Genetic screen for novel Polycomb Group (PcG) genes and targets in *Arabidopsis thaliana*

Manuel A. López Vernaza

2009

Supervisor: Dr. W. J. Goodrich

Declaration

I hereby declare that all work described in this thesis, unless otherwise acknowledge, has been performed by myself and has not been accepted for any previous application for a degree. The information obtained from sources other than this study is acknowledged in the text or included in the references.

_____

Manuel A. López Vernaza

Acknowledgements

The work presented in this thesis was carried out at the Institute for Molecular Plant Science (IMPS) at the University of Edinburgh where I spent three magnificent years of my life. Very special people made this happened:

Firstly, I wish to express my endless gratitude to Sir Kenneth Murray and the Darwin Trust of Edinburgh for giving me the opportunity and the financial support to do my thesis in such a great place and in such outstanding conditions.

Secondly, I am most grateful to my supervisor, Dr. William Justin Goodrich, for his, support, guidance, patience, inspiration and enthusiastic way of enjoy Life. Through my thesis, his knowledge, technical advices and scientific precision were highly valuable and appreciated. A particular thankfully thought as well for all his hard work, encouragements, help and opinions during the writing of this manuscript. For all this and the many rewarding pints of beer: Thanks Justin!
I warmly thank all the members of "The big lab" and "The Goodrich's lab" in particular: Dan, Daniel, Frazer, Johanna, Kate, Oliver, Ralf, Suxin, Jaan and Jessica with who I could share many exciting, vibrant and passionate discussions and thoughts during my thesis. I would like to warmly thank as well Gwyneth, Andrew and Peter who had always time to answer my many questions. All their advices and friendly chats help me and encourage me during my thesis especially during the writing period.

Thirdly, a special thank to all my friends and relatives in Edinburgh and elsewhere for their unlimited support and their kindness: In Edinburgh: Lucio and Federica, Kim and Frazer, Anna, Laura, Martin, Stefano, Christine, Jens, Vincent, Ross, Andrea, Fiona, Daniel, Margaret, Nuala, Donal and Elaine. Elsewhere: Charles and Anne, Tristan and Adeline, Sandra and Vincent, John and Robyn, John and Jennifer, Olivier, Nick, Nicolas, Cyril, Sébastien, Stéphane, Tam and Fabien.

Fourthly, I would like to thank all my previous supervisors Agnes Fouet, Genevieve Ephritikhine, Catherine Perrot-Rechenmann, Rishikesh Bhalerao and David Leach who in many independent and personal ways showed me how enjoyable, fun, interesting, exciting, stimulating, fascinating and rewarding science can be.

Fifthly, I would like to thank my mother and my brother for their everlasting love and support through my whole life. I would not start and achieve this without them.

Finally, a special though to Lorna who transform every moment of my life into sweet magic moments of happiness and who inspired me with courage and love to be a better man everyday.

Nomenclature

In this manuscript I followed the nomenclature proposed by David Meinke and Maarten Koornneef in The Plant Journal (1997) 12(2), 247-253, in which a gene identified by a mutation in *Arabidopsis thaliana* should be named as follow:

- Mutant gene symbols should have three letters (italics) written in lower case letters (e.g. *abc*).
- The wild-type allele should be written (italics) in capital letters (e.g. *ABC*).
- The full descriptive names of the wild-type (*ALPHABETICA*) and the mutant (*alphabetica*) alleles should be written in the same manner.
- Protein products of genes should be written in capital letters without italics (e.g. ABC).
- Phenotypes may be designated by the gene symbol (no italics) with the first letter capitalized. Thus Abc+ describes the wild-type; Abc- refers to the mutant.
- Different genes with the same symbol are distinguished by different numbers (e.g. *abc1* and *abc2*).
- Different alleles of the same gene are distinguished with a number following a hyphen (e.g. *abc4-1* and *abc4-2*).

Abstract

Genetic screen for novel Polycomb Group (PcG) genes and targets in *Arabidopsis thaliana*.

Polycomb Group (PcG) proteins are responsible for post-transcriptional modifications on histone tails leading to chromatin condensation and changes in gene expression. In *Arabidopsis thaliana*, *CURLY LEAF* (*CLF*) is a member of the Polycomb Repressive Complex 2 (PRC2), which confers a repressive epigenetic mark, namely trimethylation of histone H3 at lysine 27 (H3K27me3). In the *clf* mutant, the expression of the floral organ identity gene *AGAMOUS* (*AG*) is derepressed in vegetative stages and coincides with loss of H3K27me3 at the *AG* locus. Recent whole genome profiling studies have suggested that PcG genes regulate many more developmental regulators than *AG* (about 15% of *Arabidopsis* genes). However, it remains unclear what the relevance of PcG regulation of these targets is for plant development; in addition, it is not known how changes in H3K27me3 cause gene repression in plants. To unravel the role of *CLF* in *A. thaliana*, a T-DNA mutagenesis in the *clf* background was performed to identify mutations enhancing or suppressing the Clf- phenotype, as these may identify additional PcG genes and targets. Firstly, I screened an *A. thaliana* T-DNA mutagenized population and identified four mutations suppressing the Clf- phenotype: *suppressor of polycomb 1* to *4* (*sop1*, *sop2*, *sop3* and *sop4*). Secondly, I characterized these four mutants. The *sop1* mutant had normal flowering time and the suppressed phenotype is due to a loss of function mutation in *SEPALLATA3* (*SEP3*). I established that *SEP3* is an activator and a co-factor of *AG*. Also, I found that *SEP3* is strongly mis-expressed in *clf* mutants and *SEP3* chromatin is enriched in H3K27me3, which strongly suggests that *SEP3* is a direct target of *CLF*. In addition, I showed that a mutation in *Flowering Locus T* (*FT*), which is a positive regulator of *SEP3*, suppresses the Clf- phenotype suggesting that *FT* is also a target of *CLF*. Suppressors *sop2*, *sop3* and *sop4* are late flowering, unlike *sop1*, and show increased expression of *Flowering Locus C* (*FLC*), a

MADS-box transcription factor gene that represses flowering. I found that the *sop4* mutation is likely caused by disruption of *FPA*, a predicted RNA binding protein that promotes flowering time by repressing *FLC*. Consistent with this, *sop4* mutants show high levels of *FLC*. Unexpectedly, *fpa clf* (*sop4*) mutants are much later flowering than *clf FRI* mutants, which have similarly high levels of *FLC*. This suggests that *FPA* may regulate other genes controlling flowering than *FLC*. The genes involved in *sop2* and *sop3* mutants remain to be identified.

In this thesis I brought genetic and molecular evidence showing that *CLF*, through the PRC2, controls floral induction (*FLC*), floral integration (*FT*) and floral organ formation (*SEP3* and *AG*) in *A. thaliana*.

Abbreviations

a, A	adenine
ATP	adenosine triphosphate
bp	base pair(s)
BSA	bovine serum albumin
c, C	cysteine
cDNA	complementary deoxyribonucleic acid
ChIP	chromatin immuno precipitation
chip	microarray gene chip
cm^2	squares centimeters
CSPD	Disodium 3-(4-methoxyspiro(1,2-dioxetane-3.2′-(5-chloro) tricycle [3.3.1.1] decan)-4-yl) phenyl phosphate
dCAP	derived cleaved amplified polymorphism
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
DTT	dithiothreiotol
EDTA	ethylenediaminetetraacteic acid (disodium salt)
EMS	ethylmethanesulfonate
g, G	guanine
g	gram
GFP	green fluorescent protein
H	histone
K	lysine
kb	kilobase
M	molar
m	milli (10^{-3})
min	minutes
mRNA	messenger ribonucleic acid
μ	micro (10^{-6})

n	nano (10^{-6}
PCR	polymerase chain reaction
P	pico (10^{-12})
RT-PCR	reverse transcription polymerase chain reaction
RNA	ribonucleic acid
RNAase	ribonuclease
rpm	revolutions per minute
SDS	sodium dodecyl sulfate
SNP	Single Nucleotide Polymorphism
t, T	thymine
TE	Tris-EDTA buffer
Tween	polyethylenesorbitan monolaurate
U	unit
UTR	untranslated region
UV	ultraviolet

Contents

1. Introduction

The normal development of a multicellular organism depends on the establishment in early stages of specific transcription programs leading to correct cell growth, cell differentiation and organ morphogenesis. Expression of specific genes in early stages of development fixes the body plan in animals. These genes are called "homeotic genes", and they determine which parts of the body form what body part. Once the body plan is established, cellular mechanisms are responsible for the maintenance through mitoses of the transcriptional states during the whole life of the organism. One of the mechanisms involved in the maintenance is composed by proteins originally discovered in *Drosophila melanogaster* that form complexes and modify chromatin structure to repress: Polycomb Group (PcG) or activate: Trithorax Group (trxG) gene transcription (Francis and Kingston, 2001). It is now evident that chromatin structure is closely linked to transcription programs undertook during developmental cell processes (Hsieh and Fischer, 2005).

In eukaryotic cells, genomic DNA is folded by histone and non-histone proteins to form chromatin. Despite the important amount of coding information present in the DNA sequence, additional levels of information concerning regulation of genetic expression are associated with chromatin organization. Changes in chromatin structure influences gene expression and can lead to mitotically or meiotically (especially for DNA methylation) heritable changes in gene expression. These changes are called epigenetic as unlike classical genetic mutations, no changes in DNA sequence are involved. Histone modification, DNA methylation and chromatin-remodelling mechanisms are all mechanism for epigenetic modifications of chromatin (Jenuwein and Allis, 2001). Two distinct forms of chromatin have long been recognized on cytological basis. Heterochromatin stains more intensely than euchromatin with DNA binding dyes as is more compacted. Currently, it is widely recognized that euchromatin is in general less condensed, more accessible and more

transcriptionally active. By contrast, heterochromatin is more condensed, less accessible and contains more structured nucleosomal arrays (Huisinga et al., 2006). Chromosomal regions that are characterized by a high density of repetitive sequences (e.g. telomeres) are usually associated with heterochromatin (Grewal and Jia, 2007). These regions of heterochromatin remain condensed throughout the cell cycle and are referred to as "constitutive" heterochromatin. However, another type of heterochromatin, termed "facultative", can be nucleated at specific loci and is regulated by cellular signals or gene activity (Trojer and Reinberg, 2007). In *Arabidopsis thaliana* chromosomes have a fairly simple distribution of the two different types of chromatin. Euchromatin correspond to regions rich in genes whereas, heterochromatin match with repeat-rich sequences usually in pericentromeric regions (Fransz et al., 2003). Multiple pathways of histone modification, DNA methylation and small RNA molecules contribute in higher eukaryotes to heterochromatin assembly (Maison and Almouzni, 2004). Moreover, it is known that many of the roles of the individual histone modifications are conserved between the plant and animal kingdoms (Pfluger and Wagner, 2007).

Chromatin structure is important for maintenance of cell fate and to give an epigenetic memory to daughter cells. In plants, even if cell fate seems to be more flexible than in animals, genetic analysis suggests that changes in chromatin are important for genome defence and development. Recent studies in *A. thaliana*, have provided whole genome histone modifications maps showing important biological roles for histone modifications (Turck et al., 2007; Zhang et al., 2007b; Oh et al., 2008). In these three works, the authors focus on the tri-methylation of lysine 27 of histone H3 (H3K27me3) which correspond to one of the major epigenetic repressive marks. This modification is made by a histone methyltransferase, a member of the Polycomb Group (PcG) genes.

In the following sections, *Polycomb Group* (*PcG*) genes from *D. melanogaster* and *A. thaliana* will be presented and their role in development and chromatin regulation will be discussed. The *curly leaf* (*clf*) mutant in *A. thaliana* and a genetic screen for modifiers of the Clf- phenotype will be presented in detail.

1.1. PcG genes provide a developmental memory in *D. melanogaster*

PcG genes were originally identified from genetic screens in *D. melanogaster* as they induce common developmental abnormalities. In PcG mutants of *D. melanogaster*, the spatial and temporal expression of homeotic target genes is initiated normally in early stages of development. However, in later stages the regulation fails and PcG genes targets are mis-expressed. The PcG genes are required to guarantee that transcriptional repression of target genes is maintained in cell lineages where a target was initially repressed (Goodrich and Tweedie, 2002).

1.1.1. PcG genes in D. melanogaster

PcG genes were originally identified because specific mutant phenotypes suggested that the product of these genes were needed for the repression of homeotic genes. Homeotic genes composing the ANTENNAPEDIA (ANT) and BITHORAX (BX) complexes encode transcription factors that control segment identity, or cell fate, along the anterior-posterior body axis in the developing *D. melanogaster*. Mutations in one of the homeotic genes, *ULTRABITHORAX* (*UBX*) result in transformations of the appendages of the metathoracic segment (usually the haltere and third leg) into appendages associated with the mesothoracic segment (wing and second leg) (Struhl, 1982). Weak *PcG* mutants are viable but share an extra sex combs phenotype (in males) due to mis-expression of homeotic genes of the ANT complex in meso- and meta- throracic legs. Severe *PcG* mutant are lethal (die as early larvae), and

show profound mis-expression of most of the homeotic genes of the ANT and BX complexes (Struhl and Brower, 1982; Kennison, 1995; Ringrose and Paro, 2004; Schwartz and Pirrotta, 2007).

Biochemical studies on PcG proteins revealed that they associate in multimeric complexes, which explain the common phenotypes observed. At least three complexes have been identified and purified in *D. melanogaster*: PhoRC, Polycomb Repressive Complex 1 (PRC1) and Polycomb Repressive Complex 2 (PRC2). The PhoRC complex contains the PLEIOHOMEOTIC (PHO) and dSFMBT proteins (Klymenko et al., 2006); the PRC1 contains the POSTERIOR SEX COMBS (PSC), POLYCOMB (PC), POLYHOMEOTIC (PH), SEX COMBS EXTRA (RING) proteins (Shao et al., 1999), and the PRC2 contains the EXTRA SEX COMBS (ESC), ENHANCER OF ZEST (E(Z)), POLYCOMB LIKE (PLC), P55 and SUPPRESSOR OF ZEST 12° (SU(Z)12) proteins (Czermin et al., 2002; Muller et al., 2002).
As shown in Figure 1. 1, the different complexes identified in *D. melanogaster* may have different functions in the silencing pathway. They can be involved in the recruitment to target loci, the marking of the chromatin via methylation of specific histone residues or the interpretation of the mark previously deposited.

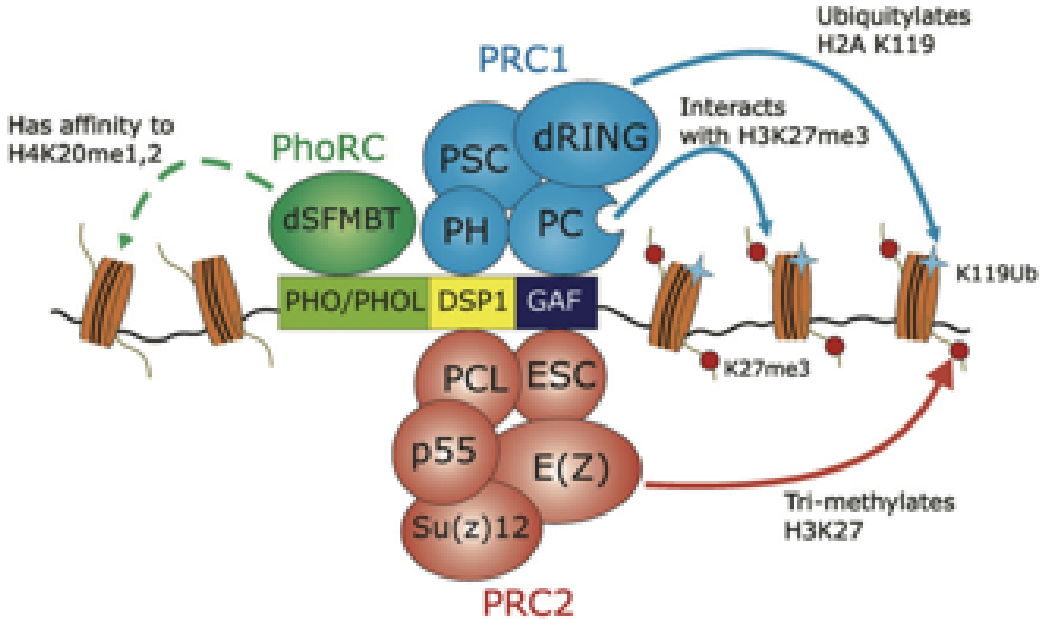

Figure 1. 1: PcG proteins form three complexes, PRC1 (blue), PRC2 (red) and PhoRC (green). In *D. melanogaster*, these complexes are recruited to the PRE by a combination of DNA binding activities carried by PHO, DSP1, and GAF. E(Z) trimethylates lysine 27 of histone H3, and the chromodomain of PC (PRC1) recognize histone H3K27me3 mark. dRING monoubiquitylates lysine 119 of histone H2A. dSFMBT has affinity for monomethyl K20 and dimethyl K20 of histone H4 although the relevance of this property for its PcG function is not clear (Schwartz and Pirrotta, 2008).

The PcG complexes have specific enzymatic activities allowing them to regulate (repress) the homeotic genes as well as other target genes (Ringrose and Paro, 2004; Schwartz and Pirrotta, 2007). *PcG* genes action (repression) is antagonistic of trithorax group (trxG) genes action (activation). trxG genes were first identified as counteractors of the PcG genes based on their ability to suppress PcG phenotypes (Pien and Grossniklaus, 2007). In *D. melanogaster*, four trxG complexes are required to maintain homeotic loci in a transcriptionally active state. This study is focused on *PcG* genes, for more extensive comprehensive discussion of the regulation of homeotic targets genes by *trxG* and *PcG* genes see reviews by Ringrose and Paro,

2004; Schuettengruber et al., 2007; Schwartz and Pirrotta, 2008 and Pien and Grossniklaus, 2007.

1.1.2. Mode of action of the PRC1 and PRC2 complexes in repressing gene expression in *D. melanogaster*

PcG target genes have *cis*-regulatory elements called PcG Response Elements (PRE) that enable PcG proteins to bind and to maintain transcriptional repression through mitosis to specific gene targets (Ringrose and Paro, 2004). The DNA sequence seems to be crucial for the binding of PcG proteins to the target loci as insertion of a PRE is sufficient to recruit PcG proteins to a novel chromosomal location. Core PRC2 members lack DNA binding specificity, however several proteins associated (GAF, PIPSQUEAK (PSQ), PHO and ZESTE (Z)) with PcG are able to bind DNA at conserved domains (PREs) (Ringrose and Paro, 2004). Nevertheless, ChIP-on-chip studies showed that PcG proteins like E(Z) and PSC are intimately associated with PREs in *D. melanogaster* (Schwartz et al., 2006). As PREs are poised with PcG and trxG proteins it has been proposed that PREs act as a platform to PcG and trxG to interact with other regulatory elements (promoter) to repress or activate transcription of target loci (Ringrose and Paro, 2004).

It has been also shown that small non coding RNAs are able to target the PRC2 components (EZH2, the mouse counterpart of E(Z)) to their targets in *D. melanogaster* and *M. musculus* (Grimaud et al., 2006; Zhao et al., 2008). Unfortunately, this studies are very speculative on the role and the function of PRC2 components in regulating/modulating the targeting and the following epigenetic silencing at specific targets.

Following the targeting, the PRC2 is the first complex to act. It has been shown that the four core components of the PRC2, E(Z), SU(Z)12, ESC, and P55 physically

interact together. The purified complex shows an activity *in-vitro* furthermore, when PcG activity is depleted *in-vivo* the level of H3K27me3 decreases in *D. melanogaster* (Cao et al., 2002). It has also been shown that in human cell culture a histone methyl transferase (HMTase) activity was associated with the PRC2 (Cao et al., 2002; Czermin et al., 2002; Kuzmichev et al., 2002; Muller et al., 2002). It is E(Z) that maintains gene repression by catalysing di- and tri-methylation of lysine 27 on histone H3 (H3K27me2 and H3K27me3). The E(Z) protein contains a SET domain and this domain was linked to the histone methyltransferase activity of the PRC2 (Czermin et al., 2002).

The PRC1 interprets the repressive marks made by the PRC2 and stabilise/maintain the transcriptional repression through mitoses. It is known that PcG proteins like PHO, DSP1 and GAF bind to target gene DNA at the PREs and recruits the PC protein (Brown et al., 2003; Schwartz and Pirrotta, 2008). It has been shown that via their chromo domain PC bind specifically H3k27me3 suggesting that the PRC1 is able to read the H3K27me3 mark (Fischle et al., 2003; Min et al., 2003). Following this interaction, the ubiquitylation of histone H2A at lysine 119 is mediated by the RING component of the PRC1 (Wang et al., 2004). Very little is known about the role of this modification. Mutations in the human homolog of the RING protein RING1B causes loss of H2A ubiquitylation and loss of repressive activity even though H3K27me3 persists (de Napoles et al., 2004). Together these results suggest that the PRC1 inhibits chromatin remodelling, although it is still unclear how the PRC1 impede access to the DNA and establish constant genetic repression.

1.2. PcG genes in *A. thaliana*

In plants, PcG members were mostly found in genetic screens based on their role in seed development and flowering. Molecular isolation showed that they all correspond to PRC2 homologs (except TERMINAL FLOWER 2 (TLF2)/LIKE HETEROCHROMATIN PROTEIN 1 (LHP1)). Most of the PcG members are represented by small gene families as showed in Table 1. 1. There are three homologues of *E(Z)* in *A. thaliana*: *CURLY LEAF* (*CLF*), *MEDEA* (*MEA*) and *SWINGER* (*SWN*). *clf* and *mea* mutations have distinct mutant phenotypes suggesting the two genes have at least partially distinct functions. *swn* mutations have no obvious phenotype, but enhance both *clf* and *mea* mutations suggesting that *swn* is redundant with *clf* and *mea* (Chanvivattana et al., 2004; Wang et al., 2006). There are three homologues of *SU(Z)12*: *EMBRYONIC FLOWER 2* (*EMF2*), *FERTILIZATION-INDEPENDENT SEED 2* (*FIS2*) and *VERNALIZATION 2* (*VRN2*). An enhanced phenotype is obtained with double mutants of the related genes *EMF2* and *VERNALIZATION2* (*VRN2*) (Schubert et al., 2005) (Figure 1. 6), suggesting that they also share some redundancy. There are five homologues of P55: *MULTICOPY SUPPRESSOR OF IRA 1* (*MSI1*), *MSI2*, *MSI3*, *MULTICOPY SUPPRESSOR OF IRA 4* (*FVE*) and *MSI5*. It has been shown that MSI1 is the only protein among the MSI1 family able to interact with other PcG proteins in *A. thaliana*. Also, *MSI1* is required to maintain the correct temporal and organ-specific expression of homeotic genes, including *AGAMOUS* and *APETALA2* (Hennig et al., 2003; Reyes and Grossniklaus, 2003). ESC only has one homolog in *A. thaliana*: *FERTILIZATION-INDEPENDENT ENDOSPERM* (*FIE*).

Table 1. 1: *A. thaliana* homologues of *D. melanogaster* PRC2 members. The full names of the proteins involved are: SUPPRESSOR OF ZEST12 (SU(Z)12), ENHANCER OF ZEST (E(Z)), EXTRA SEX COMB (ESC), EMBRYONIC FLOWER2 (EMF2), FERTILIZATION-INDEPENDENT SEED2 (FIS2), VERNALIZATION2 (VRN2), CURLY LEAF (CLF), MEDEA (MEA), SWINGER (SWN), FERTILIZATION-INDEPENDENT ENDOSPERM (FIE) and MULTICOPY SUPPRESSOR OF IRA1 (MSI1).

Drosophila **PcG protein**	***Arabidopsis*** **PcG protein**	**Protein conserved domain**	**Function**	**Developmental process**
SU(Z)12	EMF2 (Yoshida et al., 2001)	Zinc finger VEFS	Protein-protein interaction	Repression of floral genes
	FIS2 (Chaudhury et al., 1997)			Repression of *FIS*
	VRN2 (Gendall et al., 2001)			Vernalization
E(Z)	CLF (Goodrich et al., 1997)	SET C5	Histone methyltransferase Protein-protein interaction	Repression of floral genes
	MEA (Grossniklaus et al., 1998)			Repression of *PHERES* and *FIS* genes
	SWN (Chanvivattana et al., 2004)			Redundant with CLF
ESC	FIE (Ohad et al., 1996)	WD40 repeats	Protein-protein interaction	Repression of *FIS* and floral genes
P55	MSI1 (Kohler et al., 2003; Guitton et al., 2004)	WD40 repeats	Histone binding protein	Repression of *FIS* and floral genes

Plant *PcG* genes have been found as a result of their discrete functions in endosperm/embryo development, control of vernalization response and floral transition that are reviewed in the following sub-sections.

1.2.1. The *FIS* genes control endosperm/embryo development

In *A. thaliana*, double fertilisation of the ovule, gives rise to two distinct zygotic structures: the plant embryo and the endosperm, which together go on to develop into a seed. Before fertilization, the FIS PRC2-like complex composed by MEA, FIS2, FIE and MSI1 (Table 1. 1) represses proliferation of the central cell in the ovule. Mutation of only one of the component of the maternal FIS PRC2-like complex makes endosperm go through disordered divisions and seed abortion, as in the *fis2* or the *mea* mutants (Reyes and Grossniklaus, 2003). In *mea* mutant seeds, the embryo aborts at heart stage, probably due to defects in endosperm development (Grossniklaus et al., 1998). The *msi1* gametophyte shows the same endosperm phenotype, with asynchronous divisions from the egg cell having the same result (Guitton and Berger, 2005). Among the targets of the FIS complex, the PcG gene *MEA* and the MADS-box transcription factor *PHERES1* (*PHE1*) have been shown to be regulated by parental imprinting (Kohler et al., 2005; Jullien et al., 2006). Other putative targets of epigenetic regulations by the FIS PcG-like complex are the *LEAFY COTYLEDON* (*LEC*) genes: *LEC1*, *LEC2* and *FUSCA3*. Overexpression of *LEC* genes induce the presence of embryonic traits in plantlets. Indeed, in wild type plants the transcription of the *LEC* genes is repressed by the SWI/SNF2-like protein PICKLE (PKL) (Ogas et al., 1999), a CHD/Mi-2-type ATPase, suggesting that PKL can epigenetically regulate *LEC* genes. For the moment no mechanistic link has been found between PKL and the FIS PcG-like complex.

The FIS PRC2-like complex has likely a discrete function in seed development as *FIS2* and *MEA* are almost only expressed in seeds. There are very slight roles for this complex after germination.

1.2.2. Control of vernalization response

The MADS-box transcription factor *FLOWERING LOCUS C* (*FLC*) is a key regulator of flowering. The main role of the FLC protein is to represses the expression of the two floral integrators *SOC1* and *FT* (Hepworth et al., 2002; Searle et al., 2006; Li et al., 2008). *FLC* is regulated by two of the four flowering regulating pathways: the autonomous and the vernalization pathways (Komeda, 2004; Parcy, 2005) (Figure 1. 2).

The autonomous pathway consists of seven known genes: *FPA*; *FCA*; *FLK*; *FY*; *FLD*; *FVE* and *LUMINIDEPENDENS* (*LD*). Their main function is to repress *FLC* expression (Michaels and Amasino, 2001; Quesada et al., 2005). The number of members in this pathway is not fixed and seems to be more components to characterize (Terzi and Simpson, 2008).

The vernalization requirement segregates as a single gene trait mapped at the *FRIGIDA* (*FRI*) locus. The FRI protein is predicted to contain coiled-coil domains in two positions (between amino acids 55 to 100 and 405 to 450, respectively). FRI has been shown to increase RNA levels of FLC (Johanson et al., 2000). Only FRI^{+} plants can be vernalized as in a FRI^{-} background the levels of *FLC* are already low and the plants flower early. Subsequent to a vernalization treatment (cold treatment for a long period), FRI^{+} FLC^{+} plants flower earlier than non vernalized plants. It has been shown that the vernalization pathway repressing flowering requires *VRN2* (Table 1. 1) (Gendall et al., 2001; Reyes, 2006). This repression has epigenetic properties (i) mitotically stable even when plants are moved to warm and long days conditions. (ii) reversible: the progeny of vernalized plants needs vernalization to accelerate flowering. After a vernalization treatment, the *FLC* locus is epigenetically silenced. Indeed, it has been shown that *VRN2* mediates the stable repression of *FLC* in *A. thaliana* (Gendall et al., 2001). The H3K27me3 mark increases at the *FLC* locus

following a vernalization treatment, but is lost in *vrn2* mutants (Bastow et al., 2004). This modification is certainly catalysed by the VRN2 PRC2-like complex composed by CLF/SWN, VRN2, FIE and MSI1 (Table 1. 1). Mutations in *VERNALIZATION 1* (*VRN1*), which encodes a protein containing two B3 DNA-binding domains (Levy et al., 2002) and *VERNALIZATION INSENSITIVE 3* (*VIN3*), which encode a PHD-domain-containing protein (Sung and Amasino, 2004), cause insensitivity to vernalization. It has been shown that both proteins VRN1 and VIN3 are necessary to establish the silenced chromatin conformation even if the *vrn1* mutant retains some H3K27me3 (Bastow et al., 2004; Sung and Amasino, 2004; Reyes, 2006).

1.2.3. Control of flowering transition

Flowering is the developmental switch point from the vegetative to the reproductive phase. In *A. thaliana*, there are physical, chemical and biological signals required for the initiation of flowering that are integrated by four regulating pathways: the gibberellin pathway, the autonomous pathway, the vernalization pathway and the photoperiodic pathway (Komeda, 2004; Parcy, 2005) (Figure 1. 2). Two genes, *CONSTANS* (*CO*), which is specific of the photoperiod pathway and it is the output of the circadian clock specific for flowering and *Flowering Locus C* (*FLC*), are particularly important to assimilate and relay the information to the floral pathways integrators *LEAFY* (*LFY*), *Flowering Locus T* (*FT*) and *Suppressor of CO overexpression* (*SOC1*) (Figure 1. 2). Other pathways as the light quality or the ambient temperature control flowering. Overall, All the stimuli, inducing (photoperiodic pathway) or repressing flowering (autonomous pathway) are integrated and assimilated to allow the plant to flower in the best conditions for reproduction and consequently the set of seed.

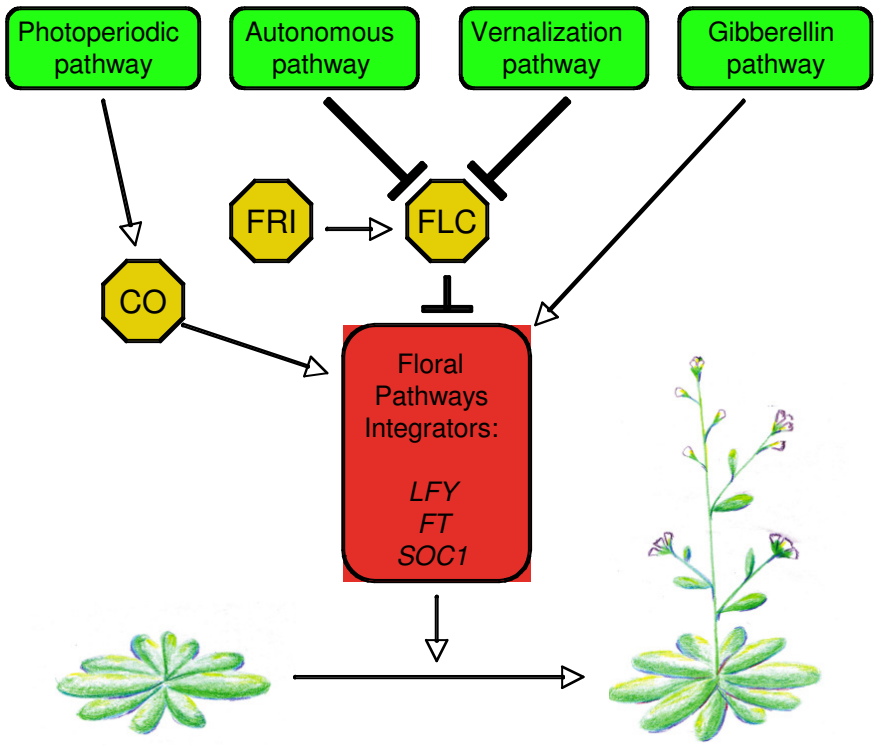

Figure 1. 2: The four genetic pathways of flowering are linked to the Floral Pathways Integrators *LEAFY* (*LFY*), Flowering Locus T (FT) and *Suppressor of CO overexpression* (*SOC1*) through the genes *CONSTANS* (*CO*) and *Flowering Locus C* (*FLC*) (Parcy, 2005).

Activation of the floral pathway integrators triggers the "flowering program" by inducing the floral meristem identity (FMI) genes, which are responsible for the formation of the first floral meristem. Among the FMI genes there are *LFY*, *APETALA1* (*AP1*), *CAULIFLOWER* (*CAL*) and *UNUSUAL FLORAL ORGANS* (*UFO*) (Krizek and Fletcher, 2005). FMI genes are expressed in early floral stages and are responsible for the floral meristem fate (Lohmann and Weigel, 2002). The principal function of the FMI genes is to activate the genes that specify floral organ identity.

The floral organ identity genes were discovered in *A. thaliana* through characterization of mutants displaying homeotic mutant phenotypes in which organs in two adjacent floral whorls develop with identity of other whorl. Based on

previous work in *Antirrhinum majus* and the effects of homeotic mutants in *A. thaliana*, the ABC model for the specification of floral organ identity was proposed in the early 1990's (Coen and Meyerowitz, 1991). This proposes three genetics functions - A, B and C - which act in a combinatorial fashion to specify organ identity (Figure 1. 3). In *A. thaliana*, the A function requires expression of *APETALA1* (*AP1*) and *APETALA2* (*AP2*) within the first and second floral whorls.

Loss of A function not only results in loss of sepal and petal identity, but can also make the flower more shoot-like, indicating a role for the A function genes in promoting identity of the floral meristem (Figure 1. 3-b). In contrast, the C function gene *AGAMOUS* (*AG*) directs development of stamens and the carpel within the central two floral whorls. Absence of C function results in a complete loss of reproductive organs with stamens homeotically replaced by petals, and carpels by sepals (Figure 1. 3-d).

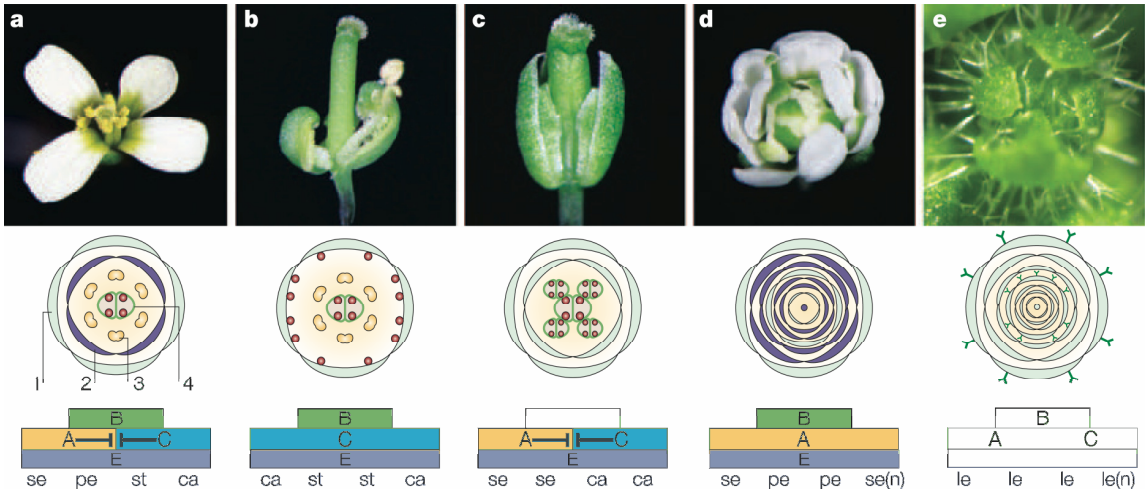

Figure 1. 3: The ABCE model for floral organ specification. a: An *Arabidopsis* flower consists of four whorls of floral organs. Each organ type is specified by a combination of activities, which are mediated by the floral homeotic genes. The so-called A function genes control the development of sepals and petals; the B function genes regulate petal and stamen development; and C function genes control the development of stamens and carpels. Furthermore, A and C function genes are predicted to act antagonistically to establish mutually exclusive domains of activity. E function genes are required for the specification of all organ types. (b-e) Flowers of floral homeotic mutants. b: *apetala2* (A function); c: *pistillata* (B function); d: *agamous* (C function); e: *sepallata1-4* quadruple mutant (E function) (Krizek and Fletcher, 2005).

AG also has a role in promoting indeterminacy of the flower because *ag* mutants produce additional whorls of alternating petals and sepals. In wild type plants *AG* switches off *WUSCHEL* (*WUS*), a major meristem maintenance gene, limiting stem cell production.

The B function genes *APETALA3* (*AP3*) and *PISTILLATA* (*PI*) regulate proper development of the petals and stamens in the second and third whorls respectively. The absence of *PI* results in the replacement of petal and stamen with sepal and carpelloid organs (Figure 1. 3-c). With the exception of *AP2*, the floral homeotic

genes all encode MADS-box transcription factors. The ability of the ABC proteins to specify floral organ identity is dependent on the presence of their flower-specific co-factors encoded by four *SEPALLATA* genes (*SEP1*, *SEP2*, *SEP3*, *SEP4*), which also encode MADS-box proteins and have collectively been termed the E function. The flower phenotype of loss of E activity is quite dramatic as the flowers have continuous whorls of leaf-like organs (Figure 1. 3-e).

Single mutants for each of the components of the EMF2 PRC2-like complex present homeotic flower defects. The EMF2 PRC2-like complex is composed by CLF, EMF2, FIE and MSI1 (Table 1. 1). Its main role is to represses flowering during vegetative development. The *clf* mutant flowers earlier than wild type plants and presents a characteristic "curly leaf" phenotype (Clf-) (Goodrich et al., 1997). This mutant background will be presented in more detail in the following section. *emf2* mutant plants flower soon after germination without making any rosette leaf. It exhibits a very early flowering phenotype and cauline leaves do not show any curling. Strong alleles like *emf2-3* have sterile flowers (Chanvivattana et al., 2004). Expression analysis showed that genes involved in the integration of flowering (*FT*, *SOC1*), in the floral meristem identity (*LFY*) and in floral organ identity (*AP1*, *AP3*, *PI* and *AG*) are highly misregulated in the *emf2* background (Moon et al., 2003). These results were confirmed by the use of floral reporter lines *pAG-I::GUS* and *pAP3::GUS* (Chanvivattana et al., 2004). Independent co-suppression of *MSI1* and *FIE* also show high mis-regulation of the floral organ identity gene *AG* (Katz et al., 2004; Bouveret et al., 2006). Also, in plants co-suppressing *FIE*, homeotic conversions could be observed (sepals and petals forming carpeloid-like organs) (Katz et al., 2004). The set of phenotypes observed suggest that PcG genes assembled in the EMF2 PRC2-like complex repress the expression of floral organ identity genes in *A. thaliana*. Moreover, combined with Yeast-two-Hybrid assays, it confirms that the members of the EMF2 PRC2-like complex physically interact together to perform this regulation (Chanvivattana et al., 2004).

1.3. The *curly leaf* (*clf*) mutant

1.3.1. The Clf- phenotype

clf plants exhibit a strong phenotype during vegetative development. A null *clf* mutation induces pleiotropic effects on leaf morphology, flower organ formation and flowering time (Goodrich et al., 1997) (Figure 1. 4). Rosette and cauline leaves are narrow and curl following the longitudinal axis. The flowers open early because the sepals fail to seal the floral bud, petals are smaller and narrower than in wild type and homeotic transformations (sepal/carpel; petal/stamen) can be observed (Goodrich et al., 1997).

Figure 1. 4: The Clf- phenotype. Wild type (Ler) and *clf-2* plants growing under Long Days (LD) conditions. *clf-2* mutants bolt and form flowers earlier than Ws plants.

Concerning the flowering time, *clf* mutants flower early in Short Day (SD) conditions, while in Long Day (LD) conditions, *clf* mutants flower slightly earlier than wild type plants (Goodrich et al., 1997) (Figure 1. 4).

1.3.2. *AG* is ectopically expressed in *clf* plants

The floral organ identity genes *AG* and *AP3* are highly mis-regulated in the *clf* mutant background (Goodrich et al., 1997; Chanvivattana et al., 2004). However, the Clf- phenotype is essentially due to the mis-regulation of *AG* as transgenic plants expressing constitutively *AG* (*35S::AG*), showed leaf curling, early flowering and the same nature of homeotic transformations of those observed in the *clf* mutant as

previously cited (Mizukami and Ma, 1992). It has been shown that CLF is necessary for the repression of *AG* during vegetative development (Schubert et al., 2006).

1.3.3. *AG* and *CLF* interact genetically

As *clf* and transgenic plants over-expressing *AG* showed very analogous phenotypes, it was thought that the Clf- phenotype was essentially due to the high mis-expression of *AG*. The double mutant *ag-3 clf-2* showed no leaf curling and also the early flowering time defects of the *clf-2* plants were partially restored (Figure 1. 5).

Figure 1. 5: Genetic interaction between *CLF* and *AG*. *ag-3* and *ag-3 clf-2* plants grew under Shorth Days (SD) conditions. *ag-3 clf-2* mutants bolt and form flowers earlier than *ag-3* mutants.

Nonetheless, the double mutant *ag-3 clf-2* persisted in having homeotic defects. This last result suggested that the Clf- phenotype was not entirely due to the ectopic expression of *AG* (Goodrich et al., 1997).

1.3.4. *CLF* is a PRC2-like member

The double mutant *clf-50 swn-3* and the single mutant *fie TK114*, in which *FIE* is expressed only during very early embryo development, but not during later seedling development exhibit similar severe phenotypes. The PRC2 activity in these mutants is considered to be lost in vegetative development (Kinoshita et al., 2001; Chanvivattana et al., 2004). The nuclear distribution of H3K27m, H3K27me2 and H3K27me3 were assessed in these mutants by immunohistolocalization. The results showed that the distribution in the nucleus of H3K27me2 and H3K27me3 were affected. A reduction of H3K27me2 and H3K27me3 in euchromatin and an enhancement of H3K27me3 at the chromocenters was observed (Lindroth et al., 2004). This results strongly suggest that SWN, CLF and FIE are part of the same regulatory complex (PRC2-like) in *A. thaliana*. It also suggests that the PRC2-like complex represses gene expression through the H3K27me3 histone modification.

1.3.5. CLF is a putative histone methyltransferase (HMTase)

It has been shown that *AG* is a direct target of *CLF* during vegetative development. Indeed, in *clf* mutants the *AG* locus shows very low H3K27me3 whereas in wild type plants, the *AG* locus is enriched for both the CLF protein and the H3K27me3 repressive epigenetic mark (Schubert et al., 2006). In addition, strong experimental evidence showed that *CLF* represses the expression of FLC, *MADS AFFECTING FLOWERING 4* (*MAF4*), *MADS AFFECTING FLOWERING 5* (*MAF5*) and *FT* through a PRC2-like complex containing *EMF2* and *FIE*. Moreover, this repression is due to the direct binding of CLF and the deposition of the repressive H3K27me3 mark at *FLC*, *MAF4*, *MAF5* and *FT* loci (Jiang et al., 2008). Taken together, these results strongly suggest that CLF has a methyl transferase (HMTase) activity within the PRC2-like in *A. thaliana*. These elements provide strong genetic and molecular evidence of the catalytic HMTase activity of CLF. However, the enzymatic activity of CLF, and the others E(Z) homologs in *A. thaliana* SWN and MEA, still has to be

experimentally proven. This can be realized by purifying complexes and showing the HMTase activity towards histones *in-vitro*. Finally, there is no evidence in *D. melanogaster*, *A. thaliana* or any other organism that CLF or any of its homologous proteins is able to methylate other proteins other than histones.

1.4.Redundancy masks the role of *CLF*

There is evidence showing that CLF and EMF2 act together in the same complex (Jiang et al., 2008). The *emf2* mutant have a much more severe phenotype than the single mutants *clf* and *swn* but it is less affected than the double mutant *clf swn*. Thus, the *swn* mutation enhance the Clf- phenotype (Chanvivattana et al., 2004). This observation suggests that *CLF* act redundantly with *SWN*, in repressing gene expression. It has been shown that *SWN* and *CLF* share the same pattern of expression during vegetative development (Goodrich et al., 1997; Chanvivattana et al., 2004). *swn* mutants do not show developmental defects during vegetative or reproductive development, *swn* plants look like wild type plants. In the double mutant *clf swn*, the aerial and root meristems fail to form organs thus, these plants exhibit a callus-like phenotype only viable in *in-vitro* conditions (Figure 1. 6) (Chanvivattana et al., 2004).

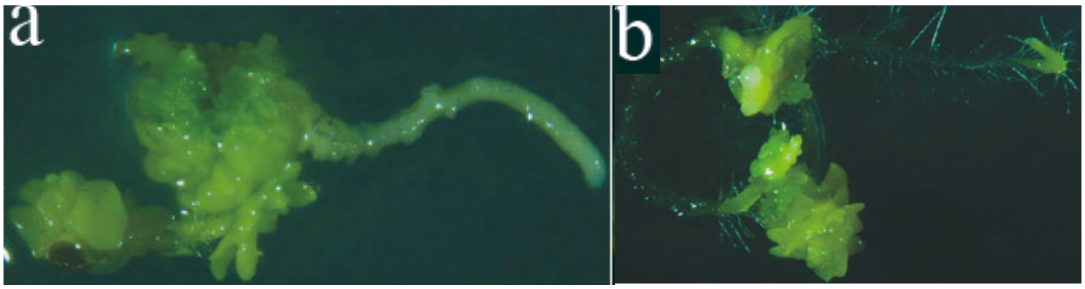

Figure 1. 6: Phenotype of double PcG gene mutants partially redundant in *A. thaliana*. (a) *clf-50 swn-3* double mutant grown in tissue culture. (b) *vrn2-1 emf2-3* double mutant grown in tissue culture. (Chanvivattana et al., 2004; Schubert et al., 2005).

As shown in Figure 1. 6, there is a similar enhancement of the Emf2- phenotype by the *vrn2* mutation showing that there is also a redundancy between *VRN2* and *EMF2* in *A. thaliana* (Schubert et al., 2005). The enhancement of the Clf- and Emf2- phenotypes by *swn* and *vrn2* mutations respectively, suggests that both *SWN*/*CLF* and *VRN2*/*EMF2* act redundantly to regulate different developmental processes in *A. thaliana*.

1.5. The PRC1 in *A. thaliana*

Proteins homologous to members of the PRC1 in *D. melanogaster* have not been found in *A. thaliana*. However, *TFL2*/*LHP1* in *A. thaliana*, has been shown to be structurally homologous to *HETEROCHROMATIN PROTEIN 1* (*HP1*) in animals (Kotake et al., 2003). In the *tfl2* mutant, PcG targets as *FT*, *AP3*, *PI* and *AG* are highly mis-regulated (Kotake et al., 2003; Nakahigashi et al., 2005) It has been proposed that LHP1/TFL2 functionally replaces the PRC1 in *A. thaliana* as it binds H3K27me3 with high affinity throughout the genome (Turck et al., 2007; Zhang et al., 2007a). Another putative substitute member for the PRC1 in *A. thaliana* is *EMBRYONIC FLOWER 1* (*EMF1*). As for *tlf2*, in *emf1* plants, the expression of PcG targets *AG* and SHOOTMERISTEMLESS (*STM*) is up regulated. The up-regulation of *AG* and *STM* in *emf1* plants is related with loss of H3K27me3 at these loci (Calonje et al., 2008). Moreover, EMF1 is able to directly interact with *AG* sequences and this interaction is EMF2 dependent. Therefore, the authors claimed that EMF1 plays a PRC1-like role in floral development (Calonje et al., 2008). However, if the levels of H3K27me3 decrease in the *emf1* mutant, could equally be that EMF1 is needed by the PRC2 to deposit the repressive mark.

Very recent computing and structure analyses of the conserved domain of the C-terminal region of the PRC1 RING protein in *D. melanogaster* showed significant structural similarity with a new ubiquitin-like domain called RAWUL. Moreover,

the RAWUL domain was found in proteins of *A. thaliana* and *Caenorhabditis elegans* suggesting a functional conservation of the PRC1 all through eukaryotes (Sanchez-Pulido et al., 2008). However, proteins bearing the RAWUL domain has not been experimentally characterized and their PRC1-like putative activity in *A. thaliana* remain to be proven.

1.6. Genetic screen to find new PcG genes

To date, several main players of the PcG genes regulatory process have not been identified in *A. thaliana*. The question of how the repressive mark made by the PRC2 is specified/targeted, interpreted, maintained and erased in *A. thaliana* is poorly undertood. For example, DNA specific sequences (PREs) permitting the targeting of the PRC2-like complexes, and proteins like PHO, PSQ or PC that bind to the PREs have not been identified in *A. thaliana*. Functionally, no homolog of the PRC1 complex was been found in the *A. thaliana* genome sequence. This suggests that the H3K27me3 deposited by the PRC2 is recognised, read and interpreted by another mechanism that firstly characterized in *D. melanogaster*. Moreover, there are very few PcG gene targets known in *A. thaliana* and their mis-regulation do not totally explain the phenotypes observed in the double mutants *clf swn* or *emf2 vrn2*. Recent whole genome profiling studies have suggested that PcG genes regulate many more developmental regulators than floral homeotic genes (about 15% of *A. thaliana* genes carry the H3K27me3) (Zhang et al., 2007b). However, the relevance of PcG genes regulation at these targets during plant development is unclear.

1.6.1. Identification of PcG and trxG genes in *D. melanogaster*

In *D. melanogaster*, genetic screens for modifiers of PcG mutant phenotypes have been exceptionally successful. Many of the *D. melanogaster* PcG genes were identified for their ability to enhance the phenotype of other PcG genes (Jurgens,

1985; Ringrose and Paro, 2004). *PSQ* for example, enhances the phenotype of *PC* and *PSC* mutants (Huang et al., 2002). Genes belonging to the trxG complexes were identified as suppressors of PcG genes phenotypes. For example, a mutation in *ABSENT SMALL HOMEOTIC 1* (*ASH1*) suppresses the Esc- phenotype (Ringrose and Paro, 2004). Using similar approaches, chromatin remodelling factors as the human homolog of *TRX* (*MLL*) which has a H3K4 methyltransferase activity associated with *HOX* gene activation, has been identified (Milne et al., 2002; Ringrose and Paro, 2004).

1.6.2. Identification of *PcG* genes in *A. thaliana*

Members of the PRC2 in *A. thaliana* have been also identified in genetic screens (Table 1. 1). For example, the *FERTILIZATION-INDEPENDENT SEED* (*FIS*) mutants: *MEA* initially named as *FIS1*, *FIE* initially named as *FIS3* and *FIS2* were isolated and identified from a mutated population as suppressors of the Pi- (Pistillata-) phenotype. The *pistillata* (*pi*) mutant has short siliques without seed. The *fis* mutants in the *pi* background have long siliques containing developing seeds, even though fertilization did not occur (Chaudhury et al., 1997). Also, *CLF* and *EMF2* were identified in independent screens for mutants with early flowering time phenotypes (Goodrich et al., 1997; Yoshida et al., 2001). *VRN1* and *VRN2* were isolated from a screen for altered vernalization response in *fca*, a member of the autonomous pathway that constitutively represses *FLC* expression (Figure 1. 2) (Komeda, 2004)) (Chandler et al., 1996).

Genetic screens for further PcG proteins in *A. thaliana* have not been as successful as in *D. melanogaster* for several reasons (i) in *A. thaliana*, each component of the PRC2 is represented by a small gene family (e. g. *E(Z)* is represented by *CLF*, *SWN* and *MEA*), (ii) the phenotypes observed in *A. thaliana* are less specific than in *D. melanogaster* and (iii) it seems that the number of targets in *A. thaliana* is way more important than in *D. melanogaster*.

1.6.3. Mutagenesis of the *clf* mutant background

Screens for modifiers of PcG genes mutant phenotypes in *D. melanogaster* have been very successful in identifying additional PcG genes and factors regulating chromatin structure and regulation. Plants as an exception to other model systems, are more tolerant to the accumulation of mutations (*clf swn* for example). Genetic screens have allowed the isolation of PcG genes directly involved in process like the endosperm/embryo development (Chaudhury et al., 1997) as presented previously. One of the key features of *clf* plants is that their phenotype corresponds to intermediate level of PcG activity *in planta* (Figure 1. 7). In the *clf* background, additional mutations might increase or decrease the PcG activity. These variations in PcG activity are reflected by the plant phenotype (Figure 1. 7). Indeed, as shown earlier the *swn* mutation enhances the Clf- phenotype by decreasing the PcG activity in the double mutant *clf swn* (Figure 1. 7).

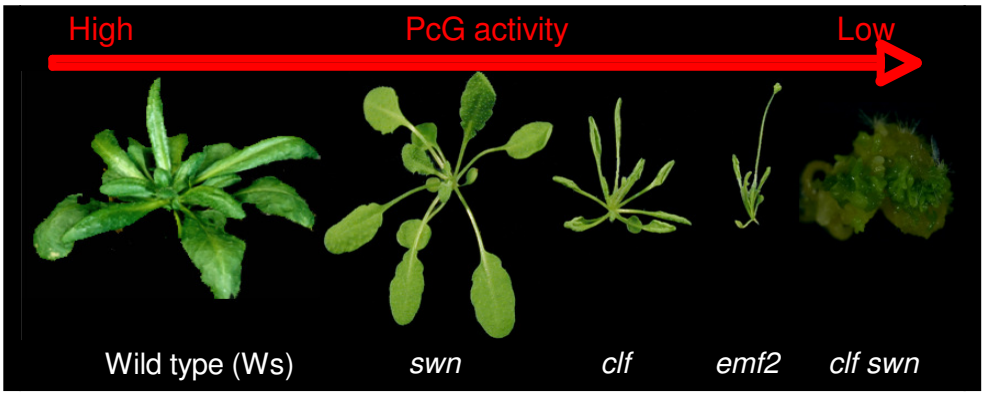

Figure 1. 7: Spectrum of phenotypes of PcG mutants reflects the level of PcG activity. Removing single PcG gene activity such as *SWN* has only subtle effects on development, whereas removing single PcG gene activity as *CLF* or *EMF2* exhibit more severe defects suggesting that they play an important role in development. The *clf swn* double mutant likely lacks all PcG activity during post-germination development. Subsequently, the double mutant *clf swn* is only viable in tissue culture.

Conversely, the *ag* mutation suppresses the Clf- phenotype as it is a target of *CLF*. Therefore, *clf* is a perfect mutant background for genetic screens for modifiers (enhancers and suppressors), as both types of mutations can be easily discriminated from the original mutant background.

Previous to this work, two different genetic screens in the *clf* mutant background were carried out in the laboratory. The first screen used a chemical ethylmethane sulfonate (EMS) mutagenesis and the second used an insertional (T-DNA) mutagenesis. These two approaches have different advantages. The T-DNA mutagenesis is less efficient but allows the isolation of tagged locus rapidly whereas the EMS mutagenesis generate more mutations. The isolation of weak mutations by map-based cloning can be very frustrating and time consuming. In the following sections, the T-DNA mutagenesis is described as a primer to this thesis.

1.6.4. T-DNA mutagenesis of Clf- plants with an activation-tagging vector

T-DNA mutagenesis needs vigorous plants able to produce abundant and fertile flowers for the floral dip transformation (Clough and Bent, 1998). *clf* plants are weak and their fertility is very reduced (Figure 1. 8 left). The T-DNA mutagenesis was done in the *clf-50* background but, the deleterious Clf- phenotype was rescued by a homozygous single copy of the *p(CLF)::CLF-GR* transgene. It can rescue the Clf- phenotype in an inducible (steroid-dependent) manner.

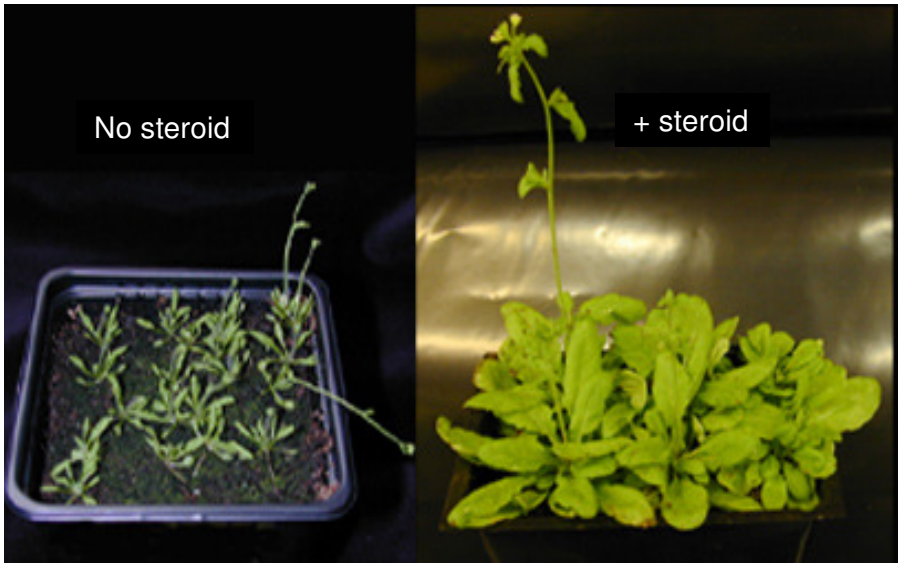

Figure 1. 8: Phenotype of the *clf-50 pCLF:CLF-GR* transgenic line with and without steroid. The CLF activity is chemically restored by the dexamethasone treatment (right) whereas without treatment (left), the CLF activity is absent. Dexamethasone (1 µM) was applied twice a week by direct spraying. All plants were grown under Long Days (LD) conditions.

Indeed, the fusion protein CLF-GR has steroid dependent CLF+ activity thanks to the binding domain of the rat glucocorticoid steroid receptor (-GR). Plants grown with the steroid dexamethasone (dex) have wild-type phenotypes whereas plants grown without dex have the Clf- severe phenotype (Figure 1. 8).

The T-DNA used for the mutagenesis is an activation tagging T-DNA carried by the plasmid: pSKI074 (Figure 1. 9) (Weigel et al., 2000). This T-DNA is able to give a dominant gain of function mutations if it is inserted upstream of a locus. This is due to the over-expression of the locus targeted. Alternatively, it can also induce a recessive loss of function mutations if the insertion is intragenic and cause disruption of the native transcript.

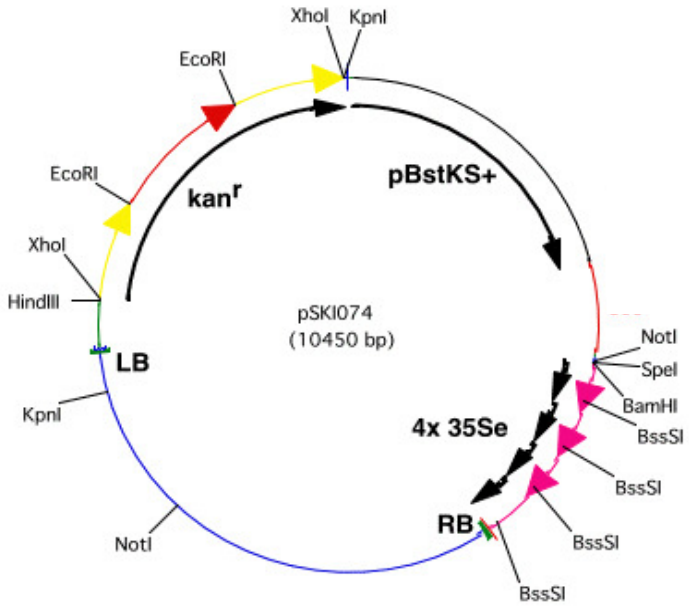

Figure 1. 9: Diagram of the pSKI074. 4x35Se: are the 4x35SCaMV enhancer sequences allowing gain of function mutations. pBstKS+: corresponds to the sequence allowing propagation in *E. coli* (containing an origin of replication and a selective marker specific for *E. coli* (ampicillin). Kan^r: is the plant selection gene for Kanamycin. LB and RB: are the Left and Right T-DNA flanking sequences. These sequences are crucial for the transfert and the integration of the T-DNA in *A. thaliana*'s genome. (http://wigelworld.org/resources/plasmid/activationtagging/pski074).

By using the activation tagging strategy, it is possible to create very different types of mutations. Also, if the mutation is tagged, it allows a rapid and accurate identification of locus/loci mutated. Indeed, the plasmid sequences of the pSKI074 flanking the T-DNA have several restriction enzyme sites that can be used to rescue the Right Border (RB) of the T-DNA and adjacent plant sequences from transformed plants (see Methods for detail) (Weigel et al., 2000). Finally, the rapid gene isolation is very useful for transgenic lines with the less T-DNA insertions as possible. Indeed, the genomic sequence of the insertion site can be particularly enlightening. If the T-DNA mutates in any way a protein belonging to the PcG genes or known to

be a factor involved in chromatin organization, the mutation will be studied in detail.

1.6.5. Screen for modifiers of the Clf- phenotype

In the laboratory, an *A. thaliana* mutagenized population with the activation tagging of 15,000 T2 families was available to screen. Each T2 family derived from 2-3 T1 individual plants. During my PhD, I screened 1,000 T2 families for modifiers of the Clf- phenotype. Each of the T2 families (around 50 seeds) was sowed in *in-vitro* tissue culture (MS plates). The screen isolation of enhancers and suppressors took place at the seedling stage (15 to 20 days old). In 13 and 6 independent families, suppressed and enhanced plants respectively were found.

The enhancer mutations isolated seemed rather heterogeneous and it was not clear if they were relevant to PcG function as they didn't have obvious phenotypes like the ones observed in the double mutant *clf swn* and *emf2 vrn2*. Moreover, the enhancer mutations lead to a mixture of different phenotypes that generally made the plants a bit smaller and sicker. The isolation of these mutations was not pursued further.

Alternatively, the suppressor mutations isolated showed in all cases a more homogenous phenotype. Suggesting that the mutations were less pleiotropic and in the same time more specific of the PcG silencing pathway. Attention was focused on the suppressor mutations.

1.7. Aims

Mutations that suppress the Clf- phenotype can be (i) mutations that prevent the ectopic activation of *AG* or any other target of *CLF* or (ii) mutations that act downstream of *CLF* or *AG* as previously explained and discussed in the introduction. The aims of this PhD were therefore to:

1. Identify the suppressor mutations that are likely to be T-DNA tagged.

2. Clone by plasmid rescue or genome walker plant genomic sequences flanking the T-DNA insertion.

3. Characterize the mutations suppressing the Clf- phenotype, and find a relationship between *CLF* and the new mutations in the silencing pathway.

2. Material and Methods

2.1. Plant material

All the experiments described in this thesis were realized with *A. thaliana* plants. The wild type ecotypes used were: Columbia-0 (Col-0), Landsberg *erecta* (L*er*) and Wassilewskija (Ws).

2.1.1. *curly leaf* (*clf*) mutant lines

clf-50

This allele was identified in a T-DNA mutagenesis screen in the Ws background. Molecular evidence has shown that the *CLF* locus has been deleted in this line; therefore *clf-50* represents a null allele. The *clf-50* line was provided by Dr. E. Huala.

clf-50 pCLF:CLF-GR

This transgenic line was used for the activation tagging T-DNA mutagenesis. It carries the null *clf-50* allele and a single copy of the *pCLF:CLF-GR* transgene (Schubert et al., 2006) that can rescue the mutant Clf- phenotype in an inducible (steroid dependent) manner. The fusion protein CLF-GR has steroid dependent CLF+ activity thanks to the binding domain of the rat glucocorticoid steroid receptor (-GR). Plants grown with the steroid dexamethasone (dex) have wild-type phenotypes whereas plants grown without dex have a severe Clf- phenotype.

clf-81

This line is in the Columbia background and was isolated from an EMS mutagenesis screen. It carries a mis-sense mutation that encoded the substitution R794H within the SET domain of *CLF*. It shows a severe leaf curling and an early flower plenotype (Schubert et al., 2006).

clf-28

This line is in the Columbia background and was obtained from the Salk Institute Genome Analysis Laboratory collection (SALK_139371). The T-DNA insertion is located in the fourthly exon. The severe phenotype of the double mutant *clf-28 swn-7* suggests that *clf-28* is a null allele.

FRIGIDA^{+} (*FRI*$^{+}$*) clf-52*

The *clf-52* mutation was isolated in Dr. R. Amasino's laboratory in a genetic screen for suppressors of the late flowering phenotype of *FRI*$^{+}$ in the Ws background. The *clf-52* allele is a loss of function mutation due to an insertion of a T-DNA at *CLF*'s coding region.

2.1.2. Other mutant lines

sepallata3-2 (*sep3-2*)

This allele was found in a population of *A. thaliana* (Col) plants that carries independent insertions of the autonomous maize transposable element *En-1*. The *En-1* element is inserted into *SEP3*, 48 bases after the start codon, in the first exon, with the 3' end towards the start of the gene. This allele was kindly provided by Prof. M. Yanofsky.

flowering locus t-10 (*ft-10*)

This allele of *FT* was isolated from a T-DNA library generated by GABI-Kat, in which a T-DNA is inserted in the first intron of *FT* in position +844bp. It is a recessive loss of function mutation.

FRI⁺ Col-0, *FRI⁺ Ler* and *FRI⁺ Ws*

The active FRI allele from the San Feliu 2 (Sf-2) accession was introgressed in the wild type ecotypes Col-0, Ler and Ws. All three lines were kindly provided by Prof. R. Amasino.

flowering locus c-3 (*flc-3*)

This *FLC* allele was obtained in a fast-neutron mutagenesis. It carries a small deletion. The *flc-3* allele is in the *FRI⁺* Col-0 background. This line was kindly provided by Prof. R. Amasino.

fpa-6, *fpa-7* and *fpa-8*

The *fpa-6* allele is a lesion caused by a T-DNA insertion/deletion in the Ws background. It was shown that a 2.5 Kbp deletion removed the entire 5' regulatory region and the first exon of *FPA* (Schomburg et al., 2001). The *fpa-7* allele is in the Col-0 background and is a T-DNA insertion in intron 5 of FPA (line SALK_021959). This allele was obtained from Dr. G. Simpson's lab and corresponds to the SALK insertion referenced SALK_138449. The *fpa-8* allele is also in the Col-0 background and was obtained from a Ethyl-Methane Sulfonate (EMS) mutagenesis. It is a missense mutation: TRP98-> Stop. It is a recessive loss of function mutation.

fpa-8 35S:FPA-YFP

The translational *FPA-YFP* protein fusion was cloned into the binary vector *pBarM:35S*. The resulting construct was transformed into *fpa-8* mutants using the floral dip method (Baurle et al., 2007).

agamous-2 (*ag-2*)

This allele of *AGAMOUS* (*AG*) was found in a T-DNA mutagenesis. Sequence analysis showed that the T-DNA is inserted in the second intron of *AG* and

interferes with normal transcription and splicing of *AG* (Yanofsky et al., 1990). It is a recessive loss of function.

2.2. Bacterial material and microbiology techniques

2.2.1. Strain of *Escherichia coli* (*E. coli*) used for cloning: DH5α

The strain DH5α (F- φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F)U169 *rec*A1 *end*A1 *hsd*R17(r_k^-, m_k^+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ$^-$) was used in this study to clone plant genomic DNA fragments and PCR fragments from genome walker. This strain allows cloning and conservation of foreign DNA fragments, without genetic modifications like recombination, as the *REC-A* activity is crucial for recombination between DNA molecules. Moreover, this strain carries part of the LAC-Z gene necessary for the α complementation for the white-blue selection of clones that carry an insert.

2.2.2. Cloning plasmids into *E. coli* DH5α

In order to clone the vectors obtained by plasmid rescue, commercial electroporation competent DH5 α *E. coli* cells from Stratagene (ElectroTen-Blue®) were used. For each transformation, 40µL of E. coli cells were thawed on ice for ten minutes before 10µL of ligation was added, and mixed gently by flicking the tube. The mixture was then transfered to a chilled electroporation cuvette that has a 0.1 cm gap. The samples were pulsed once at the recommended settings (applied volts (0.1 cm gap): 2250V; field strength (0.1-cm gap): 22.5 KV/cm; resistance: 200Ω and capacitance: 25µF). Then, the cuvette was quickly removed and 1mL of sterile SOC medium (nutrient-rich bacterial growth medium used for microbiological culture), at 37°C was added to resuspend the cells. The cells were incubated for 45 minutes (min) at 37°C with shaking at 225 rpm to allow expression of the appropriate antibiotic resistance or selectable marker gene carried by the plasmid. The cells were

then pelleted by centrifugation at 4000 x g for 30 seconds, the supernatant removed and the cells resuspended in 50 μL of SOC. The entire volume was then plated on LB agar plates with the appropriate antibiotic.

2.2.3. PCR products cloned into plasmids and propagated in *E. coli*

The Fast and Easy GeneJET™ kit (Fermentas) was used to clone PCR products. This kit contains a specific plasmid (pJET1/blunt cloning vector) in which the gene for a lethal restriction enzyme for all *E.coli* strains commonly used for cloning is present. Ligation of a PCR fragment into the cloning site disrupts this lethal gene. As a result, only cells with recombinant plasmids can be propagated. The cloning steps were carried out according to manufacturers' guidelines (Fermentas).

2.2.4. Culture of *E.coli* cells on plates

Luria Broth (LB), a bacterial growth medium used for microbiological culture plus agar was autoclaved, cooled to approximately 55°C, the appropriate antibiotics were added (ampicillin 100 mg/mL) and the media poured into Petri dishes. When the agar had set, *E. coli* cells were spread over the surface. The *E. coli* was then incubated at 37°C overnight.

2.2.5. Culture of *E. coli* in liquid medium

Antibiotics were used at the same concentration as used for agar plates (ampicillin 100 mg/mL). Liquid cultures were incubated at 37°C with shaking at 225 rpm; 5 mL cultures were inoculated with a single colony and grown overnight.

2.2.6. Conservation of *E. coli* clones

For long term storage of bacterial clones, glycerol stocks containing equal volumes of 80% glycerol:bacteria solution were made, snap frozen in N_2(l) and stored at -80˚C.

2.3.Plant culture and genetic analysis of *A. thaliana*

2.3.1. Growth conditions for *A. thaliana* in-vitro and on soil

For growth in sterile *in-vitro* conditions, seed were sterilised by two different methods:

(i) For the T-DNA screen, seeds were exposed overnight to chlorine fumes. This was done by mixing 100mL sodium hypochlorite (NaClO) solution (6%) with 10mL of concentrated HCl (36%) in a dessicator jar. (ii) For a small number of sterilisations, seeds were sterilised by rinsing in 70% ethanol (EtOH), 0.05% Triton X-100 for 10 min. The seeds were then washed with 95% EtOH and dispensed onto sterile 3MM filter paper in the sterile hood to allow the EtOH to evaporate. Sterile seeds were grown on Murashige and Skoog (MS) nutrient medium agar plates (0.5X MS salts and vitamin mix (Gibco/BRL, Gaithersburgh, MD, USA), 0.6% sucrose, 1% microagar (Detroit, MI, USA); pH 5.7, with the appropriate antibiotic if necessary). Plates were stratified at 4˚C for two days to synchronise germination before transferring them to the growth chambers where they were grown in Long Day (LD: 21˚C; 16h of light/day and 50-65% humidity) or Short Day (SD: 21˚C; 9h of light/day and 50% humidity) conditions. To score kanamycin resistant ratios, seeds were sterilized and grown on plates with 50 mg/mL of kanamycin. The resistant or sensitive phenotype was scored after 14 days of growth in *in-vitro* tissue culture.

For growing plants on soil, the seeds where planted on 2:1:1 mixture of Levington F2 compost, sand and fine grit. The seeds were planted out and stratified as previously described. The seed trays were covered with clear plastic domes during the first week to maintain high levels of humidity. The conditions for growth were either LD (16h of light/day) or SD (9h of light/day) both at a temperature of 21˚C and 50% humidity.

2.3.2. Genetic crosses

Crosses between plant lines were made by first emasculating a flower that had not yet dehisced on the acceptor plant (female). Carpels were then hand pollinated one day after with pollen from two/three flowers from the donor plant (male). The emasculation and the pollinisation were done with sterile (rinsed with 70% EtOH) tweezers. Mature seedpods were collected and seed were conserved in dry conditions.

2.3.3. Flowering time assessment

Counting the number of rosette-leaves plants possessed at the time of a 1 cm bolt was formed assessed the flowering time in long days. In short day (SD) conditions the same approach was used, but leaf counting was carried out on a weekly basis to make sure no leaves were omitted. The number of leaves of twelve plants minimum was counted for each genotype in each experiment. Average and standard error (SE) were calculated as follow: Average = $\sum x_i/N$ and SE = $(\sum(x_i-x_m)^2)/(N\sqrt{(N-1)})$. Where x_i is the number of leaves from a single plant for a specific genotype, N is the total number of plants for a specific genotype and x_m is the average number of leaves for a specific genotype.

2.3.4. Leaf area calculation

The leaf area was calculated using a freeware called ImageTool (IT). It was developed in the Department of Dental Diagnostic Science at The University of Texas Health Science Center, San Antonio, Texas. The program was developed by C. Donald Wilcox, S. Brent Dove, W. Doss McDavid and David B. Greer. (http://ddsdx.uthscsa.edu/dig /itdesc.html).

2.3.5. Starch assay

The presence of starch in leaves was qualitatively determined by staining for 30 minutes with a solution containing iodine (5.7 mM) and potassium iodide (43.4 mM) in HCI (0.2 N). Leaves were decolorized before staining by soaking in 96% ethanol for about 6 h (Caspar et al., 1985).

2.4. DNA techniques

2.4.1. *A. thaliana* genomic DNA extraction

Leaf samples were collected, frozen and stored at -80˚C prior to DNA extraction in Eppendorf tubes. Frozen leaves were crushed to a fine powder using a mini-pestle inside the tube. 500µL of extraction buffer (50mM EDTA, 0.1M NaCl, 0.1M pH8 Tris-HCl, 1% SDS) was then added and mixed with the powder. Samples were then placed in liquid nitrogen before thawing them, first on the bench at room temperature, and then at 65˚C for 5 minutes. 500µL of phenol/chloroform (equilibrated phenol:chloroform:isoamylalcohol 25:24:1), were added to the sample to remove contaminants from the extraction preceding vortex agitation and rest for five minutes on the bench. The upper (aqueous) phase was removed and the phenol/chloroform extraction repeated. To the second aqueous extract (400µL), 50µL of 3M NaAc (pH5.2) and 350µL of isopropanol were added. The sample was

inverted gently to precipitate DNA, then centrifuged at 14,000rpm for five minutes. The pellet was rinsed with 70% EtOH, then air-dried before being resuspended in 50µL of R40 [TE (10mM Tris-HCl pH8, 1mM EDTA) containing 5µg/mL of ribonuclease A (pre-boiled (100˚C) for five minutes to remove DNase activity and stored as a 1mg/mL stock)], and stored at -20˚C. The genomic DNA obtained was used for Southern-blot analysis and diluted ten times for PCR analysis.

2.4.2. PCR reaction

For each PCR reaction, a master mix was made by mixing the following constituents and then 1µL of plasmid/genomic DNA was added. 2µL of 10X PCR buffer (500mM KCl, 100mM Tris-HCl pH9, 1% Triton X-100), 2µL of 25mM $MgCl_2$, 1µL of 10µM forward primer, 1µL of 10µM reverse primer, 0.5µL of 10mM dNTPs (10mM of each: dATP, dCTP, dGTP, dTTP) and 0.2µL of Taq DNA polymerase enzyme (Promega, Madison, USA), 12.3µL of dH_2O. Oligonucleotide stocks were either already available to use in the laboratory or ordered from VH Bio (UK). The primers were designed using the Primer3 software available online at the following URL: http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi. Each primer had an optimal ratio of 30-50% dGTP/dCTP and a G- clamp at the 3' end. For primer pairs, the melting temperatures were chosen to be compatible. PCR tubes were placed in a MJ Systems DNA engine thermocycler preheated to 94˚C and the following program run: 94˚C for two minutes, followed by 30-35 cycles of [94˚C for 30 seconds (denaturing), 53˚C/60˚C (depending on primer set) for 30 seconds (annealing), 72˚C for 30 seconds per expected 500 base pair (bp) (extension)], followed by 72˚C for five minutes to complete extension and ending by 4˚C indefinitely.

2.4.3. Primers used in this study

In the following table, I listed all the primers used in this study and to which all the experiments described in this manuscript are related. The name, the sequence and the purpose are listed below.

Table 2. 1: List of oligonucleotides used in this study to genotype mutants, RT-PCR and ChIP PCRs.

Primer name	SEQUENCE (5'-3')	Purpose
3AGL9 (primer n˚1)	AGTGGCCAGTGAGTGCTCT	Mapping *sep3-7*
3aAGL9 (primer n˚2)	TGGTTAGGGTTCAGCTGGAG	Mapping *sep3-7*
3bAGL9 (primer n˚3)	TTGGTTACATTACTTTTAGATGAGGAA	Mapping and genotyping *sep3-7*
3cAGL9 (primer n˚4)	CTTTTCAGCGTGCAATGAAA	Mapping *sep3-7*
3dAGL9 (primer n˚5)	TGAGGATATAGCACCTGGCC	Mapping *sep3-7*
3eAGL9 (primer n˚6)	TGGACTTCACTTTGTTGTTGCT	Mapping *sep3-7*
3fAGL9 (primer n˚7)	TGTTCGTCCAAGTTACTGCAA	Mapping *sep3-7*
3gAGL9 (primer n˚8)	TCACCTTCAGGATCAGGTCC	Mapping *sep3-7*
3hAGL9 (primer n˚9)	TTGGGAAACATTTCATGCTG	Mapping *sep3-7*
3iAGL9 (primer n˚10)	CCTTTATGGATGGCCAAGAA	Mapping *sep3-7*
3jAGL9 (primer n˚11)	CCCCACTCAATAAACCCAGA	Mapping and genotyping *sep3-7*
5aAt1g54930 (primer n˚12)	CACAAAGCTTCACCTGTTTCTC	map T-DNA insertion in *At1g54940*
5bAt1g54930 (primer n˚13)	GAATCGATTCCTTTCCCACA	map T-DNA insertion in *At1g54940*
5cAt1g54930 (primer n˚14)	GGTTGATGACAATGGAAGCA	map T-DNA insertion in *At1g54940*
5dAt1g54930 (primer n˚15)	CTGTCTCTGTCTCAACAATAATTCTT	map T-DNA insertion in *At1g54940*
5eAt1g54930 (primer n˚16)	TTTTGAAACAGCAAAAGAATTTTT	map T-DNA insertion in *At1g54940*
5fAt1g54930 (primer n˚17)	TCCATCTTCAACGTTACGGA	map T-DNA insertion in *At1g54940*
At1g54940-SI_R (primer n˚18)	AGGCTTCTGACGAGTGTAGTACTGT	map and genotype T-DNA insertion in *At1g54940*
SEQ	AAACAATACAAAGACAGATAAAGC	Primer used for

		first sequence plasmid rescue plasmids
LB 4	GAAACAGAATACCCGCGAAA	locate at LB in pSKI074
FPAseq	TTTGGAAGCGAGCAAACTCT	Mapping *fpa-10*
FPA F	TCACAAAACTTTGATGATGCTTG	Genotyping *fpa-10*
FPA R2	CTCGATTTTCCCAAACTTGC	Genotyping *fpa-10*
AB32	GTTCTCGCGATCTATATCTTCGCG	genotype *clf-81*
Ab33	GATGTTTCTGGTTGGGGAGCT	genotype *clf-81*
En8130	GAGCGTCGGTCCCCACACTTCTATAC	genotype *sep3-2*
9 X12	GTTGGGAAAATCGTACGAGGCTTCACCTAGT	genotype *sep3-2*
9 ms2	GTCACTTGCCTATTGA	genotype *sep3-2*
JH2295	TAAGCTCAATGATATTCCCGTACA	Genotyping *ft-10*
JH2296	CAGGTTCAAAACAAGCCAAGA	Genotyping *ft-10*
JH2297	CCCATTTGACGTGAATGTAGACAC	Genotyping *ft-10*
(GS)flc3 F	CAAAGACGCTCGTCAGCGG	genotyping *flc-3*
(GS)flc3 R	CAAATGAAAACCCAGGTAAGG	genotyping *flc-3*
clf28F	CTGCCAGTTCAGGAATGGTT	genotyping *clf-28*
clf28R	GAAGGGAGCTCTCTGCTTGAT	genotyping *clf-28*
LBb1.3	ATTTTGCCGATTTCGGAAC	genotyping *clf-28*
(GS)fpa7 F	CATTTTATAATAACGCTGCGGACATCTAC	genotyping *fpa-7*
(GS)fpa7 tdna	TACTCCAATGGGTGTCGATGAGAGGT	genotyping *fpa-7*
(GS)fpa7 R	GACGGGTCACTTGGAGTACATCAG	genotyping *fpa-7*
SY1	TTCGCAAAAGAGAGGATACTGC	Genotyping *clf-50 CLF::GR*
SY2	GTGTTCTTCCGTGTGGAGAGTA	Genotyping *clf-50 CLF::GR*
SY3	TGGTCATCGAGGTGTAAGTTTC	Genotyping *clf-50 CLF::GR*
genewalker LB-1	GTTTCTCATCTAAGCCCCCATT	Genome walker. located at LB in pSKI074
genewalker LB-2	ACGYGAATGTAGACACGTCGAA	Genome walker. located at LB in pSKI074
genewalker AP-1	GTAATACGACTCACTATAGGGC	Genome walker. Adapter
genewalker AP-2	TACGACTCACTATAGGGCACGC	Genome walker. Adapter
TUBF	GTTCTTGATAACGAGGCCTT	RT-PCR on *TUBULIN*
TUBR	ACCTTCTTCCTCATCCTCG	RT-PCR on *TUBULIN*
AGQPCR1F	TCCGAGTATAAGTCTAATGCC	RT-PCR on *AGAMOUS*
AGQPCR1r	GCCTATATTACACTAACTGGAGAG	RT-PCR on *AGAMOUS*

RT SEP3_F	TGACGTTTGCAAAGAGAAGG	RT-PCRs on *SEPALLATA3*
RT SEP3_R	CTGCTCCCATTCCATCTTGT	RT-PCRs on *SEPALLATA3*
AGL24 F	TTCAGTTCTTTGCGATGCTG	RT-PCR on *AGL24*
AGL24 R	TCAAGGGGAGTTCCACTGTC	RT-PCR on *AGL24*
FUL F	CTCTGTTCTCTGCGATGCTG	RT-PCR on *FRUITFUL*
FUL R	CCCAACTCTCTCCACAAAGC	RT-PCR on *FRUITFUL*
SEP1 F	TCAACAACAAACCTGCCAAA	RT-PCR on *SEPALLATA1*
SEP1 R	ATGTAACCGTTTCCCTGCTG	RT-PCR on *SEPALLATA1*
SEP2 F	TGGCTCCATTGAAGTCAACA	RT-PCR on *SEPALLATA2*
SEP2 R	CTGAGCACACTGGATGGCTA	RT-PCR on *SEPALLATA2*
SEP4 F	TTTCTCTAACCGTGGCAAGC	RT-PCR on *SEPALLATA4*
SEP4 R	TTCCTGAATTGGAGGGTTTG	RT-PCR on *SEPALLATA4*
SHP1 F	GTTGCCCTCGTCATCTTCTC	RT-PCR on *SHATTERPROF1*
SHP1 R	CCGGATTCGTAAACTGTCGT	RT-PCR on *SHATTERPROF1*
SHP2 F	CCATCACCGAAGCTAATACTCA	RT-PCR on *SHATTERPROF2*
SHP2 R	CACACTCGATTCTTGTTGCTG	RT-PCR on *SHATTERPROF2*
AGL6 F`	TGGTGTCAGGAGGTGACAAA	RT-PCR on *AGL6*
AGL6 R	CCCAACCTTGGACGAAATTA	RT-PCR on *AGL6*
FLCRtm-PCRFor	CGGTCTCATCGAGAAAGCTC	RT-PCR on *FLC*
FLCRtm-PCRRev	CCACAAGCTTGCTATCCACA	RT-PCR on *FLC*
RT PCR-At1g54930_F	CTCCGCCGGACATTCACT	RT-PCR on *At1g54930*
RT PCR-At1g54930_R	CTCACTGTTGGAGACCGATCC	RT-PCR on *At1g54930*
RT PCR-At1g54940_F	CTCCGGAATCATGGTAATGC	RT-PCR on *At1g54940*
RT PCR-At1g54940_F	ATGGCTCGAGAACCTGATGT	RT-PCR on *At1g54940*
AG 3´gene F	TGTTGGGACCAACTTGTGTG	Semi-quantitative ChIP PCR on *At4g18970*
AG 3´gene R	CCCTTGGCACACTAGCAAAT	Semi-quantitative ChIP PCR on *At4g18970*
AG 4LB	ATTCGACACGCAATTTCCA	Semi-quantitative ChIP PCR on

		AGAMOUS
AG 4RB	ATCTTGCGCTCAATTCCAAC	Semi-quantitative ChIP PCR on *AGAMOUS*
ChIP SEP3 F	CTTTTGATTCTGGGGGTCGT	Semi-quantitative ChIP PCR on *SEPALLATA3*
ChIP SEP3 R	GATGAATCCCATCCCCAAGT	Semi-quantitative ChIP PCR on *SEPALLATA3*
actin F	CGTTTCGCTTTCCTTAGTGTTAGCT	Semi-quantitative ChIP PCR on *ACTIN* (reference)
actin R	AGCGAACGGATCTAGAGACTCACCTTG	Semi-quantitative ChIP PCR on *ACTIN* (reference)

2.4.4. Digestion of DNA with restriction enzymes

Restriction digestions were carried out according to manufacturers' guidelines (Promega/NEB) at suggested temperatures using the restriction enzyme, appropriate buffer, Bovine Serum Albumin (BSA) where required and water. Digests were incubated from one to three hours (five hours for Southern blot) in the smallest volume possible (commonly 10μL). In double digests, the two enzymes were added at the same time if requiring the same buffer. If not, the DNA was first cut with the enzyme requiring a lower salt concentration then the reaction's salt concentration was adjusted for optimal cutting of the second enzyme.

2.4.5. dCAPs genotyping for *clf-81*

In the case of *clf-81*, it was not possible to genotype by a simple PCR as it is not a T-DNA insertion or a deletion line, but a single nucleotide substitution that creates a restriction site for the endonuclease restriction enzyme *Hha1*. The PCR product obtained using the two specific primers AB32 and AB33 (230bp) is digested by *Hha1* (for conditions see above) and ran on an electrophoretic gel of 3% agarose (for

conditions see below). The bands visible on the gel correspond to the wild type allele of CLF (uncut: 230bp) and the *clf-81* allele (cut: 210bp).

2.4.6. Agarose gel electropheresis of DNA

DNA fragments were separated using electrophoresis. 1% (0.7% for Southern-blot) agarose gels were made by dissolving agarose powder in 1X TBE buffer (0.89 M Tris base, 0.89 M Boric acid and 0.02 M EDTA). Ethidium bromide (10 mg/ml stock) was added (5 μL per 100 mL of gel) then, the gel was poured into a casting tray and run in the appropriate tanks. DNA samples were loaded after addition of 1/10th volume of loading buffer (40% w/v sucrose, 0.25% bromophenol blue). 0.5 μg of 1 Kbp ladder (NEB, manufacturer's guidelines). Gels were run in electrophoresis tanks at room temperature at 5 volts/cm (distance between the two electrodes) for diagnostic tests or 2 volts/cm when run overnight for a Southern-blot separation (see section 2.4.7). After electrophoresis, separated DNA fragments were viewed on a transilluminator under UV light.

2.4.7. Analysis of genomic DNA by Southern blotting

Genomic DNA was first digested with restriction endonucleases. 15.5 μL of DNA (approximately 2 μg), 2.5 μL of spermidine (1M), 2.5 μL of endonuclease buffer (10X), 2 μL of endonuclease (10-20 U/μL) and H_2O to a final volume of 25 μL were mixed in a microcentrifuge tube and incubated at 37°C for 3 hours. The samples were then loaded on to a 0.7% agarose gel and run at 30-35 V overnight to separate the fragments according to size. The digested DNA was vizualized in a transilluminator and the gel was trimmed to a convenient size. Then it was immersed in an HCl solution (0.25 M) for 15 minutes, to depurinate the DNA, which aids in the subsequent transfer of large DNA molecules. The gel was then immersed in a denaturation solution (1.5 M NaCl, 0.5 M NaOH) for 15 minutes to denature DNA prior to transfer, and in a neutralisation solution (1.5 M NaCl, 0.5 M Tris.HCl

pH 7.2, 1 mM EDTA) for 15 minutes. DNA was transferred overnight by capillarity onto a membrane (Hybond-N) by Southern blotting. The following day the membrane was washed for 10 minutes in SSC 2X and DNA was cross-linked to the membrane by baking the membrane for 2 hours at 80°C. To make the probe, the DIG DNA labeling kit® from ROCHE was used. The probe was made by labelling the DNA products of PCR reactions incorporating dNTPs and Digoxigenin-dUTP (Dig). 50 µL reactions were made in the same way as PCR reactions using 1.5 µL of Dig labelling mix. Incorporation can be seen by slower migration of a small aliquot of Dig labelled product next to a standard reaction as large Dig-dUTP slows migration. The next step is to hybridise the probe to the membrane. To limit unspecific hybridisation of the DNA probe to the membrane, a blocking step at 65°C with Hybe buffer (SSC (5X), N-lauryl sarcosine (0.1%), SDS (0.02%) and milk powder (1% w/v)) for two hours with gentle shaking is necessary. The probe was then denaturated by boiling for three minutes and hybridised to the membrane by incubation with Hybe buffer (10 mL/10 cm^2) at 65°C overnight. The next day the membrane was washed twice for 5 minutes at room temperature with Wash 1 (SSC (2X) and SDS (0,1%)), and twice in stringent Wash 2 (SSC (0.5%) and SDS (0.1%)) at 68°C. The membrane was then blocked in 50 mL of maleic acid blocker (blocking reagent from ROCHE (10% w/v), maleic acid (100 mM pH7) and NaCl (150 mM)) for 1 hour at room temperature to prevent unspecific binding of antibodies to the membrane. The membrane is then incubated for 30 minutes in 20 mL maleic acid blocker at room temperature with 2 µL anti-DIG antibody conjugated to Alkaline Phosphatase (anti-DIG-AP). The membrane was washed twice for 15 minutes at room temperature in 100 mL maleic acid 1X buffer (maleic acid (100 mM pH7.5), NaCl (150 mM) and Tween20 (0.3%) and equilibrated in detection buffer (Tris-HCl (100 mM pH9.5) and NaCl (100 mM)) for 3 minutes at room temperature, wet with the chemiluminescent detection substrate, CSPD (250 µM in detection buffer), sealed in a bag and exposed to a Kodak X-OMAT AR film (Sigma-Aldrich) for 2 hours. The film was then developed using a Konica developing machine (Konica).

2.4.8. Recovery of plant genomic sequences by plasmid rescue

The T-DNA in pSKI074 contains pUC19 sequences with a bacterial origin of replication and an ampicillin resistance gene for plasmid rescue (pBstKS+). These sequences are flanked by several restriction enzyme sites that can be used for rescue of T-DNA and adjacent plant sequences from transformed plants. The plasmid rescue technique needs four steps: (i) digestion to produce fragments containing genomic plant DNA, the Right Border (RB) of the T-DNA and the pBstKS+ sequences; (ii) precipitation of the DNA fragments and ligation in a large volume to allow the fragments to self-ligate; (iii) precipitation in a small volume and finally (iv) transformation of the plasmids into *E. coli* by electroporation.

Digestion: The DNA was digested as follow: 15µl DNA (2-3 µg), 5µl Buffer H 10X (Promega Madison, USA), 0.5µl spermidine (100mM), 4.5µl *Eco*RI 10 U/µL (Promega Madison, USA) and up to 50µL of H_2O. The digestion was incubated at 37°C overnight.

Precipitation: The restriction mixture was topped up to 200µl with ddH_2O. 200µl of phenol/chloroform was then added and then centrifuged at 14000rpm in a centrifuge for five minutes. Approximately 190µl of the upper (soluble) phase was then transferred to a clean eppendorf and the lower phase discarded. 20µl of 3M NaAc pH5.2 and 175µl of isopropanol were then added and the tubes were then mixed by inverting and stored for 20 minutes at -20°C, before being centrifuged for 5 minutes at 14000rpm. The supernatant was then discarded and the pellet was then rinsed by addition of 500µL of cold 70% ethanol and centrifugation at 14000rpm for 10 minutes. The supernatant was removed with a pipette and the pellets left to air dry. 179µl of ddH_2O was then added and the pellets resuspended by incubation at 37°C for 2 minutes. 1µl of T4 ligase (NEB) and 20 µl of T4 ligase buffer 10X (NEB) were then added to the resuspended DNA and left to ligate overnight at 16°C.

Precipitation: The ligation was then precipitated as above, but then resuspended in 10 µl of ddH_2O.

Transformation: The resuspended DNA was then used to transform ElectroTen-Blue® Electroporation competent DH5α cells as detailed in section 2.2.2.

2.4.9. The genome walker strategy

The genome walker is a PCR-based method where after digestion by specific (blunt) enzymes an adaptator is ligated to the DNA and allows walking in uncloned genomic DNA (Siebert et al., 1995). Genome Walker libraries were constructed by using Universal Genome Walker kit according to manufacturer's instruction (BD Biosciences Clontech). *A. thaliana sop3* genomic DNA (0.25–0.5 μg) in each reaction was digested at 37°C overnight with a restriction enzyme. Four enzymes (*Eco*RI, *Eco*RV, *Hin*DIII and *Vsp*I) were used in four reactions, respectively. After purification by phenol/chloroform and precipitation by ethanol, the digested DNA was ligated to Genome Walker adapters AP1 and AP2 at 16°C overnight. Primers for PCR-based DNA walking in Genome Walker Libraries were T-DNA-specific genewalker-LB1 and genewalker-LB2. Two Genome Walker PCR reactions were carried out successively using a Phusion Taq polymerase (NEB), which has high fidelity and efficiency in the amplification of DNA template. genewalker-LB1 and AP1 were used in primary PCR, and genewalker-LB2 and AP2 were used in secondary (nested) PCR. In a 50 μL PCR mix, 1 μL of each Genome Walker DNA library was used as templates in the primary PCR, and 2 μL of primary PCR products was used as templates in secondary PCR. The PCR conditions used are the same specified in section 2.4.2. The PCR products obtained were cloned as specified in section 2.2.3.

2.4.10. Plasmid preparations from *E. coli*

1.5 mL of bacterial culture was poured into a microcentrifuge tube and centrifuged at 12000 x g for 1 minute. The supernatant was removed and the bacterial pellet resuspended by vigorous vortexing in 150 μL of resuspension buffer (Tris.HCl (5M

pH8), EDTA (10 mM pH8) and RNAse (100 μg/mL)). 150 μL of lysis buffer (NaOH (0.2N) and SDS (0.1%)) was added and mixed by gentle inversion, followed by 200 μL of ice cold neutralisation buffer (potassium acetate (3M) and glacial acetic acid (11.5%)). The tube was gently inverted to mix contents, centrifugated at 12000 x g for 10 minutes and the supernatant removed to a fresh tube. Contaminants were removed by a phenol/cloroform treatment. The DNA was then precipitated by addition of 2.5 volumes of ethanol and 0.3 volumes of sodium acetate (3M). The precipitated DNA was then pelleted by centrifugation at 18000 x g for 15 minutes. The ethanol was removed and the pellet air dried and dissolved in 50 μL TE (Tris (10 mM pH8) and EDTA (1 mM)).

2.4.11. Sequencing of DNA

The QIAEX II Extraction kit (Qiagen) was used to clean plasmid DNA from *E. coli* in order to remove contaminants. This was carried out according to the manufacturer's guidelines by treating the sample, made up to 100 μL with H_2O, as a 100 mg gel slice. DNA was then quantified on an agarose gel and 300-400 ng were used for a sequencing reaction. For a sequencing reaction the following was added to a PCR tube: DNA (300-400 ng), sequencing primer (0.8 pM) and up to 6 μL with ddH_2O. The sequencing services of the School of Biological Science (SBS) at the King's Buildings campus carried out the dideoxy sequencing reactions. Samples were analyzed using ABI Prism 3100 (Abgene, Epsom, UK). Sequences were analyzed using 4Peaks software (4Peaks by A. Griekspoor and Tom Groothuis, mekentosj.com).

2.5.RNA techniques

2.5.1. RNA extraction

Plant tissue was harvested and snap frozen in liquid nitrogen. 100 mg of seedlings were placed in a microcentrifuge tube with 1 mL of TRIzol® reagent (Invitrogen). The tissue was macerated using a homogeniser. The samples were centrifuged for 2 minutes at 18000 x g and the supernatant removed to a fresh tube. 200 μL of chloroform was added, the samples vortexed for 15 seconds and then spun at 18000 x g for 15 minutes at 4°C. The aqueous phase was removed to a fresh tube, mixed with 500 μL of isopropanol and finally RNA precipitated by centrifugation for 10 minutes at 18000 x g. The supernatant was removed and the pellet washed with 70% ethanol. The RNA pellet was dried and dissolved in water that had been treated with DEPC to inactivate RNAses. Finally, the samples were quantified using a nanodrop spectrophotometer (Thermo scientific). The quality of the RNAs was also controlled in a 2% agarose gel.

2.5.2. Semi-quantitative RT-PCR

Total RNA was treated with RQ1 RNase-free DNase (Promega) as follow: 1 μL of RQ1 buffer (10X), 1 μL of RQ1 RNase-free DNase, 1 μg of total RNA and up to 10 μL of ddH_2O. The reaction was incubated for 30 minutes at 37°C. Subsequently, the RNA was precipitated using a phenol/chloroform technique. The samples were resuspended in 10 μL of ddH_2O and quantified again using a nanodrop spectrophotometer (Thermo scientific). Next, 500 ng of DNase treated RNA were reverse transcribed using the Im-Prom-II Reverse Transcription kit (Promega) following the manufacturer's instructions. Oligo-dT_{15} was used to prime the reaction. 1 μL of the cDNA product from the reverse transcription reaction was used in 20 μL PCR reaction.

The RT-PCR settings were carried out on several dilutions of the cDNA. The annealing temperature, the number of cycles and the concentration of the reference gene primers, *TUBULIN* (*TUB*), were optimized for each couple of primers. At the optimal temperature, in the linear phase of the PCR amplification and with equivalent amplification efficiencies for the two couples of primers, the conditions obtained allow us to quantify the amplification product of the target gene and to compare that with the reference (*TUB*). Each reaction was then repeated three times and the PCR products (target gene and reference (*TUB*)) were migrated on an agarose gel and the intensity of the bands obtained was quantified using free software: ImageJ 1.38x (Wayne Rasband, National Institutes of Health, USA http://rsb.info.nih.gov/ij/). The expression of the target gene compared to the expression of *TUB* in each sample was then calculated as follow: (average target gene signal)/(average ß–tubulin signal). Standard error were calculated as follow: $(\sum(\text{intensity } PCR_x - \text{average intensity})^2)/(3\sqrt{2})$ where PCR_x is the indivivual intensity for each of the 3 reactions performed.

2.6. Protein techniques

2.6.1. Protein extraction

Plant inflorescence or seedlings were collected and frozen in liquid nitrogen ($N_2(l)$). A mortar and pestle was pre-chilled using $N_2(l)$, then used to grind tissue to a fine powder in 200 µL of extraction buffer (Tris (0.05M pH7.5), NaCl (0.15M), dithiothreitol (DTT) (0.1M filter-sterile pH8), one tablet of protease inhibitor cocktail (ROCHE) in 10 mL of ddH_2O). For total protein extraction, 1/5 volume of 6X protein loading buffer (Tris.HCl (0.35M pH6.8), SDS (10.28% w/v), glycerol (36% w/v) DTT (0.6M) and bromophenol blue (0.012% w/v)) was added and samples stored at -80°C until required.

2.6.2. Western-blotting

Protein gels were prepared and run as follows using the Biorad mini-protean 3 system. To make the gel, "resolving" gel (acryl/bis-acrylamide 37:5:1 (10% w/v), Tris (0.375M pH8.8), SDS (0.1% w/v), APS (0.1% w/v) and Temed (0.01% w/v)) was overlaid with 1/5 depth of "stacking" gel (bis/acrylamide (6% w/v), Tris (0.125M pH6.8, SDS (0.1% w/v), APS (0.1% w/v) and TEMED (0.01% w/v)). 15 μL of protein samples were denaturated by boiling for five minutes, then were centrifuged for five minutes at 14,000 rpm prior to loading on to the gel. A protein size marker ladder was also loaded. The gel was run for one hour at a fixed voltage of 120V at room temperature in running buffer (TRIS (0.6%) glycine (2.28%), SDS (1% w/v) at pH8.3). Separated proteins were then transferred to Hybond Nitrocellulose membrane in a Biorad mini-protean 3 transfer cassette and tank; buffer containing glycine (1.44%), Tris (0.3%) and methanol (20% v/v), was used to soak the membrane and the gel prior to transfer. Transfer was carried out at 4°C in transfer buffer (glycine (1.44%), TRIS (0.3%) and methanol (20% v/v), at 65V for two hours. Protein transfer and the ladder position was monitored by rinsing the membrane briefly in Ponceau stain (Ponceau (0.1% w/v) and acetic acid (5% w/v)). The membrane was blocked by rinsing for two hours at room temperature in PBST milk (milk powder (5%), phosphate buffered saline solution with Tween20 (PBST) (NaCl (0.8%), KCl (0.02%), Na_2HPO_4 (0.144%), KH_2PO_4 (0.024%) and Tween20 (0.2%) at pH 7.2). The membrane was incubated in 1/1000 primary (1°) antibody, (anti-AG C-terminus rabbit serum) in PBST milk, overnight at 4°C on a rolling incubator. The membrane was rinsed for one hour with several changes of PBST, and then incubated in a 1/4000 dilution of secondary (2°) antibody in PBST milk, (anti rabbit IGG horseradish peroxidase (Amersham Biosciences)) at room temperature for one hour on a rolling incubator. The membrane was again rinsed in PBST with four changes of PBST over an hour, and then the position of antibody attachment detected using an Enhanced Chemiluminescence (ECL) reaction. The membrane was incubated at room temperature in equal volumes of ECL detection reagents 1

and 2 (Amersham Biosciences) for two minutes, then exposed to photographic film (Kodak X-OMAT AR, Sigma-Aldrich) in the dark for five minutes, thirty minutes and one hour (to completion of light emanation). The film was developed as for a Southern blot in section 2.4.7.

2.7.ChIP techniques

2.7.1. Chromatin Immunoprecipitation

Approximately 2,000 *A. thaliana* seeds from wild type and *clf-50* were sterilized and grown 8 to 10 days in MS agar media on Petri plates. Crosslinking of the chromatin was achieved by vacuum infiltrating a solution of 1% formaldehyde at room temperature for ten min. The cross-linking reaction was stopped by adding glycine to a final concentration of 0.1M and incubating for five min. The seedlings were rinsed three times with water and then ground to a fine powder in liquid nitrogen. Chromatin was isolated, and resuspended in Nuclei lysis buffer (Tris-HCl (50 mM pH8), EDTA (10 mM), SDS (1%)) and sonicated in a Ultrasonic Homogenizer 3000 (Biologics, Inc) to achieve an average fragment size of approximately 0.8-1.5 kb. Any remaining cellular debris were removed by centrifugation. The chromatin solution was then diluted 10-fold with ChIP dilution buffer (Triton X-100 (1.1%), EDTA (1.2 mM), Tris-HCl (16.7 mM pH8), NaCl (167 mM)) and precleared by incubating 40µl of protein-A agarose beads (50% suspension in ChIP dilution buffer) (Roche biochemicals) for 1 hour at 4°C. Histone-DNA complexes were immunoprecipitated with H3K27me3 antibodies overnight at 4°C. An equal amount of chromatin solution was not treated with antibody and, thus, served as the mock antibody control. The modified histone-DNA complexes were then extracted by incubating 30 µl of 50% protein A-agarose beads with the chromatin-antibody solution for one hour at 4°C. After several washes, the immunocomplexes were eluted twice from the beads using 250ul of elution buffer (SDS (1%), NaHCO3 (0.1M)). The samples were reverse-crosslinked by incubating for 6-8 hours at 65°C and then treated with

proteinase-K to remove all proteins. DNA was purified by phenol-chloroform extraction and recovered by ethanol precipitation in the presence of 1 μg of glycogen. Precipitated DNA was resuspended in 100 μl of TE buffer and used for PCR analysis. A small aliquot of untreated sonicated chromatin was reverse cross-linked for use as the total input DNA control. Immunoprecipitated DNA was analyzed by PCR described in the following section. Reactions were performed in 25 μl with 0.4 μl of immuno-precipitated DNA. The amplified DNA was visualized on 1.5% agarose gel stained with ethidium bromide.

2.7.2. Semi-quantitative ChIP PCR

In all ChIP experiments, I amplified DNA using two primer pairs in duplex PCR reactions: one pair was specific for the gene of interest (*SEP3*), whereas the other pair was specific for a gene that is not expected to be enriched (reference). As different primer pairs seldom amplify with equal efficiency in duplex PCR, I compared the relative amounts of the two products amplified from IP DNA against the relative amounts amplified for input (pre-IP) DNA. The ratio reflects the enrichment for a sequence following IP and corrects for any differences in the efficiency of the two primers or of the amounts of chromatin isolated. Fragment intensity was measured with the program ImageJ 1.38x (Wayne Rasband, National Institutes of Health, USA http://rsb.info.nih.gov/ij/). The relative enrichment was determined as follows: (intensity of band of fragment of interest (in IP)/intensity of control band (in IP; actin))/(intensity of band of fragment of interest (in input)/intensity of control band (in input; actin)). Several dilutions of input were used in amplification to reveal whether amplification was in a linear range. If PCR products were present in the 'beads' control, intensities of amplifications were subtracted from the sample (IP with specific antibodies) values (Schubert et al., 2006).

3. *SUPPRESSOR OF POLYCOMB 1* (*SOP1*) is a loss of function mutation in the MADS-box gene *SEPALLATA3* (*SEP3*)

suppressor of polycomb 1 (*sop1*) mutants were first isolated by Dr. S. Yang from screens where plants were grown on tissue culture MS plates *in-vitro* and then transferred onto soil. The suppressed plants did not show the characteristic curled leaves of *cfl50/- pCLF:CLF-GR* progenitor plants. Seeds from Sop1- and Sop1+ plants, from the original screened family were available in the laboratory for further analysis.

In this chapter, the *SOP1* gene was identified and the phenotype of Sop1- plants was analyzed. Finally, the relationship between the mutation causing the Sop1- phenotype and the *clf* mutation was dissected.

3.1. Identification of the gene disrupted by the *sop1* mutation

Seeds from Sop1+, siblings of Sop1- plants were grown on soil under normal conditions and a family segregating for the Sop1- phenotype was obtained. The DNA from Sop1+ and Sop1- plants was extracted. Southern-blot analysis was performed to determine the number of T-DNA insertions in these plants and to establish if the suppressed phenotype was linked with a distinct T-DNA insertion. The mutated locus was therefore identified and further characterization via plasmid rescue and sequencing respectively. Finally, to support the experiments previously described, the suppression of the Clf- phenotype was verified by crossing an independent allele of the identified mutation with the *clf-81* allele in the Columbia background.

3.1.1. Identification by Southern-blot analysis of the T-DNA insertion responsible for the Sop1- suppressed phenotype

Southern-blot analysis was performed using a probe detecting CaMV35S enhancer sequences on plants showing the two different phenotypes: Sop1- (suppressed) and Sop1+ (Clf- phenotype) as shown in Figure 3.1-A. The Southern-blot was carried out on a limited numbers of plants (n=6), as with the segregating material it would be straightforward to determine the insertion responsible for the Sop1- phenotype. Sop1- plants [-] have in Southern-blot either one band at 11Kbp (lane 1) or two bands at 11 and 7Kbp (lanes 2, 3 and 5), see red arrows in Figure 3.1-B. In lane 1 only one band is visible, this could be because only one band can be detected as there is only one T-DNA insertion or because we reach the detection limits of the Southern-blot. Nonetheless, all the Sop1- plants show at least the band at 11Kbp. Sop1+ plants [+] showed no bands (lanes 4 and 6).

To confirm the results of the Southern-blot, seeds from the six plants analyzed were harvested and grown on soil. The phenotypes observed in the next generation plants were: 100% suppressed (Sop1-) for progeny of plants from lanes 1, 2, 3 and 5 and 100% unsuppressed (Sop1+) for progeny of plants from lanes 4 and 6.

The Southern-blot analysis allowed to identify of a T-DNA insertion (band at 11Kbp), linked to the suppressed phenotype observed as it co-segregates with Sop1- plants.

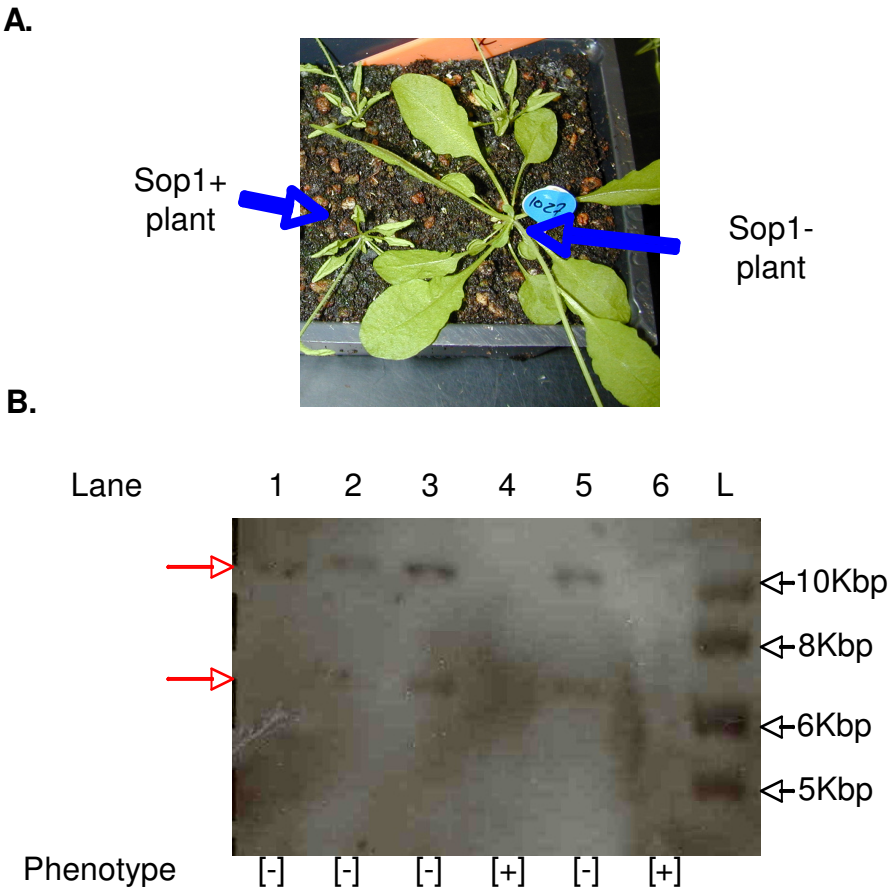

Figure 3.1: Analysis of plants segregating for the *sop1* mutation. A. Phenotypes obtained in a segregating family: Sop1- and Sop1+ plants. B. Southern-blot with *Eco*RI digested DNA from 6 independent plants (lanes 1 to 6). The phenotype of each plant is shown under the lane in which its DNA was loaded ([-]: suppressed phenotype; [+]: unsuppressed phenotype). The red arrows indicate the 2 different bands detected at 11Kbp and 7Kbp. The probe used detects CaMV35S enhancer sequences of the pSKI074 vector used for the mutagenesis. L corresponds to the 1Kbp ladder used to assess the size of the bands obtained (sizes showed on the right hand side of the figure).

3.1.2. The mutation responsible for the suppressed Sop1- phenotype is recessive

To better characterize the *sop1* mutation, seeds segregating for the Sop1- phenotype were sown on MS plates and the phenotype was scored. Among 73 germinated plants, 58 showed the Clf- phenotype and 15 the Sop1- suppressed phenotype. The ratio observed between Clf- and Sop1- plants (3:1) fits with the Mendelian segregation ratio of a single recessive mutation. Although in Figure 3.1 for suppressed [S] plants in lanes 2, 3 and 5 two T-DNA insertions were detected. One explanation is that the two T-DNA insertions detected in the Southern-blot are at different loci. Another explanation could be that the T-DNA insertion is complex and the RB sequences were duplicated at some point in the mutagenesis/T-DNA insertion. However, to definitively genetically characterize the *sop1* mutantion, a genetic cross was made between a *sop1* homozygous plant and the progenitor *clf50 pCLF:CLF-GR* line. The F1 generation was Sop1+ in phenotype and in the F2 generation ¼ of the plants showed a Sop1- phenotype. This result confirms that the mutation responsible for the Sop1- phenotype is recessive in the *clf* background.

3.1.3. Fine mapping of the *SOP1* locus disrupted by the T-DNA insertion

Mapping the mutation in detail was necessary to discriminate between the two hypotheses formulated previously about the nature of the T-DNA insertion. The suppressed plant analyzed in lane 1 of the Southern-blot in Figure 3.1-B had only one T-DNA insertion (band at 11Kbp). Its DNA was used to plasmid rescue the Right Border (RB) of the T-DNA and adjacent plant sequences as previously explained in chapter 2 section 2.4.8 (Weigel et al., 2000). The plant sequence found by plasmid rescue was the 3' region of the *At1g24260* locus (*SEPALLATA3* (*SEP3*). Taking advantage of the complete genome sequence from The Arabidopsis Information Resource (TAIR), primers to sequence the plasmid previously obtained

were designed. Their location in the genome is specified in Figure 3.2. The mutation responsible for the sop1 phenotype is a complex T-DNA insertion in which the DNA sequence of two adjacent loci: *SEP3* and *At1g24265* is modified (Figure 3.2-B). First, a 4.3 Kb deletion of the 5′ sequence and the first five exons of *SEP3* was identified (Figure 3.2-A). Second, the *At1g24265* locus is interrupted by the insertion of the T-DNA between the 6th and the 7th exon (Figure 3.2-B). The left border (LB) of the T-DNA was not mapped. From here in this study the deletion mutation at the *SEP3* locus will be named: *sep3-7*.

The *SEPALLATA3* (*SEP3*) gene, *SEP3* is a member of the MADS-box transcription factor family. *SEP3* is redundant with *SEP1* and *SEP2*. Flowers of *sep1 sep2 sep3* triple mutants show a conversion of petals and stamens to sepal ((Pelaz et al., 2000). The *At1g24265* locus encodes for an unknown protein similar to bZIP transcription factor, bZIP56, in *Glycine max*. According to The Arabidopsis Information Resource (TAIR), there is no mutant characterized in *A. thaliana* for this locus.

sop1 mutant plants therefore carry three mutations: *clf50*, *sep3-7* and *At1g24265*. The suppressed phenotype can be explained by either of the two double mutants: *clf50 sep3-7* and *clf50 At1g24265* or the triple mutant: *clf50 sep3-7 At1g24265*.

It has been shown that the *ag* mutation is able to suppresses the Clf- phenotype (Goodrich et al., 1997). Also, it is known that *AG* interacts with *SEP3* in a Yeast-two-Hybrid assay (Y2H) (Fan et al., 1997). Taking these elements together, it seems that *sep3-7* might be responsible for the suppressed phenotype of *sop1* plants. The relationship between the *sep3-7* mutation and the Sop1- phenotype was therefore analyzed.

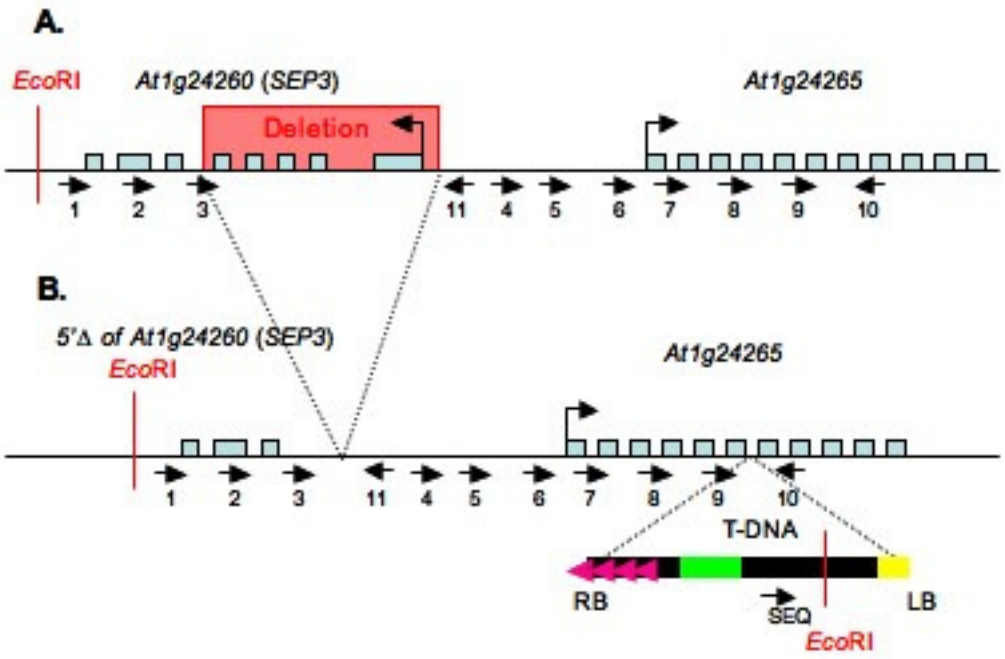

Figure 3.2: Diagram of the T-DNA insertion at *SEP3* found in Sop1- plants. A.: Location of the deletion within the first 5 exons of *SEP3*. Note the opposite transcription direction of *SEP3* and *At1g24265*. B.: genomic rearrangement of the *SEP3* locus and location of the T-DNA insertion in *At1g24265*. The *Eco*RI restriction sites allowing the recovery of *SEP3* genomic sequences are shown in red. Arrows numbered 1 to 11 correspond to the primers used to map the *sop1* mutation see chapter 2 section 2.4.3 for details. Note primers 10 and 11 are in opposite direction.

3.1.4. Cosegregation of the *sep3-7* mutation with the Sop1- phenotype

F2 plants from a cross between a *sop1* mutant and the progenitor line *clf50 pCLF:CLF-GR* were grown on soil under normal conditions. If the deletion of *SEP3* is responsible for the Sop1- phenotype, *sop1* mutants should be homozygous for the *sep3-7* mutation, i. e. *sop1* and *sep3-7* should cosegregate. The phenotype of each F2 plant was scored and DNA extractions were made from ten F2 plants.

As the mutation was mapped, the identification of the *sep3-7* deletion by PCR seemed simple and straightforward using primers numbered 1 (3'AGL9), 2 (3'aAGL9), 3 (3'bAGL9) and 11 (3'jAGL9), see Figure 3.2 for details. The three different primers pairs (1&11; 2&11 and 3&11), covering *sep3-7* would amplify in PCR a short band for the *sep3-7* allele and a bigger band for the wild type allele *SEP3*. Unfortunately, only one combination of primers (3&11) gave products and these were not of the expected sizes: the *sep3-7* mutant band at 0.4 Kbp and wild type band at 4.6 Kbp (Figure 3.3-A). The PCR bands sizes obtained were: for the *sep3-7* allele a band at 1.6 Kbp and for the wild type allele a band at 4 Kbp (Figure 3.3-B). The PCR products were not cloned and sequenced as using homozygous *sop1*, wild type (Ws) and *clf50 pCLF:CLF-GR* DNA controls, discrimination between homozygous *sep3-7*, heterozygous *sep3-7*/*SEP3* mutants and wild type plants was possible (Figure 3.3-B). The difference in length of the PCR products might be explained as the ecotype used here is Ws and small differences might exist with Col0.

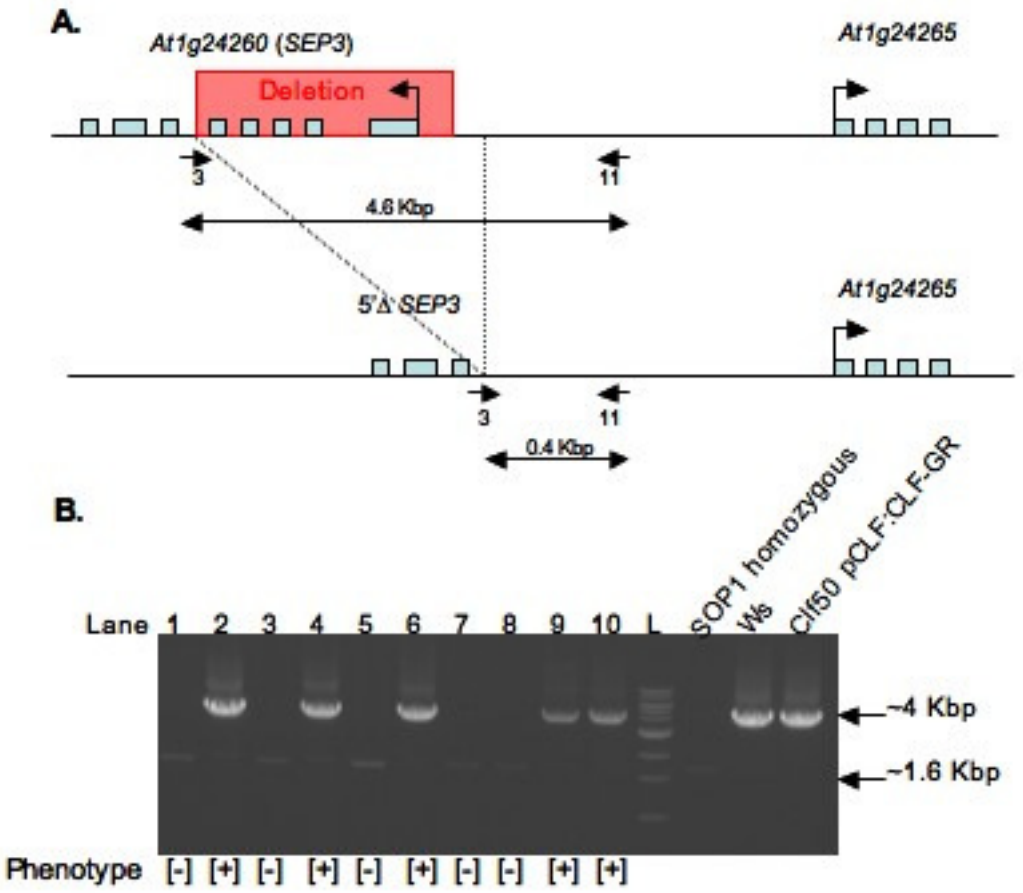

Figure 3.3: Genotyping the *sep3-7* deletion in segregating F2 plants. A. Diagram showing the localisation of primers number 3 (3'bAGL9) and number 11 (3'jAGL9) used for the genotyping PCR. B. Genotyping PCR result for 10 plants (lanes 1 to 10). The band around 4 Kbp correspond to the wild type product (controls on the right hand side: Ws and *clf50 pCLF:CLF-GR*). The lower band around 1.6 Kbp corresponds to the *sep3-7* allele (control on right hand side: *sop1* homozygous). Samples in lanes 1, 3, 5, 7 and 8 corresponds to homozygous *SEP3* plants. Samples in lanes 9 and 10 correspond to homozygous *sep3-7* plants. Samples in lanes 2, 4 and 6 correspond to heterozygous *sep3*-7/*SEP3* plants.

The PCR results obtained in Figure 3.3-B show that all Sop1+ plants were homozygous *SEP3* or heterozygous *sep3*-7/*SEP3*. Whereas all Sop1- plants were homozygous *sep3-7* as predicted if the *sop1* mutation is linked with the *sep3-7* allele. Moreover, to further check the cosegregation, seeds from the Sop1+ plants tested

were harvested and grown on soil to determine whether parents were homozygous *SOP1* or heterozygous *sop1*/*SOP1*. The phenotypes observed in the following generations were: (i) 100% Sop1+ from homozygous *SEP3* plants, (ii) ¾ Sop1+ and ¼ Sop1- from heterozygous *sep3*-7/*SEP3* plants and 100% Sop1- from homozygous *sep3-7* plants. This result confirms cosegregation of the Sop1- phenotype and the *sep3-7* mutation. Nevertheless, these results cannot distinguish whether the Sop1- phenotype is caused by the *sep3-7* mutation or a closely linked mutation. To this end an allelism test with an independent *sep3* mutant will be necessary.

3.1.5. Analysis of the double mutant *sep3-2 clf81*

To conclusively demonstrate that *sep3-7* is responsible for the suppressed phenotype of Sop1- plants, I tested whether an independent *sep3* mutation suppresses a *clf* mutation. This will allow the exclusion of the Sop1- phenotype as a result of a mutation linked to the deletion at *SEP3* (*sep3-7*). *sep3-2* and *clf-81*, two well-characterized mutations in the Columbia background were crossed and the phenotype of the double mutant *sep3-2 clf-81* was assessed. The *sep3-2* allele is an insertion of an *En-1* transposable element 48 bases after the *SEP3* start codon. *sep3-2* mutants have no discernable phenotype due to the redundancy between *SEP3* and the *SEP1*, *SEP2* and *SEP4* genes (Pelaz et al., 2000) see also Figure 3.5-B. The *clf-81* allele was isolated from an EMS mutagenesis screen and carries a point mutation in the SET domain of the *CLF* gene (Schubert et al., 2006) see also Figure 3.5-B. It has a strong leaf curling phenotype (Figure 3.5-B) suggesting it is a severe *clf* allele. Consistent with that, *clf-81* carries a missense mutation in a highly conserved residue (R794H) within the SET domain. The *sep3-2* mutation was detected using duplex PCR. The *clf-81* mutation has to be detected by derived Cleaved Amplified Polymorphic Sequence (dCAPS). In this technique, a restriction enzyme recognition site including the specific Single Nucleotide Polymorphism (SNP) of *clf-81* is introduced into the PCR product by a primer containing one or more mismatches to template DNA. The PCR product modified in this manner is then subjected to

restriction enzyme digestion and the presence or absence of the SNP is determined by the resulting restriction pattern. For genotyping details on these two alleles and primers see chapter 2 sections 2.4.3 to 2.4.5.

Twenty-four F1 plants from the cross between *sep3-2* and *clf81* were obtained and had a wild type phenotype. Seeds from the F1 generation plants were harvested and the progeny of individuals F1 plants was sown on soil under normal conditions. In the F2 generation, the direct visual identification of the double mutants was difficult as they are predicted to be rare (1/16) and to have wild type phenotype. Therefore, homozygous *sep3-2* heterozygous *clf-81* plants were first identified by PCR-based genotyping. The seeds of these plants were harvested and F3 plants segregating ¼ *sep3-7 clf-81* double mutants were genotyped and scored phenotypically. Ten F3 plants were examined to find the double mutant *sep3-2 clf81* (Figure 3.4). Samples from plants analyzed in lanes 3 and 9 (highlighted in red circles in Figure 3.4), are double mutant *sep3-2 clf81* as they only posses the bands corresponding to the 2 mutants alleles *sep3-2* and *clf81*.

Figure 3.4: Genotyping of *sep3-2* and *clf-81/CLF*. In the upper gel the upper band corresponds to the *clf-81* allele and the lower band to the wild type *CLF* allele. In the lower gel the single band detected corresponds to the *sep3-2* mutant allele. In lane 6 Col0 DNA was used as a control. No amplification for *sep3-2* and/or *clf-81* mutations was observed; only the wild type *CLF* allele was amplified. In lanes 4 and 7, only the amplification of the wild type band for CLF was detected. These plants are *sep3-2* homozygous mutant. Samples in lanes 1, 2, 5, 8 and 10 showed two bands (the mutant and the wild type), for the *CLF* locus. These plants are homozygous *sep3-2* and heterozygous *clf-81/CLF*

Having genotyped seedlings, I scored the phenotype of adult F3 plants and found perfect cosegregation of the suppressed phenotype with the double mutant *sep3-2 clf-81*. This result confirms that *sep3* suppresses *clf* (Figure 3.5). In Figure 3.5-B, the double mutant *sep3-2 clf-81* is flanked by the two single mutants *sep3-2* and *clf-81*. First of all, the phenotype of the double mutant *sep3-2 clf-81* is quite different from the two single mutants *clf-81* and *sep3-2*. It loses the characteristic leaf curling of the single *clf-81* mutant and displays a strong serration phenotype compared to the *sep3-2* mutant (Figure 3.5-A for details). The serration phenotype was not observed in *sop1* plants. In the double mutant *sep3-2 clf-81* the leaf size seems to be the same as in the single *sep3-2* mutant. Finally, the double mutant does not flower as early as the *clf-81* single mutant (Figure 3.5-B).

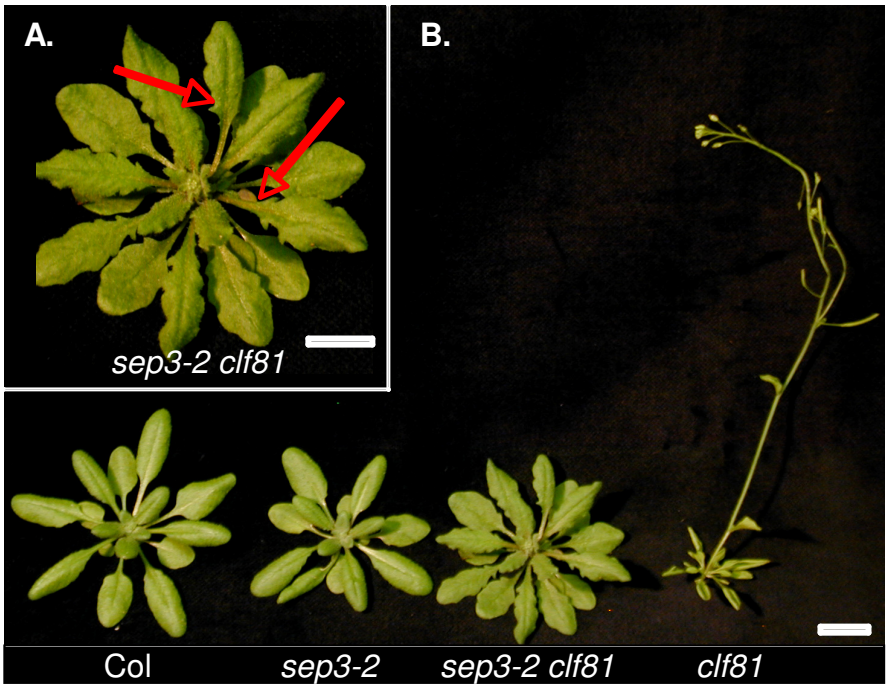

Figure 3.5: Phenotype of the double mutant *sep3-2 clf-81*. A. Detail of the leaf phenotype of the double mutant *sep3-2 clf-81*. B. Comparison between wild type *Arabidopsis* ecotype Col, the single mutants *sep3-2*; *clf-81* and the double mutant *sep3-2 clf-81*. The plants were growing during 4 weeks in LD conditions. Scale bar = 1cm.

To conclude, the *sep3-2* mutation suppresses the Clf- phenotype induced by the *clf-81* mutation. Also, the *sep3-2* mutation induces a serration phenotype that was not observed on Sop1- plants in the Ws background.

3.2. Characterization of the double mutant *clf-50 sop1*

The main trait allowing the isolation of Sop1- plants was the suppression of the curled leaf phenotype of the *clf-50 pCLF:CLF-GR* progenitor line. The identification of the deletion in *SEP3* (*sep3-7*) and the following analysis of the double mutant *sep3-2 clf-81* showed that the suppression by *sep3-2* of the *clf-81* mutant phenotype has subtle characteristics, for example leaf serration was observed. In this section,

the Sop1- phenotype was characterized in more detail, at the vegetative and flower stages and the flowering-time phenotype of Sop1- plants was analyzed. Molecularly, the expression of *AGAMOUS* (*AG*), which is known to be up-regulated in *clf* mutants was tested. Finally, the expression levels of *SHATTERPROF* (*SHP2*), a target of *AG* were measured in Sop1- plants.

3.2.1. Phenotypic characterization

During vegetative development, an increase in leaf serration depth is the only phenotype that distinguishes *sop1* and *sep3-2 clf-81* mutants from *CLF* plants. The double mutants *sep3-2 clf-81* in the Columbia background show a strong leaf serration compared to the single mutants *sep3-2* and *clf-81* and also compared to the wild type plants Col0 (Figure 3.5). In the Wassilewskija (Ws) ecotype, leaves from Sop1- plants appear flat and seem to have approximately the same size and shape than wild type leaves (Figure 3.6-A). A closer examination on same aged leaves of Ws and Sop1- plants confirms that there is a light enhancement of the leaf serration in Sop1- plants compare to wild type (Figure 3.6-B and C). Two serration tips were observed for Ws plants versus 3 for Sop1- plants as shown in Figure 3.6-B and C. However, the depth of the serration sinus in Sop1- plants is bigger than in wild type plants (Figure 3.6-C).

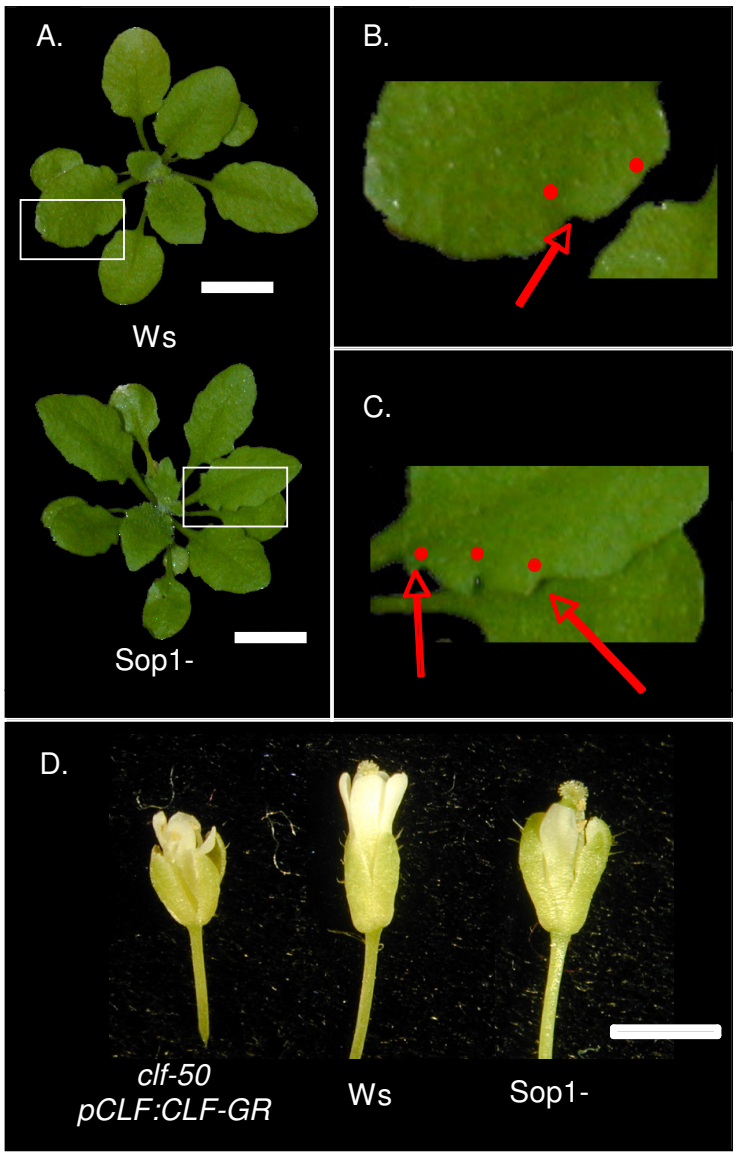

Figure 3.6: Phenotype of Ws and Sop1- leaves and flowers. A: Wild type Ws and Sop1- plants. Scale bar = 1cm. B: Close-up of inset of Ws in (A.) showing the serration detail of a Ws leaf. C: Close-up inset of Sop1- in (A.) showing the serration detail of a Sop1- leaf. Red arrows show the serration sinus and the red dots show the number of sinuses. D: Flower phenotype of Sop1- plants compared with *clf-50 pCLF:CLF-GR* and Ws. Scale bar = 1mm. Flowers were photographed after 21 and 28 days of growth on soil under LD conditions.

Flowers from Sop1- and the double mutant *sep3-2 clf-81* were dissected to identify morphological abnormality. Flowers from *clf-50 pCLF:CLF-GR*, *clf-81*, *sep3-2* single

mutants and Ws, Col0 wild type plants were observed as a control. *sep3-2* flowers did not show a flower phenotype, coherent with the redundancy of *SEP* genes in *A. thaliana* (Pelaz et al., 2000). Flowers of *clf-50 pCLF:CLF-GR* and *clf-81* mutants showed the Clf- flower phenotype previously described (Goodrich et al., 1997) in which flowers open early because the sepals fail to seal the floral bud. The phenotype was stronger for flowers of *clf-81* but no homeotic transformations (sepal/carpel; petal/stamen) for both alleles were observed in our growth conditions. In Figure 3.6-D flowers from Sop1- plants (*sep3-7 clf-50* double mutant) show the same morphological abnormalities than the progenitor line *clf-50 pCLF:CLF-GR*: default to seal the floral bud. A slight irregularity observed was that flowers of *sop1* mutant seem smaller than the progenitor line's flowers (Figure 3.6-D).

The absence of an enhanced floral phenotype in Sop1- plants confirm the redundancy of *SEP3* with *SEP1* and *SEP2* (Pelaz et al., 2000). Some SEP (SEP1, SEP2 and SEP4) activity remains in Sop1- plants enough to compensate the lack of the SEP3 protein, Sop1- plants then exhibit the Clf- flower phenotype.

3.2.2. The *sep3-7* mutation suppresses the early flowering phenotype of Clf- plants

It is known that Clf- plants flower earlier than wild type plants and we showed that *sep3-2* suppresses the early flowering phenotype of the *clf-81* mutant as shown in Figure 3.5. To know if *sep3-7* suppresses, partially or totally, the early flowering phenotype of the *clf-50* mutation; a time to flowering experiment was conducted under LD and SD conditions (Figure 3.7). Under LD and SD conditions wild type control plants (Ws) flowered after 19.8 ± 0.53 and 41.6 ± 0.78 rosette leaves respectively showing that the two treatments (LD & SD) were effective. The progenitor line, *clf-50 pCLF:CLF-GR*, flowered after 8.3 ± 0.22 rosette leaves under LD and 20.1 ± 0.37 rosette leaves under SD conditions as expected for an early

flowering genotype. Sop1- plants flowered at the same time as Ws plants: under LD conditions after 21.1 ± 0.54 rosette leaves and in SD conditions after 44.2 ±0.93 rosette leaves (Figure 3.7).

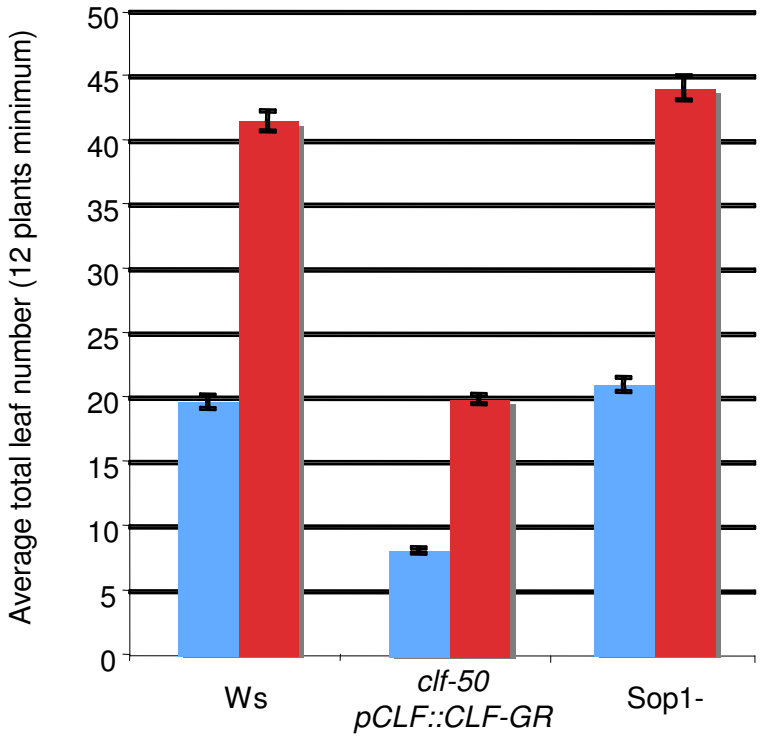

Figure 3.7: Flowering time experiment under LD and SD conditions. The results show the average of rosette leaves at bolting (12 plants minimum). Red: SD conditions. Blue: LD conditions. Genotypes tested: Ws, *clf50 pCLF:CLF-GR* and Sop1-. Bars show Standard Error (SE).

From this experiment I could conclude that *sep3-7* fully suppresses the early flowering phenotype due to the *clf-50* mutation and restores a wild type flowering-time phenotype. Also, it has been proposed that *SEP3* and *AG* can interact together in the same complex (Castillejo et al., 2005) so, the early flowering phenotype might be due to a mis-regulation of a target of the complex formed by *SEP3* and *AG*.

3.2.3. *AG* is highly expressed in Sop1- plants.

At the molecular level, it has been shown that *AG* is highly misexpressed in leaves of *clf-2* mutants (but not in wild type leaves). Also, the *ag-3* mutation suppresses the curled leaf phenotype of *clf-2* indicating that one of the factors causing the Clf- phenotype is the ectopic expression of *AG* (Goodrich et al., 1997) (chapter 1 section 1.3.3). In order to know if the loss of the curled leaf in Sop1- plants is linked to a down regulation of *AG* by *SEP3*, semi-quantitative RT-PCRs were conducted on 15 day old seedlings of Ws, *clf-50 pCLF:CLF-GR* and Sop1- plants (Figure 3.8). In Ws plants, a very low level of expression of *AG* can be detected in this experiment. It is because in LD conditions plants undergo the floral transition around ten days after germination (Bradley et al., 1997). Then, the expression detected is certainly from the Shoot Apical Meristem (SAM). In the progenitor line, *clf-50 pCLF:CLF-GR*, the level of expression of *AG* relative to tubulin is approximately 50 times higher than in Ws plants (Figure 3.8). In Sop1- plants, *AG* expression's level is 17 times higher than in the Ws control relative with the tubulin expression. The expression of *AG* in Sop1- plants is approximatively 66% lower than in the progenitor line *clf-50 pCLF:CLF-GR* (Figure 3.8). This is lower than in the progenitor line but significantly higher than in wild type plants.

From this experiment, I could conclude that the loss of the Clf- phenotype in Sop1- plants does not correlate with a total down regulation of *AG*.

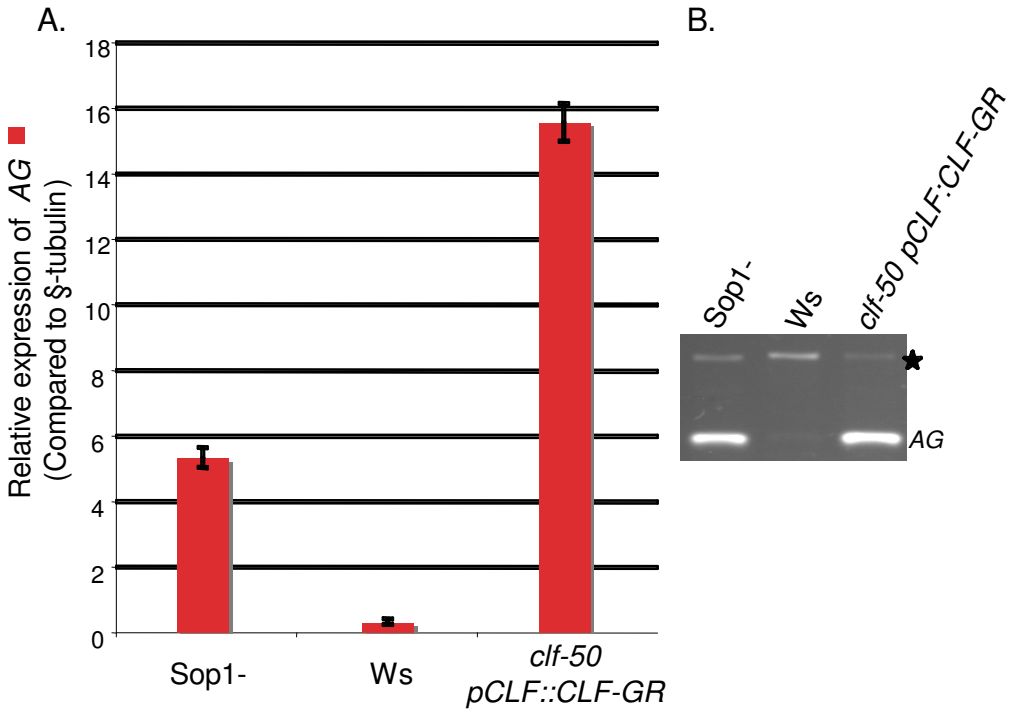

Figure 3.8: Expression of *AG* in Sop1- plants after 30 PCR cycles. A. Quantification of the transcript levels of *AG* in Sop1-, Ws and *clf50 pCLF:CLF-GR* plants The quantification corresponds to the average of the ratio (target gene signal/ß-tubulin signal) of three different RT-PCR reactions. Bars show Standard Error (SE). B. Example of RT-PCR of *AG* on Sop1-, Ws and *clf50 pCLF:CLF-GR* cDNA samples. Duplex PCRs were performed to amplify ß-tubulin (marked with ★) as an internal control with the specific primers for *AG*. The plants analysed were grown on plates and the mRNA was extracted from 15 day old seedlings.

It has been shown that ectopic expression of *SEP3* (*35S:SEP3*) is sufficient to activate *AP3* and *AG* in leaves (Castillejo et al., 2005). The authors made the hypothesis that *SEP3* might be an activator of *AG* in *A. thaliana*.

The results previously exposed strongly support the same conclusion as the expression levels of *AG* in Sop1- plants is reduced of 66%compared to the levels observed for the progenitor line *cl-f50 pCLF:CLF-GR* (Figure 3.8).

3.2.4. *SEP3* is a co-factor of *AG* in *Arabidopsis*

It has been shown that *SHP2*, a MADS-box transcription factor involved in carpel development is an *AG* target (Savidge et al., 1995). Also, *SHP2* is a downstream target of a complex formed by SEP and AG (Castillejo et al., 2005). Moreover, complexes containing the SHP2 protein and involved in ovule differentiation can only be formed in the presence of SEP proteins (Battaglia et al., 2006). This suggests that *SEP3* is a co-factor of *AG* and that *SHP2* is one of the targets of this MADS-box complex.

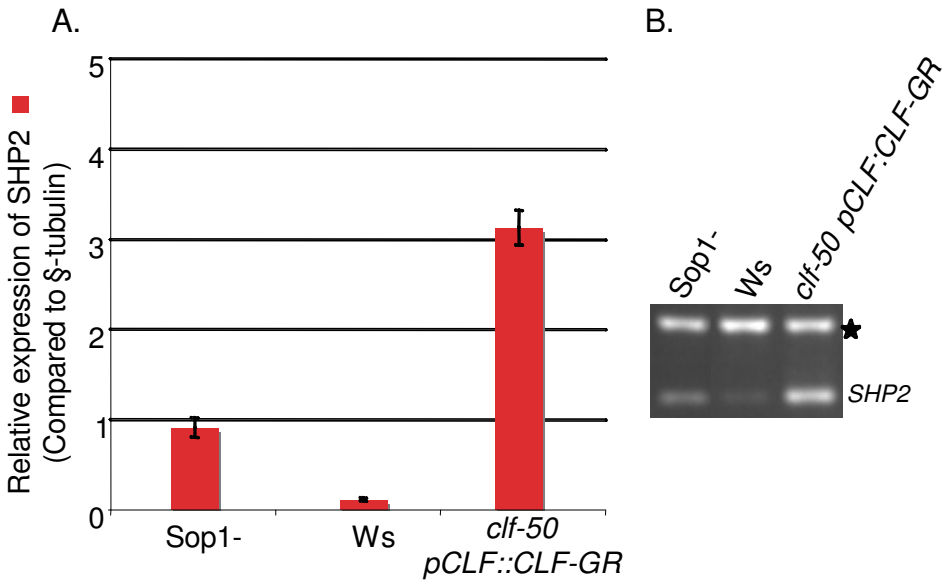

Figure 3.9: Expression of SHP2 in Sop1- plants after 32 PCR cycles. A. Quantification of the transcript levels of *SHP2* (red) in Sop1-, Ws and *clf50 pCLF:CLF-GR* plants. The quantification corresponds to the average of the (target gene signal/ß-tubulin signal) ratio of three different RT-PCR reactions. Bars show Standard Error (SE). B. Example of RT-PCR of *SHP2* on Sop1-, Ws and *clf50 pCLF:CLF-GR* cDNA samples. Duplex PCRs were performed to amplify ß-tubulin (marked with ★) as an internal control with the specific primers for *SHP2*. The plants analyzed were grown on plates and the mRNA was extracted from 15 day old seedlings.

I tested this hypothesis by analyzing the expression level of *SHP2* in Sop1-, Ws and *clf-50 pCLF:CLF-GR* plants (Figure 3.9). As previously described in Figure 3.8 for *AG*, in Figure 3.9 there is a reminiscent expression of *SHP2* in Ws plants. This is certainly due to the same reasons previously cited. In the progenitor line, the level of expression of *SHP2* is approximately 30 times higher than in Ws plants. However, in Sop1- plants *SHP2* expression is only ten times higher than in Ws plants.

The presence of the AG protein was tested in Sop1- plants (Figure 3.10). In wild type flower buds, the AG protein is detected at 30 kDa whereas in the *ag* mutant it is not. Wild type seedlings as expected did not show the AG protein. In Sop1- seedlings the AG protein is present. The level of expression of the AG protein is surprisingly similar to the level observed for the progenitor line *clf-50 pCLF:CLF-GR*. As no accurate protein quantification was performed, this result is expected and the only certitude that we can state is that the AG protein is present in Sop1- plants.

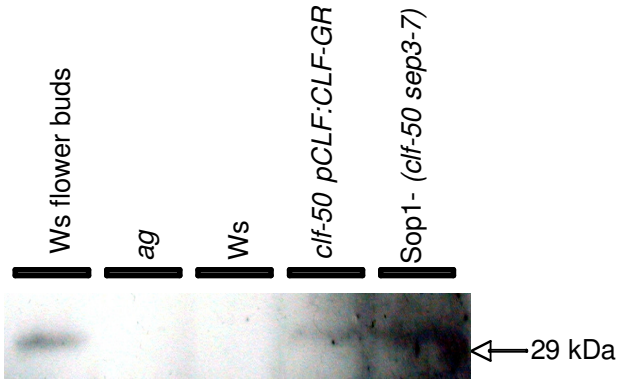

Figure 3.10: Western-blot analysis of AG in Sop1- plants. Total protein was extracted from 14 days old seedlings grown on soil on LD conditions. The AG protein is detected at 30 kDa by an antibody directed against the carboxy terminus of AG. The positive control used was wild type flower buds. The negative control (*ag* mutant) does not present the expected band at 30 kDa.

From this mRNA and protein expression analysis I could deduce that *SEP3* is necessary for the activation of *SHP2* by *AG*. Also, the results in Figure 3.8 and in Figure 3.10 where the presence of *AG* at the level protein was assessed support the hypothesis that SEP3 is a co-factor of AG in *A. thaliana*.

3.2.5. The *FT* mutation suppresses the Clf- phenotype

It has been shown that the expression of *SEP3* in mature rosette leaves is regulated by *FLOWERING LOCUS T* (*FT*) (Teper-Bamnolker and Samach, 2005). Indeed, transgenic lines carrying a *35S:FT* construct show a curled leaf phenotype and expression of *SEP3* was highly mis-expressed. In addition, the curled leaf phenotype due to the ectopic expression of *FT* is suppressed by the *sep3-2* mutation. To know if the Clf- phenotype observed in *clf* mutants is due to *SEP3* or *FT*, the *ft-10* mutation was introduced in the double mutant *clf-28 flc-3*. We used the double mutant *clf-28 flc-3* background to show that the suppression is independent of *FLC*. Indeed, *clf-28* allele has a relative weak leaf curling due to *FLC* up-regulation. Thus, the double mutant *clf-28 flc-3* has a severe leaf curling and an early flowering phenotype. So, I tested if *ft-10* can suppress the severe phenotype of the double mutant *clf-28 flc-3*. The F1 generation from the cross between the single mutant *ft-10* and the double mutant *clf-28 flc-3* exhibited a wild type phenotype. Plants displaying the Clf- phenotype were selected on the F2 generation and the seeds harvested. In the F3 generation, plants were genotyped for the *ft-10* and *flc-3* mutations (Figure 3.11-A.). The wild type DNA control (Col0) shows a band at 1.3 Kbp and at 0.5 Kbp corresponding to the wild type allele of *FT* and *FLC* respectively. All the plants tested were homozygous for the *flc-3* mutation.

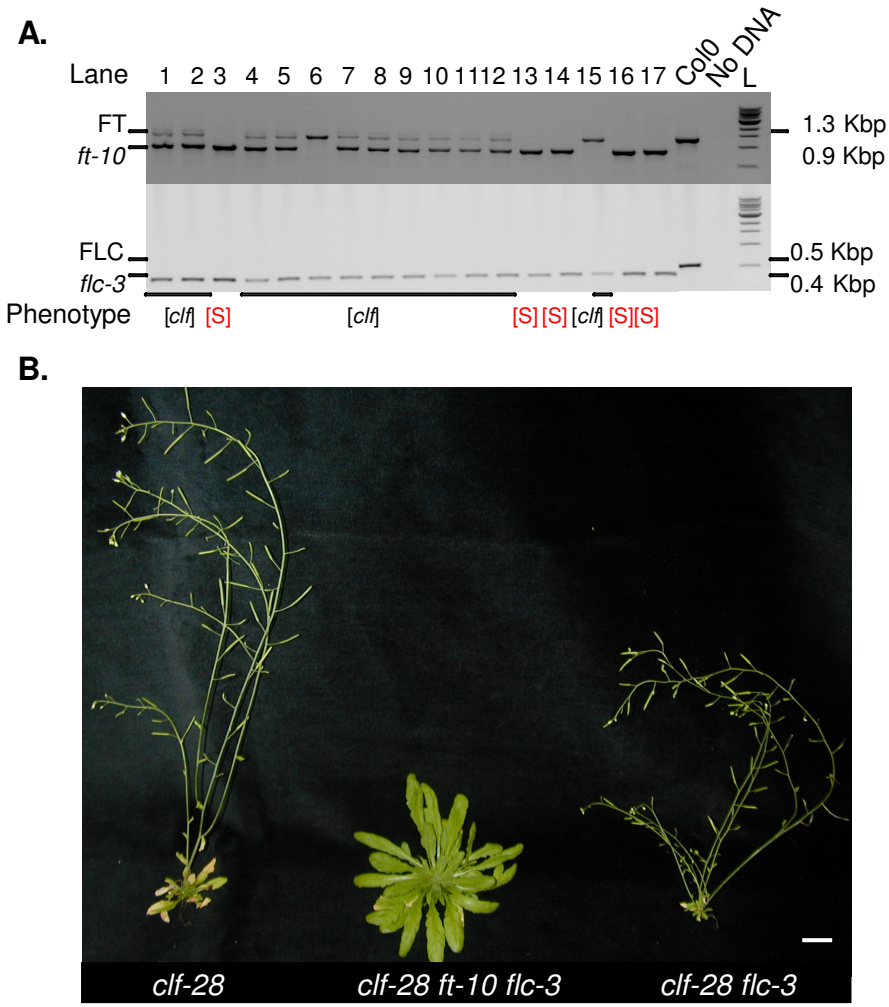

Figure 3.11: The *ft-10* mutation suppresses the Clf- phenotype. A. Genotyping PCR for *flc-3 ft-10* plants on 17 F3 plants. Genotyping primer sets (JH2295, JH2296, and JH2297) and (flc-3-F and flc-3-R) were used to identify homozygous *ft-10 flc-3* mutations respectively (Yoo et al., 2005). Homozygous *ft-10* plants produced a single band of 926 bp in size, whereas wild-type plants produced a band of 1,392 bp in size. Homozygous *flc-3* plants produced a single band of 400 bp in size, whereas wild-type plants produced a band of 500bp in size. L, One-kilobase ladder. B. Phenotype of the triple mutant *clf-28 flc-3 ft-10*. Plants were growing during four weeks in LD conditions. Scale bar = 1cm.

Twelve out of seventeen plants showed a Clf- phenotype and had at least one copy of the wild type allele of *FT* as shown in Figure 3.11 A. (lanes 1, 2, 4, 5, 6, 7, 8, 9, 10,

11, 12 and 15). Five plants showed a suppressed phenotype and are *ft-10* homozygous as shown in Figure 3.11 A. in lanes 3, 13, 14, 16 and 17. The single mutant *clf-28* and the double mutant *clf-28 flc-3* show characteristic traits of the Clf- phenotype. In Figure 3.11 B. the triple mutant *clf-28 flc-3 ft-10* displays a suppressed Clf- phenotype and is late flowering compare with *clf-28* and *clf-28 flc-3* genotypes. At the moment, it is not known if the triple mutant *clf-28 flc-3 ft-10* flowers later than wild type Col-0 plants.

A whole genome expression analysis performed in the laboratory showed that the expression of *FT* in *clf-28* is 9.5 times higher compared to wild type plants Col0 (Dr. F. Thorpe personal communication). Also, the CLF and H3K27me3 ChIP on chip database show that the FT locus presents enrichment for CLF and H3K27me3 (Dr. O. Clarenz personal communication and (Zhang et al., 2007b)). Recently, it was shown that during vegetative development, *CLF* and *FIE* strongly repress the expression of *FT* and that CLF also directly interacts with and mediates the deposition of H3K27me3 into FT chromatin (Jiang et al., 2008). As the over-expression of *FT* leads to the over-expression of *SEP3* (Teper-Bamnolker and Samach, 2005); an explanation of the Clf- phenotype could be that the over-expression of *FT* induces the expression of *SEP3* which is likely to be an activator and a co-factor of *AG* as it was previously proposed in this study and in the litterature (Castillejo et al., 2005).

Taking all together, these results suggest that in absence of *CLF*, *AG* maybe activated by *FT* via SEP3. In this case, *FT* might act indirectly on *AG*.

3.3.SEP3 is a novel Pc-G target

To know if *SEP3* is directly regulated by CLF as *AG* is, the expression of *SEP3* in the *clf* mutant background was considered. The enrichment in CLF and H3K27me3 was

also analyzed at the *SEP3* locus using the ChIP-chip dataset generated in the laboratory. Semi-quantitative ChIP PCR confirmed the results obtained by ChIP-chip.

3.3.1. *SEP3* is mis-expressed in the Clf- background

The microarray expression analysis performed in the laboratory by Dr. F. Thorpe, gave a "snap-shot" picture of the different genetic mis-regulations generated by the lack of the CLF protein. After ten days of *in-vitro* culture, *SEP3* is the fourth most up-regulated loci in *clf-28* compared to wild type Col-0 plants. *SEP3* is up-regulated 16 fold in *clf-28* relative to Col, which genotype is CLF (Dr. F. Thorpe personal communication). In order to confirm this result for *clf-50 pCLF:CLF-GR* and Sop1- plants, semi-quantitative RT-PCR on *SEP3* was performed (Figure 3.12). The expression of *SEP3* in wild type Ws plants is not detectable for the same reasons as previously described in section 3.2.3. There is no *SEP3* expression in Sop1- plants. This is expected, as the *sep3-7* mutation is a deletion. The expression of *SEP3* in *clf-50 pCLF:CLF-GR* plants is 8.6 times higher than the expression observed for wild type Ws plants in the same growth conditions. In the two experiments (microarrays and semi-quantitative RT-PCR), *SEP3* is highly up-regulated in the *clf* mutant background.

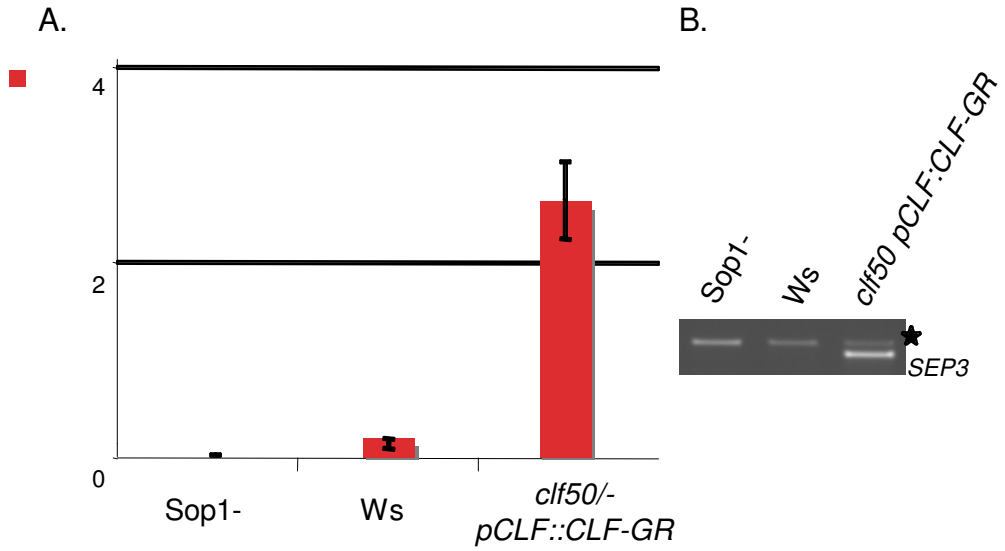

Figure 3.12: Expression of *SEP3* in *clf-50 pCLF:CLF-GR* and Sop1- plants. A. Quantification of the transcript levels of *SEP3* (red) in Sop1-, Ws and *clf50 pCLF:CLF-GR* plants The quantification corresponds to the average of the (target gene signal/ß-tubulin signal) ratio of three different RT-PCR reactions. Bars show Standard Error (SE). B. Example of RT-PCR of *SEP3* on Sop1-, Ws and *clf50 pCLF:CLF-GR* cDNA samples. Duplex PCRs were performed to amplify ß-tubulin (marked with ★) as internal control with the specific primers for *SEP3*. The plants analysed were grown on plates and the mRNA was extracted from 15 day old seedlings.

The level of mis-expression of *SEP3* in the two experiments is not the same (16 versus 8.6 times). This difference might be due to the higher detection sensitivity of the microarray technique. Moreover, the alleles used in the two experiments are different, *clf-28* is a T-DNA insertion and *clf-50* is a deletion. In the case of *clf-28*, residual transcription might occur and a truncated CLF protein might enhance *SEP3*'s mis-regulation.

Finally, the results confirm the total absence of expression of *SEP3* in Sop1- plants due to the *sep3-7* mutation and we confirm the up-regulation of the *SEP3* locus observed in the microarray experiment. We can then conclude that *CLF* is one of the factors regulating *SEP3* expression in *A. thaliana*.

3.3.2. CLF binds to *SEP3*

CLF must bind the *SEP3* locus to directly regulate its expression. The ChIP-chip dataset, generated in the laboratory by Dr. O. Clarenz in collaboration with the laboratory of Dr. S. Jacobsen, was analyzed to determine (i) if CLF binds *SEP3* sequences and (ii) if there is an enrichment of H3K27me3 at the *SEP3* locus. As there is no specific antibody for CLF, an epitope-tag protein GFP-CLF was engineered. Chromatin from 35S:GFP-CLF *clf-28* seedlings was immunoprecipitated using an antibody against GFP. The chromatin obtained was hybridized with whole-genome tiling arrays as described in (Zhang et al., 2007b). The analysis of the dataset shows enrichment for the repressive H3K27me3 histone modification at *FUSCA3 (FUS)*, *AGAMOUS (AG)*, *MEDEA (MEA)*, *PHERES1 (PHE)*, *SHOOT MERISTEMLESS (STM)* and *AGAMOUS-LIKE19 (AGL19)*, which are known to be Pc-G target in *A. thaliana* (Kohler et al., 2003; Jullien et al., 2006; Makarevich et al., 2006; Schonrock et al., 2006; Schubert et al., 2006). For the *SEP3* locus enrichment on CLF and H3K27me3 was observed (Figure 3.13). The entire locus has high levels of H3K27me3 as previous cited targets (Zhang et al., 2007b). Also, the binding of CLF is dispersed through the whole *SEP3* locus as shown in Figure 3.13.

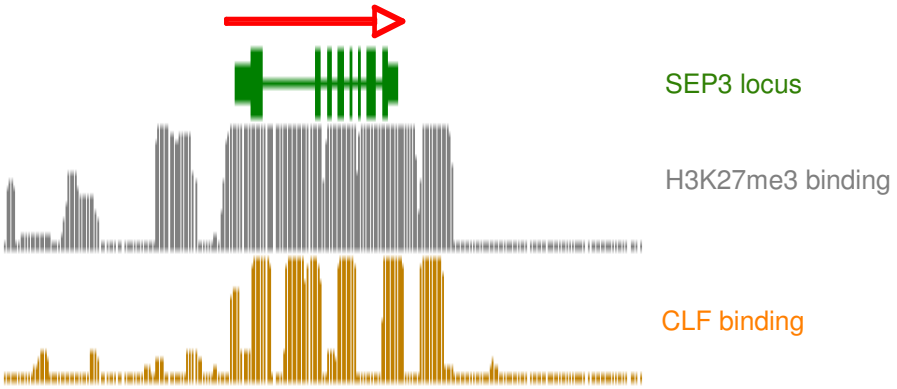

Figure 3.13: ChIP-chip results for the *SEP3* locus (red arrow indicate direction of transcription). The *SEP3* locus is shown as green boxes (exons) and lines (introns), H3K27me3 binding is shown as vertical grey bars and CLF binding is shown as vertical orange bars

These results show that CLF binds *SEP3 in vivo*, and is dispersed all over the *SEP3* locus. This suggests that *SEP3* is a direct target of CLF as *SEP3* sequences are specifically enriched on H3K27me3.

3.3.3. The *SEP3* locus is enriched on H3K27me3 in wild type *A. thaliana* plants

To confirm that the repressive H3K27me3 histone mark is present at the *SEP3* locus, chromatin from Ws and *clf-50* seedlings was tested by semi-quantitative ChIP PCRs. The IPs were performed with antibodies against H3K27me3 from wild-type Ws, and *clf-50* chromatin of ten days old plants. As a control I used *AG* sequences known to be enriched for H3K27me3 (*AG4*) and *At4g18970* 5′ sequences known not to be enriched for H3K27me3 (*AG3*) in wild type plants (Figure 3.14-A.).

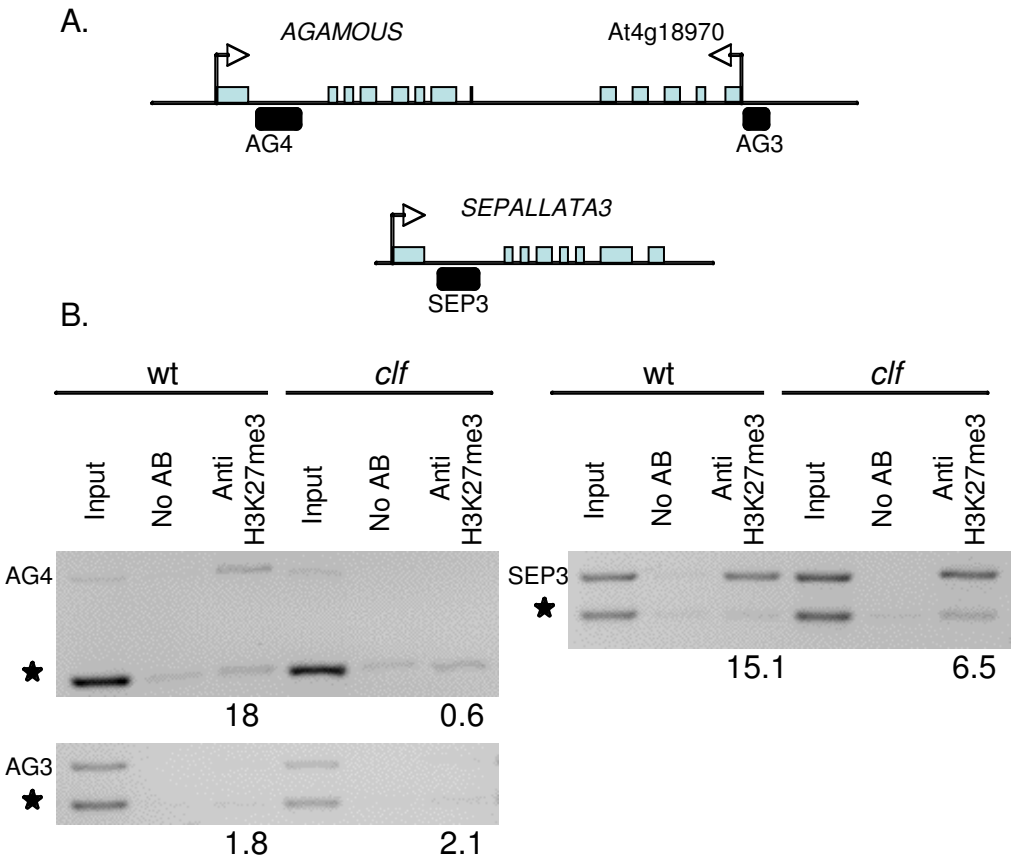

Figure 3.14: H3K27me3 patterns of *SEP3* in wild type and *clf-50* mutant plants. A: Location of the PCR products for *AG4*, *AG3* and *SEP3*. B: IPs were performed without antibody ("No AB") or with an antibody against H3K27me3 ("Anti-H3K27me3") from wild type (Ws) and *clf-50* chromatin extracted from 10 days old seedlings. Precipitated DNA was amplified by PCR, regions amplified are shown in A. Calculation of relative enrichment is described in chapter 2 section 2.7.2. Semi-quantitative ChIP duplex PCRs were used to amplify *ACTIN* (marked with a star) as internal controls. Numbers below the lane indicate the ratio of the intensity of *AG4*, *AG3* or *SEP3* products, respectively compared to *ACTIN* intensity after IP normalized to the ratio before IP ("Input"). Each experiment was performed three times; representative experiment is shown

The results for the control sequences *AG4* and *AG3* in Figure 3.14-B show enrichment of 30 times higher in Ws than in *clf-50* for *AG4* and no enrichment for Ws (1.8) and *clf-50* (2.1) for *AG3* (Figure 3.14-B). For the tested sequence of *SEP3*, the semi-quantitative PCR show that the *SEP3* locus is enriched in H3K27me3 three times more in Ws than in *clf-50*. This result confirms the patterns observed on ChIP-chip for *SEP3*. Moreover, Figure 3.14-B confirms that the first intron of *SEP3* is highly enriched in H3K27me3 as shown in Figure 3.13.

Using ChIP experiments I could confirm that the repressive H3K27me3 mark is present at the *SEP3* locus. This result plus the analysis of the ChIP on chip dataset show that CLF bind the *SEP3* locus and that the *SEP3* is enriched in H3K27me3. So, *SEP3* is a target of CLF.

3.4. Summary

In this chapter I describe the identification of the mutation responsible of the suppressed phenotype of Sop1- plants. It is a loss of function mutation in *SEP3*. *SEPALLATA3* is a member of the MADS-box transcription factor family redundant with SEP1 and SEP2, which are involved in floral organ identity at the latest stages of flower formation. Here, I provide genetic and molecular evidence showing that *SEP3* acts during floral organ formation as an activator and a co-factor of *AG* in *A. thaliana*. In addition, we show that *SEP3*'s chromatin is enriched in H3K27me3 and that CLF binds the *SEP3* locus. The last results provide sufficient elements to pointing out SEP3 as a direct target of CLF in *A. thaliana*.

4. *SUPPRESSOR OF POLYCOMB 4* (*SOP4*) is a loss of function mutation in *FPA*

The *suppressor of polycomb 4* (*sop4*) mutation was isolated from the same screen, in tissue culture plates, as *sop1*. Sop4- plants showed a strong suppression of the Clf- phenotype as shown in Figure 4.5 in section 4.2.1. Only seeds from *sop4* homozygous mutant plants were available in the laboratory for further analysis.
In this chapter, the *SOP4* gene was identified and the Sop4- phenotype was characterized. As previously undertaken for the Sop1- plants, the relationship between the mutation causing the Sop4- phenotype and the *clf* mutation was studied.

4.1. Identification of the mutation responsible for the Sop4- phenotype

For this suppressor, a segregating family was not available in the laboratory. Thus, it was decided to accelerate the identification of the T-DNA responsible for the Sop4- phenotype, by adopting a "shot-gun" strategy. Southern-blot analyses were not conducted, however mutant genomic DNA was used to perform plasmid rescue experiments. The suppressed phenotype could be linked with a T-DNA insertion using genotyping PCR. Also, in collaboration with the laboratory of Dr. G. Simpson, western-blot analysis confirmed the absence of the protein encoded by the ORF where the T-DNA was inserted in Sop4- plants.

4.1.1. Plasmid rescue using a "shot-gun" strategy

Genomic DNA from Sop4- plants was used to plasmid rescue the Right Border (RB) of the T-DNA and adjacent plant sequences as described in chapter 2 section 2.4.8. Plasmid DNA was extracted from transformed *E. coli* colonies obtained, and digested with the restriction enzyme *Eco*RI and size separated by electrophoresis (Figure 4.1). This digestion allowed us to identify several plasmids presumably corresponding to different T-DNA insertions at different loci. Fragments sizes will change according to the position of *Eco*RI sites in the plant DNA flanking the T-DNA insertion.

The digestions revealed that only one T-DNA insertion from Sop4-'s DNA could be recovered by plasmid rescue. The size of the plasmids obtained in lanes one to seven is the same, approximately 5.8Kbp (Figure 4.1).

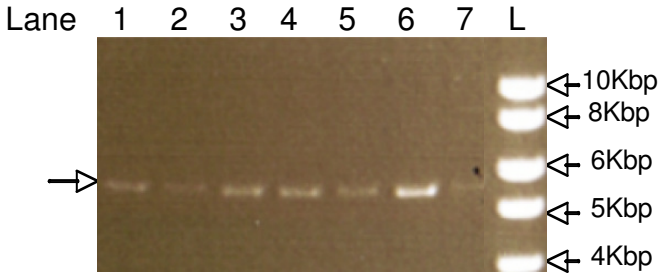

Figure 4.1: Analysis of plasmid rescue clones from Sop4- plants. Lanes 1 to 7 correspond to 7 independent plasmid tested by *Eco*RI. In all lanes a plasmid of approximately 5.8Kbp is present. L corresponds to the 1Kbp ladder used to assess the size of the bands obtained (sizes showed on the right hand side of the figure).

The "SEQ" primer located in the T-DNA between the pBstKS+ sequences and the closest *Eco*RI restriction site to the RB (Figure 4.2) was used to identify plant genomic sequences by sequencing. Three plasmids (lanes 1; 4 and 5 from Figure 4.1) were sequenced in order to find the *A. thaliana* genomic sequence where the T-DNA was inserted. The plant sequence found for the three different plasmids was in the

first intron of *FPA* (*At2g43410*). *FPA* is a member of the flowering autonomous pathway, which consists on seven genes: *FPA*; *FCA*; *FLK*; *FY*; *FLD*; *FVE* and *LUMINIDEPENDENS* (*LD*). They promote flowering by repressing the accumulation of *FLC*'s mRNA in *A. thaliana* (Michaels and Amasino, 2001; Quesada et al., 2005). *FPA* and *FCA* are two plant-specific RNA binding proteins containing a RNA Recognition Motif (RRM) domain (Schomburg et al., 2001; Quesada et al., 2005).

4.1.2. Fine mapping of the T-DNA insertion in *FPA*

Using the complete genome sequence from The *Arabidopsis* Information Resource (TAIR) the primer "FPA-seq" (Figure 4.2) in the first intron of *FPA* was designed and the insertion point of the T-DNA was finely mapped as shown in Figure 4.2. The T-DNA is inserted in the first intron, 1072bp after the transcription start of *FPA*. Consequently, the T-DNA insertion identified here is a novel *fpa* allele hereafter called *fpa-10*.

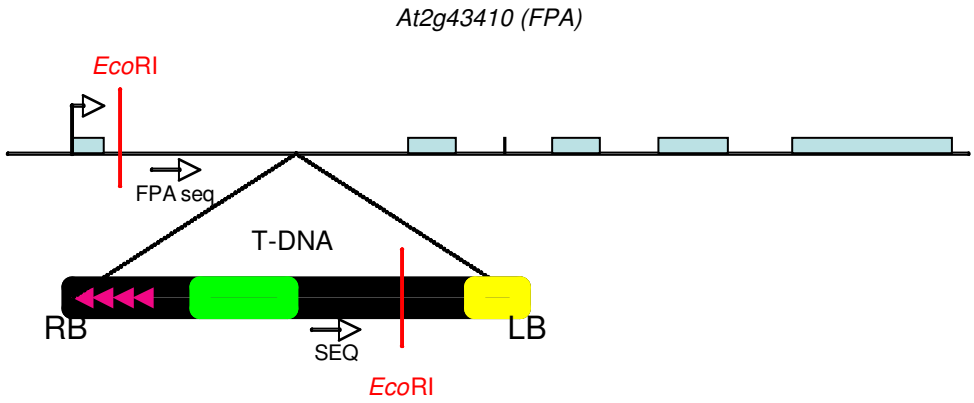

Figure 4.2: Diagram of the T-DNA insertion at the FPA locus found in Sop4- plants. The T-DNA insertion is located in the first intron of *FPA*. Arrows correspond to primers "FPA-seq" and "SEQ" used to map the *fpa-10* mutation. The *Eco*RI restriction sites allowing the recovery of FPA genomic sequences are shown in red. For details see primers table chapter 2 section 2.4.3.

At this stage, I could not eliminate the possibility that the Sop4- phenotype was provoked by another T-DNA insertion in the genome. Although from the "shot-gun" strategy only the insertion in *FPA* was identified. Nonetheless, the *fpa-10* mutation was the best candidate to be responsible for the suppressed phenotype of Sop4- plants as they are late flowering, as are *fpa* mutants.

4.1.3. The mutation responsible for the Sop4- phenotype is recessive

Sop4- plants were crossed with the progenitor line *clf-50 pCLF:CLF-GR* to genetically characterize the *sop4* mutation. All F1 plants had Sop4+ phenotype, indicating that the *sop4* mutation is recessivein the *clf* background. The segregation in the F2 generation was 3:1 expected for a single recessive mutation. The phenotype of the

progeny of five independent F2 families was scored and on average 10.2 ± 0.4 Sop4- plants and 30.4 ± 2.3 Clf- plants were obtained out of 40 plants sown on soil (Table 4.1). In all the lines tested, segregation occurred at a ratio of 3:1 (Clf- to Sop4-) fitting with the Mendelian segregation ratio of a single recessive mutation. Heterogeneity chi square values indicated no difference between the lines tested and the 3:1 ratio expected. The segregation results show that only one mutation (ratio 3:1) is responsible for the Sop4- phenotype and is recessive in the *clf* mutant background.

Table 4.1: Segregation analysis show *sop4* is a single recessive mutation. The phenotype of plants in the F2 generation from the cross between Sop4- and the progenitor line *clf-50 pCLF:CLF-GR* was assesed. For the statistical analysis, a Chi Square test for 3:1 segregation of the Clf- and Sop4- phenotypes was performed. NS: not significant at P=0.05.

	clf **plants**	***sop4*** **plants**	**total plants**	χ^2 **(3:1)**	**P**
Line 1	25.0	9.0	34.0	0.04	0.84
Line 2	33.0	10.0	43.0	0.07	0.79
Line 3	35.0	11.0	46.0	0.03	0.86
Line 4	33.0	10.0	43.0	0.07	0.79
Line 5	26.0	11.0	37.0	0.44	0.51

4.1.4. Cosegregation of *fpa-10* with the Sop4- phenotype

Genomic DNA was extracted from Clf- and Sop4- plants from the F2 progeny previously cited in section 4.1.3. Genotyping PCRs with specific *FPA* and T-DNA primers were performed in order to confirm that *fpa-10* cosegregates with the Sop4- phenotype (Figure 4.3). In Figure 4.3-A the location of the primers used and the length of the expected amplification products at the *FPA* locus are shown. The PCR bands obtained did not show the expected size, as for the *fpa-10* allele a band at 3 Kbp was obtained instead of the 0.9 Kbp band expected. This is certainly due to a

complex T-DNA insertion where sequences from the binary vector or a plant genomic duplication, were added at the Left Border (LB) of the T-DNA insertion site. For the wild type *FPA* allele, the expected band at 1.4 Kbp was obtained (Figure 4.3-B). Using homozygous *sop4*, wild type (Ws) and *clf-50 pCLF:CLF-GR* DNA controls, homozygous *fpa-10*, heterozygous *fpa-10/FPA* and wild type plants were discriminated (Figure 4.3-B). The PCR results in Figure 4.3-B show that all Sop4+ (+) plants are either homozygous *FPA* (lanes 1, 4 and 6) or heterozygous *fpa-10/FPA* (lane 7, which has a low intensity band at 3 Kbp) at the *FPA* locus. Whereas, all Sop4- plants are homozygous for the *fpa-10* allele (lanes 2, 3 and 5) as predicted if the *fpa-10* mutation is responsible for the Sop4- phenotype. In addition, the cosegregation of the Sop4- phenotype was confirmed in the following generation as the phenotypes observed were: (i) 100% Sop4+ from homozygous *FPA* plants (in lanes 1, 4 and 6). (ii) ¾ Sop4+ and ¼ Sop4- from heterozygous *fpa-10/FPA* plants (in lane 7) and 100% Sop4- from homozygous *fpa-10* plants (in lanes 2, 3 and 5).

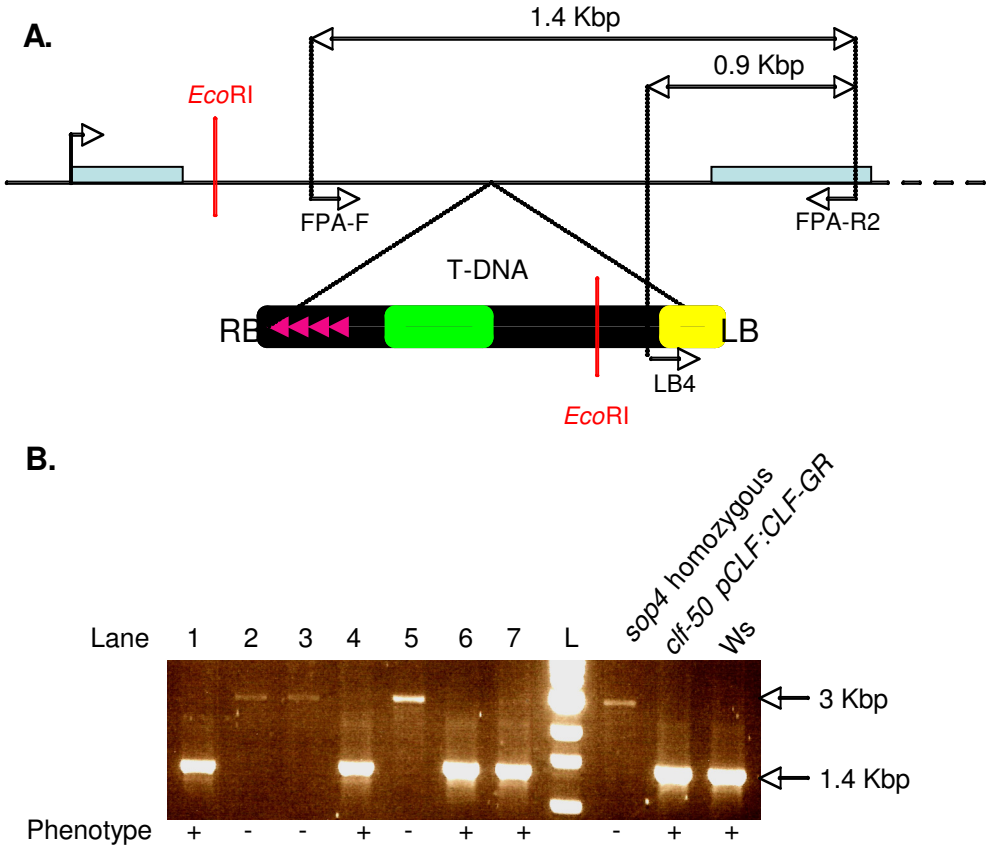

Figure 4.3: Genotyping of segregating F2 plants at the *FPA* locus. A: Location of primers FPA-F; FPA-R2 at the *FPA* locus and LB4 in the T-DNA. PCR products will amplify a fragment of 1.4 Kbp for the wild type *FPA* allele and 0.9 Kbp for the *fpa-10* mutation. The *Eco*RI restriction sites allowing the recovery of FPA genomic sequences are shown in red. B: Results for 7 plants (lanes 1 to 10). The band at 1.4 Kbp correspond to the wild type product (controls on the right hand side: *clf-50 pCLF:CLF-GR* and Ws). The higher band at 3 Kbp corresponds to the *fpa-10* allele (control on the right hand side: *sop4* homozygous). Samples in lanes 2, 3 and 5 correspond to homozygous *fpa-10* plants. Samples in lanes 1, 4 and 6 correspond to homozygous *FPA* plants. Sample in lane 7 correspond to heterozygous *fpa-10*/*FPA* plants (band at 3 Kbp has a low intensity). The plant phenotype (+) corresponds to Sop4+ and (-) to Sop4- plants.

These results strongly suggest that the mutation responsible for the Sop4- phenotype is the T-DNA insertion in the first intron of *FPA*, as the the *fpa-10* mutation cosegregates with the Sop4- phenotype.

4.1.5. The FPA protein is not detected in *fpa-10* plants

In collaboration with the laboratory of Dr. G. Simpson, FPA protein levels in Sop4- plants were monitored by western-blot. Total proteins were extracted from seedlings grown for 14 days on soil under LD conditions and western-blot analysis was conducted by Dr. L. Terzi using antibodies directed against the carboxy terminus of FPA and detecting FPA at 98 kDA.

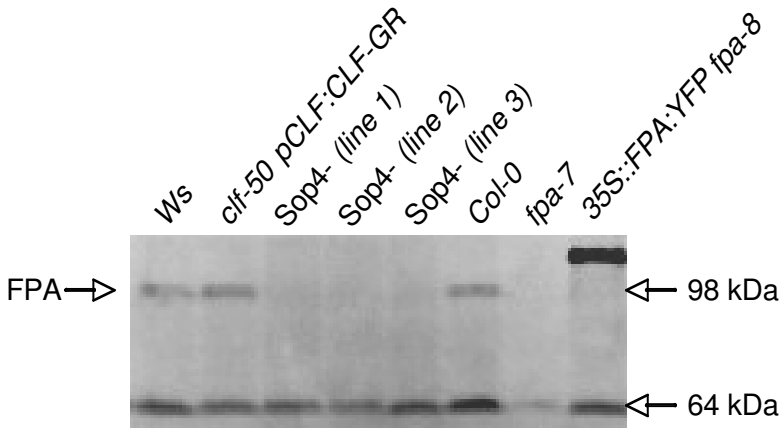

Figure 4. 4: Western-blot analysis of FPA protein levels in Sop4- plants. Total protein was extracted from 14 days old seedlings grown on soil on LD conditions. The FPA protein is detected at 98 kDa by an antibody directed against the carboxy terminus of FPA. All the plant extracts showed a band at 64 kDa corresponding to a non specific band. Data kindly provided by Dr. L. Terzi & Dr. G. Simpson.

The western-blot in Figure 4. 4 shows that the FPA protein is present at 98 kDA in all the control lines used: Ws, *clf-50 pCLF:CLF-GR* and Col-0. Also, the *fpa-7* mutant in the Columbia background did not show the expected band at 98 kDA and the

over-expression of the fusion protein FPA:YFP showed a band higher than 98 kDa (Figure 4. 4). Three independent Sop4- lines (1, 2 and 3) were tested for the presence of the FPA protein. The FPA protein was not detected at 98 kDa in any of the lines tested as expected (Figure 4. 4).

We could conclude that the FPA protein cannot be detected in *fpa-10* plants with the antibodies used. Collating the results of this section, the *fpa-10* mutation certainly suppresses the Clf- phenotype by depleting the FPA activity in *A. thaliana* plants.

4.2. Characterization of the *sop4* mutation

As Sop1- plants, the principal phenotypic trait permitting the isolation of Sop4- plants was the suppression of the leaf curling of the *clf-50 pCLF:CLF-GR* progenitor line. In addition, Sop4- plants also showed a very late flowering phenotype. In the following section, the Sop4- phenotype was characterized in more detail at the vegetative stage and the flowering-time phenotype of Sop4- plants was analyzed. The stability of the mRNA of AG was tested in Sop4- plants and finally several MADS-box transcription factors were tested as possible targets for FPA.

4.2.1. Phenotype of Sop4- plants at the early stage of development

The phenotype exhibit by Sop4- plants in early stages of development is very close to wild type (Ws) plants. Indeed, at this point the only feature distinguishing Sop4- plants from Clf- plants is the total suppression of the leaf curling Figure 4. 5. Sop4- plants are late flowering and show increased number of rosette leaves relative to wild type Ws (Figure 4. 5).

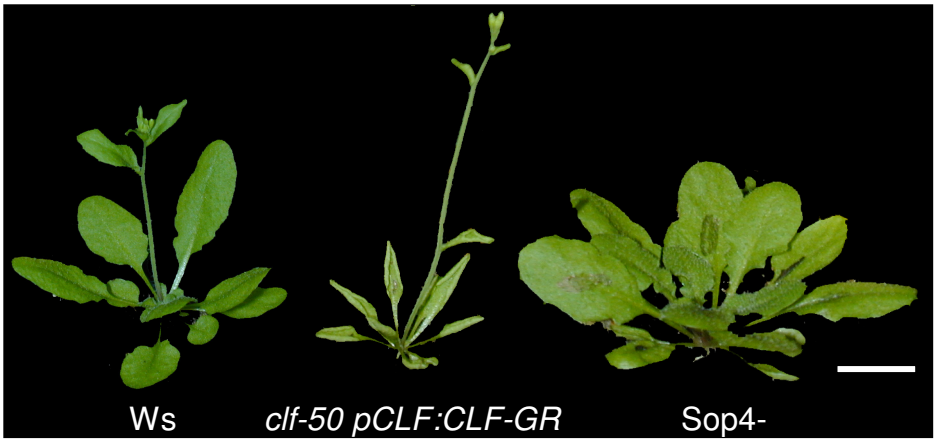

Figure 4. 5: Phenotype of Sop4- plants at late vegetative stages. Comparison between wild type Arabidopsis ecotype Ws, the progenitor line *clf-50 pCLF:CLF-GR* and Sop4- plants. Plants were growing during 3 weeks on soil in LD conditions. Scale bar = 1 cm.

4.2.2. Sop4- plants are late flowering due to a high levels of *FLC*

FPA is known to interact with genes involved in flowering, including known targets of the PRC2 as *FLC*. It has been reported that the late flowering phenotype of a null allele of *fpa* was abolished by a null allele of *flc* (Michaels and Amasino, 2001), indicating that the effects of FPA on flowering are exclusively mediated by *FLC*. Although, *fpa* like other autonomous pathway mutants such as *fca*, is not needed for the vernalization response as *fpa* mutants flower as wild type if vernalized. Moreover, *FLC* activity is post-transcriptionally repressed by *FPA*, although the specifics of this interaction are not known (Schomburg et al., 2001). Thus, *FPA* act upstream of *FLC* and promote flowering by inhibiting *FLC* expression. In order to (i) test if the double mutation *fpa-10 clf-50* in Sop4- plants has normal vernalization response as it would be expected for *fpa* mutants and (ii) to quantify the late-flowering phenotype of Sop4- plants a time to flower experiment was conducted under LD conditions with vernalization (V) and without 4 weeks of vernalization (NV) respectively (Figure 4. 6).

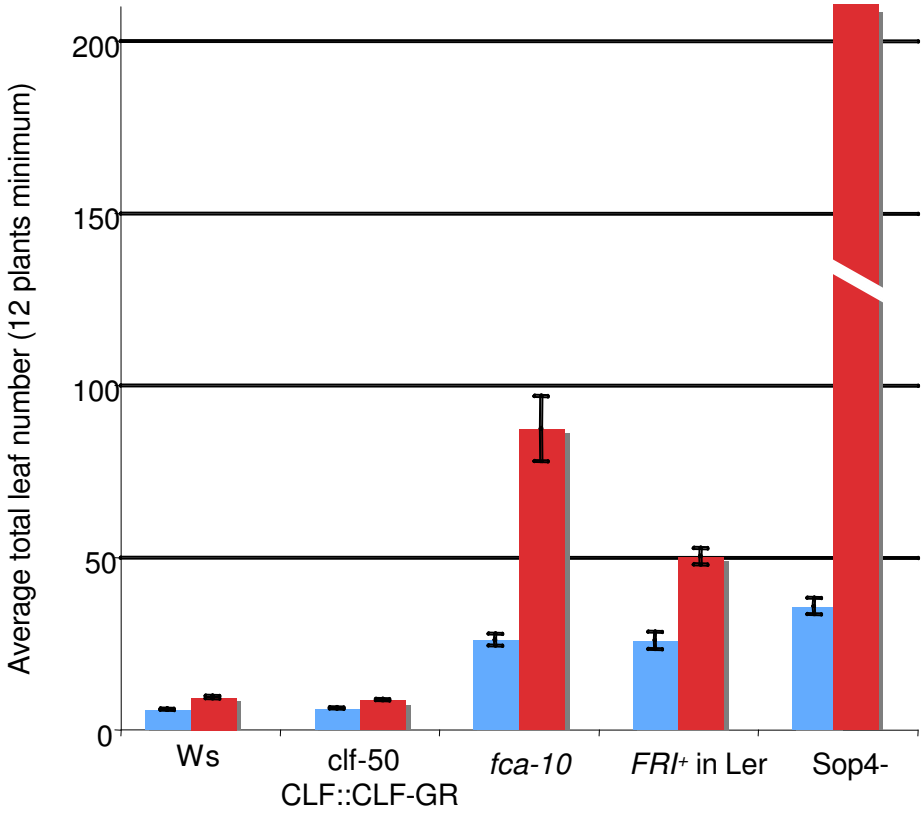

Figure 4. 6: Flowering time experiment under LD conditions with or without vernalization. The results show the average of resette leaves at bolting (12 plants minimum). Red: with vernalization (V). Blue: without vernalization (NV). Genotypes tested: Ws, *clf-50 pCLF:CLF-GR*, *fca-10*, *FRI+* in Ler and Sop4- (*clf-50 fpa-10*). Plants were vernalized for 4 weeks in the dark at 4˚C. Bars show Standard Error (SE).

Under these conditions, wild type (Ws), *fca-10* and a *FRI*$^{+}$ in *Landsberg erecta* background (Ler) were used as controls. Ws plants flowered after 5.67 ± 0.14 (V) and 9.25 ± 0.36 (NV); the *fca-10* mutant flowered after 26 ± 1.7 (V) and 87.4 ± 9.43 (NV) and *FRI*$^{+}$ in Ler flowered after 25.8 ± 2.5 (V) and 50.3 ± 2.4 (NV) rosette leaves. This result showed that the vernalization treatment was effective in considerably accelerating flowering, although the response was not saturated, as vernalized *FRI*$^{+}$ and *fca-10* were still somewhat late flowering (Figure 4. 6). The progenitor line, *clf-50*

pCLF:CLF-GR, flowered after 6.17 ± 0.17 and Sop4- (*fpa-10 clf-50*) plants after 35.8 ± 2.3 rosette leaves with vernalization. Without vernalization the time to flower for Sop4- plants increased exceptionally as over 240 rosette leaves were counted (still not flowered) compared with the progenitor line which only needed 8.55 ± 0.16 rosette leaves to flower.

From this experiment, it can be concluded that Sop4- plants have normal vernalization responses, as expected if the late flowering is caused by an *fpa* mutation and acts through *FLC*.

To molecularly confirm the conclusions related with the time to flowering experiment, the expression of *FLC* in 15 days old Sop4- plants was monitored by semi-quantitative RT-PCR (Figure 4. 7). The expression of *FLC* compared to ß-tubulin in the progenitor line, *clf-50 pCLF:CLF-GR*, is slightly higher (three times) than in Ws plants (Figure 4. 7-A). In Sop4- plants, *FLC*'s expression compared to ß-tubulin are 15 and 6 times higher than in Ws and *clf-50 pCLF:CLF-GR* respectively (Figure 4. 7-A).

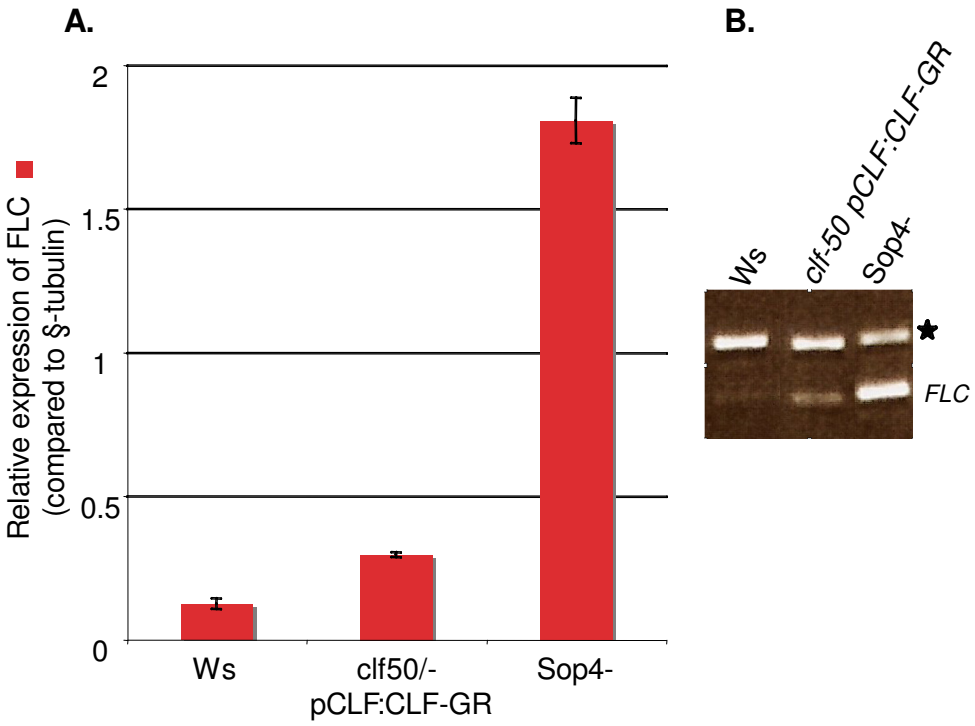

Figure 4. 7: Expression of *FLC* in Sop4- plants. A. Quantification of the transcript levels of *FLC* in Ws, *clf-50 pCLF:CLF-GR* and Sop4- plants. The quantification corresponds to the average of the ratio (target gene signal/ß-tubulin signal) of three different RT-PCR reactions. Bars show Standard Error (SE). B. Example of RT_PCR of *FLC* on Ws, *clf-50 pCLF:CLF-GR* and Sop4- cDNA samples. Duplex PCRs were performed to amplify ß-tubulin (marked with ★) as internal controls with the specific primers for *FLC*. The plants were grown on soil in LD conditions and the mRNA were extracted from 15 days old seedlings.

The semi-quantitative RT-PCRs results corroborate that there was an accumulation of *FLC*'s mRNA in Sop4- plants as shown in Figure 4. 7-A and B.

In the light of these results, it can be concluded that the late flowering phenotype observed for Sop4- plants is certainly due to high level of *FLC* in Sop4- plants.

4.2.3. *AG* is not a target of *FPA*

It has been shown that plants overexpressing *FPA* in SD conditions flower earlier than wild type and the *flc* null mutant suggesting that *FPA* can promote flowering in other ways than by its interaction with *FLC* (Schomburg et al., 2001). Also, parallels have been proposed between the post-transcriptional processing of *FLC* and *AG* (Simpson, 2004). For example, the 5'-structure similarities between the two genes at their unusually large 5' introns (*FLC* intron 1 is 3.5kb, *AG* intron 2 is 3.0kb) suggest that RNA-binding proteins like FPA may regulate *FLC* and *AG* by a post-transcriptional mechanism. As one of the factors responsible for the Clf- phenotype is the ectopic expression of *AG* (Goodrich et al., 1997) (chapter 1 section 1.3.2) and because in Sop4- plants the Clf- phenotype is suppressed, *AG* as a target of FPA was tested.

Firstly, the expression of *AG* by semi-quantitative RT-PCR was monitored in Ws, *clf-50 pCLF:CLF-GR* and Sop4- plants to check if *AG*'s expression remains high. Figure 4. 8-A shows that the level of expression of *AG* related to ß-tubulin in *clf-50 pCLF:CLF-GR* was approximately 20 times higher than in Ws plants. Also, the expression of *AG* compared to ß-tubulin in Sop4- plants was analogous to the expression detected in the progenitor line *clf-50 pCLF:CLF-GR* as shown in Figure 4. 8-A and -B. From this experiment it can be concluded that *AG* continues to be highly expressed in Sop4- plants and also that the *fpa-10* mutation does not change the level of expression of *AG* in Sop4- plants. The levels of the AG protein were assessed in Ws, *clf-50 pCLF:CLF-GR*, *ag* and Sop4- by western-blot. Total protein extracts were obtained from seedlings grown for ten days on soil under LD conditions and western-blot analysis was performed (section 2.6.1 and 2.6.2) using a polyclonal serum against the C-terminal of the AG protein provided by Prof. E. Meyerowitz and Dr. I. Ito (Figure 4. 9).

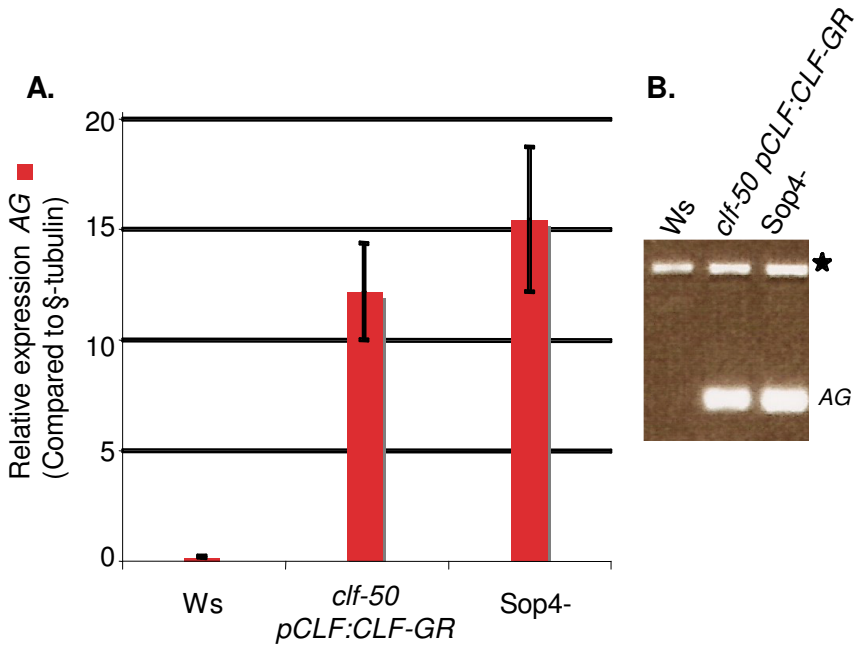

Figure 4. 8: Expression of *AG* in Sop4- plants. A: Quantification of the transcript levels of *AG* in Ws, *clf-50 pCLF:CLF-GR* and Sop4- plants. The quantification corresponds to the average of the ratio (target gene signal/ß-tubulin signal) of three different RT-PCR reactions. Bars show Standard Error (SE). B. Example of RT-PCR of *AG* on Ws, *clf-50 pCLF:CLF-GR* and Sop4- cDNA samples. Duplex PCRs were performed to amplify ß-tubulin (marked with ⋆) as internal control with the specific primers for *AG*. The plants were grown on soil in LD conditions and the mRNA were extracted from 15 days old seedling

Wild type (Ws) flowers buds and *ag* homozygous seedlings were used as positive and negative controls respectively (Figure 4. 9). In the positive control (flower buds from Ws plants), the strong band at 29kDa corresponds to the presence of the AG protein, which is not present in *ag* samples as expected. The samples corresponding to wild type Ws seedlings did not show the presence of the AG protein at 29 kDa confirming the low levels of expression of AG obtained in Figure 4. 8.

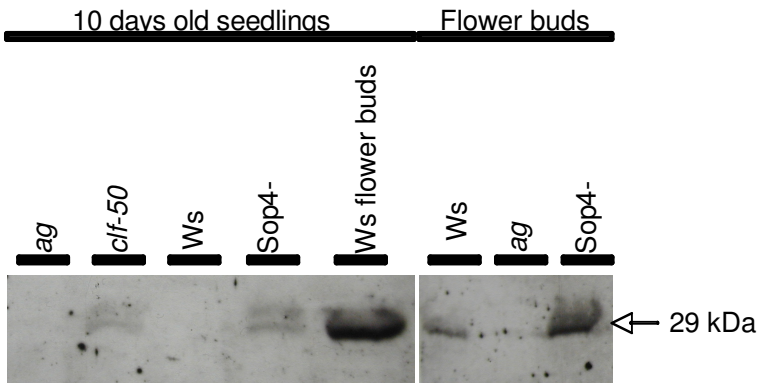

Figure 4. 9: Western-blot analysis of AG in *sop4* plants Total protein was extracted from 14 days old seedlings grown on soil on LD conditions. The AG protein is detected at 30 kDa by an antibody directed against the carboxy terminus of AG. For the detection, a rabbit serum was kindly provided by Prof. E. Meyerowitz and Dr. T. Ito.

For *clf-50 pCLF:CLF-GR* and Sop4- samples, the AG protein was detected at 29 kDa with approximately the same intensity as shown in Figure 4. 9. The presence of the AG protein at floral stages in Sop4- plants was checked by western-blot with samples from flower buds of Ws, *ag* (homozygous mutant) and Sop4- plants as shown in the right hand side of Figure 4. 9. This experiment show that the level of the AG protein in Sop4- plants was not changed compared with the progenitor line, *clf-50 pCLF:CLF-GR* and also that AG was strongly and ectopically expressed in Sop4- plants.

From this experiments it was concluded that the mRNA of *AG* was not a target of FPA for post-transcriptionally processing despite the similarities between *FLC* and *AG* transcripts.

4.2.4. *SHP2*, *AGL24*, *FUL* and *SEP3* are down regulated in Sop4- plants

It has been previously shown in this thesis (chapter 3 section 3.3) that *SEP3* is a direct target of CLF and that a loss of function mutation in *SEP3* suppresses the Clf- phenotype. Also, data was provided to support the fact that *SEP3* is a co-factor of *AG* and that they act together to activate downstream common targets as *SHP2*. Previously in this chapter, it was shown that *AG*'s mRNA is not a target of FPA as the AG protein is present in Sop4- seedlings (Section 4.2.3). However, down regulation in Sop4- plants of another known *CLF* target (*SEP3*) due to impairment in SEP3's mRNA processing could also explain the Sop4- phenotype. Moreover, recent studies propose and/or indicate that the association of floral MADS proteins (such as AG and SEP3) into ternary higher-order MADS complexes might be the principal mode of combinatorial control for floral-organ specification (Honma and Goto, 2001; Parenicova et al., 2003; Krizek and Fletcher, 2005). For example, ternary MADS complexes AP3–PI–AP1 and AP3–PI–SEP3 are sufficient to activate the *AP3* promoter whereas the AP3-PI heterodimer binds to the *AP3* promoter but are unable to activate transcription (Honma and Goto, 2001).

Also, protein-protein yeast two-hybrid "interactome" of all members of the *A. thaliana* MADS-box transcription factor family has been published. This study focus in two sub-networks, one for the flower induction and one for the flower organ formation pathway, which seem to be linked and in which the usual suspects *AG* and *SEP3* can be included (de Folter et al., 2005).

In the light of all these elements, it was decided to test the expression level of *SEP3* and also putative MADS-box transcription co-factors of *AG* and *SEP3* in Sop4- plants (Figure 4.10). The hypothesis was that in Sop4- plants a member of a MADS complex is inactive, so that the Clf- phenotype is suppressed because AG is no

longer active although present in leaves of Sop4- plants. The physical interaction in yeast two-hybrid with either *AG* and/or *SEP3* and the existing interaction in other species with *AG* and/or *SEP3* were the criteria used to select the targets to test. Among the 100 MADS box transcription factors analyzed (de Folter et al., 2005), only eight MADS box transcription factors seem to be able to interact with *AG* and/or *SEP3*: *AGAMOUS LIKE 6* (*AGL6*); *SHATTERPROF 1* and *2* (*SHP1*); (*SHP2*); *SEPALLATA 1, 2* and *4* (*SEP1,2,4)*; *FRUITFUL* (*FUL)* and *AGAMOUS LIKE 24* (*AGL24)* (de Folter et al., 2005). In Figure 4.10, the expression levels of *SEP3*, *AGL6*, *SHP1*, *SHP2*, *SEP1*, *SEP2*, *SEP4*, *FUL* and *AGL24* in 15 days old Ws, *clf-50 pCLF:CLF-GR* and Sop4- seedlings was monitored.

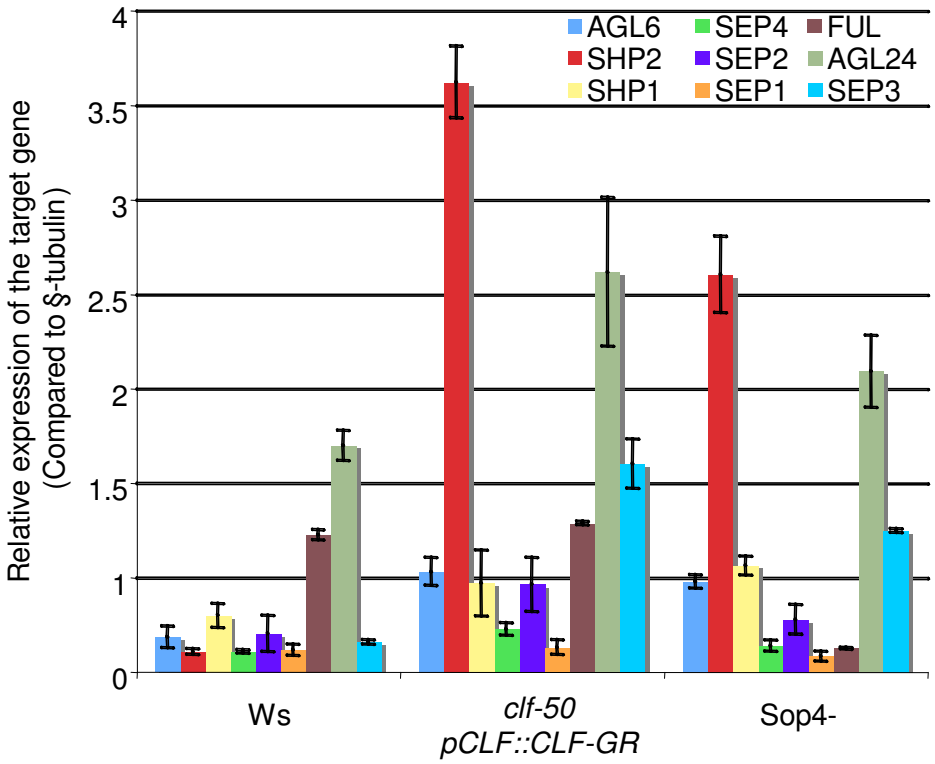

Figure 4.10: Expression of MADS-box transcription factors *AGL6*; *SHP2*; *SHP1*; *SEP4*; *SEP2*; *SEP1*; *FUL*; *AGL24* and *SEP3* in Sop4- plants. The quantification corresponds to the average ratio (target gene [*AGL6*; *SHP2*; *SHP1*; *SEP4*; *SEP2*; *SEP1*; *FUL*; *AGL24* and *SEP3*] signal/ß-tubulin signal) of three different RT-PCR reactions. Bars show Standard Error (SE). The plants analyzed were grown on plates and the mRNA was extracted from 15 days old seedlings.

All the expression levels correspond to the relative expression compared with ß-tubulin in the same sample. Three different pattern of expression can be seen: (i) Low expression in Ws plants then the expression is higher and similar in *clf-50 pCLF:CLF-GR* and Sop4- plants (*AGL6* and *SHP1*). (ii) Low expression in Ws plants then the expression is very high in *clf-50 pCLF:CLF-GR* but reduced in Sop4- plants (*SHP2, SEP2, SEP3, SEP4* and *AGL24*). (iii) Stable expression in Ws and *clf-50 pCLF:CLF-GR* plants but the expression weakens in Sop4- plants (*SEP1* and *FUL*)

(Figure 4.10). If FPA was not able to correctly process the mRNA of one of these genes, the expected result was a significant mis-regulation of the transcript in Sop4- plants compared to *clf-50 pCLF:CLF-GR* plants. Accordingly with this, the most interesting MADS box transcription factors was those showing patterns (ii) and (iii): *SHP2*, *SEP1*, *SEP2*, *SEP3*, *SEP4*, *AGL24* and *FUL*. Indeed, in this two patterns there was a difference in expression between the progenitor line *clf-50 pCLF:CLF-GR* and Sop4- plants. Among these genes, the ones exhibiting a considerable mis-regulation (reduction) in Sop4- plants were: *AGL24*, *SHP2*, *SEP3* and *FUL* (Figure 4.10).

The expression in Sop4- plants of *AGL24* compared to the progenitor line was reduced by approximately 20%. This decrease did not seem to be significant enough to suppress the clf phenotype, as *AGL24* expression was already high in Ws seedlings (Figure 4.10). Concerning the expression of *SHP2* and *SEP3* in Sop4- plants, their expression was reduced by 33% compared to the *clf-50 pCLF:CLF-GR* progenitor line. *SHP2* mis-regulation in *clf-50 pCLF:CLF-GR* has the biggest amplitude and the decrease rate (33%) match with the decrease observed in Sop1- (*clf-50 sep3-7*) plants (chapter 3 section 3.2.4). It was assumed that the reduction of the expression of *SHP2* in Sop4- plants might be related with the decrease of the expression of *SEP3*. These results suggest that *SEP3* may be a target of *FPA*. Finally, *FUL* is the MADS box transcription factor which expression in Sop4- plants compared to the *clf-50 pCLF:CLF-GR* progenitor line was reduced the most: 85% (Figure 4.10). It has been shown that a loss of function mutation in *FUL* suppresses the curled leaves and the early flowering phenotype observed in transgenic lines overexpressing *FT* (Teper-Bamnolker and Samach, 2005). Therefore, this result suggested that *FUL* might be also a target of *FPA*.

From this experiment, it can be concluded that *SHP2*, *FUL* and *SEP3* are the most probable targets of FPA. *AGL24* does not seem to be a target of FPA in the conditions described in this thesis. Therefore, the deficient processing of the mRNA

of *SHP2* and/or *FUL* and/or *SEP3* by FPA could be the source for the Sop4- phenotype. Alternatively, high *FLC* levels in Sop4- plants may lessen the Clf- phenotype. In that case the effect of *fpa-10* may be indirect.

4.3. High levels of *FLC* partially suppress the *Clf-* phenotype

It was shown that loss of FT activity suppressed the Clf- phenotype (chapter 3 section 3.2.5). And, it is known that FLC represses *FT* expression (Searle et al., 2006; Li et al., 2008).Then, a loss of function mutations of *FLC* enhances the Clf- phenotype. Conversely, the over-expression of FLC suppresses the Clf- phenotype. The double mutant *clf-52 FRI*$^{+}$ has theoretically the same effect on *FLC* than the double mutant *fpa-10 clf-50* (*sop4*): the expression of FLC is highly mis-expressed. In the following section, the suppression of the Clf- phenotype in *fpa-10 clf-50* (*sop4)* and in *clf-52 FRI*$^{+}$ plants was studied. Leaf curling, leaf area and the time to flower were analyzed in both backgrounds compared with the progenitor lines *clf-50 pCLF:CLF-GR*, *fpa-6*, *FRI*$^{+}$ and wild type Ws plants. Finally, the first steps to obtain the triple mutant *flc-3 fpa-7 clf-28* are presented here.

4.3.1. The clf-52 FRI^{+} line

The FRIGIDA active allele, *FRI*$^{+}$, has been shown to increase *FLC* mRNA levels (Michaels and Amasino, 1999). Therefore, *FRI*$^{+}$ plants have a late flowering phenotype due to a high level of active *FLC*. The *clf-52* mutation was isolated in Dr. R. Amasino's laboratory in a genetic screen for suppressors of the late flowering phenotype of *FRI*$^{+}$ in the Ws background. The *clf-52* allele is a loss of function mutation due to an insertion of a T-DNA in *CLF*'s open reading frame. In Figure 4. 11, the suppression of the Clf- phenotype in the double mutant *clf-52 FRI*$^{+}$ is shown. The characteristic curled leaves of Clf- plants as in the progenitor line used in this

study, *clf-50 pCLF:CLF-GR*, disappeared. Instead, the *clf-52 FRI*$^{+}$ line exhibit flat leaves and also, the early flowering phenotype of Clf- plants is inhibited, as shown in Figure 4. 11.

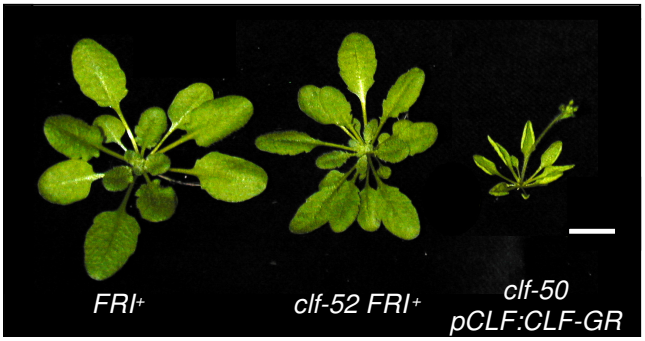

Figure 4. 11: The *clf-52* allele suppressed the *FRI*$^{+}$ late flowering phenotype: Phenotype of the double mutant *clf-52 FRI*$^{+}$. Plants were grown during four weeks in LD conditions. Scale bar = 1 cm.

Accordingly with the known function of active alleles of FRI, it was supposed that in *clf-52 FRI*$^{+}$ plants the Clf- phenotype due to the *clf-52* mutation is suppressed because the levels of *FLC* are very high. Dr. R. Amasino kindly gave this line to us to study the level of suppression of the Clf- phenotype.

4.3.2. *FRI*$^{+}$ and *fpa-10* suppress the leaf phenotype of Clf- plants

In order to compare the leaf curling suppression in the two lines *clf-52 FRI*$^{+}$ and *clf-50 fpa-10*, the leaf phenotype: size and flatness was qualitatively and quantitatively assessed. The single mutant *clf-52* was not available for this analysis but as it is a loss of function mutation it was assumed that its phenotype was similar to the phenotype observed for the progenitor line *clf-50 pCLF:CLF-GR*. In Figure 4. 12-A, the leaf phenotype of *clf-52 FRI*$^{+}$, *clf-50 fpa-10*, *fpa-6*, *FRI*$^{+}$, the progenitor line *clf-50 pCLF:CLF-GR* and wild type Ws plants can be observed.

clf-52 FRI^{+} and *clf-50 fpa-10* seem to have slightly smaller leaves than wild type Ws, *fpa-6* and *FRI*$^{+}$. Additionally, the progenitor line *clf-50 pCLF:CLF-GR* has the smallest leaves of all the six genotypes observed (Figure 4. 12-A).

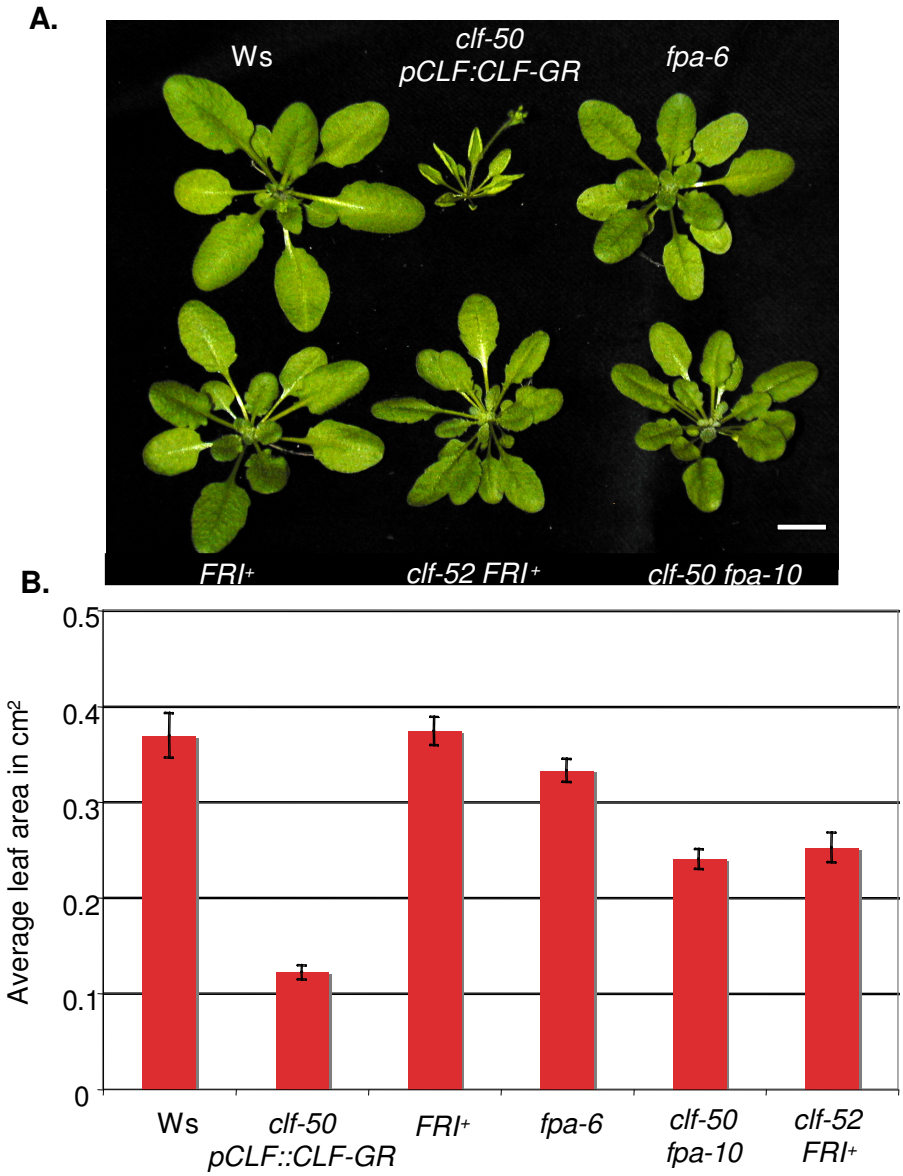

Figure 4. 12: Suppression of the Clf- phenotype by *FRI*+ and *fpa-10*. A. FRI+ and fpa-10 mutations suppress the Clf- phenotype. The phenotype of the lines *clf-52 FRI*+ and *clf-50 fpa-10* is shown. Plants were grown during four weeks in LD conditions. White bar = 1 cm. B. Quantification of the leaves are of *clf-52 FRI*+ and *clf-50 fpa-10* compared with *FRI*+, *fpa-10*, the progenitor line *clf-50 pCLF:CLF-GR* and wild type Ws plants. The five biggest leaves of three different plants were dissected and the area calculated using ImageTool (IT) software (chapter 2 section 2.3.4). Bars show Standard Error (SE).

Concerning the flatness of the leaf, both lines *clf-52 FRI*$^{+}$ and *clf-50 fpa-10* show a total reduction of the Clf- phenotype. The reduction of the leaf area in both lines *clf-52 FRI*$^{+}$ and *clf-50 fpa-10* was quantified and compared between the two genotypes. In Figure 4. 12-B, quantification for the leaf area of *clf-52 FRI*$^{+}$, *clf-50 fpa-10*,*fpa-6*, *FRI*$^{+}$, the progenitor line *clf-50 pCLF:CLF-GR* and wild type Ws plants is shown. *clf-50 pCLF:CLF-GR* plants showed the lowest leaf area among the genotypes tested: 0.12 ± 0.01 cm^2. The leaf area for Wild type Ws, *FRI*$^{+}$ and *fpa-6* was very similar: 0.35 ± 0.02 cm^2. This area was 1.4 times bigger than the leaf area of both lines *clf-52 FRI*$^{+}$ and *clf-50 fpa-10* which have the same leaf area: 0.25 ± 0.02 cm^2 (Figure 4. 12-B).

This experiment showed that in, *clf-52 FRI*$^{+}$ and *clf-50 fpa-10*, the leaf curling phenotype is suppressed. Also, the leaf flatness phenotype seemed to be induced similarly in both genotypes, as the results are very alike.

4.3.3. The suppression of the early flowering phenotype of Clf- plants by *FRI*$^{+}$ and *fpa*-10 is different

The flowering time phenotype of loss of function mutations in *FPA* and *FRI* has been described extensively in the literature. *fpa-6* was characterized as a late flowering mutation in the Ws background due to an up-regulation of *FLC* (Michaels and Amasino, 2001; Schomburg et al., 2001). Concerning *FRI*, it has been shown that over-expression of *FLC* delays flowering in a *FRI*$^{-}$ background (Michaels and Amasino, 1999). Also, *FRI*$^{+}$ plants do not have a late flowering phenotype in the absence of *FLC* (Michaels and Amasino, 2001). Then, it appeared that the main function of *FRI* is to up-regulate *FLC*, which then delayed flowering. In order to know how *FPA* and *FRI* regulate the flowering time phenotype in the *clf* background, a time to flower experiment in LD conditions between wild type Ws and the progenitor line *clf-50 pCLF:CLF-GR*; the single mutants *fpa-6* and *FRI*$^{+}$ in the

Ws background and the *clf-52 FRI*+ and *clf-50 fpa-10* lines was performed (Figure 4. 13).

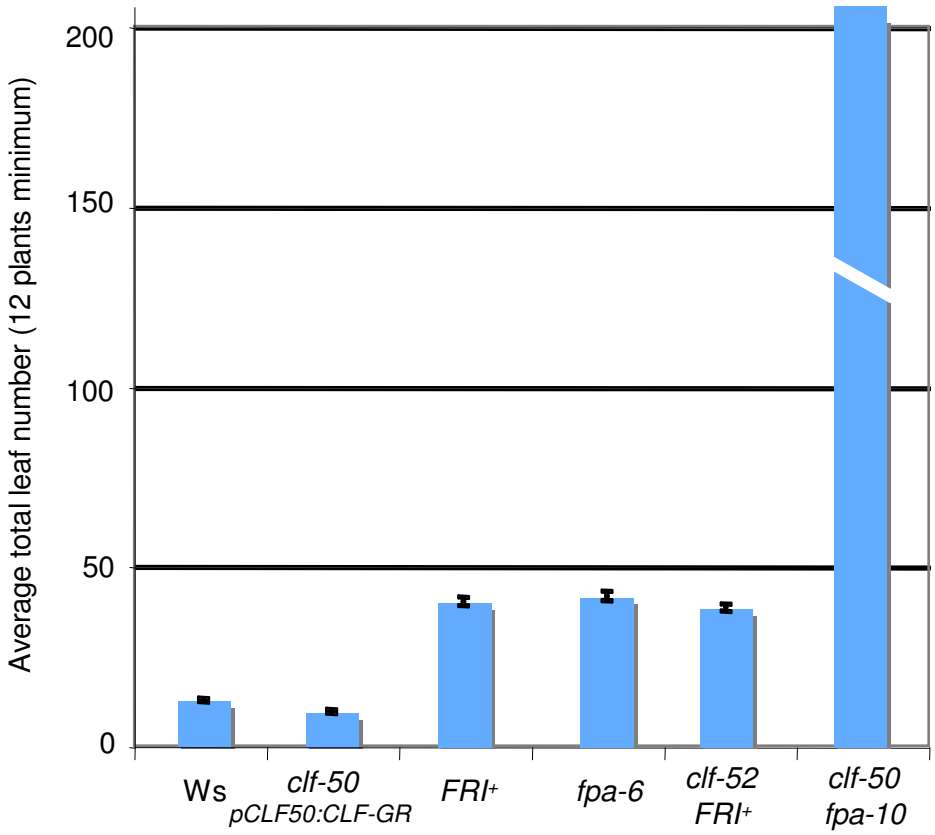

Figure 4. 13: Flowering time experiment under LD conditions. The results show the average of rosette leaves at bolting (12 plants minimum) in LD conditions. Genotypes tested: Ws, *clf-50 pCLF:CLF-GR*, *clf-52 FRI*+ and *clf-50 fpa-10*. Bars show Standard Error (SE).

Under LD conditions, wild type control plants (Ws) flowered after 12.7 ± 0.5 rosette leaves. The progenitor line, *clf-50 pCLF:CLF-GR*, flowered after 9.6 ± 0.5 rosette leaves. As expected, the progenitor line *clf-50 pCLF:CLF-GR* flowered earlier than wild type plants (Ws) in LD conditions. Single mutants *FRI*+ and *fpa-6* plants flowered at approximately the same time after 40.2 ± 1.2 and 41.7 ± 1.3 rosette leaves respectively. *clf-52 FRI*+ flowered slightly earlier than the single mutant *fpa-6* and the

FRI+ line after 38.5 ±0.9 rosette leaves in average. The double mutant *clf-50 fpa-10* did not flower after 200 (still counting), rosette leaves (Figure 4. 13).

We notice that in our conditions the suppression of the late flowering phenotype of *FRI*+ plants by a mutation in *CLF* is very weak. Also, I did not do this experiment in presence and absence of dex then I could not compare the two mutations *fpa-6* and *fpa-10* concerning the flowering time.

Nevertheless, this experiment showed that the single mutations *fpa-10* and *FRI*+ suppress differently the early flowering phenotype induced by the *clf* mutation. Indeed, the time to flower of the double mutant *clf-50 fpa-10* under LD conditions is at least four times longer than the time to flower needed by the *clf-52 FRI*+ line (Figure 4. 13). As segregation test showed that *fpa-10* is the only mutation induced by the activation tagging approach, these results suggest that the suppression of the Clf- phenotype by the *fpa-10* mutation involves additional factors other than *FLC* certainly regulated by FPA.

4.3.4. Obtaining the triple mutant *clf-28 fpa-7 flc-3*

To improve understanding of the suppressing features of Sop4- plants, I decided to make the triple mutant *clf-28 fpa-7 flc*-3. Indeed, in this triple mutant the loss of function mutation of *FLC* should allow us to clearly differentiate which trait of the Clf- phenotype the *fpa-10* mutation suppress through a high mis-regulation of *FLC* and more interestingly which other features are suppressed by an additional factor. To this end, the double mutant *fpa-7 flc-3* (kindly provided by Dr. G. Simpson) was crossed with the double mutant *clf-28 flc-3* available in the laboratory. In the F1 generation 36 wild type plants with an early flowering phenotype were obtained. The early flowering phenotype was caused by the homozygous *flc-3* mutation. The F2 progeny of two independent F1 plants were grown on soil and genomic DNA

was extracted from 24 independent plants. Genotyping PCRs for the *clf-28* and *fpa-7* single mutations were then performed (Figure 4. 14).

The 24 selected plants exhibited either the Clf- (-) or a suppressed phenotype Sop4- (+). As expected plants in lanes 2, 3, 7, 9, 11, 12 and 19-21 displaying the Clf- phenotype (-) were homozygous for the *clf-28* mutation and suppressed plants (+) in lanes 1, 4-6, 8, 10, 14-18, 22 and 23 were either heterozygous *clf-28/CLF* or homozygous *CLF* (Figure 4. 14).

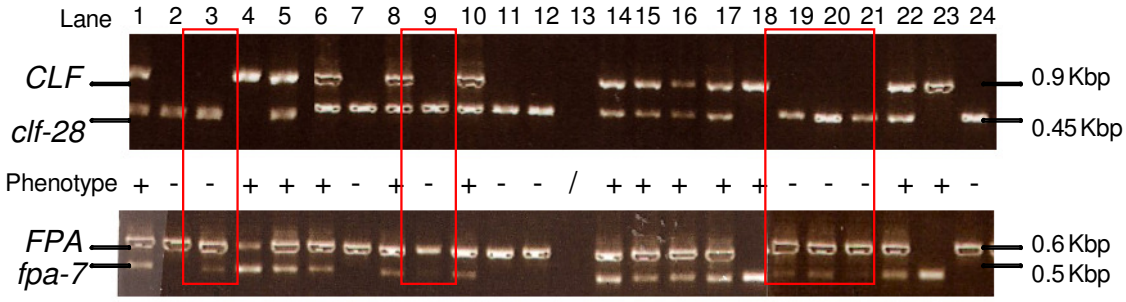

Figure 4. 14: Genotyping PCR for *clf-28* and *fpa-7* single mutations. The genotyping PCR where conducted on 24 F2 plants. Genotyping primer sets (clf28F, clf28R and LbB1.3) and (fpa 7 F, fpa 7 R and fpa 7 tdna) were used to identify homozygous *clf-28 fpa-7* plants. Homozygous *clf-28* plants produced a single band of 450 bp in size, whereas wild type plants produced a band of 900 bp in size. Homozygous *fpa-7* plants produced a single band of 500 bp in size, whereas wild type plants produced a band of 600 bp in size. The phenotype was scored as follow: Clf- plants (-) and Sop4- plants (+).

Five independent plants showed in red boxes (lanes 3, 9, 19-21) in Figure 4. 14 were homozygous *flc-3 clf-28* and heterozygous *fpa-7/FPA*. The F3 seeds of these plants

have been currently harvested and put on soil in order to find the triple mutant *clf-28 fpa-7 flc-3*.

From this experiment it was concluded that the triple mutant clf-28 fpa-7 flc-3 would be available soon for further analysis in the laboratory. Unfortunately I was not able due to constraints on time to complete the triple mutant.

4.4. Summary

In this chapter the mutation most likely to be responsible for the suppressed phenotype of Sop4- plants was identified. It is a loss of function mutation in the *FPA* locus of *A. thaliana*. The new allele found was called *fpa-10*. The FPA protein has a RNA binding domain and is a member of the autonomous flowering pathway. FPA main known function is to repress *FLC* expression. Therefore, *fpa* plants have a late flowering phenotype. Here, it was shown that the *fpa-10* mutation is a null allele of *FPA,* as the FPA protein was not detected in Sop4- plants. In addition, the double mutant *clf-50 fpa-10* has a very late flowering phenotype due to a high mis-expression of *FLC*. Concerning the loss of the curling in leaves, the AG protein still expressed at very high levels suggesting that AG is not post-transcriptionally regulated by FPA. We provide expression analysis suggesting that either *SHP2* or *SEP3* or *FUL* mRNAs could be potential targets of FPA. We compared the suppression of the Clf- phenotype by *fpa-10* and FRI^{+}. In the *clf* mutant background, both *fpa-10* and FRI^{+} alleles suppress the leaf curling phenotype characteristic of Clf- plants. But, only the *fpa-10* mutation was able to induce a very late flowering phenotype in LD conditions. Finally, the triple mutant *clf-28 fpa-7 flc-3* was made in the Col-0 background available in the laboratory for further analysis of the suppression of the Clf- phenotype by the *fpa-7* mutation.

5. Characterization of *suppressor of polycomb 2* and *3* (*sop2* and *sop3*)

The *suppressor of polycomb 2* (*sop2*) and *suppressor of polycomb 3* (*sop3*) mutations were identified in the same screen as *sop1* and *sop4*. The suppression of the Clf- phenotype by *sop2* and *sop3* mutations is very strong and both mutants are late flowering.

In this chapter, the suppressed phenotype observed in Sop2- and Sop3- plants is described. In addition, expression analysis of *FLC* and *AG* in the mutant background was done to characterize these mutants at the molecular level. Also, using a plasmid rescue strategy, an attempt to identify the T-DNA insertions responsible for the suppressed phenotypes was conducted.

5.1. The phenotype of Sop2- and Sop3- plants

The mutations, *sop2* and *sop3*, suppressing the Clf- phenotype showed a very late flowering phenotype on soil. In order to identify the pathways in which *sop2* and *sop3* mutations were involved, a time to flower experiment was performed and expression analysis by semi-quantitative RT-PCR on known targets of CLF such as *FLC* and *AG* was performed.

5.1.1. *sop2* and *sop3* efficiently suppress the Clf- phenotype

The two mutants *sop2* and *sop3* showed a very strong suppression of the Clf- phenotype where the leaves are totally flat and well expanded. In addition, strongly serrated leaves were observed for the two mutants (Figure 5.1). At late vegetative stages of development both mutants exhibit a great number of leaves due to the late flowering phenotype (Figure 5.1).

Figure 5.1: Phenotype of Sop2- and Sop3- plants. Comparison between the progenitor line *clf-50 pCL:CLF-GR*, Sop2- and Sop3- plants. Plants were grown during four weeks on soil in LD conditions. Scale bar = 1cm.

From Figure 5.1 it was concluded that the two mutants efficiently and strongly suppress the Clf- phenotype. Also the two mutations, *sop2* and *sop3*, seem to equally have a serration phenotype.

5.1.2. Sop2- and Sop3- plants respond differently to vernalization treatments

Sop2- and Sop3- plants show many similarities in different traits such as the suppression of the Clf- phenotype, the leaf serration and the late flowering phenotype. In order to better characterize these two mutants, a time to flower experiment under LD conditions with or without vernalization was performed to quantify the delay in flowering for each of the mutants (Figure 5.2). Under these conditions, wild type (Ws), *fca-10*, *flc-5* and *FRI*$^{+}$ in Ler were used as controls. Ws plants flowered after 11.8 ± 0.3 and 9.9 ± 0.4; the *fca-10* mutant flowered after 51.7 ± 1.6 and 21 ± 0.9 and *FRI*$^{+}$ in Ler flowered after 47.9 ± 0.5 and 24.9 ± 0.3 rosette leaves without and with vernalization respectively. Moreover, the *flc-5* mutant flowered

after 8 ± 0.2 rosette leaves without vernalization and 9.2 ± 0.2 rosette leaves with vernalization.

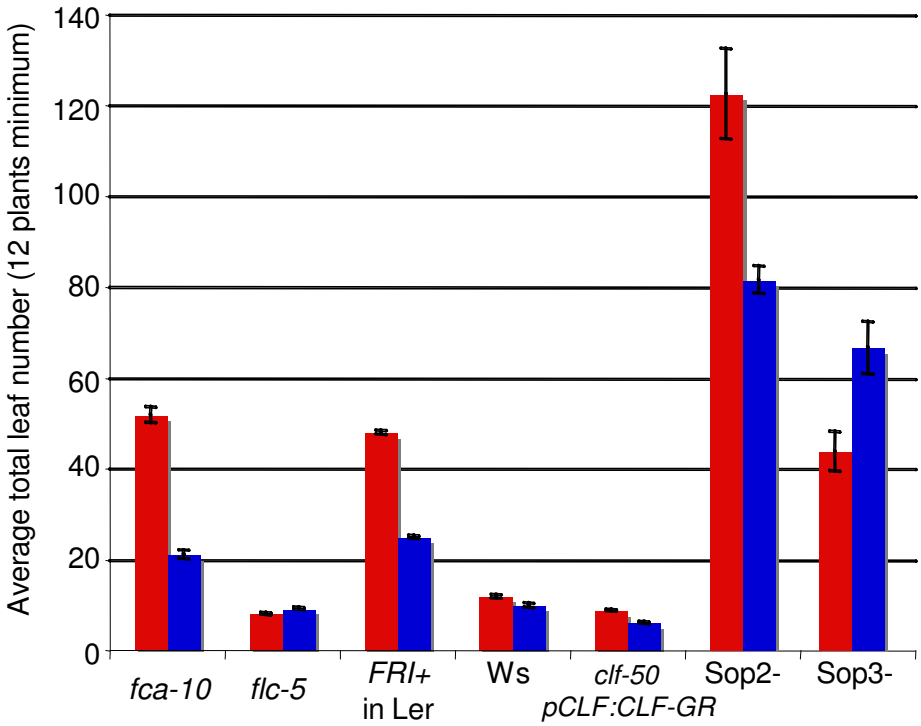

Figure 5.2: Flowering time experiment under LD conditions with or without vernalization. The results show the average of rosette leaves at bolting (12 plants minimum). In red: without vernalization. In blue: with vernalization. Genotypes tested: *fca-10*, *flc-5*, *FRI*$^{+}$ in Ler, Ws, *clf-50 pCLF:CLF-GR*, Sop2- and Sop3-. Plants were vernalized for four weeks in the dark at 4˚C. Bars show standard Error (SE).

This result showed that the vernalization treatment in this experiment was effective in accelerating flowering time of *fca-10* and *FRI*$^{+}$, two mutants known to be vernalization responsive (Figure 5.2). However, the treatment was not saturating, as both mutants remained later flowering than wild type (Ws) even after vernalization treatment (Figure 5.2).

The progenitor line, *clf-50 pCLF:CLF-GR*, flowered after 8.7 ± 0.2 and 6 ± 0.2 whereas Sop2- plants flowered after 122.5 ± 10 and 81.5 ± 3 and Sop3- plants flowered after 43.7 ± 4.3 and 66.5 ± 5.8 rosette leaves without and with vernalization respectively. For Sop2- plants, a delayed time to flower higher than wild type, *fca-10* and FRI^{+} in Ler controls without or with vernalization was observed. Nonetheless, Sop2- plants respond to the vernalization treatment. Sop3- plants did not react to the vernalization treatment as expected. Indeed, these plants flowered later (66.5 ± 5.8 rosette leaves) with vernalization than without (43.7 ± 4.3 rosette leaves) whereas the opposite effect is typically observed.

From this experiment, it was concluded that *sop2* and *sop3* are likely to be mutations at different loci, which mainly differ in their effect on the flowering time phenotype in the *clf* mutant background. Moreover, Sop2- plants were late flowering but responsive to vernalization treatments, which suggested that the suppressor mutation *sop2* has not occurred in any member of the vernalization pathway. Finally, unlike *sop2*, the *sop3* mutation appears to affect an element in the vernalization pathway. It is important to remember that a mutant like *constans* (*co*), will exhibit a very late flowering phenotype and will not respond to a vernalization treatment.

At this point, it seems that the *sop3* mutation impairs either the photoperiod and/or the vernalization pathway.

It is known that vernalization promotes flowering mainly by repressing *FLC* expression (Sheldon et al., 1999). *VRN2* is a key regulator of the vernalization response and a member of the PRC2 in *A. thaliana* (chapter 1 section 1.2.2). *vrn2* mutants are vernalization insensitive thus, *vrn2* mutants showed little or no vernalization response. The VRN2 activity is essential for the stable repression of

FLC as in the *vrn2-1* mutant it was shown that *FLC* is highly mis-expressed (Gendall et al., 2001). In Figure 5.2, it was shown that both Sop2- and Sop3- plants flowered very late and that *sop3* plants are insensitive to vernalization treatments. As both mutants are late flowering, the expression of *FLC* was tested (up-regulated) compared to the progenitor line, *clf-50 pCLF:CLF-GR*, in Sop2- and Sop3- plants (Figure 5.3).

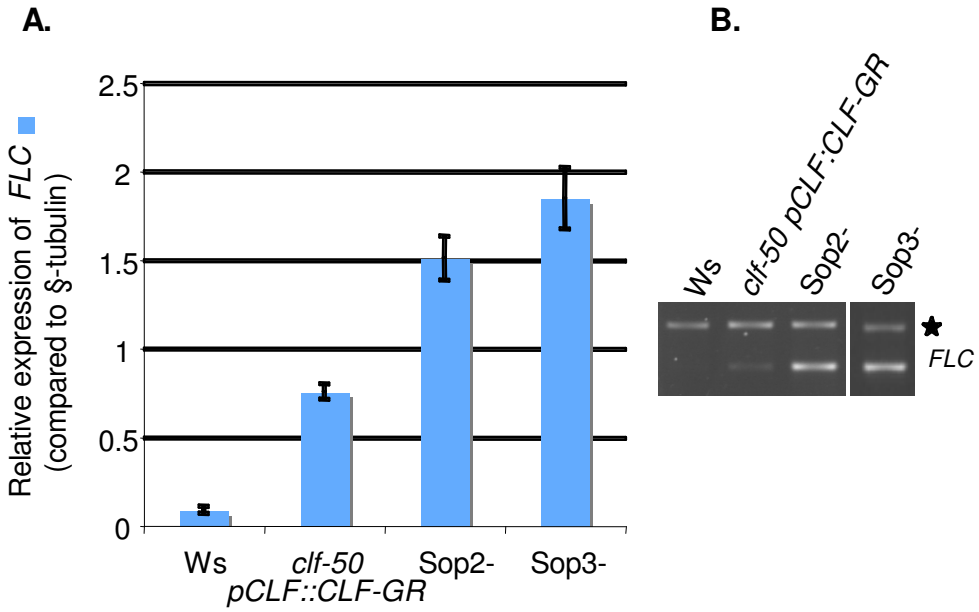

Figure 5.3: Expression analysis of *FLC* in Sop2- and Sop3- plants by semi-quantitative RT-PCR. A: Quantification of the transcript levels of *FLC* in Ws, *clf-50 pCLF:CLF-GR*, Sop2- and Sop3- plants. The quantification corresponds to the average ratio (target gene signal/ß-tubulin signal) of 3 different RT-PCR reactions. Bars show Standard Error (SE). B: Example of RT-PCR of *FLC* on Ws, *clf-50 pCLF:CLF-GR*, Sop2- and Sop3- cDNA samples. Duplex PCRs were performed to amplify ß-tubulin (marked with ⋆) as internal control with specific primers for *FLC*. The plants were grown on soil in LD conditions and mRNA was extracted from 15 days old seedlings.

The expression of *FLC* compared to ß-tubulin in the progenitor line, *clf-50 pCLF:CLF-GR*, is eight fold higher than in Ws plants (Figure 5.3). In Sop2- and Sop3- plants, *FLC*'s expression compared to ß-tubulin is 2 and 2.5 times higher than in *clf-50 pCLF:CLF-GR* respectively (Figure 5.3). The *FLC* expression analysis in Sop2- and Sop3- plants confirmed that the late flowering phenotype of these mutants is linked with a high mis-regulation of *FLC*.

5.1.3. The expression of *AG* is different in Sop2- and Sop3- plants compared with the progenitor line *clf-50 pCLF:CLF-GR*

As indicated previously in chapter 3 (section 3.2.5) and in chapter 4 (section 4.3), the up-regulation of *FLC* is able to suppress the Clf- phenotype. Also, it is known that high mis-expression of *AG* is closely linked with the leaf curling of Clf- plants as described in chapter 3 (section 3.2.3) and in the literature (Goodrich et al., 1997). The expression of *AG* in Sop2- and Sop3- plants was monitored to check if *AG*'s expression is still highly mis-regulated in the double mutant backgrounds.

The expression of *AG* was assessed by semi-quantitative RT-PCR in Ws, *clf-50 pCLF:CLF-GR*, Sop2- and Sop3- seedlings (Figure 5.4). The level of expression of *AG* related to ß-tubulin in *clf-50 pCLF:CLF-GR* is approximately 25 times higher than in Ws plants. The expression of *AG* compared to ß-tubulin in Sop2- seedlings is about twice the level in the progenitor line. Concerning Sop3- seedlings, the expression of *AG* compared to ß-tubulin is down regulated by half, compared with the progenitor line *clf-50 pCLF:CLF-GR* (Figure 5.4). Nonetheless, the level of expression of *AG* in Sop3- plants was approximately ten times higher than in Ws plants. This level of expression of *AG* in Sop3- plants might be high enough to induce some leaf curling, as normally there is no *AG* expression in leaves.

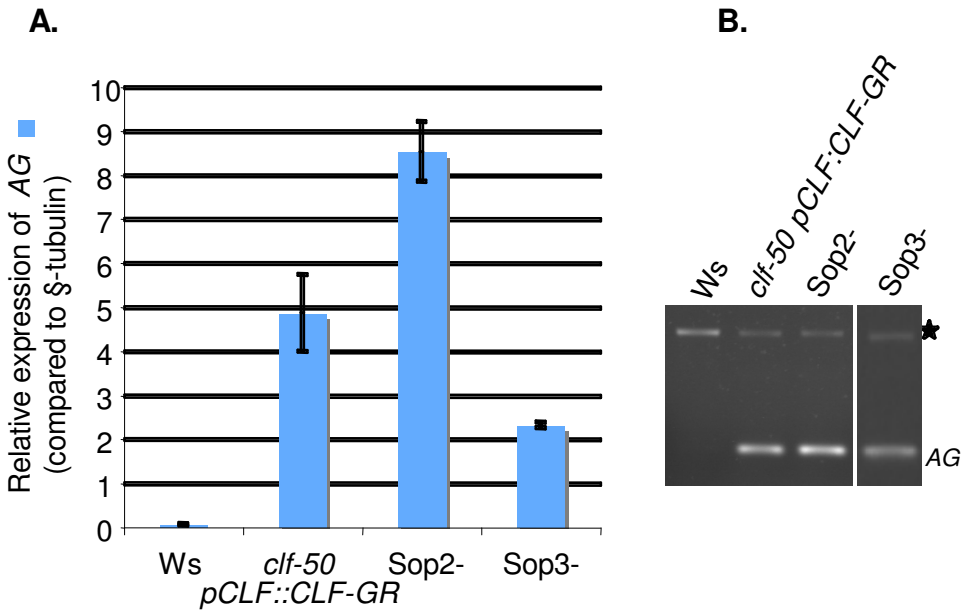

Figure 5.4: Expression analysis of *AG* in Sop2- and Sop3- plants by semi-quantitative RT-PCR. A: Quantification of the transcript levels of *AG* in Ws, *clf-50 pCLF:CLF-GR*, Sop2- and Sop3- plants. The quantification corresponds to the average ratio (target gene signal/ß0tubulin signal) of three different RT-PCR reactions. Bars show Standard Error (SE). B: Example of RT-PCR of *AG* on Ws, *clf-50 pCLF:CLF-GR*, Sop2- and Sop3- cDNA samples. Duplex PCRs were performed to amplify ß-tubulin (marked with ⋆) as internal control with specific primers for *AG*. The plants were grown on soil in LD conditions and mRNA was extracted from 15 days old seedlings.

From this experiment it was concluded that in Sop2- and Sop3- plants, the expression of *AG* is mis-regulated. Thus, the previous results, obtained for Sop2- and Sop3- in section 5.1.2, indicating that the suppression of the Clf- phenotype in Sop2- and Sop3- plants involves different pathways were molecularly confirmed.

5.1.4. The time to flower phenotype of Sop3- plants becomes more severe between generations

During the characterization of Sop3- plants, it was observed that Sop3- plants from the T4 generation flowered later than Sop3- plants from the previous generation (T3). This observation was quantified by assessing at the same time the time to flower of Sop3- plants from different generations (T3, T4 and T5). In addition, the expression of *FLC* for the different generations of Sop3- plants was analyzed. The progeny of all the Sop3- lines used in this experiment exhibited only the Sop3- phenotype suggesting that all the lines studied were homozygous for the mutation responsible of the Sop3- phenotype. In Long Days conditions (LD), the controls plants, wild type Ws, flowered after 14.4 ± 0.3 and the progenitor line, *clf-50 pCLF:CLF-GR*, flowered after 9.25 ± 0.3 rosette leaves (Figure 5.5-A). Sop3- plants from the T3 generation flowered after 33.4 ± 1.6; the T4 generation flowered after 50.1 ± 3.9 and the T5 generation flowered after 128.8 ± 9.9 rosette leaves (Figure 5.5-A). This experiment showed that there was an increase of the time to flower for Sop3- plants between generations in long day conditions without vernalization.

As previously shown, the expression of *FLC* in Sop3- plants was mis-regulated (see section 5.1.2, Figure 5.3). In order to linked the increase of the time to flower in different generations to an increase of the expression of *FLC* between generations, its expression was assessed by semi-quantitative RT-PCR in wild type Ws, the progenitor line *clf-50 pCLF:CLF-GR*, Sop3-T3, Sop3-T4 and Sop3-T5 seedlings (Figure 5.5-B).

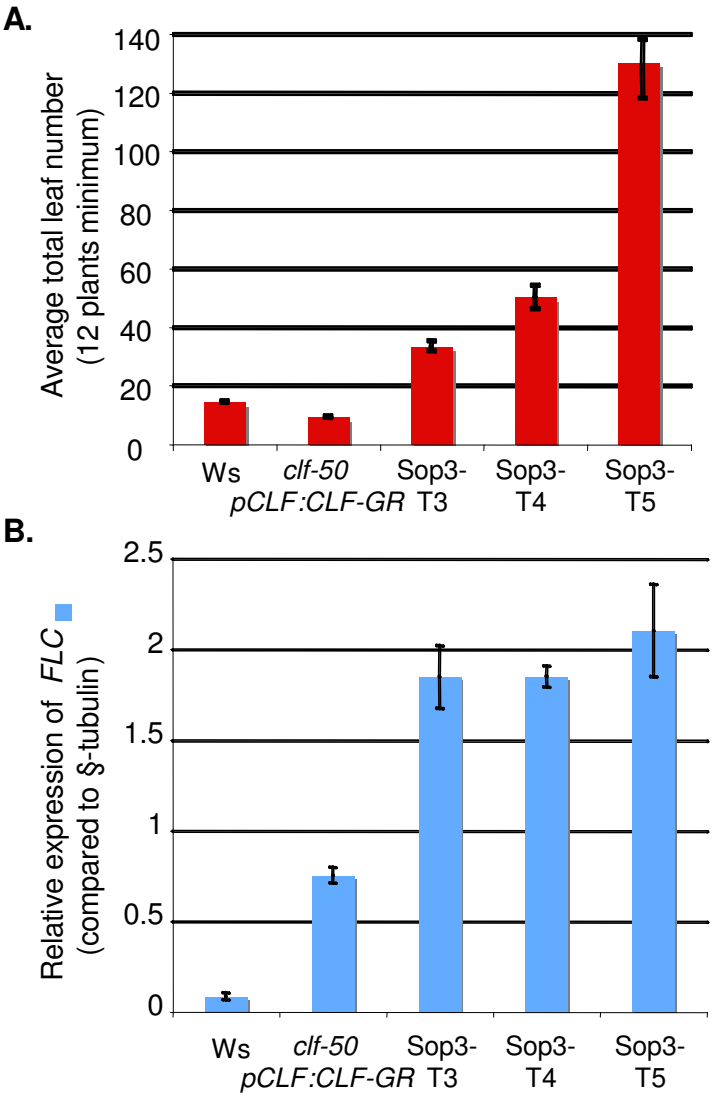

Figure 5.5: Flowering time and expression analysis of *FLC* by semi-quantitative RT-PCR on three successive generations of *sop3* plants. Genotypes tested: Ws, *clf-50 pCLF:CLF-GR*, Sop3-T3, Sop3-T4 and Sop3-T5. A: Flowering time experiment under LD conditions. Results show the average of rosette leaves at bolting (12 plants minimum) for each of the genotypes tested. B: Quantification of the transcript levels of *FLC* for each of the genotypes tested. The quantification corresponds to the average ratio (target gene signal/ß-tubulin signal) of three different RT-PCR reactions. For the semi-quantitative RT-PCR, plants were grown on soil in LD conditions and mRNA was extracted from 15 days old seedlings. Bars in both experiments show standard Error (SE).

The level of expression of *FLC* related to ß-tubulin in *clf-50 pCLF:CLF-GR* plants is approximately eight times higher than in Ws plants. The expression of *FLC* compared to ß-tubulin in Sop3-T3, Sop3-T4 and Sop3-T5 seedlings is approximately two and a half higher than the level in the progenitor line. The level of expression of *FLC* in Sop3- plants from different generations is consistent with the level of mis-regulation observed in a previous experiment (see section 5.1.2, Figure 5.3). In this experiment it was shown that in Sop3- plants the expression level of *FLC* is highly mis-regulated but this mis-regulation does not increase between generations.

These experiments showed that the regulation of the time to flower in Sop3- plants involves a mechanism allowing the induction of flowering, partially independent of *FLC* and the effect is cumulative with generations.

5.2. *sop2* is an untagged mutation

For this suppressor, no seeds could be recovered from T2 Clf- siblings. As a result, there were no segregating families to study the co-segregation of the mutant phenotype with a specific T-DNA. Nonetheless, Southern-blot analyses were conducted and plants with one T-DNA insertion were found. Using plasmid rescue the T-DNA insertion detected in Southern-blot was identified. However, the T-DNA insertion found did not cause the Sop2- phenotype.

5.2.1. Identification of a single T-DNA insertion in Sop2- plants

Southern-blot analysis was performed using a probe detecting CaMV35S enhancer sequences on eleven T3 plants (progeny of a suppressed T2 plant), showing the suppressed Sop2- phenotype (Figure 5.6). After two generations of self-pollinisation, the plants studied in Southern-blot were considered to be homozygous for the mutation responsible for the Sop2- phenotype. Among all the plants tested, plants

had either one (lanes 1, 3, 7, 8, 9, 10 and 11) or two (lanes 2, 4 and 5) bands (T-DNA insertions) as showed by the red arrows in FFigure 5.6. All the plants tested have the same band at 10Kbp (low intensity for lanes 2, 4, 6 and 10), which is therefore a candidate to cause the Sop2- suppressed phenotype. The low intensity in some lanes is certainly due to a less DNA loaded for these lanes.

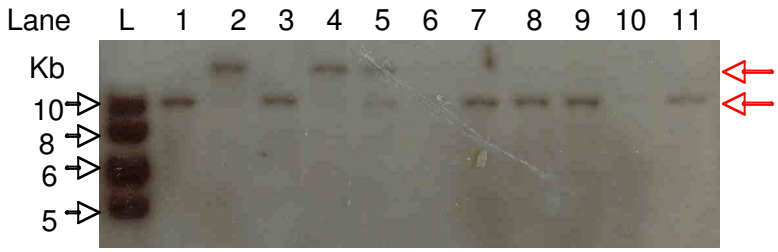

Figure 5.6: Southern-blot analysis on Sop2- plants. Southern-blot with *Eco*RI digested DNA from 11 independent plants (lanes 1 to 11). The red arrows indicate the band detected at 10 Kb. The probe used detects CaMV35S enhancer sequences of the pSKI074 vector used for the mutagenesis. L corresponds to the 1Kbp ladder used to assess the size of the bands obtained (sizes showed on the left hand side of the figure).

The Southern-blot analysis allowed the identification of a T-DNA insertion (band at 10Kbp), potentially linked to the suppressed phenotype observed as it co-segregates with Sop2- suppressed plants.

5.2.2. Fine mapping of the single T-DNA insertion in Sop2- plants

The suppressed plant analyzed in lane 1 of the Southern-blot in Figure 5.6 had only one T-DNA insertion (band at 10Kbp). Its DNA was used to plasmid rescue the Right Border (RB) of the T-DNA and adjacent plant sequences (chapter 2 section 2.4.8 (Weigel et al., 2000)). The plant sequence found by plasmid rescue was the 3' region of the *At1g54930* locus, which codes for a protein containing a zinc knuckle

(znkn; CCHC-type) domain (Figure 5.7). Taking advantage of the complete genome sequence from The *Arabidopsis* Information Resource (TAIR), primers to sequence the plasmid previously obtained were designed. Their location in the genome is specified in Figure 5.7. As the size of the plasmid was large (about 10 kbp), a fragment of plant genomic DNA of 5 kbp approximately was recovered. That is why the T-DNA insertion was located quite far from the first sequences obtained by plasmid rescue in the first exon of the adjacent coding region of *At1g54930*: the *At1g54940* locus. This locus encodes for a *PLANT GLYCOGENIN STARCH INITIATION PROTEIN-4* (*PGSIP-4*).

It has been reported that in *A. thaliana*, Plant Glycogenin Starch Initiation Proteins (PGSIPs) exist in a small gene family comprising six members (from *PGSIP-1* to *PGSIP-6*) (Chatterjee M., 2005). *PGSIP-1* was identified from yeast and animal glycogenin sequences.

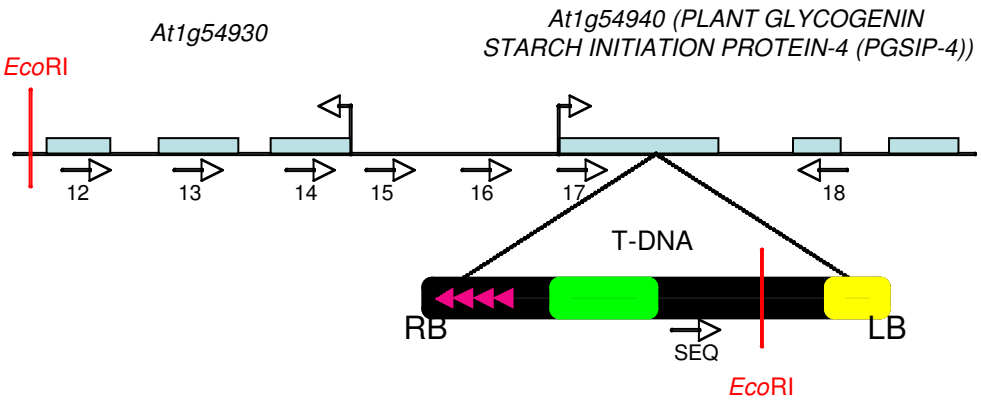

Figure 5.7: Diagram of the T-DNA insertion in Sop2- plants. The T-DNA insertion is located in the first exon of *PLANT GLYCOGENIN STARCH INITIATION PROTEIN-4* (*PGSIP-4*). Arrows correspond to primers used to map the insertion and genotype by PCR the *sop2* mutation. The *Eco*RI restriction site allowing the recovery of *At1g54930* genomic sequences are shown in red. For details see primers table (chapter 2 section 2.4.3).

Using a RNA interference (RNAi) knockout strategy, it was shown that repression of the expression of *PGSIP-1* leads to a reduced starch phenotype in leaves of transgenic plants indicating that this protein plays a role in the starch biosynthesis pathway and may have a starch priming function as well as the other members of the PGSIP family (Chatterjee M., 2005).

5.2.3. *PGSIP-4* and *At1g54930* are over- expressed in Sop2- plants

First, the expression of *PGSIP-4* was tested to confirm if it was down regulated in Sop2- plants due to the T-DNA insertion in the first exon of *PGSIP-4*. Surprisingly, the expression of *PGSIP-4* compared to ß-tubulin in Sop2- plants was nine fold higher than in Ws and *clf-50 pCLF:CLF-GR* plants (Figure 5.8).

The expression of *At1g54930* was also controlled. As for *PGSIP-4*, the expression of *At1g54930* compared to ß-tubulin in Sop2- plants is up regulated. It is six fold higher than in Ws and *clf-50 pCLF:CLF-GR* plants (Figure 5.8).

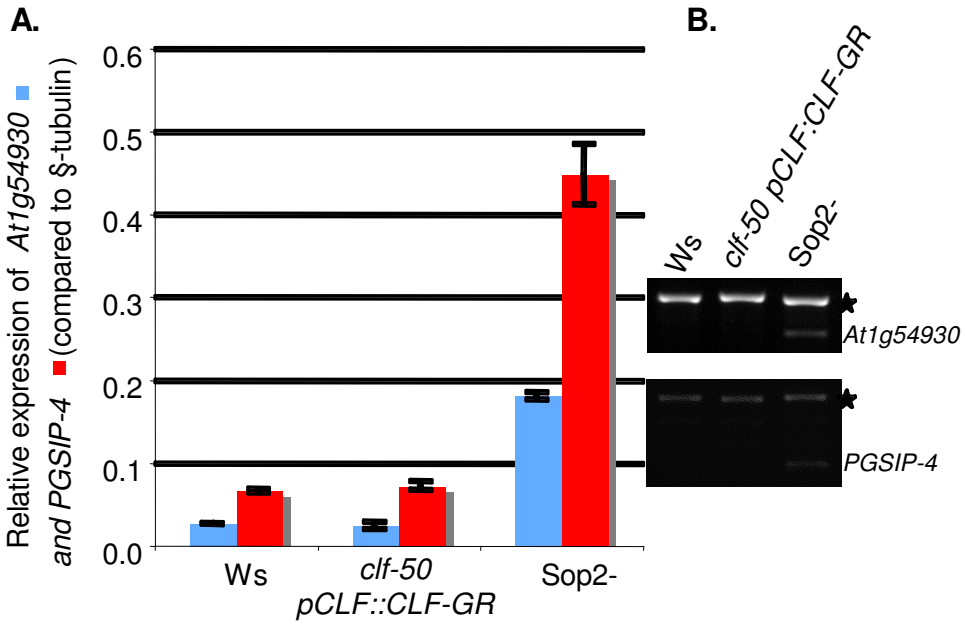

Figure 5.8: Expression analysis of *At1g54930* and *PGSIP-4* in Sop2- plants by semi-quantitative RT-PCR. A. Quantification of the transcript levels of *At1g54930* and *PGSIP-4* in Ws, *clf-50 pCLF:CLF-GR* and Sop2- plants. The quantification corresponds to the average ratio (target gene signal/ß-tubulin signal) of three different RT-PCR reactions. Bars show Standard Error (SE). B: Example of RT-PCR of *At1g54930* and *PGSIP-4* on Ws, *clf-50 pCLF:CLF-GR* and Sop2- cDNA samples. Duplex PCRs were performed to amplify ß-tubulin (marked with ⋆) as internal control with specific primers for *At1g54930* and *PGSIP-4*. The plants were grown on soil in LD conditions and mRNA was extracted from 15 days old seedlings.

From this experiment it can be concluded that the expression of both genes: *At1g54930* and *PGSIP-4* is up regulated in Sop2- plants. The *CaMV35S* enhancing sequences at the RB of the T-DNA are certainly the cause of the over expression of *At1g54930* and *PGSIP-4* in Sop2- plants.

5.2.4. Phenotypic assay for the over expression of *PGSIP-4*

It was shown in the literature that repression of *PGSIP-1* leads to a reduced starch phenotype in leaves (Chatterjee M., 2005). As *PGSIP-4* and *PGSIP-1* belong to the same gene family, in order to phenotypically confirm that the over expression of *PGSIP-4* increase starch levels in leaves in Sop2- plants, the levels of starch in Sop2- leaves were qualitatively tested and compared with the starch levels in leaves from Ws and *clf-50 pCLF:CLF-GR* plants (Figure 5.9).

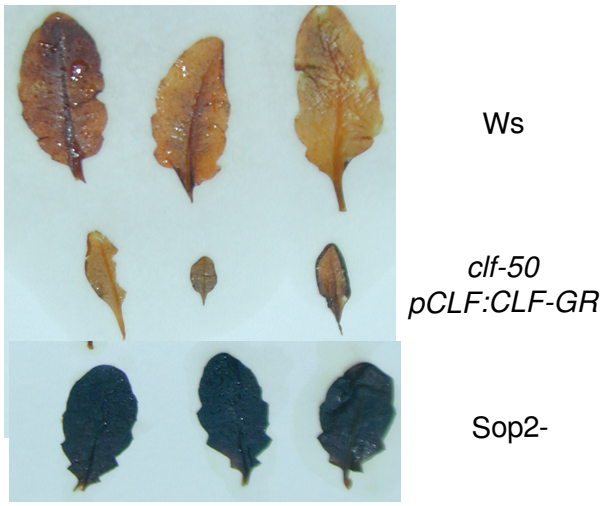

Figure 5.9: Qualitative starch assay by staining Arabidopsis leaves with iodine solution at the end of the light period. Plants were five weeks old. In the assay performed the starch was visualised as a violet-dark coloration by staining with an iodine solution followed by a decolouration with 96% ethanol (chapter 2 section 2.3.5).

The leaves from Sop2- plants are coloured in an intense violet-dark colour corresponding to high contents in starch. Leaves from Ws and *clf-50 pCLF:CLF-GR* plants exhibit an orange colour corresponding to a low content in starch. This colour difference indicates that levels of starch are higher in Sop2- leaves than in Ws and *clf-50 pCLF:CLF-GR* leaves.

In this experiment the molecular phenotype observed in Sop2- plants where *PGSIP-4* is over expressed could be linked with a high starch phenotype. *PGSIP-4* was confirmed to be involved in starch biosynthesis and may also have a starch priming function.

One reason why the T-DNA insertion in *PGSIP-4* induces more enzyme activity is because the first exon of *PGSIP-4* is not coding then an active protein is over-expressed due to the T-DNA insertion. Overall, it is very surprising and confusing to see the enzyme activity increased if the enzyme is knocked down.

5.2.5. The Sop2- phenotype does not cosegregate with the T-DNA insertion in *PGSIP-4*

The *sop2* mutant was crossed with wild type and with the progenitor line *clf-50 pCLF:CLF-GR*. From the cross between Sop2- and Ws plants, 18 F1 plants with a wild type phenotype were obtained. This result suggests that the Sop2- late flowering phenotype is recessive. Unfortunately, the cross between Sop2- and the progenitor line failed. Recessivity/dominance of the *sop2* mutation in the *clf* mutant background could not be tested at this point of the thesis. Nonetheless, two F2 families from independent F1 plants from the cross between Sop2- and Ws plants were grown on soil. After three weeks, the curly leaf phenotype was scored for each plant ([+] for curled leaves and [-] for flat leaves (suppression of the Clf- phenotype) as shown in Figure 5.10-B), genomic DNA was extracted and 24 plants were genotyped by PCR (Figure 5.10). In addition to the *At1g54940* locus, plants were genotyped for the *clf-50* mutation.

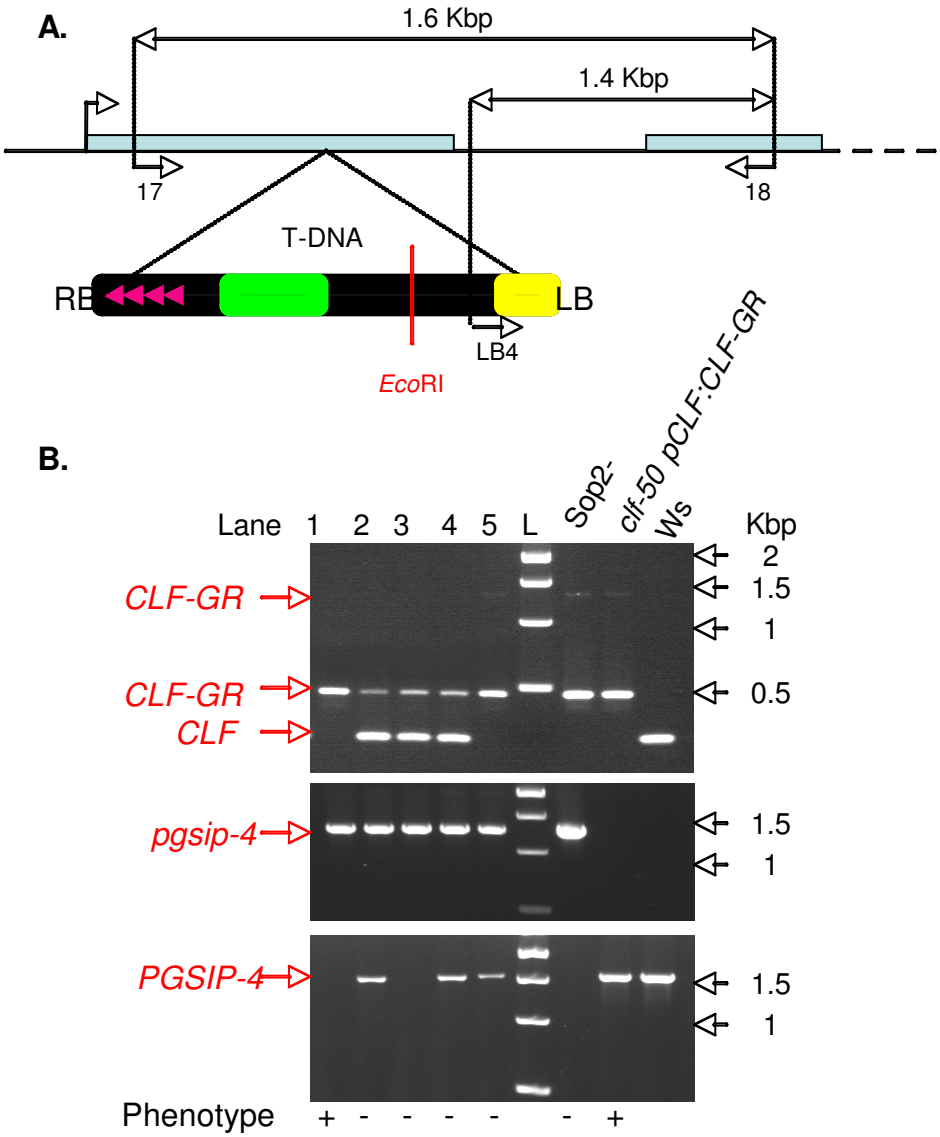

Figure 5.10: PCR Genotyping of the *clf-50* and the T-DNA insertion in *PGSIP-4* mutations in segregating F2 plants. A: Location of primers 17: 5fAt1g54930; 18: At1g54940-SI-R at the *PGSIP-4* locus and LB4 in the T-DNA used to genotype the T-DNA insertion in *PGSIP-4*. PCR products will amplify a fragment of 1.6 Kbp for the wild type *PGSIP-4* allele and 1.4 Kbp for the *pgsip-4* mutation. The *Eco*RI restriction sites allowing the recovery of *At1g54930* genomic sequences are shown in red. B: PCR genotyping results for five plants (lanes 1 to 5). The top gel corresponds to a duplex PCR to genotype for the *clf-50* mutation. Three bands can be distinguished at 0.25, 0.45 and 1.3 kbp. Bands at 0.45 and 1.3 Kbp correspond to the transgene *pCLF:CLF-GR* (in red). Heterozygous or homozygous *CLF+* have only the band at 0.25 Kbp (in red). *clf-50* plants show no amplification at 0.25 Kbp with the primers

used. The middle and lowers gels show the PCR amplifications obtained for *pgsip-4* (1.4 Kbp in red) and *PGSIP-4* (1.6 Kbp in red) respectively. The expected sizes for the PCR amplifications for the three different controls are shown in the right hand side of the three different gels. The curled leaf phenotype was scored for each plant ([+] for curled leaves and [-] for flat leaves.

This analysis allowed finding *clf-50* homozygous plants with no curled leaves (CLF^+ = [-] phenotype). These were presumably Sop2- plants. As a controls, Sop2- homozygous mutant, the progenitor line *clf-50 pCLF:CLF-GR* and wild type (Ws) DNA were used. In Figure 5.10-B a representative sample of the PCR amplification obtained on 25 plants tested is shown. In the upper gel, in lanes 1 and 5 plants are *clf-50* homozygous as no amplification at 0.25 Kbp was detected and PCR bands were detected at 0.45 Kbp (strong) and at 1.3 Kbp (low intensity). Plants tested in lanes 2, 3 and 4 are CLF^+/CLF^+ or $CLF^+/clf\text{-}50$ and carry the *pCLF:CLF-GR* transgene (Figure 5.10-B). In the middle gel all the plants tested have the band corresponding to the T-DNA insertion mutation at the *PGSIP-4* locus. In the bottom gel, plants tested in lanes 2, 4 and 5 have the band corresponding to a wild type *PGSIP-4* allele. In relation with the insertion at the *PGSIP-4* locus, plants tested in lanes 1 and 3 are homozygous for the T-DNA insertion whereas plants tested in lanes 2, 4 and 5 are heterozygous for the T-DNA insertion (Figure 5.10-B).

The results from the genotyping show that the plant in lane 5 is *clf-50* homozygous but has flat leaves as the phenotype scored is [-] (Figure 5.10-B). Therefore, the plant in lane 5 is a Sop2- mutant, as the Clf- phenotype is suppressed. The same plant is also *PGSIP-4*/*pgsip-4* heterozygote, therefore the *sop2* mutation is tagged and linked to the T-DNA insertion in *PGSIP-4* only if *pgsip-4* is a dominant mutation. Unfortunately, the plant in lane 1 which genotype is *clf-50 pgsip-4* homozygous and did not exhibit suppression of the Clf- phenotype as expected if the *pgsip-4* mutation was dominant (Figure 5.10-B).

In the light of these results, it could be concluded that the T-DNA insertion in *PGSIP-4* is not linked with the Sop2- phenotype. Consequently, the *pgsip-4* mutation is not responsible for the suppressed phenotype of Sop2- plants.

5.3. *sop3* is an uncharacterized mutation

For this suppressor as for Sop2-, no seeds could be recovered from T2 Clf- siblings. No segregating families to study the co-segregation of the mutant phenotype with a specific T-DNA were available. Nevertheless, Southern-blot analyses were conducted and plants with two T-DNA insertions were found. Using plasmid rescue and genome walker T-DNA insertions detected in Southern-blot were identified. However, as for *sop2*, the *sop3* mutation was shown not to be tagged as the Sop3- phenotype does not cosegregate with the kanamycin resistant gene carried by the T-DNA.

5.3.1. Identification of two T-DNA insertion in Sop3- plants

Southern-blot analysis was performed using a probe detecting *CaMV35S* enhancer sequences of nine T3 sibling Sop3- plants (Figure 5.11). Among all the plants tested, plants had either two (lanes 1, 2, 3, 4, 7 and 8) or five (lanes 5, 6 and 9) bands (T-DNA insertions) as showed by the red arrows in Figure 5.11. All the plants tested have the same bands at approximately 10.5 Kbp and 7 Kbp, which seemed to be linked with the Sop3- suppressed phenotype.

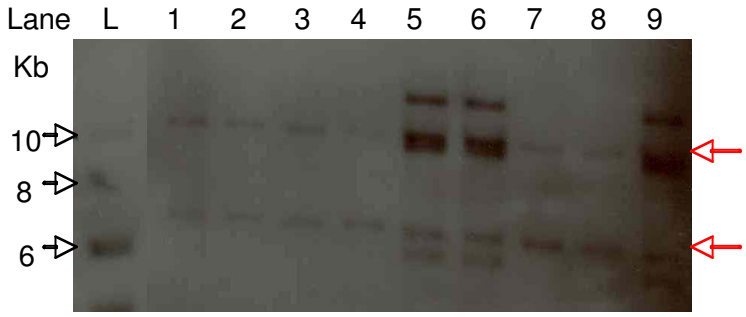

Figure 5.11: Southern-blot analysis of homozygous Sop3- plants. Southern-blot with *Eco*RI digested DNA from nine independent plants (lanes 1 to 9). All the plant used were homozygous for the *sop3* mutation. The red arrows indicate the bands detected at 10.5 and 7 Kbp. The probe used detects CaMV35S enhancer sequences of the pSKI074 vector used for the mutagenesis. L corresponds to the 1Kbp ladder used to assess the size of the bands obtained (sizes showed on the left hand side of the figure.

The Southern-blot analysis allowed the identification of two T-DNA insertions (bands at 10.5 and 7 Kbp). These two bands were potentially linked to the suppressed phenotype, as they were present in all Sop3- plants tested by Southern-blot.

5.3.2. Mapping of the locus disrupted by the T-DNA insertion in Sop3- plants

The suppressed plant analyzed in lane 1 of the Southern-blot in Figure 5.11 had two T-DNA insertions (bands at 10.5 and 7 Kbp). Its DNA was used for plasmid rescue of the Right Border (RB) of the T-DNA and adjacent plant sequences (chapter 2 section 2.4.8 (Weigel et al., 2000)). Only the band at 7 Kbp could be recovered by this method. In addition, a genome walker strategy was used as well to clone adjacent plant genomic sequences at the Left Border (LB) of the T-DNA (chapter 2 section 2.4.9). The plant sequence found by plasmid rescue was the 5′ region of the

At1g04380 locus, which encodes a protein similar to a 2-oxoglutarate-dependent dioxygenase (Figure 5.12-A).

Sequence analysis indicated that a Left Border (LB) sequence was present and adjacent to the Right Border (RB) (Figure 5.12-A), indicating that the insert was complex and contained at least one partial T-DNA. Using the genome walker strategy, three different *Arabidopsis* genomic sequences were recovered. (i) The 3' region of the *At1g04380* locus, which is the same locus as the one identified by plasmid rescue. (ii) The 5' region of *At5g66550*, which encode a protein with a Maf domain. This domain has been shown to be a putative inhibitor of septum formation in eukaryotes, bacteria, and archaea (Minasov et al., 2000). (iii) The last exon of *At1g59830*, which encode one of two isoforms of the catalytic subunit of protein phosphatase 2A (Perez-Callejon et al., 1998) (Figure 5.12-B). Collating the results obtained from the plasmid rescue and from the genome walker are consistent with a deletion at *At1g04380* as outlined in Figure 5.12-C.

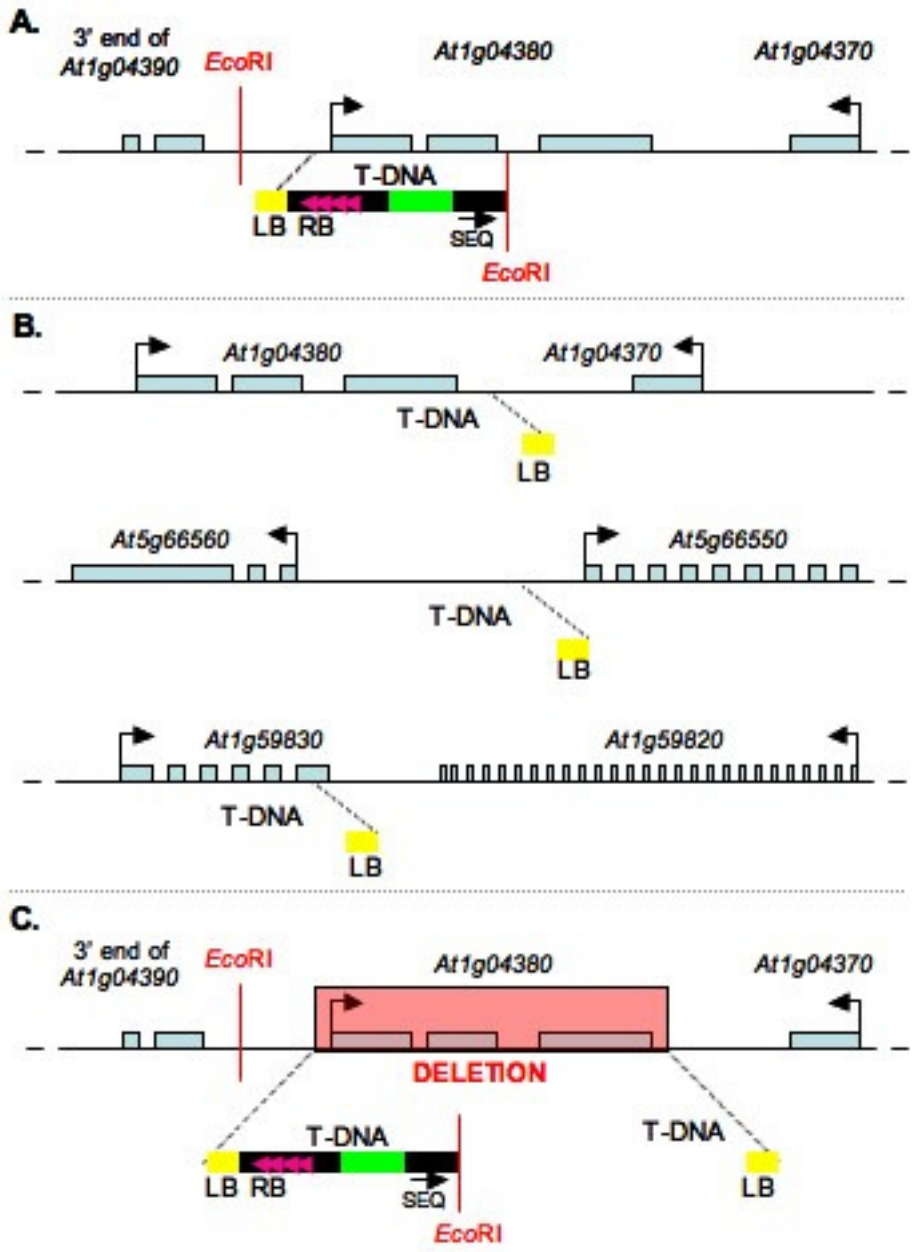

Figure 5.12: Diagram of the different locations of the T-DNA insertions obtained from Sop3- plants. A: T-DNA insertion in the 5' region of *At1g04380* recovered by plasmid rescue. The *Eco*RI restriction sites allowing the recovery of the genomic sequence are shown in red. B: Three different T-DNA insertions recovered from the LB of the T-DNA by genome walker. The 3' region of *At1g04380*, the 5' region of *At5g66550* and the last exon of *At1g59830*. C: A putative rearrangement of the *At1g04380*. The pink box correspond to a deletion of the whole *At1g04380* locus.

From Sop3- plants, three loci were recovered but only two bands were detected on Southern-blot. So, either one insert is partial (lacks the RB sequences), or a re-arrangement (inversion, duplication, translocation, aborted insertions causing deletions) happened during the T-DNA insertion. At this stage any of the three T-DNA inserts identified can be responsible for the Sop3- phenotype.

5.3.3. The *sop3* mutation could be tagged by a partial T-DNA insertion

It is known that the T-DNA integrates in the genome in an orientated manner. The first element to be integrated in the genome is the Left Border (LB), and then the T-DNA itself is integrated in the host genome. The process finishes by the incorporation of the Right Border (RB) (Tzfira et al., 2004). As three T-DNA insertion of the Left Border were identified by genome walker (which only requires LB sequences, unlike the plasmid rescue which requires majority of the T-DNA especially the pBstKS+ backbone), it was not sure that the whole T-DNA is present for these insertions. Presumably they are not as two bands were identified with a RB probe but there are three insertions at unlinked loci. Prior to any genotyping by PCR of any of the putative T-DNA insertions identified previously by plasmid rescue and genome walker, the presence of the full T-DNA was tested by sowing seeds from Sop3- plants in MS plates with kanamycin. The phenotype of the lines tested for the resistance to kanamycin was also scored in MS plates without the antibiotic.

The results in Table 5.1 show that in the T2 generation, more than one T-DNA is present in *sop3* plants. Indeed, the ratio between Kan^R and Kan^S plants is close to 10:1, which is between 3:1 (one insertion) and 15:1 (two insertions). This confirms the Southern-blot patterns seen previously where the minimum number of T-DNA insertion was two. In the T3 and T4 generations some Sop3- plants were found that carried only one T-DNA insertion, as the ration between the Kan^R and the Kan^S

plants is 3:1. For the same T3 and T4 lines 100% of the plants exhibit the Sop3- phenotype.

Table 5.1: Kanamycin resistance test for five independent lines in theT2, T3 and T4 generations. Seedlings were grown on MS plates with and without 50 g/mL of Kanamycin. The phenotypes (Kan^R, Kan^S and *sop3*) were scored after 14 days of growth in tissue culture conditions. The generation tested is shown in brackets. For the statistical analysis, a Chi Square test for 3:1 segregation of the Kan^R and Kan^S phenotypes was performed. NS: not significant at P=0.05.

Line	**Kan^R**	**Kan^S**	**Total plants**	**$\chi 2$ (3:1)**	**P**	**Phenotype**
sop3 line 1 (T2)	87	9	96	12.5	NS	100% Sop3-
sop3 line 2 (T3)	38	16	54	0.62	0.43	100% Sop3-
sop3 line 3 (T3)	69	22	91	0.03	0.86	100% Sop3-
sop3 line 4 (T4)	72	19	91	0.82	0.36	100% Sop3-
sop3 line 5 (T4)	98	31	129	0.06	0.8	100% Sop3-

These results show a progressive silencing of the gene conferring the kanamycin resistance throughout generations. This might be due to some post-transcriptional gene silencing, a phenomena currently seen in transgenic plants.

Nonetheless, all together these results show that Sop3- plants could be tagged by a partial insert yet to be identified. The T-DNA insertions identified by genome walker in the 5′ region of *At5g66550* and in the last exon of *At1g59830* would be the most promising candidates for further studies.

5.4. Summary

In this chapter two suppressor of the Clf- phenotype found in a screen made in tissue cultures plates: *suppressor of polycomb 2* (*sop2*) and *suppressor of polycomb 3* (*sop3)* were presented. Both suppressors showed the same strong suppression of the leaf curling, characteristic of the *clf* mutant background, and a very late flowering phenotype fitting with a high mis-expression of *FLC*. By testing the ability of these suppressors to respond to a vernalization treatment and by assessing the expression level of *AG*, it was shown that the mutations responsible for the Sop2- and Sop3- phenotypes are distinct. The *sop3* mutation in particular is unique as Sop3- plants flower later with vernalization treatment, the expression of *AG* is down regulated compared with the progenitor line *clf-50 pCLF:CLF-GR* and the time to flower increases between generations. Unfortunately, it was not possible to identify the mutations responsible for Sop2- and Sop3- suppressed phenotype. However, using genome walker, candidate mutations for the *sop3* mutation, in the 5' region of *At5g66550* and in the last exon of *At1g59830* could be found. These are interesting targets for future isolation of the *sop3* mutation.

6. Discussion

In this study, four mutations that suppress the Clf- phenotype were found and in two cases the genes disrupted were identified and characterized (*sep3-7* and *fpa-10*). I showed that *SEPALLATA3* is a direct target of *CLF* and that it is an activator and a co-factor of *AG*. In addition, I showed that *FT*, which positively regulates *SEP3*, suppresses the Clf- phenotype suggesting that *FT* is also a target of *CLF*. I also showed that the suppression of the Clf- phenotype by *fpa-10* is linked to a high mis-expression of *FLC*. Surprisingly, the *clf-50 fpa-10* mutant flowers much later than the *clf-52 FRI*$^{+}$ mutant, which have equal high levels of *FLC*, suggesting that *FPA* may regulate other genes controlling flowering in addition to *FLC*. Expression analyses in the double mutant *clf-50 fpa-10* suggest that *SHP2*, *SEP3* and/or *FUL* could be targets of *FPA*. Finally, the two unidentified mutations, *sop2* and *sop3*, suppressing the Clf- phenotype show delayed time to flower phenotypes suggesting a shared role with *CLF* in the control of flowering in *A. thaliana*. Taken together, these results suggest that *CLF*, through the Polycomb Repressive Complex 2 (PRC2), regulates floral induction (*FLC*), floral integration (*FT*) and floral organ formation (*SEP3* and *AG*) in *A. thaliana*.

6.1. *SEP3* is a novel direct target of *CLF*

Several studies based on *in-situ* hybridisations and northern-blot analysis indicate that the expression of *SEP3* (just before the expression of B and C MADS-box genes),

is exclusive to flowers, from stage 2 to 16 (Rounsley et al., 1995; Pelaz et al., 2000). *SEP3* belongs to a largely redundant sub-family of MADS-box genes, named *SEPALLATA 1* (*SEP1*), *SEP2*, *SEP3* and *SEP4*. They are important as they are crucial for the specification of the identity of all floral organs (Pelaz et al., 2000; Ditta et al., 2004). The *SEP* genes were defined as "class E genes" in the floral model for specification of floral organ identity (chapter 1 section 1.2.3) (Pelaz et al., 2001a; Krizek and Fletcher, 2005). Disruption of *SEP1*, *SEP2* and *SEP3* leads to the development of sepals rather than petals, stamens and carpels (that is why the name "SEPALLATA") (Pelaz et al., 2000). The complete disruption of the class E gene function (in *sep1 sep2 sep3 sep4* quadruple mutants) leads to the transformation of all floral organs into vegetative leaves (Ditta et al., 2004). It was shown that the class E floral homeotic protein SEP3 interacts with the class B proteins AP3 and PI in yeast three-hybrid assays (Honma and Goto, 2001). Moreover, two studies have shown that *SEP3* is required to induce a leaf curling phenotype in *A. thaliana* (Castillejo et al., 2005; Teper-Bamnolker and Samach, 2005). Whole genome ChIP-on-chip analysis on wild type (Col-0) plants showed that there is enrichment for CLF (Clarenz O. and Zhang X. unpublished results), and H3K27me3 (Zhang et al., 2007b) at *SEP3* sequences. Using the comprehensive expression atlas, *Arabidopsis thaliana* Tiling Array Express (At-TAX), which is based on whole-genome tiling arrays (Laubinger et al., 2008), it was observed that *SEP3* is expressed at very low levels during vegetative development and is highly expressed in flower development. In this thesis, the previous expression studies and the *in-silico* observation were

partially confirmed by showing that the expression of *SEP3* in wild type (Ws) seedlings is very low during vegetative development. Furthermore, in Clf- seedlings the expression of *SEP3* is highly mis-regulated during vegetative development (chapter 3). Therefore, SEP3 is needed for the leaf curling in *A. thaliana*. Chromatin IP (ChIP) permit to show that there is enrichment for H3K27me3 at *SEP3* sequences in wild type (Ws) plants. Consistent with the mis-expression of *SEP3* in *clf* mutant seedlings, H3K27me3 at *SEP3* is reduced in the *clf-50* mutant. This result shows that the H3K27me3 is functionally relevant for the repression of *SEP3* at vegetative stages of development.

The requirement of SEP3 for the leaf curling phenotype in Clf- plants could be because (i) SEP3 is needed to activate the expression of *AG* in the *clf* mutant or (ii) SEP3 is needed for the activity or stability of the AG protein in the *clf* mutant. To discriminate between the two possibilities, the expression levels of *AG* at the RNA and protein level were tested in *clf-50 sep3-7* plants. The AG protein is still present in the double mutant *clf-50 sep3-7* even though the mRNA expression is lower than in *clf-50* plants. Therefore, SEP3 is required as a co-factor for AG to cause leaf curling in leaves, and to some extent to express at high levels *AG*. This result confirms previous studies where the AG/SEP3 complex was shown to auto regulate *AG*'s expression (Gomez-Mena et al., 2005). How the AG/SEP3 complex induces leaf curling remains unclear, as the leaves do not show homeotic transformations. It has been shown that *35S::FT* activates *SEP3* expression in wild type leaves so

presumably the repression accomplished by PcG gene complexes can be overcome by the over-expression of *FT* (Teper-Bamnolker and Samach, 2005). Normally, in long day conditions *FT* is expressed in leaves but do not activate *SEP3* in wild type leaves suggesting that the CLF repression is confined to the meristem. Meaning that the transcriptional induction of *SEP3* by FT is spatially and temporally regulated in the meristem. One of the factors acting in this regulation is the epigenetic repression of *SEP3* by the PRC2.

The *E(Z)* homologous *SWINGER* (*SWN*), which is partially redundant with *CLF* (Chanvivattana et al., 2004) might also be involved in the repression of *SEP3* as a residual H3K27me3 was observed at *SEP3* sequences in *clf-50* plants. Although, whole genome expression analysis performed in the laboratory did not show that in the *swn* mutant *SEP3* was mis-regulated. In the double mutant *clf swn*, *SEP3* is 5.8 fold mis-regulated (Dr. F. Thorpe personal communication), suggesting that *SWN* plays a role in the regulation of *SEP3*. Together, these findings show that the PRC2 is responsible for the repression of the *SEP3* locus during vegetative development. Also, the repression is mediated by the H3K27me3 repressive epigenetic mark. CLF regulates then the correct spatial and temporal expression of *AG* and *SEP3*, two crucial MADS-box transcription factors, during vegetative development.

6.2.Non-redundant roles for *SEP3* in leaves and flowers

Among the small sub-family of redundant MADS-box transcription factors *SEPALLATA1* (*SEP1*), *SEP2*, *SEP3* and *SEP4*, *SEP3* is the one that is characterized the best. This is due to a non redundant role of *SEP3* in flower development. For example, *sep3* mutants exhibit phenotypes with close resemblance of the phenotypes observed for intermediates *ap1* mutants (Pelaz et al., 2001b). Indeed, petals show a partial transformation towards sepals with the presence of stomata structures while, in the triple mutant *sep1 sep2 sep4* the floral organ development is like in wild type plants (Pelaz et al., 2001b). Moreover, *SEP3* is enough to induce carpel and ovule development in the double mutant *sep1 sep2* whereas *SEP1* or *SEP2* fails to promote ovule development in *sep2 sep3* and *sep1 sep3* mutant backgrounds respectively (Favaro et al., 2003). Finally, *SEP3* has a transcriptional activation potential activity that *SEP1* and *SEP2* do not have (Honma and Goto, 2001). Therefore, it has been proposed by several authors that *SEP3* plays several key roles in (i) floral meristem identity (Honma and Goto, 2001; Pelaz et al., 2001b; Castillejo et al., 2005), (ii) floral organ identity (Pelaz et al., 2000; Pelaz et al., 2001a; Ditta et al., 2004) and (iii) ovule development (Favaro et al., 2003). The major role of *SEP3* is confirmed by phylogenetic studies showing that close to the emergence of the angiosperms two SEP lineages appeared, one contains *SEP3* and another in which two more duplications happened leading to *SEP1*, *SEP2* and *SEP4* (Zahn et al., 2005).

It is interesting to point out that *SEP4* is also mis-expressed in Clf- plants (Dr. F. Thorpe personal communication). Although, in the double mutant *clf sep3* the leaf

curling is totally abolished. This suggests firstly, that *SEP4* cannot mediate the leaf curling without *SEP3* and secondly, that the *SEP3* function is non redundant in leaves of Clf-plants.

6.3.Repression of flowering by *CLF*

Polycomb proteins are one of the cellular systems that maintain transcription repressed through mitoses at target genes. This repression occurs in specific time and space during the organism life cycle. Loss of function mutations in *CLF*, a member of PRC2-like complexes in *A. thaliana* show weak homeotic transformations and an early flowering time phenotype (Goodrich et al., 1997; Chanvivattana et al., 2004). Previous studies showed that *CLF* is involved in the regulation of the MADS-box transcription factors AGAMOUS (AG) and *APETALA3* (*AP3*) (Goodrich et al., 1997; Katz et al., 2004; Schubert et al., 2006). *AG* and *AP3* are key flowering factors as they regulate floral organ identity. Thus, it has been proposed that *CLF* is a repressor of floral homeotic genes during flower development (Reyes and Grossniklaus, 2003; Pien and Grossniklaus, 2007). In this study I suggest that the role of *CLF* in regulating flowering is broader than formerly thought. I propose that *CLF* also regulates genes involved in the induction and the integration of exogenous and endogenous signals leading to flower formation. I provide genetic and expression results suggesting that *Flowering Locus T* (*FT*) and *Flowering Locus C* (*FLC*) are regulated by *CLF* (see chapters 3 and 4). A very recent study (Jiang et al., 2008) reveals with strong evidence that PRC2-like complex subunits *CLF*, *EMF2* and *FIE* repress the expression of *FLC*, *MADS AFFECTING FLOWERING 4* (*MAF4*),

MADS AFFECTING FLOWERING 5 (*MAF5*) and *FT*. Moreover, the authors showed that CLF directly binds to and mediates the deposition of the repressive mark H3K27me3 at these four loci (Jiang et al., 2008). The different phases of flower development (induction, integration and floral organ specification) and their links have been extensively described in the literature (Komeda, 2004; Krizek and Fletcher, 2005; Parcy, 2005). In this thesis, I confirm few of the findings of Jiang et al., 2008 (*FLC* and *FT*) and showed that the same regulating system, here the PRC2-like complex containing CLF, regulate early genes (repression of flowering by *FLC*) as well as intermediate regulators (integration and floral stimulation by *FT*) and finally late effectors (induction of floral organogenesis by *AG*). Since *FT*, which is not a MADS-box transcription factor is also regulated by *CLF*, it seems that *CLF* targets are not just MADS-box transcription factors. One possibility could be that in the absence of *CLF*, PRC2-like targets as *AG*, *SEP3* or *AP3* lose the H3K27me3. Thus, the induction of their transcription could be dependent on *FT*.

6.4.Role of *FPA* in regulating *FLC*

FPA is a member of the autonomous pathway, which promotes flowering in *A. thaliana* (Chapter 1 section 1.2.2). It has been shown that the *fpa* mutation causes late flowering in long and short days (Schomburg et al., 2001). Moreover, *FPA* effects on flowering seem to be exclusively mediated by *FLC* as the double mutant *fpa flc* flower normally (Dr. J. Goodrich personal communication). *FPA* encodes a plant-specific RNA binding protein containing a RNA Recognition Motif (RRM) domain.

(Schomburg et al., 2001; Quesada et al., 2005). Another member of the autonomous pathway, *FCA*, also posses a RRM domain (Quesada et al., 2005). It has been shown that FCA forms a complex with FY to regulate *FLC* and its own expression. FCA acts by controlling the polyadenylation sites of its targets. Although, it is only known for its autoregulation (Simpson et al., 2003). As *FPA* has a strong sequence (RRM domain) and function (repression of *FLC*) homology with *FCA*, it was hypothesized that FPA could also control polyadenylation site selection. However, the mechanism of *FPA*'s regulation of *FLC* is unknown as there are no modifications in *FLC*'s splicing in the *fpa* background. Very recently, it has been proposed that *FCA* and *FPA* act together with the RNA Polymerase IVa largest subunit (*NRPD1a*) and the *RNA DEPENDENT RNA POLYMERASE 2* (*RDR2*) and redundantly with each other to amplify siRNAs derived from transgenes. Thus, *FCA* and *FPA* would have a role in the perception and interpretation of the silencing signal at single-copy loci in *A. thaliana* (Baurle et al., 2007). Although, phenotypic relevance for these latest results is not totally clear. However, some genes silenced by small interfering RNA (siRNA) and DNA methylation are derepressed in the *fpa* background (Baurle et al., 2007). This suggests that *FPA* might play a role in epigenetic silencing through the siRNA pathway. The main target of *FPA* would be *FLC*. The results obtained in this study are consistent with the previous studies on *FPA* as in the double mutant *clf-50 fpa-10* the expression of *FLC* is highly mis-regulated (see chapter 4). Thus, *FPA* is necessary for the correct regulation of *FLC*. In this study it was shown that the high levels of

FLC are responsible of the suppression of the Clf- phenotype as shown in *clf-50 fpa-10* and *clf-52 FRI*$^{+}$.

6.5.*FPA* may act independently of *FLC* to promote flowering

The double mutant *clf-50 fpa-10* flowers much later (over 200 rosette leaves before bolting) than the *clf-52 FRI*$^{+}$ line (around 40 rosette leaves before bolting), although both lines exhibit high levels of *FLC*. This result suggests that *CLF* and *FPA* may act redundantly to control flowering genes other than *FLC*. In this study, I tested eight MADS-box transcription factors able to interact with AG and/or SEP3 in *A. thaliana*: *AGL6*, *SHP1*, *SHP2*, *SEP1*, *SEP2*, *SEP4*, *FUL* and *AGL24* (chapter 4 section 4.2.4). The expression of *FUL* in *clf-50 fpa-10* plants showed a decrease of 80% compared to *clf-50* plants. This result suggests that the down regulation of *FUL* in *clf-50 fpa-10* might be responsible for the late flowering phenotype observed. It will be very interesting to check the expression of FUL in a high FLC background. For example the clf-52 FRI+ line would give and idea of the FUL regulation in such conditions.

It has been reported that the levels of *FLC* is higher in the double *clf-81 fca-9* (Jiang et al., 2008). It would be interesting to know what happens to the levels of expression of *FLC* in Sop4- plants with and without dex. This experiment should allow us to compare the regulation of *FLC* by the FCA and FPA proteins.

Another candidate possibly responsible for the delayed flowering phenotype in Sop4- plants is *FWA*. *FWA* encodes a homeodomain-containing transcription factor and the transition to flowering in *A. thaliana* is delayed in *fwa* plants. The DNA sequence of wild type and *fwa* mutant alleles was identical in the genomic region of *FWA*. The *FWA* gene is ectopically expressed in *fwa* mutants and silenced in wild-type plants. This silencing is associated with methylation of two direct repeats in the 5' region of the gene (Soppe et al., 2000; Chan et al., 2005). Preliminary expression analysis of *FWA* showed that it is not mis-expressed in Sop4- plants.

Together, the phenotypic and molecular observations made on the double mutant *clf-50 fpa-10* are consistent with recent findings suggesting that *FPA* has many other targets than *FLC*. However, the true role of *FPA* is masked by the redundancy with other regulators as *FCA*. In addition, *FPA* may have a more general role in epigenetic silencing (Baurle et al., 2007).

6.6.sop3

Sop3- plants show a unique and very interesting late flowering phenotype. The time to flower for Sop3- plants with vernalization is greater (~66 rosette leaves) than without vernalization (~44 rosette leaves). Also, I observed that the time to flower of Sop3- plants increased between generations although, the expression of *FLC* did not. The mutant *vrn2-1* showed moderate resemblance on the responsiveness to vernalization with Sop3- plants. In Long Day (LD) conditions, *vrn2-1* plants flower

slightly late with vernalization treatments (~12 rosette leaves) than without vernalization treatment (~8 rosette leaves) (Gendall et al., 2001). Moreover, mutations in *METHYLTRANSFERASE 1* (*MET1*) have been shown to provoke similar transgenerational increase of the phenotype as in Sop3- plants. Weak *met1* mutations initially have no morphological phenotype, but *met1* plants show progressively more severe phenotypes following self pollinisation due to the accumulation of numerous epimutations (Kankel et al., 2003; Chan et al., 2005). So, the mutation responsible for the Sop3- phenotype might disrupt an epigenetic regulator in *A. thaliana*.

6.7. Validity of modifier screen for identifying PcG components

Despite pleiotropic effects of *PcG* gene mutants on development, it is very difficult to discriminate between direct effects and indirect consequences of developmental changes in these mutants. This is one of the reasons explaining why only few *PcG* gene direct targets have been identified in *A. thaliana* (Pien and Grossniklaus, 2007). Also, targets of *CLF* can act antagonistically (e.g. *FLC* and *FT*). So, the phenotypic outcome of removing the PcG activity is unpredictable. Indeed, *clf* mutants are early flowering even though *FLC* is highly mis-expressed. Whole genome approaches (Chip-on-chip) have identified many H3K27me3 decorated genes as potential novel direct targets but their relevance is limited as it is not known if they are mis-

expressed or not in *PcG* mutant backgrounds (Turck et al., 2007; Zhang et al., 2007b). Similarly expression studies or genetic screens can identify genes important for the Clf- phenotype but it is not known if they are direct targets or factors acting in the same silencing pathway. The combination of genetic and whole genome approaches can decipher PcG regulating pathways. In this study, by combining genetic screens with the available ChIP-on-chip data, it was possible to identify a novel direct target of *CLF* that mediates the Clf- phenotype (*SEP3*).

Previously in chapter 1 section 1.7, mutations that could suppress the Clf- phenotype were described as (i) direct targets of *CLF* as *AG* or (ii) proteins that act downstream of *CLF*. Therefore, the predictions made for a direct target (*SEP3*), and for a protein that prevents the activation of a *CLF* target (*FPA*), to suppress the Clf- phenotype are confirmed. Although, *FPA* may have more directs roles depending on how it acts with *CLF* to regulate other floral transition factors.

FPA and *SEP3* act at the extremities of the flowering pathway in *A. thaliana*. *SEP3* is a downstream floral meristem and floral organ identity gene (Krizek and Fletcher, 2005) whereas *FPA* is one of the numerous genes of the autonomous pathway involved in the repression of *FLC* and the initiation of flowering (Komeda, 2004). The mutations identified are relevant in the frame of the screen that I made as *SEP3* and *FPA* play important roles in the induction and the establishment of flower organs in *A. thaliana*.

6.8. Ancient role for PcG in preventing precocious flowering

In *A. thaliana*, it has been shown that plants with high mis-expression of both *SEP3* and *LFY* developed flower organs on leaves (Castillejo et al., 2005). The same kind of observation was made by over-expressing *PI*, *AP3*, *AG* and *SEP3* which are MADS-box transcription factors members of the ABC model for flower formation (Krizek and Fletcher, 2005), resulting in the transformation of cauline leaves into stamens or staminoid organs (Honma and Goto, 2001). In primitive plants as ferns (sporophylls), there are no flower organs and reproductive structures develop on leaves. The spatial and temporal repression of *SEP3* along with homeotic genes of the ABC model for flower formation as *AG* and floral integrators as *FT* and *LFY* could be the main mechanism to prevent flowers to form on vegetative organs as leaves. Moreover, as homeotic genes involved in the ABC floral model are well conserved, their repression at vegetative stages by polycomb proteins is certainly an ancient role of the PRC2-like complex containing CLF in plants.

6.9. Future work

In this work I showed that *SEP3* is a target of *CLF* (chapter 3 section 3.3). However, I could not experimentally show the enrichment of CLF at the *SEP3* locus in wild type (Ws) plants compare to the *clf-50* plants. Using anti GFP antibodies, enrichment in CLF-GFP fusion protein at *AG* and *STM* sequences has been shown by ChIP

(Schubert et al., 2006). Using the same approach it should be possible to confirm the ChIP-on-chip results for *SEP3* and show enrichment for CLF-GFP at the *SEP3* locus.

It would be very interesting to test if *FT* is responsible for the activation of *SEP3* and *AG* in the *clf* background. Expression analysis for *SEP3* and *AG* in the double mutant *clf ft* and in the triple mutant *clf ft flc* will be essential to answer that question. It would probably also be relevant for the *PcG* gene silencing pathway to define the exact pattern of expression of *SEP3* in wild type and in the *clf* mutant. A reporter gene protein fusion, *SEP3::GFP* par example, will be useful to see if SEP3 is present in the vasculature like FT.

The *sop2* mutation has to be identified. This can be done using two different methods. First, genome walker can be used to identify Left Border (LB) sequences corresponding to partial T-DNA insertions that could not be detected by Southern-blot with the CaMV35S enhancer sequence probe used in this study. The results obtained for *sop3* confirm that the technique is fairly robust and can rescue sequences flanking pSKI074 insertions. Secondly, a map-based cloning strategy can be adopted. By crossing Sop2- plants with a strong *clf* allele in Col-0 background like *clf-81* or *clf-24*, the mutation responsible for the Sop2- phenotype could be mapped and therefore cloned. The second strategy demands more time, handwork and money. Nevertheless, with the new whole genome sequencing techniques

available handwork and time are not limitations to use this approach, as it is certain of success.

The *sop3* mutation needs to be identified as well. To this end, Sop3- plants (Ws background) have been crossed with *clf-24* plants (Col-0 background) and the F2 generation is ready for realizing a map-based cloning strategy. Alternatively, oligonucleotides on both sides of the putative T-DNA insertions identified by genome walker could be designed and genotyping PCRs could be performed at the 3 loci (*At1g04380*, *At5g66550*, *At1g59830*) potentially linked with the T-DNA insertion responsible for the Sop3- phenotype.

In order to totally confirm that *FPA* suppresses the Clf- phenotype, independent alleles of *FPA* and *CLF* have to be cross in a *FLC* background. Moreover, I already cross *clf-28* and *fpa-7* in an *flc-3* background to see if the suppression of the Clf- phenotype was *FLC* dependent. The results from these genetic crosses would be available soon in the laboratory and will show if the suppression of the Clf- phenotype by *FPA* goes partially or totally through *FLC*.

As *FUL* was proposed to be a target of *FPA* in this study, it would be interesting to see if FPA associates with *FUL*'s chromatin or RNA using ChIP or RNA IP (RIP). Also, by looking at altered splicing or polyadenylation sites for the *FUL* transcript in the *fpa* mutant background, it would be possible to see an interaction between the

FPA protein with *FUL*. Finally, it has been shown that the *35S::FPA-YFP* complement the *fpa-8* mutant but also slightly promotes flowering. It would be interesting to see if the constitutive expression of *FPA* has any effect on *FUL*'s transcript by RT-PCR.

Recent studies showed that *LIKE HETEROPROTEIN 1* (*LHP1*) specifically binds to H3K27me3 and is responsible for long term transcription repression in *A. thaliana* (Turck et al., 2007; Zhang et al., 2007a). At the *FT* locus, loss of *LHP1* leads to *FT* derepression and early flowering (Gaudin et al., 2001). It has been proposed that LHP1 is able to "read" the repressive mark made by the PRC2. These findings combined with genome-wide analyses (Turck et al., 2007; Zhang et al., 2007b) have shown that PRC2 complexes regulate specific small domains in *A. thaliana* (Pfluger and Wagner, 2007). One of the main questions today is what are the kinetics and dynamics of the histone modification landscape and how these changes can control cell proliferation, differenciation and de-differenciation. Genetic screens for suppressors in an *lhp1* background might help in the identification of factors interpreting, maintaining and modifying the H3K27me3 repressive mark. By looking at the expression of floral model genes as *FLC*, *FT*, *SEP3* and *AG* or genes involved in meristem maintenance as *STM* and the histone modifications associated, this should give an idea of the changes involved at particular spatial and temporal stages of development.

7. References

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

Battaglia, R., Brambilla, V., Colombo, L., Stuitje, A.R., and Kater, M.M. (2006). Functional analysis of MADS-box genes controlling ovule development in Arabidopsis using the ethanol-inducible alc gene-expression system. Mech Dev **123,** 267-276.

Baurle, I., Smith, L., Baulcombe, D.C., and Dean, C. (2007). Widespread role for the flowering-time regulators FCA and FPA in RNA-mediated chromatin silencing. Science **318,** 109-112.

Bouveret, R., Schonrock, N., Gruissem, W., and Hennig, L. (2006). Regulation of flowering time by Arabidopsis MSI1. Development **133,** 1693-1702.

Bradley, D., Ratcliffe, O., Vincent, C., Carpenter, R., and Coen, E. (1997). Inflorescence commitment and architecture in Arabidopsis. Science **275,** 80-83.

Brown, J.L., Fritsch, C., Mueller, J., and Kassis, J.A. (2003). The Drosophila pho-like gene encodes a YY1-related DNA binding protein that is redundant with pleiohomeotic in homeotic gene silencing. Development **130,** 285-294.

Calonje, M., Sanchez, R., Chen, L., and Sung, Z.R. (2008). EMBRYONIC FLOWER1 participates in polycomb group-mediated AG gene silencing in Arabidopsis. Plant Cell **20,** 277-291.

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

Caspar, T., Huber, S.C., and Somerville, C. (1985). Alterations in Growth, Photosynthesis, and Respiration in a Starchless Mutant of Arabidopsis thaliana (L.) Deficient in Chloroplast Phosphoglucomutase Activity. Plant Physiol **79,** 11-17.

Castillejo, C., Romera-Branchat, M., and Pelaz, S. (2005). A new role of the Arabidopsis SEPALLATA3 gene revealed by its constitutive expression. Plant J **43,** 586-596.

Chan, S.W., Henderson, I.R., and Jacobsen, S.E. (2005). Gardening the genome: DNA methylation in Arabidopsis thaliana. Nat Rev Genet **6,** 351-360.

Chandler, J., Wilson, A., and Dean, C. (1996). Arabidopsis mutants showing an altered response to vernalization. Plant J **10,** 637-644.

Chanvivattana, Y., Bishopp, A., Schubert, D., Stock, C., Moon, Y.H., Sung, Z.R., and Goodrich, J. (2004). Interaction of Polycomb-group proteins controlling flowering in Arabidopsis. Development **131,** 5263-5276.

Chatterjee M., B.P., Vyas D., Coates S. and Barsby T. (2005). Reduced expression of a protein homologous to glycogenin leads to reduction of starch content in Arabidopsis leaves. Plant Science **168,** 501-509.

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

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

Coen, E.S., and Meyerowitz, E.M. (1991). The war of the whorls: genetic interactions controlling flower development. Nature **353,** 31-37.

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

de Folter, S., Immink, R.G., Kieffer, M., Parenicova, L., Henz, S.R., Weigel, D., Busscher, M., Kooiker, M., Colombo, L., Kater, M.M., Davies, B., and Angenent, G.C. (2005). Comprehensive interaction map of the Arabidopsis MADS Box transcription factors. Plant Cell **17,** 1424-1433.

de Napoles, M., Mermoud, J.E., Wakao, R., Tang, Y.A., Endoh, M., Appanah, R., Nesterova, T.B., Silva, J., Otte, A.P., Vidal, M., Koseki, H., and Brockdorff, N. (2004). Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. Dev Cell **7,** 663-676.

Ditta, G., Pinyopich, A., Robles, P., Pelaz, S., and Yanofsky, M.F. (2004). The SEP4 gene of Arabidopsis thaliana functions in floral organ and meristem identity. Curr Biol **14,** 1935-1940.

Fan, H.Y., Hu, Y., Tudor, M., and Ma, H. (1997). Specific interactions between the K domains of AG and AGLs, members of the MADS domain family of DNA binding proteins. Plant J **12,** 999-1010.

Favaro, R., Pinyopich, A., Battaglia, R., Kooiker, M., Borghi, L., Ditta, G., Yanofsky, M.F., Kater, M.M., and Colombo, L. (2003). MADS-box protein complexes control carpel and ovule development in Arabidopsis. Plant Cell **15,** 2603-2611.

Fischle, W., Wang, Y., Jacobs, S.A., Kim, Y., Allis, C.D., and Khorasanizadeh, S. (2003). Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. Genes Dev **17,** 1870-1881.

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

Fransz, P., Soppe, W., and Schubert, I. (2003). Heterochromatin in interphase nuclei of Arabidopsis thaliana. Chromosome Res **11,** 227-240.

Gaudin, V., Libault, M., Pouteau, S., Juul, T., Zhao, G., Lefebvre, D., and Grandjean, O. (2001). Mutations in LIKE HETEROCHROMATIN PROTEIN 1 affect flowering time and plant architecture in Arabidopsis. Development **128,** 4847-4858.

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

Gomez-Mena, C., de Folter, S., Costa, M.M., Angenent, G.C., and Sablowski, R. (2005). Transcriptional program controlled by the floral homeotic gene AGAMOUS during early organogenesis. Development **132,** 429-438.

Goodrich, J., and Tweedie, S. (2002). Remembrance of things past: chromatin remodeling in plant development. Annu Rev Cell Dev Biol **18,** 707-746.

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

Grewal, S.I., and Jia, S. (2007). Heterochromatin revisited. Nat Rev Genet **8,** 35-46.

Grimaud, C., Bantignies, F., Pal-Bhadra, M., Ghana, P., Bhadra, U., and Cavalli, G. (2006). RNAi components are required for nuclear clustering of Polycomb group response elements. Cell **124,** 957-971.

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

Guitton, A.E., and Berger, F. (2005). Loss of function of MULTICOPY SUPPRESSOR OF IRA 1 produces nonviable parthenogenetic embryos in Arabidopsis. Curr Biol **15,** 750-754.

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

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

Hepworth, S.R., Valverde, F., Ravenscroft, D., Mouradov, A., and Coupland, G. (2002). Antagonistic regulation of flowering-time gene SOC1 by CONSTANS and FLC via separate promoter motifs. Embo J **21,** 4327-4337.

Honma, T., and Goto, K. (2001). Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. Nature **409,** 525-529.

Hsieh, T.F., and Fischer, R.L. (2005). Biology of chromatin dynamics. Annu Rev Plant Biol **56,** 327-351.

Huang, D.H., Chang, Y.L., Yang, C.C., Pan, I.C., and King, B. (2002). pipsqueak encodes a factor essential for sequence-specific targeting of a polycomb group protein complex. Mol Cell Biol **22,** 6261-6271.

Huisinga, K.L., Brower-Toland, B., and Elgin, S.C. (2006). The contradictory definitions of heterochromatin: transcription and silencing. Chromosoma **115,** 110-122.

Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. Science **293,** 1074-1080.

Jiang, D., Wang, Y., Wang, Y., and He, Y. (2008). Repression of FLOWERING LOCUS C and FLOWERING LOCUS T by the Arabidopsis Polycomb repressive complex 2 components. PLoS ONE **3,** e3404.

Johanson, U., West, J., Lister, C., Michaels, S., Amasino, R., and Dean, C. (2000). Molecular analysis of FRIGIDA, a major determinant of natural variation in Arabidopsis flowering time. Science **290,** 344-347.

Jullien, P.E., Katz, A., Oliva, M., Ohad, N., and Berger, F. (2006). Polycomb group complexes self-regulate imprinting of the Polycomb group gene MEDEA in Arabidopsis. Curr Biol **16,** 486-492.

Jurgens, G. (1985). A group of genes controlling the spatial expression of the bithorax complex in Drosophila. . Nature **316,** 153-155.

Kankel, M.W., Ramsey, D.E., Stokes, T.L., Flowers, S.K., Haag, J.R., Jeddeloh, J.A., Riddle, N.C., Verbsky, M.L., and Richards, E.J. (2003). Arabidopsis MET1 cytosine methyltransferase mutants. Genetics **163,** 1109-1122.

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

Kennison, J.A. (1995). The Polycomb and trithorax group proteins of Drosophila: trans-regulators of homeotic gene function. Annu Rev Genet **29,** 289-303.

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

Klymenko, T., Papp, B., Fischle, W., Kocher, T., Schelder, M., Fritsch, C., Wild, B., Wilm, M., and Muller, J. (2006). A Polycomb group protein complex with sequence-specific DNA-binding and selective methyl-lysine-binding activities. Genes Dev **20,** 1110-1122.

Kohler, C., Page, D.R., Gagliardini, V., and Grossniklaus, U. (2005). The Arabidopsis thaliana MEDEA Polycomb group protein controls expression of PHERES1 by parental imprinting. Nat Genet **37,** 28-30.

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

Komeda, Y. (2004). Genetic regulation of time to flower in Arabidopsis thaliana. Annu Rev Plant Biol **55,** 521-535.

Kotake, T., Takada, S., Nakahigashi, K., Ohto, M., and Goto, K. (2003). Arabidopsis TERMINAL FLOWER 2 gene encodes a heterochromatin protein 1 homolog and represses both FLOWERING LOCUS T to regulate flowering time and several floral homeotic genes. Plant Cell Physiol **44,** 555-564.

Krizek, B.A., and Fletcher, J.C. (2005). Molecular mechanisms of flower development: an armchair guide. Nat Rev Genet **6,** 688-698.

Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (2002). Histone methyltransferase activity associated with a human

multiprotein complex containing the Enhancer of Zeste protein. Genes Dev **16,** 2893-2905.

Laubinger, S., Zeller, G., Henz, S.R., Sachsenberg, T., Widmer, C.K., Naouar, N., Vuylsteke, M., Scholkopf, B., Ratsch, G., and Weigel, D. (2008). At-TAX: a whole genome tiling array resource for developmental expression analysis and transcript identification in Arabidopsis thaliana. Genome Biol **9,** R112.

Levy, Y.Y., Mesnage, S., Mylne, J.S., Gendall, A.R., and Dean, C. (2002). Multiple roles of Arabidopsis VRN1 in vernalization and flowering time control. Science **297,** 243-246.

Li, D., Liu, C., Shen, L., Wu, Y., Chen, H., Robertson, M., Helliwell, C.A., Ito, T., Meyerowitz, E., and Yu, H. (2008). A repressor complex governs the integration of flowering signals in Arabidopsis. Dev Cell **15,** 110-120.

Lindroth, A.M., Shultis, D., Jasencakova, Z., Fuchs, J., Johnson, L., Schubert, D., Patnaik, D., Pradhan, S., Goodrich, J., Schubert, I., Jenuwein, T., Khorasanizadeh, S., and Jacobsen, S.E. (2004). Dual histone H3 methylation marks at lysines 9 and 27 required for interaction with CHROMOMETHYLASE3. Embo J **23,** 4286-4296.

Lohmann, J.U., and Weigel, D. (2002). Building beauty: the genetic control of floral patterning. Dev Cell **2,** 135-142.

Maison, C., and Almouzni, G. (2004). HP1 and the dynamics of heterochromatin maintenance. Nat Rev Mol Cell Biol **5,** 296-304.

Makarevich, G., Leroy, O., Akinci, U., Schubert, D., Clarenz, O., Goodrich, J., Grossniklaus, U., and Kohler, C. (2006). Different Polycomb group complexes regulate common target genes in Arabidopsis. EMBO Rep **7,** 947-952.

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

Michaels, S.D., and 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. Plant Cell **13,** 935-941.

Milne, T.A., Briggs, S.D., Brock, H.W., Martin, M.E., Gibbs, D., Allis, C.D., and Hess, J.L. (2002). MLL targets SET domain methyltransferase activity to Hox gene promoters. Mol Cell **10,** 1107-1117.

Min, J., Zhang, Y., and Xu, R.M. (2003). Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. Genes Dev **17,** 1823-1828.

Minasov, G., Teplova, M., Stewart, G.C., Koonin, E.V., Anderson, W.F., and Egli, M. (2000). Functional implications from crystal structures of the conserved Bacillus subtilis protein Maf with and without dUTP. Proc Natl Acad Sci U S A **97,** 6328-6333.

Mizukami, Y., and Ma, H. (1992). Ectopic expression of the floral homeotic gene AGAMOUS in transgenic Arabidopsis plants alters floral organ identity. Cell **71,** 119-131.

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

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

Nakahigashi, K., Jasencakova, Z., Schubert, I., and Goto, K. (2005). The Arabidopsis heterochromatin protein1 homolog (TERMINAL FLOWER2) silences genes within the euchromatic region but not genes positioned in heterochromatin. Plant Cell Physiol **46,** 1747-1756.

Ogas, J., Kaufmann, S., Henderson, J., and Somerville, C. (1999). PICKLE is a CHD3 chromatin-remodeling factor that regulates the transition from embryonic to vegetative development in Arabidopsis. Proc Natl Acad Sci U S A **96,** 13839-13844.

Oh, S., Park, S., and van Nocker, S. (2008). Genic and global functions for Paf1C in chromatin modification and gene expression in Arabidopsis. PLoS Genet **4,** e1000077.

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

Parcy, F. (2005). Flowering: a time for integration. Int J Dev Biol **49,** 585-593.

Parenicova, L., de Folter, S., Kieffer, M., Horner, D.S., Favalli, C., Busscher, J., Cook, H.E., Ingram, R.M., Kater, M.M., Davies, B., Angenent, G.C., and Colombo, L. (2003). Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in Arabidopsis: new openings to the MADS world. Plant Cell **15,** 1538-1551.

Pelaz, S., Tapia-Lopez, R., Alvarez-Buylla, E.R., and Yanofsky, M.F. (2001a). Conversion of leaves into petals in Arabidopsis. Curr Biol **11,** 182-184.

Pelaz, S., Ditta, G.S., Baumann, E., Wisman, E., and Yanofsky, M.F. (2000). B and C floral organ identity functions require SEPALLATA MADS-box genes. Nature **405,** 200-203.

Pelaz, S., Gustafson-Brown, C., Kohalmi, S.E., Crosby, W.L., and Yanofsky, M.F. (2001b). APETALA1 and SEPALLATA3 interact to promote flower development. Plant J **26,** 385-394.

Perez-Callejon, E., Casamayor, A., Pujol, G., Camps, M., Ferrer, A., and Arino, J. (1998). Molecular cloning and characterization of two phosphatase 2A catalytic subunit genes from Arabidopsis thaliana. Gene **209,** 105-112.

Pfluger, J., and Wagner, D. (2007). Histone modifications and dynamic regulation of genome accessibility in plants. Curr Opin Plant Biol **10,** 645-652.

Pien, S., and Grossniklaus, U. (2007). Polycomb group and trithorax group proteins in Arabidopsis. Biochim Biophys Acta **1769,** 375-382.

Quesada, V., Dean, C., and Simpson, G.G. (2005). Regulated RNA processing in the control of Arabidopsis flowering. Int J Dev Biol **49,** 773-780.

Reyes, J.C. (2006). Chromatin modifiers that control plant development. Curr Opin Plant Biol **9,** 21-27.

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

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

Rounsley, S.D., Ditta, G.S., and Yanofsky, M.F. (1995). Diverse roles for MADS box genes in Arabidopsis development. Plant Cell **7,** 1259-1269.

Sanchez-Pulido, L., Devos, D., Sung, Z.R., and Calonje, M. (2008). RAWUL: a new ubiquitin-like domain in PRC1 ring finger proteins that unveils putative plant and worm PRC1 orthologs. BMC Genomics **9,** 308.

Savidge, B., Rounsley, S.D., and Yanofsky, M.F. (1995). Temporal relationship between the transcription of two Arabidopsis MADS box genes and the floral organ identity genes. Plant Cell **7,** 721-733.

Schomburg, F.M., Patton, D.A., Meinke, D.W., and Amasino, R.M. (2001). FPA, a gene involved in floral induction in Arabidopsis, encodes a protein containing RNA-recognition motifs. Plant Cell **13,** 1427-1436.

Schonrock, N., Bouveret, R., Leroy, O., Borghi, L., Kohler, C., Gruissem, W., and Hennig, L. (2006). Polycomb-group proteins repress the floral activator AGL19 in the FLC-independent vernalization pathway. Genes Dev **20,** 1667-1678.

Schubert, D., Clarenz, O., and Goodrich, J. (2005). Epigenetic control of plant development by Polycomb-group proteins. Curr Opin Plant Biol **8,** 553-561.

Schubert, D., Primavesi, L., Bishopp, A., Roberts, G., Doonan, J., Jenuwein, T., and Goodrich, J. (2006). Silencing by plant Polycomb-group genes requires dispersed trimethylation of histone H3 at lysine 27. Embo J **25,** 4638-4649.

Schwartz, Y.B., and Pirrotta, V. (2007). Polycomb silencing mechanisms and the management of genomic programmes. Nat Rev Genet **8,** 9-22.

Schwartz, Y.B., and Pirrotta, V. (2008). Polycomb complexes and epigenetic states. Curr Opin Cell Biol **20,** 266-273.

Schwartz, Y.B., Kahn, T.G., Nix, D.A., Li, X.Y., Bourgon, R., Biggin, M., and Pirrotta, V. (2006). Genome-wide analysis of Polycomb targets in Drosophila melanogaster. Nat Genet **38,** 700-705.

Searle, I., He, Y., Turck, F., Vincent, C., Fornara, F., Krober, S., Amasino, R.A., and Coupland, G. (2006). The transcription factor FLC confers a flowering response to vernalization by repressing meristem competence and systemic signaling in Arabidopsis. Genes Dev **20,** 898-912.

Shao, Z., Raible, F., Mollaaghababa, R., Guyon, J.R., Wu, C.T., Bender, W., and Kingston, R.E. (1999). Stabilization of chromatin structure by PRC1, a Polycomb complex. Cell **98,** 37-46.

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

Siebert, P.D., Chenchik, A., Kellogg, D.E., Lukyanov, K.A., and Lukyanov, S.A. (1995). An improved PCR method for walking in uncloned genomic DNA. Nucleic Acids Res **23,** 1087-1088.

Simpson, G.G. (2004). The autonomous pathway: epigenetic and post-transcriptional gene regulation in the control of Arabidopsis flowering time. Curr Opin Plant Biol **7,** 570-574.

Simpson, G.G., Dijkwel, P.P., Quesada, V., Henderson, I., and Dean, C. (2003). FY is an RNA 3' end-processing factor that interacts with FCA to control the Arabidopsis floral transition. Cell **113,** 777-787.

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

Struhl, G. (1982). Genes controlling segmental specification in the Drosophila thorax. Proc Natl Acad Sci U S A **79,** 7380-7384.

Struhl, G., and Brower, D. (1982). Early role of the esc+ gene product in the determination of segments in Drosophila. Cell **31,** 285-292.

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

Teper-Bamnolker, P., and Samach, A. (2005). The flowering integrator FT regulates SEPALLATA3 and FRUITFULL accumulation in Arabidopsis leaves. Plant Cell **17,** 2661-2675.

Terzi, L.C., and Simpson, G.G. (2008). Regulation of flowering time by RNA processing. Curr Top Microbiol Immunol **326,** 201-218.

Trojer, P., and Reinberg, D. (2007). Facultative heterochromatin: is there a distinctive molecular signature? Mol Cell **28,** 1-13.

Turck, F., Roudier, F., Farrona, S., Martin-Magniette, M.L., Guillaume, E., Buisine, N., Gagnot, S., Martienssen, R.A., Coupland, G., and Colot, V. (2007). Arabidopsis TFL2/LHP1 specifically associates with genes marked by trimethylation of histone H3 lysine 27. PLoS Genet **3,** e86.

Tzfira, T., Li, J., Lacroix, B., and Citovsky, V. (2004). Agrobacterium T-DNA integration: molecules and models. Trends Genet **20,** 375-383.

Wang, D., Tyson, M.D., Jackson, S.S., and Yadegari, R. (2006). Partially redundant functions of two SET-domain polycomb-group proteins in controlling initiation of seed development in Arabidopsis. Proc Natl Acad Sci U S A **103,** 13244-13249.

Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R.S., and Zhang, Y. (2004). Role of histone H2A ubiquitination in Polycomb silencing. Nature **431,** 873-878.

Weigel, D., Ahn, J.H., Blazquez, M.A., Borevitz, J.O., Christensen, S.K., Fankhauser, C., Ferrandiz, C., Kardailsky, I., Malancharuvil, E.J., Neff, M.M., Nguyen, J.T., Sato, S., Wang, Z.Y., Xia, Y., Dixon, R.A., Harrison, M.J., Lamb, C.J., Yanofsky, M.F., and Chory, J. (2000). Activation tagging in Arabidopsis. Plant Physiol **122,** 1003-1013.

Yanofsky, M.F., Ma, H., Bowman, J.L., Drews, G.N., Feldmann, K.A., and Meyerowitz, E.M. (1990). The protein encoded by the Arabidopsis homeotic gene agamous resembles transcription factors. Nature **346,** 35-39.

Yoo, S.K., Chung, K.S., Kim, J., Lee, J.H., Hong, S.M., Yoo, S.J., Yoo, S.Y., Lee, J.S., and Ahn, J.H. (2005). CONSTANS activates SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 through FLOWERING LOCUS T to promote flowering in Arabidopsis. Plant Physiol **139,** 770-778.

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

Zahn, L.M., Kong, H., Leebens-Mack, J.H., Kim, S., Soltis, P.S., Landherr, L.L., Soltis, D.E., Depamphilis, C.W., and Ma, H. (2005). The evolution of the SEPALLATA subfamily of MADS-box genes: a preangiosperm origin with multiple duplications throughout angiosperm history. Genetics **169,** 2209-2223.

Zhang, X., Germann, S., Blus, B.J., Khorasanizadeh, S., Gaudin, V., and Jacobsen, S.E. (2007a). The Arabidopsis LHP1 protein colocalizes with histone H3 Lys27 trimethylation. Nat Struct Mol Biol **14,** 869-871.

Zhang, X., Clarenz, O., Cokus, S., Bernatavichute, Y.V., Pellegrini, M., Goodrich, J., and Jacobsen, S.E. (2007b). Whole-genome analysis of histone H3 lysine 27 trimethylation in Arabidopsis. PLoS Biol **5,** e129.

Zhao, J., Sun, B.K., Erwin, J.A., Song, J.J., and Lee, J.T. (2008). Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science **322,** 750-756.