOp:AG-I::GUS* is not expressed in young leaves. (B) *pOp:AG-I::GUS* was introduced into *clf-2/+* (Landsberg *erecta* background) plants and primary transformants were selected for resistance to BASTA® (PPT) herbicide. β-glucuronidase activity was analyzed in 7 day old T2 *clf-2* plants showing ectopic expression of *pOp:AG-I::GUS* occurred in young leaves. (C) Magnified from Figure A revealing no β-glucuronidase activity in the vasculature of leaves and (D) is magnified from Figure B revealing that β-glucuronidase activity occurs in the vasculature of leaves in *clf-2* plants. (E) β-glucuronidase activity was analyzed in 28 day old *pOp:AG-I::GUS* plant showing that *pOp:AG-I::GUS* expression is restricted to the carpels and stamens of flowers. (F) β-glucuronidase activity was observed throughout the whole of *clf-2* plants. Scale bar 1 cm.

4.iv. Activation *pOp:AG-I::GUS* in leaves under the late leaf promoter *p650::LhG4*

The *p650::LhG4* driver line is comprised of a late leaf promoter of the *At1g13650* gene driving the *LhG4* trans-factor. *p650::LhG4* is activated when leaves are 1 mm long, at the distal end first and then throughout the leaf (Y. Eshed pers. comm.). To confirm that *p650::LhG4* was not expressed in young leaves, homozygous *p650::LhG4* plants were crossed to plants carrying the reporter construct, *pOp::GUS*, that has *Lac Op* sites cloned upstream of the *GUS* reporter gene. β-glucuronidase activity was analyzed in F3 plants that were homozygous for both *p650::LhG4* and *pOp::GUS*. Figure 4.4 A shows β-glucuronidase activity in a 7 day old *p650::LhG4 pOp::GUS* plant revealing *p650::LhG4* expression around the edges of the cotyledons and first leaves. Figure 4.4 C shows β-glucuronidase activity in a 30 day old plant with expression in the older rosette leaves, cauline leaves, but little expression in the main vasculature, petioles and inflorescence meristem. Figure 4.4 F shows GUS expression in minor vasculature and the ground tissue of the leaves. Expression is leaving from the distal end of the main vein and the petioles. Figure 4.4 I shows *p650::LhG4* is expressed in the distal regions of the sepals, but not in other floral organs. Taken together these results confirm that *p650::LhG4* is not expressed in young leaves and is expressed very widely in older rosette leaves and sepals.

p650::LhG4 homozygotes were crossed to T3 *CLF+* plants homozygous for *pOp:AG-I::GUS*. Strikingly, *CLF* mediated silencing of *AG* is reversed. Figure 4.4 D shows activation of *pOp:AG-I::GUS* under the *p650::LhG4* driver line in the vasculature of rosette leaves, cauline leaves, petioles and flowers (Figure 4.4 J). Figure 4.4 B shows *pOp:AG-I::GUS* under the late leaf driver line *p650::LhG4* is not activated in 7 day old plants. Figure 4.4 F *p650::LhG4 pOp::GUS* is uniformly expressed in leaves but not expressed in the major vasculature or petiole, whereas *pOp:AG-I::GUS* is activated mostly in the vasculature and petioles (Figure 4.4 G). *p650::LhG4* driving *pOp:AG- I::GUS* results in expression that is localised mainly in

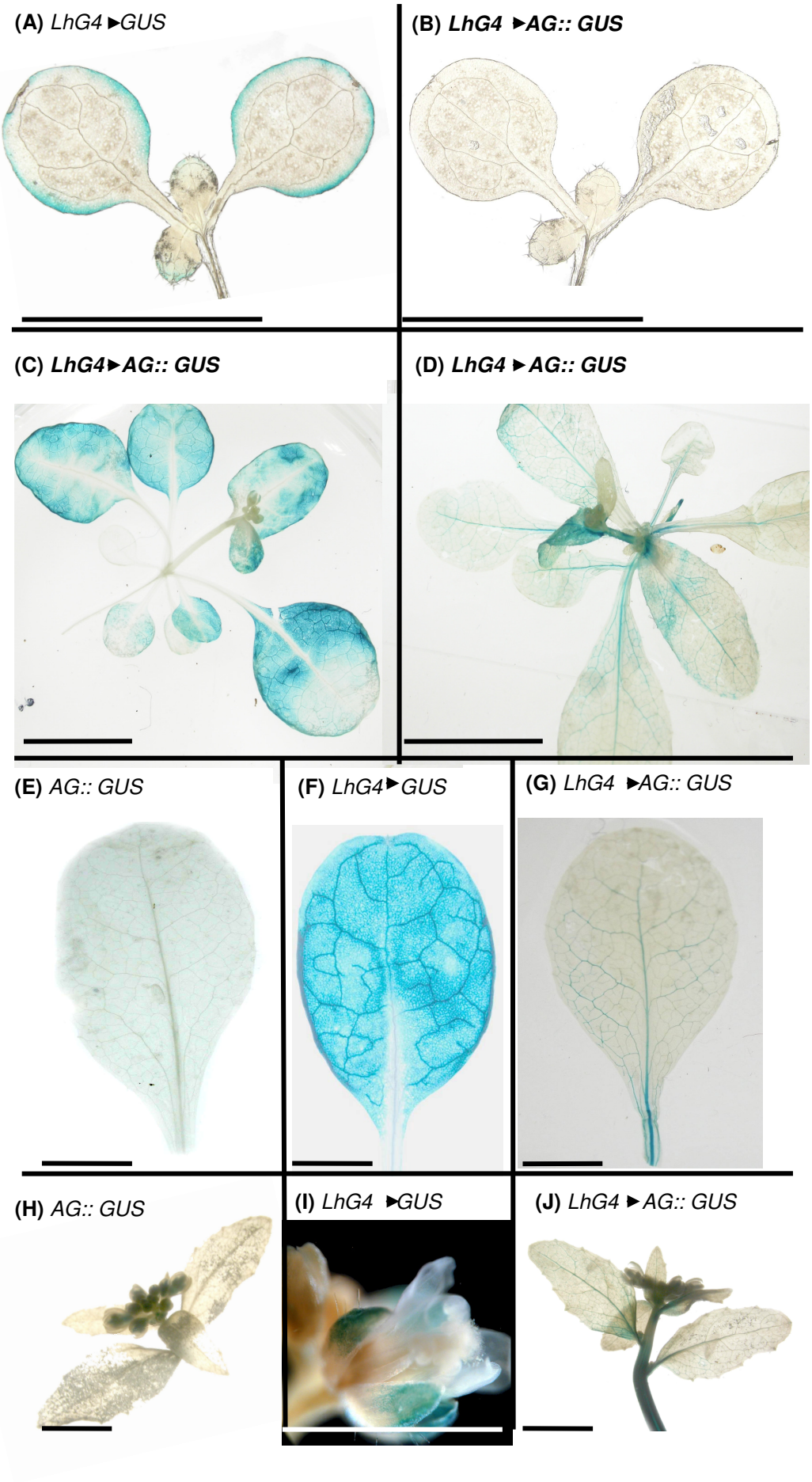

Figure 4.4 Expression of the *p650::LhG4* driver line in wildtype plants and overcoming Pc-G mediated silencing of *AG* in leaves

Figure 4.4 Expression of the *p650::LhG4* driver line in wildtype plants and overcoming Pc-G mediated silencing of *AG* in leaves

(A) 7 day old *p650::LhG4 pOp::GUS* F2 plants showing that *p650::LhG4* is not expressed in very young leaves or leaf primordia, but expressed at the distal end of the expanding leaf. (B) 7 day old *pOp:AG-I::GUS p650::LhG4* F2 plants showing that *pOp:AG-I::GUS* expression does not occur in young leaves. (C) *p650::LhG4 pOp::GUS* is stably expressed in older leaves with little difference between cauline and rosette leaves. (D) 21 day old *pOp:AG-I::GUS p650::LhG4* F2 plants showing activation of *pOp:AG-I::GUS* in the vasculature of rosette leaves, cauline leaves, petioles and flowers. (E) *pOp:AG-I::GUS* plants showing no expression of *pOp:AG-I::GUS* in leaves. (F) *p650::LhG4 pOp::GUS p650::LhG4* is expressed throughout the lamina of older rosette leaves, in vasculature of the leaf, including minor and major veins of young rosette leaves but not in the proximal mid rib. (G) GUS expression in *pOp:AG-I::GUS p650::LhG4* planrs occurs mostly in the vasculature and main leaf veins. (H) *pOp:AG-I::GUS* plants showing GUS expression is restricted to stamens and carpels of flowers. (I) *p650::LhG4 pOp::GUS* F2 plants showing the *LhG4* trans-factor *p650::LhG4* is expressed in the sepals of flowers. (J) β-glucuronidase activity was analyzed in 28 day old *pOp:AG-I::GUS p650::LhG4 F2* plants showing GUS expression in cauline leaves. Scale bar 1 cm.

the vasculature of young rosette leaves, petioles and inflorescence meristem of 28 day old plants (Figure 4.4). Taken together, these results show that the activation of *pOp:AG-I::GUS* by *p650::LhG4* occurs exclusively in the vasculature, rather then the entire expression domain of *p650::LhG4*, with little or no activation in the mesophyll or interveinal areas. Pc-G mediated silencing does not appear to be overcome in the majority of the *p650* expression domain, illustrating that Pc-G mediated silencing is not reversible in all cell types in older leaves.

4.v. Activation of *pOp:AG-I::GUS* in young leaves under the early leaf promoter *pAINTEGUMENTA::LhG4*

The *pAINTEGUMENTA::LhG4 (pANT::LhG4)* driver line is comprised of the promoter of *AINTEGUMENTA* (*ANT*). *ANT* is transiently expressed in the primordia of lateral organs (leaves, sepals and petals etc.) but not in older organs (Mizukami & Fischer 2000). To confirm *pANT::LhG4* expression, plants homozygous for the *pANT::LhG4* driver line were crossed to lines carrying the *pOp::GUS* reporter transgene. Figure 4.5 A shows β-glucuronidase activity in 10 day old *pANT::LhG4 pOp::GUS* plants showing the expression domain of the *pANT::LhG4* driver line in the region of the SAM and leaf primordia. Figure 4.5 C shows β-glucuronidase activity in a 14 day old plant, showing that *pANT::LhG4* is not expressed in leaves. These results confirm that *pANT::LhG4* is expressed in leaf primordia but not in older leaves. *pANT::LhG4* homozygotes were crossed to Landsberg *erecta* T3 plants homozygous for *pOp:AG-I::GUS*. F2 progeny were grown and β-glucuronidase activity was analyzed. Figure 4.5 B shows β-glucuronidase activity in 10 day old *pOp:AG-I::GUS pANT::LhG4* plant revealing that activation of *pOp:AG-I::GUS* occurs under the early leaf promoter *pANT::LhG4*, linked to the region of *pANT::LhG4* expression. Figure 4.5 D shows β-glucuronidase activity in 14 day old *pOp:AG-I::GUS pANT::LhG4* plants revealing that *pOp:AG-I::GUS* is activated in leaf primordia, the domain of *pANT::LhG4*, but not in older leaves.

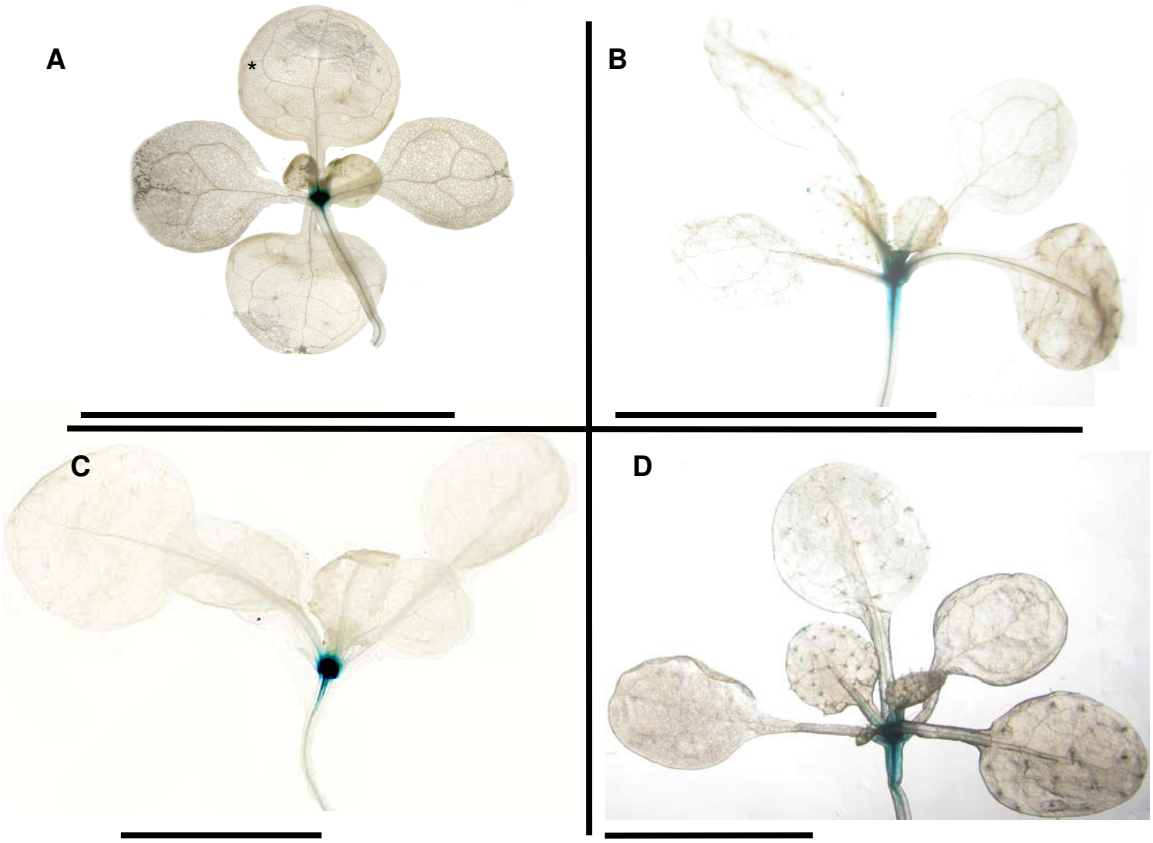

Figure 4.5 Expression of the *pANT::LhG4* driver line in the region of young leaf primordia and transient activation of *pOp:AG-I::GUS* by *pANT::LhG4*

(A) β-glucuronidase activity was analyzed in *pANT::LhG4 pOp::GUS* 10 day old F2 plants. The expression domain of the *pANT::LhG4* driver line is restricted to the SAM region. (B) β-glucuronidase activity was analyzed in 10 day old *pOp:AG-I::GUS pANT::LhG4* plants showing activation of *pOp:AG-I::GUS* in the *pANT::LhG4* domain, but not in older leaves. (C) β-glucuronidase activity was analyzed in 14 day old *pANT::LhG4 pOp::GUS* plants showing *pANT::LhG4* is not expressed in older leaves. (D) β-glucuronidase activity was analyzed in a 14 day old *pOp:AG-I::GUS pANT::LhG4* plants showing that transient activation of *pOp:AG-I::GUS* is not stable in older leaves, but remains active in lateral organs under *pANT::LhG4*. Scale bar 1 cm.

Activation of *pOp:AG-I::GUS* in leaf primordia driven by *pANT::LhG4* is not mitotically stable once expression of *pANT::LhG4* has ceased. This indicates that activation of a Pc-G target, *AG*, can occur early in leaf development, but it is not stable during leaf development.

4.vi. Conclusion

To investigate Pc-G mediated silencing of *AG*, a target gene of *CLF*, *AG* was introduced into the *LhG4/ pOp* transactivation system. *pOp:AG-I::GUS* expression was confirmed in stamens and carpels and repressed in leaves by *CLF*, faithfully mimicking endogenous *AG* expression. The late leaf driver line, *p650::LhG4* was used to test whether Pc-G mediated silencing could be overcome in older leaves. Surprisingly, activation of *pOp:AG-I::GUS* occurred in older leaves under the *LhG4* late leaf promoter, *p650::LhG4*. However, *pOp:AG-I::GUS* was not activated in the full expression domain of *p650::LhG4*, but outside the activator domain. This implies that Pc-G mediated silencing of target genes differs in different cell types in leaves.

Under the early leaf promoter, *pANT::LhG4*, *pOp:AG-I::GUS* was transiently activated in the leaf primordia. *pOp:AG-I::GUS* was activated in leaf primordia, in the entire expression domain of *pANT::LhG4*, however, activation was not stable in older leaves. The significance of this result is that Pc-G silencing appears to be overcome everywhere in the *pANT::LhG4* expression domain. This result also illustrates that transient activation of *pOp:AG-I::GUS* is not mitotically stable in the absence of *LhG4* trans-factor. This reveals that Pc-G mediated silencing is reversible in young developing leaves. Transient activation is not epigenetically stable and Pc-G silencing can be re-established once the *LhG4* trans-factor is no longer present.

5. Does activation of Pc-G target *APETALA1* require the removal of H3K27me3?

5.i. Introduction

A key feature of epigenetic silencing is that although stable and heritable, modifications are reversible; however, how Pc-G mediated silencing of target genes is reset remains unclear. Floral homeotic genes are regulated by Pc-G proteins, silenced during vegetative development and are expressed in particular cell types following floral induction. The transition from vegetative to reproductive development is regulated by exogenous signals and endogenous pathways. In response to long days (LD), CONSTANS (*CO*) the nuclear zinc-finger protein activates the floral pathway integrator *FLOWERING LOCUS T* (*FT*) that encodes a RAF-kinase-inhibitor-like protein (Samach et al. 2000). FT travels from the leaves to the shoot apical meristem and interacts with the bZIP transcription factor, *FD*, promoting *AP1* expression. *AP1* promotes the switch from inflorescence to floral meristem by initiating the development of new tissues, floral organs and specific cell types (Bowman et al. 1993). The vegetative meristem is then converted to an inflorescence meristem (Abe et al. 2005; Wigge et al. 2005) (Figure 1.8). Plants that express FD under the 35*S* RNA promoter of the cauliflower mosaic virus (*35S::FD*) show that FT can interact with FD in LD throughout the entire plant (not just the SAM) (Wigge et al. 2005).

The repressive histone modification associated with Pc-G silencing, H3K27me3, is abundant throughout the genome (Zhang et al. 2007). Studies suggest that H3K27me3 is dispersed over promoter sequences and spreads over target loci (Schubert et al. 2006; Sheldon et al. 2008; Zhang et al. 2007). It is not known how Pc-G directed H3K27me3 marks are lost, whether passively, through loss or replacement at mitosis, or actively through the activity of a demethylase. The transactivation of *AG* under the *LhG4/ pOp* transactivation system (discussed in Chapter 4) occurred despite the presence of *CLF+*. However, we could not tell whether the Pc-G directed H3K27me3 remains at the *AG* locus as we did not have sufficient

material to perform ChIP. The artificial nature of the *LhG4/ pOp* transactivation system also limited our analysis. The placement of the transgene within the genome would firstly have to characterized, in order to assess how alike the chromatin structure is to endogenous *AG* and how transcription factors that activate *AG* differ to the GAL4 of the transactivation system.

Like *AG*, *AP1* is a target of *CLF* and enriched with H3K27me3, the repressive mark associated with Pc-G gene mediated silencing (Zhang et al. 2007). In the presence of Pc-G proteins, activation of target genes can occur and so this raises the following question: does activation of a Pc-G silenced target require the removal of H3K27me3? By using the *35S::FD* background, activity can be characterized under natural circumstances of endogenous *AP1* (and associated histone modifications). The aim of this chapter is to elucidate the histone dynamics of the floral homeotic gene *AP1* in *35S::FD* to address whether activation of a Pc-G target requires the removal of H3K27me3. This chapter describes how we characterized the effects of photoperiod on flowering in *35S::FD* plants and the activation of *AP1* in leaves, relating these phenotypic changes to changes in histone modifications by ChIP analysis.

5.ii. Photoperiod enhances the *35S::FD* phenotype

Under short days (SD) *35S::FD* show no gross morphological phenotype. Figure 5.1 A shows *35S::FD* plants grown in SD, appearing wildtype (Figure 5.1 B). When *35S::FD* plants are grown in LD, they are dwarfed, with inwardly curling leaves and are early flowering (Figure 5.1 D). Changes in photoperiod affect the *35S::FD* phenotype (rapidly). Figure 5.1 E shows a 10 day old *35S::FD* plant 3 days following shift from SD to LD, leaf curling has occurred compared to *35S::FD* sibling grown only in SD (Figure 5.1 F). These results imply that under LD, FD can interact with FT throughout the whole plant, not just in the SAM. The *35S::FD* system was chosen for the current study due to the rapid activation of Pc-G target, *AP1*, in leaves following shift from SD to LD.

5.iii. Rapid activation of *AP1::GUS* upon day length shift in *35S::FD*

The expression of the floral meristem identity gene *AP1* is regulated by FT and FD (Abe et al. 2005). The *AP1::GUS* reporter was introduced into the *35S::FD* background. *35S::FD*

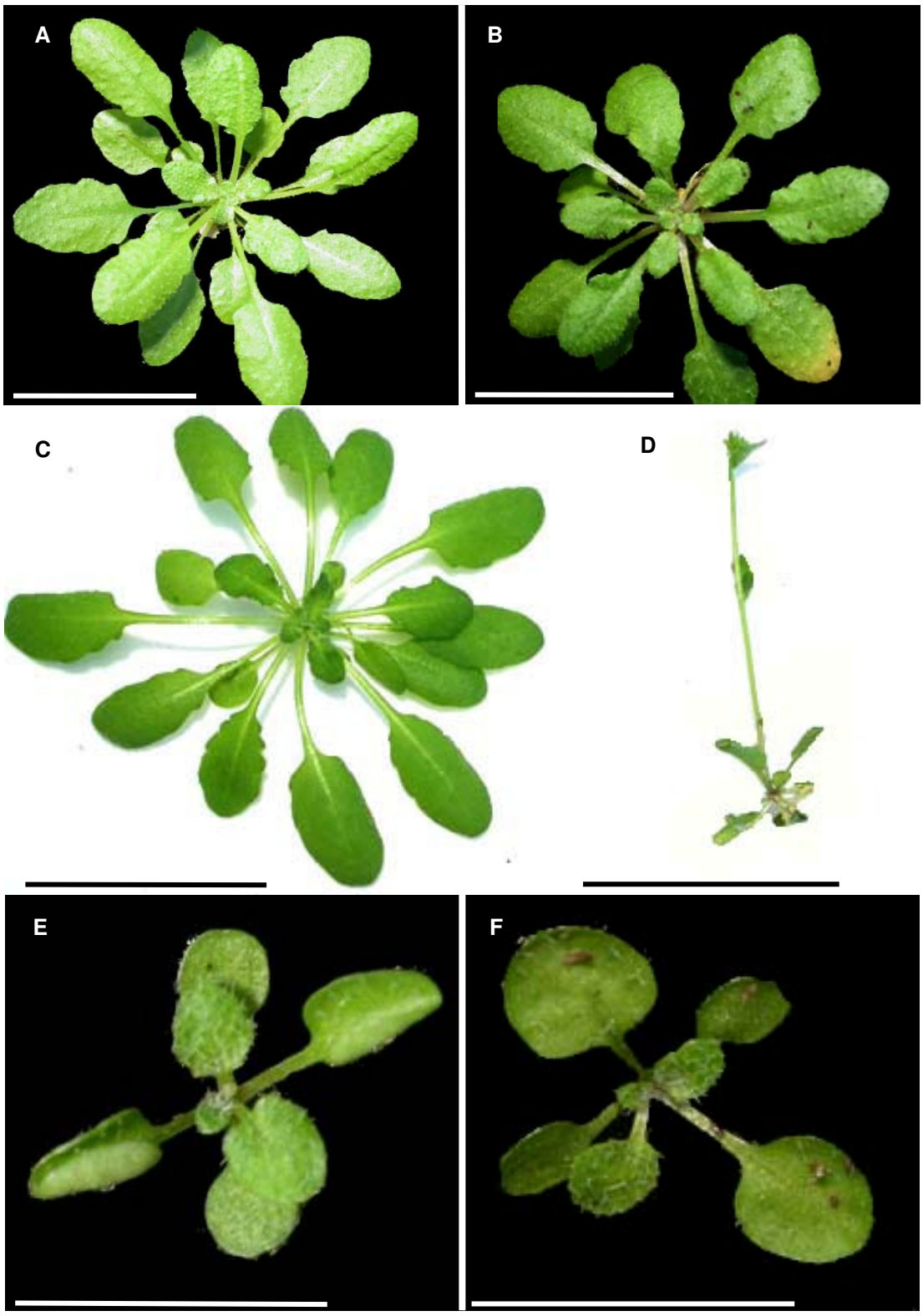

Figure 5.1 *35S::FD* plants grown under SD and LD

(A) *35S::FD* grown in SD has a wildtype phenotype, with no distinguishable differences to (B) wildtype to 25 day old wildtype plants grown in SD. (C) Wildtype plants grown in LD. (D) *35S::FD* grown in LD are dwarfed with curled leaves and early flowering. (E) *35S::FD* phenotype is rapidly induced upon shift from SD to LD. (F) *35S::FD* sibling grown only in SD. Scale bar 1 cm.

plants with the *AP1::GUS* reporter construct, have a wildtype phenotype when grown in SD (Figure 5.2 A and B). Upon shift to LD, *AP1::GUS* is activated in leaves of *35S::FD AP1::GUS* plants (Figure 5.2 D). In 30 day old *35S::FD AP1::GUS* plants expression of *AP1::GUS* is uniform over the leaf and does not differ in young and older rosette leaves (Figure 5.3 A). Activation of *AP1::GUS* in *35S::FD AP1::GUS* plants occurs in uncurled rosette leaves and cauline leaves (Figure 5.3 A). β-glucuronidase activity was analyzed in *35S::FD AP1::GUS* plants 5 days following shift from SD to LD. Activation of *AP1::GUS* was observed upon shift to LD in young and older leaves and flowers. These results reveal that in LD, *AP1::GUS* is activated and expressed in both young and old leaves.

5.iv. Chromatin immunoprecipitation at the *AP1* locus in *35S::FD* plants

To determine whether the activation of *AP1* requires the removal of H3K27me3, ChIP was performed on *35S::FD* plants that had been grown in SD only and SD shifted to LD. As *AP1* is a target of *CLF* and is enriched for the repressive mark, H3K27me3 (Zhang et al. 2007), the rapid activation of *AP1* upon shift from SD to LD in *35S::FD* plants could be associated with changes in histone modifications. ChIP was performed using antibodies against H3K27me3, the histone modification associated with Pc-G mediated silencing and H3K4me3, associated with trxG proteins and gene activation. In this ChIP experiment, DNA was amplified two primer pairs in duplex PCR reactions (as described in Schubert *et al* 2006). One primer pair was specific for *AP1* and the other for actin that is not expected to be enriched. The relative amounts of the two PCR products amplified from the IP DNA were compared to the relative amounts amplified from the input. This corrects for the differences in the efficiency of the two primers or for differing amounts of chromatin. Figure 5.4 shows the result of the ChIP analysis on *35S::FD* plants at the endogenous *AP1* locus. For *35S::FD* plants grown in SD, there is enrichment for H3K27me3 at the *AP1* locus consistent with repressed *AP1* expression in leaves. However, when the plants were shifted from SD to LD we can see that the amount of enrichment for H3K27me3, surprisingly does not change. The results of the ChIP revealed that there were no differences in enrichment for H3K27me3 and H3K4me3 at the *AP1* locus in *35S::FD* grown in SD only and SD shifted to LD (Figure 5.4).

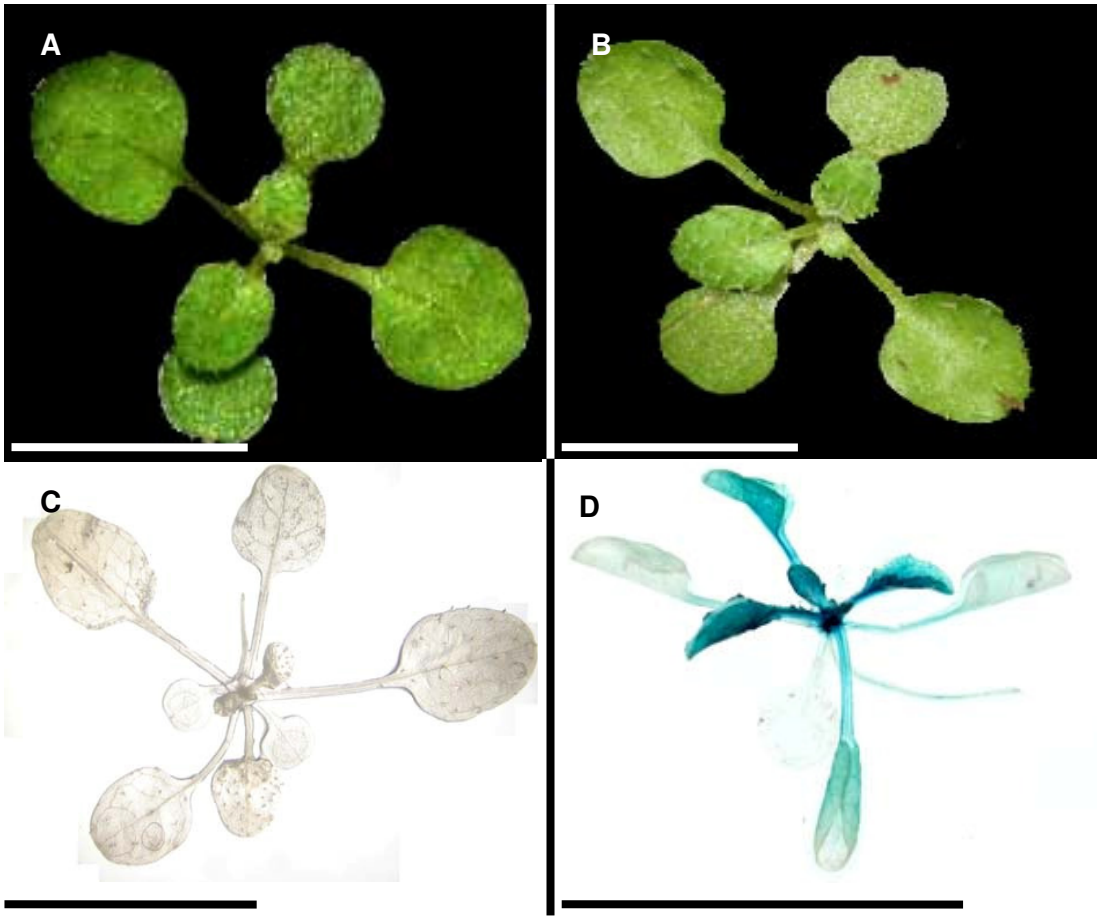

Figure 5.2 *AP1* is activated following shift from SD to LD in *35S::FD* plants

(A) 10 day old *35S::FD* plants carrying the *AP1::GUS* reporter construct grown under SD have a wildtype phenotype. (B) 10 day old wildtype (Landsberg *erecta*) plants grown under SD. (C) Under SD, *AP1::GUS* is not expressed in *35S::FD* plants. (D) Activation of *AP1::GUS* in *35S::FD* plants following 5 days shift from SD to LD. Scale bar 1 cm.

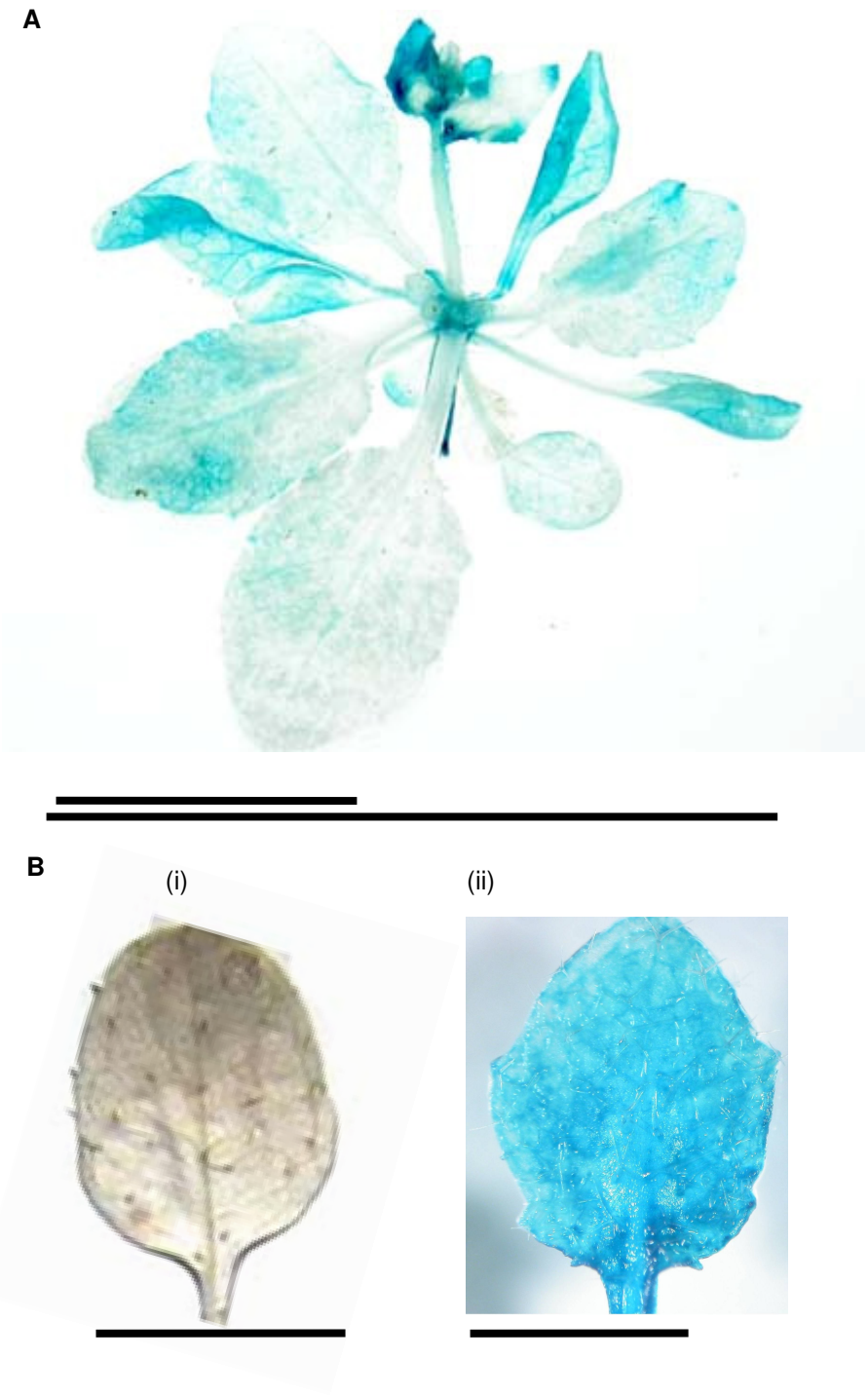

Figure 5.3 GUS expression in *35S::FD* plants carrying the *AP1::GUS* reporter upon shift from SD to LD

(A) 30 day old *35S::FD* plants carrying the *AP1::GUS* reporter construct following 5 days shift from SD to LD showing activation of *AP1::GUS* in older rosette leaves, cauline leaves and throughout the whole plant including flowers. (B) 25 day old *35S::FD AP1::GUS* plants grown in SD (i) and following 7 days after shift to LD (ii) showing that GUS expression is uniform in older leaves. Scale bar 1 cm.

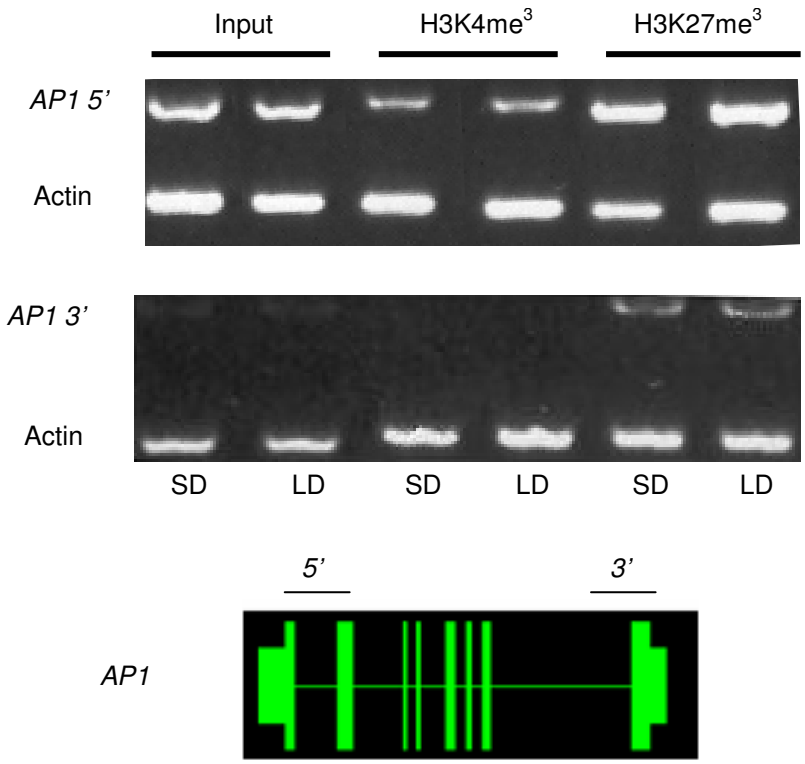

Figure 5.4 Histone modifications at the *AP1* locus before and after shift from SD to LD

35S::FD plants were grown in SD for 7 days post germination then either shifted to LD or left in SD. 10 day old plants from each growth condition were bulked and harvested. ChIP analysis was carried out with antibodies against H3K27me3 and against H3K4me3. Duplex PCR was carried out using primers designed to amplify around 100 bp at the 5' and 3' ends of *AP1* locus (shown on diagram). Primers for Actin and primers for *AP1* were used in the same PCR reaction and the relative efficiency of the two primers sets is established using the input. If there is enrichment of H3K27me3 and H3K4me3 at *AP1,* the relative amounts of the two products should be increased relative to the ratio in the input.

The same ChIP samples were also analyzed using primers for the *AG* locus as experimental controls using a known target gene that is repressed in leaves. These experiments show enrichment for H3K27me3 and no enrichment for H3K4me3 (Figure 5.5).

5.v. Activation of *AP1* in leaves in a *35S::FD* background

ChIP analysis was carried out on *35S::FD* to determine the histone modifications at the endogenous *AP1* locus, but not at the *AP1::GUS* reporter transgene. Activation of *AP1::GUS* was confirmed by analyzing β-glucuronidase activity in *35S::FD AP1::GUS* plants. To confirm that endogenous *AP1* expression is up-regulated following shift from SD to LD, RT-PCR was preformed. Expression of endogenous *AP1* was analyzed in *35S::FD* plants grown in SD and shifted to LD. Figure 5.6 shows that there is faint *AP1* expression in *35S::FD* grown in SD. This might reflect low level of *AP1* in the *35S::FD* background. Yet, *AP1* expression was strongly up-regulated in *35S::FD* plants following shift from SD to LD. These results confirmed what was observed using the *AP1::GUS* reporter construct in *35S::FD*, *AP1* is activated and expressed upon shift to LD.

5.vi. Conclusion

In LD, FT is expressed and interacts with FD, promoting *AP1* expression that converts the SAM to an inflorescence meristem. In SD, *35S::FD* appears wildtype, as FT is not expressed; however, upon shift to LD, activation of *AP1* occurs in leaves, due to the interactions FT and FD. The *35S::FD* phenotype is rapidly induced upon day length shift. Activation of *AP1* in leaves in LD was independently confirmed by analyzing β-glucuronidase activity of the *AP1::GUS* reporter and RT-PCR of endogenous *AP1* expression. To establish whether activation of a Pc-G target requires the removal of H3K27me3, ChIP was performed at the *AP1* locus. ChIP analysis was performed on *35S::FD* either grown in SD only or shifted from SD to LD. Surprisingly; activation of *AP1* did not require the removal of H3K27me3, nor the gain of H3K4me3. These results illustrate that H3K27me3 is not sufficient to silence *AP1*. The activation of endogenous *AP1* by its own transcription factors, FD and FT in leaves, indicates that the *AP1* promoter sequence is accessible to transcription factors despite the presence of Pc-G proteins as well as the repressive mark, H3K27me3.

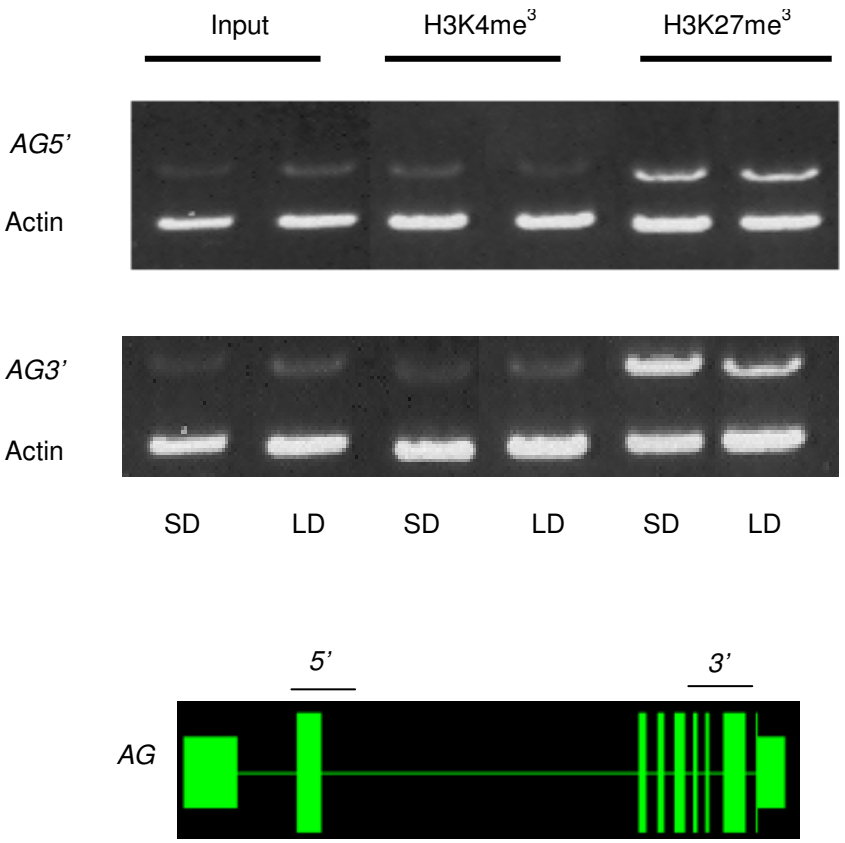

Figure 5.5 Histone modifications at the *AG* locus before and after shift from SD to LD

35S::FD plants were grown in SD for 7 days post germination, and then either kept in SD or shifted to LD. 10 day old plants from each growth condition were bulked and harvested. ChIP analysis was carried out with antibodies against H3K27me3 and H3K4me3. Duplex PCR was carried out using primers designed to amplify around 100 bp at the 5' and 3' ends of *AG* locus (shown on diagram) and actin.

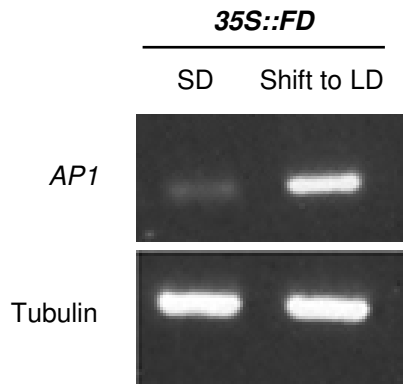

Figure 5.6 Expression of *AP1* native gene before and after shift from SD to LD

35S::FD plants were grown in SD for 7 days post germination and either kept in SD or shifted to LD for 3 days. RNA was extracted from bulked 10 day old seedlings, cDNA was synthesized and semi-quantitative PCR was performed. *AP1* expression was compared to the amplification of Tubulin.

6. Using targeted expression to dissect the roles of *CLF* and *SWN* during plant development

6.i. Introduction

The roles of *CLF* and *SWN* in late leaf and flower development remain unknown due to the compensatory effects of partially redundant Pc-G genes. The severity of *clf- swn-* mutants limits the analysis of loss of Pc-G function phenotypes, as although *clf- swn-* mutants germinate, post-germination development arrests, beginning with the loss of root identity and root-to-callus transformation occurs (Figure 6.1 A, arrow). The callus-like tissue produced can grow indefinitely on sterile tissue-culture, however, no true plant-like organs and differentiated cell types are distinguishable (Figure 6.1).

Microarray analysis, ChIP and ChIP on chip experiments reveal that *CLF* and *SWN* target many genes (Thorpe and Goodrich unpublished, Zhang et al. 2007, Clarenz and Goodrich unpublished). Further understanding of the biological relevance of these candidate target genes and identifying new targets is difficult due to the downstream effects of target gene mis-expression. Promoter::GUS reporter constructs provide evidence to suggest that the loss of *CLF* and *SWN* leads to ectopic expression of floral homeotic genes; *AG, AP1, PI* and *FUSCA3* (*FUS3*) (Figures 4.1 and 6.1). The expression of floral targets that are normally repressed during vegetative development and expressed upon floral induction indicates that one of the developmental roles of Pc-G is to repress floral development. Furthermore, ChIP analysis revealed targets of *CLF* included genes involved in leaf polarity, branching and stomatal development, suggesting many novel roles for Pc-G proteins, however, the roles of *CLF* and *SWN* in leaves cannot be seen as *clf- swn-* mutants do not have leaves.

The aim of this chapter is to create two approaches to gain insight into the developmental roles of Pc-G proteins in plants, particularly in late leaf and flower development. This chapter

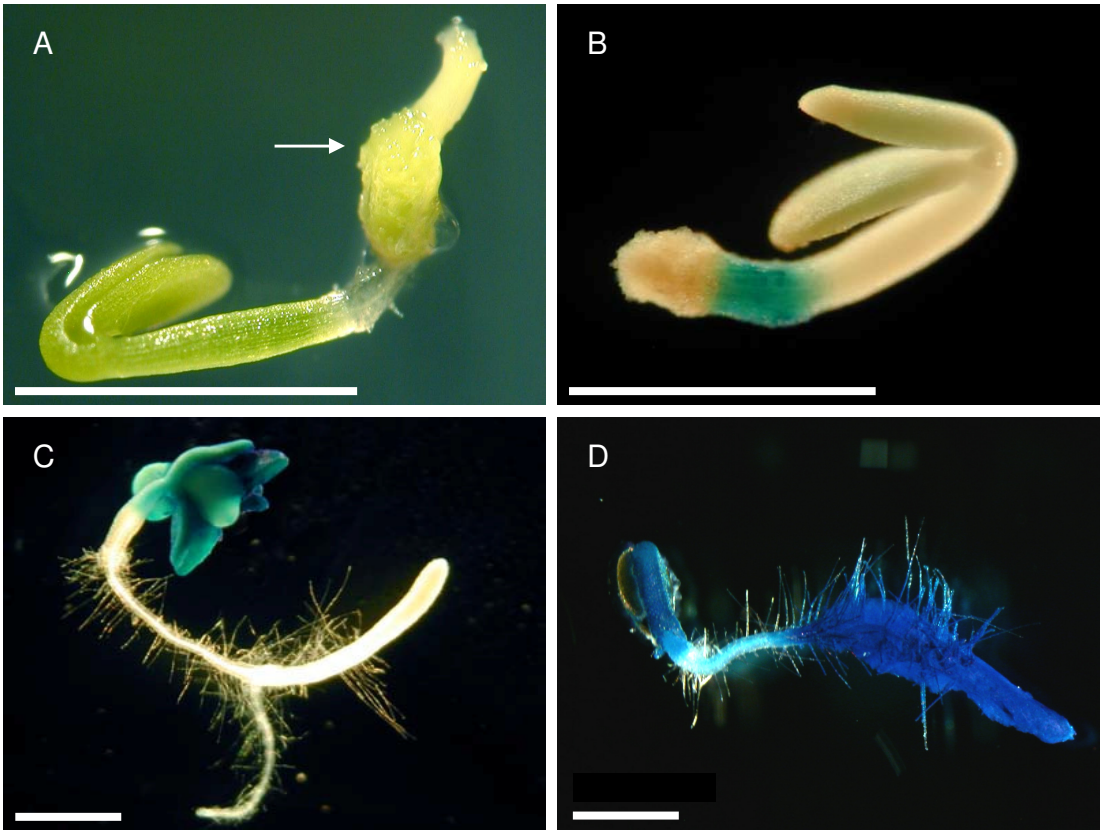

Figure 6.1 Ectopic expression of several target genes in *clf- swn-* mutants

AP1::GUS, *PI::GUS* and *FUS3::GUS* reporter transgenes were introduced into a *clf- swn-* background. These genes are not expressed in wildtype seedlings (data not shown). (A) *clf- swn-* mutants germinate, however, post-germination development arrests and root-to-callus transformation occurs, marked by arrow. (B) Ectopic expression of floral homeotic gene *AP1* in the root of *clf- swn-*. (C) Ectopic expression of floral homeotic gene *PI* in the leaf-like organs of young *clf- swn-*. (D) Ectopic expression of the seed storage gene *FUS3* throughout the whole of *clf- swn-*. Scale bar 1 cm. From Clarenz and Goodrich, unpublished.

describes the cloning and characterization of two approaches developed to target *CLF* expression to dissect the roles of *CLF* and *SWN* during plant late leaf and flower development. The first approach described in this chapter is to provide spatial and temporal control of *CLF* by introducing *CLF* into the *LhG4/ pOp* bipartite transactivation system. The second approach described is to provide clonal analysis of *CLF* by introducing *CLF* into the *CRE lox* recombination system.

6.ii. Approach I: Transactivating *CLF* to complement *clf-2* and *clf-2 swn-7*

Approach I provides spatial and temporal control of *CLF+* activity in *clf- SWN+* and *clf- swn-* backgrounds. In order to establish whether providing *CLF+* activity could complement *clf-2* and *clf-2 swn-7*, *CLF* was introduced into the *LhG4/ pOp* bipartite transactivation system (Moore et al. 1998) (Figure 6.2 A). *LhG4* driver lines of the promoters *pSHOOTMERISTEMLESS* (*pSTM::LhG4*) and *pKNOTTED*-like from *Arabidopsis thaliana* 2 (*pKNAT2::LhG4*) were used to transactivate *CLF* early in development. The *p650::LhG4* driver line will be used to transactivate *CLF* in late leaf development in *clf-2* and *clf-2 swn-7* mutants.

To transactivate *CLF* using the *LhG4/ pOp* transactivation system*, Lac Op* sites were cloned upstream of *CLF* cDNA inserted as a *Kpn*I and *BamH*1 fragment (Figure 6.2 B). *pOp::CLF* was cloned as a *Not*I fragment inserted into the binary vector *pMLBART* that confers BASTA® (phosphinothricin-tripeptide (PTT)) herbicide resistance in plants (Figure 6.2 C). Due to the low fertility of *clf-2* homozygotes, *pMLBART::pOp::CLF* was introduced into *clf-2/+* heterozygotes by floral dip transformation. As *clf-2 swn-7* mutations are lethal on soil, *pMLBART::pOp::CLF* was introduced into *clf-2/+ swn-7* by floral dip transformation. Primary transformants were selected on BASTA® (PTT) herbicide resistance and independent seed lines were harvested. T2 seed was selected on hygromycin and BASTA® (PTT) herbicide and *clf-2* homozygotes were selected by phenotype on soil. *pMLBART::pOp::CLF* did not complement *clf-2* and *clf-2 swn-7* phenotypes. T3 seed was selected on BASTA® (PTT) herbicide to identify T2 plants that were homozygous for *pMLBART::pOp::CLF*. To provide

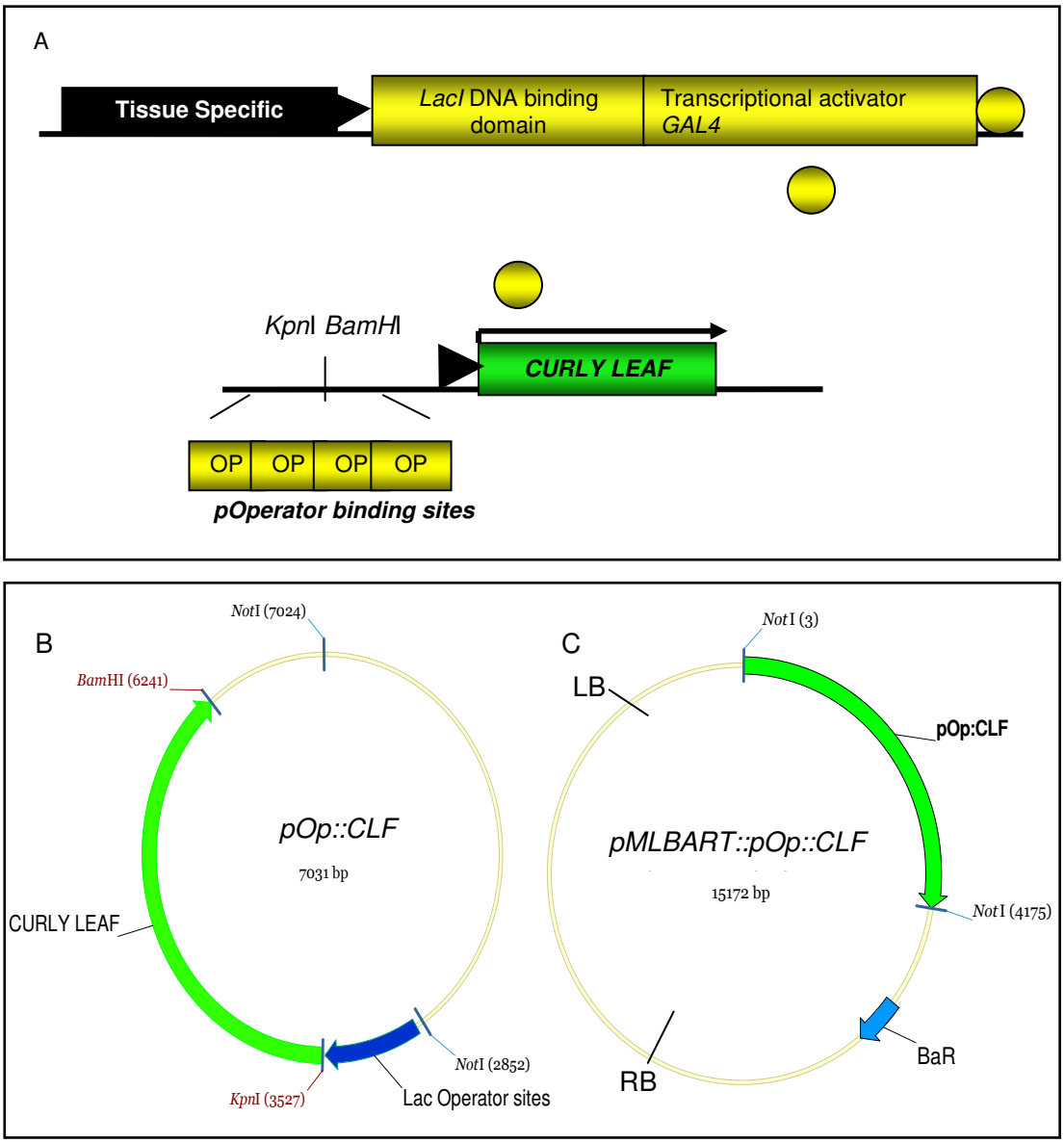

Figure 6.2 The *LhG4/ pOp* transactivation system and *pOp:CLF* and *pMLBART:pOp:CLF* transgene maps

(A) The *LhG4* transactivation system is a two component system. The driver line is comprised of a tissue specific promoter fused to the *LacI* DNA binding domain that is transcriptionally fused to the *GAL4* transcription-activation domain-II from *Saccharomyces cerevisiae*. Plants carrying this transgene are crossed to plants containing *pOp::CLF*. (B) The *pOp::CLF* transgene contains *Lac Op* sites cloned upstream of *CLF* cDNA inserted as a *Kpn*I and *BamH*1 fragment. (C) *pOp::CLF* was cloned as a *Not*I fragment inserted into the binary vector *pMLBART* that confers BASTA® (PTT) herbicide resistance in the transgenic plants.

CLF+ transiently early in plant development, the *pSTM::LhG4* driver line was chosen to transactivate *CLF*. The *pKNAT2::LhG4* driver line was chosen to transactivate *CLF* to provide *CLF+* activity in the peripheral zone of the shoot apical meristem, and inflorescence stems. In order to test whether providing *CLF+* continuously during late leaf development would complement *clf-2* and *clf-2 swn-7* mutants, the *p650::LhG4* driver line was chosen to transactivate *CLF*.

pSTM::LhG4, *pKNAT2::LhG4* and *p650::LhG4* homozygous lines were crossed to *clf-2* and *clf-2/+ swn-7*. F1 progeny were selected on hygromycin and BASTA® (PPT) herbicide and seeds were collected. F2 progeny were selected on hygromycin and BASTA® (PTT) herbicide, *clf-2* phenotypes were seen on soil. F3 progeny of the independent F2 lines were selected on BASTA® (PTT) herbicide to identify lines that were homozygous for the *LhG4* driver line. These plants were then crossed to *clf-2*/+ plants homozygous for *pMLBART::pOp::CLF*. The F1 progeny of this cross were selected on hygromycin and BASTA® (PTT) herbicide.

To confirm the expression of the *LhG4* driver lines, *pSTM::LhG4*, *pKNAT2::LhG4* and *p650::LhG4* were crossed to *pOp::GUS*. Figure 6.3 A shows β-glucuronidase activity in a *pSTM::LhG4 pOp::GUS* seedling, confirming *pSTM* is expressed in the shoot apical meristem (SAM) and is not expressed in leaves. Figure 6.3 B shows β-glucuronidase activity in a *pKNAT2::LhG4 pOp::GUS* 14 day old plant, confirming *pKNAT2* is expressed in the SAM. For the confirmation of the late leaf promoter *p650::LhG4*, refer to Chapter 4, Figure 4. 4 for the *p650::LhG4* driver line expression.

6.iii. Approach II: Introducing *CLF* into the *CRE lox* recombination system

Approach II was chosen in order to generate mosaic plants. Mosaic analysis has been little applied to address epigenetic questions in plants. As plant cell fate appears to be less rigidly determined compared to animal cell fate, it is relevant to address whether plant Pc-G proteins have similar mitotic stability and autonomy of action as animal Pc-G proteins.

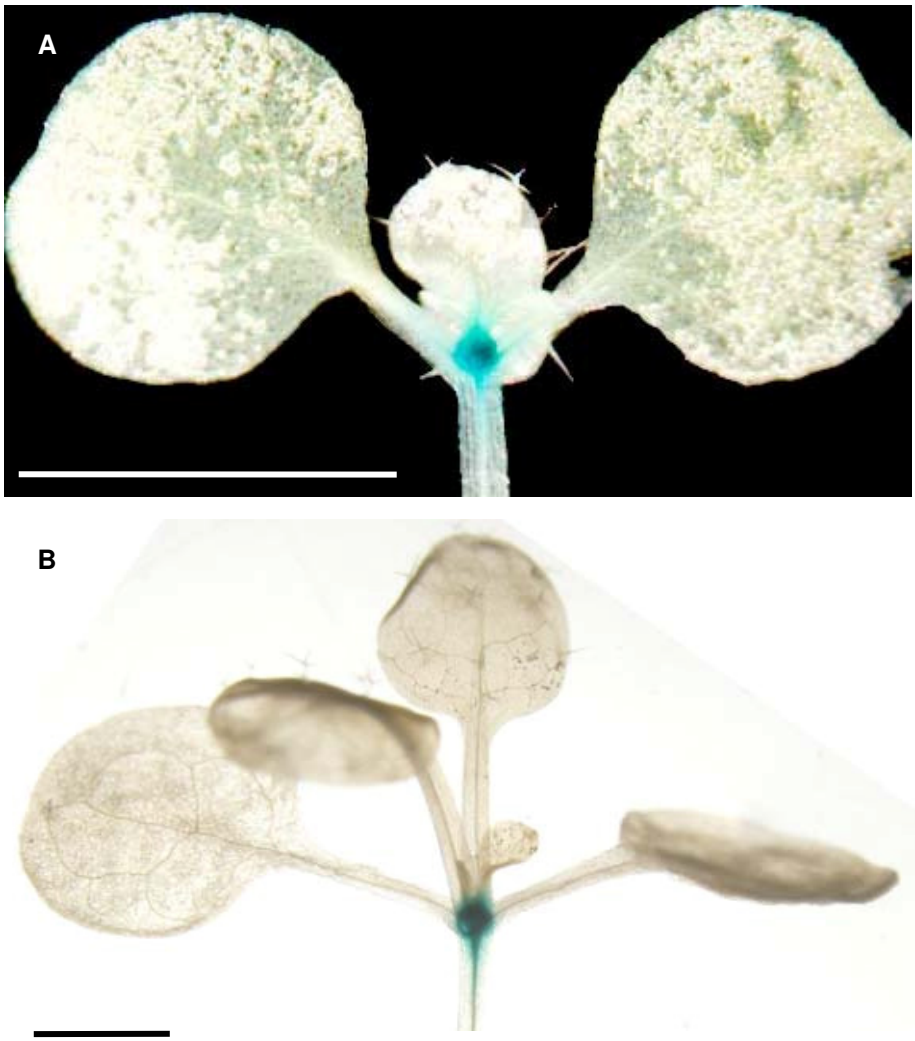

Figure 6.3 *pSTM::LhG4* and *pKNAT2::LhG4* expression in the SAM

pSTM::LhG4 and *PKNAT2::LhG4* driver lines were crossed to the reporter transgene *pOp::GUS*. (A) F2 seedlings show *pSTM::LhG4* is expressed in the SAM and not in the leaves. (B) 25 day old F2 plant showing *pKNAT2::LhG4* is expressed in the SAM and not expressed in leaves.

Approach II differs to approach I as *CLF+* activity is precisely controlled by heat-shock. The stochastic nature of *CRE* recombinase provides the opportunity to dissect *CLF* action in organs and cell types that have lost *CLF+* activity, *clf-* cells will be surrounded by *CLF+* cells in the same tissue and organs. The expression of GFP_{ER} allows us to use florescence microscopy to identify cells in which *CRE lox* recombination events have occurred by.

In order to create marked clonal sectors that are *clf-* in a *CLF+* background, *CLF* was introduced into the *CRE lox* recombination system. The 7.5 Kb genomic fragment that spans the *CLF* locus is sufficient to complement *clf-* mutants (Schubert et al. 2006) and therefore, has all the necessary regulatory sequences for *CLF* function, was amplified by PCR using high fidelity Phusion® DNA polymerase, was cloned into pGEMTeasy® and cloned as a *Not*I insert into the binary vector *pCB1*. The *pCB1* binary vector contains the recombination cassette consisting of a *CRT1* conferring resistance to kanamycin resistance flanked by two bacteriophage *lox* repeats (Figure 6.4). This separates and prohibits the *35S* constitutive promoter from inducing *GAL4VP16* transcription and transactivating GFP_{ER} (Figure 6.4 A).

6.iv. *pCB1::CLF* complements *clf-28* and *clf-28 swn-7* mutants

Due to the low fertility *clf-28* homozygotes and the lethality of *clf-28 swn-7* homozygotes on soil, *pCB1::CLF* was introduced into *clf-28/+* and *clf-28/+ swn-7* plants by floral dip transformation. Primary transformants were selected on BASTA® (PTT) herbicide to select *pCB1::CLF* transgene and independent plants were harvested. T2 seed of independent transgenic lines were selected on kanamycin for *clf-28*. *pCB1::CLF* complemented the *clf-28* phenotype (Figure 6.5) therefore *clf-28* homozygotes cannot be identified by phenotype. T3 lines showing 100 % kanamycin resistance, therefore homozygous for *clf-28*, were scored for *clf-28* phenotype to identify lines carrying a single locus insertion of *pCB1::CLF*.

6.v. Creating *clf-* deletion sectors by heat-shock induction

In order to create *clf-* sectors, *pG7HSCREII* in a Columbia-0 background was crossed to *clf-28* and *clf-28/+ swn-7.* The *pG7HSCREII* transgene contains the *CRE* recombinase driven under the *Arabidopsis HSP18.2* heat-shock promoter. Upon heat-shock, *CRE* recombinase

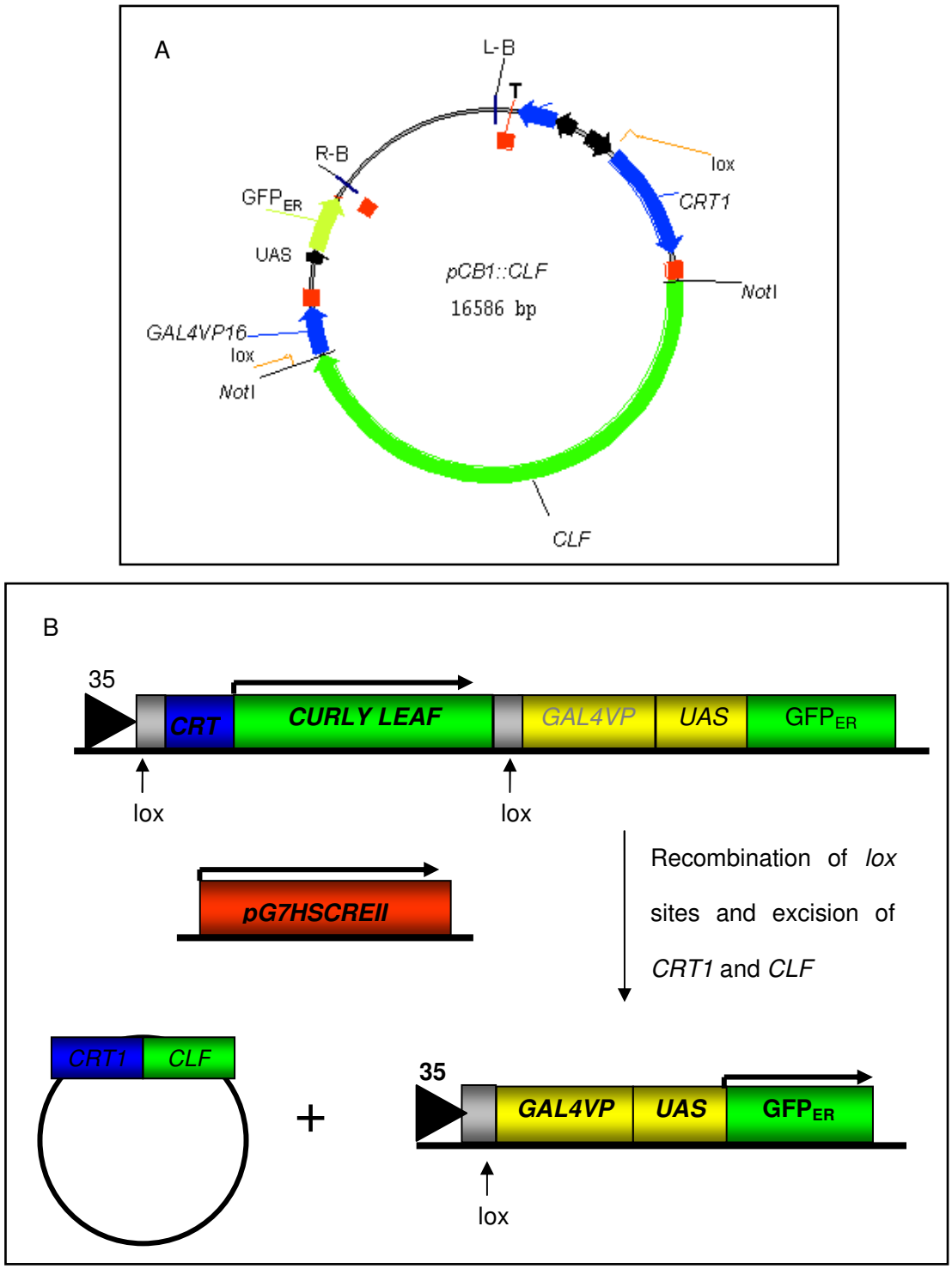

Figure 6.4 *pCB1::CLF* construct and the *CRE lox* recombination system used to create *clf-* cells

(A) The 7.5 Kb full length genomic *CLF* was cloned into *pCB1* inserted at a *Not*I site. *pCB1::CLF* contains the recombination cassette consisting of the *CRT1* resistance gene flanked by two bacteriophage *lox* repeats. This stops the *35S* constitutive promoter from inducing *GAL4VP16*. (B) Upon heat-shock the *CRE* recombinase mediates stochastic *lox* recombination and two products are created. A circular piece of DNA with the excision product, *CRT1* and *CLF*. The *35S* promoter drives *GAL4VP16* and activates GFP_{ER} Cells that lose *CLF+* activity are marked by GFP_{ER}.

Figure 6.5 *pCB1::CLF* complements *clf-28* mutants

The progeny of primary transformants (that had been selected on BASTA® (PPT) herbicide resistance), were selected on kanamycin. T2 lines that conferred 100% kanamycin resistance (homozygous for *clf-28*) were transferred to soil, grown in LD at 24 ℃ and scored for the wildtype phenotype, showing complementation by *pCB1::CLF*. (A) Complementation of homozygous *clf-28* by *pCB1::CLF* and (B) sibling homozygous *clf-28* that does not contain the *pCB1::CLF* transgene. Scale bar 1 cm.

mediates stochastic *lox* recombination excising *CLF* from *pCB1::CLF* resulting in the loss of *CLF+* activity (Figure 6.4 B). The *35S* constitutive promoter then drives *GAL4VP16* expression that in turn activates expression of endoplasmic reticulum ($_{ER}$)-localised green fluorescent protein (GFP_{ER}). Cells that have lost *CLF+* activity are marked by GFP_{ER} and easily identified by fluorescence microscopy.

6.vi. Conclusion

Novel roles for *CLF* and *SWN* in leaf and flower development are difficult to address due to the severity of *clf- swn-* mutants. Two approaches were created to further the current understanding of the roles of *CLF* and *SWN* in leaf and flower development. *CLF* was introduced into the *LhG4/ pOp* transactivation system in order to transactivate *CLF* early and late in plant development in a Pc-G null background. The CRE lox recombination system was designed to generate GFP_{ER} marked sectors that lose *CLF+* activity in *clf-28* and *clf-28 swn-7* mutant backgrounds. It is hoped that clonal analysis in leaves and flowers will provide insight into the development roles of *CLF* and *SWN*. Due to time limitations, these systems were not further characterized in this thesis.

7. Discussion

7.i. Main aims of this thesis

The main aims of this thesis were to;

I) Investigate the role of *CLF* in regulating flowering time in natural populations of *A. thaliana*.

II) Determine whether Pc-G mediated silencing is reversible, testing i) if Pc-G mediated silencing could be overcome during late leaf development and ii) whether transient activation of *AG* in leaf primordia is epigenetically stable.

III) Address whether the activation of a Pc-G target, *AP1*, required the removal of H3K27me3, testing whether H3K27me3 mediated by Pc-G complexes is sufficient to silence.

IV) To develop two approaches to dissect the roles of *CLF* and *SWN* in leaf and flower development by targeted expression.

7.ii. A polymorphism in *CLF* does not regulate flowering time in natural populations of *A. thaliana*

Through genome-wide association mapping, a polymorphism in *CLF* was identified with a high pairwise haplotype score (a system that based on the estimation of the shared length around a polymorphism compared between individual haplotypes) suggesting it could be under natural selection. *A. thaliana* accessions that contain the ThrGly polymorphism in *CLF* showed a reduced SD to LD flowering time ratio, indicating that they are less sensitive to photoperiod. We were able to confirm that the Edinburgh-0 accession, that contains the ThrGly insertion, has a reduced SD to LD flowering time ratio (expressed as a ratio between the number of days to flowering under SD to LD) under standard growth conditions (Figure 3.2). *CLF* alleles were cloned from Columbia-0 and Edinburgh-0 and compared in an isogenic background, which was homozygous for *clf-28*, a null allele of *CLF.* Our data did not support the hypothesis that the flowering time phenotype was caused by the ThrGly polymorphism in *CLF*.

There is considerable natural variation in flowering time in *A. thaliana,* mostly due to allelic variation in *FRI* (Johanson et al. 2000). As previously mentioned in the Introduction, *FRI* regulates flowering by promoting *FLC* expression. *FRI+* alleles are late flowering (due to *FLC* expression), whereas, *fri-* alleles are early flowering and have been reported to be under selection (Johanson et al. 2000). The Columbia-0 accession is *fri-* and as such the different alleles of *CLF* were compared in a *fri-* background. It is of interest to note that the accessions carrying the ThrGly insertion are in a *FRI+* background. In *FRI+* backgrounds, it is plausible to consider that the need for epigenetic silencing of floral target genes could differ to that in a *fri-* background due to the need of vernalization. Shindo *et al* reported variation in the vernalization response showing that accessions required different periods of cold to ensure *FLC* silencing (Shindo et al. 2006). *FRI+* accessions require long vernalization, including Edinburgh-0. Although, *clf-28* is in Columbia-0, a *fri-* accession, it shows a vernalization effect, flowering earlier with 4 weeks vernalization treatment (J. Goodrich, unpublished data). It would be of interest to repeat this investigation in a *FRI+* background either, by inserting a *FRI+* allele into *clf-28* or using a *clf-* homozygous null in a *FRI+* background when such becomes available.

7.iii. Pc-G mediated silencing of *AG* is reversible

The floral homeotic gene *AG* is regulated by *CLF*, being repressed in leaves and expressed in stamens and carpels of flowers (Goodrich et al. 1997). In *clf-* mutants, silencing of *AG* is lost and *AG* becomes activated in leaves and organs (Figure 4.1 D and F). The mechanisms involved in maintaining and resetting Pc-G mediated silencing are not fully understood. *pOp:AG-I::GUS* is repressed in leaves by *CLF*, mimicking endogenous *AG* expression. Under the late leaf driver line, *pOp:AG-I::GUS* is activated in rosette leaves, petioles, flowers and expressed strongly in the vasculature, particularly the main veins in rosette leaves (Figure 4.4). Yet, the late leaf driver line *p650::LhG4* is expressed uniformly in older rosette leaves (Figure 4.4). As *pOp:AG-I::GUS* was not activated in the full expression domain of *p650::LhG4*, this suggests that Pc-G mediated silencing in different cell types is either independent of other cell types or regulation of target gene expression differs in cell types. In

clf- mutants ectopic expression of *AG* is mainly in the major veins in older leaves (Goodrich et al. 1997). We suspect that this could reflect where the floral promoter, FT. FT is strongly expressed in the vasculature and together with the *p650::LhG4* driver line, this may activate *pOp:AG-I::GUS*. The *LhG4* transactivation system could provide the opportunity for proteins to access *AG*, which could block CLF activity. If CLF is no longer present, the mechanism that prevents transcription factors gaining access to *AG* would be lost, thereby allowing gene activation. It would be necessary to grow *pOp:AG-I::GUS p650::LhG4* lines in SD, where *FT* is not expressed, to confirm whether FT is necessary for *p650::LhG4* to reverse Pc-G silencing of *pOp:AG-I::GUS*. To establish if CLF is lost from *AG* when transactivated under the *LhG4* trans-factor, ChIP analysis could be performed.

Activation of *pOp:AG-I::GUS* under the early leaf promoter *pANT::LhG4* reveals that Pc-G silencing is reversible in plants, however, transient activation is not mitotically stable in the absence of the *LhG4* trans-activator (Figure 4.5). *pANT::LhG4* is not expressed in older leaves (Figure 4.5). β-glucuronidase activity was analyzed in *pOp:AG-I::GUS pANT::LhG4* plants, surprisingly showing that activation occurs throughout the *pANT::LhG4* expression domain (Figure 4.5). This indicates that Pc-G mediated silencing is overcome and reversible early in development, but transient activation of a Pc-G target is not mitotically stable in the absence of the inducer. These results suggest that the *LhG4/ pOp* transactivation system can overcome Pc-G mediated silencing of *AG* during early leaf development, yet, Pc-G silencing is re-established in the absence of the inducer in older leaves. To propose that *CLF* re-establishes silencing of *pOp:AG-I::GUS* in leaves when activated in leaf primordia, the system must be compared in a *clf-* background. However, as *pOp:AG-I::GUS* is activated in *clf-2* mutants, it would be difficult to distinguish the β-glucuronidase activity of *pOp:AG-I::GUS* activated by *pANT::LhG4* and activation of *pOp:AG-I::GUS* due to the loss of *CLF+* activity. Transient activation of *pOp:AG-I::GUS* could be further investigated using the Dexamethasone inducible *LhGR/ pOp* transactivation system. This would allow transient activation to be controlled at any stage of plant development through the application and removal of steroid. *LhGR* lines have been crossed to *pOp:AG-I::GUS* to test whether

transient activation of *pOp:AG-I::GUS* can lead to stable activation following withdrawal of the steroid. However, the limitations of this system include the uneven uptake of steroid and the persistence of the steroid would lead to variation in the results.

The stable activation of Pc-G target *AG* illustrates Pc-G silencing is reversible. Is this activation of *AG* associated with changes in mitotically heritable histone modifications? Chromatin remodelling has been defined as 'any event that alters the nuclease sensitivity of a region of chromatin' (Aalfs & Kingston 2000). Compacted chromatin might result in blocking transcription factors to DNA, associated with the 'closed' chromatin conformation. This model appears over simplistic and there has been very little evidence to support it. A previous study by McCall *et al* suggests Pc-G mediated silencing does not exclusively block all DNA-binding proteins (McCall & Bender 1996). In *Drosophila* Pc-G mediated silencing inhibited *GAL4*-dependent transcription, however it did not inhibit the bacteriophage T7 RNA polymerase (T7RNAP) that was used to probe DNA accessibility in the Bithorax complex (McCall & Bender 1996). Pc-G mediated silencing could be specific and inhibit promoter and transcription factor interactions, or block trxG proteins gaining access to target genes.

Previous work by Schubert *et al* (Schubert et al. 2006) using a Dexamethasone inducible *CLF* transgene (*pCLF::CLF-GR*) demonstrated that the removal of *CLF* late in development resulted in the stable activation of *AG*. This illustrates that activators of *AG* are present in leaves but *CLF* normally prevents them from activating *AG* there. Therefore, activators that are normally present in leaves are unable to gain access to *AG* in leaves, yet the *LhG4* trans-factor is. To deduce whether Pc-G repression is easily overcome by the high levels of expression of the artificial *LhG4* trans-factor, we can envisage the 5 x *Lac Op* sequences are bombarded by *GAL4* molecules and varying amounts are able to actually gain access to the *Lac Op* sites and transactivate *AG* (Figure 7.1). Pc-G mediated silencing is associated with covalent histone modifications believed to remodel chromatin, rendering transcription sites inaccessible to transcription factors (Ringrose & Paro 2004). Yet, the activation of *pOp:AG-I::GUS* by *LhG4* trans-factor indicates that *Lac Op* sites are accessible to *GAL4* molecules.

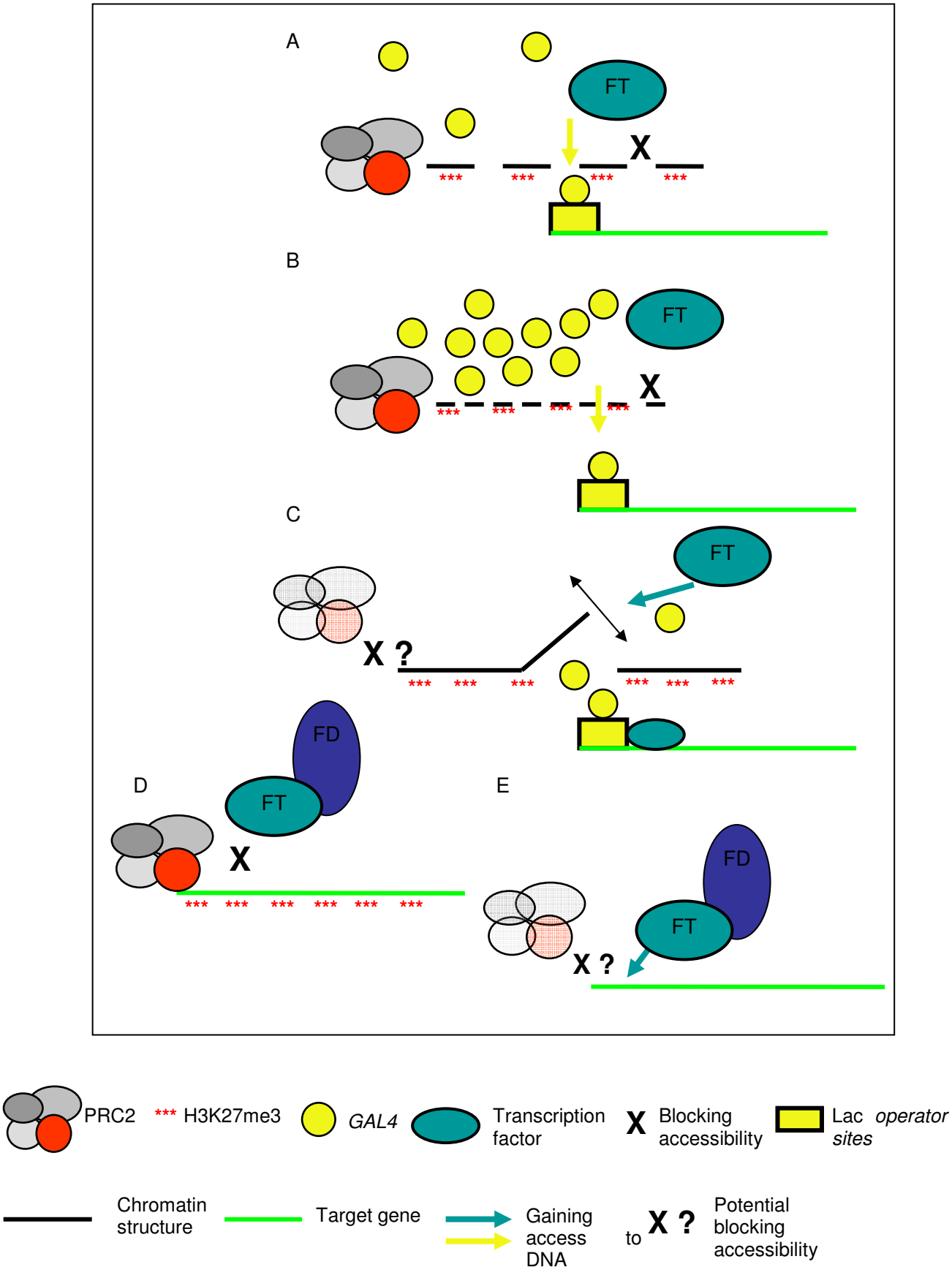

Figure 7.1 Model for overcoming Pc-G mediated silencing

Figure 7.1 Model for overcoming Pc-G mediated silencing

(A) PRC2 mediated chromatin remodelling leads to chromatin conformations in which transcription factors are not able to gain access, but smaller proteins, for example *GAL4* can. (B) Pc-G mediated silencing renders target genes inaccessible, but high concentration of protein (for every 10 molecules, 1 can gain access) overcoming silencing. (C) *GAL4* gain access either when PRC2 complex is reset during mitosis, or by blocking the accessibility of PRC2 to re-establish the repressive mark, thereby allowing transcription factors to gain access. (D) Target gene silencing is maintained in a silent state by Pc-G directed H3K27me3 marks. (E) Transcription factors gain access to Pc-G marked sequences when the epigenetic code is being reset, for example during S-phase. The binding of transcription factors during this time prevents the Pc-G marks being re-established, resulting in loss of silencing.

Accessibility of *GAL4* to *Lac Op* sites would need to be quantified in order to establish the ease in which *GAL4* molecules bind to the *Lac Op* sites. The genomic location of the transgene *pOp:AG-I::GUS* within the genome could also be confirmed, to compare the higher order structure of the sequence to that of endogenous *AG*, presumably, heterochromatin. For this, ChIP analysis could be performed using antibodies against histone H3, with primers specific for the transgenic *AG* and endogenous *AG*.

Transient activation of *pOp:AG-I::GUS* in leaf primordia was not mitotically stable in the absence of *pANT::LhG4* inducer. We hypothesize that H3K27me3 is not removed when a Pc-G target is activated (based on the results of Chapter 5) and that Pc-G silencing is overcome, but upon the removal of *LhG4* trans-factor, silencing is reset, assisted by the presence of H3K27me3. Alternatively, H3K4me3 is recruited to *AG* and either competes with or removes H3K27me3 (Schatlowski et al. 2008), then in contrast to popular belief the active state may not be mitotically heritable. To address whether activation of a Pc-G target results in the loss of H3K27me3 and/ or gain of H3K4me3, the histone dynamics at *pOp:AG-I::GUS* could be characterized by ChIP analysis in both wildtype and *clf*- mutant backgrounds. Further questions to consider are; once Pc-G silencing is re-established following transient activation, are histone modifications reset to the same as in a continually maintained Pc-G target? The current view of Pc-G and trxG silencing associated with histone modifications being epigenetically maintained may no longer be accurate.

7.iv. Activation of Pc-G target, *AP1*, does not require the removal of H3K27me3

In LD, the floral promoter FT is expressed and interacts with FD, regulating the expression of floral meristem gene, *AP1*. The enhancement of the *35S::FD* phenotype upon shift from SD to LD illustrates that FT can interact with FD, not only in the inflorescence meristem, but when FD is expressed throughout the whole plant. Target genes that are normally confined to the floral meristem become activated in leaves in *35S::FD* background when shifted to LD. *AP1* is maintained in a silent state throughout vegetative development by Pc-G directed

H3K27me3. How FT and FD activate the floral meristem gene *AP1* is unknown. The activation of endogenous *AP1* under its own transcription factors FT and FD provided the opportunity to address whether activation of a Pc-G target required the removal of H3K27me3.

AP1 is a target of *CLF* and is enriched for H3K27me3 (Zhang et al. 2007) and becomes activated in *35S::FD* plants in LD (Figure 5.2 and 5.6). The activation of *AP1* occurs in the presence of *CLF*, in agreement with the conclusions drawn from the activation of *AG* discussed in Chapter 4. ChIP analysis at the endogenous *AP1* locus, was performed in *35S::FD* upon shift from SD to LD, revealed that H3K27me3 was not lost upon shift to LD, nor accompanied by the gain of H3K4me3. The activation of *AP1* appears to be independent of its histone modifications, revealing that the activation of a Pc-G target does not require the removal of H3K27me3. This suggests that this could be a general mechanism in which histone modifications are independent to gene transcription if transcription factors are present. This highlights the flaws and simple nature of the proposed model for covalent histone modifications epigenetically maintaining gene expression states by modifying chromatin to form 'open' and 'closed' conformations. The accessibility of transcription sites of a Pc-G silenced target gene may not be as restricted as previously thought. (Figure 7.1)

It would be necessary to repeat the ChIP on *35S::FD* following shift to LD over a longer time course, to test whether histone modifications are passively removed over many cell divisions. If specific regions of H3K27me3 are lost, transcription factors that are present can gain access to the transcriptional start site, yet H3K27me3 is still present at the *AP1* locus. This hypothesis may explain how activation of a Pc-G target can occur. Also, we could use antibodies against heterochromatin and histone H3 in order to test whether the chromatin state alters, allowing accessibility to transcription factors. Profiling histone enrichment spanning the entire gene would also be important to further understand the relationship between relative enrichment of histone modifications and gene transcription.

The activation of a Pc-G target under endogenous transcription factors does not require the removal of H3K27me3. This implies H3K27me3 is not sufficient for silencing. We predict that because *AP1* is stably activated but retains H3K27me3, this may facilitate silencing of *AP1* when *35S::FD* plants are returned to SD. In order to confirm this, *35S::FD* plants would need to be shifted back to SD and *AP1* expression quantified. Analyzing β-glucuronidase activity in *35S::FD AP1::GUS* plants following shift to SD after LD, would not correlate with expression of *AP1*, due to the persistence of the GUS protein. To test whether *AP1* is stably silenced following the shift back to SD, RT-PCR of *AP1* over a time course would need to be performed in order to confirm resetting *AP1* silencing. However, it would be difficult to ascertain whether this is due to the presence of H3K27me3 or the absence of FD/ FT interactions. The results of Chapter 5, as well as Chapter 4 illustrate that Pc-G mediated silencing is reversible, indicating Pc-G silencing is not as stably inherited as once believed, at least in these circumstances.

The activation of endogenous *AP1* by its own transcription factors, FD and FT in leaves, indicates that the *AP1* promoter sequence is accessible despite the presence of Pc-G proteins as well as the repressive mark, H3K27me3. One hypothesis that could explain these results could be associated with the accessibility of the promoter during the cell cycle. Taken together, the fact that Pc-G proteins are lost during the cell cycle and histones are randomly distributed following cell division (Schubert et al. 2006), it is possible transcription factors gain access prior to Pc-G silencing being re-established following cell division. Given that activation of Pc-G targets does not require the removal of H3K27me3 indication that the histone mark itself may not be sufficient to silence.

Pc-G mediated silencing in plants could be similar to that in animals, whereby H3K27me3 recruits other chromatin modifiers to target genes, promoting stable target gene silencing. Although, plant PRC1 complexes have not yet been identified, it is possible that other proteins could fulfill the function. In this view, H3K27me3 is still necessary to promote silencing by establishing target gene recognition.

7.v. Using targeted expression to dissect the roles of *CLF* and *SWN* during leaf and flower development

The two approaches have been designed to enable the dissection of the roles of *CLF* and *SWN* late in leaf and flower development. By introducing *CLF* into the *LhG4* transactivation system and *CRE lox* recombination system provides the opportunity to identify new targets and novel developmental roles for *CLF* and *SWN*. These systems provide the opportunity to address whether Pc-G mediated silencing has similar mitotic stability and autonomy of action in plants as in animals. This has not been carried out due to time restraints. However, the systems created and discussed in Chapter 6 will provide the Goodrich group with the necessary tools to characterise new phenotypes and associated with the loss or gain of *CLF*+ activity.

The *CRE lox* recombination system has previously been demonstrated to be effective analysis of roots (Heidstra et al. 2004). It is easier to identify sectors in uncovered transparent cell types like in roots, unlike leaves and flowers. The advantages of using the *CRE lox* recombination system is the generation of marked sectors that lose *CLF* activity at different developmental stages. The *clf*- sectors created will be marked by GFP_{ER} allowing visualization of cells that have lost *CLF+* activity. Also, when *pCB1::CLF* is excised, all *CLF*+ activity is lost, whereas RNAi and steroid-inducible systems can be permeable and one cannot be sure that all *CLF+* activity has been removed. It will be necessary to distinguish plants that have a single locus insertion of *pCB1::CLF* as well as a single copy at the single locus insertion site. Due to stochastic recombination excision events, it is unlikely more than one copy of *pCB1::CLF* would be excised at the same time. Thus, the compensatory effects of *pCB1::CLF* would complement cells if one copy remained. Plants showing 3:1 segregation for *pCB1::CLF* will be analyzed by Southern blot to establish the number of transgene insertions. The lines that have been established to have single locus insertions of *pCB1::CLF* will be crossed to the heat-shock inducible line. As recombination events are stochastic and mitotically independent, it is hoped that by heat-shocking young seedlings, large clonal *clf*- sectors will be created.

Creating *clf-* cells that are surrounded by *CLF+* cells will provide insight into whether *CLF+* action is cell intrinsic. If *CLF+* activity is cell intrinsic, silencing of *AG* in a *CLF+* petal sector will not be affected by the activation of *AG* in neighbouring *clf-* cells. The next step would be to characterise the effects on candidate target genes. Promoter::GUS reporter constructs could be introduced into the lines either by genetic crossing of reporter lines or introducing transgenes by floral dip transformation. B-glucuronidase activity would be analyzed in the transgenic plants, showing whether the promoter::GUS reporter transgene is activated within and around the *clf-* sectors created. To monitor the effects on target gene expression, reporter::histone2B:YFP (YFP_{H2B}) constructs could also be created that would allow sectors to be analyzed for GFP_{ER} (loss of *CLF+* activity) and YFP_{H2B} (target gene) expression. The florescent makers can be distinguished by confocal microscopy by intracellular location (cytoplasmic or nuclear) as well as having different spectral properties. To identify post-mitotic stages of leaf development, *CYCLIN::GUS* reporter lines could be introduced into the transgenic lines by genetic crossing.

7.vi. What is the mechanism of Pc-G mediated silencing and how reversible is it?

Alike in *Drosophila*, Pc-G mediated silencing in plants could provide the establishment of the initial repressive mark; H3K27me3. This histone modification could act as a guide to recruit other chromatin modifiers that promote stable long term silencing. The results described in Chapter 4 provide evidence to contradict the stability of Pc-G silencing in plants. The results described in Chapters 4 and 5 provide evidence against the Pc-G silencing in plants and suggest that transcription factors are able to access Pc-G marked sequences. It is possible that access is limited to periods when the epigenetic code is being reset, for example during S-phase. The binding of transcription factors during this time could prevent the Pc-G marks being re-established resulting in loss of silencing (Figure 7.2).

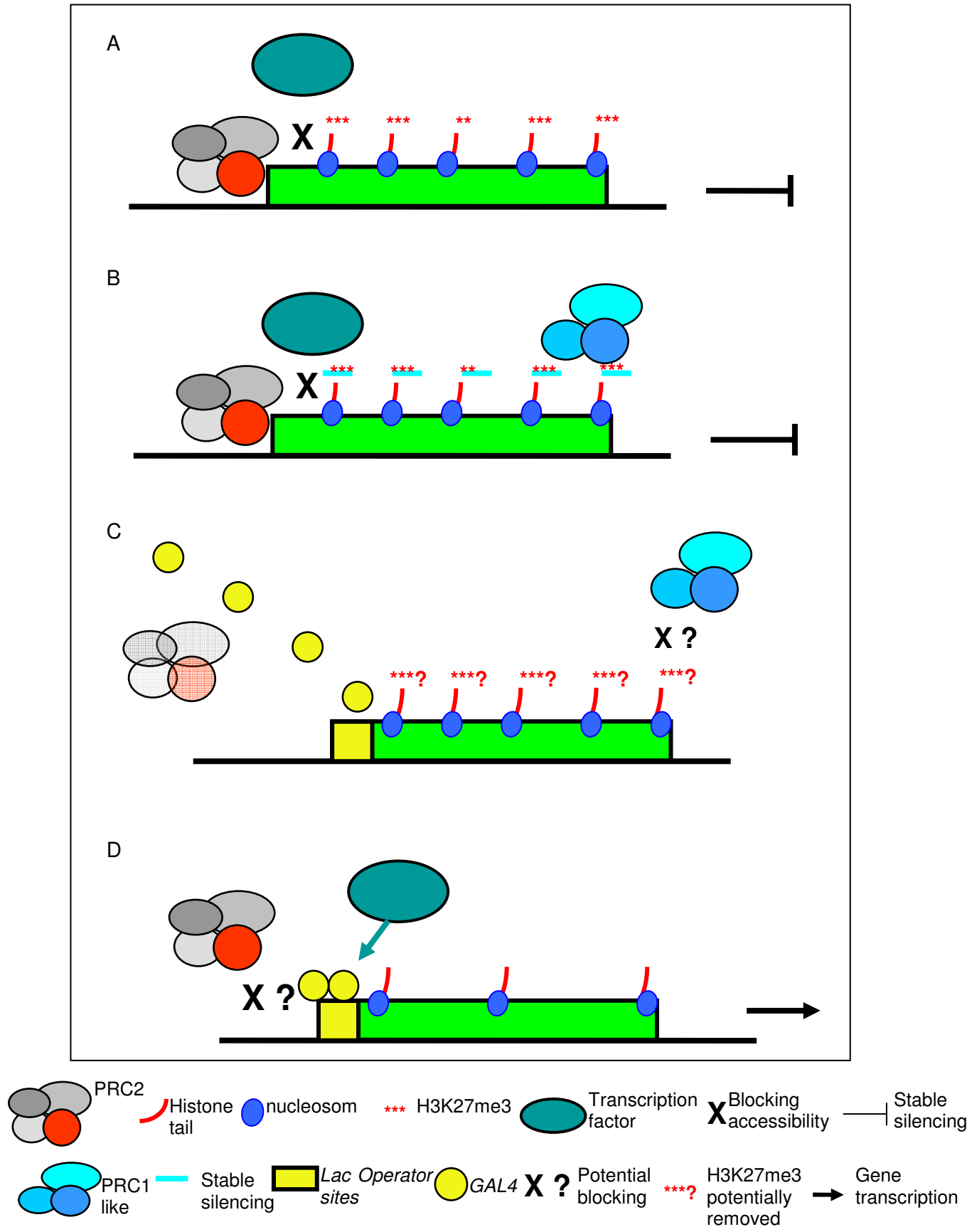

Figure 7.2 Reversing Pc-G mediated silencing

Figure 7.2 Model for reversing Pc-G mediated silencing

(A) The PRC2 complex is recruited and directly binds to target gene DNA, directing H3K27me3, promoting chromatin remodelling. This blocks accessibility to transcription factors. (B) H3K27me3 mark could act as a guide that recruits other chromatin modifiers to promote stable long term silencing in plants, similar to that in animals. (C) *GAL4* molecules can gain access to Pc-G marked sequences when the epigenetic code is being reset, for example during S-phase. This then blocks PRC2 from re-establishing the repressive mark, which hinders the recognition of other chromatin modifiers to target sequences. (D) The binding of *GAL4* during this time could prevent the Pc-G marks from being re-established resulting in loss of silencing. Transcription factors are then able to gain access leading to gene activation,

Although stable activation of a Pc-G silenced gene has been shown in Chapter 4, we have not shown whether CLF and H3K27me3 have been lost, or shown the gain of trxG and H3K4me3. It is not clear whether activation of Pc-G silenced genes occurs by a reversal of the process by which the genes were initially silenced or by a change that allows expression despite the 'marks of silencing' still being present.

The results described in Chapter 5 show that activation of a Pc-G target gene does not require the removal of H3K27me3. This result not only suggests that this repressive histone modification may not indicate silenced gene transcription states, but that a specific mechanism to remove the mark (for example a H3K27me3 demethylase) may not be required to allow gene transcription. A detailed analysis comparing histone modifications and gene transcription of reactivated genes would be required to determine if this result is general or specific to this system.

The accessibility of transcription factors to Pc-G modified chromatin in the presence of H3K27me3 provides evidence that Pc-G silencing may not be as stable as in *Drosophila*. However, this may in part be associated with the activity of other chromatin modifiers that interpret H3K27me3 differently. Pc-G silencing in plants could provide dynamic control of gene transcription, allowing gene activation or gene silencing to be reassessed following environmental cue or internal signal.

8. References

Aalfs, J. D. & Kingston, R. E. 2000, "What does 'chromatin remodeling' mean?", *Trends Biochem.Sci.*, vol. 25, no. 11, pp. 548-555.

Abe, M., Kobayashi, Y., Yamamoto, S., Daimon, Y., Yamaguchi, A., Ikeda, Y., Ichinoki, H., Notaguchi, M., Goto, K., & Araki, T. 2005, "FD, a bZIP Protein Mediating Signals from the Floral Pathway Integrator FT at the Shoot Apex", *Science*, vol. 309, no. 5737, pp. 1052-1056.

An, H., Roussot, C., Suarez-Lopez, P., Corbesier, L., Vincent, C., Pineiro, M., Hepworth, S., Mouradov, A., Justin, S., Turnbull, C., & Coupland, G. 2004, "CONSTANS acts in the phloem to regulate a systemic signal that induces photoperiodic flowering of Arabidopsis", *Development*, vol. 131, no. 15, pp. 3615-3626.

Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jorgensen, H. F., John, R. M., Gouti, M., Casanova, M., Warnes, G., Merkenschlager, M., & Fisher, A. G. 2006, "Chromatin signatures of pluripotent cell lines", *Nat Cell Biol*, vol. 8, no. 5, pp. 532-538.

Balasubramanian, S., Sureshkumar, S., Agrawal, M., Micheal, T. P., Wessinger, C., Maloof, J. N., Clark, R., Warthmann, N., Chory, J., & Weigel, D. 2006, "The PHYTOCHROME C photoreceptor gene mediates natural variation in flowering and growth responses of *Arabidopsis thaliana*", *Nature Genetics*, vol. 38, no. 6, pp. 711-715.

Bastow, R., Mylne, J. S., Lister, C., Lippman, Z., Martienssen, R. A., & Dean, C. 2004, "Vernalization requires epigenetic silencing of FLC by histone methylation", *Nature*, vol. 427, no. 6970, pp. 164-167.

Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., & Plath, K. 2006, "A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells", *Cell*, vol. 125, no. 2, pp. 315-326.

Bowman, J. L., Alvarez, J., Weigel, D., Meyerowitz, E. M., & Smyth, D. R. 1993, "Control of flower development in Arabidopsis thaliana by APETALA1 and interacting genes", *Development*, vol. 119, no. 3, pp. 721-743.

Boyer, L. A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L. A., Lee, T. I., Levine, S. S., Wernig, M., Tajonar, A., Ray, M. K., Bell, G. W., Otte, A. P., Vidal, M., Gifford, D. K., Young, R. A., & Jaenisch, R. 2006, "Polycomb complexes repress developmental regulators in murine embryonic stem cells", *Nature*, vol. 441, no. 7091, pp. 349-353.

Calonje, M., Sanchez, R., Chen, L., & Sung, Z. R. 2008, "EMBRYONIC FLOWER1 Participates in Polycomb Group-Mediated AG Gene Silencing in Arabidopsis", *The Plant Cell*, vol. 20, no. 2, pp. 277-291.

Carpenter, R. & Coen, E. S. 1990, "Floral homeotic mutations produced by transposon-mutagenesis in Antirrhinum majus", *Genes and Development*, vol. 4, no. 9, pp. 1483-1493.

Carroll, S. B., DiNardo, S., O'Farrell, P. H., White, R. A., & Scott, M. P. 1988, "Temporal and spatial relationships between segmentation and homeotic gene expression in Drosophila embryos: distributions of the fushi tarazu, engrailed, Sex combs reduced, Antennapedia, and Ultrabithorax proteins", *Genes and Development*, vol. 2, no. 3, pp. 350-360.

Chandler, J., Wilson, A., & Dean, C. 1996, "Arabidopsis mutants showing an altered response to vernalization", *Plant J.*, vol. 10, no. 4, pp. 637-644.

Chanvivattana, Y. 2002, *Interaction of Polycomb-group genes regulating flowering in Arabidopsis,* PhD, University of Edinburgh.

Chaudhury, A. M., Ming, L., Miller, C., Craig, S., Dennis, E. S., & Peacock, W. J. 1997, "Fertilization-independent seed development in Arabidopsis thaliana", *Proc.Natl.Acad.Sci.U.S.A*, vol. 94, no. 8, pp. 4223-4228.

Chen, L., Cheng, J. C., Castle, L., & Sung, Z. R. 1997, "EMF genes regulate Arabidopsis inflorescence development", *Plant Cell*, vol. 9, no. 11, pp. 2011-2024.

Coen, E. 1999, *The art of genes* Oxford University Press, Oxford.

Coen, E. S. & Meyerowitz, E. M. 1991, "The war of the whorls: genetic interactions controlling flower development", *Nature*, vol. 353.

Finnegan, E. J. & Dennis, E. S. 2007, "Vernalization-Induced Trimethylation of Histone H3 Lysine 27 at FLC Is Not Maintained in Mitotically Quiescent Cells", *Current Biology*, vol. 17, no. 22, pp. 1978-1983.

Francis, N. J. & Kingston, R. E. 2001, "Mechanisms of transcriptional memory", *Nat.Rev.Mol.Cell Biol.*, vol. 2, no. 6, pp. 409-421.

Gendall, A. R., Levy, Y. Y., Wilson, A., & Dean, C. 2001, "The VERNALIZATION 2 Gene Mediates the Epigenetic Regulation of Vernalization in Arabidopsis", *Cell*, vol. 107, no. 4, pp. 525-535.

Goodrich, J., Puangsomlee, P., Martin, M., Long, D., Meyerowitz, E. M., & Coupland, G. 1997, "A Polycomb-group gene regulates homeotic gene expression in Arabidopsis", *Nature*, vol. 386, no. 6620, pp. 44-51.

Goodrich, J. & Tweedie, S. 2002, "REMEMBRANCE OF THINGS PAST: Chromatin Remodeling in Plant Development", *Annu.Rev.Cell Dev.Biol.*, vol. 18, pp. 707-746.

Greb, T., Mylne, J. S., Crevillen, P., Geraldo, N., An, H., Gendall, A. R., & Dean, C. 2007, "The PHD Finger Protein VRN5 Functions in the Epigenetic Silencing of Arabidopsis FLC", *Current Biology*, vol. 17, no. 1, pp. 73-78.

Grossniklaus, U., Spillane, C., Page, D. R., & Kohler, C. 2001, "Genomic imprinting and seed development: endosperm formation with and without sex", *Curr.Opin.Plant Biol.*, vol. 4, no. 1, pp. 21-27.

Grossniklaus, U., Vielle-Calzada, J. P., Hoeppner, M. A., & Gagliano, W. B. 1998, "Maternal control of embryogenesis by MEDEA, a polycomb group gene in Arabidopsis", *Science*, vol. 280, no. 5362, pp. 446-450.

Guitton, A. E., Page, D. R., Chambrier, P., Lionnet, C., Faure, J. E., Grossniklaus, U., & 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*, vol. 131, no. 12, pp. 2971-2981.

Heidstra, R., Welch, D., & Scheres, B. 2004, "Mosaic analyses using marked activation and deletion clones dissect Arabidopsis SCARECROW action in asymmetric cell division", *Genes and Development*, vol. 18, no. 16, pp. 1964-1969.

Jack, T., Fox, G. L., & Meyerowitz, E. M. 1994, "Arabidopsis homeotic gene APETALA3 ectopic expression: transcriptional and posttranscriptional regulation determine floral organ identity", *Cell*, vol. 76, no. 4, pp. 703-716.

Johanson, U., West, J., Lister, C., Michaels, S., Amasino, R., & Dean, C. 2000, "Molecular analysis of FRIGIDA, a major determinant of natural variation in Arabidopsis flowering time", *Science*, vol. 290, no. 5490, pp. 344-347.

Jones, R. S. & Gelbart, W. M. 1990, "Genetic analysis of the enhancer of zeste locus and its role in gene regulation in Drosophila melanogaster", *Genetics*, vol. 126, no. 1, pp. 185-199.

Jones, R. S. & Gelbart, W. M. 1993, "The Drosophila Polycomb-group gene Enhancer of zeste contains a region with sequence similarity to trithorax", *Mol.Cell Biol.*, vol. 13, no. 10, pp. 6357-6366.

Krizek, B. A. & Fletcher, J. C. 2005, "Molecular mechanisms of flower development: an armchair guide", *Nat Rev Genet*, vol. 6, no. 9, pp. 688-698.

Lee, T. I., Jenner, R. G., Boyer, L. A., Guenther, M. G., Levine, S. S., Kumar, R. M., Chevellar, B., Johnstone, S. E., Cole, M. F., Isono, K., Koseki, H., Fuchikami, T., Abe, K., Murray, H. L., Zucker, J. P., Yuan, B., Bell, G. W., Herbolsheimer, E., Hannett, N. M., Sun, K., Odom, D. T., Otte, P. A., Volkert, T. L., Barterl, D. P., Melton, D. A., Gifford, D. K., Jeanisch, R., & Young, R. 2006, "Control of Developmental Regulators by Polycomb in Human Embryonic Stem Cells", *Cell*, vol. 125, pp. 301-313.

Lehmann, R. & Nusslein-Volhard, C. 1987, "hunchback, a gene required for segmentation of an anterior and posterior region of the Drosophila embryo", *Developmental Biology*, vol. 119, no. 2, pp. 402-417.

Lewis, E. B. 1978, "A gene complex controlling segmentation in Drosophila", *Nature*, vol. 276, no. 5688, pp. 565-570.

Li, E. 2002, "Chromatin Modification and Epigenetic Reprogramming in Mammalian Development", *Nature Genetics*, vol. 3, pp. 662-673.

Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F., & Richmond, T. J. 1997, "Crystal structure of the nucleosome core particle at 2.8 A resolution", *Nature*, vol. 389, no. 6648, pp. 251-260.

Luo, M., Bilodeau, P., Koltunow, A., Dennis, E. S., Peacock, W. J., & Chaudhury, A. M. 1999, "Genes controlling fertilization-independent seed development in Arabidopsis thaliana", *Proc.Natl.Acad.Sci.U.S.A*, vol. 96, no. 1, pp. 296-301.

McCall, K. & Bender, W. 1996, "Probes of chromatin accessibility in the Drosophila bithorax complex respond differentyl to Polycomb-mediated repression", *EMBO J.*

Michaels, S. D. & Amasino, R. M. 2001, "Loss of FLOWERING LOCUS C Activity Eliminates the Late-Flowering Phenotype of FRIGIDA and Autonomous Pathway Mutations but Not Responsiveness to Vernalization", *The Plant Cell*, vol. 13, no. 4, pp. 935-942.

Mizukami, Y. & Fischer, R. L. 2000, "Plant organ size control: AINTEGUMENTA regulates growth and cell numbers during organogenesis", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 2, pp. 942-947.

Moore, I., Galweiler, L., Grosskopf, D., Schell, J., & Palme, K. 1998, "A transcription activation system for regulated gene expression in transgenicáplants", *Proceedings of the National Academy of Sciences*, vol. 95, no. 1, pp. 376-381.

Nordborg, M., Borevitz, J. O., Bergelson, J., Berry, C. C., Chory, J., Hagenblad, J., Kreitman, M., Maloof, J. N., Noyes, T., Oefner, P. J., Stahl, E. A., & Weigel, D. 2002, "The extent of linkage disequilibrium in Arabidopsis thaliana", *Nat Genet*, vol. 30, no. 2, pp. 190-193.

Ohad, N., Yadegari, R., Margossian, L., Hannon, M., Michaeli, D., Harada, J. J., Goldberg, R. B., & Fischer, R. L. 1999, "Mutations in FIE, a WD polycomb group gene, allow endosperm development without fertilization", *Plant Cell*, vol. 11, no. 3, pp. 407-416.

Olins, A. L. & Olins, D. E. 1974, "Spheroid chromatin units (v bodies)", *Science*, vol. 183, no. 122, pp. 330-332.

Ringrose, L. & Paro, R. 2004, "Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins", *Annu.Rev.Genet.*, vol. 38, pp. 413-443.

Russo, V. E. A., Martienssen, R. A., & Riggs, A. D. 1996, *Epigenetic mechanisms of gene regulation* Cold spring Harbor Laboratory Press, Cold Spring Harbor, NY..

Salah El-Din El-Assal, Alonso-Blanco, C., Peeters, A. J. m., Raz, V., & Koornneef, M. 2001, "A QTL for flowering time in *Arabidopsis* reveals a nmovel allele of *CRY2*", *Nat Gene*, vol. 29, pp. 435-440.

Saleh, A., Al-Abdallat, A., Ndamukong, I., varez-Venegas, R., & Avramova, Z. 2007, "The Arabidopsis homologs of trithorax (ATX1) and enhancer of zeste (CLF) establish 'bivalent chromatin marks' at the silent AGAMOUS locus", *Nucleic Acids Research*, vol. 35, no. 18, pp. 6290-6296.

Samach, A., Onouchi, H., Gold, S. E., Ditta, G. S., Schwarz-Sommer, Z., Yanofsky, M. F., & Coupland, G. 2000, "Distinct Roles of CONSTANS Target Genes in Reproductive Development of Arabidopsis", *Science*, vol. 288, no. 5471, pp. 1613-1616.

Saurin, A. J., Shao, Z., Erdjument-Bromage, H., Tempst, P., & Kingston, R. E. 2001, "A Drosophila Polycomb group complex includes Zeste and dTAFII proteins", *Nature*, vol. 412, no. 6847, pp. 655-660.

Schatlowski, N., Creasey, K., Goodrich, J., & Schubert, D. 2008, "Keeping plants in shape: Polycomb-group genes and histone methylation", *Seminars in Cell & Developmental Biology*, vol. In Press, Corrected Proof.

Schubert, D., Primavesi, L., Bishopp, A., Roberts, G., Doonan, J., Jenuwein, T., & Goodrich, J. 2006, "Silencing by plant Polycomb-group genes requires dispersed trimethylation of histone H3 at lysine 27", *EMBO J.*

Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B., & Cavalli, G. 2007, "Genome Regulation by Polycomb and Trithorax Proteins", *Cell*, vol. 128, no. 4, pp. 735-745.

Schwartz, Y. B. & Pirrotta, V. 2008, "Polycomb complexes and epigenetic states", *Current Opinion in Cell Biology*, vol. 20, no. 3, pp. 266-273.

Shao, Z., Raible, F., Mollaaghababa, R., Guyon, J. R., Wu, C. T., Bender, W., & Kingston, R. E. 1999, "Stabilization of chromatin structure by PRC1, a Polycomb complex", *Cell*, vol. 98, no. 1, pp. 37-46.

Sheldon, C. C., Hills, M. J., Lister, C., Dean, C., Dennis, E. S., & Peacock, W. J. 2008, "Resetting of FLOWERING LOCUS C expression after epigenetic repression by vernalization", *Proceedings of the National Academy of Sciences* p. 0711453105.

Shindo, C., Lister, C., Crevillen, P., Nordborg, M., & Dean, C. 2006, "Variation in the epigenetic silencing of FLC contributes to natural variation in Arabidopsis vernalization response", *Genes and Development*, vol. 20, no. 22, pp. 3079-3083.

Sieburth, L. E. & Meyerowitz, E. M. 1997, "Molecular dissection of the AGAMOUS control region shows that cis elements for spatial regulation are located intragenically", *Plant Cell*, vol. 9, no. 3, pp. 355-365.

Steeves, T. A. & Sussex, I. M. 2002, *Patterns in plant development* Cambridge University Press, Cambridge.

Stewart, R. N. & Burk, L. G. 1970, "Independence of tissues derived from apical layers in ontogeny of the tobacco leaf and ovary", *American Journal of Botany*, vol. 57, no. 8, pp. 1010-1016.

Stinchcombe, J. R., Weinig, C., Ungerer, M., Olsen, K. M., Mays, C., Halldorsdottir, S. S., Purugganan, M. D., & Schmitt, J. 2004, "A latitudinal cline in flowering time in Arabidopsis thaliana modulated by the flowering time gene FRIGIDA", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 13, pp. 4712-4717.

Strahl, B. D. & Allis, C. D. 2000, "The language of covalent histone modifications", *Nature*, vol. 403, no. 6765, pp. 41-45.

Struhl, G. & Brower, D. 1982, "Early role of the esc+ gene product in the determination of segments in Drosophila", *Cell*, vol. 31, no. 1, pp. 285-292.

Sung, S. & Amasino, R. M. 2004, "Vernalization and epigenetics: how plants remember winter", *Curr.Opin.Plant Biol.*, vol. 7, no. 1, pp. 4-10.

Toomajian, C., Hu, T. T., Aranzana, M. J., Lister, C., Tan, C., Zheng, H., Zhao, K., Calabrese, P., Dean, C., & Nordborg, M. 2006, "A Nonparametric Test Reveals Selection for Rapid Flowering in the *Arabidopsis* Genome", *PLoS.Biol.*, vol. 4, no. 5, pp. 0732-0738.

van den Berg, C., Willemsen, V., Hage, W., Weisbeek, P., & Scheres, B. 1995, "Cell fate in the Arabidopsis root meristem determined by directional signalling", *Nature*, vol. 378, no. 6552, pp. 62-65.

Waddington, C. H. 1947, "Organisers & genes", *Cambridge University Press*.

Wang, J., Mager, J., Chen, Y., Schneider, E., Cross, J. C., Nagy, A., & Magnuson, T. 2001, "Imprinted X inactivation maintained by a mouse Polycomb group gene", *Nat.Genet.*, vol. 28, no. 4, pp. 371-375.

Weigel, D., Alvarez, J., Smyth, D. R., Yanofsky, M. F., & Meyerowitz, E. M. 1992, "LEAFY controls floral meristem identity in Arabidopsis", *Cell*, vol. 69, no. 5, pp. 843-859.

Weigel, D. & Meyerowitz, E. M. 1993, "Activation of Floral Homeotic Genes in Arabidopsis", *Science*, vol. 261, no. 5129, pp. 1723-1726.

Wigge, P. A., Kim, M. C., Jaeger, K. E., Busch, W., Schmid, M., Lohmann, J. U., & Weigel, D. 2005, "Integration of Spatial and Temporal Information During Floral Induction in Arabidopsis", *Science*, vol. 309, no. 5737, pp. 1056-1059.

Yoshida, N., Yanai, Y., Chen, L., Kato, Y., Hiratsuka, J., Miwa, T., Sung, Z. R., & Takahashi, S. 2001, "EMBRYONIC FLOWER2, a novel polycomb group protein homolog, mediates shoot development and flowering in Arabidopsis", *Plant Cell*, vol. 13, no. 11, pp. 2471-2481.

Zhang, X., Clarenz, O., Cokus, S., Bernatavichute, YV., Pellegrini, M., Goodrich, J., & Jacobsen, S. E. 2007, "Whole-Genome Analysis of Histone H3 Lysine 27 Trimethylation in *Arabidopsis*", *PLoS.Biol.*, vol. 5.

9. Appendix

9.i. Alignment of CLF to *Petunia* and *Antirrhinum-like* CLF

GenBank association numbers CAA71599 (*CLF*), AB098523 (*PhCLF1*), AB098524 (*PhCLF2*) and AB098525 (*PhCLF3*). *AntCLF1* and *AntCLF2* sequences were taken from the *Antirrhinum majus* genetic and genomic database (www.antirrhinum.net).

```
                  1                                                  50
     CLF     (1) MASEASPSSSATRSEPPKDSPAEERGP---ASKEVSEVIESLKKKLAADR
 AntCLF1     (1) -MSAPKAS--TSRSDRQVAGRQETRR----SKQKISAVINSFSKKFASDS
 AntCLF2     (1) -MSAPKASPAASGSDHTVAGQQGSP-----SAREVLAVIDSFKKEVSNRR
  PhCLF1     (1) --SESNGSDSQLQP-LNNHLPDVQPEGTTVEPDEVLSVIDSLKEKVAFER
  PhCLF2     (1) -MAASNGSLSVPRSQITDELMFDTPEEATVETDEVLSVIELLKKEVASAR
  PhCLF3     (1) MISSTSKSSESSPISTKSHGEHERDS----SS-SLTYGINQLKRKIQADR

                  51                                                 100
     CLF    (48) CISIKKRIDENKKNLFAITQSFMRSSMERG----GSCKDGSDLLVKRQRD
 AntCLF1    (44) SSYIEKRMVENIEKGLDVAKNLYKLSLDRRNFKIIGDDRSIDLLSKRQKD
 AntCLF2    (45) SEYIQKRMNENSEKVLDVTKDLYKLSVERKSFKITDADRSVDLLAKRQKD
  PhCLF1    (48) AHYIKKRVEGNTQKLEDLTKDLYNLATERKSLEVFGADRTVDLLAKRQKD
  PhCLF2    (50) ADYVKKRAEKNAQKLDDVTKDLFKSSTERRNLVIHGAD-SVDLLSKRQQD
  PhCLF3    (46) VLSVRDKHEENKRNLEKHVSELYLLATSRSDTVKNSGS--GKMLSLRVAN

                  101                                                150
     CLF    (94) SPGMKSGIDESN----NNRYVEDGPASSGMVQGSSVPVKISLRPIKMPDI
 AntCLF1    (94) AVDMHNDIATKSEDTESSSSREDGYASSAVLLGASITVKNAVLPINLPEV
 AntCLF2    (95) ATDMHNGIGTSNGDNESSSTQEDGYASSAILLGSSIAVKNAVRPIKLPEV
  PhCLF1    (98) AIDMQNGIDTSHGDDDSNSSEDDGYATSAILLGSSIAVKNAVRPIKLPEV
  PhCLF2    (99) AIDMHNGIDSSNGDNDSNSSEDDGYASSAILLGSSIAVKNAVRPITLPEM
  PhCLF3    (94) PLCKVGGLLQGSG------DRDYANGEEVVSS----------TTAKLPVI
```

```
                  151                                              200
    CLF   (140)   KRLSPYTTWVFLDRNQRMTEDQSVVGRRRIYYDQTGGEALICSDSEE---
AntCLF1   (144)   QRLPPYTSWIFLDRNQRMPEDQSVVGRKRIYYDQNGGEALICSDSEEGMT
AntCLF2   (145)   KRLPPYTSWIFLDRNQRMTEDQSVLGRRRIYYDQSGGEALICSDSEE---
 PhCLF1   (148)   KRIPPYTTWIFLDRNQRMTEDQSVVGRRRIYYDQNGGETLICSDSDE---
 PhCLF2   (149)   RRLPPYTTWVFLDRNQRMTEDQSVVGRRRIYYDQNGGEALICSDSEE---
 PhCLF3   (128)   EKIPPYTTWIFLDRNQRMAEDQSVVGRRRIYYDKHGSEALICSDSEE---

                  201                                              250
    CLF   (187)   -----------------EAIDDEEEKRDFLEPEDYIIRMTLEQLGLSDSV
AntCLF1   (194)   IEEVGLSDTTLDVLARYQAIDDEDSKRVFGESEDYIL-------------
AntCLF2   (192)   -----------------EANEDEDSKKEFGESEDYILRMTINEAGLSHTT
 PhCLF1   (195)   -----------------EVLDEEEEKKVFAEPEDYVLRMTIEEVGLSNTV
 PhCLF2   (196)   -----------------EGLEDEEEKKEFVESEDFMLRMAIKQVGLSDTV
 PhCLF3   (175)   -----------------DIAEPEEGKREFSEGEDKILWMASQEFGLSEEV

                  251                                              300
    CLF   (220)   LAELASFLSRSTSEIKARHGVLMKEK-------EVSESGDNQAESSLLNK
AntCLF1   (231)   -------LCRKSSEIKGRYEDLVKSGEAIY---LVNHGTDGGFVNPHINQ
AntCLF2   (225)   LDLLAQFLSRKSTDIKGRYEDLIKNGNALC---HVSNGNTGGSVNPYLDK
 PhCLF1   (228)   LDLLGKCLSRKPSDVKARYEDLVKEDNAGT---SKNQ-YMESSLDLYLAK
 PhCLF2   (229)   LDLLAQCLSRKPSELKARYEDIVKEENACV---SKNE-SIEGTVDFFLDK
 PhCLF3   (208)   LDILAHYVGGTTSEILERCNVLSEKHQDTDGKSLKDSGESGSRGTIFLDK

                  301                                              350
    CLF   (263)   DMEGALDSFDNLFCRRCLVFDCRLHGCSQDLIFPAEKPAPWCPPVDENLT
AntCLF1   (271)   DLDAALDSFDNWFCRRCFVFDCKLHGCSQDLVFPSEKQSSWSCLDVEKNP
AntCLF2   (272)   DLDAAMDSFDNLFCRRCLVFDCRLHGCSQDLVFPPEKQLSWSGLDVEKEP
 PhCLF1   (274)   DLDAALDSFDNLFCRRCLVFDCRLHGCSQDLIFPAEKQLPWHCSNADMEP
```


PhCLF2 (275) DIDAALDSFDNLFCRRCLVFDCRLHGCSQDLILTAEKQSAWHSPDADKEP
PhCLF3 (258) SLSAASDSFDNLFCRRCLVFDCRLHGCSQMLIDASEKQPYSSDSEDDGKP
351 400
CLF (313) CGANCYKTLLKSGRFPGYGTIEGKTGTSSDG--AGTKTTPTKFSSKLNGR
AntCLF1 (321) CGPDCHRLVLNAETKNKLSYEQADLEEKDSP------ASTVTASDMHKMK
AntCLF2 (322) CGPNCHKLAIFEEKAASASNGAGMQMQKRE----------QG-GLSVEK
PhCLF1 (324) CGPNCYSLAKKFESNATVISPQCASHEEKNVLPSDVASNTQLPGRKHVSR
PhCLF2 (325) CGPNCYRLAIKKESKATLTPPQLAIHGENPVQPSEVANSTQVAGRKHVSR
PhCLF3 (308) CSDQCYLKVKG----------
401 450
CLF (361) KPKTFPSESASSNEKCALETSDSENGLQQDTNSDKVSSSPKVKGSG-RRV
AntCLF1 (365) KSKSCQGE----------SSSSSVIKPVNDVTLSLKGPAAAKCG------
AntCLF2 (360) KSKSCQSE----------SASSSEIKSVNDITSSLKGSAARKFG------
PhCLF1 (374) RSKSSQGEGAPSNAKAVSESSDSEIRPINDVTSNKCSSSPTKSKSDSKDG
PhCLF2 (375) RSKSFQTESASSNAKNISESSDSEIRPIKDITSVKWTVSPSKTKSDCNGD
PhCLF3 (319) ----------AGDPSKHSTVDLPQGPGTGDPEEHT------
451 500
CLF (410) GRKRNKNRVAERVPRKTQKRQKKTEASDSDSIASGSCSPSDAKH-----K
AntCLF1 (399) SDKRSSKRVADNVLVAKQKRQKKTAASDSDPLAGGRFSLVNLGN---SSQ
AntCLF2 (394) SDKRSSKRIAESVLVAQKKRQKKTVASESDPVAGCSVSLKDLNLRNGSRR
PhCLF1 (424) SNKRNSKRIAEHVLVASKNKQKKMTALETDSVASGSLGSKGLNLHSISRK
PhCLF2 (425) SNKRNNKRIAEPVIAAIKKRQKKMTPMEPDS---GNQASKDLNLCSNSHK
PhCLF3 (344) -DGKMTHGASDSICTTMEKSDLVSDDQQDSSCK----------R
501 550
CLF (455) DNEDATSSSQKHVKSGNSGKSRKNGTPAEVSNNSVKDDVPVCQSNEVASE
AntCLF1 (446) EEEDIST--SLKVKYESPGKPRSKECPVLESEKSLPGVAAASASNEKVSN

```
AntCLF2  (444) ENEDSSS--SQKLKSTSSRKSRKRGSPAPESDKMLQGEGAIGASNESEGN
 PhCLF1  (474) ENGDVSSP-SQKAQCHSAKRSRRKNSPVTDSENSLQGKALDCQLIEATSE
 PhCLF2  (472) DVEDVSSS-SQRAPRHNGRRSRRKDCAVLSSENSLQGEGSSCQYKEATSQ
 PhCLF3  (377) RKLSVPTTVSVGAEDGSESNEISIITNDYVSHSPAPDDIGYNHSISLHKT

                      551                                       600
    CLF  (505) LDAPGSDESLRKEEFMGETVSRGRLATNKLWRPLEKSLFDKGVEIFGMNS
AntCLF1  (494) -LTMATSNTLKKEEFIDENKYPQELIDNKGWKPLEKALYEKGIQIFGRNS
AntCLF2  (492) QPVDSIGDILKKEEFVDENMHKQEVVDNKSWKPFEKALYDKGIEIFGRNS
 PhCLF1  (523) KPAINCDGMSRKNEYVGENNCKQEIDGIKSWRPIEKALFEKGLEMFGRSS
 PhCLF2  (521) KCGMSSEDTLRKNEFVDENNCKQKIDGDKSWRPLEKALFEKGLEMFGRSS
 PhCLF3  (427) GDSARSEGEDTKMEIVKQASC---LKNLQEWKPLEKELYSKGVEIFGRNS

                      601                                       650
    CLF  (555) CLIARNLLSGFKSCWEVFQYMTCSENKASFFGGDGLN--PDGSSKFDING
AntCLF1  (543) CLFARNLMSGLKSCSEVFGYMHRAEIKLFSQSTDGLN-LPGSCSKVESN-
AntCLF2  (542) CLIARNLMNGLKSCVEVFRYMHHSEKKMFSESNDGRD-PADSHLKADGN-
 PhCLF1  (573) CLIARNLMNGLKTCWEVFQYMNNSGNKLFSGAGDGMNGIFEGGSNGDGQ-
 PhCLF2  (571) CMIARNLMNGLKTCGEVFQYMNNS-EDMLSRVGYGVNGMLEGSSRGDAN-
 PhCLF3  (474) CLIARNLLPGLKTCMEVSSYMDGG-----AAAQRGSSARLFSEDNGNADM

                      651                                       700
    CLF  (603) NMVNNQVRRRSRFLRRRGKVRRLKYTWKSAAYHSIRKRITEKKDQPCRQF
AntCLF1  (591) ETVGGGARRRSKNLRRKGRVRRLKYTGKSAGYQSFRKRISERKDQPCTQY
AntCLF2  (590) ETLGGGARRRSRFLRRKGRVRRLKYTWKSAGYQSFRKRITERKDQPCRQY
 PhCLF1  (622) ENMGNEPRRRSKFLRRRGRVRRLKYTWKSAGYHAIRKRISERKDQPCRQF
 PhCLF2  (619) GIVGNAARRGSKFLRRRGRVRRLKYSWKSAGYHAFRKRISERKDQPCRQY
 PhCLF3  (519) DYMEPDMPTKSRFLRRRGRTRKLKYSSKSAGHPSMWRRIADGKNQSCIQY

                      701                                       750
```

```
    CLF  (653) NPCNCKIACGKECPCLLNGTCCEKYCGCPKSCKNRFRGCHCAKSQCRSRQ
AntCLF1  (641) NPCGCQSLCGKECPCLVNGTCCEKYCGCPKSCKNRFRGCHCAKSQCRSRQ
AntCLF2  (640) NPCGCQSACGKECPCLVNGTCCEKYCGCPKNCKNRFRGCHCAKSQCRSRQ
 PhCLF1  (672) NPCGCQGPCGKECPCIVNATCCEKYCGCPKSCKNRFRGCHCAKSQCRSRQ
 PhCLF2  (669) NPCNCQAPCGKECPCIVNGTCCEKYCGCPN-CKNRFRGCHCAKSQCRSRQ
 PhCLF3  (569) NPCGCQPTCGKDCPCLQNGTCCEKYCGCSKSCKNRFRGCHCAKSQCRSRQ

                   751                                       800
    CLF  (703) CPCFAADRECDPDVCRNCWVIGGDGSLGVPSQRGDNYECRNMKLLLKQQQ
AntCLF1  (691) CPCFAADRECDPDVCRNCWISCGDGSLGVPSQRGDNYECRNMKLLLRQQQ
AntCLF2  (690) CPCFAADRECDPDVCRNCWISCGDGSLGVPSQRGDNYECRNMKLLLKQQQ
 PhCLF1  (722) CPCFAAGRECDPDVCRNCWISCGDGTLGIPPQRGDSHECKNMKLLLKQQQ
 PhCLF2  (718) CPCFAADRECDPDVCRNCWISCGDGTLGTPSQRGDNYECRNMKLLLKQQQ
 PhCLF3  (619) CPCFAAGRECDPDVCRNCWVSCGDGSSGEPPRQGE-GQCGNMRLLLRQQQ

                   801                                       850
    CLF  (753) RVLLGISDVSGWGAFLKNSVSKHEYLGEYTGELISHKEADKRGKIYDRDEN
AntCLF1  (741) RVLLGRSGVSGWGAFLKNSVGKHEYLGEYTGEVISHHEADKRGKIYDRDEN
AntCLF2  (740) RILLGRSDISGWGAFLKNSVNKHEYLGEYTGELISHREADKRGKIYDRDEN
 PhCLF1  (772) KVLLGRSDVSGWGAFLKNSVGKHEYLGEYTGELISHREADKRGKIYDRDEN
 PhCLF2  (768) RVLLGRSDVSGWGAFLKNSVGKHEYLGEYTGEIISHHEADKRGKIYDREED
 PhCLF3  (668) RILLAKSHVAGWGAFLKNPVNKNDYLGEYTGELISHREADKRGKIYDRAN

                   851                                       900
    CLF  (803) CSFLFNLNDQFVLDAYRKGDKLKFANHSPEPNCYAKVIMVAGDHRVGIFA
AntCLF1  (791) SSFLFNLNDQFVLDAYRKGDKLKFANHSPNPNCYAKVILVAGDHRVGIFS
AntCLF2  (790) SSFLFNLNDQFVLDAYRKGDKLKFANHSPDPNCYAKVLMAAGDHRVGIFA
 PhCLF1  (822) SSFLFNLNDQFVLDAHRKGDKLKFANHSPVPNCYAKVMMVAGDHRVGIFA
 PhCLF2  (818) SSFLFNLNDQFVLDAYRKGDKLKFANHSPAPNCYAKVIMVAGDHRVGIFA
 PhCLF3  (718) SSFLFDLNDQYVLDAYRKGDKLKFANHSSNPNCYAKVMLVAGDHRVGIFA
```


901
950
CLF (853) KERILAGEELFYDYRYEPDRAPAWAKKPEAPGSKKDENVTPSVGRPKKLA
AntCLF1 (841) KERINAGEELFYDYHYQPEKAPAWAKKPGGASGAKREDAAPSNGRAKKHA
AntCLF2 (840) KERICAGEELFYDYRYEPDRAPAWAKKPE-ASGAKREEGAPSNGRAKKH-
PhCLF1 (872) NERICAGEELFYDYRYEPDSAPAWARKPE-ASGPRKDDAAPSSGRARKHT
PhCLF2 (868) KERICAGEELFYDYRYEADKAPAWARKPE-ASGTKKDDAAPSSGRAKKHT
PhCLF3 (768) KEHIEASQELFYDYRYGPDQAPIWARKPE---GTKREDSPVPPGRPKKHQ

Seminars in Cell & Developmental Biology 19 (2008) 547–553

Contents lists available at ScienceDirect

Seminars in Cell & Developmental Biology

journal homepage: www.elsevier.com/locate/semcdb

Review

Keeping plants in shape: Polycomb-group genes and histone methylation

Nicole Schatlowski[b], Kate Creasey[a], Justin Goodrich[a,*], Daniel Schubert[b,*]

[a] Institute of Molecular Plant Science, University of Edinburgh, Mayfield Road, Edinburgh, United Kingdom
[b] Institute of Genetics, Heinrich-Heine-University Duesseldorf, Universitaetsstr. 1, Germany

ARTICLE INFO

Article history:
Available online 31 July 2008

Keywords:
Polycomb
Histone methylation
Flowering
ChIP on chip
Resetting

ABSTRACT

The Polycomb-group proteins (Pc-G) are transcriptional repressors that control many key developmental transitions in plants. Recent whole genome profiling of the histone methylation marks that are characteristic of the Pc-G have suggested many novel targets of Pc-G regulation. However, the number and nature of targets also suggests that Pc-G regulation may be much more dynamic than hitherto assumed. We discuss the role of histone methylation in gene regulation, the complexes involved in setting and reading the mark and how activity states may be reset during development.

Contents

1. Introduction

The Polycomb-group (Pc-G) genes are a large and miscellaneous group of genes that were originally identified genetically, through their role in regulating homeotic gene expression in *Drosophila*, and were subsequently found to regulate development in other animals and plants. They mediate transcriptional silencing of numerous target genes, and this repression has been considered to be epigenetic, firstly because in some cases it has been shown to be propagated through numerous rounds of mitotic cell division in the absence of the original silencing signal, and secondly because it is also reversible, for example between generations [1–3]. A major breakthrough came in 2002 when several groups showed that four of the Pc-G proteins associate in a complex – often termed Polycomb Repressive Complex 2 (PRC2) – that methylates histone H3 in animals [4–7]. Subsequently, it has emerged that this modification – specifically, trimethylation of lysine 27 on histone H3 (H3K27me3) – is a universal mark for Pc-G activity in plants and animals. Although it remains unclear whether this mark is a cause or a consequence of Pc-G mediated silencing [8], it has provided a vital molecular lever for analysing Pc-G action. Here, we discuss recent progress in understanding which genes carry this mark in

Abbreviations: AG, AGAMOUS; BPC, BASIC PENTACYSTEINE; CLF, CURLY LEAF; CYP71, CYCLOPHILIN71; EMF1, EMBRYONIC FLOWER1; EMF2, EMBRYONIC FLOWER2; FIE, FERTILIZATION INDEPENDENT ENDOSPERM; FIS2, FERTILIZATION INDEPENDENT SEED2; FLC, FLOWERING LOCUS C; ICU2, INCURVATA2; MEA, MEDEA; MSI1, MULTICOPY SUPPRESSOR OF IRA1; PHE, PHERES; PRC1, Polycomb Repressive Complex 1; PRC2, Polycomb Repressive Complex 2; STM, SHOOTMERISTEMLESS; SWN, SWINGER; TFL2, TERMINAL FLOWER2; VEL, VERNALIZATION5/VIN3-LIKE; VIN3, VERNALIZATION INSENSITIVE3; VRN1, VERNALIZATION1; VRN2, VERNALIZATION2.

* Corresponding authors.
E-mail addresses: Justin.Goodrich@ed.ac.uk (J. Goodrich), Daniel.Schubert@uni-duesseldorf.de (D. Schubert).

1084-9521/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.semcdb.2008.07.019

plants, how it is deposited and interpreted, and how it may be removed.

2. Writing and reading the H3K27me3 mark

In *Drosophila*, the recruitment of Pc-G complexes to their targets is mediated by *cis*-acting sequences termed Polycomb Response Elements (PRE's, see [8,9] for review). No such elements have yet been identified in any other species including plants, although there has been some progress in defining *cis*-acting regions of the Pc-G targets *AGAMOUS* (*AG*) and *FLOWERING LOCUS C* (*FLC*) that are required for repression by Pc-G [10,11]. Analysis of two other H3K27me3 targets *SEEDSTICK* and *INNER NO OUTER* [12,13] identified the plant-specific GA/TC-binding protein BASIC PENTACYSTEINE1 (BPC1) as an important regulator. GA-repeats can be found in regulatory sequences of many other Pc-G targets such as *AG*, *FLC* and *SHOOTMERISTEMLESS* (*STM*), however a link to Pc-G silencing has not been revealed yet, as loss of BPC proteins only leads to mild phenotypes [12].

In animals, the PRC2 contains four core subunits: ENHANCER OF ZESTE [E(Z)], EXTRA SEX COMBS (ESC), SUPRESSOR OF ZESTE 12 (Su(z)12) and p55. Several PRC2 variants have been described which differ in associated proteins and their activity or specificity in histone methylation [14,15]. Although PRC2 complexes have yet to be purified from plant extracts, the four core components are conserved and there is strong genetic and biochemical evidence to suggest that they associate to form at least three similar complexes in *Arabidopsis*: the FIS (FERTILIZATION INDEPENDENT SEED) complex which controls seed development, the VRN (VERNALIZATION) complex which mediates the vernalization response and the EMF (EMBRYONIC FLOWERING) complex which represses precocious flowering and flower development (Fig. 1, see [16,17] for review). Whereas the *Arabidopsis* homologues of ESC and p55 [FERTILIZATION-INDEPENDENT ENDOSPERM (FIE) and MULTICOPY SUPPRESSOR OF IRA1 (MSI1), respectively] are likely a common component of all three complexes, the three different homologues of E(Z) [CURLY LEAF (CLF), SWINGER (SWN) and MEDEA (MEA)] and of SU(Z)12 [EMBRYONIC FLOWER2 (EMF2), VRN2 and FIS2] show differences in expression and target gene specificity that suggest that they are specific for the different complexes which have partially discrete functions [18,19]. Wood et al. [20] suggest that the VERNALIZATION INSENSITIVE3 (VIN3) protein may represent an additional, novel component of the VRN complex as VIN3 was found to co-immunoprecipitate with VRN2 and several of the other PRC2 core subunits. It is unlikely however that VIN3 is an integral PRC2 component as it is only expressed during cold (vernalization) treatments and therefore might mediate recruitment of a vernalization-specific complex during cold treatments. *VIN3* is a member of a small gene family, termed the *VEL* genes, one of whose members – *VERNALIZATION5* (*VRN5*) – is expressed more broadly and shows a more pleiotropic mutant phenotype raising the possibility that the other VEL proteins are components of the EMF and FIS PRC2-like complexes [21,22].

Different histone modifications do not all act independently, but rather can antagonize or promote one another [23]. Two genes have now been identified that may facilitate H3K27me3 methylation in plants by catalysing other modifications. The *PROTEIN ARGININE METHYLTRANSFERASE 5* (*AtPRMT5*) gene is required for silencing of the floral repressor *FLC* and encodes a histone arginine methyltransferase that catalyzes symmetric di-methylation of arginine 3 on histone H4 (H4R3sme2) both *in vivo* and *in vitro* [24–26]. Interestingly, in *atprmt5* mutants *FLC* does not acquire H3K27me3 during vernalization, indicating a crosstalk between histone arginine and lysine methylation. *CYCLOPHILIN71* (*CYP71*) is a likely component of Pc-G silencing as loss of function mutants show mis-expression of Pc-G target genes and subtle reductions in H3K27me3 [27]. The CYP71 protein binds histone H3 and is predicted to have proline isomerase activity. A link to histone methylation is suggested by a study in yeast showing that isomerization of proline 38 of histone H3 towards the *cis*-form inhibits lysine 36 methylation, most likely by changing the local conformation of the histone tail [28]. One possibility therefore is that *CYP71* facilitates H3K27me3 by regulating the isomerization of the nearby proline residue (H3P30). Both the *cyp71* and *atprmt5* phenotypes are mild relative to those of Pc-G mutants, suggesting that they are beneficial but not obligatory for H3K27me3 or that they act redundantly with other proteins.

In flies and mammals, a distinct Pc-G protein complex (PRC1) can bind H3K27me3 via the chromodomain protein POLYCOMB (PC) and is thought to confer stable, long-term silencing [9]. However, the PRC1 components are not conserved in plants, suggesting that an alternative plant-specific mechanism for interpretation of the H3K27me3 mark may have evolved. Consistent with this, several recent studies indicate that the *Arabidopsis* chromodomain protein LIKE HETEROCHROMATIN PROTEIN1 (LHP1) [also known as TERMINAL FLOWER2 (TFL2)] might fulfil an equivalent role to the PC protein. Firstly, *tfl2/lhp1* mutants show overlapping phenotypes with several Pc-G mutants [29–31]. Secondly, TFL2/LHP1 binds H3K27me3 *in vitro* (but also H3K9me2 and H3K9me3) and genome-wide profiling shows that it co-localizes with H3K27me3. Lastly, H3K27me3 levels are maintained in the *tfl2/lhp1* mutant indicating that TFL2/LHP1 acts downstream of methylation. Collectively, this is compelling evidence that TFL2/LHP1 reads the H3K27me3 mark [32,33], see also the review by Turck et al. in this issue]. However, the *tfl2/lhp1* mutant phenotype is very mild relative to severe Pc-G mutants, so other proteins must also play a role. These need not be chromodomain proteins, as other unrelated domains such as the plant homeodomain (PHD) have also been shown to bind methylated histones [34].

The plant-specific EMBRYONIC FLOWER1 (EMF1) protein was recently suggested to be a functional analogue of the *Drosophila* PRC1 component POSTERIOR SEX COMBS (PSC) as both proteins can inhibit transcription *in vitro* [35,36]. EMF1 is required for floral repression during vegetative development and *emf1* plants display a similar phenotype as the PRC2 mutant *emf2* [37]. EMF1 binds directly to several Pc-G target genes and at least for *AG*, binding depends on EMF2. However, biochemical analyses indicate an interaction with the PRC2 member MSI1 and H3K27me3 at Pc-G target genes is strongly reduced in *emf1* mutants, thus it is not entirely clear at which step of Pc-G silencing EMF1 acts [35]. EMF1 also binds DNA in a non-sequence specific manner, similar to VERNALIZATION1 (VRN1) which is required for the vernalization response [35,38]. Like EMF1, VRN1 is plant-specific and might act in a PRC1-like fashion: whereas H3K27me2/3 at *FLC* is not gained during vernalization in the PRC2 mutant *vrn2*, the modification is largely maintained in *vrn1* mutants indicating that VRN1 acts downstream of PRC2 [39–41].

3. Whole genome profiling suggests numerous novel Pc-G targets

In *Drosophila* and mammals, whole genome analyses of the H3K27me3 distribution revealed that the modification is not restricted to single genes but may cover numerous adjacent genes and extend over hundreds of kilobases [42–46]. The distribution of PRC2 components correlates well with that of H3K27me3, although in Drosphila they are more restricted to PRE elements whereas in mammals they are more broadly dispersed. Intriguingly, PRC2-mediated repression can be reversible as during differentiation of

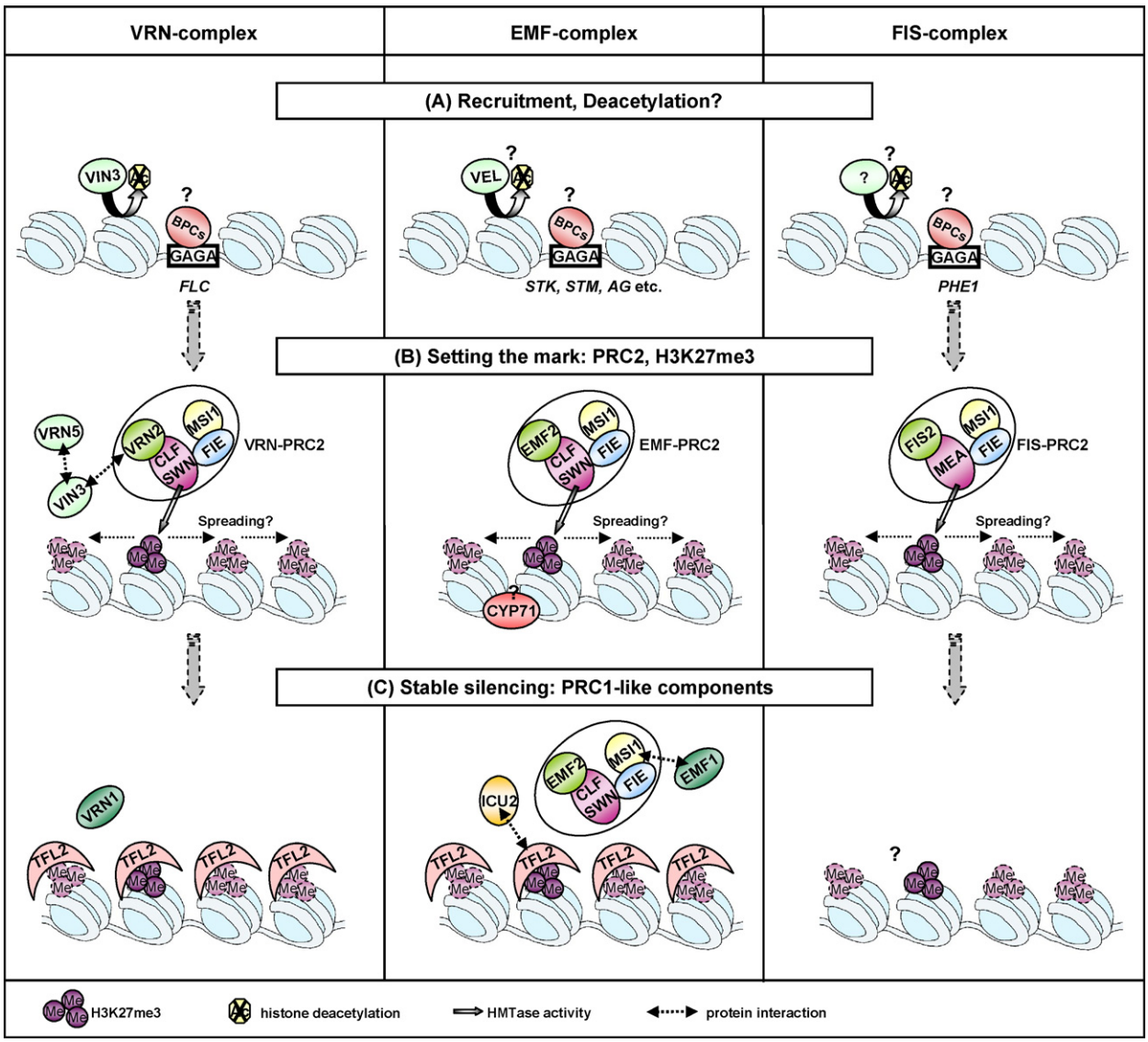

Fig. 1. Mechanisms and complexes of Pc-G silencing in *Arabidopsis*. Three PRC-like complexes have been proposed based on genetic and biochemical evidence: the VERNALIZATION (VRN), the EMBRYONIC FLOWER (EMF) and the FERTILIZATION INDEPENDENT SEED (FIS) complexes. (A) The PRC complexes are recruited to their target genes by unknown mechanism, which could involve BPC proteins through binding of GAGA-DNA-repeats. VIN3 and possibly homologous VEL proteins are required for histone deacetylation and might recruit PRC2 by protein interactions. (B) The H3K27me3 mark is set by the PRC2 complex catalyzed by the SET domain proteins CLF/SWN or MEA. The H3K27me3 mark spreads over the target loci by unknown mechanisms and proteins. CYP71 might be involved histone H3 isomerization. (C) TFL2, a potential PRC1-like component, binds to H3K27me3 and might mediate long term silencing. ICU2, a DNA polymerase subunit interacts with TFL2 and might be involved in epigenetic inheritance. EMF1 and VRN1 act downstream of PRC2. EMF1 interacts with the PRC2 component MSI1 *in vitro*.

ES cell lines, H3K27me3 and PRC2 components are removed from some target genes, as they are activated in specific cell lineages [42,46]. In plants, chromatin immunoprecipitation (ChIP) studies of several known plant Pc-G target genes showed that they carried H3K27me3 that was widely dispersed over the promoter and transcribed regions. Furthermore, Pc-G proteins were also bound to these targets and the methylation was lost in Pc-G mutants, consistent with the mark being deposited by Pc-G proteins [39–41,47]. Two groups have now extended these studies by probing tiling arrays with DNA obtained from ChIP ("ChIP on chip"), thereby profiling the distribution of H3K27me3 across the entire *Arabidopsis* genome [48] or chromosome four [32]. Perhaps the most striking feature of these studies is the large number of genes – roughly 4400, corresponding to 15% of *Arabidopsis* genes – that carry extensive H3K27me3. It is not yet certain whether all of these are Pc-G targets as mutants with severely depleted Pc-G activity retain some H3K27me3 [49], raising the possibility that other (non Pc-G) proteins can catalyze this modification in plants. Further profiling is therefore needed to show that plant PRC2 components are also located at the H3K27me3 targets, and that the methylation is lost in mutants which lack PRC2 activity. Nonetheless, it is likely that a significant proportion will correspond to Pc-G targets. Bioinformatic analysis of the H3K27me3 targets shows that they are enriched for transcription factors and signalling components such as the *CLAVATA3/ENDOSPERM SURROUNDING REGION-LIKE* (*CLE*) genes, and tend to be expressed in tissue specific fashion and at low levels relative to the genome average. As well as including all previously known Pc-G targets, many novel genes are revealed as potential Pc-G targets, for example those controlling leaf polarity (*YABBY* and *KANADI* genes), epidermal development (*PDF2* and *ATML1*) and stomatal development (*FAMA*, *MUTE*, *SPEECHLESS*) amongst others. Although some of these targets fit the classical model of Pc-G

mediated silencing in that they are silenced in particular domains early in development and remain off throughout, others show more dynamic expression patterns suggesting that Pc-G silencing may be more labile and reversible than often assumed. A limitation to the current studies is that they were obtained using whole seedlings, i.e., chromatin from a mixture of many different cell types and developmental stages. It will therefore be important to profile specific cell types, or a developmental series, in order to see how H3K27me3 and other modifications are affected during changes in cell fate and target gene expression.

4. Propagating the mark

The whole genome profiling studies confirm that H3K27me3 is widely dispersed over the transcribed and promoter regions of target genes in plants, but – in contrast to animals where it extends over large intergenic regions – it is relatively gene specific [32,41,48]. In both kingdoms, this argues against DNA sequence-specific recruitment of H3K27me3 but predicts spreading or looping by unknown mechanisms from a specific DNA element, such as the PRE in flies. Two studies support spreading of H3K27me3 in plants: Finnegan and Dennis [50] carefully analysed H3K27me3 at *FLC* during vernalization and revealed that H3K27me3 is found in a small region of *FLC* during cold treatment and then spreads over the entire gene after transfer to warm temperature. The spreading may be important for the stability of silencing, as when old leaves were vernalized spreading does not occur and silencing was not maintained. Schubert et al. [41] used reporter transgenes for the Pc-G target gene *AG* and demonstrated recruitment of H3K27me3 to the transgene and spreading over transgenic, non-plant sequences. As profiling studies show that H3K27me3 generally co-localizes with the transcribed region, one possibility is that spreading depends on RNA Polymerase II. Although participation of the transcriptional machinery in silencing may be rather counter-intuitive, there is a precedent in fission yeast where transcriptional silencing via RNA-interference dependent heterochromatin formation involves RNA polymerase II [51,52]. Alternatively, H3K27me3 might spread during DNA replication as chromatin is assembled behind the replication fork. Consistent with this, spreading is impaired in old leaves, where cell division and DNA replication is no longer occurring [50]. A potential link between DNA replication and Pc-G silencing comes from the somewhat surprising finding that a DNA polymerase is implicated in Pc-G action. *INCURVATA2* (*ICU2*) encodes the putative catalytic subunit of DNA polymerase α, and whilst severe mutants are lethal, the weak *icu2-1* allele gives a similar mutant phenotype to the Pc-G mutants *curly leaf* and *tfl2/lhp1* and shows ectopic expression of several Pc-G target genes [53,54]. The ICU2 protein interacts *in vitro* with TFL2/LHP1 and it is possible that the *icu2-1* mutation disrupts the interaction with this or another Pc-G component that is important for recruiting Pc-G during or after DNA replication to ensure mitotic stability of silencing and therefore epigenetic inheritance [54]. As H3K27me3 is mostly confined to single genes in *Arabidopsis*, spreading during replication would require insulators that prevent spreading of H3K27me3 into neighboring genes.

What might the biological relevance of dispersed H3K27me3 be? Random assortment of parental nucleosomes between daughter chromatids following DNA replication could result in stochastic loss of the modification on any one daughter. This is less likely to happen if H3K27me3 is present on numerous nucleosomes at one locus and thus would help to ensure stable inheritance of the modification and thus epigenetic fidelity. This is supported by a recent modeling study, which suggested that bistable epigenetic inheritance is facilitated by spread of H3K27me3 between non-adjacent nucleosomes, as might occur by DNA looping for example [55].

5. Silencing genes—the mark is not sufficient

Although the levels of H3K27me3 and H3K4me3 generally correlate well with target gene activity states, it is unclear whether they are causal. Studies using *AGAMOUS* reporter transgenes showed that when certain regulatory sequences were deleted the transgene still acquired high levels of H3K27me3 despite the fact that it was strongly mis-expressed in leaves, suggesting that the mere presence of H3K27me3 is not sufficient to inhibit transcription [10,41]. This is also supported by studies in *vrn1* and *tfl2/lhp1* mutants: although target genes were not silenced in the mutants they still carried H3K27me3 [32,39,40].

A detailed comparison of H3K27me3, Pc-G protein binding and expression of the *Drosophila* Pc-G target gene *Ultrabithorax* (*Ubx*) surprisingly showed that in the active state binding of Pc-G proteins was only slightly reduced [56]. In addition, H3K27me3 was only lost from the *Ubx* promoter and the 5′ coding region but not from an upstream control region. Thus, this study reveals a complex regulation mechanism that deviates from the simple "Pc-G present – Gene off" dogma [56]. Similar experiments will be required to determine how the presence and distribution of H3K27me3 and Pc-G proteins at plant Pc-G targets correlates with their activity.

6. Kicking against the PRCs—the antagonists of Pc-G silencing

The Trithorax-group genes (Trx-G) function as antagonists of Pc-G mediated silencing and maintain the activity of Pc-G targets in cell lineages where they were activated early in development. They were initially identified from genetic screens in *Drosophila* for factors that are required to maintain the activity of homeotic genes and comprise a functionally hetereogenous group of proteins (see [9,57] for review). One of the main activities of the *Trx-G* genes is correlated with H3K4 methylation, particularly H3K4me3, a mark associated with gene expression. Profiling studies using murine cell lines have shown that H3K4me3 and H3K27me3 are not mutually exclusive. In undifferentiated Embryonic Stem (ES) cell lines, many of the Pc-G/Trx-G target genes that are involved in cell fate specification are "bivalently" marked – i.e carry both these marks – and are transcriptionally silent. Bivalency is thought to mark genes poised for stable activation or repression later in development, as in differentiated cell lines the bivalent state is resolved so that the same targets carry one or other mark according to their activity state in a given cell line [42,58]. In *Arabidopsis*, there are five TRITHORAX homologues, of which ARABIDOPSIS HOMOLOG OF TRITHORAX (ATX1) is best characterized. ATX1 is required for the activity of several Pc-G targets including *AG* and *FLC* and is correlated with H3K4me3 at these targets [59,60]. Intriguingly, Saleh et al. [60] found that *AG* was enriched for both H3K4me3 and H3K27me3, yet stably silenced during vegetative development, and Pien et al. [59] also found evidence for both marks at *FLC*, the relative levels changing as FLC expression itself changes during floral transition. In another study of *FLC*, it was shown that different lengths of vernalization treatment result in different relative amounts of H3K4me3 and H3K27me3 that correlate the level of FLC activity [50]. It is not yet clear whether simultaneous detection of H3K4me3 and H3K27me3 in these analyses is equivalent to "bivalency" as a mixture of cell types and differentiation states were characterized. However, these studies do show that Pc-G/Trx-G regulation need not be all or nothing and can change during somatic development.

7. Resetting Pc-G silencing—removing the mark?

A defining characteristic of epigenetic changes is that they can be reversed. In animals, this often occurs during gametogenesis or early embryogenesis and presumably is needed to reprogram gene expression states and confer pluripotency. In plants, the vernalization-induced silencing of *FLC* is also reset between generations so that each new generation requires vernalization afresh, but it was not known exactly where or when this occurred. A recent study indicates that the silenced *FLC* gene is re-activated both around the time of male gametogenesis and also early in embryogenesis [61]. However, it is also likely that resetting of Pc-G target genes occurs at many other stages in plant development. Although some targets such as *MEDEA*, *PHERES1* and *FUSCA3* are stably silenced after germination [47,62] and therefore fit well with the commonly held view that Pc-G maintain repression states throughout somatic development, other genes such as *AG* and further floral homeotic genes are silent during vegetative development, yet need to be activated in floral organs. In addition, rapid activation of several Pc-G target genes is observed when plants are switched from short day to long day conditions. This includes the Pc-G target gene *FLOWERING LOCUS T* which is relatively rapidly activated upon transfer to inductive conditions [63]. Thus, the floral transition and more generally switches in meristem identity could represent major resetting events of Pc-G silencing in plant development.

Genome-wide analyses of Pc-G target genes during *Drosophila* development revealed that Pc-G proteins are not constitutively bound to all target genes [43,44]. In vertebrates, binding appears to be even more dynamic as many Pc-G target genes in embryonic stem cells are activated during differentiation and lose Pc-G binding [42,45]. Thus, Pc-G silencing appears to be much more dynamic than previously thought and does not only serve to ensure stable and long-term memory of repressed states but may be a more general mechanism to lock expression states at a certain stage of development.

What could the mechanisms be that activate silenced Pc-G targets? Although H3K27me3 is not sufficient for silencing of Pc-G targets [39–41,56], it may be inhibitory for transcription and activation of genes likely involves H3K27me3 removal. In dividing cells, resetting might occur passively through dilution, for example if PRC2 components are downregulated or otherwise not recruited to their targets after mitosis. At least CLF is not present on metaphase chromosomes, thus, prevention of recruitment of PRC2 after mitosis could result in loss of silencing [41]. However, the rapid changes that can occur for example with floral induction argue that active resetting may occur in the absence of cell division. Active removal of H3K27me3 (and other methylated histone residues) might occur by two mechanisms: exchange of the entire histone H3 with the replication-independent replacement histone H3.3 and removal of specific methylation sites by histone demethylases (HDMs) [64]. Exchange of histones would likely be a general resetting and activating step as all modifications are lost. Following the dynamics of the histone variants in plant development could reveal whether they play a role in major resetting events like the floral transition or during generation of the gametophytes and has been initiated for a male specific H3.3 [65].

In contrast, histone demethylases allow removal of specific methylgroups from histones. Lysine-specific demethylase 1 type (LSD1) histone demethylases can remove methyl groups from mono- or dimethylated residues whereas Jumonji-type (JmjC) histone demethylases can also target diverse trimethylated lysine residues (reviewed in [66]). Most interestingly, among the JmjC-domain containing HDMs, antagonists of Trx-G and Pc-G proteins have been identified in animals. Proteins from the JARID class of JmjC HDMs can remove methyl groups from the Trx-G mark H3K4me2/3, co-purify with Pc-G proteins and bind to a subset of Pc-G target genes [67,68]. JMJD3 class proteins are H3K27me2/3 antagonists, associate with Trx-G components and are required for activation of some Pc-G target genes [69,70]. Thus, a coordinated process of histone methylation and demethylation might ensure proper gene-regulation of Pc-G/Trx-G target genes.

In *Arabidopsis*, the JARID class of JmjC HDMs is conserved whereas no homologues of H3K27 demethylases are present (D. Schubert, unpublished). Several putative *Arabidopsis* HDMs are involved in the regulation of flowering time and regulate the Pc-G target *FLC* [71–74]. Interestingly, many putative HDMs are expressed strongly in meristems during floral induction [75] which could indicate a contribution of this class of proteins during this fundamental developmental process. Although no obvious H3K27me3 demethylases are encoded in the *Arabidopsis* genome the apparently dynamic regulation of H3K27me3 would predict presence of this class of enzymes. These might be identified in genetic screen for resetting mutants. The JmjC class proteins are a type of 2-oxoglutarate-iron-dependent dioxygenase that act via hydroxylation of the methyl-group on lysines [76]. As there are a large number of similar enzymes predicted in plants that lack a JmjC domain, it is possible that one of these could provide the demethylase activity.

8. Conclusions and summary

Whole genome profiling of plant Pc-G components should help refine the numerous direct targets of Pc-G repression that are suggested by whole genome profiling of H3K27me3. Pc-G silencing appears to be much more dynamic than previously thought and for many targets it is likely that silencing is not permanent but rather can be reversed during flowering or other developmental transitions. The mechanism of Pc-G silencing and the extent to which histone methylation is a cause or consequence of changes in gene activity remain unclear. To address these issues, it will be necessary to profile how Pc-G binding and other marks alter during cell fate transitions. In addition, rigorous biochemical purification and characterization of plant Pc-G complexes may help to clarify how H3K27me3 is read and interpreted.

Acknowledgement

N.S. and D.S. are supported by a SFB junior research grant from the German Research Society.

References

[1] Gendall AR, Levy YY, Wilson A, Dean C. The VERNALIZATION 2 gene mediates the epigenetic regulation of vernalization in *Arabidopsis*. Cell 2001;107:525–35.
[2] Goodrich J, Tweedie S. Remembrance of things past: chromatin remodeling in plant development. Annu Rev Cell Dev Biol 2002;18:707–46.
[3] Muller J. Transcriptional silencing by the Polycomb protein in *Drosophila* embryos. EMBO J 1995;14:1209–20.
[4] Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, et al. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science 2002;298:1039–43.
[5] Czermin B, Melfi R, McCabe D, Seitz V, Imhof A, Pirrotta V. *Drosophila* enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. Cell 2002;111:185–96.
[6] Kuzmichev A, Nishioka K, Erdjument-Bromage H, Tempst P, Reinberg D. Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. Genes Dev 2002;16:2893–905.
[7] Muller J, Hart CM, Francis NJ, Vargas ML, Sengupta A, Wild B, et al. Histone methyltransferase activity of a *Drosophila* Polycomb group repressor complex. Cell 2002;111:197–208.
[8] Ringrose L, Paro R. Polycomb/Trithorax response elements and epigenetic memory of cell identity. Development 2007;134:223–32.
[9] Schuettengruber B, Chourrout D, Vervoort M, Leblanc B, Cavalli G. Genome regulation by polycomb and trithorax proteins. Cell 2007;128:735–45.

[10] Sieburth LE, Meyerowitz EM. Molecular dissection of the AGAMOUS control region shows that cis elements for spatial regulation are located intragenically. Plant Cell 1997;9:355–65.
[11] Sheldon CC, Conn AB, Dennis ES, Peacock WJ. Different regulatory regions are required for the vernalization-induced repression of FLOWERING LOCUS C and for the epigenetic maintenance of repression. Plant Cell 2002;14:2527–37.
[12] Kooiker M, Airoldi CA, Losa A, Manzotti PS, Finzi L, Kater MM, et al. BASIC PENTACYSTEINE1, a GA binding protein that induces conformational changes in the regulatory region of the homeotic *Arabidopsis* gene SEEDSTICK. Plant Cell 2005;17:722–9.
[13] Meister RJ, Williams LA, Monfared MM, Gallagher TL, Kraft EA, Nelson CG, et al. Definition and interactions of a positive regulatory element of the *Arabidopsis* INNER NO OUTER promoter. Plant J 2004;37:426–38.
[14] Kuzmichev A, Jenuwein T, Tempst P, Reinberg D. Different EZH2-containing complexes target methylation of histone H1 or nucleosomal histone H3. Mol Cell 2004;14:183–93.
[15] Nekrasov M, Klymenko T, Fraterman S, Papp B, Oktaba K, Kocher T, et al. Pcl-PRC2 is needed to generate high levels of H3-K27 trimethylation at Polycomb target genes. EMBO J 2007;26:4078–88.
[16] Schubert D, Clarenz O, Goodrich J. Epigenetic control of plant development by Polycomb-group proteins. Curr Opin Plant Biol 2005;8:553–61.
[17] Kohler C, Villar CB. Programming of gene expression by Polycomb group proteins. Trends Cell Biol 2008;18:236–43.
[18] Chanvivattana Y, Bishopp A, Schubert D, Stock C, Moon YH, Sung ZR, et al. Interaction of Polycomb-group proteins controlling flowering in *Arabidopsis*. Development 2004;131:5263–76.
[19] Wang D, Tyson MD, Jackson SS, Yadegari R. Partially redundant functions of two SET-domain polycomb-group proteins in controlling initiation of seed development in *Arabidopsis*. Proc Natl Acad Sci USA 2006;103:13244–9.
[20] Wood CC, Robertson M, Tanner G, Peacock WJ, Dennis ES, Helliwell CA. The *Arabidopsis thaliana* vernalization response requires a polycomb-like protein complex that also includes VERNALIZATION INSENSITIVE 3. Proc Natl Acad Sci USA 2006;103:14631–6.
[21] Greb T, Mylne JS, Crevillen P, Geraldo N, An H, Gendall AR, et al. The PHD finger protein VRN5 functions in the epigenetic silencing of *Arabidopsis* FLC. Curr Biol 2007;17:73–8.
[22] Sung S, Schmitz RJ, Amasino RM. A PHD finger protein involved in both the vernalization and photoperiod pathways in *Arabidopsis*. Genes Dev 2006;20:3244–8.
[23] Fischle W, Wang Y, Allis CD. Histone and chromatin cross-talk. Curr Opin Cell Biol 2003;15:172–83.
[24] Pei Y, Niu L, Lu F, Liu C, Zhai J, Kong X, et al. Mutations in the Type II protein arginine methyltransferase AtPRMT5 result in pleiotropic developmental defects in *Arabidopsis*. Plant Physiol 2007;144:1913–23.
[25] Schmitz RJ, Sung S, Amasino RM. Histone arginine methylation is required for vernalization-induced epigenetic silencing of FLC in winter-annual *Arabidopsis thaliana*. Proc Natl Acad Sci USA 2008;105:411–6.
[26] Wang X, Zhang Y, Ma Q, Zhang Z, Xue Y, Bao S, et al. SKB1-mediated symmetric dimethylation of histone H4R3 controls flowering time in *Arabidopsis*. EMBO J 2007;26:1934–41.
[27] Li H, He Z, Lu G, Lee SC, Alonso J, Ecker JR, et al. A WD40 domain cyclophilin interacts with histone H3 and functions in gene repression and organogenesis in *Arabidopsis*. Plant Cell 2007;19:2403–16.
[28] Nelson CJ, Santos-Rosa H, Kouzarides T. Proline isomerization of histone H3 regulates lysine methylation and gene expression. Cell 2006;126:905–16.
[29] Gaudin V, Libault M, Pouteau S, Juul T, Zhao G, Lefebvre D, et al. Mutations in LIKE HETEROCHROMATIN PROTEIN 1 affect flowering time and plant architecture in *Arabidopsis*. Development 2001;128:4847–58.
[30] Mylne JS, Barrett L, Tessadori F, Mesnage S, Johnson L, Bernatavichute YV, et al. LHP1, the *Arabidopsis* homologue of HETEROCHROMATIN PROTEIN1, is required for epigenetic silencing of FLC. Proc Natl Acad Sci USA 2006;103: 5012–7.
[31] Sung S, He Y, Eshoo TW, Tamada Y, Johnson L, Nakahigashi K, et al. Epigenetic maintenance of the vernalized state in *Arabidopsis thaliana* requires LIKE HETEROCHROMATIN PROTEIN 1. Nat Genet 2006;38:706–10.
[32] Turck F, Roudier F, Farrona S, Martin-Magniette ML, Guillaume E, Buisine N, et al. *Arabidopsis* TFL2/LHP1 specifically associates with genes marked by trimethylation of histone H3 lysine 27. PLoS Genet 2007;3:e86.
[33] Zhang X, Germann S, Blus BJ, Khorasanizadeh S, Gaudin V, Jacobsen SE. The *Arabidopsis* LHP1 protein colocalizes with histone H3 Lys27 trimethylation. Nat Struct Mol Biol 2007;14:869–71.
[34] Mellor J. It takes a PHD to read the histone code. Cell 2006;126:22–4.
[35] Calonje M, Sanchez R, Chen L, Sung ZR. EMBRYONIC FLOWER1 participates in polycomb group-mediated AG gene silencing in *Arabidopsis*. Plant Cell 2008;20:277–91.
[36] King IF, Francis NJ, Kingston RE. Native and recombinant polycomb group complexes establish a selective block to template accessibility to repress transcription in vitro. Mol Cell Biol 2002;22:7919–28.
[37] Aubert D, Chen L, Moon YH, Martin D, Castle LA, Yang CH, et al. EMF1, a novel protein involved in the control of shoot architecture and flowering in *Arabidopsis*. Plant Cell 2001;13:1865–75.
[38] Levy YY, Mesnage S, Mylne JS, Gendall AR, Dean C. Multiple roles of *Arabidopsis* VRN1 in vernalization and flowering time control. Science 2002;297:243–6.
[39] Sung S, Amasino RM. Vernalization in *Arabidopsis thaliana* is mediated by the PHD finger protein VIN3. Nature 2004;427:159–64.
[40] Bastow R, Mylne JS, Lister C, Lippman Z, Martienssen RA, Dean C. Vernalization requires epigenetic silencing of FLC by histone methylation. Nature 2004;427:164–7.
[41] Schubert D, Primavesi L, Bishopp A, Roberts G, Doonan J, Jenuwein T, et al. Silencing by plant Polycomb-group genes requires dispersed trimethylation of histone H3 at lysine 27. EMBO J 2006;25:4638–49.
[42] Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 2006;125:315–26.
[43] Negre N, Hennetin J, Sun LV, Lavrov S, Bellis M, White KP, et al. Chromosomal distribution of PcG proteins during *Drosophila* development. PLoS Biol 2006;4:e170.
[44] Schwartz YB, Kahn TG, Nix DA, Li XY, Bourgon R, Biggin M, et al. Genome-wide analysis of Polycomb targets in *Drosophila melanogaster*. Nat Genet 2006;38:700–5.
[45] Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 2007;448:553–60.
[46] Bracken AP, Dietrich N, Pasini D, Hansen KH, Helin K. Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. Genes Dev 2006;20:1123–36.
[47] Makarevich G, Leroy O, Akinci U, Schubert D, Clarenz O, Goodrich J, et al. Different Polycomb group complexes regulate common target genes in *Arabidopsis*. EMBO Rep 2006;7:947–52.
[48] Zhang X, Clarenz O, Cokus S, Bernatavichute YV, Pellegrini M, Goodrich J, et al. Whole-genome analysis of histone H3 lysine 27 trimethylation in *Arabidopsis*. PLoS Biol 2007;5:e129.
[49] Lindroth AM, Shultis D, Jasencakova Z, Fuchs J, Johnson L, Schubert D, et al. Dual histone H3 methylation marks at lysines 9 and 27 required for interaction with CHROMOMETHYLASE3. EMBO J 2004;23:4286–96.
[50] Finnegan EJ, Dennis ES. Vernalization-induced trimethylation of histone H3 lysine 27 at FLC is not maintained in mitotically quiescent cells. Curr Biol 2007;17:1978–83.
[51] Djupedal I, Portoso M, Spahr H, Bonilla C, Gustafsson CM, Allshire RC, et al. RNA Pol II subunit Rpb7 promotes centromeric transcription and RNAi-directed chromatin silencing. Genes Dev 2005;19:2301–6.
[52] Kato H, Goto DB, Martienssen RA, Urano T, Furukawa K, Murakami Y. RNA polymerase II is required for RNAi-dependent heterochromatin assembly. Science 2005;309:467–9.
[53] Serrano-Cartagena J, Candela H, Robles P, Ponce MR, Perez-Perez JM, Piqueras P, et al. Genetic analysis of incurvata mutants reveals three independent genetic operations at work in *Arabidopsis* leaf morphogenesis. Genetics 2000;156:1363–77.
[54] Barrero JM, Gonzalez-Bayon R, del Pozo JC, Ponce MR, Micol JL. INCURVATA2 encodes the catalytic subunit of DNA Polymerase alpha and interacts with genes involved in chromatin-mediated cellular memory in *Arabidopsis thaliana*. Plant Cell 2007;19:2822–38.
[55] Dodd IB, Micheelsen MA, Sneppen K, Thon G. Theoretical analysis of epigenetic cell memory by nucleosome modification. Cell 2007;129:813–22.
[56] Papp B, Muller J. Histone trimethylation and the maintenance of transcriptional ON and OFF states by trxG and PcG proteins. Genes Dev 2006;20: 2041–54.
[57] Kennison JA, Tamkun JW. Trans-regulation of homeotic genes in *Drosophila*. New Biol 1992;4:91–6.
[58] Azuara V, Perry P, Sauer S, Spivakov M, Jorgensen HF, John RM, et al. Chromatin signatures of pluripotent cell lines. Nat Cell Biol 2006;8:532–8.
[59] Pien S, Fleury D, Mylne JS, Crevillen P, Inze D, Avramova Z, et al. ARABIDOPSIS TRITHORAX1 dynamically regulates FLOWERING LOCUS C activation via histone 3 lysine 4 trimethylation. Plant Cell 2008.
[60] Saleh A, Al-Abdallat A, Ndamukong I, varez-Venegas R, Avramova Z. The *Arabidopsis* homologs of trithorax (ATX1) and enhancer of zeste (CLF) establish 'bivalent chromatin marks' at the silent AGAMOUS locus. Nucleic Acids Res 2007;35:6290–6.
[61] Sheldon CC, Hills MJ, Lister C, Dean C, Dennis ES, Peacock WJ. Resetting of FLOWERING LOCUS C expression after epigenetic repression by vernalization. Proc Natl Acad Sci USA 2008;105:2214–9.
[62] Jullien PE, Katz A, Oliva M, Ohad N, Berger F. Polycomb group complexes self-regulate imprinting of the Polycomb group gene MEDEA in *Arabidopsis*. Curr Biol 2006;16:486–92.
[63] Corbesier L, Vincent C, Jang S, Fornara F, Fan Q, Searle I, et al. FT protein movement contributes to long-distance signaling in floral induction of *Arabidopsis*. Science 2007;316:1030–3.
[64] Bannister AJ, Schneider R, Kouzarides T. Histone methylation: dynamic or static? Cell 2002;109:801–6.
[65] Ingouff M, Hamamura Y, Gourgues M, Higashiyama T, Berger F. Distinct dynamics of HISTONE3 variants between the two fertilization products in plants. Curr Biol 2007;17:1032–7.
[66] Lan F, Nottke AC, Shi Y. Mechanisms involved in the regulation of histone lysine demethylases. Curr Opin Cell Biol 2008.
[67] Pasini D, Hansen KH, Christensen J, Agger K, Cloos PA, Helin K. Coordinated regulation of transcriptional repression by the RBP2 H3K4 demethylase and Polycomb-Repressive Complex 2. Genes Dev 2008;22:1345–55.
[68] Lee MG, Norman J, Shilatifard A, Shiekhattar R. Physical and functional association of a trimethyl H3K4 demethylase and Ring6a/MBLR, a polycomb-like protein. Cell 2007;128:877–87.

[69] Lee MG, Villa R, Trojer P, Norman J, Yan KP, Reinberg D, et al. Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. Science 2007;318:447–50.
[70] Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, Lee TI, et al. Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature 2006;441:349–53.
[71] He Y, Michaels SD, Amasino RM. Regulation of flowering time by histone acetylation in *Arabidopsis*. Science 2003;302:1751–4.
[72] Jiang D, Yang W, He Y, Amasino RM. *Arabidopsis* relatives of the human lysine-specific Demethylase1 repress the expression of FWA and FLOWERING LOCUS C and thus promote the floral transition. Plant Cell 2007;19:2975–87.
[73] Noh B, Lee SH, Kim HJ, Yi G, Shin EA, Lee M, et al. Divergent roles of a pair of homologous jumonji/zinc-finger-class transcription factor proteins in the regulation of *Arabidopsis* flowering time. Plant Cell 2004;16:2601–13.
[74] Krichevsky A, Gutgarts H, Kozlovsky SV, Tzfira T, Sutton A, Sternglanz R, et al. C2H2 zinc finger-SET histone methyltransferase is a plant-specific chromatin modifier. Dev Biol 2007;303:259–69.
[75] Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, et al. A gene expression map of *Arabidopsis thaliana* development. Nat Genet 2005;37:501–6.
[76] Trewick SC, McLaughlin PJ, Allshire RC. Methylation: lost in hydroxylation? EMBO Rep 2005;6:315–20.