HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Functional analysis of SCI1 – A PWWP domain protein involved in Polycomb group mediated gene regulation in *Arabidopsis*

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Mareike L. Hohenstatt

aus Boulder

Düsseldorf, Mai 2012

aus dem Institut für Genetik
der Heinrich-Heine Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Referent: Dr. D. Schubert
Korreferent: Professor Dr. R. Simon

Tag der mündlichen Prüfung: 06.06.2012

Part of this work was published in the following publications:

Lafos,M., Kroll,P., Hohenstatt,M.L., Thorpe,F.L., Clarenz,O., and Schubert,D. (2011). Dynamic regulation of H3K27 trimethylation during Arabidopsis differentiation. PLoS. Genet. *7*, e1002040.

Schatlowski,N., Stahl,Y., Hohenstatt,M.L., Goodrich,J., and Schubert,D. (2010). The CURLY LEAF interacting protein BLISTER controls expression of polycomb-group target genes and cellular differentiation of Arabidopsis thaliana. Plant Cell *22*, 2291-2305.

TABLE OF CONTENTS

1 INTRODUCTION 4

1.1 EPIGENETICS 4

1.2 CHROMATIN STRUCTURE AND HISTONE MODIFICATIONS 4

1.3 POLYCOMB GROUP PROTEINS ARE KEY REGULATORS OF EUKARYOTIC DEVELOPMENT .. 6

1.3.1 PcG complexes and gene regulation in *Drosophila* 7

1.3.2 PcG complexes and gene regulation in *Arabidopsis* 9

1.4 FLOWERING TIME REGULATION BY POLYCOMB GROUP PROTEINS 15

1.5 PROTEIN DOMAINS IMPLICATED IN BINDING OF HISTONE MODIFICATIONS 18

1.6 AIMS OF THIS STUDY 21

2 MATERIALS AND METHODS 22

2.1 MATERIALS 22

2.1.1 Chemicals 22

2.1.2 Buffer 22

2.1.3 Media 22

2.1.4 Enzymes 22

2.1.5 Antibodies and peptides 23

2.1.6 Molecular size standards 23

2.1.7 Membrane and Paper 23

2.1.8 Oligonucleotides 23

2.1.9 Constructs and Vectors 26

2.1.10 Microorganism 29

2.1.11 Plants 29

2.1.12 Software 30

2.2 METHODS 31

2.2.1 Basic molecular methods 31

2.2.1.1 Cloning and vector generation 31

2.2.2 Quantitative real time PCR (qRT-PCR) 31

2.2.3 Bacteria methods 31

2.2.3.1 Growth conditions and transformation 31

2.2.3.2 Preparation of plasmid DNA from *E. coli* 31

2.2.4 Yeast methods 32

2.2.4.1 Growth conditions 32

2.2.4.2 Yeast two-hybrid interaction studies 32

2.2.5 Plant methods 32

2.2.5.1 Plant growth conditions 32

2.2.5.2 Seed sterilization 32

2.2.5.3 Plant transformation 32
2.2.5.4 Isolation of genomic DNA from *Arabidopsis* 33
2.2.5.5 RNA extraction and cDNA synthesis 33
2.2.5.6 ß-Glucuronidase assay 33
2.2.5.7 Flowering time analysis 33
2.2.5.8 Root length measurement 33
2.2.5.9 Plant nuclei isolation for nuclear size measurements 33
2.2.5.10 Ploidy analyses 34
2.2.6 Protein extraction and analysis from *E. coli* and *Arabidopsis* 34
2.2.6.1 Protein expression and purification from E. coli 34
2.2.6.2 Histone enriched protein extraction from *Arabidopsis thaliana* 34
2.2.6.3 Immunoblot procedures 34
2.2.6.4 NuclearProtein extraction, immunoprecipitation and LC-MS/MS analysis 35
2.2.7 Chromatin Immunoprecipitation (ChIP) 36
2.2.8 Immunoprecipitation of GST-fusion proteins with biotinylated peptides 36
2.2.9 ModifiedTM Histone Peptide Array 36
2.2.10 Luciferase assays 36
2.2.11 Microscopic analyses 37

3 RESULTS 38

3.1 SCI1 IS A PLANT SPECIFIC PWWP DOMAIN PROTEIN 39
3.2 YEAST TWO-HYBRID INTERACTION STUDIES WITH SCI1 40
3.2.1 SCI1 was isolated in a yeast-two-hybrid screen with the PRC2 component CLF 40
3.2.2 SCI1 interacts with several PRC2 components in yeast 40
3.2.3 Quantification of the SCI1 yeast two-hybrid interactions via ONPG assay 42
3.3 *IN PLANTA* ANALYSIS OF THE SCI1 PROTEIN 42
3.3.1 Intracellular localization of SCI1, CLF and SWN *in planta* 42
3.3.2 Interaction studies of SCI1 with CLF and SWN *in planta* 45
3.3.3 Immunoprecipitation with SCI1-GFP reveals possible complex partners of SCI1 46
3.4 ANALYSIS OF *SCI1* MUTANT ALLELES IN *ARABIDOPSIS THALIANA* 49
3.5 EXPRESSION ANALYSIS OF *SCI1, SCI2* AND *SCI3 IN PLANTA* 51
3.6 ANALYSIS OF FLOWERING TIME AND ROOT PHENOTYPES IN *SCI1* MUTANTS 52
3.6.1 Loss of SCI1 function results in an early flowering phenotype 52
3.6.2 *FLC* expression is downregulated in *sci1-1* mutant background 54
3.6.3 Primary root growth is reduced in *sci1-1* mutants 56
3.7 SCI1 ACTS REDUNDANTLY WITH SCI2 AND SCI3 57
3.7.1 Phenotypic analysis of *SCI* double and triple mutants 58
3.7.2 *SCI* triple mutants are seedling lethal 61
3.8 GENETIC INTERACTION OF *SCI1* WITH PRC2 COMPONENT *CLF* 62

3.8.1 Phenotypic and flowering time analysis of *sci1 clf-28* double mutants 62
3.8.2 SCI1 dependent regulation of PcG target genes 64
3.8.3 ChIP analysis of histone methylation levels in *sci1-1* mutants 68
3.8.4 Global levels of histone modifications in *sci1-1 clf-28* and other PcG mutants 69
3.9 SCI1 REGULATES NUCLEAR SIZE SIMILAR TO LINC1 72
3.10 FUNCTIONAL ANALYSES OF THE SCI1 PWWP-DOMAIN 74
3.10.1 The SCI1 PWWP domain exhibits in vitro histone peptide binding capacity 75

4 DISCUSSION 78

4.1 SCI1 INTERACTS WITH CORE MEMBERS OF THE *ARABIDOPSIS* PRC2 COMPLEX 79
4.2 SUB-CELLULAR LOCALIZATION OF SCI1 *IN PLANTA* 80
4.3 *IN PLANTA* CO-LOCALIZATION AND INTERACTION STUDIES 83
4.4 SCI1 AND ITS INTERACTION PARTNER LINC1 REGULATE NUCLEAR SIZE IN *ARABIDOPSIS* 85
4.5 SCI1 GENETICALLY INTERACTS WITH CLF 86
4.5.1 The SCI1 expression pattern overlaps with those of CLF, SWN and MEA 86
4.5.2 SCI1 and its homologs SCI2 and SCI3 exhibit redundant functions 87
4.5.3 Mutation of *SCI1* severely enhances the *clf* phenotype 88
4.6 SCI1 IS A NOVEL REGULATOR OF *FLC* 91
4.6.1 Analysis of histone modification levels in *sci1-1* and PcG mutants 92
4.7 FUNCTIONAL ANALYSIS OF THE SCI1 PWWP DOMAIN 94
4.8 CONCLUDING REMARKS 97

5 SUMMARY 100

5.1 ZUSAMMENFASSUNG 102

6 LITERATURE 104

7 APPENDIX 120

7.1 ABBREVIATIONS 120
7.2 LIST OF FIGURES 121
7.3 LIST OF TABLES 122
7.4 LIST OF PUTATIVE COMPLEX PARTNERS OF SCI1 123

8 EIDESSTATTLICHE ERKLÄRUNG 125

9 ACKNOWLEDGEMENTS 126

1 INTRODUCTION

1.1 Epigenetics

Genetic information in eukaryotic organisms is encoded in the macromolecule desoxyribonucleic acid (DNA) (Avery et al., 1944), and passed on through the germline to the next generation. Many phenotypic traits are caused by allelic differences due to mutations in the DNA sequence, and mendelian inheritance depends on the equal distribution of the parental alleles to the gametes and hence to the progeny.

In the more recent past additional levels of inheritance were discovered, that do not depend on changes of the DNA sequence. These 'epigenetic' phenomena are defined as "mitotically and/or meiotically heritable changes in gene function that cannot be explained by changes in DNA sequence" (Russo et al., 1996). These epigenetic phenomena are for example caused by post-translational histone modifications including histone methylation, acetylation, phosphorylation and ubiquitination. Furthermore DNA methylation, RNA based mechanisms, nucleosome exchanges, and prions were described to play a role in the establishment and maintenance of epigenetic states (Halfmann and Lindquist, 2010; Feng et al., 2010). These modifications are involved in defining the expression state of genes, and maintaining it through several rounds of cell division, or even to the next generation. The advantage of epigenetic gene regulation is that an organism can respond far more flexible to environmental and endogenous cues, as the mediated modifications are stable but not irreversible.

Epigenetic gene regulation is highly important for the development of eukaryotic organisms as it was reported to play a decisive role in processes like gene silencing, position effect variegation (PEV), imprinting, paramutation, transgene silencing, X-chromosome inactivation, parent-of origin effects, cell differentiation and reprogramming and therefore also in diseases like cancer (Feng et al., 2010; Berdasco and Esteller, 2010; Avramova, 2009). Most of these processes originate in the modification of chromatin, a highly complex structure of DNA and associated proteins.

1.2 Chromatin structure and histone modifications

In eukaryotic cells the DNA has to be highly organized in order to fit into the nucleus. The nucleosomal core particle contains 147 bp of DNA, which is wrapped in 1,65 turns around a histone octamer. This octamer consists of a central heterotetramer of the histones H3 and H4 and is flanked by two heterodimers of H2A and H2B. The nucleosomal core particles are

separated by 10-60 bp of 'linker-DNA', and this conformation is called the 10 nm chromatin fiber (Peterson and Laniel, 2004; Luger et al., 1997). The linker-DNA is associated with histone H1 (linker histone) which stabilizes a further compaction of the DNA to ~30 nm thick fibers through interaction with the nucleosomal core particle. The nucleosomal core particle, linker DNA and histone H1 together, represent the basic repeat unit of chromatin, the nucleosome (Happel and Doenecke, 2009). Further condensation of the chromatin to 100-400 nm fibers in interphase chromatin and even higher compaction in metaphase chromosomes can be observed *in vivo* (Peterson and Laniel, 2004).

Each core histone consists of a C-terminal histone fold domain and an N-terminal tail domain, that is rich in basic amino acids, positively charged and extends from the surface of the nucleosome (Peterson and Laniel, 2004; Hansen, 2002). These histone tails are involved in controlling the folding of chromatin into higher order structures, as the removal of the histone tails *in vitro* prevents a compaction beyond the 10 nm fiber (Hansen, 2002). Furthermore the amino-terminal tails of the core histones are subject to extensive post-translational modifications that influence chromatin structure as well as transcriptional gene activity. Since nucleosomal DNA is not freely accessible to regulatory proteins, chromatin structure has to be changed dynamically upon endo- or exogenous cues. For example the covalent modifications on histone tails, replacement of histones, histone variants and DNA cytosine methylation alter the histone-DNA interaction and create or block protein binding sites (Pfluger and Wagner, 2007).

The modifications discovered on histone tails by mass spectrometry and specific antibodies are acetylation, methylation, phosphorylation, ubiquitination, sumoylation, ADP ribosylation, deimination and proline isomeration, although not all histone modifications are implicated in epigenetic gene regulation (Kouzarides, 2007; Fuchs et al., 2006). The most diverse of these modifications is methylation, as arginine residues (R) are found to be mono- and di-methylated whereas lysine residues (K) are mono-, di- and tri-methylated. Depending on the specific lysine or arginine residue and the number of methyl groups added, these modifications can be either associated with transcriptional activation or repression (Kouzarides, 2007). While methylation of histone tails has mainly been connected to transcriptional regulation, phosphorylation of histone tails at serine (S) and threonine (T) residues is involved in a multitude of pathways, among them chromosome condensation, cell division, transcriptional regulation and DNA damage repair. There is evidence that phosphorylation serves as a platform for recruitment and release of effector proteins, and thus establishes a crosstalk between histone modifications (Banerjee and Chakravarti, 2011). For

instance it was shown in mammalian cell culture experiments that H3S28 phosphorylation on the silenced PcG target gene α-globulin leads to a reduction of H3 lysine27 tri-methylation (H3K27me3) and an induction of H3K27 acetylation (H3K27ac) which results in gene activation (Lau and Cheung, 2011). This example clearly demonstrates the complexity achieved by the combination of several histone modifications.

On each of the core histones several amino acids are subject to post-translational modifications, and many of the modifications have been associated either with an active or repressive state of chromatin. While the effect of most individual modifications is conserved between plants and animals, each kingdom also harbors unique modifications (Pfluger and Wagner, 2007; Zhang et al., 2007b).

It was proposed that the combination and/or sequential arrangement of histone modifications on one or more histone tails can be regarded as the so called 'histone code', which can be read by other proteins leading to a number of different downstream events (Strahl and Allis, 2000). More recent studies have revealed that not only the combination of histone modifications have to be taken into account, as in different contexts and timing the same modification can be involved in gene activation or repression. Furthermore all modifications on histone tails are also removable which adds to the complexity of the system (Alvarez-Venegas, 2010). It is therefore not sufficient to map histone modification patterns on genes of interest, but essential to simultaneously explore the recruitment, regulation and interactions of the complexes that mediate the modifications (Lee et al., 2010).

1.3 Polycomb group proteins are key regulators of eukaryotic development

Among the proteins involved in epigenetic gene regulation, Polycomb group (PcG) proteins are key regulators which are of great importance for the development of eukaryotic organisms (Grimaud et al., 2006). In a variety of organisms PcG and Trithorax group (TrxG) proteins have been thoroughly analyzed. Initially identified in *Drosophila melanogaster*, PcG proteins are described to maintain a repressive state of homeotic (Hox) gene expression, genes that control embryonic positional identities, while the antagonistically acting TrxG proteins maintain Hox gene activity. PcG genes were named after the *Drosophila* protein *Polycomb* (*Pc*), as flies with a mutation *Pc* display global transformations of embryonic segments into the most posterior segment which results in lethality. Hence PcG genes were defined as genes that, when individually mutated, show homeotic transformations similar to those observed in *Polycomb* (*Pc*) mutants or which can enhance the phenotype of *Pc* mutant alleles (Simon and Tamkun, 2002).

1.3.1 PcG complexes and gene regulation in *Drosophila*

The molecular analysis of PcG and TrxG proteins revealed that both groups act as large multimeric complexes indirectly regulating transcription. This is accomplished by deposition of post-translational modifications, which results in modulation of chromatin structure (Grimaud et al., 2006). In case of PcG proteins, initially two main complexes were identified and analyzed, POLYCOMB REPRESSIVE COMPLEX 1 (PRC1) and PRC2. The biochemical purification of PRC1 identified the core subunits of the complex including Polycomb (Pc), Polyhomeotic (Ph), Posterior sex combs (Psc) and dRing1. Additional proteins like Sex combs on midleg (Scm), Zeste (z) and general transcription factors are also associated with PRC1 and are thought to be exchanged to obtain specificity of the complex at certain developmental time points (Shao et al., 1999; Saurin et al., 2001).

PRC2 consists of four core members in *Drosophila*, Enhancer of zeste (E(z)), Extra sex combs (Esc), p55 and Suppressor of zeste12 (Su(z)12). The PRC2 complex mediates the tri-methylation of H3K27 through its subunit E(z) that contains a catalytic SET domain. The functionality of the SET domain as well as the presence of Esc is essential for the tri-methylation of H3K27 (Cao and Zhang, 2004; Czermin et al., 2002; Muller et al., 2002; Cao et al., 2002; Kuzmichev et al., 2002). Furthermore it was observed that the repressive tri-methylation of H3K27 is stimulated by already pre-existing H3K27me3 marks, while it is inhibited by tri-methylation of H3K4 and di- and tri-methylation of H3K36, modifications associated with gene activity (Schmitges et al., 2011). This finding illustrates the interplay of opposing chromatin modifications that ensures that the correct target genes are methylated.

The H3K27me3 mark is specifically bound by the PRC1 subunit Pc via its chromodomain, which leads to an inhibition of nucleosome remodeling and transcription, and to a compaction of chromatin *in vitro* (Fischle et al., 2003c; Shao et al., 1999; Lo and Francis, 2010; Francis et al., 2004).

The PRC1-like complex dRing associated factors (dRAF) includes dRing1 and Psc in *Drosophila* and the subunits Ring1A and Bmi-1 in mammals that comprise catalytic ring-domains and have been shown to act as ubiquitin E3 ligases mediating the mono-ubiquitination of lysine residue 119 of histone H2A (H2A K119ub) (Muller and Verrijzer, 2009; de Napoles et al., 2004; Cao et al., 2005). Mutation of Bmi-1 did not affect the level of H3K27me3, while the disruption of PRC2 catalytic activity resulted in the loss of H3K27me3 and H2A K119ub. This result places PRC2 upstream of PRC1 recruitment, H2A ubiquitination and gene silencing. H2A ubiquitination might either directly or indirectly block

the recruitment of the basic transcriptional machinery to promoters, which in turn would result in stable gene silencing (Cao et al., 2005; Muller and Verrijzer, 2009).

More recently an additional Polycomb complex was identified that harbors ubiquitin hydrolase activity with specificity towards H2A monoubiquitination. The Polycomb repressive deubiquitinase (PR-DUB) complex discovered in *Drosophila* consists of the ubiquitin hydrolase Calypso and Additional sex combs (Asc), and is essential for the repression of PcG targent genes (Scheuermann et al., 2010). Thus it seems to be critical to tightly balance H2A monoubiquitination levels in order to maintain the PcG mediated repressed chromatin state. This applies at least to a subset of PcG target genes, as a recent study could show by mutation of the dRing subunit, that there are two classes of PcG targets of which only one is dependent on H2A ubiquitination (Scheuermann et al., 2010; Gutierrez et al., 2012).

The fourth PcG complex Pleiohomeotic repressive complex (PhoRC) includes the proteins Pleiohomeotic (Pho) and dSfmbt (Smc-related gene containing four malignant brain tumor (MBT) domains) and was identified to be essential for HOX gene silencing in *Drosophila*. The Pho subunit comprises sequence specific DNA binding ability, while the SFmbt subunit was shown to bind mono- and dimethylated H3K9 and H4K20 residues (Klymenko et al., 2006). In *Drosophila* the recruitment of the PcG machinery is achieved through cis-acting DNA elements, the Polycomb response elements (PRE), which are required for PcG mediated HOX gene silencing (Chan et al., 1994; Simon et al., 1993). These PREs contain binding sites for Pho and Pho-like (Phol) (Brown et al., 2003; Brown et al., 1998), and it has been shown that PRC1, PRC2 and PhoRC are associated with these sites (Papp and Muller, 2006). Furthermore Pho was shown to directly bind to E(z) and Esc and Phol to Esc which might lead to a stabilization of PRC2 binding to target chromatin (Brown et al., 2003; Wang et al., 2004)

It was discovered that from the PRE sequences long non-coding RNA (ncRNA) transcripts are produced which are probably involved in the recruitment of PcG complexes to its target genes (Hekimoglu and Ringrose, 2009; Margueron and Reinberg, 2011). One prominent example of ncRNA recruitment is the trans-acting HOTAIR ncRNA which interacts with PRC2 and is essential for PcG mediated gene silencing of the HOXD gene cluster in fibroblast cells (Rinn et al., 2007).

Intracellular localization studies of PcG proteins in *Drosophila* revealed that they are concentrated in nuclear foci which are called PcG bodies (Pirrotta and Li, 2011). The amount and size of these PcG bodies is variable and dependent on the cell type. One possible

explanation for this specific localization is the fact that PcG binding sites in the *Drosophila* genome tend to occur in linear clusters. But there is also evidence for interaction of PcG target genes that map at large distances from each other, and of the involvement of PREs which are also able to interact. Insulators are known to mediate chromatin looping and might play a role as well as ncRNAs, which could serve as scaffolds for higher order assembly of PcG proteins (Pirrotta and Li, 2011). The PcG bodies were shown to move, meet and split dynamically during development in agreement with the finding that PcG target genes are not stably associated with PcG bodies (Cheutin and Cavalli, 2012; Pirrotta and Li, 2011).

The concentration of PcG proteins at specific regions within the nucleus appears to be furthermore regulated by additional proteins in a cell type dependent manner. In myoblast cells the homeodomain protein Msx1 specifically localizes to the nuclear periphery. It was shown that Msx1 associates with the PRC2 complex and promotes the spatial redistribution of the repressive H3K27me3 to the nuclear periphery which results in the cell type specific silencing of Msx1 target genes (Wang et al., 2011).

PcG mediated gene regulation is tightly controlled by the crosstalk of the different complexes, the timing of modification setting and removal, and by additional factors like long ncRNAs conferring specificity toward the large amount of target genes at different developmental time points. Furthermore the counteracting PcG and Trx systems regulate each other as the mediated modifications oppose the activity of the antagonistic system (Klymenko and Muller, 2004).

1.3.2 PcG complexes and gene regulation in *Arabidopsis*

Organ development in plants is in contrast to *Drosophila* and mammals not restricted to embryogenesis, but takes place throughout the life cycle originating from undifferentiated meristems, and in case of the shoot apical meristem it subsequently gives rise to vegetative and reproductive organs (Avramova, 2009). Similar to PcG proteins in *Drosophila* and mammals, the plant counterparts are found to regulate homeotic genes in *Arabidopsis.* These belong to the homeobox and MADS box transcription factor families, and are involved in several developmental processes in sporophytic and reproductive development, for instance leaf, flower and seed formation and morphology (Avramova, 2009; Katz et al., 2004). Forward genetic screens for regulators of flower development, seed formation and the vernalization response, which is the acquirement of the competence to flower after a prolonged period of cold treatment, identified the first PcG proteins in plants with homology

to PRC2 subunits (Gendall et al., 2001; Goodrich et al., 1997; Schubert et al., 2005; Yoshida et al., 2001).

All four subunits of the PRC2 complex from *Drosophila* are conserved in *Arabidopsis*, and except for the single copy gene *FERTILIZATION INDEPENDENT ENDOSPERM* (*FIE*) which is the ortholog of Esc, all subunits are encoded by small gene families (Schubert et al., 2005; Ohad et al., 1999).

The catalytic subunit E(z) harboring the histone methyltransferase activity has three orthologs in *Arabidopsis* encoded by the genes *CURLY LEAF* (*CLF*), *SWINGER* (*SWN*) and *MEDEA* (*MEA*) (Goodrich et al., 1997; Grossniklaus et al., 1998; Luo et al., 1999). There are three orthologous proteins to the Su(z)12 subunit encoded by *VERNALIZATION 2* (*VRN2*), *EMBRYONIC FLOWER 2* (*EMF2*) and *FERTILIZATION INDEPENDENT SEED 2* (*FIS2*) (Luo et al., 1999; Gendall et al., 2001; Yoshida et al., 2001). The genes *MULTICOPY SUPPRESSOR OF IRA 1-5* (*MSI1-5*) have sequence homology to the WD40 protein p55. MSI1 was initially identified to be required for epigenetic maintenance of reproductive development, and was then found to be a part of the *Arabidopsis* PRC2 complex (Kohler et al., 2003; Hennig et al., 2003). MSI1 is also involved in chromatin remodeling complexes apart from PRC2, and was suggested to be involved in the maintenance of epigenetic information by targeting silencing complexes to chromatin during nucleosome assembly (Hennig et al., 2005).

The fact that three of the PRC2 subunits in *Arabidopsis* are encoded by multiple genes suggests that the composition of the complex might be variable. Furthermore the genes show differences in expression pattern as *MEA* and *FIS2* are expressed in seeds, *VRN2* is preferentially expressed in vegetative tissue whereas *CLF*, *SWN*, *EMF2*, *MSI1* and *FIE* are ubiquitously expressed (Hennig et al., 2003; Chanvivattana et al., 2004; Gendall et al., 2001; Goodrich et al., 1997; Yoshida et al., 2001). Moreover there is genetic evidence that the homologous genes *CLF*, *SWN* and *MEA* have redundant as well as discrete functions as mutations in the single genes cause distinct phenotypes, while double mutants of *clf* and *swn* show a severe phenotype with callus-like growth habit unable to reproduce and complete the plant life cycle (Chanvivattana et al., 2004; Schubert et al., 2005). Genetic studies suggest the existence of at least three PRC2 complexes in *Arabidopsis*: the VRN2-PRC2 complex which is involved in the vernalization response, the FIS2-PRC2 complex required to inhibit fertilization independent seed development and the EMF2-PRC2 complex, that is needed for the suppression of precocious flowering and for floral organ development (Figure 1.1) (Guitton and Berger, 2005). Hence the different complexes control important transitions at

different stages of plant development and are therefore crucial to accomplish the complete life cycle (Guitton and Berger, 2005; Butenko and Ohad, 2011).

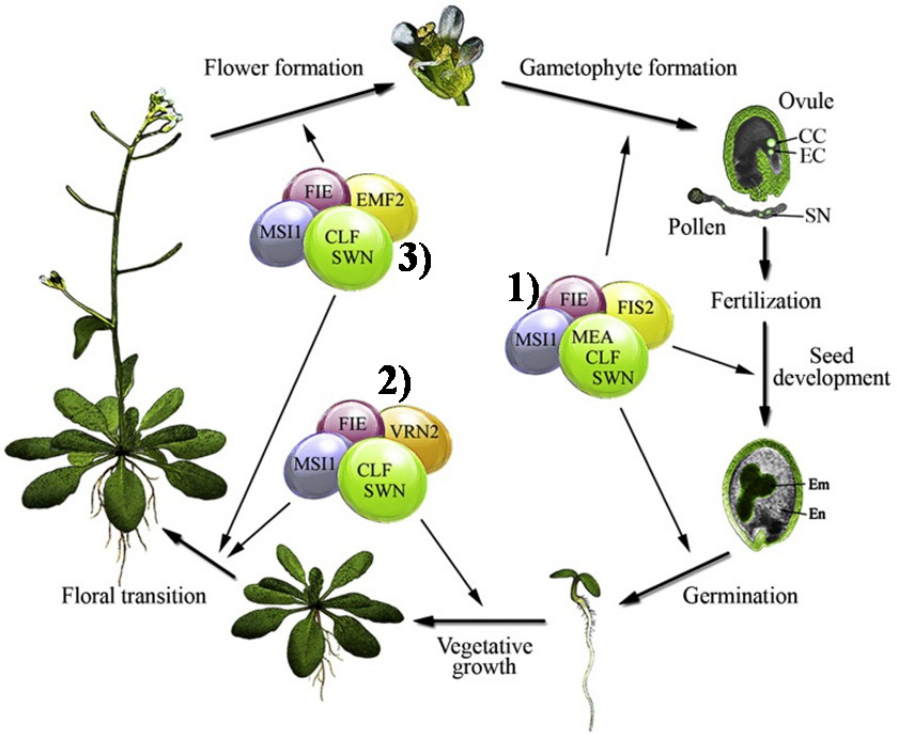


Figure 1.1: PcG complexes control developmental transitions in *Arabidopsis thaliana*
PcG complexes of different composition are important at several stages of plant development that are depicted in the outer circle. The compositions of the complexes regulating the indicated developmental transitions are displayed in the middle. The FIS2-PRC2 (1) regulates gametophyte development and inhibits fertilization independent seed development, the VRN2-PRC2 (2) complex is involved in the vernalization response, the competence to flower after prolonged cold treatment, and is therefore active during vegetative growth. The EMF2-PRC2 (3) complex is needed for the suppression of precocious flowering and for floral organ development (modified from (Butenko and Ohad, 2011)).

Biochemical purification was achieved for the VRN2-PRC2 complex comprising the core subunits VRN2, FIE, SWN and MSI1 as well as the PHD finger proteins VERNALIZATION 5 (VRN5), VERNALIZATION INSENSITIVE (VIN3) and VERNALIZATION 5/ VIN3-LIKE (VEL1) (De Lucia et al., 2008). The FIS2-PRC2 complex consisting of FIS2, MEA or SWN as well as FIE and MSI1 was purified from flowers and young siliques (Kohler et al., 2003).

While the EMF2-PRC2 complex has not yet been purified, there is a large amount of genetic evidence for its existence. The complex is believed to involve CLF or SWN, EMF2, FIE and MSI1 and the mutation of either of the genes except for *SWN* results in an early flowering phenotype (Chanvivattana et al., 2004; Katz et al., 2004; Hennig et al., 2003; Yoshida et al., 2001; Kinoshita et al., 2001). Beside the repression of flowering time regulators such as *FLOWERING LOCUS T* (*FT*) and *AGAMOUS LIKE 19* (*AGL19*), the complex also regulates

floral identity genes like *AGAMOUS* (*AG*), *APETALA1* (*AP1*), *APETALA3* (*AP3*) and *PISTILLATA* (*PI*) whose deregulation causes homeotic transformations of the floral organs (Schonrock et al., 2006; Moon et al., 2003; Chanvivattana et al., 2004).

The catalytic activity of the *Drosophila* PRC2 complex conducted by the histone methyltransferase (HMT) subunit E(z) was recently also affirmed for the *Arabidopsis* VRN2-PRC2 and EMF2-PRC2 complexes with CLF as catalytic HMT subunit (Cao et al., 2002; Czermin et al., 2002; Muller et al., 2002; Schmitges et al., 2011). Furthermore global analysis of histone modification on total histone extracts showed that the mutation of *CLF* as well as the reduction of *EMF2* levels reduces the amount of H3K27 tri-methylation to approximately 50%, while the simultaneous mutation of *CLF* and *SWN* results in a total loss of this modification (Jiang et al., 2008; Lafos et al., 2011). Additionally, a local loss of H3K27me3 accompanied with target gene derepression was observed by target gene analysis in *clf* and *mea* mutants (Makarevich et al., 2006; Schubert et al., 2006).

Even though the phenotypes of the PcG single mutants like *clf* are caused by the misexpression of few target genes (Goodrich et al., 1997; Lopez-Vernaza et al., 2012), whole genome analyses have revealed the existence of approximately 8000 genes carrying the H3K27me3 mark in *Arabidopsis* (Lafos et al., 2011; Zhang et al., 2007a; Oh et al., 2008). It was shown that H3K27me3 correlates with target gene repression, and that, among these target genes, are protein coding genes as well as microRNA coding genes and transposable elements. This discloses a high variety of pathways in which PcG mediated gene regulation is involved (Lafos et al., 2011). Whole genome analyses of H3K27me3 target genes in undifferentiated meristem cells compared to differentiated leaf cells revealed several hundred dynamically regulated target genes that either gain or lose the modification in course of differentiation (Lafos et al., 2011). The dynamic regulation of PcG target genes requires specific H3K27me3 demethylase activity which was recently described in plants. The Jumonji domain protein RELATIVE OF EARLY FLOWERING 6 (REF6) shows H3K27me3 demethylase activity, and overexpression of REF6 results in ectopically increased H3K27me3 levels and in decreased mRNA expression of hundreds of target genes (Lu et al., 2011). H3K27me3 is preferentially enriched in gene bodies within euchromatin, whereas H3K4me3 is enriched near the transcription start sites and 5' regions of active genes (Lafos and Schubert, 2009; Zhang et al., 2007a; Oh et al., 2008).

The recruitment of PRC2 to its target genes has not been solved in *Arabidopsis* as classical PREs have not been identified. Nevertheless cis-acting DNA elements in the DNA sequence of *FLC*, *AG* and *FT* as well as GA-repeat motifs in several PcG target genes have been

described (Meister et al., 2004; Adrian et al., 2010; Kooiker et al., 2005; Sheldon et al., 2002; Sieburth and Meyerowitz, 1997). The binding of the GA-repeats in the promoters of the PcG target genes SEEDSTICK (STK) and INNER NO OUTER (*INO*) by the plant specific protein BASIC PENTACYSTEINE 1 (BPC1) is essential for restricting the expression of these genes to carpels and ovules (Meister et al., 2004; Kooiker et al., 2005). GA-repeat elements are not only associated with transcriptional repression, but were also found to be essential for gene activation for example in the promoter region of the PcG target gene *LEAFY COTYLEDON 2* (*LEC2*) (Berger et al., 2011). Promoter analyses revealed a repressive, cis-acting element in the *LEC2* promoter which was essential for H3K27me3 deposition at *LEC2*, and furthermore introduction of this element in the promoter of an unrelated gene is sufficient for silencing (Berger et al., 2011). As the LEC2 promoter exhibits a binding site for GA-repeat binding proteins as well as a repressive element essential for PcG mediated H3K27me3, there could be an interaction of BPC and PcG silencing in plants even though no evidence for this has been observed to date (Berger et al., 2011; Schatlowski et al., 2008).

Just as the recruitment of PRC2 to its target chromatin, the proteins involved in reading the H3K27me3 mark are not conserved in *Arabidopsis*. The *Drosophila* PRC1 subunit Pc is responsible for the binding of the complex to H3K27me3 and does not have any sequence homologs in *Arabidopsis*. Nevertheless there is evidence for a plant specific mechanism involving TERMINAL FLOWER 2 / LIKE HETEROCHROMATIN PROTEIN 1 (TFL2/LHP1), that has been shown to bind to H3K27me3 *in vitro* and co-localizes with this mark on a genome wide level (Turck et al., 2007; Zhang et al., 2007b). The finding that *lhp1* mutants show misexpression of several PcG target genes, and that LHP1 function is required for stable silencing of FLC and FT, supports the hypothesis that LHP1 might fulfill a PRC1 like function in plants (Adrian et al., 2010; Mylne et al., 2006; Sung et al., 2006; Kotake et al., 2003). The fact that the phenotype of *lhp1* mutant plants is relatively mild compared to PRC2 loss of function mutants suggests the involvement of other not yet identified proteins in the recognition of trimethylated H3K27 (Farrona et al., 2008).

The proteins EMBRYONIC FLOWER 1 (EMF1) and VERNALIZATION 1 (VRN1) likely also have a function in PcG mediated gene regulation downstream of PRC2. EMF1 binds DNA in an unspecific manner and is required for the suppression of several PcG target genes. EMF1 associates with the chromatin of *AG*, *AP3* and *PI,* and the binding to *AG* chromatin is dependent on EMF2. Furthermore transcriptome data of *emf1* and *emf2* mutants show a high overlap of misexpressed target genes (Calonje et al., 2008; Kim et al., 2010; Moon et al., 2003). VRN1 is like VRN2 required for the vernalization response and therefore for the

repression of *FLOWERING LOCUS C* (*FLC*) in response to cold treatment. VRN1 similarly to EMF1 harbours DNA binding capacity and *vrn1* mutants do not show altered H3K27me3 levels at *FLC* chromatin, which places VRN1 downstream of PRC2 function (Levy et al., 2002; Bastow et al., 2004).

While the H3K27me3 binding protein Pc is not conserved in plants, the Ring1 component of the *Drosophila* PRC1 complex that mediates H2A K119 ubiquitination has two orthologs in *Arabidopsis*, AtRING1a and AtRING1b. These proteins, like Ring1, possess an amino-terminal RING domain as well as a carboxy-terminal RAWUL domain, a novel ubiquitin-like protein domain (Xu and Shen, 2008; Sanchez-Pulido et al., 2008). The *Arabidopsis* RING1a and RING1b proteins have been shown to be essential for the repression of class I Knotted-like homeobox (KNOX) transcription factors, hence for the maintenance of meristem function and for the repression of embryonic traits during vegetative growth (Chen et al., 2010; Xu and Shen, 2008). Furthermore interaction of AtRING1a and AtRING1b proteins with CLF, LHP1 and the three *Arabidopsis* orthologs of the mammalian PRC1 protein BMI1 were shown in yeast. Since double mutants of *ring1a* and *ring1b* do not cause a reduction of H3K27me3 levels at selected PcG target genes, it was proposed that they, together with LHP1 and the BMI1 orthologs, fulfill a PRC1-like function downstream of PRC2 (Chen et al., 2010; Xu and Shen, 2008). While previous studies were not able to detect H2A monoubiquitination activity of the PRC1-like components, AtBMI1a and AtBMI1b were shown to mediate this modification on the histone variant H2A.1 (Bratzel et al., 2010). In agreement with these findings the overexpression of the third homolog AtBMI1c, which is expressed during reproductive and seed development, was reported to have monoubiquitin ligase activity *in vivo* and *in vitro* (Bratzel et al., 2012; Li et al., 2011).

In conclusion there is accumulating evidence that PRC1 function is conserved in plants, even though not all *Drosophila* and mammalian proteins have sequence orthologs in the plant kingdom. Plant specific proteins, which might be hidden due to redundancy among gene families, are likely to be involved in PcG mediated gene silencing to meet the specific needs of plant development. These needs differ from mammalian and insect development and include among others, meristem maintenance, phase transitions, organ morphology and stress and pathogen response. In *Arabidopisis* the key transition in the plant life cycle is flowering time which is in part controlled by PcG proteins and has a great agricultural and therefore economical relevance.

1.4 Flowering time regulation by Polycomb group proteins

Flowering time is one of the most important developmental transitions in plants, and therefore it has been intensely studied. This provided a good understanding of how the different environmental and endogenous cues are integrated to cause the switch from vegetative to reproductive growth. Several of the regulatory pathways converge on the PcG target and MADS box transcription factor FLC (Crevillen and Dean, 2011). High levels of FLC delay the transition from vegetative to reproductive growth by repression of the floral integrators SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 (SOC1) and FLOWERING LOCUS T (FT) (Dennis and Peacock, 2007). The major pathways regulating flowering time are the photoperiod pathway, the vernalization pathway, the gibberellin pathway, the autonomous pathway, ambient temperature and an endogenous pathway involving plant age (Srikanth and Schmid, 2011; Crevillen and Dean, 2011). A scheme of the pathways involved in regulation of flowering time is depicted in
Figure 1.2.

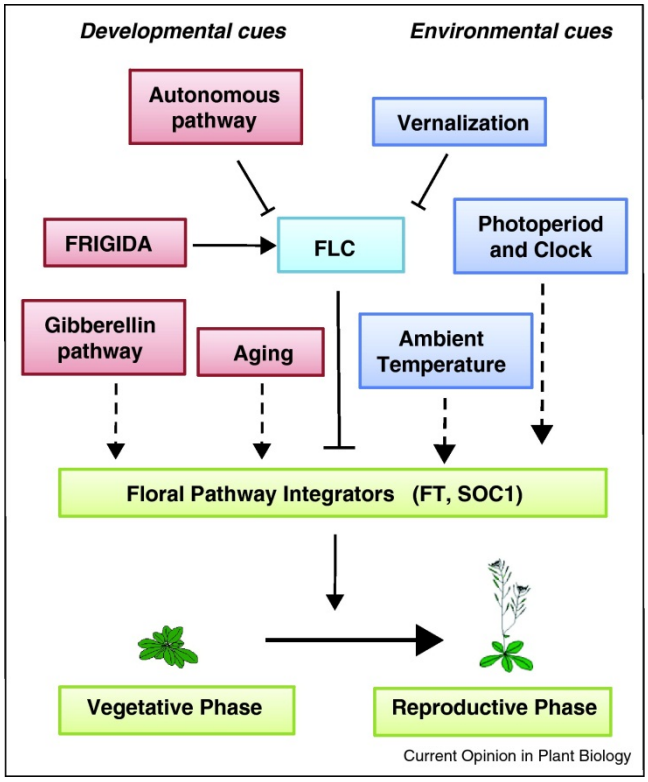


Figure 1.2: The regulation of the switch from vegetative to reproductive development in *Arabidopsis*
Figure from (Crevillen and Dean, 2011)

The expression of *FLC* is controlled by various chromatin complexes mediating histone modifications, histone variant exchange and chromatin remodeling. The transcriptional

activation of *FLC* involves the histone modifications H3K4me3 and H3K36me3, the activity of the RNA polymerase associated factor 1 complex (PAF1C) regulating transcriptional elongation, transcriptional activators like FRIGIDA, H2B ubiquitination and chromatin remodeling through the SWR1 complex (Crevillen and Dean, 2011). On the other hand *FLC* expression is restricted by activity of the autonomous pathway involved in deacetylation and H2A ubiquitination and the vernalization pathway that mediates *FLC* repression in response to cold treatment (Gendall et al., 2001; Jeon and Kim, 2011). In winter annual plants vernalization is the acquisition of the competence to flower after prolonged cold treatment as experienced in winter, and is required to ensure timely flowering upon return to warm conditions. This is an epigenetic phenomenon as the initial signal (cold treatment) takes place several weeks before the plants start to flower, hence the signal has to be 'memorized' and inherited through several mitotic cell divisions (Gendall et al., 2001). The dominant locus *FRIGIDA* (*FRI*) plays an important role in conferring a vernalization requirement in *Arabidopsis* accessions and acts as a transcriptional activator of *FLC* (Koornneef et al., 1994). The vernalization response involves activity of the VRN2-PRC2 complex that mediates the repressive histone modification H3K27me3, the DNA binding protein VRN1 and the PHD finger proteins VIN3 and VRN5 (Gendall et al., 2001; Levy et al., 2002; Greb et al., 2007; Sung and Amasino, 2004). The PHD finger protein VIN3 is transiently induced upon cold treatment which is essential for *FLC* repression as *vin3* mutants do not show *FLC* repression after cold treatment (Sung and Amasino, 2004). Furthermore VIN3 directly binds to *FLC* chromatin and was shown to interact with its homolog VERNALIZATION 5 / VIN3 LIKE1 (VRN5/VIL1) and components of the PRC2 complex (Sung and Amasino, 2004; De Lucia et al., 2008; Greb et al., 2007). The interaction of VIN3 and VRN5 during prolonged periods of cold appears to be essential for the vernalization induced chromatin modifications and thus for epigenetic silencing of *FLC* (Greb et al., 2007). VIN3 and VRN5 belong to a family of 5 PHD finger proteins, and beside VIN3 and VRN5 also VIL2 was reported to be involved in flowering time regulation as it is required for repression of the *FLC*-clade member *MADS AFFECTING FLOWERING5* (*MAF5*) in concert with PRC2 in a photoperiod dependent manner (Kim and Sung, 2010). The PRC2 component VRN2 and the putative PRC1-like protein VRN1 are not required for the initial silencing of *FLC*, but in *vrn1* and *vrn2* mutants the silencing is not maintained upon transfer of the plants to warm conditions (Sung and Amasino, 2004). It was furthermore shown that H3K27me3 levels at the transcriptional start site of *FLC* increase in response to vernalization (Finnegan and Dennis, 2007). Beside VRN1 and VRN2 also LHP1 binds to the chromatin harboring the H3K27me3 mark and is essential

for the maintenance of *FLC* silencing after transition of the plants to warm conditions (De Lucia et al., 2008; Sung et al., 2006; Mylne et al., 2006).

In summary, vernalization induced silencing of FLC involves VIN3 and VRN5, which are essential for the perception of the cold signal, the VRN2-PRC2 complex that mediates H3K27me3, and the PRC1 like proteins VRN1 and LHP1 which ensure the maintenance of *FLC* silencing.

In addition to the protein complexes discussed above also non-coding RNA (ncRNA) transcripts play a role in the silencing of *FLC* upon cold treatment. It has been reported that the first response to cold conditions is the transcription of the *FLC* antisense transcript COOLAIR, that is proposed to result in transcriptional silencing of *FLC*, which is then preceded by the long term silencing mechanisms discussed above (Swiezewski et al., 2009). Another report found the COLD ASSISTED LONG INTRONIC NONCODING RNA (COLDAIR) to be essential for vernalization-mediated silencing of *FLC* through recruitment of PRC2 to the *FLC* locus (Heo and Sung, 2011). The transiently transcribed ncRNA specifically interacts with the CXC domain of CLF and is correlated with the increase of PRC2 abundance at *FLC* chromatin during prolonged cold treatment. Knockdown experiments of COLDAIR transcript resulted in decreased levels of PRC2 abundance and H3K27me3 at *FLC* suggesting that COLDAIR function is mainly the recruitment of PRC2 to *FLC* during the cold (Heo and Sung, 2011).

Beside the abundance of repressive tri-methylation of H3K27 which accumulates upon vernalization, *FLC* expression is regulated by several other histone modifications, for instance H3K4me3, H3K36me3 and histone acetylation, marks which are correlated with gene expression (Dennis and Peacock, 2007). In case of repression of *FLC* the H3 acetylation is removed by a putative histone deacetylase complex containing the autonomous pathway WD40 protein FVE/MSI4 (Ausin et al., 2004). Recent evidence suggests that FVE/MSI4 is a component of several multimeric complexes. Beside deacetylation it was found to be associated with a CUL4-DDB1 ubiquitin ligase complex and the CLF-PRC2 complex, suggesting multiple roles for FVE/MSI4 in flowering time regulation (Ausin et al., 2004; Jeon and Kim, 2011; Jeon and Kim, 2011; Pazhouhandeh et al., 2011).

Since flowering time is such an important trait for the success of the plant life cycle, it is controlled by multiple epigenetic pathways that converge on the regulation of key players such as *FLC* and *FT*. This in turn has the effect that mutations in genes involved in basic epigenetic mechanisms like the PcG genes or other histone modifying enzymes often have an effect on flowering time.

1.5 Protein domains implicated in binding of histone modifications

The active or repressive state of chromatin is in part established by a combination of covalent histone modifications (chapter 1.2). The so called 'histone code' requires proteins that are capable of binding and interpreting these modifications, and that subsequently affect chromatin structure and the level of transcription either in a direct or indirect manner. A number of evolutionary conserved domains have been identified to interact with specific histone modifications (Strahl and Allis, 2000; Turner, 2002).

Most of these domains descend from a common ancestor and belong to the so called 'Royal family' which includes Tudor, Chromatin-binding (Chromo), Malignant Brain Tumor (MBT), PWWP and Agenet domains. The domains belonging to the 'Royal family' are specific for eukaryotes, and the common ancestor of these domains likely harbored methyl-substrate binding ability like the Heterochromatin protein 1 (HP1) chromodomain or the Survival motor neuron (Smn) tudor domain (Maurer-Stroh et al., 2003). Initially the chromodomain of the *Drosophila* protein HP1 was shown to associate with methylated H3K9 essential for the formation of heterochromatin (Lachner et al., 2001; Bannister et al., 2001). The *Arabidopsis* genome encodes for one HP1 homolog, and TFL2/LHP1 comprises a chromodomain that has been shown to bind methylated H3K9 and H3K27 peptides *in vitro*, and furthermore co-localizes with genes carrying the H3K27me3 mark *in vivo* (Turck et al., 2007; Gaudin et al., 2001). Beside the 'Royal family' domains, others have been identified to bind histone modifications: the bromodomain that binds acetylated lysine residues, BRCT domains involved in the recognition of phosphorylated serine and threonine residues and PHD finger domains associating with methylated lysines (Yap and Zhou, 2010). The PHD finger proteins BPTF (bromodomain and PHD finger transcription factor), a protein of the NURF (Nucleosome remodeling factor) complex, and the tumor suppressor gene INHIBITION OF GROWTH 2 (ING2) have both been shown to bind to H3K4me3 marks, and have a function in associated gene activation (Shi et al., 2006; Wysocka et al., 2006). The hydrophobic pocket necessary for binding of the trimethyl group is not conserved in the VIN3-like PHD finger proteins from *Arabidopsis* suggesting an altered binding specificity of the proteins (Wysocka et al., 2006; Shi et al., 2006; Greb et al., 2007).

Another interesting module is the CW domain named for the conserved cysteine and tryptophan residues that form a four cysteine zinc-finger motif (Hoppmann et al., 2011; Perry and Zhao, 2003). Recently the CW domain of the *Arabidopsis* ASH1 HOMOLOG2 (ASHH2) histone methyltransferase was found to specifically bind methylated H3K4 residues. ASHH2 is therefore able to act on H3K4me marked genes mediating trimethylation of H3K36, and

thereby combining the ability to 'read' and 'write' histone modifications in one protein (Hoppmann et al., 2011).

The PWWP domain was first identified in course of the characterization of the mammalian gene Wolf-Hirschhorn syndrome candidate 1 (WHSC1), and was named for a conserved proline-tryptophan-tryptophan-proline motif (Stec et al., 2000). The core domain consists of a five-stranded ß-sheet packed against a helical bundle and was initially characterized as a nonspecific DNA-binding domain, for instance in the mammalian DNA methyltransferase 3b (DNMT3B) and the Hepatoma derived growth factor (HDGF) (Yap and Zhou, 2010; Lukasik et al., 2006; Qiu et al., 2002). Due to its similarity to the tudor- and chromodomains it was speculated that PWWP domains may also have chromatin targeting ability which was subsequently demonstrated for the PWWP domains of mammalian DNMT3A and DNMT3B proteins (Maurer-Stroh et al., 2003; Ge et al., 2004).

Further studies identified the PWWP domain protein 1 (Pdp1) protein from fission yeast and Bromodomain and PHD finger containing 1(Brpf1) from zebrafish to bind to H4K20me1 and H3K36me3 marks via their PWWP domains, respectively (Wang et al., 2009; Vezzoli et al., 2010). The *Arabidopsis* genome contains at least 19 proteins (http://www.uniprot.org) that contain a PWWP domain, among them the serine threonine kinase ATM (ARABIDOPSIS THALIANA ATAXIA_TELANGIECTASIA MUTATED) and the ARABIDOPSIS TRITHORAX homologs 1-5 (ATX1-5).

Studies of the human DNMTL, a protein with sequence similarity to DNMT3A and DNMT3B, showed, that histone modifications also function to restrict proteins from binding to marked chromatin. DNA methylation patterns depend on DNMTL even though it does not contain enzymatic activity. While DNMTL binding to histone H3 tails is inhibited by methylation of H3K4, it is insensitive toward other modifications. DNA associated with chromatin that is not marked by H3K4me3 is methylated through binding of DNMTL to the histone tail and recruitment of the methyltransferase DNMT3B (Ooi et al., 2007).

The proximity of amino acids subject to covalent modifications on the histone tails led to the hypothesis of the existence of binary switches such as the 'methyl/phos switch' (Fischle et al., 2003a; Fischle et al., 2003b). In this mechanism the phosphorylation of an amino acid adjacent to a methyl mark leads to a loss of binding of effector proteins to this methyl mark, and to subsequent recruitment of effector proteins to the phosphorylation mark. The phosphorylation does not necessarily affect the methylation mark but overrules its downstream effects by causing e.g. a change in transcriptional activity or general chromatin state (Fischle et al., 2003a; Winter et al., 2008). Such 'methyl/phos switches' were proposed

for two positions on histone H3 tails, H3K9/H3S10 and H3K27/H3S28. In case of H3K9/H3S10 it was shown that the dissociation of HETEROCHROMATIN PROTEIN 1 from tri-methylated H3K9 during mitosis is dependent on H3S10 phosphorylation (Fischle et al., 2005; Hirota et al., 2005).

Additionally it was shown recently that the PRC2 complex is displaced from chromatin with the double mark H3K27meH3S28pho, which results in a release of transcriptional silencing (Gehani et al., 2010; Lau and Cheung, 2011).

The complexity of the epigenetic histone code requires many more proteins that interpret the combination and sequential arrangement of histone modifications and transmit these signals for example to the transcription machinery, chromatin remodeling complexes or histone modifying enzymes. As PWWP domains have been implicated in DNA as well as histone binding, they may also play a role in the recruitment or dismissal of chromatin modifying enzymes to or from their specific target genes.

1.6 Aims of this study

The purpose of this study is the functional characterization of the plant specific PWWP domain protein SWINGER/CURLY LEAF INTERACTOR 1 (SCI1). Since SCI1 was identified as interactor of the Polycomb group (PcG) protein CURLY LEAF (CLF), it will be analyzed whether SCI1 is involved in PcG mediated gene regulation.

A protein involved in PcG mediated gene regulation should fulfill several criteria. It will be analyzed whether SCI1 is able to interact with additional PcG members in yeast and in plants, as core members of the PcG complexes show several interactions among each other. Furthermore *sci1* mutant alleles will be identified and the morphological as well as molecular phenotypes will be analyzed. Double mutant analyses with PcG mutants like *clf* will be conducted to analyze possible genetic interactions between *SCI1* and *CLF*. The expression level of PcG target genes will be analyzed in single and double mutant backgrounds. Furthermore global and local histone modification levels will be analyzed in *sci1* mutant alleles in order to reveal whether SCI1 is involved in the deposition or maintenance of the repressive H3K27me3 mark.

SCI1 belongs to a plant specific gene family with three members, thus possible redundant functions of SCI1, SCI2 and SCI3 will be studied by establishment of double and triple mutants. In parallel the spatio-temporal expression pattern and the sub-cellular localization of SCI1 will be investigated by RNA and reporter gene analysis. Transgenic lines expressing a tagged SCI1 protein will be used for an immunoprecipitation experiment followed by mass spectrometry analysis, in order to isolate *in vivo* interaction partners of SCI1.

The most prominent feature of the SCI1 protein is the amino-terminal PWWP domain which is implicated in binding of modified histones. Thus *in vitro* experiments will be performed to analyze the binding ability and specificity of the SCI1 PWWP domain to modified histone peptides.

The functional analysis of SCI1 will reveal whether this protein is involved in PcG mediated gene regulation in *Arabidopsis*, and whether SCI1 exhibits additional functions in plant development. The analysis of the SCI1 PWWP domain in respect to histone binding ability will be essential to discover SCI1 function in *Arabidopsis*.

2 MATERIALS AND METHODS

2.1 Materials

2.1.1 Chemicals

Chemicals used in this study were ordered from the following companies:
Biozym (Oldendorf), Duchefa (Haarlem, NL), Eurogentec (Köln), Fluka (Neu-Ulm), Invitrogen™ (Karlsruhe), Merck-Eurolab (Darmstadt), Roche Diagnostics (Mannheim), Roth (Karlsruhe), Serva (Heidelberg), Sigma-Aldrich® (Deisenhofen), Thermo Fischer Scientific (Schwerte).

2.1.2 Buffer

Buffers were prepared as described in (Ausubel, 1996)or (Sambrook et al., 1989)

2.1.3 Media

Bacteria:	LB:	1 % NaCl; 0,5% yeast extract; 1 % bacto tryptone, adjust to pH 7,0 Add 1 % agar for solid media
	YEB	0,5 % sucrose; 0,1 % yeast extract; 0,5 % bacto peptone; 0,5 % beef extract; 2 mM $MgCl_2$; adjust to pH 7,2; 1 % agar (for solid media)
Yeast:	YPD:	1 % yeast extract; 2 % bacto peptone; 2 % glucose Add 1 % agar for solid media
	SD:	6,7 g/l yeast nitrogen base without amino acids; 10 % sterile 10x dropout solution; 2 % glucose, adjust pH to 5,8; 2 % agar (for solid media)
Plant	½ MS:	2,2g MS; 0,05 % MES; 0,5 % sucrose; adjust to pH 5,7 Add 0,8 % agar for solid media; 1,2 % agar for vertical growth assays

2.1.4 Enzymes

Restriction enzymes were ordered from Fermentas (St. Leon Rot) and New England Biolabs (Frankfurt). DNA/RNA modifying enzymes were obtained from Invitrogen™ (Karlsruhe). All enzymes were used according to the manufacturers' instructions.

2.1.5 Antibodies and peptides

Name	Company (Cat. #)	Dilution	Host	Purpose
Primary antibodies				
anti-H3	Abcam (ab1791)	1:5000	rabbit	western blot
anti-H3K27me3	Millipore (07-449)	1:5000	rabbit	western blot, ChIP
anti-H3K27me2	Millipore (07-452)	1:5000	rabbit	western blot
anti-H3K27me2	Active Motif (39245)	1:1000	rabbit	western blot
anti-H3K27me1	Millipore (07-448)	1:5000	rabbit	western blot
anti-H3S28pho	Millipore (07-145)	1:2500	rabbit	western blot
anti-H3K4me3	Diagenode(pAB-003-050)	1:5000	rabbit	western blot, ChIP
anti-GST	Sigma-Aldrich (G7781)	1:2500	rabbit	western blot
anti-GFP	Abcam (ab290)	1:2500	rabbit	western blot
Secondary antibodies				
anti-rabbit-HRP conjugated	Sigma-Aldrich (A9169)	1:60000	goat	western blot
Biotinylated peptides				
H3 (1-21)	Millipore (12-403)	1 µg		*in-vitro* IP
H3 (21-44))	Millipore (12-404)	1 µg		*in-vitro* IP
H3K4me3	Millipore (12-564)	1 µg		*in-vitro* IP
H3K9me3	Millipore (12-568)	1 µg		*in-vitro* IP
H3K27me1	Millipore (12-567)	1 µg		*in-vitro* IP
H3K27me2	Millipore (12-566)	1 µg		*in-vitro* IP
H3K27me3	Millipore (12-565)	1 µg		*in-vitro* IP
H3S28pho (AC-ATKAARK-S(p)-APATGGVKKPHRYRPGGK-(Biot)-NH2)	Intavis	1 µg		*in-vitro* IP
Other peptides				
chicken histone octamers	Abcam (ab45275-3)		chicken	western blot
Magnetic beads				
Dynabeads® M-280 Streptividin	Invitrogen (112-05D)			*in-vitro* IP
Dynabeads® Protein G	Invitrogen (100-03D)			ChIP

Table 2.1: Antibodies and peptides

2.1.6 Molecular size standards

PageRuler^{TM} Prestained Protein Ladder (#SM0671), Fermentas (St. Leon Rot):

2.1.7 Membrane and Paper

Western blot analysis was performed using Roti-PVDF membrane (Roth; Karlsruhe) and 3mm Whatman paper (Maidstone, England).

2.1.8 Oligonucleotides

The oligonucleotides used in this study are listed in Table 2.2 to Table 2.5, sorted by their application. They were ordered from Eurogentec (Köln).

Name	Locus	Stock-no.	Sequence (5′ → 3′)
I. cloning			
At3g03140-Prom-F	At3g03140	JD2	CTAACTTCACAGCACGGCTCTGAGG
At3g03140-ATG-F	At3g03140	JD3	ATGGCAAGTCCAGGATCAGGTGC
At3g03140-Stop-R	At3g03140	JD4	TTGAACTCTTCTTCTCTCGTTAAAGGC
SCI1-PWWP-2-R	At3g03140	LI6	TTAGCATCAAGAGAGTCATCTCTAGC
SCI1-FLAG-rev	At3g03140	QF6	TCATTTGTCATCGTCGTCCTTGTAGTCTTGAA CTCTTCTTCTCTCG
SCI1-HA-rev	At3g03140	QF7	TCAAGCGTAATCTGGAACATCGTATGGGTAT TGAACTCTTCTTCTCTCG
SCI1-319-for	At3g03140	QE3	ATGGCCATTCTCCATGCCCTTG
SCI1-691-for	At3g03140	QE4	ATGAGTGGGGATCATAGCATGGAG
SCI1-1059-for	At3g03140	QE5	ATGAGTACAAGCAGTGAGGAAGATC
SCI1-1537-for	At3g03140	QE6	ATGGGTCAGGAGTTTGATGTTTCTG
SCI1-1900-for	At3g03140	QE7	ATGTCTCTGGCAGATCAGGGAAGG
SCI1-401rev	At3g03140	QE8	TTGCATCAAGAGAGTCATCTCTAGC
SCI1-801rev	At3g03140	QE9	CTCATGTAGGTCTGAAACATCA
SCI1-1201rev	At3g03140	QF1	AGATCTTCTTGGAAGAGTCCGGA
SCI1-1601rev	At3g03140	QF2	GCTTCCTCGCACTTCTAAATCC
SCI1-2001rev	At3g03140	QF3	TGGTCGCGGAATTCGCAATG
SCI1-2301rev	At3g03140	QF4	TTGAACTCTTCTTCTCTCGTTAAAG
At3g21295-Prom-F	At3g21295	JD8	GTACCGTGCATCTTCATGGATGG
At3g21295-Stop-R	At3g21295	JE1	TCCTGTATTACCTGCAGATGGTAATG
SWN-ATG-F	At4g02020	II7	ATGGTGACGGACGATAGCAAC
SWN-SET-R	At4g02020	II8	GCATTGCCCTTCTCCGCGTC
CLF-A-ATG-F	At2g23380	NB2	AATGGCGTCAGAAGCTTCGCC
CLF-SET-R	At2g23380	NB3	TTTCATATTCCTGCACTCATAATTA
TFL2-CHROMO-F	At5g17690	II9	CGGAAGCGCAAATATGCAGGTCC
TFL2-Stop-R	At5g17690	JA1	AGGCGTTCGATTGTACTTGAGATG
CHROMO-TFL2-F	At5g17690	LI4	GATGAAGGGTTTTATGAAATTG
CHROMO-TFL2-R	At5g17690	LI5	TTACTTCAAACTTCCCTCAAAGGC
FVE-ATG-F	At2g19520	NA9	ATGGAGAGCGACGAAGCAGCAGC
FVE-Stop-R	At2g19520	NB1	AAAGACAAACCAGCCTTACAAAAACC
attR1Xho-F	GW-cassette	KB5	ATCTCGAGCGAATTCCCATCAACAAGTTTG
attR2-Xho-R	GW-cassette	KB7	ATCTCGAGTCGACGGATCCCCATCAACCAC
II. site directed mutagenesis			
SCI1_W27-A_F	At3g03140	PH3	GCTCGTGGGCGCCGGGGAG
SCI1_W27-A_R	At3g03140	PH4	CATTCCTCCTCCTCACCCACACAATCG
SCI1_W63-A_F	At3g03140	PH1	GTGCTTAGGGATGCGTACAATTTAGAG
SCI1_W63-A_R	At3g03140	PH2	ATTAAAATACGAATGCTTTCAGTAATC

Table 2.2: Oligonucleotides for cloning

Name	Stock-no.	forward	reverse
WT fragment			
sci1-1	BH2/BH1	GCAGCACAAGGGAAAAAGAA	GTGCCCTTGTCTTTTGGCTA
sci1-3	JD3/KC8	ATGGCAAGTCCAGGATCAGGTGC	AGAATGGCATCTTCTCTCCGAGC
sci1-GABI	BH2/BH1	GCAGCACAAGGGAAAAAGAA	GTGCCCTTGTCTTTTGGCTA
SCI1::SCI1 -GFP	BH2/JA6	GCAGCACAAGGGAAAAAGAA	TCAAATAGCGTCCTCCTGCT
SCI1::SCI1 -GUS	BH2/JA6	GCAGCACAAGGGAAAAAGAA	TCAAATAGCGTCCTCCTGCT
sci2-2	JA4/JA5	AGAGGAGCCCGAATCTTCAT	TCCTAACGCCGCAGAGTAGT
sci3-2	JD9/JF3	ATGGGTAGTAGTGATGAGCGAAACTG	GACGGAGAGCAATCCCTGCTACT
clf-28	AF2/AF3	CTGCCAGTTCAGGAATGGTT	GAAGGGAGCTCTCTGCTTGAT
swn-7	IH4/IH5	GGATAAGCAGAATACCGAGGAATTTTC	GTCTTTTAGAATTGGGACCTCACGC
flc-3 (del.)	AD2/AD3	GCAATAGTTCAATCCGTATCG	GCGTCACAGAGAACAGAAAGC
fve-3	NF6/NF7	AGAGAGGGAAAATGGAGAGCGAC	CTTTCTTTCATTTTAACATCGCTCAG TT
Insertion / transgene			
sci1-1	AB6/BH1	GCCTTTTCAGAAATGGATAAATAGCCTT GCTTCC	GTGCCCTTGTCTTTTGGCTA
sci1-3	JD3/KC8	Point mutation, cut mutant fragment with NlaIII	
sci1-GABI	BH2/BI4	GCAGCACAAGGGAAAAAGAA	CTGGGAATGGCGAAATCAAGGCATC
SCI1::SCI1 -GFP	BH2/HG6	GCAGCACAAGGGAAAAAGAA	GCTTCATATGATCAGGGTAACG
SCI1::SCI1 -GUS	BH2/OD3	GCAGCACAAGGGAAAAAGAA	CAACGCTGATCAATTCCACA
sci2-2	JA4/AB4	AGAGGAGCCCGAATCTTCAT	TTGGGTGATGGTTCACGTAGTGGG
sci3-2	JD9/AB6	ATGGGTAGTAGTGATGAGCGAAACTG	GCCTTTTCAGAAATGGATAAATAGC CTTGCTTCC
clf-28	AF2/AB4	CTGCCAGTTCAGGAATGGTT	TTGGGTGATGGTTCACGTAGTGGG
swn-7	IH4/AB4	GGATAAGCAGAATACCGAGGAATTTTC	TTGGGTGATGGTTCACGTAGTGGG
flc-3	AD2/AD3	Deletion, WT fragment larger than mutant fragment	
fve-3	NF6/NF7	Point mutation, cut WT fragment with AlwNI	

Table 2.3: Oligonucleotides for genotyping mutant and transgenic plant lines

Name	Locus	Stock-no.	Sequence (5´ → 3´)
I. H3K27me3 and H3K4me3 ChIP			
ACTIN2	AT5G09810	GI5	TAGTGAAAAATGGCCGATGG
		GI6	CCATTCCAGTTCCATTGTCA
AG	AT4G18960	AH7	TGGGTACTGAGAGGAAAGTGAGA
		AH8	TGAGTGATTGCCCAACTTGA
II. H3K27me3 ChIP			
FLC	AT5G10140	FC7	GACCAGGCTGGAGAGATGAC
		FC8	GTTTCCAGTGGCCTTTTCAA
FUS3	AT1G24260	LF9	GTGGCAAGTGTTGATCATGG
		LG1	AGTTGGCACGTGGGAAATAG
SEP3	AT1G24260	LH7	GGGTTTCCAATTTTGGGTTT
		LH8	GATGAATCCCATCCCCAAGT
III. H3K4me3 ChIP			
FLC	AT5G10140	GI3	GTCGCTCTTCTCGTCGTCTC
		GI4	AGGGGGAACAAATGAAAACC
SEP3	AT1G24260	LH9	TGACGTTTGCAAAGAGAAGG
		LI1	GCATGCTCGAACTACTGCAA

Table 2.4: Oligonucleotides for ChIP analysis

Name	Locus	Stock-no.	Sequence (5′ → 3′)
I. semiquantitative RT-PCR			
SCI1-I.-F	At3g03140	JD3	ATGGCAAGTCCAGGATCAGGTGC
SCI1- I.-R	At3g03140	KC8	CACGATCTGGAACTCCGGTGAAGC
SCI1-II.-F	At3g03140	BH2	GCAGCACAAGGGAAAAAGAA
SCI1- II.-R	At3g03140	QF3	TGGTCGCGGAATTCGCAATG
SCI1-III.-F	At3g03140	QE7	ATGTCTCTGGCAGATCAGGGAAGG
SCI1- III.-R	At3g03140	JD4	TTGAACTCTTCTTCTCTCGTTAAAGGC
SCI2-F	At1g51745	JD1	ATGGCAAGTCCAGGATCAGGTGC
SCI2-R	At1g51745	JD6	ATGGAGAGTAATGATGACCGAAACTTGG
SCI3-F	At3g21295	KF1	CTAGGTCGTGACGATGCTAGCG
SCI3-R	At3g21295	KF2	GATGCAGTAGCCTTTGCCGTCGC
eIF4A-F	AT1G54270	GA9	TTCGCTCTTCTCTTTGCTCTC
eIF4A-R	AT1G54270	GB1	GAACTCATCTTGTCCCTCAAGTA
II. qRT-PCR			
AG	AT4G18960	FE3	ATCCGATCCAAGAAGAATGAG
		FE4	TTTCAGCTATCTTTGCACGA
AGL19	AT4G22950	OB1	TTGCTTGGAGAAGGCATTGATGCA
		OB2	GATTCCTCTCCTCTGCCTTCAAC
AGL24	AT4G24540	OB3	CTGAAAAGAAGGGCGAGTGTG
		OB4	CTTTCTCTCTCTCAGATTCATTCCC
FLC	AT5G10140	JG9	TTGAACTTGTGGATAGCAAGCTT
		JH1	CGGTCTTCTTGGCTCTAGTCA
FLD	AT3G10390	UD6	GCTTCTATTAAAGCCGAGCGA
		UD7	TAATCCTTCTGATCATCTTCCGCT
FLK	AT3G04610	UD8	CTGATACCACCGCTCTAGTC
		UD9	ACCCATTATGCAACTCATCTG
FT	AT1G65480	LC3	GATCCAGATGTTCCAAGTCC
		LC4	ACAATCTCATTGCCAAAGGT
FVE	AT2G19520	UD2	GGCTCGCTAACCATAACCT
		UD3	ACACTTCCATCAGTTTGCTC
LINC1	AT1G67230	UC5	AAGAGTTCAGATGTGGAGAGG
		UC6	CTTCTCGTTCAGCAATGTAGG
nblack (TIP41-like)	AT4G34270	EA5	GTGAAAACTGTTGGAGAGAAGCAA
		EA6	TCAACTGGATACCCTTTCGCA
PI	AT5G20240	LA6	AAATCTGATGGCTGTCGAGCAC
		LA7	TCTGGTGGTCTCGGACTTTGTC
PRMT5	AT4G31120	UE1	CCTTGCAGTTATGGCTAAGG
		UE2	AGACGAAACGAATTCCACAG
SEP3	AT1G24260	LA8	AGAGCTCTCAGGACACAGTTTATGCT
		LA9	GCATGCGTTCCTTACTCTGAAGAT
STM	AT1G62360	DB8	AATTGTCAGAAGGTTGGAGC
		DB9	CGTTGAAGGAAGACCATAGC

Table 2.5: Oligonucleotides for expression analysis

2.1.9 Constructs and Vectors

All vectors were created using the GATEWAY® cloning system (Invitrogen™) or by using restriction endonucleases and ligation. All methods were performed according to manufacturer's instructions.

Empty vectors used for cloning purposes:

Name	Reference / Origin	Features
Entry vectors		
pCR®8/GW/TOPO	Invitrogen	
Destination vectors		
pGADT7-GW	(Horak et al., 2008)	GAL4-AD (GW vector)
pGBKT7-GW	(Horak et al., 2008)	GAL4-BD (GW vector)
pGADT7	Clontech	GAL4-AD
pGEX4T3-GW	modified pGEX4T3, GE Healthcare	Ptac::GST (GW vector)
pGEX4T3	GE Healthcare	Ptac::GST
pMDC7-Cterm-GFP	(Bleckmann et al., 2010)	i35S::CtermGFP (GW vector)
pMDC7-Cterm-mCherry	(Bleckmann et al., 2010)	i35S::Cterm mCherry (GW vector)
pMDC7-Nterm-GFP	modified pMDC7	i35S::NtermGFP (GW vector)
pMDC7-Nterm-mCherry	modified pMDC7	i35S::Nterm mCherry (GW vector)
pMDC107	(Curtis and Grossniklaus, 2003)	proend::gene-GFP (GW vector)
pMDC163	(Curtis and Grossniklaus, 2003)	proend::gene-GUS (GW vector)
pYS39	(Schatlowski et al., 2010)	i35S::LUC-C
pYS40	(Schatlowski et al., 2010)	i35S::LUC-N

Table 2.6: Vectors used for cloning purposes

Created entry clones:

Name	Locus	Oligos	Entry clone no.
Entry clones			
SCI1-CDS in pCR®8/GW/TOPO	At3g03140	JD3 + JD4	1
SCI1pro:SCI1 in pCR®8/GW/TOPO	At3g03140	JD2 + JD4	2
SCI1pro:SCI1(W27A) in pCR®8/GW/TOPO	At3g03140	PH3 + PH4	3
SCI1pro:SCI1(W63A) in pCR®8/GW/TOPO	At3g03140	PH1 + PH2	4
SCI1pro:SCI1-FLAG in pCR®8/GW/TOPO	At3g03140	JD2 + QF6	5
SCI1pro:SCI1-HA in pCR®8/GW/TOPO	At3g03140	JD2 + QF7	6
SCI1-PWWP-domain in pCR®8/GW/TOPO	At3g03140	JD3 + LI6	7
SCI1-PWWP-domain (W27A) in pCR®8/GW/TOPO	At3g03140	PH3 + PH4	8
SCI1 incl. NLS in pCR®8/GW/TOPO	At3g03140	JD3 + QF2	9
SCI1ΔPWWP in pCR®8/GW/TOPO	At3g03140	QE3 + QF4	10
SCI1-FRAGMENT 1 (aa 1-137) in pCR®8/GW/TOPO	At3g03140	JD3 + QE8	11
SCI1-FRAGMENT 2 (aa 107-274) in pCR®8/GW/TOPO	At3g03140	QE3 + QE9	12
SCI1-FRAGMENT 3 (aa 231-403) in pCR®8/GW/TOPO	At3g03140	QE4 + QF1	13
SCI1-FRAGMENT 4 (aa 354-541) in pCR®8/GW/TOPO	At3g03140	QE5 + QF2	14
SCI1-FRAGMENT 5 (aa 513-670) in pCR®8/GW/TOPO	At3g03140	QE6 + QF3	15
SCI1-FRAGMENT 6 (aa 634-769) in pCR®8/GW/TOPO	At3g03140	QE7 + QF4	16
SCI3pro:SCI3 in pCR®8/GW/TOPO	At3g21295	JD8 + JE1	17
CLFΔSET in pCR®8/GW/TOPO	At2g23380	NB2 + NB3	18
SWNΔSET in pCR®8/GW/TOPO	At4g02020	II7 + II8	19
TFL2ΔCHROMO in pCR®8/GW/TOPO	At5g17690	II9 + JA1	20
TFL2-CHROMO-domain in pCR®8/GW/TOPO	At5g17690	LI4 + LI5	21
FVE-CDS in pCR®8/GW/TOPO	At2g19520	NA9 + NB1	22

Table 2.7: Created entry clones

Other entry clones:

Name	Locus	Order -no.	Entry clone no
Entry clones			
VRN2-cDNA in pEntr223	AT4G16845	G09871	23
EMF2-cDNA in pEntr223	AT5G51230	G25372	24
MSI1-cDNA in pEntr223	AT5G58230	U16887	25
TFL2-cDNA in pEntr223	AT5G17690	G82682	26
FVE-cDNA in pENTR/SD-dTopo	At2g19520	U15529	27
SWN-cDNA in pEntr223	AT4G02020	G14590	28

Table 2.8: Other entry clones used in this study

Created destination vectors:

Name	Vector backbone	Derived from entry clone no.
Yeast interaction studies		
pGAD-SCI1-CDS	pGADT7-GW	1
pGBD-SCI1-CDS	pGBKT7-GW	2
pGAD-SWNΔSET	pGADT7-GW	19
pGBD-SWNΔSET	pGBKT7-GW	19
pGAD-VRN2-cDNA	pGADT7-GW	23
pGAD-FVE-cDNA	pGADT7-GW	22
pGAD-TFL2-cDNA	pGADT7-GW	26
pGAD-MSI1-cDNA	pGADT7-GW	25
In vitro pulldown		
GST-SCI1-PWWP-domain	pGEX4T3-GW	7
GST-SCI1-PWWP-domain (W27A)	pGEX4T3-GW	8
GST-TFL2-CHROMO-domain	pGEX4T3-GW	21
Localization studies		
SCI1pro:SCI1-GFP	pMDC107	2
SCI1pro:SCI1-GUS	pMDC163	2
i35S::SCI1-CDS-GFP	pMDC7-Cterm-GFP	1
i35S::SCI1-CDS-mCherry	pMDC7-Cterm-mCherry	1
SCI3pro:SCI3-GFP	pMDC107	17
SCI3pro:SCI3-GUS	pMDC163	17
i35S::CLFΔSET-GFP	pMDC7-Cterm-GFP	18
i35S::CLFΔSET-mCherry	pMDC7-Cterm-mCherry	18
i35S::SWNΔSET-GFP	pMDC7-Cterm-GFP	19
i35S::SWNΔSET-mCherry	pMDC7-Cterm-mCherry	19
Split-Luciferase assay		
N-LUC SCI1-CDS	pYS40	1
C-LUC CLFΔSET	pYS39	18
C-LUC SWNΔSET	pYS39	19

Table 2.9: Created destination vectors

Other vectors used in this study:

Name	Vector backbone	Reference / Origin
Yeast interaction studies		
pAD-SCI1(trunc)	pAD424 (Clontech)	Provided by Daniel Schubert
pBD-CLFΔSET	pGBT9 (Clontech)	(Chanvivattana et al., 2004)
pAD-MEA CDS	pAD424 (Clontech)	Provided by Claudia Köhler
Split-Luciferase assay		
N-LUC H2A	pYS40	(Schatlowski et al., 2010)
C-LUC H2B	pYS39	(Schatlowski et al., 2010)
Renilla-LUC (R-LUC)	pYS Y3 (provided by Yvonne Stahl)	

Table 2.10: Other vectors used in this study

2.1.10 Microorganism

The *Escherichia coli* strains were used for cloning, plasmid amplification and protein expression. *Agrobacterium tumefaciens* strains were employed to stably or transiently transform *Arabidopsis thaliana* and *Nicotiana benthamiana*. *Saccharomyces cerevisiae* strains listed were used for yeast two-hybrid experiments.

Strain	Description
E. coli	
DB3.1	F– *gyr*A462 *end*A1 Δ(sr1-*rec*A) *mcr*B *mrr hsd*S20(rB–, mB–) *sup*E44 *ara*-14 *gal*K2 *lac*Y1 *pro*A2 *rps*L20(SmR) *xyl*-5 λ– *leu mtl*1
DH5α	F– Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (rK–, mK+) *pho*A *sup*E44 λ– *thi*-1 *gyr*A96 *rel*A1
BL21(DE3)	F– ompT gal dcm lon hsdSB(rB- mB-) λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5]
Rosetta(DE3)pLys	F^- ompT $hsdS_B(R_B^- m_B^-)$ gal dcm λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5]) pLysSRARE (Cam^R)
A. tumefaciens	
GV3101 (pMP90)	*C58C1, pMK90*, *Rif*^{r}, *Gm*^{r} (Koncz, 1986)
GV3101 (pMP90RK)	*C58C1, pMK90RK*, *Rif*^{r}, *Gm*^{r}, *Km*^{r} (Koncz, 1986)
S. cerevisiae	
AH109	*MATa, trp1-901, leu2-3,112, ura3-52, his3-200, gal4Δ, gal80Δ, LYS2::GAL1*$_{UAS}$ *-GAL1*$_{TATA}$*-HIS3, MEL1, GAL2*$_{UAS}$ *-GAL2*$_{TATA}$*-ADE2, URA3::MEL1*$_{UAS}$*-MEL1*$_{TATA}$*-lacZ*
YST1	*MATa,ura3-52, his3-200, ade2-101, trp1-901, leu2-3,112, gal4Δmet-, gal80Δ, URA3::GAL1*$_{UAS}$ *-GAL1*$_{TATA}$*-lacZ, MEL1*

Table 2.11: Microorganisms used in this study

2.1.11 Plants

The *Arabidopsis thaliana* plants used in this study are either in Columbia (Col-0) or Wassilewskaja (WS) ecotype. Plant carrying a T-DNA insertion, deletion or point mutation and transgenic plants created in this study are listed in Table 2.12. Seeds were ordered from the Nottingham Arabidopsis Stock Center (NASC).

Name	Locus	Insertion / Description	NASC-ID
sci1-1	At3g03140	Sail_342_C09	N815951
sci1-3	At3g03140	Tilling allele: Point mutation in PWWP-domain:	N93526
sci1-GABI	At3g03140	GABI_219C02	N420954
SCI1::SCI1-GFP	At3g03140	Translational fusion of gen. locus of SCI1 to GFP (pMDC107)	
SCI1::SCI1-GUS	At3g03140	Translational fusion of gen. locus of SCI1 to GUS (pMDC163)	
sci2-2	At1g51745	Salk 136093	N636093
sci3-2	At3g21295	Sail_828_A07	N836957
SCI3::SCI3-GFP	At3g21295	Translational fusion of gen. locus of SCI3 to GFP (pMDC107)	
SCI3::SCI3-GUS	At3g21295	Translational fusion of gen. locus of SCI3 to GUS (pMDC163)	
clf-28	At2g23380	Salk 139371	N639371
swn-7	At4g02020	Salk 109121	N609121
vrn2-1	At4g16845	(Gendall et al., 2001)	
emf2-10	At5g51230	(Chanvivattana et al., 2004)	
flc-3	At5g10140	deletion (fast neutron from R. Amasino,)	
fve-3	At2g19520	Point mutation (Kim et al., 2006)	
FRI(SF2)FLC-LUC	At4g00650	FRI(SF2) + FLC-LUC in Col-0 background (Mylne et al., 2004)	
linc1-1	At1g67230	SALK_025347 (Dittmer et al., 2007)	N525347
LINC1::LINC1-GFP	At1g67230	(Dittmer and Richards, 2008)	

Table 2.12: Mutant and transgenic *Arabidopsis thaliana* lines

2.1.12 Software

Cloning purposes

Vector NTI, Invitrogen

Perl Primer: http://perlprimer.sourceforge.net/

Protein analysis

Protein domains: Expasy-Prosite: http://prosite.expasy.org/

Protein Subcellular Localization Prediction: http://wolfpsort.org/

Uni-Prot database: http://www.uniprot.org

Sequence Analysis

NCBI BLAST: http://blast.ncbi.nlm.nih.gov/Blast.cgi

KALIGN: http://www.ebi.ac.uk/Tools/msa/kalign/

BoxShade: http://www.ch.embnet.org/software/BOX_form.html

Analysis of Arabidopsis mutant lines

TAIR: http://arabidopsis.org/

T-DNA express: http://signal.salk.edu/cgi-bin/tdnaexpress

Arabidopsis root and nuclei measurements

Image J: http://rsbweb.nih.gov/ij/

2.2 Methods

2.2.1 Basic molecular methods

Basic molecular methods like DNA amplification by PCR, semi-quantitative RT-PCR, DNA/RNA separation, phenol/chloroform extractions and DNA precipitation were performed using protocols from (Sambrook et al., 1989).

2.2.1.1 Cloning and vector generation

DNA amplification for cloning purposes was performed using the Phusion polymerase (Finnzymes) according to manufacturers' instructions. Cloning was performed using the Gateway technology (Invitroten), and created entry and destination vectors are listed in Table 2.7 and Table 2.9. Site directed mutagenesis was performed in a one-step mutagenesis protocol as described in (Zheng et al., 2004)

2.2.2 Quantitative real time PCR (qRT-PCR)

qRT-PCR analyses was performed with technical triplicates and at least two biological replicates using oligonucleotides listed in Table 2.5 for expression analyses and Table 2.4 for ChIP analyses. The Mesa Blue Sybr Mix (Eurogentec) was used for amplification in a Chromo4 real-time PCR cycler (Biorad). Expression levels were normalized to the reference gene At4g34279 (nblack) (Czechowski et al., 2005).

2.2.3 Bacteria methods

Bacteria strains used in this study are listed in Table 2.11.

2.2.3.1 Growth conditions and transformation

E. coli cells were generally grown at 37°C, and *A. tumefaciens* strains at 28°C in liquid cell culture or on solid media (2.1.3). Chemical competent *E.coli* and A. tumefaciens cells were prepared as described in (Inoue et al., 1990) and (Hofgen and Willmitzer, 1988), respectively. Competent *E.coli* cells were transformed using the heat shock method described in (Sambrook et al., 1989), while competent *A. tumefaciens* were transformed by a method that includes freezing in liquid nitrogen followed by thawing at 37°C (Hofgen and Willmitzer, 1988).

2.2.3.2 Preparation of plasmid DNA from *E. coli*

Plasmid DNA extraction was carried out using the peqGOLD Plasmid Miniprep Kit (Peqlab, Erlangen) or the QIAGEN Plasmid Midi Kit (Qiagen, Hilden)

2.2.4 Yeast methods

2.2.4.1 Growth conditions

All yeast cultures were grown at 28°C, untransformed AH109 and YTS1 cultures were grown on solid or liquid YPD media supplemented with adenine. Yeast cultures transformed with either Gal4-BD or Gal4-AD constructs were grown on selective SD media lacking tryptophan (W) or leucine (L), respectively. After mating, the strains were grown on the control media SD (-LW) as well as on media additionally lacking histidine (-LWH) or histidine and adenine (-LWHA) selecting for interaction and resulting reporter gene activity.

2.2.4.2 Yeast two-hybrid interaction studies

All yeast techniques were performed as described in the 'Yeast Protocols Handbook' (Clontech Laboratories, Inc., protocol PT3024-1, version PR13103). The yeast strains YST1 and AH109 were transformed with Gal4-BD and Gal4-AD constructs. After mating, yeast two-hybrid studies were performed, either by dilution series on selective media or in liquid culture via the ONPG assay.

2.2.5 Plant methods

Arabidopsis thaliana lines used in this study are listed in Table 2.12.

2.2.5.1 Plant growth conditions

Arabidopsis thaliana plants were grown in long day (LD) conditions (16 h light/8 h dark cycles at 20°C) or short day (SD) conditions (8 h light/ 16 h dark cycles at 20°C) on ½ MS medium, and were transferred to soil after 10-14 days of growth.

Nicotiana benthamiana plants were grown on soil in long day conditions.

2.2.5.2 Seed sterilization

Arabidopsis thaliana seeds were sterilized with Ethanol: Seed were incubated in 70% Ethanol + 0,05% TritonX-100 for 5 min followed by incubation in 96% Ethanol for 5 min and subsequent drying of the seeds. Sees were sown on ½ MS medium if necessary supplemented with antibiotics for transgene detection.

2.2.5.3 Plant transformation

Agrobacterium tumefaciens mediated transformation of *Arabidopsis thaliana* was performed after the floral dip method (Clough and Bent, 1998).

Transient transformation of *N. benthamiana* leaf epidermis cells was performed using four week old plants that were infiltrated as described in (Bleckmann et al., 2010).

2.2.5.4 Isolation of genomic DNA from *Arabidopsis*

Genomic DNA was isolated using a modified protocol from (Dellaporta and Hicks, 1983)

2.2.5.5 RNA extraction and cDNA synthesis

Total RNA was isolated using the RNeasy Plant Mini Kit (Qiagen, Hilden), followed by DNaseI (Fermentas) treatment according to manufacturers' instructions. The DNase treated RNA was then transcribed to cDNA using SuperScriptII (Invitrogen) for further analysis via RT-PCR or qRT-PCR

2.2.5.6 ß-Glucuronidase assay

ß-Glucuronidase (GUS) activity in transgenic *Arabidopsis* lines was performed as described in (Colon-Carmona et al., 1999)

2.2.5.7 Flowering time analysis

The number of rosette leaves of *Arabidopsis* plants grown in LD or SD conditions was determined of approximately 20 plants per genotype. Compared plants were always grown simultaneously.

2.2.5.8 Root length measurement

Plants used for root length measurements were sown on ½ MS medium with 1,2% of agar, and placed vertically in the growth chamber in LD conditions. Every second day the plates were scanned using Canon Scan8800F, and roots were measured using the Image J software.

2.2.5.9 Plant nuclei isolation for nuclear size measurements

Protocol was adjusted from (Schubert et al., 1993): 3-4 whole seedlings were fixed in 4% Paraformaldehyde (in H_2O) for 20 min. After rinsing in cold Tris-Buffer (10 mM Tris HCl pH 7,5;M 10 mM Na_2 EDTA; 100 mM NaCl), the tissue was chopped with a razor blade in 400µl LB01-Buffer (15 mM Tris HCl pH7,5; 2 mM Na_2-EDTA; 0,5 mM Spermin x 4HCl; 80 mM KCl; 20 mM NaCl; 0,1% Triton X-100) in a glass petri-dish. The suspension was filtered through a 50 µm mesh, and 12 µl of the nuclei suspension was mixed with 12 µl of SO-Buffer (100 mM Tris; 50 mM KCl; 2 mM $MgCl_2$; 0,05% TWEEN-20; 5% sucrose) directly on a microscopic slide. Nuclei were dried overnight and stored at 4°C until analysis. For DAPI staining 1µg/ml DAPI in VECTASHIELD (Vector Laboratories) was applied to the dried nuclei followed by microscopic analyses (Axioskop, Zeiss).

2.2.5.10 Ploidy analyses

Ploidy analyses were carried out with whole seedlings 10 dag (8 seedlings per sample) and cotyledons (12 per sample). Tissue was chopped in nuclei isolation buffer (15 mM Tris; 2 mM Na_2-EDTA; 0,5 mM Spermin; 80 mM KCL; 20 mM NaCl; 15 mM ß-Mercaptoethanol; 0,1%TritonX-100; pH adjusted to 7,5), filtered through a 50 µm mesh, stained with Propidiumiodide (0,005 mg/ml) and the DNA content was measured using a Fluorescence Activated Cell Sorter (FACS Aria BD Bioscience).

2.2.6 Protein extraction and analysis from *E. coli* and *Arabidopsis*

2.2.6.1 Protein expression and purification from E. coli

E. coli strains used for expression of protein domains were Rosetta and BL21 (Table 2.11), transformed with constructs based on the pGEX4T3 vector backbone (Table 2.6). GST-fusion protein expression was induced with 0,5 mM IPTG for 3 hours at 28°C. Proteins were purified using the MagneGST Pulldown system (Promega) according to the manufacturers' instructions.

2.2.6.2 Histone enriched protein extraction from *Arabidopsis thaliana*

Histone enriched protein extracts were prepared with an adjusted protocol from (Yan et al., 2007): Approx. 3 g of whole seedlings 10 dag (LD) were homogenized in liquid N_2 and suspended in 30 ml of histone extraction buffer (0.25 M sucrose; 60 mM KCl; 15 mM NaCl; 5 mM $MgCl_2$; 1 mM $CaCl_2$; 15mM Pipes pH 6,8; 0.8% Triton X-100 and protease inhibitor cocktail (Sigma)). After centrifugation at 10000 g for 20 min, the pellet was resuspended in 6 ml of 0,4 M H_2SO_4 and centrifuged at 22000 g for 5 min. Then, 12 vol. of acetone were added to the supernatant and left at −20°C overnight to precipitate proteins. The proteins were then centrifuged at 8000 g for 15 min and resuspended with 4 M urea to a final volume of 1,2 ml. Neutralization of pH was achieved by adding NaOH.

2.2.6.3 Immunoblot procedures

The western blot procedures were performed as described in (Ausubel, 1996). Protein extracts were mixed with 6x SDS sample buffer (0,3 M Tris-HCl (pH6,8); 10 % (w/v) SDS; 30 % (v/v) glycerol; 0,6 M DTT; 0,01 % (w/v) bromophenol blue) and incubated at 95 °C for 5 min before loading on 15% (global histone analyses) or 10% SDS Polyacrylamide gels. Gel Electrophoresis was performed using a Mini-PROTEAN tetra cell (Biorad). Proteins were transferred to PVDF membrane (Roth) using a Trans-Blot SD semi dry transfer cell (Biorad).

The chemoluminescent substrate used for immonodetection with HRP coupled secondary antibody was SuperSignal West Pico (Thermo Fischer Scientific).

2.2.6.4 NuclearProtein extraction, immunoprecipitation and LC-MS/MS analysis

Nuclear proteins were isolated from 3 g of whole seedlings including root tissue grown in LD conditions. Nuclei were extracted without prior fixation of the tissue as described in (Kaufmann et al., 2010). The further procedure was performed as described in (Smaczniak et al., 2012): Nuclei were resuspended in 1 ml of lysis buffer (µMACS GFP Isolation Kit; Miltenyi Biotec) with Complete Protease Inhibitor Mixture (Roche) followed by sonication and preclearing by sonification. The nuclear protein extract was collected in a 2 ml tube followed by immunoprecipitation with 50 µl of µMACS anti-GFP MicroBeads (Miltenyi Biotec) for 1 h at 4 °C. Then the mixture was applied onto a calibrated µ Column (Miltenyi Biotec) and allowed to run through by a gravity flow. Immobilized beads were washed 6 times with 200 µl of lysis buffer and 2 times with 200 µl of wash buffer 2 (Miltenyi Biotec). Proteins were eluted from the beads with 50 µl of 8 M urea. *Arabidopsis thaliana* Col-0 wild-type plants were used as a control IP experiments. The control plants were processed and analyzed in the same way as the SCI1-GFP plants. To compare GFP samples with wild-type controls, we used two biological replicates each analyzed in two technical replicates. The IP eluate was diluted 4x with 50 mM ammonium bicarbonate (Fluka), followed by cysteine residues reduction with 7.5 mM DTT (Fluka) and carbamidomethylation with 15 mM iodoacetamide (GE Healthcare). Proteins were digested with 7,5 µg of trypsin (Trypsin Gold, Mass Spectrometry Grade; Promega) overnight at 37°C. The digest was purified and desalted using Oasis HLB µElution columns (Waters) before MS analysis on LTQ Orbitrap XL mass spectrometer (Thermo Scientific). Protein identification was performed by the database search with the Andromeda search engine.

The LC-MS/MS analysis and subsequent statistical data analysis was performed by C. Smaczniak (Smaczniak et al., 2012): The spectra acquired from the LTQ Orbitrap mass spectrometer were loaded to the MaxQuant protein quantification software (version 1.1.1.36; Max Plank Institute). After data processing, label-free quantification (LFQ) values were taken for further analysis. To distinguish specifically interacting proteins from the background, protein abundances were compared between the sample and the control groups, and False Discovery Rate (FDR) was used to control the error of the Student's t-test statistics used in this comparison. The proteins are arranged according to the ratio which is a measure for the enrichment of peptides in the GFP sample compared to the Col-0 control.

2.2.7 Chromatin Immunoprecipitation (ChIP)

ChIP analyses were carried out using seedlings 10 dag as described previously (Schubert et al., 2006). Antibodies used in this study were H3K27me3 (Millipore 07-449) and H3K4me3 (Diagenode pAB-003-050) (3 µg per sample), 20 µl/sample of Protein G Dynabeads (Invitrogen) were used. Input and Immunoprecipitated (IP) DNA was diluted 1:10, and 1 µl was used for qRT-PCR. Oligonucleotides are listed in Table 2.4, and values of IP are shown in reference to the Input values (%Input). To account for differences in immunoprecipitation efficiency between *sci1-1* and Col-0 samples, the %Input values were normalized to reference gene values. In case of the H3K4me3 Chip, the highly and ubiquitously expressed Actin gene which carries H3K4me3 served as reference locus (Lafos et al., 2011; Oh et al., 2008), while values of the H3K27me3 Chip were normalized to those of the FUSCA3 gene (At3g26790) which is expressed seed specifically and not expressed in seedlings were it is targeted by H3K27me3 (Makarevich et al., 2006).

2.2.8 Immunoprecipitation of GST-fusion proteins with biotinylated peptides

Protein domains expressed as GST-fusion proteins (vector backbone pGEX4T3 Table 2.9) in *E. coli* were used for immunoprecipitation experiments with biotinylated histone peptides listed in Table 2.1. Method was modified from (Shi et al., 2006): GST-fusion proteins were expressed and purified as described (2.2.6.1). 1 µg of fusion protein was incubated with 1 µg of biotinylated peptide in 300 µl Gozani-buffer (50mM Tris-HCl pH7,5; 300mM NaCl; 0,1% Igepal; 1mM PEFA and 1:100 plant protease inhibitor cocktail (Sigma-Aldrich P9599)) over night at 4°C on a rotating platform. Per sample 15 µl of Streptavidin magnetic beads (Invitrogen, prepared as instructed) were added and incubated for 1 h at 4°C on a rotating platform. Beads were washed 3 times with 500 µl Gozani-buffer at 4°C, then the bound proteins were eluted with 50 µl 0,1% SDS and incubation at 95°C for 5 min. The eluted proteins were studied by immunoblot analysis described in chapter 2.2.6.3.

2.2.9 ModifiedTM Histone Peptide Array

The GST-SCI1-PWWP fusion protein was hybridized to a ModifiedTM Histone Peptide Array (Active Motif) to reveal binding specificity to modified histone peptides. The hybridization was done in collaboration with A. Jeltsch (University of Bremen) as described previously (Dhayalan et al., 2010).

2.2.10 Luciferase assays

Luciferase reporter gene activity in plants carrying the FLC-LUC reporter gene was measured by spraying the plants with the substrate solution (5 mM D-Luciferin Potassium Salt in H_2O)

followed by 5 min of incubation in darkness and imaging of the luminescence with the NightOwl system (Berthold).

The Split-Luciferase assay was performed as described previously (Schatlowski et al., 2010). *Nicotiana benthamiana* epidermal leaf cells were transiently transformed with the Split-Luciferase constructs listed in Table 2.9. After induction of the ß-estradiol inducible constructs, the substrate (10 µM ViviRen; Promega) was infiltrated in the epidermal leaf cells. After 5 min of darkness the luminescence was measured with the NightOwl system (Berthold).

2.2.11 Microscopic analyses

Normarsky microscopy was carried out using the Axioskop (Zeiss), pictures were taken with the Zeiss Axiocam and analyzed using the Axio Vision software.

Scanning electron microscopy (SEM) was performed using the LEO (Zeiss) microscope and software.

Confocal microscopy was performed using the LSM510 Meta or the LSM710 from Zeiss. For analysis the software ZEN 2011 (Zeiss) was used. Excitation and emission spectra are listed below.

Fluorophore/Dyes	Excitation	Emission
GFP	488 nm argon laser	meta channel 496 nm – 550 nm
mCherry	561 nm cw laser diode	meta channel 571 nm – 636 nm (leaves) LP 575 nm (root)
DAPI	405 nm cw laser diode	BP 480 nm – 520 nm
PI	561 nm cw laser diode	LP 575nm

Table 2.13: Excitation and emission spectra

3 RESULTS

Polycomb-group (PcG) proteins are epigenetic regulators which were first discovered in *Drosophila melanogaster* as regulators of homeotic genes (Schwartz and Pirrotta, 2007; Schwartz and Pirrotta, 2007; Klymenko and Muller, 2004). PcG proteins are organized in high molecular weight complexes in *Drosophila*, namely POLYCOMB REPRESSIVE COMPLEX1 (PRC1), PRC2, PhoRC and PR-DUB, of which PRC2 is highly conserved in plants (Schatlowski et al., 2008; Scheuermann et al., 2010). Some aspects of PcG function has evolved independently in different organisms, and as many *Drosophila* PcG proteins have no sequence homologs in plants, it is likely that numerous plant specific proteins contribute to PcG mediated gene silencing in *Arabidopsis*.

In order to identify novel PcG proteins in plants, a yeast two-hybrid screen was carried out using the *Arabidopsis* PRC2 member CURLY LEAF (CLF) as bait (Schatlowski et al., 2010). One of the CLF interactors identified in the screen was the PWWP domain protein SWINGER/CURLY LEAF-INTERACTOR1 (SCI1), which is encoded by the gene locus At3g03140.

In general, *Arabidopsis* PcG mutants are characterized by pleiotropic and flowering time phenotypes, reduced H3K27me3 levels and misexpression of PcG target genes. To assess whether SCI1 has a function in PcG mediated gene regulation, interaction studies of SCI1 with additional PcG proteins were performed. Mutants of *SCI1* were analyzed to reveal PcG resembling and independent phenotypes. In order to discover possible genetic interactions of *SCI1* and *CLF*, double mutant analyses were carried out. Furthermore the global level of histone modifications in the mutants was analyzed, to reveal whether SCI1 is essential for PRC2 function in *Arabidopsis*. As *clf* and other PcG mutants show misexpression of meristem identity and homeotic genes, the expression of a subset of PcG target genes was analyzed in *sci1* single and double mutants. Furthermore the H3K27m3 and H3K4me3 levels at PcG target gene loci were analyzed in *sci1-1* compared to Col-0 by Chromatin Immunoprecipitation (ChIP).

SCI1 belongs to a plant specific gene family with three members that possibly exhibit redundant functions. Therefore double and triple mutants were obtained and phenotypically analyzed. Furthermore the expression pattern and the intracellular localization of the three SCI proteins were studied. The most prominent feature of the SCI1 protein is the amino-terminal PWWP domain which is implicated in chromatin binding (Maurer-Stroh et al., 2003;

Wang et al., 2009; Dhayalan et al., 2010). Histone peptide binding studies of the SCI1 PWWP domain were performed in order to reveal the binding specificity of the SCI1 PWWP domain.

3.1 SCI1 is a plant specific PWWP domain protein

Blast analyses revealed that *SCI1* belongs to a gene family of three members in *Arabidopsis thaliana*. The two homologs are accordingly named *SCI2* (At1g51745) and *SCI3* (At3g21295) (Figure 3.1 B). SCI1 is a plant specific protein with orthologs in monocots like *Oryza sativa* and *Brachypodium distachyon* as well as in dicots like *Medicago truncatula, Vitis vinifera* and *Populus trichocarpa*. *SCI1* encodes a protein of 769 amino acids (aa) and the protein comprises a predicted PWWP-domain at the amino-terminus of the protein (aa 13-75) and a nuclear localization signal (NLS; aa 393-409) in the central region (Figure 3.1 A and B). The greatest homology between the three SCI proteins can be found at the amino-terminus containing the predicted PWWP domain that is implicated in nucleic acid and chromatin binding (Figure 3.1 B) (Maurer-Stroh et al., 2003).

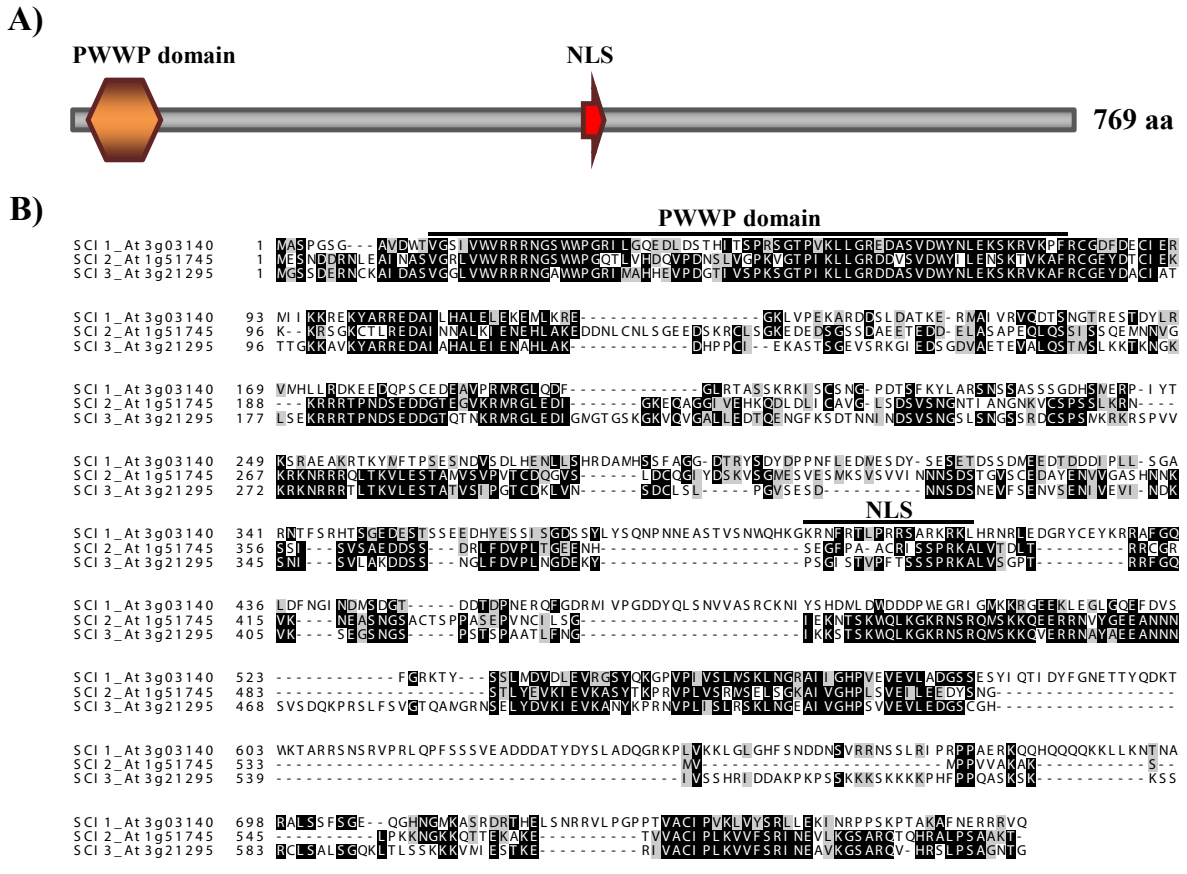


Figure 3.1: Protein structure of SCI1 and sequence alignment of SCI1, SCI2 and SCI3
(A) Protein structure of SCI1 with predicted amino-terminal PWWP domain (orange hexagon) and central NLS (red arrow). (B) Amino acid alignment of SCI1 and its homologs SCI2 and SCI3 in *Arabidopsis* showing a high level of conservation in the N-terminal region. Identical residues are shaded in black, residues with high similarity are shaded in grey and the predicted SCI1 PWWP domain and NLS are indicated above the sequence.

3.2 Yeast two-hybrid interaction studies with SCI1

SCI1 was isolated in a yeast two-hybrid (Y2H) screen as interactor of the PRC2 component CLF, and it can be speculated that SCI1 is involved in PcG mediated gene silencing in *Arabidopsis*. An interaction of SCI1 with additional PcG proteins would confirm this hypothesis, therefore the first approach to characterize SCI1 was to perform Y2H interaction studies with further PcG proteins and flowering time regulators.

3.2.1 SCI1 was isolated in a yeast-two-hybrid screen with the PRC2 component CLF

The Y2H screen that identified SCI1 as interaction partner of CLF was carried out using a truncated CURLY LEAF protein as bait. This bait protein lacked the C-terminal catalytic SET-domain, as the SET domain reduced the interaction strength with known interactors in yeast, and was fused to the Gal4-DNA-binding domain (CLFΔSET-BD) (Chanvivattana et al., 2004). In this screen 75 clones were identified that showed growth on selective media. Among these were established interaction partners of CLF like EMF2 and VRN2, which are components of the *Arabidopsis* PRC2 complex (Chanvivattana et al., 2004).

The initial SCI1(trunc)-AD (Gal4-activation domain) clone identified in the screen contained a truncated SCI1 protein devoid of the PWWP domain and starting with amino acid 273 of the 769 aa containing full-length protein. Therefore the interaction domain of SCI1 with CLF is located in the central or C-terminal region of the protein that contains the NLS and excludes the PWWP domain of SCI1.

3.2.2 SCI1 interacts with several PRC2 components in yeast

To confirm the interaction of SCI1 with CLF, and to test for further interactions with proteins involved in PcG mediated gene silencing and proteins regulating flowering time, the full length cDNA of SCI1 was cloned and fused to the Gal4-BD and Gal4-AD domains. Expression clones of the PRC2 components CLFΔSET, SWNΔSET, MEA, VRN2, EMF2 and MSI1, the MSI1 homolog FVE/MSI4 and the PRC1-like protein LHP1/TFL2 were generated. The interactions between SCI1 and these proteins were tested in a growth assay using dilution series on drop out media selecting for expression of the auxotrophic markers HIS3 and ADE2. The medium lacking the amino acids leucine and tryptophan (LW) was used as control to confirm the presence of the AD and BD constructs. The selective media additionally lack histidine (LWH) or histidine and adenine (LWHA), and select for HIS3 and ADE2 reporter gene activity. In case of an interaction of the AD and BD fusion proteins, the two reporter

genes are activated and the yeast cells are able to grow on selective media. In case of weak interaction, growth may only be visible on the less stringent LWH selective medium.

The truncated version of SCI1 (SCI1(trunc)-AD) was able to interact with CLFΔSET and SWNΔSET, while no interaction could be detected with TFL2, VRN2 and EMF2 (Figure 3.2 A and data not shown). The interaction assay could not be carried out using MSI1-BD construct, as this construct activated reporter gene expression when co-expressed with the empty AD-vector. The SCI1-BD construct containing the full length SCI1 protein showed additional interactions with MEA, VRN2 and FVE/MSI4, furthermore homodimerization of SCI1 could be observed (Figure 3.2 A). The interaction of SCI1-BD with VRN2-AD appears to be weaker than those with the other PcG proteins, as the yeast cells only showed growth on LWH medium. The interaction of CLFΔSET and SWNΔSET with VRN2 served as positive controls to ensure the functionality of the truncated constructs. Additionally it was shown that SWNΔSET has the ability to homodimerize in yeast, which contradicts a previous study carried out with the SWN full length protein (Figure 3.2 A) (Luo et al., 2000).

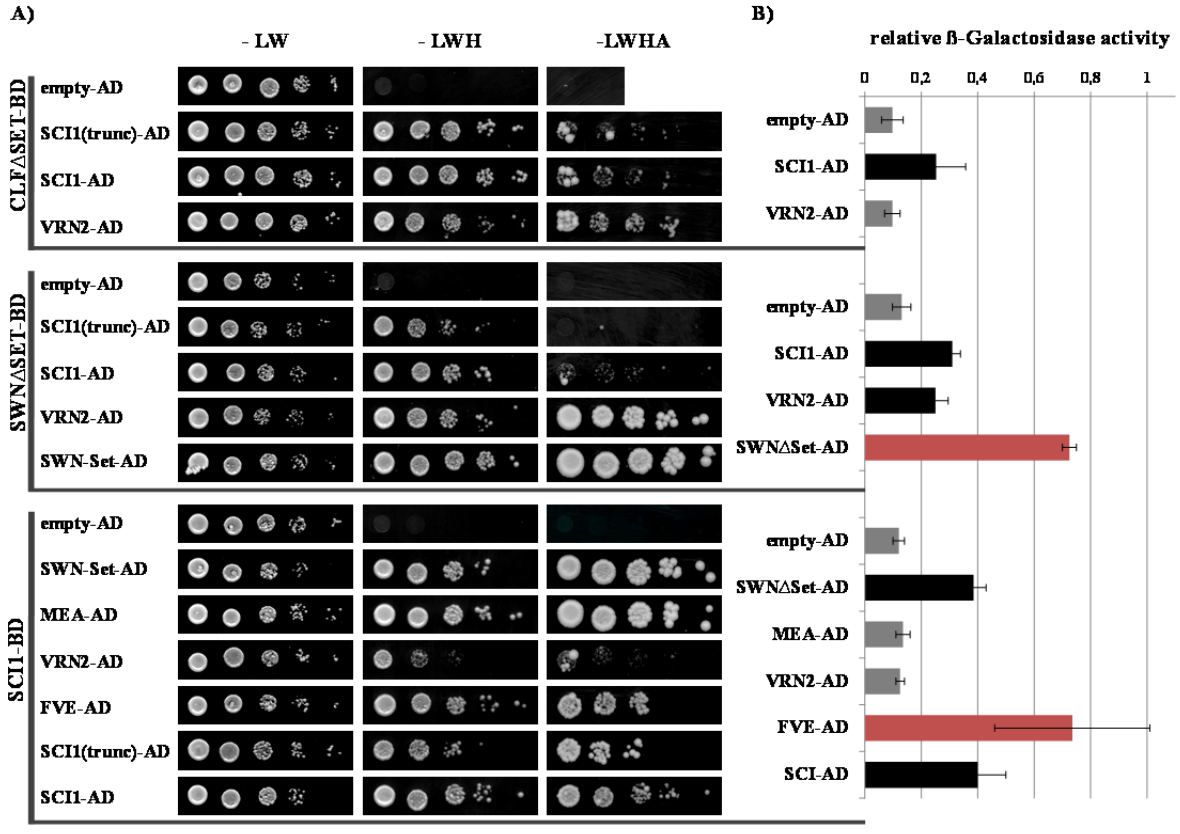


Figure 3.2: Yeast two hybrid interactions of SCI1 with several PcG proteins
A) Dilution growth assay of yeast cells transformed with Gal4-BD and Gal4-AD constructs on control medium selecting for presence of pGBKT7 and pGADT7 plasmids (-LW), selective medium (-LWH) and more stringent selective medium (-LWHA) selecting for HIS3 or HIS3 and ADE2 reporter gene activity, respectively. Negative controls are yeast cells transformed with the respective BD constructs and the pGADT7 vector (empty-AD).
B) Quantification of the protein-protein interactions via ONPG assay that measures the relative ß-Galactosidase activity. Grey bars represent the background activity; black and red bars show positive interactions and strong interactions, respectively. Average values are derived from at least two independent experiments; error bars represent standard deviation between the replicates.

3.2.3 Quantification of the SCI1 yeast two-hybrid interactions via ONPG assay

The colorimetric ONPG assay is used for quantification of protein-protein interactions in yeast by measuring the relative ß-Galactosidase activity. This enzyme is encoded by the *lacZ* reporter gene which is expressed upon protein interaction of the AD and BD constructs in the yeast two-hybrid system, and converts the transparent substrate ONPG to a yellow dye which is measured spectrophotometrically. The results of the ONPG assay largely affirmed the interactions detected in the growth assay (Figure 3.2 A and B). The interaction of SCI1-BD with FVE-AD as well as the homodimerization of SWNΔSET proved to be very strong, and the interactions of CLFΔSET-BD with SCI1-AD, SWNΔSET-BD with SCI1-AD and VRN2-AD as well as of SCI1-BD with SWNΔSET-AD and SCI1-AD were confirmed. The interaction of CLFΔSET-BD with VRN2-AD and of SCI1-BD with MEA-AD could not be observed in the ONPG assay, suggesting that these interactions are too weak or transient to be detected with this method (Serebriiskii and Golemis, 2000). In case of VRN2 previous studies used the VRN2-VEFS domain showing interaction with CLF and SWN. Similarly the EMF2 full length protein does not interact with CLFΔSET in yeast, while interaction can be detected with truncations of EMF2 (Chanvivattana et al., 2004). Since the interaction of SCI1 with CLFΔSET and SWNΔSET was observed using truncated proteins that lack the catalytic SET domains, the use of the full length MEA protein might be the reason that the interaction of SCI1-BD with MEA-AD could not be observed in the ONPG assay.

In conclusion it could be shown in the yeast two-hybrid interaction studies that SCI1 is able to interact with the three histone methyltransferase subunits of the plant PRC2 complex CLF, SWN and MEA, as well as weakly with the PRC2 subunit VRN2. Additionally an interaction with the MSI1 homolog FVE/MSI4 was observed, which was shown to act as a component of the autonomous flowering promotion pathway. This might suggest an involvement of SCI1 in flowering time regulation.

3.3 *In planta* analysis of the SCI1 protein

3.3.1 Intracellular localization of SCI1, CLF and SWN *in planta*

The intracellular localization of proteins *in planta* is important to gain insight into possible protein function, and to interpret results from yeast or *in vitro* experiments. Therefore fusion proteins of SCI1, CLF and SWN with fluorescent proteins were constructed and transformed into plants. The fluorescence was observed *in vivo* by confocal microscopy. Co-localization studies of SCI1 with CLF and SWN were performed to analyze whether the proteins show

overlapping intracellular localization patterns, which is essential for the interaction *in planta*. Furthermore truncations of the SCI1 protein were analyzed in order to identify protein domains important for specific sub-cellular localization patterns (Figure 3.2 A-D) Hence the full length cDNA of SCI1 (I. SCI1cDNA), the N-terminus of SCI1 including the PWWP domain (II. SCI1-incl.NLS), the C-terminal part of SCI1 devoid of the PWWP domain (III. SCI1ΔPWWP), and a central protein fragment lacking the PWWP domain but including the NLS (IV. SCI1-frag.4) were cloned as translational GFP fusion proteins under the control of a ß-estradiol inducible 35S-promoter (Bleckmann et al., 2010). Furthermore CLFΔSET (V) and SWNΔSET (VI) were fused to the fluorophore mCherry. These constructs were transiently transformed into *Nicotiana benthamiana* epidermal leaf cells (Figure 3.3 B and C). Upon induction of the construct, one could observe a nuclear localization of all constructs tested. Nevertheless there were differences in the localization patterns, as SCI1cDNA-GFP localizes to the nucleoplasm as well as to nuclear speckles in *N. benthamiana.* The localization ranged between nucleoplasmic expression with few speckles and exclusive localization to speckles (Figure 3.3 B I, I', I''). The construct SCI1-incl.NLS-GFP shows a more uniform expression in the nucleoplasm and less nuclear speckles and the SCI1ΔPWWP-GFP fusion protein localizes to the nucleoplasm and to the cytosol and no nuclear speckles are visible. SCI1-frag4-GFP mainly localizes to nuclear speckles and the nucleolus (Figure 3.3 B II-IV).

The co-localization studies performed in *N. benthamiana* showed that, compared to SCI1-GFP and SCI1incl.NLS-GFP, the fusion proteins CLFΔSET-mCherry and SWNΔSET-mCherry display a uniform nucleoplasmic localization (Figure 3.3 C). When CLFΔSET-mCherry or SWNΔSET-mCherry fusion proteins are co-expressed with SCI1-GFP, one can observe that the expression patterns of the fusion proteins overlap in the nucleoplasm and not in the nuclear speckles, which exclusively comprise SCI1-GFP expression (Figure 3.3 C I+V and I+VI). When co-expressing SCI1-incl.NLS-GFP with CLFΔSET-mCherry or SWNΔSET-mCherry, a stronger overlap can be observed, as the amount of nuclear speckles is reduced (Figure 3.3 C II+V and II+VI).

The intracellular localization of SCI1 in *Arabidopsis thaliana* was analyzed by construction of a translational fusion of the SCI1 genomic locus to GFP (SCI1pro::SCI1-GFP), which complements the *sci1* mutant phenotype (Figure 3.7). This construct was transformed into *Arabidopsis* plants, and the intracellular localization was analyzed in homozygous lines by confocal microscopy (Figure 3.3 D I-IV).

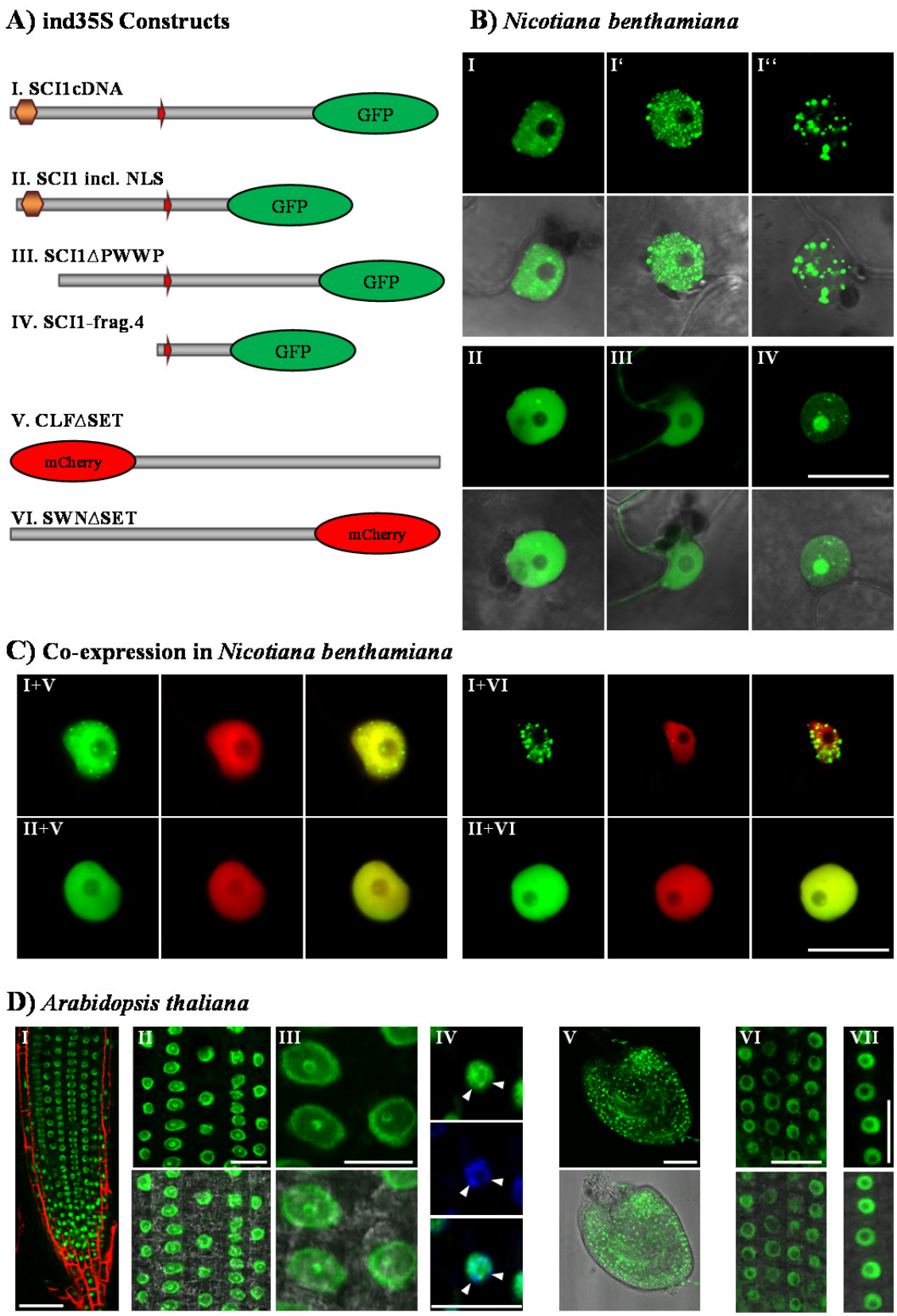


Figure 3.3:Intracellular localization of SCI1, SCI1 truncations, CLFΔSET and SWNΔSET
(A) Schemes of GFP and mCherry fusion proteins used for the transient localization studies in *N. benthamiana*. All constructs were expressed under the control of a ß-Estradiol inducible 35S-promoter. (B) Confocal analysis of the transiently expressed constructs, depicted in (A), in *N.benthamiana* leaf epidermal cells. I-I'' show different localization patterns of SCI1-GFP in *N. benthamiana*, in II-IV the intracellular localization of the constructs SCI1incl.-NLS-GFP, SCI1ΔPWWP-GFP and SCI1frag.4-GFP are depicted. White scale bar represents 20µm.

(C) Co-localization studies of indicated GFP and mCherry fusion proteins depicted in A) in *N. benthamiana* leaf epidermal cells. White scale bar represents 20µm. (D) Expression analysis by confocal microscopy of the following transgenes in *A. thaliana* (Col-0): (I-III) SCI1pro::SCI1-GFP in root cells, IV) SCI1pro::SCI1-GFP in root cells with DAPI co-staining, arrowheads point to position of chromocenters; V) SCI1pro::SCI1-GFP in unfertilized ovules; (VI) 35S::GFP-CLF in root cells (*clf-50*, WS) and (VII) SWN::SWN-GFP in root cells. White scale bars in D) represent: I) 50µm; II) 20µm III-IV) 10µm; V) 50µm; VI-VII) 20µm.

SCI1-GFP fluorescence in *Arabidopsis* root cells was observed in the nucleus and to a lesser extent in the cytoplasm. In contrast to the transient expression in *N. benthamiana* no nuclear speckles were visible. Inside the nucleus the localization of SCI1-GFP appears to be similar to proteins localizing to the nuclear periphery (Figure 3.3 D II-III). DAPI co-staining showed, that SCI1-GFP fluorescence does not overlap with the chromocenters which are comprised of heterochromatin (Figure 3.3 D IV), thus the SCI1-GFP protein localizes to regions in the nucleus that are of euchromatic composition. The localization of the fusion protein does not change in the *sci1-1* single or *sci*-triple mutant backgrounds (Chapter 3.4 and 3.7). In unfertilized ovules, a nuclear SCI1-GFP fluorescence can be observed in the entire tissue (Figure 3.3 D V).

The intracellular localization of CLF and SWN in *Arabidopsis* was analyzed using stable lines comprising a 35S::GFP-CLF and a SWN::SWN-GFP transgene, respectively. In root cells both fusion proteins localize to the nucleus while the GFP-CLF fusion protein also shows cytoplasmic localization (Figure 3.3 D VI-VII).

In conclusion it can be noted that SCI1, CLF and SWN co-localize to the nucleus in plants, which is essential for protein-protein interaction. The fact that SCI1-GFP localizes to so far unknown nuclear speckles in *N. benthamiana* and to the nuclear periphery in *Arabidopsis* root cells might suggest, that the interaction of SCI1 with CLF and SWN only takes place at the interface of these sub compartments.

3.3.2 Interaction studies of SCI1 with CLF and SWN *in planta*

In order to confirm the interaction of SCI1 with CLF and SWN *in planta*, which is important to show that results from yeast studies reflect the native situation, the SCI1cDNA, CLFΔSET and SWNΔSET constructs from the Y2H and intracellular localization studies were fused to the N- and C-terminal part of the luciferase *Renilla reniformis* (Rluc) to be used in the split-luciferase assay (Figure 3.2; Figure 3.3) (Schatlowski et al., 2010). In case of an interaction of the tested proteins, the N- and C-terminal parts of Rluc come into close proximity, which restores the enzyme activity. The substrate Coelenterazine is catalyzed by the luciferase, what leads to the emission of photons that can be measured as luminescence.

The fusion proteins were transiently expressed in *Nicotiana benthamiana* leaf epidermal cells, the full length luciferase served as positive transformation control. The proteins of H2A and H2B fused to N-Luc and C-Luc fragments, respectively, are known to interact when co-expressed in *N. benthamiana* and were therefore used as positive controls (Figure 3.4) (Fujikawa and Kato, 2007). The combination of SCI1-N-Luc with H2B-C-Luc, as well as CLFΔSET-C-Luc and SWNΔSET-C-Luc with H2A-N-Luc, served as negative controls. In case of co-expression of SCI1-N-Luc with CLFΔSET-C-Luc and SWNΔSET-C-Luc, no detectable luciferase activity could be observed (Figure 3.4). This result was also observed when SCI1 was fused to C-Luc and CLFΔSET and SWNΔSET were fused to N-Luc (data not shown), suggesting that the full length SCI1 protein does not interact with the truncated CLFΔSET and SWNΔSET proteins in the split-luciferase assay. This finding can have several reasons, one of them being the intra-nuclear localization of SCI1 to nuclear speckles (Figure 3.3 A and B). The analyses of SCI1 truncations performed in *N. benthamiana* suggest that the speckle formation is a feature of the functional SCI1 protein. As CLFΔSET and SWNΔSET uniformly localize to the nucleoplasm in *N. benthamiana,* the interaction with SCI1 might not be able to take place under these conditions (Figure 3.3 A and B). Another problem might be the nature and the position of the tagged luciferase protein, which could sterically inhibit the protein-protein interaction. Further interaction analyses with complementing and therefore functional fusion proteins in *Arabidopsis thaliana* and *in vitro* pulldown experiments will have to prove whether SCI1 is able to interact with CLF and SWN *in planta.*

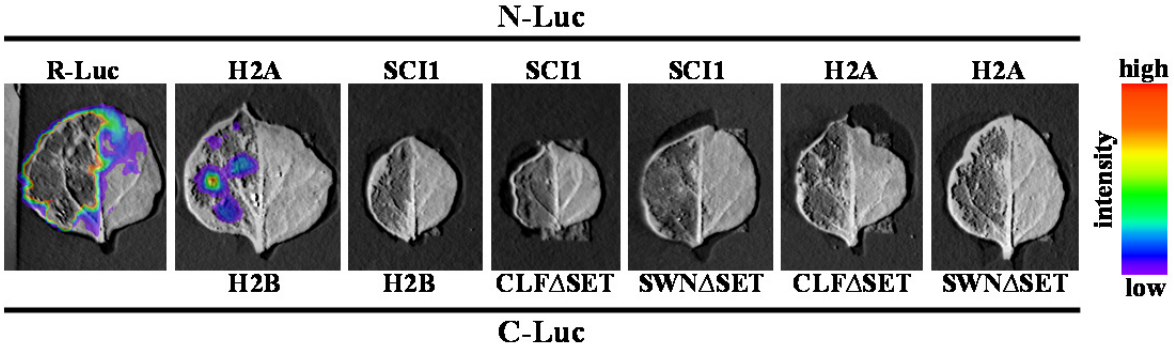


Figure 3.4: Split-luciferase complementation assay with SCI1, CLF and SWN
Epidermal cells of *N.benthamiana* leaves were co-transformed with indicated N-Luc and C-Luc constructs. After substrate infiltration into the left abaxial side of the leaves, luminescence was monitored using the Nightowl luciferase imaging system (Berthold); the intensity is represented by a false color code. The intensity of the full length *Renilla reniformis* Luciferase (Rluc) appears translucent, since the intensity is higher than the scale.

3.3.3 Immunoprecipitation with SCI1-GFP reveals possible complex partners of SCI1

The interaction studies performed in yeast showed that SCI1 interacts with PcG proteins as well as with the autonomous pathway protein FVE/MSI4 (Figure 3.2). These interaction

partners, as most chromatin regulating proteins, have been shown to act as part of multimeric complexes in *Arabidopsis* (Guitton and Berger, 2005; Jeon and Kim, 2011; Pazhouhandeh et al., 2011). SCI1 might either be associated to the PRC2 complexes or be an integral member, but SCI1 could also be part of a different complex that associates with the proteins mentioned above.

In order to discover the *in planta* interaction partners of SCI1, an immunoprecipitation (IP) experiment with nuclear protein extract from SCI1pro::SCI1-GFP and Col-0 plants using an anti-GFP antibody was performed. After tryptic digestion of the isolated proteins, the extracts were analyzed with liquid chromatography-mass spectrometry (LC-MS/MS). The statistical analysis of four replicates yielded 110 proteins significantly enriched in the SCI1pro::SCI1-GFP samples compared to the Col-0 wildtype samples. The significantly enriched proteins are shown in Table 7.1 and Table 7.2, and selected putative interaction partners in Table 3.1. The interaction partners are listed according to the ratio, a measure for the certainty of enrichment in the IP sample compared to the control.

#	Ratio	AGI-code	name	protein data
1	820,30	At2g40930	UBP5	Ubiquitin carboxyl-terminal hydrolase 5
2	82,77	At1g79350	EMB1135	embryo development ending in seed dormancy
5	21,58	At3g03140	SCI1	SCI1
9	15,61	At5g40340		PWWP domain-containing protein
11	14,85	GFP	GFP	GFP-like fluorescent chromoprotein FP506 (zFP506)
PcG or PcG-like proteins				
21	8,08	At5g58230	MSI1	WD-40 repeat-containing protein MSI1 (MULTICOPY SUPPRESSOR OF IRA 1)
16	11,88	At4g06634		Uncharacterized zinc finger protein
37	4,98	At1g79020		Enhancer of polycomb-like protein
Chromatin remodeling				
8	16,23	At5g14170	CHC1	SWI/SNF complex component SNF12 homolog
15	12,04	At5g18620		Putative chromatin remodelling complex ATPase chain ISWI
19	10,33	At2g47620	SWI3A/CHB1	SWI/SNF complex subunit SWI3A (AtSWI3A)
Nuclear envelope				
27	6,50	At5g65770	LINC4	LITTLE NUCLEI4 (Putative nuclear matrix constituent protein 1-like protein)
46	3,63	At1g67230	LINC1	LITTLE NUCLEI1
ARID Transcription factors				
6	20,91	At2g17410	ARID3	AT-rich interactive domain (ARID)-containing protein 3
12	14,16	At1g76510	ARID5	ARID domain-containing protein 5
13	12,55	At3g43240	ARID4	ARID domain-containing protein 4
20	9,83	At1g20910	ARID6	ARID domain-containing protein 6
Histone proteins				
24	6,76	At3g20670		Probable histone H2A.2 (HTA13)
32	5,86	At1g51060		Probable histone H2A.1 (HTA10)
36	5,04	At5g02560		Probable histone H2A.4 (HTA12)
38	4,94	At5g27670		Probable histone H2A.5 (HTA7)
48	3,41	At1g08880		Probable histone H2AXa (HTA5)
50	3,33	At4g40030		Histone H3

Table 3.1: List of selected proteins identified as putative complex partners of SCI1
Putative complex partners are listed according to the ratio and are depicted including ranked number, ratio value, AGI code, gene name and predicted protein function as retrieved form Uni-Prot and TAIR databases. Both parts of the bait protein, SCI1 and GFP, were identified and marked in red and green. PRC2 member MSI1 is marked in blue and LINC1 in orange.

The fact that SCI1 (#5, red) and free GFP (#11, green) were identified proves that SCI1-GFP was expressed and present in the nuclear protein extract after immunoprecipitation with anti-GFP antibodies (Table 3.1). Therefore the other proteins identified are potential complex or interaction partners of SCI1. The protein most enriched in the SCI1-GFP sample was the Ubiquitin carboxyl-terminal hydrolase 5 (UBP5 #1) which is predicted to be involved in ubiquitin mediated protein degradation. Furthermore a protein of unknown function encoded by At1g79350 (#2) was found to interact with SCI1, which is involved in embryo development and seed dormancy. The identification of a PWWP domain protein of unknown function (#9) encoded by At5g40340 suggests that SCI1 interacts with additional chromatin or DNA binding proteins. The protein MSI1 is a component of the PRC2 complex, and was isolated in the immunoprecipitation (#21), which confirms the hypothesis that SCI1 interacts with PRC2. Furthermore two uncharacterized proteins that show homology to PcG associated proteins from *Drosophila* (#16; 37) were isolated as SCI1 interactors. Further proteins involved in chromatin remodeling and several ARID-domain transcription factors were found to interact with SCI1 (Table 3.1).

Unfortunately neither CLF nor SWN were identified, and therefore the interaction with SCI1 could not be confirmed in this experiment. A parallel experiment using nuclear extract from SWN::SWN-GFP transgenic plants was carried out, in which the bait protein SWN could be detected, but no other PRC2 members (data not shown). This result implies that under the applied experimental conditions the PRC2 complex might not be stable, and thus also not detectable in the SCI1-GFP samples. Nevertheless, another parallel experiment with an unrelated putative RNA processing factor as bait protein (TGH-GFP) was successful, and yielded a list of totally different interacting proteins (data not shown). This finding suggests that the interaction partners found with SCI1-GFP are genuine and not an artifact caused by protein binding to the GFP-tag.

SCI1-GFP was furthermore found to interact with LITTLE NUCLEI1 and 4 (LINC1 and 4; # 27; 46), and LINC1-GFP was described to localize to the nuclear periphery in *Arabidopsis* similar to SCI1-GFP, and was described to be involved in nuclear size regulation (Dittmer et al., 2007; vanZanten M. et al., 2011). Additionally also several histone proteins were detected, which implies that the PWWP domain of SCI1 indeed harbors histone binding capacity. The interaction of these proteins with SCI1 will have to be proven in separate experiments. The complex isolation approach yielded interesting candidates which might help to decipher SCI1 function in plant development.

3.4 Analysis of *sci1* mutant alleles in *Arabidopsis thaliana*

The isolation of *sci1* mutant alleles in *Arabidopsis* is important to study the gene function by analysis of morphological and molecular mutant phenotypes. Furthermore *sci1* mutant alleles are a powerful tool to investigate genetic interactions, for example with the protein interaction partners identified in yeast, CLF, SWN, MEA and FVE/MSI4 (Figure 3.2).

In the model plant *Arabidopsis* an extensive database of T-DNA insertion mutants is available that can be screened for insertions in the gene of interest (http://signal.salk.edu/cgi-bin/tdnaexpress) (Alonso et al., 2003). In case of *SCI1* the available T-DNA insertion lines *sci1-1, sci1-2* and *sci1-gabi* were analyzed. The *sci1-2* insertion line showed an abberant segregation probably due to a background mutation and was not analyzed further. In all three cases the T-DNA insertion is localized in the fourth exon, and as *sci1-gabi* was only recently identified, most of the further studies were carried out using the *sci1-1* allele (Figure 3.5 A and B).

Since the T-DNA insertion of all three identified alleles is localized in the last exon of SCI1, it cannot be excluded that these do not resemble null alleles. In order to gain a *sci1* null mutant, TILLING (Targeting Induced Local Lesions IN Genomes) mutants were ordered from the Seattle Tilling Project (STP) and analyzed (http://tilling.fhcrc.org) (Till et al., 2003). In this approach six lines with point mutations in the N-terminus of SCI1 were identified, of which five alleles could be verified (*sci1-3* to *sci1-5, sci1-7* and *sci1-8*) (Figure 3.5). The *sci1-3* allele exhibits a point mutation (R46Stop) in the PWWP-domain of SCI1 that leads to a premature stop codon. The point mutations in the other TILLING alleles cause amino acid exchanges indicated in Figure 3.5 A), but do not lead to a premature stop codon. As the characterization of the additional TILLING alleles revealed that none of them showed a severe phenotype, only the potential null allele *sci1-3* was further analyzed (Figure 3.5 B).

To investigate whether the *sci1* mutants exhibit a reduction or even loss of *SCI1* transcript, semi-quantitative RT-PCR analyses were carried out (Figure 3.5 C). The amplified DNA fragments are indicated in Figure 3.5 A) by black lines (I. to III.), and the housekeeping gene *eIF4A* served as loading control. All three *sci1* mutant alleles exhibit *SCI1* transcript at the 5'end of *SCI1* (I.) prior to the insertions / point mutation at a similar level as the Col-0 wildtype. SCI1 transcript in the fourth exon (II.) is not detectable in *sci1-1* and *sci1-gabi,* but present in *sci1-3*. At the 3' end of *SCI1* (III.), *sci1-1*and *sci1-3* show reduced transcript levels while *sci1-gabi* shows a similar transcript level compared to Col-0.

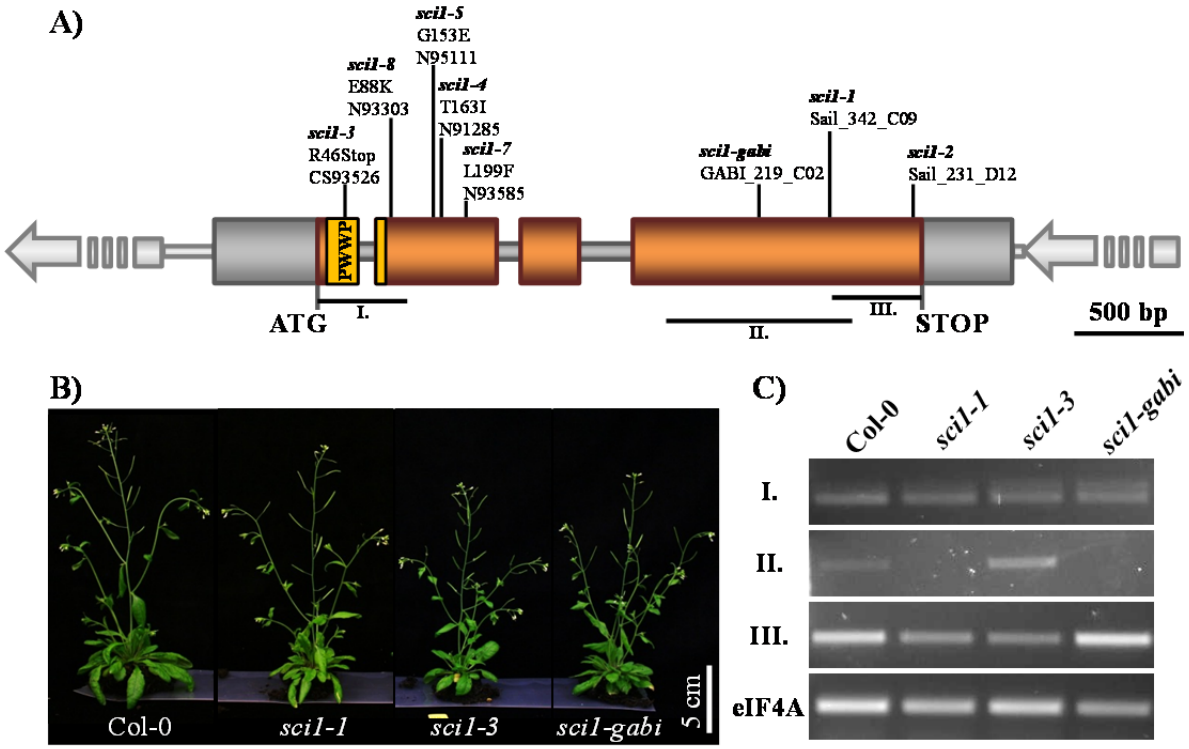


Figure 3.5: Position, phenotype and transcript abundance in *sci1* insertion mutants
(A) Genomic locus of *SCI1,* exons are depicted in orange, area encoding the PWWP domain in yellow, introns and 5' and 3' UTR in dark grey. Neighboring genes are shown in light grey (not in scale). The position and accession numbers of different *sci1* T-DNA insertion and TILLING alleles are indicated above. The amino acid exchanges in the TILLING alleles caused by point mutations are denoted. Black lines I.-III. resemble DNA fragments amplified in C) by semi-quantitative RT-PCR. (B) Phenotype of *sci1-1, sci1-3* and *sci1-gabi* compared to Col-0. Plants were grown in LD conditions. (C) Semi-quantitative RT-PCR on RNA from whole seedlings 10 dag in order to analyze the *SCI1* transcript level in the *sci1* mutant alleles. Amplified regions are indicated in A), the housekeeping gene *eIF4A* (AT1G54270) served as loading control.

The presence of transcript at the 5' region can be explained by endogenous promoter activity which is not disrupted by the insertions / point mutation, the transcript detected at the 3' region (III.) might derive from read-out transcriptional activity of the inserted T-DNAs that carry constitutive promoters. The finding that *sci1-1* and *sci1-gabi* do not exhibit detectable transcript in region II shows that in these mutants no full length *SCI1* transcript can be found, which suggests a severe reduction of SCI1 activity in these mutants. The partial transcripts found 5' and 3' of the insertions may encode for partial SCI1 protein fragments potentially providing residual protein function. In contrast *sci1-3* most likely is a null mutant due to the premature stop codon in the SCI1 PWWP domain. Since all three *sci1* alleles exhibit similar morphological phenotypes (Figure 3.7), it can be proposed that they are *sci1* loss of function alleles

3.5 Expression analysis of *SCI1, SCI2* and *SCI3 in planta*

To examine the expression pattern of *SCI1* and its homologs *SCI2* and *SCI3* throughout plant development, RNA was isolated from several different organs of Col-0 wildtype plants (seedlings, rosette leaves, roots, cauline leaves, flowers and siliques). The RNA was transcribed to cDNA, and transcript abundance of *SCI1, SCI2* and *SCI3* was tested via semi-quantitative RT-PCR compared to the housekeeping gene *eIF4A* that served as loading control (Figure 3.6 C). The three genes are expressed in all tissues tested. While *SCI1* and *SCI2* transcript levels are relatively low in rosette leaves, cauline leaves and siliques, *SCI3* exhibits more homogenous expression levels in all tissues with a peak in flowers.

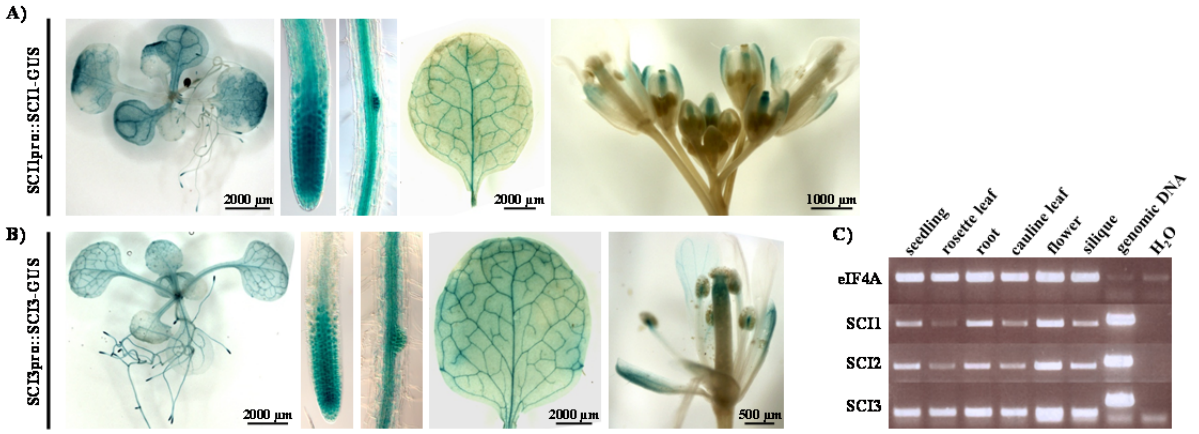


Figure 3.6: Expression pattern of *SCI1, SCI2* and *SCI3*
(A) ß-Glucuronidase activity in seedlings, primary root tip, emerging side root, leaf and inflorescence of SCI1pro::SCI1-GUS reporter line. (B) ß-Glucuronidase activity in SCI3pro::SCI3-GUS reporter line. (C) Transcript abundance of *SCI1*, *SCI2* and *SCI3* compared to the housekeeping gene eIF4A (AT1G54270) was analyzed via semi-quantitative RT-PCR. RNA was isolated from different organs of Col-0 plants grown in LD conditions.

To gain a more precise insight into the expression pattern of SCI1 and its homolog SCI3, the whole genomic loci of SCI1 (from 1534 bp upstream of ATG) and SCI3 (from 1062 bp upstream of ATG) were translationally fused to a C-terminal uidA gene, and introduced into Col-0 wildtype plants (SCI1pro::SCI1-GUS; SCI3pro::SCI3-GUS). Several transformants were analyzed for segregation of the transgene, and transgenic lines exhibiting only one locus of transgene insertion were used for further studies. ß-Glucuronidase activity observed in SCI1pro::SCI1-GUS and SCI3pro::SCI3-GUS transgenic lines were quite similar (Figure 3.6). In young seedlings SCI1-GUS and SCI3-GUS expression is observable the vasculature of the root and the shoot, the cotyledons and rosette leaves. ß-Glucuronidase activity can furthermore be observed in the primary root tip as well as in newly formed side roots and the root vasculature. In flowers, both transgenes are expressed in sepals and carpels, additionally

SCI3-GUS activity is observed at the tip of the filaments. The expression of SCI1-GUS is lost in older carpels.

In summary, SCI1 and SCI3 are expressed throughout plant development, mostly restricted to the vasculature. The reporter gene activity in meristematic regions and floral organs suggests that *SCI1* and *SCI3* are involved in cell differentiation, organ identity and plant reproduction. The overlapping expression patterns of *SCI1* and *SCI3* observed in the reporter gene study, and of all three *SCI* genes detected via semi-quantitative RT-PCR, is consistent with a potential redundant function of SCI1, SCI2 and SCI3.

3.6 Analysis of flowering time and root phenotypes in *sci1* mutants

3.6.1 Loss of SCI1 function results in an early flowering phenotype

The onset of flowering and thus the switch from vegetative to reproductive growth is tightly regulated in plants. Histone modifications e.g. mediated by PcG proteins play an important role in the regulation of flowering time, and most PcG mutants show an alteration of flowering time compared to the wildtype. In general, the induction of early flowering can be caused by ectopic expression of the floral promoters *FT* and *SOC1,* meristem identity genes like *APETALA1* (*AP1*) and *LEAFY* (*LFY*) or the floral-organ identity gene *AGAMOUS* (*AG*). Furthermore the down regulation of floral repressors like *FLOWERING LOCUS C* (*FLC*) or its homologs *MADS AFFECTING FLOWERING 1-5* (*MAF1-5*), which among others causes the induction of the floral promoters *FT* and *SOC1,* results in an early flowering phenotype (Koornneef et al., 1998; Yoo et al., 2011; Dennis and Peacock, 2007). PcG proteins are directly involved in the repression of *AG* and *FT* as well as of the floral repressor *FLC,* and thus prevent early as well as late flowering (Goodrich et al., 1997; Jiang et al., 2008). Consequently it is very interesting to assess whether the PRC2 interacting protein SCI1 is also affected in flowering time. To test this, the number of rosette leaves formed before bolting was determined in *sci1* mutants compared to the Col-0 wildtype. To exclude that an alteration in flowering time is caused by misregulation of the photoperiod pathway, the measurement was performed in inductive long day (LD) as well as in non-inductive short day conditions (SD).

sci1-1 shows an early flowering phenotype in LD as well as in SD conditions, and therefore appears to be involved in the regulation of flowering time in *Arabidopsis* (Figure 3.7). To exclude an allele specific effect, two additional *sci1* alleles were analyzed and showed a similar phenotype. The SCI1pro::SCI1-GFP transgene is able to rescue the early flowering

phenotype in LD conditions in respect to rosette leaf number. Taken together it is likely that the early onset of flowering in *sci1* mutant alleles is caused by a disruption of *SCI1* gene function.

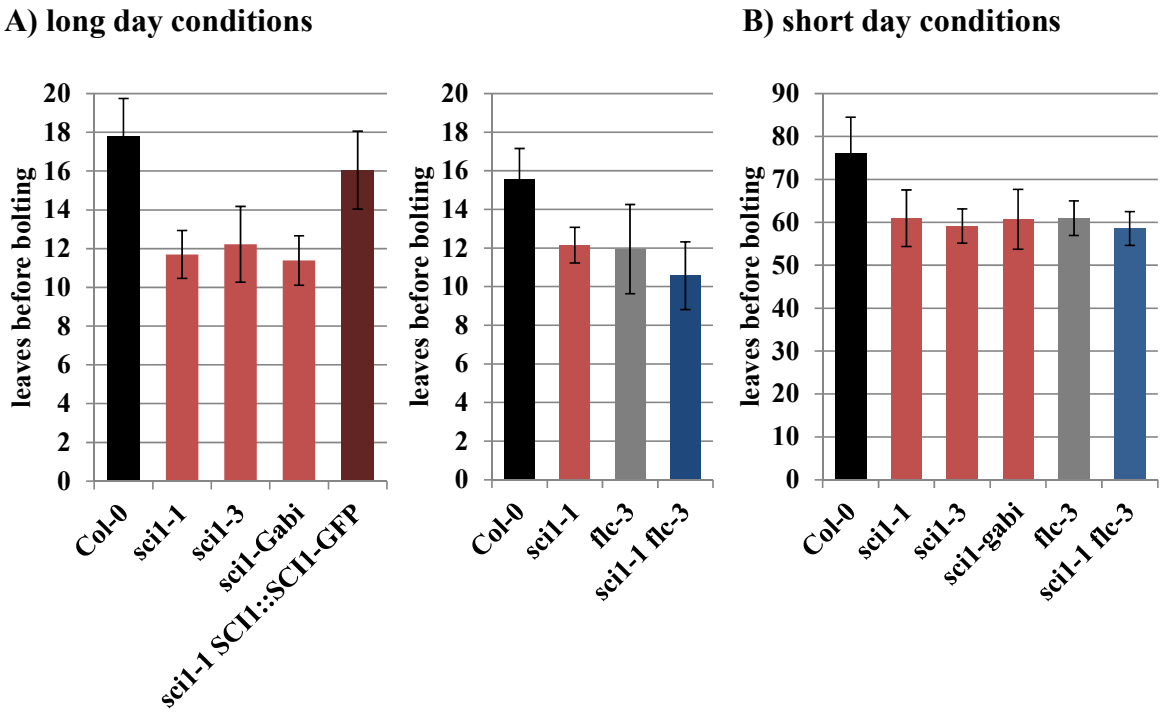


Figure 3.7: Flowering time analysis of *sci1* and *flc* mutants in LD and SD conditions
The average rosette leave number prior to bolting of Col-0 (black) and indicated mutant plants grown in long day and short day conditions are depicted in histograms. Mean values and the standard deviation were obtained from 13 to 24 individual plants. In A) the diagram on the right hand side shows the average and standard deviation of three independent experiments.

One key regulator of flowering time in *Arabidopsis* is the MADS box transcription factor *FLC* which encodes for a potent repressor of the floral transition. *FLC* expression is for example regulated by the autonomous pathway and the vernalization pathway, the latter involving PcG mediated tri-methylation of H3K27 (Amasino and Michaels, 2010). The null mutant *flc-3* shows a reduction of rosette leaves in LD and SD conditions similar to the reduction observed in *sci1* mutants (Figure 3.7) (Michaels and Amasino, 1999; Martin-Trillo et al., 2006). The analysis of *sci1-1 flc-3* revealed a similar reduction of rosette leaf formation as the respective single mutants in both conditions, which implies that SCI1 and FLC act epistatically.

3.6.2 *FLC* expression is downregulated in *sci1-1* mutant background

The results of the flowering time analysis implicate a misregulation of *FLC* in *sci1* mutants. To address this hypothesis, a reporter gene analysis and qRT-PCR analyses were carried out. In the transgene approach a line harboring a translational fusion of *FLC* with a firefly *LUCIFERASE* reporter gene (*FLC-LUC*) in *FRIGIDA*+ (*FRI*+) background was crossed with *sci1-1* mutants (Mylne et al., 2004). The luciferase reporter is inserted in the sixth exon of *FLC* under control of the endogenous promoter, and while the reporter is expressed similar to *FLC* the transgene is not functional (Mylne et al., 2004). FRIGIDA is a coiled coil protein that acts dominantly to elevate *FLC* expression, and thereby prolongs the vegetative growth phase (Johanson et al., 2000; Simpson and Dean, 2002).

Luciferase activity was monitored at different developmental time points in *FRI*+, *sci1-1* FRI+, *sci1-1* and Col-0 plants all carrying the *FLC-LUC* reporter gene (Figure 3.8). Col-0 plants without transgene served as negative control.

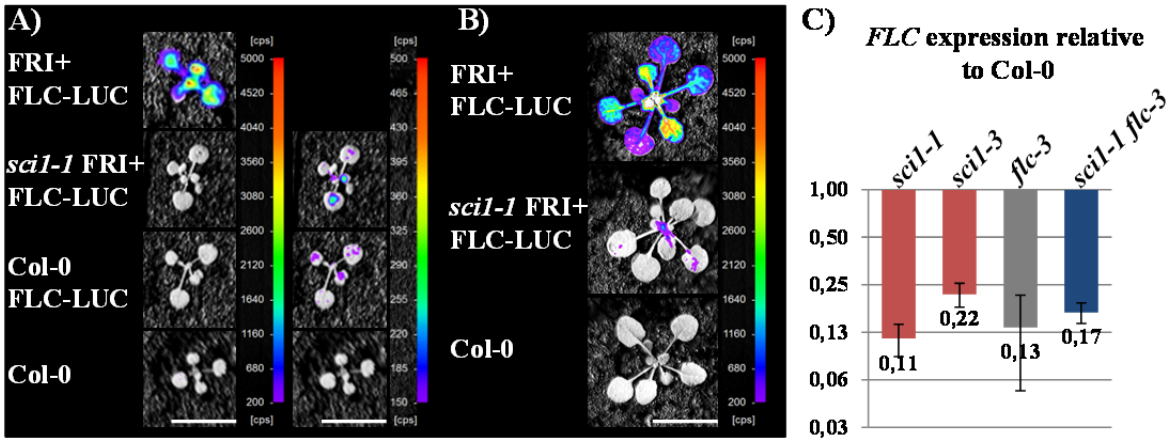


Figure 3.8: Reduced FLC expression in *sci1* mutant background
FLC-LUC reporter gene activity was visualized using the Nightowl Imaging system (Berthold, Bad Wildbad). A corresponding heat map is depicted to the right of each experiment; the false color code represents luciferase intensity in counts per second (cps) (note different scaling). (A) Seedlings of 21 dag were imaged with an exposure time of 5 min (left) and 10 min (right) (4x4 binning). (B) Seedlings of 28 dag were imaged with an exposure time of 5 min (1x1binning). White scale bars represent 1,5 cm. (C) qRT-PCR based expression analysis of *FLC* in *sci1* and *flc* single and double mutants compared to Col-0. cDNA was prepared from whole seedlings 10 dag grown in long day conditions. Expression values were normalized to At4g34279 (nblack) and are shown relative to Col-0. The average and standard deviation were obtained from at least two biological replicates each consisting of three technical replicates.

During vegetative growth *FLC* is expressed in the vasculature of the leaves and in the apex similar to *SCI1* (Greb et al., 2007), and the spatial expression pattern does not change in *sci1-1 FRI*+ background compared to *FRI*+ (Figure 3.8 A). However, strong reduction of reporter gene activity was observed in *sci1-1 FRI*+ background, which also led to a precocious termination of reporter gene expression (Figure 3.8 A and B).

In order to quantify the reduction of *FLC* transcript in *sci1* mutants compared to the *flc-3* null mutants, qRT-PCR analyses were carried out with cDNA obtained from 10 day old seedlings grown in LD conditions. This analysis revealed that, compared to Col-0, the expression of *FLC* is up to nine fold reduced in the *sci1* single mutants. This is comparable to the level detectable in *flc-3* null mutants, and *sci1-1 flc-3* double mutants exhibit similar *FLC* expression levels as the single mutants (Figure 3.7). These results coincide with the reporter gene analysis and strongly suggest that the reduction of *FLC* expression is causal for the early flowering phenotype of *sci1* mutants.

The finding that *FLC* transcript is strongly reduced in *sci1* mutant background could be caused by several scenarios. On the one hand the regulation might be direct, for example by binding of SCI1 to *FLC* chromatin. On the other hand, SCI1 might indirectly activate *FLC* expression by repressing a repressor or activating an activator. To test the possibility of an indirect regulation of *FLC, sci1-1* mutants were crossed in *FRIGIDA* (*FRI*+) and *fve-3* mutant backgrounds. While FRI is a strong activator of *FLC*, the autonomous pathway component FVE/MSI4 represses *FLC* and therefore flowering time.

The introduction of the dominant *FRI*+ allele into Col-0 plants results in an extremely late flowering phenotype (Figure 3.9 A). *flc* mutants are epistatic to *FRI*+, as flowering time of *flc FRI*+ plants is similar to that of *flc* single mutants (Michaels and Amasino, 1999; Willmann and Poethig, 2011). The introduction of *sci1-1* into *FRI*+ background results in an intermediate flowering time between *sci1-1* and *FRI*+ (Figure 3.9 A and B). This result coincides with the FLC-LUC reporter gene analysis (Figure 3.8 A and B) and suggests that there is residual *FLC* expression in *sci1-1* FRI+ plants, and that SCI1 and FRIGIDA regulate *FLC* independently of each other.

If SCI1 regulated *FLC* through the autonomous pathway component FVE/MSI4, an epistatic flowering time phenotype of the s*ci1-1 fve-3* double mutants would be expected. Nevertheless, the double mutants show an intermediate phenotype compared to the single mutants, suggesting that SCI1 regulates *FLC* independently of FVE/MSI4 (Figure 3.9 A). In previous analyses it was shown that the flowering time of *fve* mutants is affected by *flc* in a dosage dependent manner (Sanda and Amasino, 1996). This implies that the intermediate flowering time of *sci1-1 fve-3* mutants is caused by a reduction rather than a loss of *FLC* expression in *sci1-1* mutants.

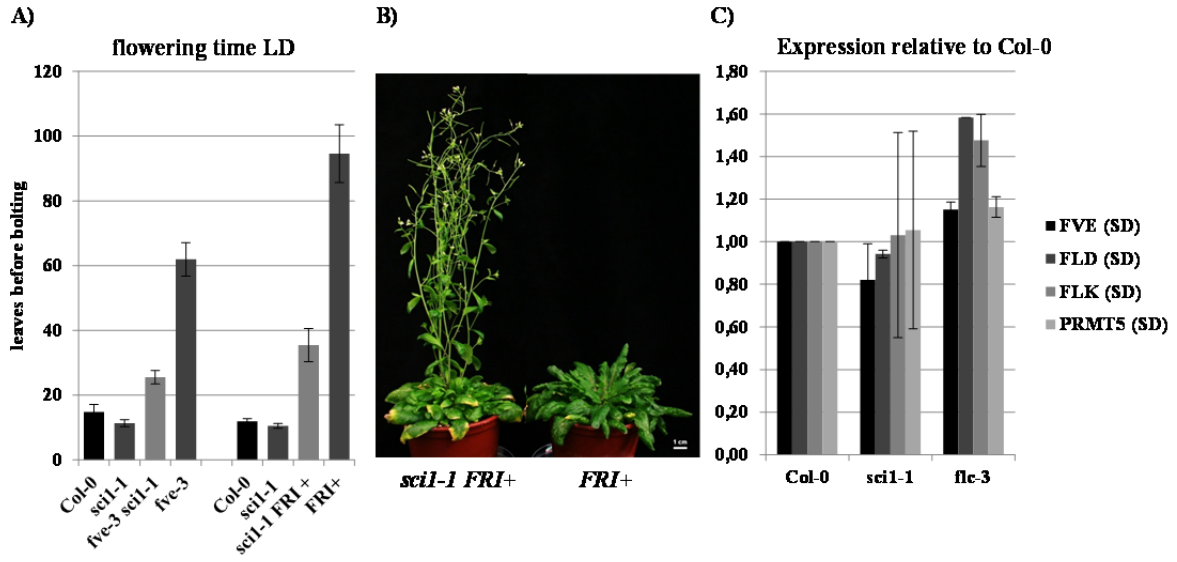


Figure 3.9: Flowering time analysis of *sci1-1 FRI+* and *sci1-1 fve-3* mutants and expression analysis of *FLC* repressors
(A) Flowering time analysis of plants grown in LD conditions, average numer of rosette leaves is depicted, error bars represent standard deviation. (B) *FRI+* and *sci1-1 FRI+* plants 73 dag grown in LD conditions. (C) qRT-PCR based expression analysis of *FVE/MSI4, FLD, FLK* and *PRMT5* in *sci1-1* and *flc-3* single single mutants compared to Col-0. cDNA was prepared from whole seedlings 10 dag grown in short day conditions. Expression values were normalized to At4g34279 (nblack) and are shown relative to Col-0. The average and standard deviation were obtained from at least two biological replicates, each consisting of three technical replicates.

The transcript level of genes involved in the repression of *FLC* was tested by qRT-PCR in *sci1-1* and *flc-3* mutants compared to Col-0 wildtype plants, in order to analyze whether SCI1 regulates *FLC* through these repressors. *FVE/MSI4* encodes for a WD40 protein involved in deacetylation and ubiquitination of the *FLC* locus, and the autonomous pathway component FLOWERING LOCUS D (FLD) is also involved in the deacetylation of *FLC* (Ausin et al., 2004; Jeon and Kim, 2011; He et al., 2003). The RNA binding protein FLOWERING LOCUS KH DOMAIN (FLK) prevents the accumulation of *FLC* mRNA, and the PROTEIN ARGININE METHYLTRANSFERASE 5 (PRMT5) is important for proper pre-mRNA splicing, that affects several RNA processing factors involved in flowering time as well as *FLK* expression, which leads to elevated *FLC* mRNA accumulation (Deng et al., 2010; Quesada et al., 2005; Lim et al., 2004). Compared to Col-0 there was no significant difference in the expression level of these genes in *sci1-1* and *flc-3* plants, suggesting that they are not involved in SCI1 dependent *FLC* regulation. In conclusion it still has to be solved by which means SCI1 is regulating *FLC* and thereby flowering time in *Arabidopsis*.

3.6.3 Primary root growth is reduced in *sci1-1* mutants

The expression analysis of SCI1 revealed a strong abundance in the primary root tip, emerging side roots and the root vasculature (Figure 3.6). These findings indicate that SCI1 is

involved in the regulation of root development, which might therefore be affected in *sci1-1* mutants. A root growth assay was carried out in LD conditions measuring the root length of vertically grown Col-0 and *sci1-1* plants every second day (Figure 3.10). At 8 dag *sci1-1* mutants roots were significantly reduced to 64,2 % of the lenth observed in wildtype (Figure 3.10 A and B). Furthermore *sci1-1* mutants show a slower increase in root length compared to Col-0 suggesting a defect either in cell proliferation or cell elongation in *sci1-1* mutants (Figure 3.10 B). In order to reveal whether the root length reduction observed in *sci1-1* mutants is caused by alterations in the root apical meristem, mpsPI (modified pseudo-Schiff propidium iodide) staining of Col-0 and *sci1-1* roots at 8 dag were performed and analyzed by confocal microscopy (data not shown) (Truernit et al., 2008). In this analysis no apparent difference of root architecture could be observed in *sci1-1* mutants, thus further analyses will be required to determine the cause of reduced root growth in *sci1-1*.

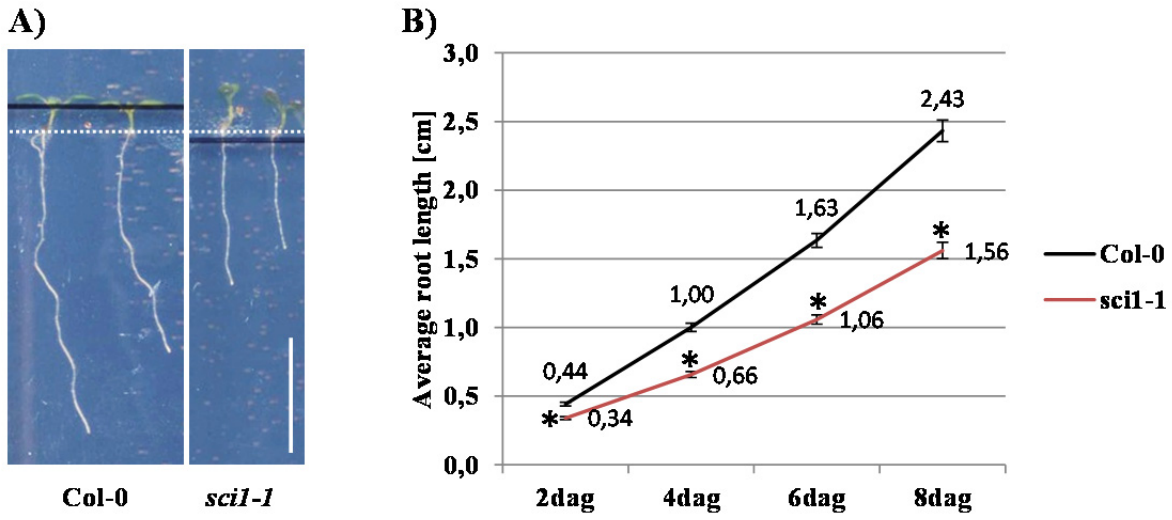


Figure 3.10: Reduction of root length in *sci1-1* mutants.
(A) Average root length of Col-0 and *sci1-1* plants grown vertically for 8 days in LD conditions. Error bars represent standard error, n= 51 Col-0 and 64 *sci1-1* plants, asterisk indicates significant difference in root length compared to Co-0 using a t-test (p ≤0,0001). (B) Example of Col-0 and *sci1-1* plants 8 dag grown in LD conditions. The white dotted line indicates root to shoot transition, the black line has no relevance. (C) Average root length of Col-0 and *sci1-1* measured every second day through 8 days of growth. Error bars represent standard error; root length was measured using Image J.

3.7 SCI1 acts redundantly with SCI2 and SCI3

sci1 mutants display a relatively mild morphological phenotype for example compared to PcG mutants like *clf-28*, *emf2-10* or the PRC2 loss of function mutant *clf-28 swn-7* (Figure 3.17). This might be due to the fact that SCI1 is encoded by a gene family with three members in *Arabidopsis*. SCI1 and the two highly homologous proteins SCI2 and SCI3 exhibit a conserved PWWP domain at the amino-terminus of the proteins and comprise a central NLS. From their similar protein architecture as well as from overlapping expression patterns

(Figure 3.6) it can be assumed, that SCI2 and SCI3 comprise redundant functions with SCI1. To investigate this hypothesis, T-DNA insertion mutants of *SCI2* and *SCI3* were obtained and analyzed (Figure 3.11 A and B). In the *sci2-2* line, the T-DNA is inserted in an intron, therefore it is likely that this line still harbors residual SCI2 function, while the insertion in *sci3-2* is in the second exon and therefore resembles more likely a loss of function allele. The weak single mutant phenotypes of *sci2-2* and *sci3-2* show a slight developmental delay or late flowering, which will have to be analyzed in more detail (Figure 3.11 B). To investigate the potential redundancy of the three *SCI* genes, double and triple mutants in all combinations were obtained and phenotypically analyzed (Figure 3.12)

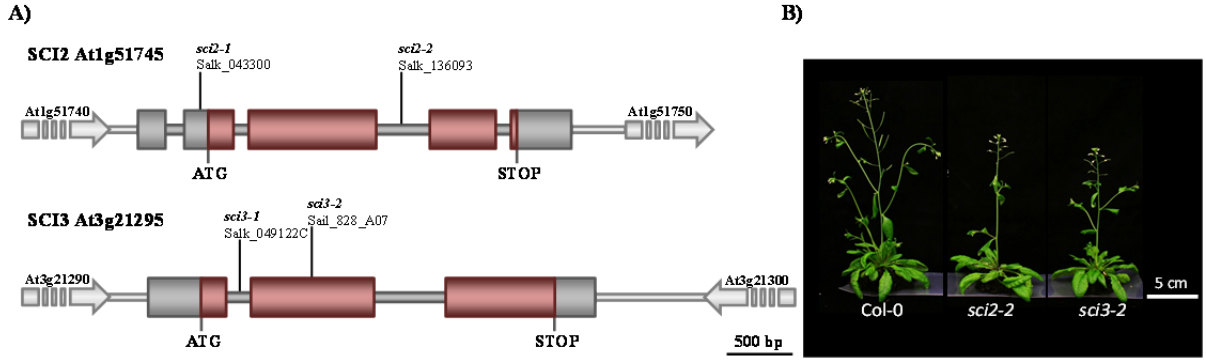


Figure 3.11: Genomic loci of *SCI2* and *SCI3* and phenotypes of T-DNA insertion mutants
(A) Genomic loci of *SCI2* (At1g51745) and *SCI3* (At3g21295) with the position and accession numbers of T-DNA insertion lines. Colored boxes resemble exons, introns as well as 5' and 3' UTR are depicted in grey. Neighboring genes (light grey) are not depicted in scale. (B) Phenotypes of the *sci2-2* and *sci3-2* mutants compared to Col-0, plants were grown in LD conditions.

3.7.1 Phenotypic analysis of *SCI* double and triple mutants

Double mutants of any *sci1*, *sci2* or *sci3* alleles show only minor morphological phenotypes like a reduction in plant size and a delay in development (Figure 3.12 A). Plants with an additional mutation of one copy of the third *SCI* gene display similar phenotypes as the double mutants. However, it cannot be excluded that the T-DNA insertion mutants comprise partial transcripts as shown for *sci1-1*, and thus potentially partial protein function (Figure 3.5). The establishment of the *sci1-1 sci3-2 sci2-2* triple mutant revealed, that loss of SCI1/2/3 function results in seedling lethality (Figure 3.12). The triple mutants show an early meristem termination, furthermore neither formation of true leaves nor root growth can be observed. Additionally, the triple mutants often show an accumulation of anthocyanin and loss of chlorophyll. The *sci1-1 sci3-2 sci2-2* triple mutants can be rescued by the introduction of SCI1pro::SCI1-GFP or SCI3pro::SCI3-GUS transgenes in the triple mutant background, suggesting that the drastic mutant phenotype is indeed due to loss of function of the three *SCI* genes (Figure 3.12).

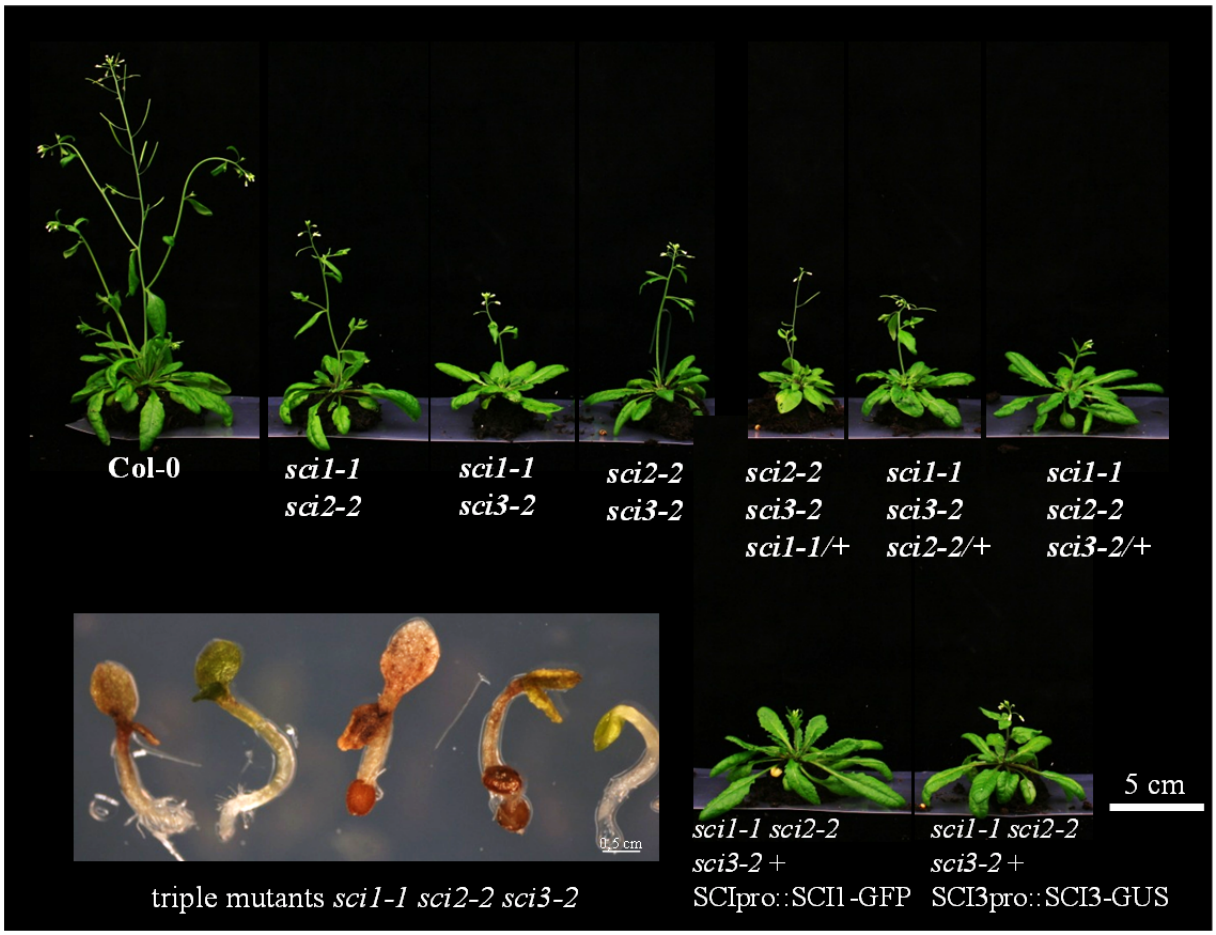


Figure 3.12: Phenotypic analysis of *SCI* double and triple mutants
Phenotypic analysis of double and triple mutant combinations of *sci1-1*, *sci2-2* and *sci3-2* alleles. Triple mutants show seedling lethality, which is rescued by introduction of the transgenes SCI1pro::SCI1-GFP or SCI3pro::SCI3-GUS. Plants were grown in LD conditions.

The analysis of the progeny of *sci1-1 sci2-2/+ sci3-2* plants revealed a strongly reduced germination rate to 67,8 %, suggesting that most of the triple mutants fail to germinate (Table 3.2 A). Further analysis of the progeny of *sci1-1 sci2-2/+ sci3-2* plants by PCR showed a distorted segregation of the *sci2-2* allele (Table 3.2 B). Triple mutants were not included in this analysis, as they germinate delayed compared to the segregating population. In the progeny, the *sci2-2* allele is predicted to segregate with a ratio of 1:2:1 (*sci2-2* : *sci2-2/+* : wildtype *SCI2*), or when the triple mutants are excluded in a ratio of 2:1 (*sci2-2/+* : wildtype *SCI2*). In the analysis 157 plants were genotyped and 89 plants were hemizygous for *sci2-2*, and 68 plants comprised a wildtype *SCI2* allele. This equals a ratio of 1,31:1 instead of the expected 2:1 ratio (sci2-2/+ : wildtype). Hence the missing *sci2-2/+* hemizygous plants either failed to germinate, or this result implicates a reduced transmission of *sci2-2* in *sci1-1 sci3-2* mutant background.

Analysis of the progeny from *sci1-1 sci3-2 sci2-2/+* plants		
(A) Germination analysis		
n	germinated	not germinated
1052	713	339
(B) Segregation analysis		
n	*sci1-1 sci3-2* ***sci2-2/+***	*sci1-1 sci3-2* ***SCI2***
157	89	68
ratio	1,3	1,0
expected ratio	2,0	1,0
(C) Reciprocal crosses		
***1) sci1-1 sci3-2 sci2-2/+* x Col-0**		
genotype of progeny	sci1-1/+ sci3-2/+ **sci2-2/+**	*sci1-1 sci3-2* ***SCI2***
n = 127	61	66
ratio	0,92	1,00
expected ratio	1,00	1,00
2)* Col-0 x *sci1-1 sci3-2 sci2-2/+		
genotype of progeny	sci1-1/+ sci3-2/+ **sci2-2/+**	*sci1-1 sci3-2* ***SCI2***
n = 157	45	112
ratio	0,40	1,00
expected ratio	1,00	1,00

Table 3.2: Analysis of the progeny from *sci1-1 sci2-2/+ sci3-2* plants
(A) The germination rate of the progeny from *sci1-1 sci2-2/+ sci3-2* plants is reduced to 67,8%. (B) Segregation analysis of the progeny of *sci1-1 sci3-2 sci2-2/+* revealed a distorted ratio of 1,31:1 (*sci2-2/+* : *SCI2*). (C) Reciprocal crosses of *sci1-1 sci3-2 sci2-2/+* with Col-0 plants showed that the transmission of the mutant *sci2-2* through the germline is reduced.

In order to analyze the transmission of the *sci2-2* allele, reciprocal crosses with Col-0 were performed (Table 3.2 C). When Col-0 pollen was used to fertilize *sci1-1 sci3-2 sci2-2/+* ovules, 61 of the resulting F1 plants were genotyped hemizygous (*sci2-2/+*) and 66 wildtype (*SCI2*). This ratio almost reflects the 1:1 ratio expected when the transmission is not affected by the mutation. On the contrary when Col-0 ovules were fertilized using pollen from *sci1-1 sci3-2 sci2-2/+* mutants, 45 hemizygous (*sci2-2/+*) and 112 wildtype (*SCI2*) plants were obtained, which equals a ratio of 0,4:1 (*sci2-2/+* : WT). This result shows that there is a severely reduced transmission of the mutant *sci2-2* allele through the germline in *sci1-1 sci3-2 sci2-2/+* background, which explains the reduced number of plants hemizygous for *sci2-2* in the segregation analysis of the progeny of *sci1-1 sci3-2 sci2-2/+* plants (Table 3.2 C).

3.7.2 *SCI* triple mutants are seedling lethal

Further analysis of the *sci1-1 sci3-2 sci2-2* triple mutants by scanning electron microscopy (SEM) analysis showed that the triple mutants exhibit severe desiccation of cells in all organs of the seedlings (Figure 3.13 C to F).

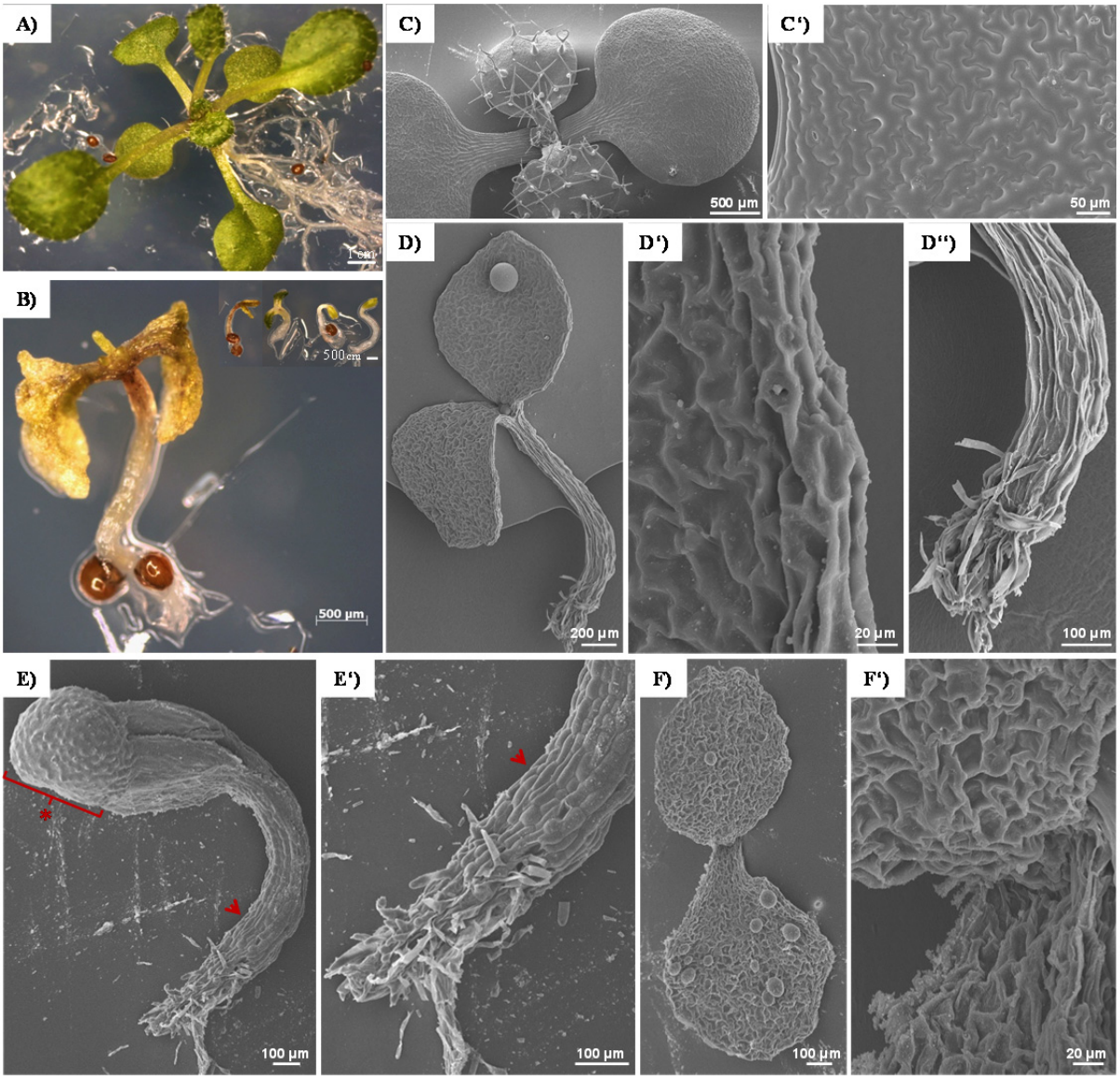


Figure 3.13: Scanning electron microscopy analysis of *sci1-1 sci2-2 sci3-2* triple mutants
(A) Segregating *sci1-1 sci3-2* double mutant compared to (B) *sci1-1 sci2-2 sci3-2* triple mutants. (C and C') SEM pictures of Col-0 wildtype seedling compared to *sci1-1 sci2-2 sci3-2* triple mutant seedlings (D-F). (D) Triple mutant seedling with desiccated cells (D'), close-up of the cotyledon and (D'') the root tip. (E) Just germinated triple mutant with wildtype-like cells in the hypocotyl (red arrowheads), close-up of the root tip in (E') shows distortion of the root apical region. Asterisk in (E) indicates the seedcoat which is still attached to the cotyledons. (F) Top view of triple mutants with desiccated cotyledons, the close-up in (F') shows the arrested shoot apical meristem.

For comparison a SEM picture of a Col-0 seedling is depicted in C, and a close-up of the cotyledon in C'. The close-up pictures D' and D'' show the cotyledon and the root tip of the triple mutant seedling depicted in D. The analysis of a just germinated triple mutant (Figure

3.13 E) revealed that the desiccation process probably starts after germination as this seedling still exhibits wildtype-like cells in the hypocotyl. Nevertheless the root apical meristem of this seedling is already distorted and starts to desiccate (Figure 3.13 E'). The arrest of the shoot apical meristem observed in the triple mutants is depicted in Figure 3.13 F and F'. This analysis clearly shows that the loss of the three *SCI* genes leads to severe problems of desiccation tolerance as well as of meristem maintenance. Since the triple mutant phenotype is too severe to be associated with a specific pathway, it will be essential in the future to study the embryonic development of the triple mutants. This analysis would help to reveal at which stage of development SCI1 function is needed first. Furthermore an inducible rescue construct would be useful to overcome the seedling lethality, and to study SCI function in adult plant organs.

3.8 Genetic interaction of *SCI1* with PRC2 component *CLF*

The phenotypic analysis of the *sci1* single and *sci1-1 sci2-2 sci3-2* triple mutants revealed that the *SCI* genes play an important role in plant development. Due to the high degree of redundancy between the three *SCI* genes, and the severity of the triple mutant phenotype, it is challenging to uncover the functional role of *SCI1* genes by phenotypic analysis. Fortunately SCI1 was found to interact with the PRC2 components CLF, SWN and MEA in yeast, suggesting an involvement in PcG mediated gene regulation (Figure 3.2). In order to study the function of SCI1 in respect to PcG mediated gene regulation, double mutants of *sci1* alleles with *clf-28* were established and analyzed.

3.8.1 Phenotypic and flowering time analysis of *sci1 clf-28* double mutants

Genetic interactions between two genes can be uncovered by the establishment of double mutants that, in case of an interaction, can show an enhancement or a suppression of the single mutant phenotypes. The *sci1-1* and *sci1-3* mutants were introduced into *clf-28* background followed by phenotypic analysis of the double mutants compared to the single mutants (Figure 3.14 A). The introduction of both *sci1* mutant alleles into *clf-28* mutant background led to a severe enhancement of the *clf* phenotype. The double mutants display stronger leaf curling and reduction of overall plant size, which is especially noticeable in short day conditions (Figure 3.14 D).

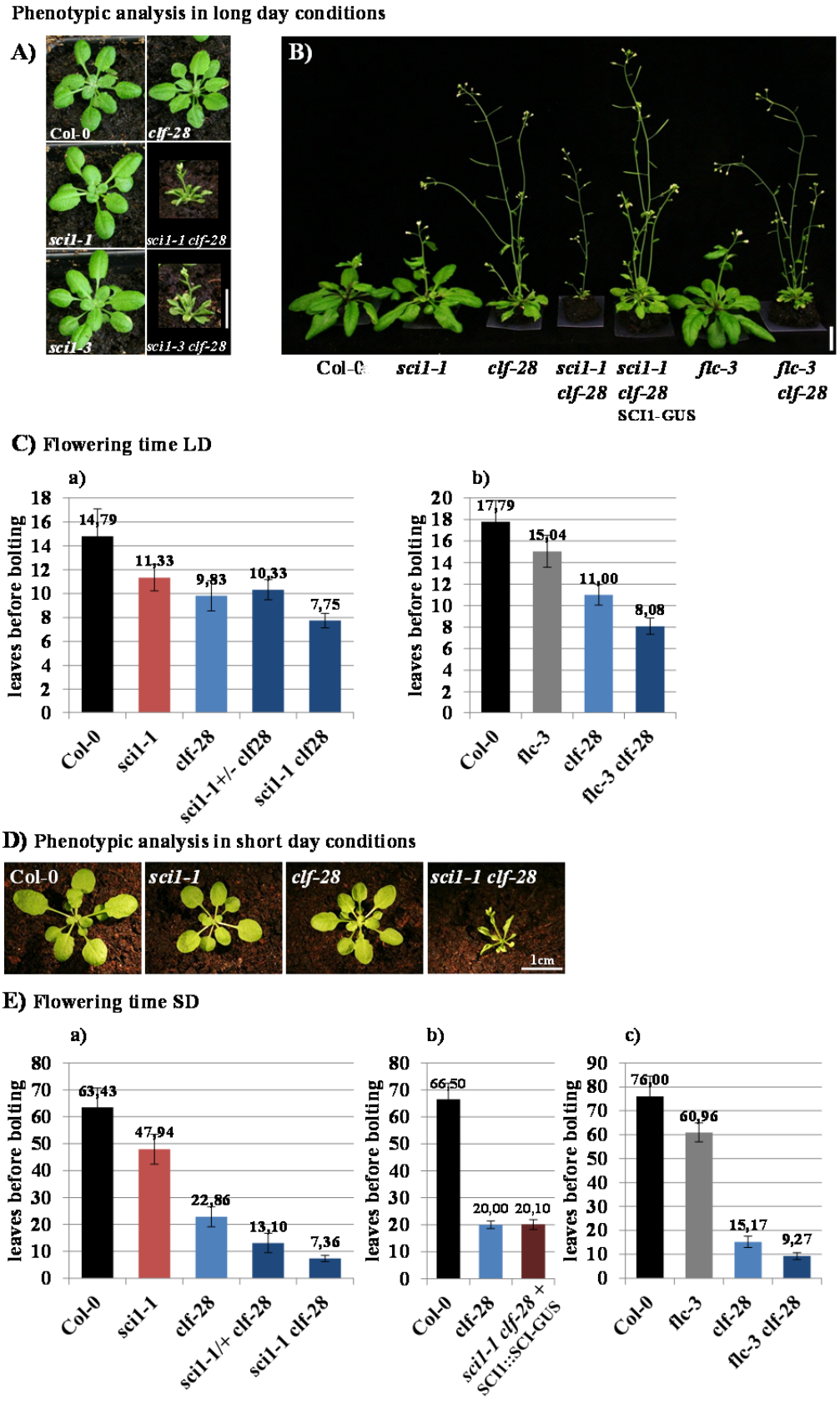


Figure 3.14: Phenotypic and flowering time analysis of *sci1 clf-28* double mutants
(A, B and D) Phenotypic analysis of indicated plants grown in long day (LD) or short day (SD) conditions, respectively, white scale bars indicate 1 cm. (C and E) Average rosette leaf number of plants grown in LD or SD conditions, error bars represent standard deviation; average was obtained from ≥15 individual plants for each genotype.

A prominent feature of the *sci1-1 clf-28* double mutant phenotype is the severe enhancement of the early flowering phenotype in LD and SD conditions (Figure 3.14 C and E). In contrast to *clf-28*, the phenotype of the double mutants is day length independent, suggesting that for example the floral promoter *FT* is no longer repressed in the double mutants grown in SD conditions. In short day conditions it can be observed that the mutation of one copy of *SCI1* already reduces the flowering time of *clf-28* to an intermediate level. This implies that SCI1 acts in a dosage dependent manner (Figure 3.14 E). The enhancement of the *clf* mutant phenotype can be complemented by introduction of a SCI1pro::SCI1-GUS transgene, which was shown phenotypically in LD conditions and by flowering time analysis in short day conditions (Figure 3.14 B and E). In respect to flowering time, the enhancement of the *clf* mutant phenotype is reminiscent of the enhancement achieved by introduction of the *flc-3* null mutant into *clf-28* background (Figure 3.14 C and E) (Lopez-Vernaza et al., 2012). Nevertheless the *sci1-1 clf-28* and *sci1-3 clf*-28 double mutants are stronger impaired in respect to leaf curling and overall plant size compared to the *flc-3 clf-28* double mutant (Figure 3.14 B). This finding suggests that other factors apart from *FLC* misregulation are involved in the *sci1-1* dependent enhancement of the *clf* phenotype. The early flowering phenotypes of PcG mutants like *clf* and *emf2* are known to be mainly caused by misexpression of the PcG target genes *AG*, *SEP3*, *FT* and *FLC* (Lopez-Vernaza et al., 2012). Therefore the transcript levels of these PcG target genes were analyzed in *sci1-1* single and double mutant backgrounds (Figure 3.15).

3.8.2 SCI1 dependent regulation of PcG target genes

The phenotypic analyses of *sci1-1 clf-28* double mutants revealed a genetic interaction of *SCI1* with *CLF,* and therefore most likely an involvement of *SCI1* in PcG mediated gene regulation (Figure 3.14). Like the PcG mutants *clf* or *emf2*, *sci1* mutant alleles show an early flowering phenotype, which is correlated with a severe downregulation of the PcG target gene *FLC* in *sci1* mutants (Figure 3.8). The early flowering phenotype of *sci1* mutants is similar to that observed in *flc-3* mutant plants. Furthermore, in double mutant analyses *SCI1* was found to act epistatically to *FLC* (Figure 3.7). The introduction of *flc-3* in *clf* mutant background leads to an enhancement of the early flowering phenotype to a similar degree as observed in *sci1-1 clf-28* mutants (Figure 3.14, Lopez-Vernaza et al., 2012). Thus the relative *FLC* transcript level of seedlings grown in LD or SD conditions was analyzed by qRT-PCR (Figure 3.15 A and D). As observed before, *FLC* expression is significantly reduced in *sci1-1* and *sci1-3* in a day length independent manner, while it is upregulated in *clf-28* background

compared to Col-0 wildtype plants (Figure 3.8; Figure 3.15 A and C). The *sci1-1 clf-28* double mutants exhibit an intermediate *FLC* expression, similar to the level observed in Col-0. The introduction of the SCI1pro::SCI1-GUS transgene in *sci1-1 clf-28* restores *FLC* expression to the level observed in *clf-28* single mutants. Thus the reduction of *FLC* expression in the double mutant is caused by loss of *SCI1* function. As shown before, the deletion mutant *flc-3* and double mutants with *sci1-1* and *clf-28* show a strong reduction of *FLC* transcript (Figure 3.8; Figure 3.15 A). The intermediate level of *FLC* expression in *sci1-1 clf-28* double mutants, that is comparable to the level observed in Col-0, is sufficient to cause the enhancement of the early flowering phenotype. In comparison, *flc-3 clf-28* double mutants display only background levels of *FLC* transcript, and FLC function is completely impaired in *flc-3* loss of function mutants.

The *clf* mutant phenotype was shown to be mainly caused by ectopic misexpression of the floral homeotic genes *AG* and *SEP3, FLC* and the floral promoter *FT* (Goodrich et al., 1997; Lopez-Vernaza et al., 2012; Jiang et al., 2008). Furthermore upregulation of several other PcG target genes was described in *clf* backgrounds, among them the MADS box transcription factors *AGAMOUS LIKE 19* (*AGL19*) and the meristem identity gene *SHOOT MERISTEMLESS* (*STM*) (Lopez-Vernaza et al., 2012; Jiang et al., 2008; Schonrock et al., 2006; Katz et al., 2004). These genes, together with the MADS box genes *AGL24* and *PISTILLATA* (*PI*), are potential candidates that could cause the enhancement of the *clf-28* phenotype in *sci1-1 clf-28* double mutants (Kim et al., 2010; Alexandre and Hennig, 2008), and were therefore analyzed by qPCR analysis (Figure 3.15).

The PcG targets *AGL19* and *AGL24* are MADS box genes that function as floral activators in response to vernalization in an FLC-independent manner (Alexandre and Hennig, 2008; Michaels et al., 2003; Schonrock et al., 2006). In qRT-PCR analysis, transcript levels of *AGL24* and *AGL19* were found to be significantly elevated in the *sci1-1 clf-28* double mutant compared to Col-0 and *clf-28,* which is restored in *sci1-1 clf-28* SCI1pro::SCI1-GUS transgenic plants (Figure 3.15 A and C). Thus loss of *SCI1* function in *clf-28* background causes the enhanced transcript levels of the floral activators. Interestingly also *flc-3 clf-28* double mutants show elevated levels of *AGL24* compared to *flc-3* single mutants, which display similar transcript levels as Col-0 (Figure 3.15 A). The four-fold enrichment of *AGL19* and *AGL24* transcript levels in *sci1-1 clf-28* compared to *clf-28* single mutants is most likely not caused by reduced *FLC* levels, as the double mutants show similar *FLC* levels as the wildtype.

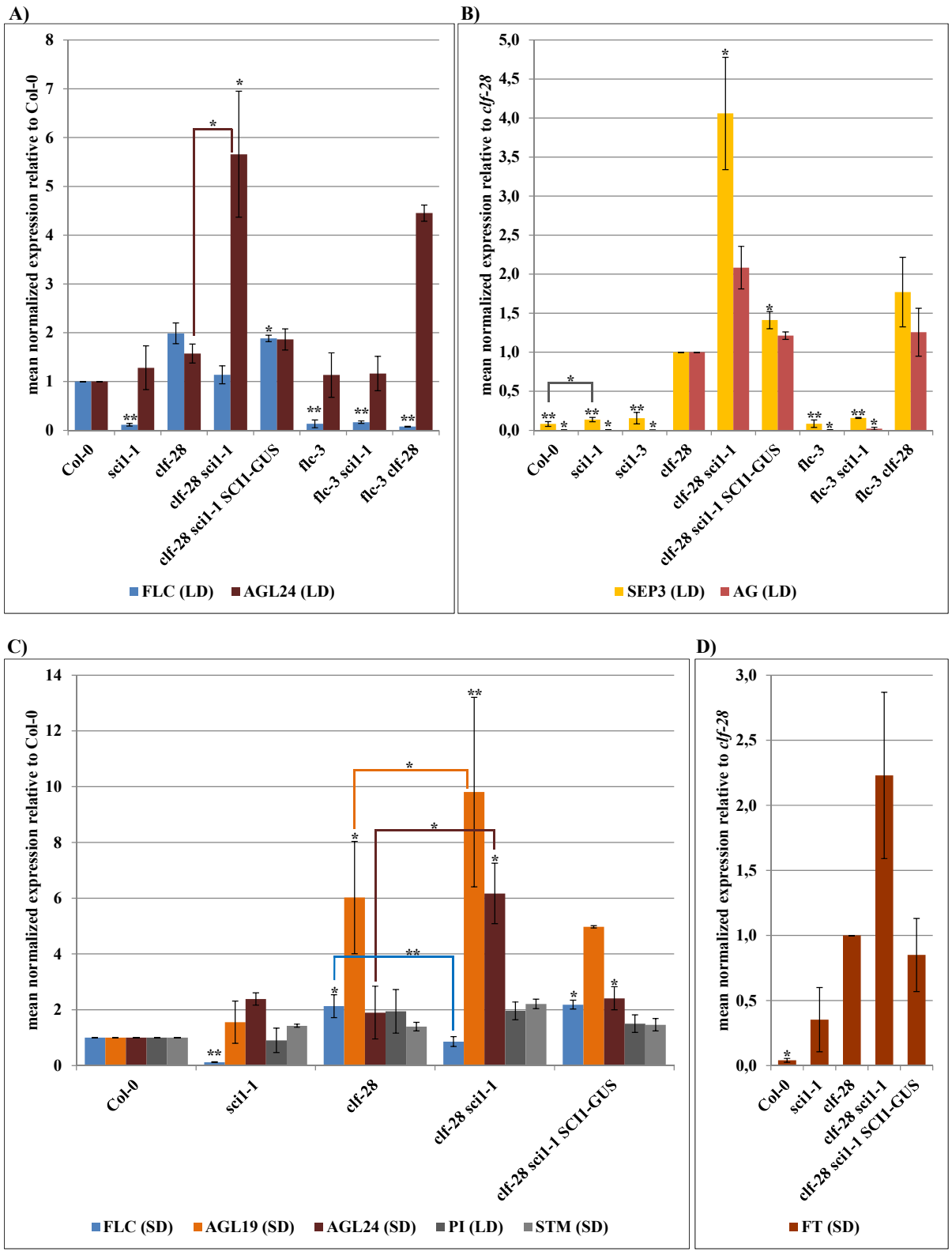


Figure 3.15: Expression analysis of PcG target genes via qRT-PCR
(A) to (D): qRT-PCR based expression analysis of *FLC, AGL19, AGL24, SEP3, AG, PI, STM* and *FT* in indicted mutants and Col-0 wildtype. RNA was extracted from whole seedlings 10 dag grown in long day (LD) or short day conditions (SD). Expression values were normalized to At4g34279 (nblack) and are shown relative to Col-0 or *clf-28,* as indicated on vertical axis. The average and standard deviation were obtained from at least two biological replicates each consisting of three technical replicates. Asterisks indicate level of significance as obtained by student ttest, *p≤0,05; **p≤0.01. Significance is shown in respect to either to Col-0 or *clf-28* (as indicated on y-axis), additional significance values are shown by asterisks above a bracket that connects the compared genotypes.

An involvement of *FLC* expression on *AGL19* and *AGL24* transcript abundance in *clf-28* background cannot be excluded and would need further investigations. While it has been shown before that *AGL19* and *AGL24* act in an *FLC* independent manner (Schonrock et al., 2006; Michaels et al., 2003), the effect of *FLC* on the transcriptional activation of these floral activators might be masked by *CLF* and thus PcG mediated gene silencing.

The analysis of *PI* and *STM* transcript levels revealed no significant difference in the tested mutant backgrounds compared to Col-0, suggesting that these genes are not involved in the *sci1-1* induced enhancement of the *clf-28* phenotype (Figure 3.15 C).

Transcript abundance of *SEP3*, *AG* and *FT* was analyzed in relation to the levels observed in *clf-28* mutants, since these genes are hardly expressed in Col-0 wildtype seedlings (Figure 3.15 B and D). Compared to Col-0, transcripts of *SEP3*, *AG* and *FT* were found to be 14-fold, 540-fold and 30-fold enriched in *clf-28* single mutant seedlings, respectively. The transcript level of *SEP3* in *sci1-1 clf-28* double mutants is four-fold increased compared to *clf-28.* In contrast, a similar misexpression is not observed in *flc-3 clf-28* double mutants. Thus *SEP3* misexpression is a good candidate to cause the enhancement of the *clf-28* phenotype observed in *sci1-1 clf-28* plants. Further evidence that *SEP3* might be responsible for the enhanced *clf-28* phenotype in *sci1-1 clf-28* is provided by the finding that *sep3* is able to suppress the *clf* phenotype (Lopez-Vernaza et al., 2012). Also transcript levels of *AG* and *FT* are enhanced in the double mutant, and in case of *AG* this is not observed in *flc-3 clf-28* mutants. Thus *AG* and *FT* misexpression may also play a role in causing the *sci1-1 clf-28* phenotype, even though in case of *FT* the elevated transcript levels might be secondary due to a *sci1-1* induced reduction of *FLC* in the double mutant compared to the *clf-28* single mutant (Figure 3.15 B and D).

In conclusion, beside the reduced *FLC* expression in *sci1-1 clf-28* compared to *clf-28*, an alteration in transcript abundance of *AGL19*, *AGL24*, *SEP3*, *AG* and *FT* might be involved in the enhancement of the double mutant phenotype in respect to both, flowering time and plant morphology. While *AGL19*, *AGL24*, and *FT* expression levels might be dependent on FLC activity, *SEP3* and *AG* are good candidates to cause the enhanced double mutant phenotype in concert with the effect of reduced *FLC* expression levels. Apart from regulation of *FLC*, *SCI1* appears to act as a repressor of the analyzed PcG target genes, at least in *clf-28* mutant background.

3.8.3 ChIP analysis of histone methylation levels in *sci1-1* mutants

The PRC2 complex mediates gene silencing by tri-methylation H3K27, and the *Arabidopsis* histone methyltransferase mutants *clf* and *clf swn* show loss of H3K27me3 levels at PcG target genes, which coincides with transcriptional de-repression (Schubert et al., 2006). *sci1* mutants show repression of *FLC,* and de-repression of several other PcG target genes (Figure 3.15). Thus levels of H3K27me3 and of the antagonistically acting H3K4me3 mark were analyzed by Chromatin Immunoprecipitation (ChIP) in *sci1-1* mutant seedlings compared to Col-0 (Figure 3.16).

In the ChIP experiments, the histone modification levels of *ACTIN*, *FLC*, *SEP3*, *AG* and *FUSCA3* (*FUS3*) were analyzed, the used oligonucleotides matched regions of the genes that were shown to be enriched for the respective modification. In the diagrams A) and D) the precipitation efficiency compared to the input (%Input) is shown for the H3K27me3 and H3K4me3 experiments, respectively. *ACTIN* is not a PcG target gene, but exhibits H3K4me3 and is ubiquitously and strongly expressed *Arabidopsis* (Zhang et al., 2007a; Oh et al., 2008; Lafos et al., 2011). Consistent with this observation, H3K4me3 levels were found to be enriched at the *ACTIN* locus (D) while no enrichment for H3K27me3 (A) could be detected in Col-0 and *sci1-1* seedlings. *FUS3* is seed specifically expressed PcG target gene that carries H3K27me3 in seedlings, which was also shown in this experiment (A) (Makarevich et al., 2006). To account for differences in precipitation efficiency between Col-0 and *sci1-1*, the H3K27me3 %Input values were normalized to those of *FUS3*, while the H3K4me3 %Input values were normalized to *ACTIN*. In the resulting diagrams (B and D) it can be observed that Col-0 and *sci1-1* seedlings show similar enrichment of H3K27me3 and H3K4me3 at the tested target genes. While *ACTIN* served as negative control in the H3K27me3 ChIP experiment, the intronic region of *AG,* which is enriched for H3K27me3 (B), served as negative control for the H3K4me3 experiment. Congruent with the expectation, *AG* showed only low levels of enrichment for the H3K4me3 mark (D) compared to *ACTIN*.

Thus it could be shown that *FLC*, *SEP3* and *AG* show enrichment of H3K27me3 in Col-0 and *sci1-1* seedlings. Furthermore *FLC* and *SEP3* exhibit H3K4me3 at regions close to the start codon, again to a similar extent in Col-0 and *sci1-1*. This finding implies that the altered expression levels of PcG targets in *sci1-1* are not due to changes in H3K4me3 or H3K27me3 levels. Nevertheless it cannot be excluded that the levels of these histone modifications are altered at different regions of the target loci, or that there are differences in specific cell types. More detailed ChIP analyses, especially of the *FLC* locus, will be essential to judge whether

altered levels of histone modifications are involved in the transcriptional misregulation of PcG target genes observed in *sci1-1* and *sci1-1 clf-28* mutants (Figure 3.15).

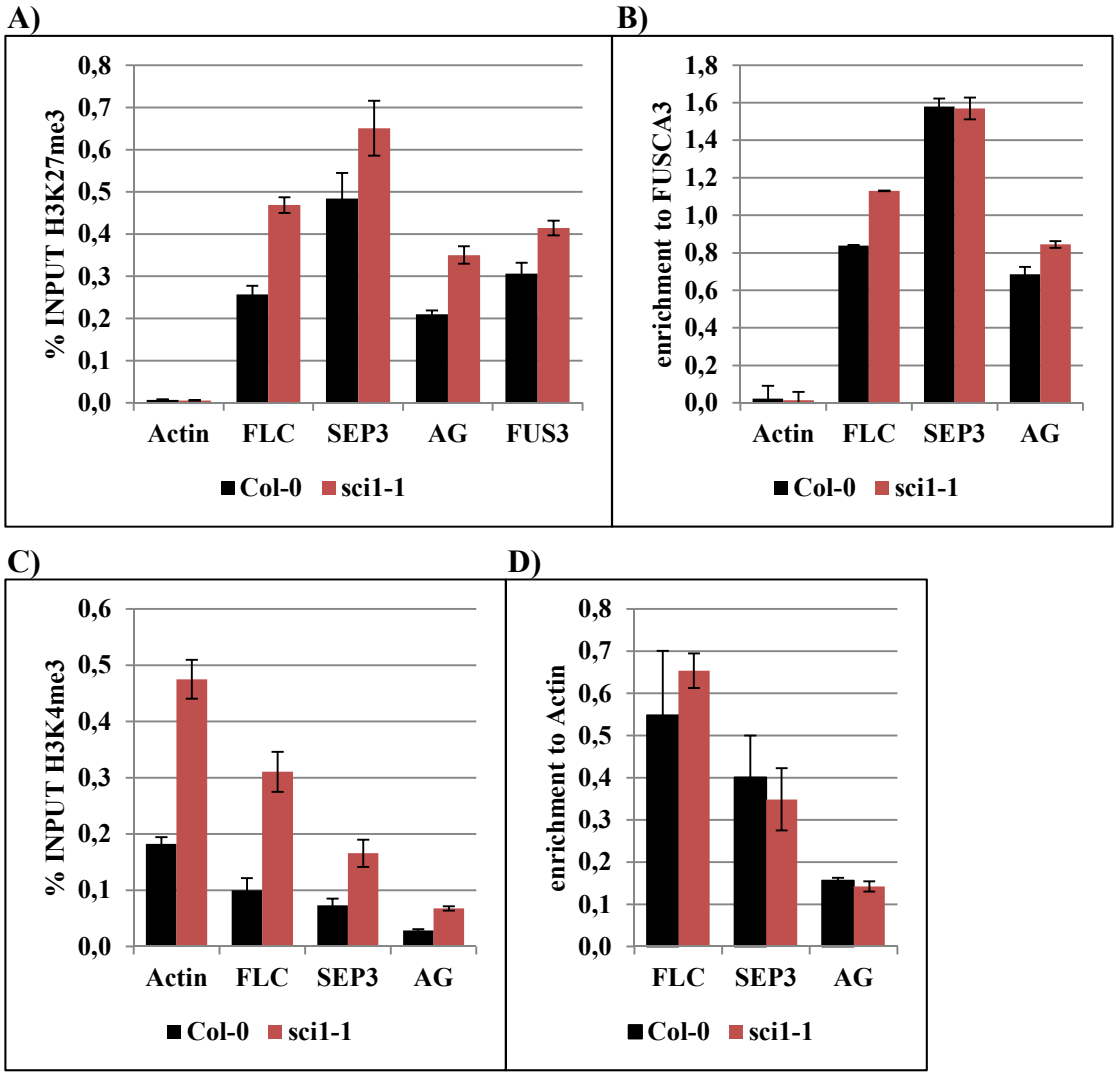


Figure 3.16: ChIP analysis of PcG target genes in *sci1-1* mutants
Chromatin Immunoprecipitation (ChIP) experiments were performed using antibodies against H3K27me3 (A and B) and H3K4me3 (C and D) on nuclear extracts from 10 day old *sci1-1* and Col-0 seedlings. Samples were analyzed by qRT-PCR using oligonucleotides matching intronic regions in case of H3K27me3, or regions close to the start codon in case of H3K4me3. The precipitation efficiency (A and C) was normalized to levels observed for FUS3 (H3K27me3) or Actin (H3K4me3) (Table 2.4). Results were obtained in a single experiment; error bars represent standard deviation of the technical qRT-PCR replicates.

3.8.4 Global levels of histone modifications in *sci1-1 clf-28* and other PcG mutants

In order to analyze whether the mutation of *SCI1* in *clf-28* mutant background causes a global change in histone modification levels, histone enriched protein extracts were examined by immunoblot analyses. Furthermore several histone modifications in PcG mutants with different levels of PcG activity were observed (Figure 3.17), and this data was recently

published (Lafos et al., 2011). The mutants analyzed were *clf-28*, *clf-28 swn-7*, *sci1-1 clf-28* and *bli* in Col-0 ecotype as well as *emf2-10* and *vrn2-1 emf2-10* in the WS ecotype (Figure 3.17 A).

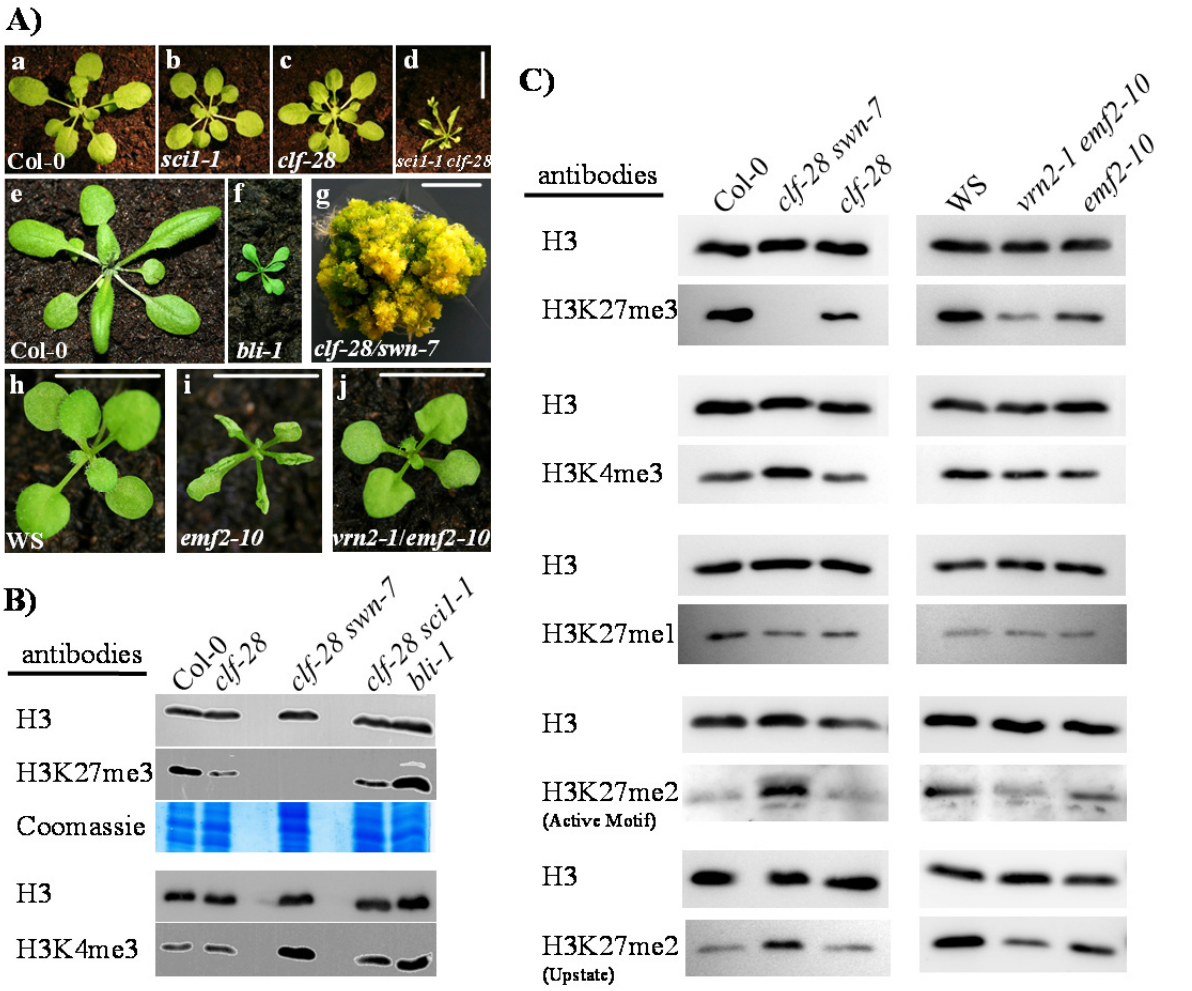


Figure 3.17: Global levels of histone modifications in PcG mutants and *sci1-1 clf-28*
(A) Morphological phenotypes of PcG mutants compared to the respective wildtype. (a-d) Plants grown in SD conditions, scale bar 1 cm; (e-f) LD conditions, (g) LD conditions, 60 dag, scale bar 0,5 cm; (h-j) short day conditions, scale bar 0,5 cm. (B) Immunoblot analysis on histone protein extract from *clf-28, clf-28 swn-7, sci1-1 clf-28* and *bli-1* compared to Col-0. Equal amounts of histones were loaded as indicated by the coomassie staining and immunoblot with H3 antibody. Immunoblot analysis of histone modifications H3K27me3 and H3K4me3 are shown beneath the respective loading control. (C) Immunoblot analysis with protein extracts from the PcG mutants *clf-28, clf-28 swn-7, emf2-10 and vrn2-1 emf2-10* compared to their respective wildtype Col-0 or WS. The histone modifications H3K27me3, H3K4me3, H3K27me1 and H3K27me2 are analyzed, the corresponding immunoblot with H3 antibody detecting the C-terminal unmodified part of H3 serves as loading control.

The *clf-28 swn-7* double mutants are severely affected in cell fate control, and grow in callus like structures (Figure 3.17 A) (Chanvivattana et al., 2004). In this global analysis, H3K27me3 was not detectable in *clf-28 swn-7* double mutants, while *clf-28* single and *sci1-1 clf-28* double mutants exhibit reduced H3K27me3 levels compared to Col-0. In contrast *bli-1*

mutants display similar levels as the wildtype (Figure 3.17 B and C). This result implies that vegetative H3K27me3 is completely dependent on the PcG proteins CLF and SWN which act in part redundantly (Chanvivattana et al., 2004). Since SWN is not able to completely fulfill CLF function, *cfl-28* single and *sci1-1 clf-28* mutants show a reduced level of H3K27me3. The mutation of *SCI1* in *clf-28* background does not result in a further loss of H3K27me3 suggesting that SCI1 acts either downstream of PRC2 or in a cell type or target gene specific manner. Furthermore SCI1 function might be masked by redundancy with the homologs SCI2 and SCI3.

The weak *emf2-10* allele shows a comparable phenotype to *clf-28* mutants, and also a similar reduction of H3K27me3 levels (Figure 3.17). H3K27me3 is further reduced in the *vrn2-1 emf2-10* double mutants, and residual levels of the modification are probably due to partial PRC2 activity caused by the weak *emf2-10* allele.

Furthermore the global levels of H3K4me3, H3K27me1 and H3K27me2 were determined, as it was not known how these modifications are affected in the different PcG mutant backgrounds. The activating mark H3K4me3 is mediated by TrxG proteins, which function as antagonists of the PcG proteins (Klymenko and Muller, 2004). While global H3K4me3 levels are not affected in *clf-28*, *sci1-1 clf-28* and *bli-1*, a strong enhancement can be observed in *clf-28 swn-7* double mutants compared to Col-0 (Figure 3.17 B and C). This enhancement seems to correlate with a complete loss of PRC2 activity, as it is not observed in the weaker *emf2-10* and *emf2-10 vrn2-1* mutants (Figure 3.17 C).

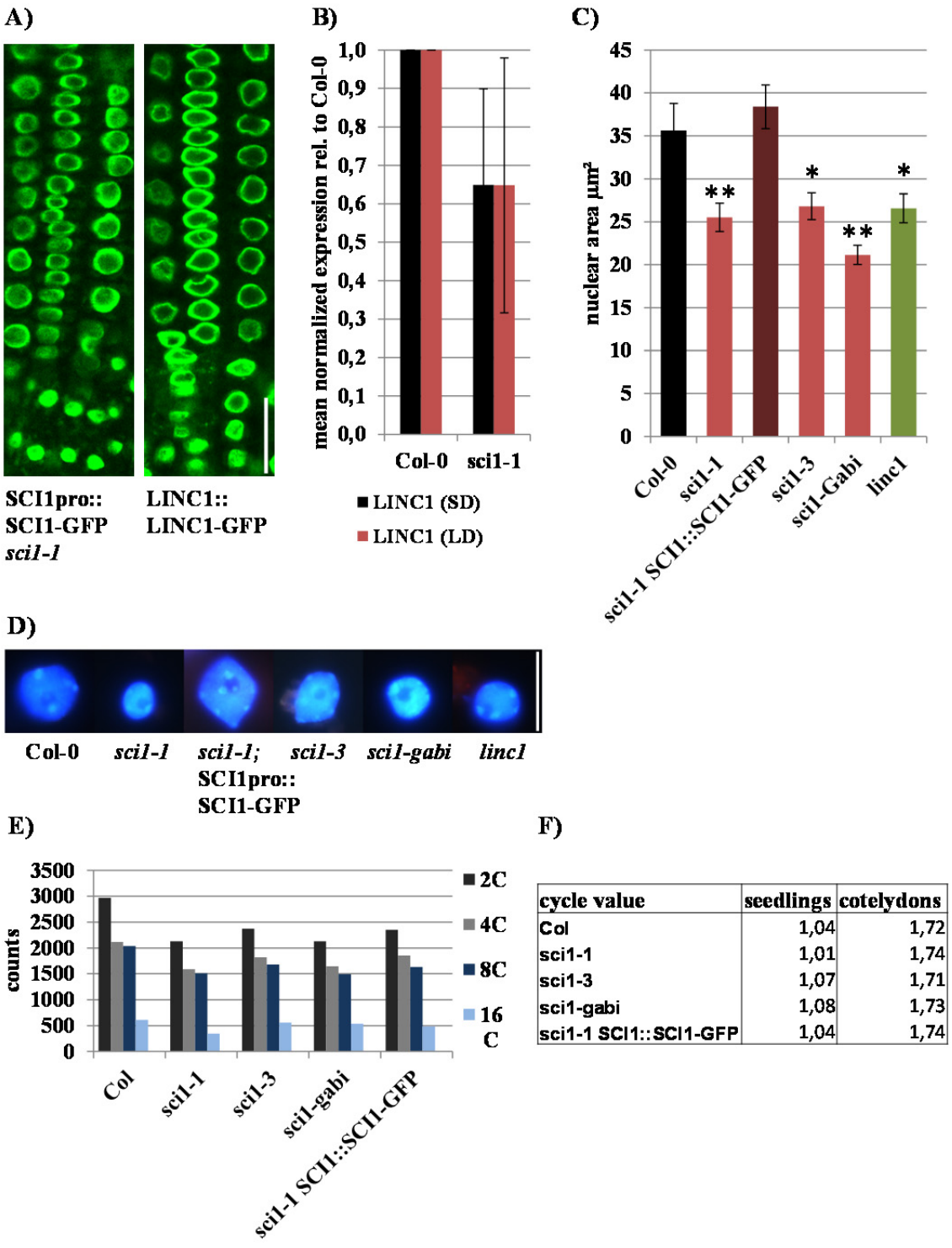
Since tri-methylation of H3K27 seems to be exclusively dependent on PRC2 activity, we investigated whether the H3K27me1 and H3K27me2 marks were changed in PcG mutant background. H3K27me1 is preferentially found in heterochromatin (Jacob et al., 2009), and global levels are not affected in the PcG mutants analyzed consistent with a role for PcG proteins in euchromatic gene regulation (Figure 3.17 C). On the other hand global levels of the euchromatic H3K27me2 mark are severely increased in *clf-28 swn-7* double mutants, while reduced in *vrn2-1 emf2-10* mutant background compared to the respective wildtype (Figure 3.17 C). In conclusion, H3K27me2 is likely not directly mediated by PRC2 activity, but is nevertheless affected in the two double mutant combinations.

These results illustrate that H3K27me3 is completely dependent on PRC2 activity, and that complete or partial loss of PRC2 function results in alterations of other histone modifications like the TrxG mediated H3K4me3 or H3K27me2 levels.

3.9 SCI1 regulates nuclear size similar to LINC1

Additional to its participation in PcG target gene and flowering time regulation, SCI1 is most likely involved in other pathways regulating plant development, which is implied by the severe *sci1-1 sci2-2 sci3-2* triple mutant phenotype (Figure 3.13). The complex isolation experiment yielded a list of promising candidate proteins which may act together with SCI1 (chapter 3.3.3). One of these putative interaction partners of SCI1, the coiled coil protein LITTLE NUCLEI1 (LINC1), shows a similar intracellular localization as SCI1-GFP in *Arabidopsis* roots (Figure 3.18 A). LINC1 is an ortholog of the Nuclear Matrix Constituent Protein 1 (NMCP1) from *Daucus carota,* and both proteins localize to the nuclear periphery (Masuda et al., 1997; Dittmer and Richards, 2008). Together with its homolog LINC2, LINC1 functions in the regulation and dynamics of nuclear size, which is reduced in *linc1* and *linc2* mutants (Dittmer et al., 2007; vanZanten M. et al., 2011). The analysis of *LINC1* transcript levels in *sci1-1* mutants (Figure 3.18 B) compared to Col-0 revealed a slight reduction of *LINC1* transcript in *sci1-1* mutant background.

The potential protein interaction, similar intracellular localization patterns of SCI1-GFP and LINC1-GFP fusion proteins and reduced transcript levels of *LINC1* in *sci1-1* imply, that SCI1 might be involved in nuclear size regulation. Isolated nuclei from whole seedlings 14 dag were stained with DAPI and microscopically analyzed (Figure 3.18 D). Area measurements of these nuclei revealed that three independent *sci1* alleles, similar to *linc1* mutants, show a significant reduction of nuclear size compared to the wildtype. This phenotype is complemented by introduction of the SCI1::SCI1-GFP transgene (Figure 3.18 C). In *linc1* mutants it was found that the size reduction is not caused by altered endoploidy levels (Dittmer et al., 2007). To analyze the level of endoreduplication in *sci1* mutants, propidium iodide stained extracts from whole seedlings and from differentiated cotyledons were analyzed by Fluorescence Activated Cell Sorting (FACS). In this analysis the ploidy levels of *sci1-1, sci1-3, sci1-gabi* and the complementing line *sci1-1* SCI1pro::SCI1-GFP are in all cases proportional to those detected in wildtype, which is shown in a histogram in Figure 3.18 E. Furthermore the cycle value, which represents the mean number of endoreduplication cycles per nucleus, was determined (Figure 3.18 F) (Barow and Meister, 2003). The cycle value is similar to Col-0 for whole seedlings and cotyledons in all genotypes tested, therefore it can be stated that the reduction of nuclear size in *sci1* mutants is not due to an alteration in endoploidy levels. Concluding SCI1 appears to be involved in nuclear size regulation, possibly in concert with LINC1, and it will be interesting to learn whether *SCI1* genetically interacts with *LINC1*.



cycle value	seedlings	cotelydons
Col	1,04	1,72
sci1-1	1,01	1,74
sci1-3	1,07	1,71
sci1-gabi	1,08	1,73
sci1-1 SCI1::SCI1-GFP	1,04	1,74

Figure 3.18: SCI1 and its interaction partner LINC1 regulate nuclear size
(A) Confocal microscopic analysis of the intracellular localization of the SCI1-GFP and LINC1-GFP fusion proteins in *Arabidopsis thaliana* root cells. Scale bar represents 20 µm. (B) qRT-PCR based expression analysis of *LINC1* in Col-0 and *sci1-1*. RNA was extracted from whole seedlings 10 dag grown in long day (LD) or short day (SD) conditions. Expression values were normalized to At4g34279 (nblack) and are shown relative to Col-0. Error bars represent standard deviation of at least two biological replicates each consisting of three technical replicates. (C) Average nuclear area (μm^2) from formaldehyde fixed and DAPI stained nuclei extracted from whole seedlings 14 dag. Error bars represent standard error (SE), at least 50 nuclei were measured per genotype. Asterisks indicate level of significance compared to Col-0 analyzed with a student ttest, *≡p≤0,05; **≡p≤0,01. (D) Representative DAPI stained nuclei for each genotype; scale bar represents 10µm. (E) Histogram of detected ploidy levels in seedlings of the indicated genotypes analyzed by FACS. (F) Cycle value which represents the mean number of endoreduplication cycles per nucleus of the analyzed genotypes in seedlings and cotyledons.

3.10 Functional analyses of the SCI1 PWWP-domain

In silico analysis of the SCI1 protein sequence predicted an amino terminal PWWP domain and a central NLS (Figure 3.1). PWWP domains have been described to confer nucleic acid binding and histone binding abilities, for example in the mammalian de novo DNA methyltransferase DNMT3A (*Mus musculus*) or the Pdp1 protein from *Schizosaccharomyces pombe* (Ge et al., 2004; Wang et al., 2009)). An alignment of the predicted SCI1 PWWP domain with those of the *Arabidopsis thaliana* proteins SCI2, SCI3 and ATX1, the mammalian DNMT3A, DNMT3B and MSH6 (*Mus musculus*), the human NSD2 and the yeast protein Pdp1 (*S. pombe*) is shown in Figure 3.19.

```
SCI1_ARATH (Q9M9N3) 1    ---MA---SPGSGAVDWTVGSIVWVRRRNGSWTPGRILGQEDLDSTHITSPRSGTPVKLLGREDA  59
SCI2_ARATH (Y1745)  1    MESND---DRNLEAINASVGRLVWVRRRNGSWPGQTLVHDQVPDNSLVGPKVGTPIKLLGRDDV  62
SCI3_ARATH (Q9LU29) 1    MGSSD---ERNCKAIDASVGGLVWVRRRNGAWPGRIMAHHEVPDGTIVSPKSGTPIKLLGRDDA  62
ATX1_ARATH (Q9C5X4) 283  YDEMVVLAATLDECQDFEPGDIVWAKLAGHAMWPAVIVDESII------GERKGLNNKVSG---- 337
DNMT3A_MUS (O88508) 270  ATKAADDEPEYEDGRGFGIGELVWGKLRGFSWWPGRIVSWWMT----------GRSRAAEG---- 320
DNMT3B_MUS (O88509) 214  AESRDGDSTEYQDDKEFGIGDLVWGKIKGFSWWPAMVVSWKAT----------SKRQAMPG---- 264
MSH6_MUS   (P54276) 76   ASSSAQA--VPPSSCDFSPGDLVWAKMEGYPWWPCLVYNHPFD----------GTFIRKKGK--- 125
NSD2_HUMAN (O96028) 204  KESCPNTGRDKDHLLKYNVGDLVWSKVSGYPWWPCMVSADPLL----------HSYTKLKGQKKS 258
PDP1_SCHPO (O59676) 34   VEESKDNLEQASADDRLNFGDRILVKAPGYPWWPALLLRRKET----------KDSLNTNSS--F  86
                                          *  :  :      **                    .

SCI1_ARATH (Q9M9N3) 60   S----VDWYNLE------KSKRVKPFRCGD--FDECIERVESSQAMIIKKREKYARREDAILHAL 112
SCI2_ARATH (Y1745)  63   S----VDWYILE------NSKTVKAFRCGE--YDTCIEKAKASSS--KKRSGKCTLREDAINNAL 113
SCI3_ARATH (Q9LU29) 63   S----VDWYNLE------KSKRVKAFRCGE--YDACIATAKATASTTGKKAVKYARREDAIAHAL 114
ATX1_ARATH (Q9C5X4) 338  GGSLLVQFFGTH-DFARIKVKQAISFIKGLLSP----SHL----------KCKQPRFEEGMQEAK 387
DNMT3A_MUS (O88508) 321  --TRWVMWFGDG-KFSVVCVEKLMPLSSFCSAFHQATYN-------------KQPMYRKAIYEVL 369
DNMT3B_MUS (O88509) 265  --MRWVQWFGDG-KFSEISADKLVALGLFSQHFNLATFN-------------KLVSYRKAMYHTL 313
MSH6_MUS   (P54276) 126  SVRVHVQFFDDSPTRGWVSKRMLKPYTGSKSKEAQKGGH--------------FYSSKSEILRAM 176
NSD2_HUMAN (O96028) 259  ARQYHVQFFGDAPERAWIFEKSLVAFEGEGQFEKLCQESA------------KQAPTKAEKIKLL 311
PDP1_SCHPO (O59676) 87   NVLYKVLFFPDF-NFAWVKRNSVKPLLDSEIAKFLGSSK-------------R---KSKELIEAY 135
                              * ::                                          .

SCI1_ARATH (Q9M9N3) 113  ELEKEMLKREGKLVPEKARDDSLDATKER- 141
SCI2_ARATH (Y1745)  114  KIENEHLAKEDDNLCNLSGEEDSKRCLSGK 143
SCI3_ARATH (Q9LU29) 115  EIENAHLAKDHPPCIEKAS-TSGE--VSRK 142
ATX1_ARATH (Q9C5X4) 388  MYLKAH------------------RLPERM 399
DNMT3A_MUS (O88508) 370  QV---------------ASSRAGKL----- 394
DNMT3B_MUS (O88509) 314  EK---------------ARVRAGKT----- 323
MSH6_MUS   (P54276) 177  QR---------------ADEALSKDTAER- 190
NSD2_HUMAN (O96028) 312  KPISGKLRAQWEMGIVQAEEAASMSVEERK 341
PDP1_SCHPO (O59676) 136  E--------------------ASKT----- 140
```

Figure 3.19: Alignment of SCI1 PWWP domain with analyzed domains from mammalian and *S. pombe* proteins
Alignment of PWWP domains from following proteins: *Arabidopsis* proteins SCI1, SCI2, SCI3 and ATX1; DNMT3A, DNMT3B and MSH6 (*Mus musculus*); NSD2 (human) and Pdp1 (*S. pombe)*. Uni-prot indentifiers are shown in brackets. The predicted PWWP domains are shaded in grey, identical amino acids in black (*) and highly similar amino acids in dark grey (:). Conserved amino acids shaded in pink were used for site directed mutagenesis experiments of the SCI1 PWWP domain in analogy to experiments with Pdp1 (*S. pombe).*

The predicted PWWP domains are indicated by grey shading, identical amino acids are shaded in black and indicated by asterisks while highly similar amino acids are shaded in dark grey and indicated by dots beneath the alignment. Functional analyses of the Pdp1 PWWP domain showed, that the amino acids shaded in pink are essential for the binding specificity to trimethylated H4K20 peptides (Wang et al., 2009). Therefore the according amino acids in

SCI1 were chosen for site directed mutagenesis to functionally analyze the SCI1 PWWP domain by complementation studies in the *Arabidopsis* and for *in vitro* histone binding assays (Figure 3.20).

In order to functionally analyze the SCI1 PWWP domain, it was fused to GST (Glutathione-S-transferase) and expressed in *E.coli*. To analyze the potential histone binding capacity of the SCI1 PWWP domain, the GST-SCI1-PWWP and the mutated GST-SCI1-PWWP-W27A fusion proteins were purified and used to perform *in vitro* immunoprecipitation experiments with modified histone peptides.

3.10.1 The SCI1 PWWP domain exhibits in vitro histone peptide binding capacity

The GST-SCI1-PWWP fusion protein expressed and purified from *E.coli* cells was hybridized to the MODified^{TM} histone binding array (Active Motif) harboring many different combinations of naturally occurring histone modifications on histone peptides (Figure 3.20 A; collaboration with A. Jeltsch) (Dhayalan et al., 2010). The GST-SCI1-PWWP fusion protein was found to bind to histone peptides, but binding is specifically inhibited by phosphorylation of H3S28 (Figure 3.20 A; red boxes). This is probably not caused by the negative charge of the phosphorylation, since histone peptides phosphorylated at H3S10 were bound by the GST-SCI1-PWWP fusion protein, the probes on the array are indicated by grey circles. Also the amino acid sequence surrounding the positions H3S10 and H3S28 is similar (ARKS), indicating that the restriction from binding to H3S28pho is a specific feature of the SCI1 PWWP domain.

The binding capacity of GST-SCI1-PWWP to unmodified and modified histone H3 peptides was independently demonstrated *in vitro*. The purified GST fusion protein was incubated with biotinylated histone peptides harboring different modifications, followed by precipitation of proteins bound to the histone peptides via streptavidin covered magnetic beads. Following immunoblot analysis with an anti-GST antibody revealed binding of the GST-SCI1-PWWP fusion protein to unmodified H3 as well as to H3K27me1, H3K27me2, H3K27me3, H3K4me3 and H3K9me3 (Figure 3.20 B). While GST alone was not able to bind to the biotinylated peptides (neg. control), the GST fusion protein with the CHROMO domain of TERMINAL FLOWER2 (TFL2/LHP1) (GST-TFL2-CHROMO) which served as positive control, showed *in vitro* binding ability to H3 peptides trimethylated at K9 and K27 consistent with previously published data (Figure 3.20 B) (Turck et al., 2007).

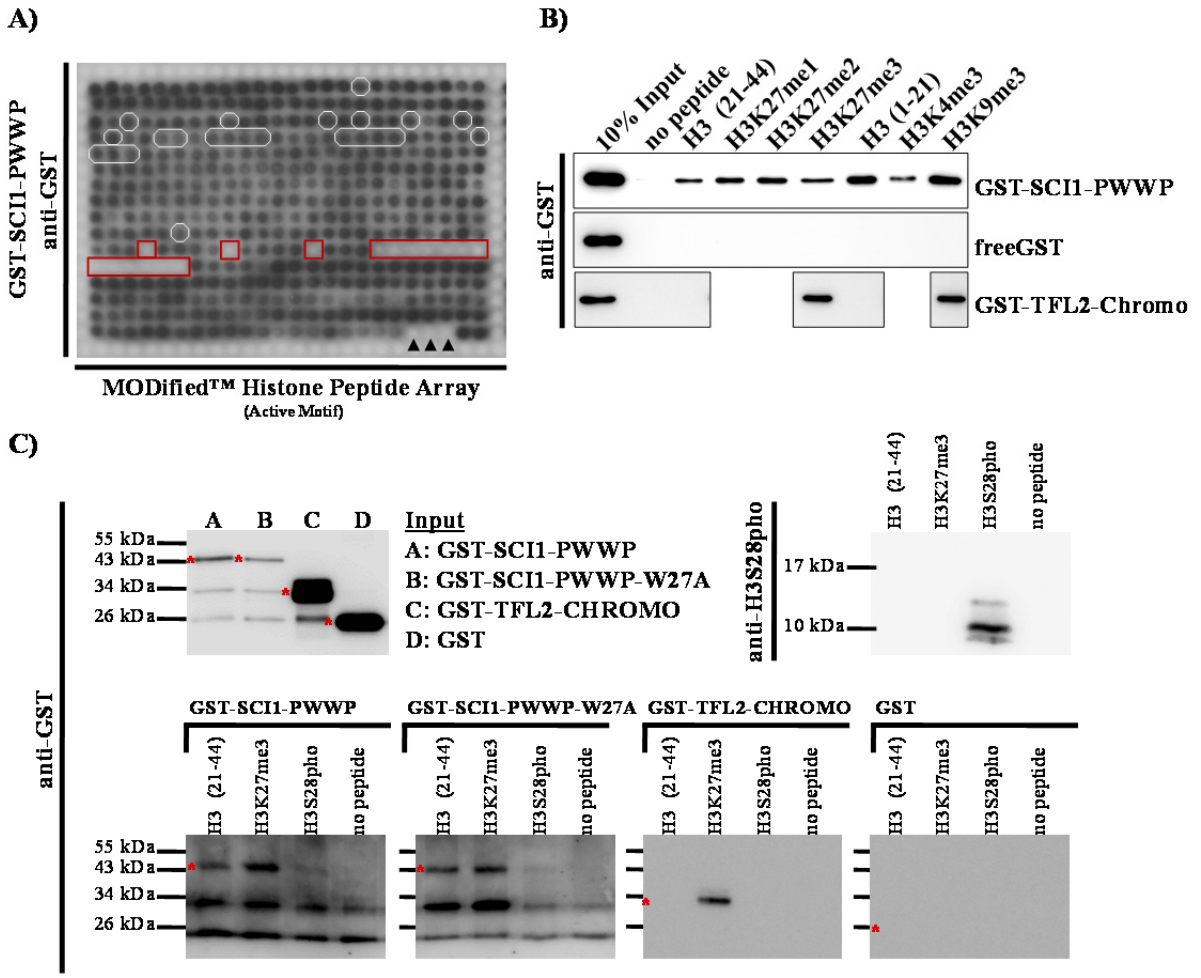


Figure 3.20: Histone peptide binding capacity of the SCI1 PWWP domain
(A) Hybridization of GST-SCI1-PWWP fusion protein to the MODified™ Histone Peptide Array (Active Motif). Each spot harbors a histone peptide modified at different amino acid residues (www. Activemotif.com). Spots with red boxes indicate peptides phosphorylated at H3S28, white circles mark spots comprising phosphorylated H3S10. Black arrowheads mark the negative controls (biotinylated control peptide, c-myc tag, HAc). (B) Immunoblot with anti-GST antibody of eluates from the *in vitro* binding experiment of GST-fusion proteins GST-SCI1-PWWP and GST-TFL2-CHROMO to biotinylated histone peptides. Free GST served as negative control. (C) *In vitro* binding studies of GST-SCI1-PWWP, GST-SCI1-PWWP-W27A and GST-TFL2-CHROMO to unmodified H3, H3K27me3 and H3S28pho. Free GST served as negative control, immunoblot analysis was performed using anti-GST and anti-H3S28pho antibodies. Red asterisks indicate the molecular weight of the GST-fusion protein analyzed, additional bands visible are either degradation products or background signal.

The omission from binding to H3S28 might be a functional feature of the SCI1 PWWP domain, and is exciting as serine 28 on histone H3 is directly adjacent to the PRC2 target site H3K27. To confirm the result of the array, the *in vitro* binding ability of GST-SCI1-PWWP to biotinylated H3S28pho peptide was analyzed (Figure 3.20 C). Similar to the result from the array hybridization, the GST-SCI1-PWWP fusion protein showed no binding to phosphorylated H3S28. The abundance of biotinylated H3S28pho peptide was tested in an additional immunoblot with an antibody against H3S28pho (Figure 3.20 C). GST and

GST-TFL2-CHROMO served as negative and positive controls, respectively. A point mutation of the conserved tryptophan 27 to alanine (SCI1-PWWP-W27A) did not alter the binding properties towards H3, H3K27me3 and H3S28pho peptides, suggesting that the mutation of the W27 residue alone is not sufficient to alter the binding specificity of the SCI1 PWWP domain. In summary, the SCI1 PWWP domain harbors histone peptide binding capacity and specifically neglects binding to peptides phosphorylated at H3S28, the amino acid adjacent to the PcG target site H3K27.

4 DISCUSSION

During the development of eukaryotic organisms, cell fate decisions are essential for proper organ and tissue formation. In plants, cell fate decisions are made throughout development, as new organs are continuously formed from pools of undifferentiated meristematic cells. The shoot apical meristem for example gives rise to rosette leaves during vegetative development, and cauline leaves and flowers during reproductive development. In order to adjust the time point of the switch from vegetative to reproductive development, plants are able to respond to endogenous and environmental cues (Srikanth and Schmid, 2011). Epigenetic marks like DNA methylation and histone modifications play an important role in the establishment and maintenance of cell fate.

Polycomb group (PcG) proteins are major regulators of chromatin state, and maintain the transcriptional repression of homeotic genes in a tissue specific manner (Chanvivattana et al., 2004; Moon et al., 2003; Schonrock et al., 2006). The general mechanisms of PcG mediated gene regulation, as well as the subunits of the POLYCOMB REPRESSIVE COMPLEX 2 (PRC2) mediating tri-methylation of H3K27, are highly conserved among the animal and plant kingdoms (Cao and Zhang, 2004; Cao et al., 2002; Goodrich et al., 1997; Luo et al., 1999; Grossniklaus et al., 1998).

The fundamental differences between plant and animal development and lifestyle, for example the fact that plants are sessile organisms and cannot escape from an unfavorable environment, demands for an adaption of molecular pathways to the special needs of the organism. Therefore it is not surprising that PcG mediated trimethylation of H3K27 and the resulting gene silencing is conserved in plants, while several key players identified in *Drosophila* are either missing, or represented by several genes in *Arabidopsis*. The PRC1 complex from *Drosophila* for example is involved in the recognition and maintenance of the H3K27me3 mark, but there are only few sequence homologs in *Arabidopsis.* Nevertheless, recent studies identified plant specific proteins that fulfill PRC1 like functions in *Arabidopsis* (Chen et al., 2010; Turck et al., 2007; Zhang et al., 2007b; Xu and Shen, 2008; Bratzel et al., 2010; Schatlowski et al., 2010).

There are several layers of PcG mediated gene regulation which are not yet sufficiently understood in plants. For example PRC2 recruitment and cell type or target gene specificity of PcG mediated gene silencing likely involves more plant specific proteins than discovered to date. This is illustrated by the large discrepancy between the number of PcG target genes identified in whole genome studies, and those target genes which were experimentally shown

to be regulated by specific complexes (Oh et al., 2008; Zhang et al., 2007a; Lafos et al., 2011; Makarevich et al., 2006). Furthermore novel plant specific proteins are likely involved in the interplay of activating and repressive histone modifications, the maintenance of the H3K27me3 mark during mitotic divisions and target gene reactivation, for example in response to abiotic stress.

The aim of this study was to functionally analyze the plant specific PWWP domain protein SWINGER/CURLY LEAF INTERACTOR 1 (SCI1, At3g03140), which was found to interact with the plant PRC2 histone methyltransferase subunits CURLY LEAF (CLF), SWINGER (SWN) and MEDEA (MEA). The genetic interaction with PcG proteins was addressed by analysis of morphological and molecular phenotypes of mutant alleles. It was analyzed whether the mutation of SCI1 influences the the expression levels of PcG target genes, and if SCI1 can be associated with PRC2 or PRC1 function. In order to identify *in vivo* interaction partners of SCI1, an immunoprecipitation experiment was carried out. SCI1 function during plant development, as well as redundancy with its two close homologs SCI2 and SCI3, was studied by mutant analyses. Furthermore a functional characterization of the SCI1 PWWP domain was performed to identify its binding properties to histone peptides.

4.1 SCI1 interacts with core members of the *Arabidopsis* PRC2 complex

SCI1 belongs to a small gene family in *Arabidopsis* with the two homologs SCI2 and SCI3, and all three proteins comprise an amino-terminal PWWP domain and a nuclear localization sequence (Figure 3.1). SCI1 is a plant specific protein with orthologs in monocots like *Oryza sativa* and *Brachypodium distachyon* as well as in dicots like *Medicago truncatula, Vitis vinifera* and *Populus trichocarpa*. No obvious homology to proteins in gymnospermae like *Pinus sylvestris* or the moss *Physcomitrella patens* were detected. This finding suggests that SCI1 function is needed for angiosperm development, whereas it is dispensable in lower plants.

In yeast cells, SCI1 is able to interact with the PRC2 core members CLF, SWN and MEA, as well as weakly with VRN2 (Figure 3.2). The fact that SCI1 shows interaction with several PcG proteins illustrates that the protein function is most likely associated with PcG mediated gene silencing, possibly as part of a specific PRC2 complex. A similar case was described for the protein BLISTER (BLI) that interacts with CLF and VRN2, and is involved in the regulation of a subset of PcG target genes (Schatlowski et al., 2010).

Furthermore SCI1 is capable to interact with FVE/MSI4, a WD40 protein involved in the regulating flowering time in *Arabidopsis* through repression of the MADS box transcription

factor *FLC* (Ausin et al., 2004). FVE/MSI4 is a component of the autonomous pathway and is involved in histone deacetylation. Furthermore FVE/MSI4 shows an interaction with CLF, and *fve/msi4* mutants exhibit reduced H3K27me3 levels as well as elevated H3 acetylation on *FLC* chromatin (Pazhouhandeh et al., 2011). Thus FVE/MSI4 function appears to be required for deposition of the H3K27me3 mark on *FLC* chromatin, which is mediated by the PRC2 complex. Additionally, FVE/MSI4 was shown to associate with the CUL4-DDB1 ubiquitin ligase that, discovered by mutant analyses, positively influences H3K27 tri-methylation of *FLC* (Pazhouhandeh et al., 2011). Hence the interaction of SCI1 with FVE/MSI4 is very interesting, as it links SCI1 function to several putative multimeric protein complexes including PRC2, which possibly interact with each other. All of these complexes are involved in the regulation of *FLC* expression and hence flowering time (Ausin et al., 2004; Pazhouhandeh et al., 2011; Jeon and Kim, 2011).

Furthermore SCI1 has the ability to form homodimers in yeast, which might be important for SCI1 function in *Arabidopsis.* It could serve to inactivate SCI1, for example in a concentration dependent manner, and SCI1 homodimerization is likely to compete with the binding of SCI1 to PRC2 components or FVE/MSI4. The ability of SCI1 to form homodimers furthermore suggests that SCI1 might also be able to interact with its homologs SCI2 and SCI3.

In conclusion it was shown in the yeast two-hybrid studies that SCI1 is able to interact with several components of the plant PRC2 complex as well as with the WD-40 protein FVE/MSI4. These proteins are all involved in the regulation of flowering time indicating a possible function for SCI1 in this process. Based on the interaction of SCI1 with several core components of PRC2 it can be proposed that SCI1 could be involved in PcG mediated gene regulation.

4.2 Sub-cellular localization of SCI1 *in planta*

Intracellular localization studies further supported the assumption of an *in planta* interaction of SCI1 with CLF and SWN, since all three proteins localize to the nucleus (Figure 3.3). The expression analyses were carried out in epidermal leaf cells of *Nicotiana benthamiana,* as well as in stable transformed *Arabidopsis* plants.

In *N. benthamiana,* SCI1-GFP fluorescence was observed in the nucleoplasm and in nuclear speckles of unknown nature (Figure 3.3). The localization to the nuclear speckles was abolished by truncation of the N-terminal PWWP domain, which led to a localization to the

nucleus and the cytoplasm. Thus the domain is important for the subcellular localization of SCI1 and might be needed to tether the protein to chromatin or other nuclear structures.

The nature of the nuclear speckles comprising SCI1-GFP is not yet known, but also plant PcG proteins were shown to localize to nuclear speckles. The PRC2 subunit VERNALIZATION 2 (VRN2) localizes to nuclear speckles when transiently expressed in onion epidermal cells (Gendall et al., 2001). Similar intranuclear speckle formation was observed for the *Arabidopsis* PRC1-like proteins LHP1 and EMF1 when expressed as GFP fusion proteins in *N. tabacum* and *Arabidopsis* or *Nicotiana benthamiana*, respectively (Calonje et al., 2008; Gaudin et al., 2001; Libault et al., 2005; Zemach et al., 2006).

Strikingly, *Drosophila* and mammalian PcG proteins localize to PcG bodies, which are observed as nuclear foci or speckles, and the number and size of them was found to vary depending on the cell type (Pirrotta and Li, 2011). Thus the nuclear speckles observed in *N. benthamiana* or onion cells comprising VRN2, LHP1, EMF1 or SCI1 might represent PcG bodies, and therefore display the localization of the plant PRC2 or PRC1 complexes in the nucleus. The fact that neither CLF nor SWN were detected in nuclear speckles might be due to the truncation of the catalytic SET domain in the fusion proteins used for sub-cellular localization studies in *N. benthamiana.*

In case of LHP1-GFP, the localization pattern is dependent on the CHROMO domain, that harbors H3K27me3 binding capacity, as well as on the cell type and the degree of differentiation (Libault et al., 2005; Turck et al., 2007). The differences observed due to cell type and differentiation status might be caused by the fact, that different subsets of target genes are silenced by the PcG machinery (Lafos et al., 2011).

EMF1 was shown to participate in PcG mediated gene silencing of *AG*, *PI* and *AP3*, and furthermore to directly bind to *FLC* and other gene loci (Kim et al., 2010; Calonje et al., 2008).

The similar sub-cellular localization patterns of SCI1 with VRN2, LHP1 and EMF1 should be further analyzed by co-localization studies. In case of co-localization, a possible *in planta* interaction could be tested by Fluorescence Resonance Energy Transfer (FRET) analysis. A possible interaction of SCI1 with EMF1 or LHP1 might link the protein to PRC1-like function. It will be important to resolve the nature of the nuclear speckles, and whether they are comparable to PcG bodies found in *Drosophila* and mammals. This specific sub-cellular localization of SCI1 could be a good starting point for such an analysis.

In *Drosophila* it has been found that the 3D organization of PcG proteins contributes significantly to their function (Bantignies and Cavalli, 2011). Some of the PcG bodies are proposed to be due to a genomic clustering of PcG target genes observed in *Drosophila* (Kharchenko et al., 2011). Furthermore it has been shown by chromosome conformation capture (3C) and DNA-fluorescent *in situ* hybridization (DNA-FISH), that PcG bodies comprise genes that map at large distances from each other, even though the association appears not to be stable (Bantignies et al., 2011; Bantignies and Cavalli, 2011). Insulator elements are known to mediate looping of chromatin, and might therefore be involved in the interaction of PcG target genes in *Drosophila* (Pirrotta and Li, 2011). Furthermore ncRNA were proposed to serve as scaffolds and nucleators for higher order assembly of PcG proteins, as it was also proposed for the assembly of stress bodies, cajal bodies and histone locus binding bodies (Mao et al., 2011). The interaction between PcG target genes suggests, that not all PcG target loci are recognized by specific DNA elements, but are rather silenced due to the proximity to a certain chromosome territory (Bantignies and Cavalli, 2011). Also in plants ncRNAs have been implicated in the regulation and recruitment of the PRC2 protein CLF to the target gene *FLC* (Heo and Sung, 2011). Further studies will have to reveal whether ncRNAs play a general role in the regulation and recruitment of PcG proteins in plants, and whether plant PcG proteins can also be organized in PcG bodies, which might occur in a cell type dependent manner.

In *Arabidopsis* the intracellular localization of SCI1-GFP was analyzed using stable transformed lines comprising the complementing SCI1pro::SCI1-GFP transgene (Figure 3.3). In roots cells, the sub-cellular localization of the SCI1-GFP fusion protein is similar to proteins that localize to the nuclear periphery, like the coiled coil protein LITTLE NUCLEI 1 (LINC1) (Figure 3.18) (Dittmer et al., 2007). In contrast, CLF-GFP and SWN-GFP fusion proteins show nucleoplasmic localisation in root cells, which partly overlaps with the pattern observed for SCI1-GFP (Figure 3.3). Like CLF, SCI1-GFP fluorescence does not co-localize with the DAPI stained heterochromatic chromocenters, suggesting a euchromatic localization and function for SCI1 in *Arabidopsis* (Figure 3.3). This finding suggests that SCI1, like PcG proteins, might have a function in gene regulation rather than in heterochromatin formation (Turck et al., 2007).

The observation that SCI1-GFP is not localized to distinct speckles in *Arabidopsis* might be due to differences in chromatin composition between *Arabidopsis* in *N. benthamiana*. The

small genome of *A. thaliana* (1C = 135Mbp) consists of only 10-15 % heterochromatin, while the 33 times larger genome of *Nicotiana tabacum* contains a considerably larger amount of repetitive sequences (Houben et al., 2003). Instead of a concentration of silenced sequences in heterochromatic chromocenters as observed in *Arabidopsis*, species with larger genomes exhibit euchromatic regions that are interspersed with repetitive sequences. Thus heterochromatic sequences are abundant throughout the nucleus (Houben et al., 2003). The localization pattern of a euchromatic protein in *N. benthamiana* might therefore appear different from the pattern observed in *Arabidopsis*. In order to assess the chromatin nature of the nuclear speckles in *N. benthamiana,* which SCI1-GFP fusion proteins localize to, it will be essential to perform co-localization studies of euchromatic and heterochromatic hallmarks with SCI1-GFP. This could for example be achieved by immunostaining of isolated nuclei from plants expressing SCI1-GFP, with anti-GFP and anti-H3K9m2 (heterochromatic) or anti-H3K4me3 (euchromatic) antibodies.

In summary, the subcellular localization studies with SCI1-GFP revealed, that SCI1 localizes to distinct nuclear speckles in *N. benthamiana* epidermal leaf cells, and to the nuclear periphery in *Arabidopsis* roots. Whether the respective domains have a common chromatin state, or whether the localization differs due to the different cell types, has yet to be studied. The localization to the nuclear periphery suggests that, at least in roots, SCI1 might regulate genes positioned in this sub-compartment. The localization to nuclear speckles observed in *N. benthamiana* similar to the PcG proteins like VRN2, LHP1 and EMF1, which furthermore resembles PcG bodies found in *Drosophila* and mammals, suggests that SCI1 might be part of a higher order complex possibly including PRC1 as well as PRC2 components.

4.3 *In planta* co-localization and interaction studies

Compared to SCI1, the CLFΔSET and SWNΔSET fusion proteins showed a uniform localization to the nucleoplasm in *N. benthamiana* (Figure 3.3). Since the localization of SCI1 to nuclear speckles appears to be important for protein function, an interaction with CLF and SWN might only take place at the interphase of the nuclear speckles and the nucleoplasm.
The constructs of CLF and SWN analyzed in the co-localization studies were both truncated missing the catalytic SET domain. This might result in a localization pattern that differs from that of the full length proteins. It would be essential to analyze the subcellular localization of complementing full length proteins in *N. benthamiana*, which might also show localization to nuclear speckles similar to SCI1. Even though the truncated constructs showed interaction

with SCI1 and other PcG proteins in yeast (Schatlowski et al., 2010; Chanvivattana et al., 2004), the truncation might result in aberrant protein folding or sub-cellular localization *in planta*.

While interaction studies performed in yeast have the advantage that also transient interactions can be detected, also artifacts can be produced that do not reflect the native situation *in planta*. Therefore interaction studies of SCI1 with CLFΔSET and SWNΔSET were carried out using the split-luciferace complementation assay in transiently transformed *N. benthamiana* epidermal leaf cells (Figure 3.4). In this assay, the interaction of the full length SCI1 construct with CLFΔSET and SWNΔSET could not be confirmed. As discussed above, this could be due to the different localization patterns of the respective proteins in *N. benthamiana,* especially to the localization of SCI1 to nuclear speckles (Figure 3.3; chapter 4.2). Further analysis using transiently transformed protoplasts from *Arabidopsis*, or most appropriate, complementing and thus functional constructs in *Arabidopsis,* might elucidate whether SCI1 is able to interact with CLF and SWN *in planta*.

In order to isolate putative complex partners of SCI1 in *Arabidopsis*, possibly including PcG proteins, an immunoprecipitation experiment was performed with nuclear protein extract from SCI1pro::SCI1-GFP transgenic plants, followed by characterization of the proteins by LC-MS/MS. Statistical analyses yielded a list of 110 potential complex partners for SCI1, which were significantly enriched compared to the Col-0 control samples (Table 7.1; Table 7.2), and a selection of interactors are listed in Table 3.1.
In this experiment MSI1, a component of the plant PRC2 complex, was isolated as interaction partner of SCI1. This finding affirmed the hypothesis that SCI1 interacts with plant PcG proteins *in planta*. Nevertheless, as MSI1 is also part of other complexes involved in chromatin remodeling like CAF-1 (Chromatin assembly factor 1), it cannot be excluded that the MSI1 found as interaction partner of SCI1 was not part of a PRC2 complex (Exner et al., 2006; Kohler et al., 2003). Furthermore the identification of homologs of PcG proteins from *Drosphila,* and of proteins involved in chromatin remodeling support the hypothesis that SCI1 function is associated with chromatin. The fact that neither CLF nor SWN were detected among the putative interaction partners might be due to the experimental conditions, in which the PRC2 complex might not be stable. Supporting this hypothesis, a parallel experiment using nuclear extract from SWN::SWN-GFP transgenic plants was carried out, and the bait protein SWN could be detected, but no other PRC2 members (data not shown).

Since several proteins of the 110 significantly enriched putative interaction partners are predicted to localize to chloroplasts, which is not the case for SCI1, these proteins might be false positives due to the extraction procedure. Further analysis of the putative complex partners will be necessary to confirm the interaction with SCI1, and whether SCI1 function is related to that of the interaction partners. In order to address these questions double mutant analyses of selected putative complex partners will be carried out as well as co-immunoprecipitation experiments.

In summary it can be stated that SCI1 co-localizes with CLF and SWN to the nucleus, but a direct interaction of the proteins could not be observed *in planta*. One possible reason for this is the distinct sub-cellular localization of SCI1 to nuclear speckles in *N. benthamiana*, and the interaction studies will have to be continued with complementing protein constructs in *Arabidopsis*. The complex isolation experiment in *Arabidopsis* yielded a list of putative complex partners of SCI1. The interaction of SCI1 with MSI1 once more suggests that SCI1 is able to interact with PcG proteins *in planta*.

4.4 SCI1 and its interaction partner LINC1 regulate nuclear size in *Arabidopsis*

The immunoprecipitation experiment furthermore identified LINC1 as interaction partner of SCI1-GFP, which is an ortholog of the Nuclear Matrix Constituent Protein 1 (NMCP1) from *Daucus carota,* and NMCP1 as well as LINC1 localize to the nuclear periphery (Dittmer and Richards, 2008; Masuda et al., 1997). Together with its homolog LINC2, LINC1 functions in the regulation and dynamics of nuclear size, which is reduced in *linc1* and *linc2* mutants (Dittmer et al., 2007; vanZanten M. et al., 2011). Despite the previously observed connection between nuclear size and DNA content, the endoploidy levels are not changed in *linc1* mutants compared to the wildtype. LINC1 and LINC2 rather exhibit a more densely packed DNA, which results in nuclear size reduction (Dittmer et al., 2007).

Similar to LINC1-GFP, the sub-cellular localization of SCI1-GFP in *Arabidopsis* root cells was observed at the nuclear periphery (Figure 3.18). Furthermore a similar reduction of nuclear size as in *linc1* mutants was observed in three independent *sci1* alleles, which is rescued by the SCI1pro::SCI1-GFP transgene (Figure 3.18). Interestingly, the endoploidy levels in *sci1* mutants are not changed compared to the wildtype. This suggests that SCI1, like LINC1 and LINC2, might be a determinant for nuclear size regulation in plants (Dittmer et al., 2007).

Compared to animal systems, there is not much known about the nuclear periphery in plants. In eukaryotes, the nuclear envelope was described to consist of the inner nuclear membrane, the outer nuclear membrane, nuclear pore complexes and the nuclear lamina, which is a meshwork of proteins that is closely attached to the inner nuclear membrane and to chromatin (Dechat et al., 2008). Lamins are intermediate filament proteins, and directly or indirectly anchor chromatin to the nuclear lamina. While it has been reported that heterochromatin associates with the nuclear periphery, this region has also been associated with gene activity (Dechat et al., 2008; Arib and Akhtar, 2011). The lamins are proposed to serve as scaffolds regulating chromatin organization, transcription and DNA replication (Dechat et al., 2009). Lamins have no sequence homologs in *Arabidopsis*, but proteins like LINC1 could fulfill analogous functions.

Furthermore, the nuclear periphery was shown to be involved in cell type specific PcG mediated gene regulation. In myoblast cells, PcG proteins are directed to the nuclear periphery through interaction with the homeodomain protein Msx1, causing an enhanced level of repressive H3K27me3 at Msx1 target genes (Wang et al., 2011).

Thus SCI1 could have a function in PRC2 recruitment to the nuclear periphery, for example in *Arabidopsis* root cells, in order to silence specific target genes, which may possibly be involved in nuclear size regulation.

To further investigate the connection of SCI1 and LINC1 function, double and triple mutants with *linc1* and *linc2* should be established and analyzed. Further, an *in planta* co-localization and interaction analyses of SCI1 and LINC1 would be necessary to affirm the interaction detected in the complex isolation experiment. In summary, it is an interesting finding that SCI1, a protein connected to PcG mediated gene regulation, also has a function in nuclear architecture and thus chromatin organization.

4.5 SCI1 genetically interacts with CLF

4.5.1 The SCI1 expression pattern overlaps with those of CLF, SWN and MEA

The expression pattern of *SCI1* and its two homologs *SCI2* and *SCI3* was studied on steady state RNA level as well as by reporter gene analysis in *Arabidopsis* plants (Figure 3.6). Transcripts of *SCI1, SCI2* and *SCI3* can be detected throughout plant development (Figure 3.6). ß-Glucuronidase activity of the complementing SCI1-GUS and SCI3-GUS transgenes was mainly observed in the vasculature of the root and the shoot, primary root tips as well as

in newly formed side roots. In flowers both transgenes were found to be expressed in sepals and carpels, SCI3-GUS activity is additionally detected at the tip of the filaments (Figure 3.6). The observed expression patterns of *SCI1, SCI2* and *SCI3,* overlap with previously reported expression patterns of *CLF* and *SWN*, both obtained by *in situ* hybridization experiments. *CLF* and *SWN* are strongly expressed in the shoot apical meristem, leaf primordia and leaves, as well as in floral meristems, the floral organs and ovules (Chanvivattana et al., 2004; Goodrich et al., 1997). SCI1 is not involved in PcG mediated gene regulation in petals and stamen, as SCI1 and CLF/SWN expression patterns do not overlap in these organs. *MEA* and *SWN* are expressed during ovule and seed development and thus overlap with *SCI1* expression (Luo et al., 2000; Grossniklaus et al., 1998). In conclusion, the expression patterns of SCI1, SCI2 and SCI3 overlap with those of the PRC2 histone methyltransferases during *Arabidopsis* development, which is congruent with a possible physical interaction of the proteins *in planta*.

4.5.2 SCI1 and its homologs SCI2 and SCI3 exhibit redundant functions

The similar protein structure as well as the overlapping expression patterns of SCI1 and its two homologs SCI2 and SCI3 imply that the proteins share redundant functions (Figure 3.1; Figure 3.6). Double mutant analyses revealed a high level of redundancy among the homologous proteins, as these plants phenotypically resembled the wildtype. Triple mutants, however, were seedling lethal, and seeds often failed to germinate, revealing an essential role for the *SCI* genes in plant development (Figure 3.12). The triple mutants display desiccation of cells and differentiation of the apical shoot and root meristems (Figure 3.13). This severe phenotype suggests that the *SCI* genes are essential from an early stage of plant development. Analysis of embryo development might discover at which time point SCI function is needed first. In contrast to the triple mutants, PcG mutants like *clf-28 swn-7,* that exhibit a loss of vegetative PRC2 activity, still proliferate and can be maintained in tissue culture. The seedling lethality of the triple mutants might be caused by misexpression of few target genes, and lines harboring an inducible SCI1 rescue construct in *sci1-1 sci2-2 sci3-2* triple mutant background might help to overcome the seedling lethality and could be useful to study SCI1 function during plant development.

4.5.3 Mutation of *SCI1* severely enhances the *clf* phenotype

In order to analyze *SCI1* function, three different *sci1* mutant alleles were analyzed, one of which exhibits a stop codon in the PWWP domain, and therefore likely resembles a null allele. Phenotypic analyses revealed mild phenotypes, including an early onset of flowering and a reduction of primary root length (Figure 3.5; Figure 3.10). Double mutants of *sci1-1* and *sci1-3* alleles with *clf-28* mutants, however, showed a strong, dosage dependent enhancement of the *clf* mutant phenotype. The enhancement was observed in respect to early flowering, leaf curling and overall plant size, and appears to be day length independent (Figure 3.14).

The *clf* mutant phenotype can be enhanced by mutation of redundantly acting proteins like SWN (Chanvivattana et al., 2004), or by regulation of PcG target genes that are causal for the phenotypic features observed in *clf* mutants.

The *clf* mutant phenotype was initially proposed to be caused by misexpression of the floral organ identity gene *AG*, as the *ag* mutant is able to suppress the *clf* phenotype (Goodrich et al., 1997). A recent suppressor screen, however, identified three mutations that suppress the *clf* phenotype despite high *AG* misexpression in vegetative organs (Lopez-Vernaza et al., 2012). These suppressors were the PcG target genes *FPA* and *FT*, both promoting the onset of flowering, and *SEP3*, which is, like *AG,* a floral organ identity gene. In all suppressor mutants, including *clf ag* double mutants, *SEP3* expression levels were strongly reduced, which implies that *SEP3* is the major determinant for the *clf* mutant phenotype (Lopez-Vernaza et al., 2012). Furthermore *FLC* expression levels are altered in *clf* mutants, and *flc clf* double mutants show an enhanced early flowering phenotype (Figure 3.14) (Lopez-Vernaza et al., 2012). The reduction of *FLC* expression appears to only affect flowering time, but no other features of the *clf* phenotype. Double and triple mutant analysis indicated that elevated levels of the floral integrator *FT,* which functions downstream of *FLC,* are causal for the early flowering time in *clf* and *clf flc* mutants (Lopez-Vernaza et al., 2012). Even though PcG proteins function as repressors of *FLC*, a downregulation of *FLC* has been observed in mutants of PcG associated proteins like INCURVATA 2 (ICU2) and BLI (Schatlowski et al., 2010; Barrero et al., 2007)

Hence the expression levels of these PcG target genes were analyzed in *sci1-1* single and *sci1-1 clf-28* double mutants, and it was shown that altered expression levels of *FLC* and *SEP3* are likely to be causal for the phenotypic enhancement observed in *sci1-1 clf-28* double mutants (Figure 3.15).

FLC expression levels are strongly reduced in *sci1* single mutants, and *SCI1* was found to function epistatically to *FLC* in respect to flowering time (Figure 3.7; Figure 3.8). In *clf-28* single mutants an elevated *FLC* transcript level can be observed, which is reduced to a level similar to Col-0 in *sci1-1 clf-28* double mutants. This reduction of *FLC* likely has downstream effects, like a further elevation of *FT* expression, which is indeed observed in *sci1-1 clf-28*, resulting in an early flowering phenotype (Figure 3.15).

However, *flc-3 clf-28* double mutants, in which *FLC* function is completely impaired, showed a similar reduction of flowering time as *sci1-1 clf-28* double mutants, even though *sci1-1 clf-28* only exhibit *FLC* expression levels similar to the wildtype (Figure 3.14; Figure 3.15) (Lopez-Vernaza et al., 2012). Thus the effect of reduced *FLC* levels is not comparable between *flc-3 clf-28* and *sci1-1 clf-28* double mutants.

Additionally, compared to *flc-3 clf-28* double mutants, a more severe enhancement of leaf curling and a stronger reduction of plant size can be observed in *sci1 clf-28* double mutants (Figure 3.14), suggesting misregulation of additional target genes such as *SEP3*.

SEP3 is significantly upregulated in *sci1-1* single and *sci1-1 clf-28* double mutants compared to Col-0, whereas no change in *SEP3* expression level was observed in *flc-3* mutants, thus *SCI1* appears to be regulating *SEP3* expression in an independent manner (Figure 3.15). This result contradicts the previous finding that *SEP3* expression level is enhanced in *flc-3*, which could be due to the fact that the previous analysis was carried out in *FRI*+ background (Deng et al., 2011), As *SEP3* transcript levels are significantly enriched in *sci1-1 clf-28* double mutants compared to *clf-28* single mutants, *SEP3* is likely the cause for the enhanced leaf curling and plant size reduction phenotype observed in the double mutants. Elevated *SEP3* levels might also be involved in the early flowering phenotype of the double mutant, as *FT* misexpression was shown to depend on *SEP3* levels in *clf* mutant background (Lopez-Vernaza et al., 2012). Collectively, *SCI1* appears to be required, in addition to *CLF* and *FLC*, to repress *SEP3* expression in *Arabidopsis* seedlings.

Further expression analyses revealed an additional upregulation of *AG*, *AGL19* and *AGL24* in *sci1-1 clf-28* double mutant seedlings, while expression levels of *PI* and *STM* are not altered (Figure 3.15). The enhanced misexpression of *AG* is probably not causal for an enhancement of the phenotype, as strong *AG* misexpression is still observed in the *sep3 clf* or *ft clf* double mutants, that show a suppression of the *clf* phenotype (Lopez-Vernaza et al., 2012).

AGL19 and *AGL24* function as floral activators in response to vernalization in an *FLC* independent manner, and transcript levels were significantly enhanced in the *sci1-1 clf-28* double mutant compared to Col-0 and *clf-28* (Figure 3.15) (Alexandre and Hennig, 2008;

Michaels et al., 2003; Schonrock et al., 2006).Thus loss of *SCI1* function in *clf-28* background causes the enhanced transcript levels of the floral activators, which in case of *AGL24* is also observed in *flc-3 clf-28* double, but not *flc-3* or *sci1-1* single mutants. The enhanced level of *AGL19* and *AGL24* transcripts might be caused by a simultaneous loss of *FLC* and *CLF* function. Whether *AGL19* and *AGL24* expression contributes to the enhancement of the phenotype observed in *sci1-1 clf-28* mutants compared to *flc-3 clf-28* could be analyzed by triple mutant analyses.

In conclusion, SCI1 functions as regulator of the PcG target genes *FLC*, *SEP3*, *AGL19*, *AGL24*, *FT* and *AG.* As some of these genes also transcriptionally regulate each other, it will have to be solved whether SCI1 directly or indirectly influences the expression of these genes. Nevertheless it can be concluded that the simultaneous misexpression of *FLC* and *SEP3* is most likely causal for the phenotype of the *sci1-1 clf-28* double mutant. This hypothesis could be tested by overexpression of *SEP3* in *flc-3 clf-28* double mutants, which should cause strong leaf curling and a reduction of overall plant size in addition to the early flowering phenotype. Furthermore *sci1-1 flc-3 clf-28* triple mutants should exhibit a phenotype stronger than *flc-3 clf-28* mutants, but probably similar to *sci1-1 clf-28* double mutants. The results of the expression analysis are summarized in Figure 4.1.

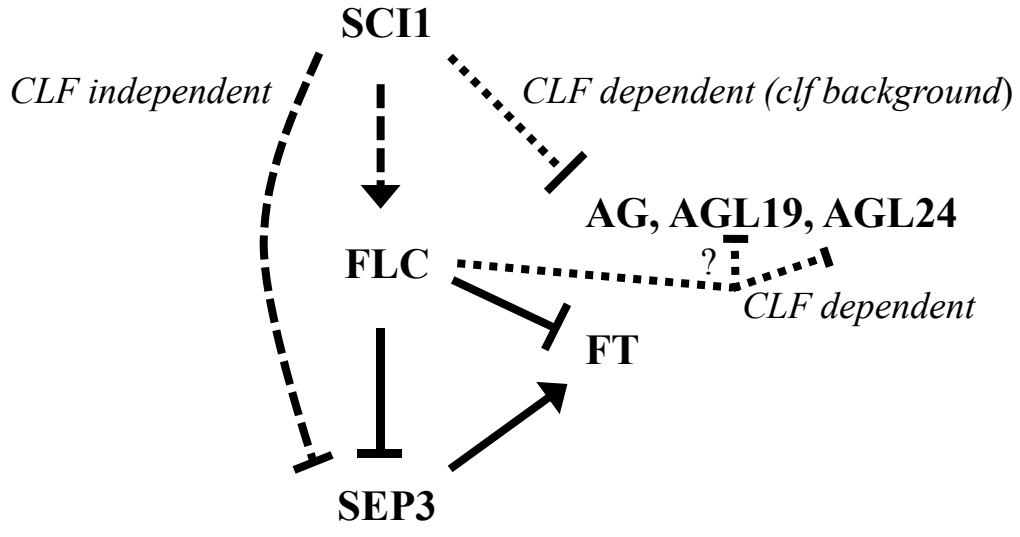


Figure 4.1: SCI1 regulates PcG target genes in a CLF dependent and independent manner
A model for SCI1 function is depicted in a simplified scheme, arrows implicate gene activation, lines with terminal bars indicate gene repression. As it is not known whether SCI1 directly or indirectly regulates the PcG target genes, dotted lines are used. Also regulation of *AGL19* and *AGL24* by FLC in *clf* mutant background is depicted in dotted lines as it is not yet proven. Previously described regulations are depicted with solid lines. SCI1 regulates *SEP3* and *FLC* expression independently of CLF, while *AG*, *AGL19* and *AGL24* are only affected in *clf* mutant background. The regulation of *FT* is probably not direct, as *FT* is regulated by FLC and SEP3.

4.6 SCI1 is a novel regulator of *FLC*

In *sci1* mutant alleles, expression analyses revealed a severe downregulation of *FLC* transcript (Figure 3.8). Double mutant analyses with the *flc-3* null allele furthermore showed that *SCI1* is epistatic to *FLC* (Figure 3.7). These results imply that SCI1 is a novel regulator of *FLC* expression, and the question arises, whether the regulation is direct or indirect. Direct regulation of *FLC* would imply that SCI1 is able to bind to the *FLC* locus, while indirect regulation would include transcriptional control of other *FLC* regulators, or the alteration of chromatin modifications at the *FLC* locus.

The possibility of a direct interaction of SCI1 with *FLC* chromatin was addressed by ChIP analysis, using a transgenic line harboring the complementing SCI1pro::SCI1-GFP transgene. SCI1-GFP binding to *FLC*, however, was not observed (data not shown), which might have several reasons. For once the SCI1-GFP fusion protein has a low stability, similar to what has been observed for CLF, with which binding studies were carried out using an overexpression line (Schubert et al., 2006). Furthermore the binding of SCI1-GFP to *FLC* might only occur in specific cell types or under specific conditions. SCI1-GFP might additionally have to compete with the binding of the potentially redundantly acting endogenous proteins SCI2 and SCI3. Thus this experiment should be repeated in *sci1-1 sci2-2 sci3-2* triple mutant background. To reveal direct target genes of SCI1-GFP, a ChIP experiment followed by hybridization of the chromatin to whole genome tiling arrays (ChIP-chip) could be carried out. This would also show whether H3K27me3 and SCI1 binding sites overlap on targeted chromatin.

CLF is able to directly bind to the *FLC* locus, and this binding is significantly decreased in *FRIGIDA*+ (*FRI*+) background (Doyle and Amasino, 2009). FRI is a coiled coil protein that acts dominantly to elevate *FLC* expression, and therefore to prolong the vegetative growth phase (Johanson et al., 2000; Simpson and Dean, 2002). The dominant allele *clf-59*, suppresses the late flowering phenotype of *FRI*+ plants in a dosage dependent manner (Doyle and Amasino, 2009). The *FLC* locus exhibits low levels of H3K27me3 in non-vernalized *FRI*+ plants compared to Col-0. The *clf-59* plants however exhibit elevated levels of H3K27me3 at *FLC* in *FRI*+ background, which leads to the repression of *FLC* expression and thus to a suppression of the late flowering phenotype.

Double mutant analysis with *sci1-1* revealed, that *sci1-1 FRI*+ plants show an intermediate flowering time compared to *FRI*+ and *sci1-1* (Figure 3.9), which is correlated with lower levels of FLC-LUC reporter gene activity in *sci1-1 FRI*+ background (Figure 3.8). It would be essential to analyze H3K27me3 levels at *FLC* in *sci1-1 FRI*+ plants to reveal whether this chromatin mark is involved in the reduction of *FLC* expression in *sci1-1* background.

Similar to *sci1-1*, also mutants of the TrxG genes *ATX1* and *ATXR7* display a partial suppression of the late flowering phenotype in *FRI*+ background, which is correlated with reduced levels of H3K4me3 at the *FLC* locus (Tamada et al., 2009; Pien et al., 2008). But in contrast to *sci1* mutants, the TrxG mutant *atx1* suppresses the *clf* phenotype, thus it is not likely that SCI1 has a general TrxG associated function (Saleh et al., 2007).

There are several other proteins which were found to suppress the *FRI*+ late flowering phenotype that can be organized in two groups (Kim and Michaels, 2006). Mutants of the first group include putative members of the PAF1 complex like *EARLY FLOWERING IN SHORT DAYS* (*EFS*) and *EARLY FLOWERING7* (*ELF7*). They suppress *FLC* expression and H3K4me3 levels in autonomous pathway mutants as well as in *FRI*+, but show other pleiotropic phenotypes as well (Kim et al., 2005; Kim and Michaels, 2006; He et al., 2004). The second group includes *SUPPRESSOR OF FRIGIDA 4* (*SUF4*), *FRIGIDA-LIKE1 (FRL1)* and *FRIGIDA ESSENTIAL1 (FES1)* which are essential for FRI dependent elevation of *FLC* levels, but have little effect on *FLC* expression in autonomous pathway mutants (Michaels et al., 2004; Schmitz et al., 2005; Kim and Michaels, 2006).

SCI1 probably belongs to the first group of *FRI*+ suppressors, since the suppression of late flowering by *sci1* is not *FRI* specific, but is also observed for the autonomous pathway mutant *fve*. The *sci1 fve* double mutants exhibit an intermediate flowering time phenotype compared to the single mutants, suggesting an intermediate level of *FLC* in the double mutant (Figure 3.9). The same result was obtained for *fve clf-59* double mutants (Doyle and Amasino, 2009). Thus the reduction of flowering time in *sci1 fve* might be due to enhanced levels of H3K27me3 at the *FLC* locus. The direct interaction of SCI1 with FVE/MSI4 in yeast suggests that SCI1 could also be involved in the acetylation status of *FLC*, which could be analyzed by ChIP. Since FLC, as master regulator of flowering time, is regulated by many parallel pathways, it cannot be excluded that SCI1 acts indirectly through one of them. But from these results it can be concluded that the novel *FLC* regulator SCI1 acts independent of FRI and FVE, and might be a regulator of the chromatin state at the *FLC* locus.

4.6.1 Analysis of histone modification levels in *sci1-1* and PcG mutants

The misexpression of PcG target genes in *sci1-1 clf-28* double mutants could be caused by altered histone modification levels at these loci. In order to analyze whether histone modifications are affected on a global scale in *sci1-1 clf-28* double mutants compared to other PcG mutants, immunoblot analysis of histone enriched protein extracts were carried out (Figure 3.17) (Lafos et al., 2011). H3K27me3 levels were found to be completely dependent

on vegetative PRC2 function, as the modification was not detectable in *clf-28 swn-7* double mutants. The reduction of H3K27me3 correlated with the severity of the phenotype of different PcG mutants. The reduction observed for *clf-28* single mutants was not enhanced in *sci1-1 clf-28* double mutants, suggesting that SCI1 acts either downstream of PRC2, functions in specific cell types or affects only a small number of target genes. Another possibility is that SCI1 function is masked by its homologs SCI2 and SCI3.

Further analyses of H3K27me1, H3K27me2 and H3K4me3 levels were carried out and novel insights could be gained concerning the effect of partial or complete loss of PRC2 function on the global abundance of these histone modifications (Figure 3.17) (Lafos et al., 2011).
In order to identify the substrate of the PRC2 complex, also the global levels of H3K27me1 and H3K27me2 were analyzed (Figure 3.17). The H3K27me1 mark is preferentially found in heterochromatic chromatin, and is not affected by mutation of PcG proteins, which is consistent with the finding that monomethylation of H3K27 is carried out by the histone methyltransferases ATXR5/6 in *Arabidopsis* (Jacob et al., 2009). In contrast, H3K27me2 is a euchromatic mark (Jacob et al., 2009), and in this study it could be shown that it is not mediated by PRC2, but that a loss of vegetative PRC2 activity leads to an elevation of H3K27me2 levels (Figure 3.17). This finding suggests that H3K27me2 is the substrate for PRC2, but this is contradicted by the fact that a partial loss of PRC2 activity, as in *vrn2-1 emf2-10* double mutants, causes a reduction of H3K27me2 compared to the wildtype (Figure 3.17). It can be concluded that PRC2 is not responsible for the dimethylation of H3K27, and the histone methyltransferase catalyzing this mark still needs to be discovered.
Surprisingly, a strong enhancement of H3K4me3 was observed in *clf-28 swn-7* double mutants (Figure 3.17). Partial loss of PRC2 function does not lead to enhanced H3K4me3 levels, suggesting that low levels of PRC2 are sufficient to avoid global accumulation of H3K4me3. The activating H3K4me3 mark is mediated by TrxG proteins that act antagonistically to PcG proteins. Mutants of the TrxG gene *ULTRAPETALA1* (*ULT1*) are able to suppress the *clf* mutant phenotype (Saleh et al., 2007; Carles and Fletcher, 2009). At the *AG* locus, *ult1* mutants show an increase of H3K27me3, while *clf* mutants display increased levels of H3K4me3. This finding fits to the observation that, on a global scale, H3K4me3 levels are strongly increased in *clf-28 swn-7* mutants (Figure 3.17), which suggests that trithorax activity is needed to balance PcG activity and vice versa (Carles and Fletcher, 2009; Schubert et al., 2006). This experiment reinforces the finding that PcG and TrxG complexes act antagonistically, and that loss of one part results in an increase of the counterpart activity.

The analysis of histone modifications at *FLC* and other loci in *sci1-1* mutants is essential to understand SCI1 function. Therefore the abundance of H3K4me3 and H3K27me3 on *FLC*, *SEP3* and *AG* in *sci1-1* mutants compared to Col-0 was furthermore tested by ChIP experiments (Figure 3.16). For both modifications no difference could be observed in *sci1-1* compared to the wildtype. At least in case of *FLC* this is surprising, as *FLC* expression is strongly reduced in the *sci1-1* single mutant. Since Col-0 plants exhibit high levels of H3K27me3 at *FLC*, these analyses should also be performed in *FRI*+ or *fve-3* backgrounds which show low levels of H3K27me3 at *FLC,* and changes would therefore be more prominent (Doyle and Amasino, 2009). It is possible that other modifications than those tested are involved in the regulation of FLC by SCI1, thus also H3K36m3, and histone acetylation will be analyzed in the future. Furthermore it is possible that SCI1 enhances *FLC* expression without altering histone modification levels, for example through interaction with the transcription machinery.

4.7 Functional analysis of the SCI1 PWWP domain

The *Arabidopsis* genome contains at least 19 PWWP domain proteins (http://www.uniprot.org), among them the TrxG proteins ATX1-5. The PWWP domain is conserved among eukaryotes, and while it was first described as unspecific DNA binding motif, the PWWP domains from the mammalian DNA methyltransferase DNMT3A/B, Pdp1 from fission yeast or from the zebrafish Brpf1 protein were shown to harbor chromatin targeting ability (Ge et al., 2004; Wang et al., 2009; Vezzoli et al., 2010).

From *in vitro* binding studies performed in this work it can be concluded that the SCI1 PWWP domain also exhibits histone peptide binding ability (Figure 3.20). The SCI1-PWWP domain was found to specifically neglect binding to peptides phosphorylated at H3S28, while other histone peptides were bound irrespective to other modifications (Figure 3.20). Congruent with these findings also several histone proteins were found as interaction partners of SCI1 in the immunoprecipitation experiment (Table 3.1).

The residue H3S28 is directly adjacent to the PRC2 target site H3K27 which is bound by the PRC1-like protein LHP1/TFL2. The proximity of amino acids that are subject to covalent modifications on the histone tails, such as at the positions H3K9/H3S10 and H3K27/H3S28, triggered the hypothesis of the existence binary switches (Fischle et al., 2003a; Fischle et al., 2003b). It was proposed, that the phosphorylation of an amino acid adjacent to a methyl mark leads to a loss of binding of effector proteins to the methyl mark. In case of H3K9/H3S10 it was shown that the dissociation of HP1 from tri-methylated H3K9 during mitosis is

dependent on H3S10 phosphorylation (Fischle et al., 2005; Hirota et al., 2005). Furthermore recent studies in mammalian cell culture could show that the PRC2 complex is displaced from chromatin with the double mark H3K27me3H3S28pho, which resulted in a release of transcriptional silencing of PcG target genes (Gehani et al., 2010; Lau and Cheung, 2011). The H3K27me3H3S28pho mark can also be detected in plant nuclei by immunostainings, and shows a euchromatic localization (C. Klose; personal communication). The double mark could therefore also be important in plants to restrict PcG complexes from binding to chromatin. The specific omission from binding to H3S28pho of SCI1 suggests that SCI1 might recognize marked chromatin and could act to restrict PRC2 or LHP1/TFL2 from binding to target genes. A model for SCI1 function combining the recognition of H3S28pho, the physical interaction with PRC2 and the regulation of PcG target genes is depicted in Figure 4.2.

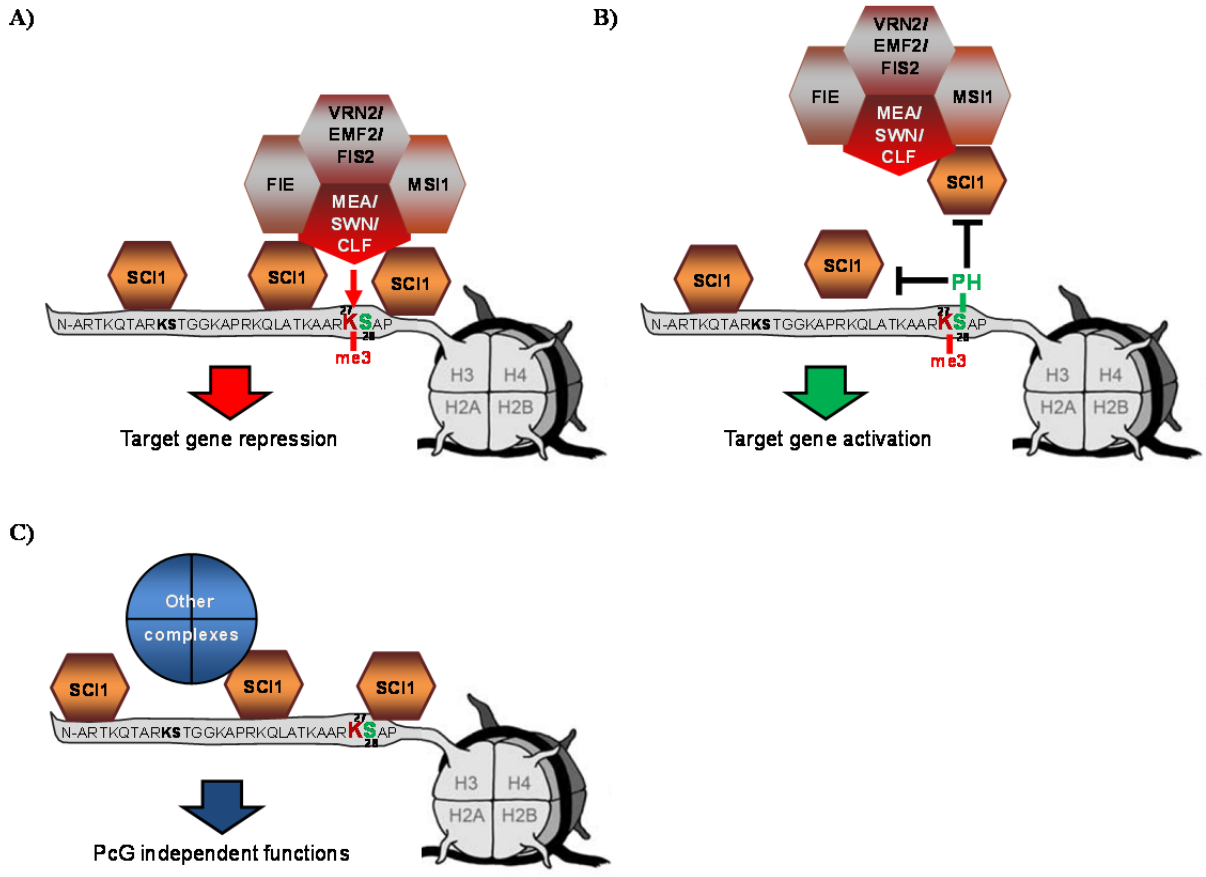


Figure 4.2: Model for SCI1 function
(A)The SCI1 protein (orange hexagon) binds to chromatin devoid of the H3S28pho mark. Because of its interaction with CLF, SWN, MEA, and MSI1, SCI1 might facilitate PRC2 binding to its target genes. This leads to trimethylation of H3K27 and subsequent target gene repression (B) In case of H3S28 phosphorylation, SCI1 is not able to bind the chromatin, which could cause the detachment of PRC2 from specific target genes, and thus a release of target gene repression. (C) Furthermore SCI1 might facilitate the access of other complexes to target chromatin e.g. through interaction with the autonomous pathway protein FVE/MSI4, to regulate chromatin independently of PcG proteins. The nucleosomes and the N-terminal sequence of histone H3 are depicted in grey; important amino acids H3K27/H3K28 are shown in red and green letters, respectively.

Histone phosphorylation was described to be involved in processes like chromosome condensation/segregation, activation of transcription, apoptosis and DNA damage repair as well as in the cross-talk of histone modifications (Houben et al., 2007). In plants phosphorylation at H3S10 and H3S28 was observed during mitosis and meiosis, where it was highly abundant in pericentromeric regions, and proposed to play a role in sister chromatid cohesion, rather than in chromosome condensation (Gernand et al., 2003; Houben and Schubert, 2003). It was shown previously that CLF-GFP is not abundant in dividing root cells (Schubert et al., 2006), and gain of H3S28pho could be the reason for dissociation of PRC2 from the chromatin during mitosis.

In mammalian cells phosphorylation of histone H3 at serine 10 and serine 28 residues in interphase chromatin is mediated by the kinases mitogen- and stress- activated protein kinase 1 and 2 (MSK1 and 2), and is involved in the transcriptional activation of genes, for example in response to stress and mitogen-stimulated signaling (Perez-Cadahia et al., 2009; Sun et al., 2007).

In *Arabidopsis* it was shown that cold, salt and hormone stress enhances the activity of the MSK homologues ATPK6/19, thus stress induced histone phosphorylation could play a role in transcriptional gene regulation in plants (Houben et al., 2007; Mizoguchi et al., 1995).

It is conceivable that specific PcG target genes could be phosphorylated at H3S28 for example in response to stress or other abiotic factors, which could lead to a transient reactivation of specific target genes. A comparison of MSK1 target genes with PcG target genes could identify genes that are silenced by the PcG machinery, and which are possibly activated by phosphorylation in response to stress.

H3S28 phosphorylation was furthermore shown to be involved in differentiation processes by regulation of PcG complexes. The *Drosophila* E(z) histone methyltransferase has two orthologs in humans, Ezh1 and Ezh2, which are differentially expressed. While Ezh2 is abundant in proliferating tissues, the homolog Ezh1 is found in non-proliferative adult organs (Shen et al., 2008; Margueron et al., 2008). The interchange of PRC2-Ezh2 with PRC2-Ezh1 was shown to be required in order to allow the activation of muscle specific genes. This interchange is dependent on Msk1 mediated H3S28 phosphorylation, which causes the displacement of PRC2-Ezh2 from specific target genes. In contrast PRC2-Ezh1 is insensitive toward H3S28pho, and was found to be essential for the proper timing of target gene activation (Stojic et al., 2011).

Also plants possess several PRC2 complexes with different composition, and while no sensitivity toward H3S28pho has been described to date, it is conceivable that apart from

differential expression patterns, also neighboring histone modifications could be involved in the target gene specificity of the different complexes. *In vitro* experiments have shown that the EMF2-PRC2 complex is inhibited by H3K4me3 while the VRN2-PRC2 is insensitive, which could be important for the establishment of bivalent marked loci that exhibit H3K4me3 and H3K27me3 on the same histone tail (Schmitges et al., 2011).

To further analyze the function of SCI1 in respect to H3S28 phosphorylation, the abundance of H3S28pho and H3K27me3H3S28pho in plant nuclei should be further analyzed by immunostainings. It would be important to analyze the histone modifications, especially the abundance of H3K27me3 and H3S28pho, at the nuclear periphery, as SCI1-GFP localizes to this region in *Arabidopsis* root cells. Furthermore it will be essential to study whether also plant PcG complexes are sensitive toward the H3K27me3H3S28pho double mark. The involvement of SCI1, SCI2 and SCI3 in the recognition of H3S28pho *in vivo* should be further analyzed. The substitution of negatively charged amino acids in the PWWP domain of SCI1 might alter its binding to histones and the restriction from binding to phosphorylated H3S28. Such a construct would be instrumental to analyze SCI1 function in relation to the H3S28 phosphorylation mark in *Arabidopsis*.

4.8 Concluding remarks

During the course of this study evidence could be gained that the function of the novel plant specific SCI1 protein is essential for plant development. Since *SCI1* belongs to a gene family of three members in *Arabidopsis* with partly redundant function, the role for SCI1 during plant development cannot easily be approached by mutant analyses. The severe, seedling lethal, triple mutant phenotype nevertheless revealed that the *SCI* genes redundantly function in meristem maintenance, desiccation tolerance and probably in cell proliferation. The analysis of *sci1* single mutant alleles furthermore disclosed a function for SCI1 in flowering time regulation by activation of the floral repressor FLC, and beside *FLC* also other PcG target genes were found to be transcriptionally regulated by SCI1. Interestingly, SCI1 appears to have a function in nuclear size regulation, possibly in concert LINC1, which was identified as interaction partner of SCI1 in an immunoprecipitation experiment.

I could show that SCI1 interacts with all three histone methyltransferase subunits of the plant PRC2 complex, and furthermore MSI1, another component of the plant PRC2, was identified in the immunoprecipitation experiment to interact with SCI1. Hence it was a major goal of this study to identify the function of SCI1 in respect to PcG mediated gene regulation.

Besides the *in vitro* interaction mentioned above, mutant analyses revealed a genetic interaction of *SCI1* with *CLF,* visible through a strong enhancement of the *clf* phenotype by mutation of *SCI1*. SCI1 was discovered to regulate the expression of several PcG target genes, among them *FLC* which causes early flowering of the *sci1* single and *sci1 clf* double mutants, as well as *SEP3*, a major determinant for the *clf* mutant phenotype that is probably responsible for the enhanced leaf curling and reduction of plant size in the double mutants. The analysis of histone modification levels on the chromatin of misexpressed target genes in *sci1* mutants showed no difference compared to the wildtype, suggesting that PRC2 function in *sci1* is not impaired, and therefore SCI1 might not be a general member of the PRC2 core complex. Nevertheless SCI1 could have a function downstream of PRC2, or might act as recruitment-factor of PRC2 to a subset of target genes.

The binding specificity analysis of the SCI1 PWWP domain uncovered further evidence, that SCI1 could be involved in regulation or recruitment of PcG complexes. The SCI1 PWWP domain harbors histone binding capacity, which is only restricted by H3S28 phosphorylation. Studies in mammals showed that simultaneous phosphorylation of H3S28 and tri-methylation of H3K27 restricts PRC2 from binding to target loci. Therefore it is a reasonable hypothesis that SCI1 might have a function in the recognition of this double modification, and could mediate the dissociation of PcG complexes from marked target genes. As there is only very little known about the effects of H3S28pho in plants, and even less about the double mark, SCI1 will be instrumental to study the role of these modifications in plant development.

Sub-cellular localization studies revealed that SCI1 is localized to the nuclear periphery in plants, about which, similar to H3S28pho, is not yet much known in plant research. The interaction of SCI1 with LINC1, and the similar regulation of nuclear size suggest that the localization of SCI1 to the nuclear periphery is important for its function. Furthermore this could imply that also H3S28pho and PcG mediated gene regulation might be involved in nuclear architecture. Enhanced PcG activity at the nuclear periphery in specific cell types has been reported in mammals, and SCI1 might function to recruit PRC2 to specific target genes, devoid of H3S28pho, at the nuclear periphery in *Arabidopsis*. The localization of SCI1 appears to be dependent on the cell type and differs between plant species, as SCI1 was found to localize to nuclear speckles in *N. benthamiana* similar to the PcG proteins VRN2, LHP1/TFL2 and EMF1. It will be essential to reveal the nature of these speckles, and to uncover if they possibly resemble PcG bodies, which have not yet been described in plants to date.

Taken together, the analysis of SCI1 did not only reveal a function in the transcriptional regulation of PcG target genes and flowering time regulation, but also uncovered a connection to nuclear size regulation and the recognition of H3S28 phosphorylation. SCI1 and its homologs SCI2 and SCI3 are likely to be involved in the interpretation of the histone modifications landscape at target gene loci. Furthermore the future analysis of the SCI proteins will possibly yield insights in novel fields of plant research, such as the role of H3S28 phosphorylation in plant development and the characterization of the nuclear periphery.

5 SUMMARY

Previous studies showed that Polycomb group (PcG) proteins that mediate the repressive tri-methylation of H3K27, and which were initially identified in *Drosophila*, are only in part conserved in plants. This might be due to differences in development and lifestyle between plants and animals, and suggests that many plant specific proteins are involved in PcG mediated gene regulation in *Arabidopsis*.

In this study the novel plant specific protein SCI1 was characterized, which is encoded by a gene family with three members in *Arabidopsis*. RNA and reporter gene analyses showed that SCI1 and its two homologs SCI2 and SCI3 are expressed throughout plant development. A high level of redundancy among the three *SCI* genes was revealed by analysis of *sci1 sci2 sci3* triple mutants, which are seedling lethal. Introduction of *SCI1* or *SCI3* transgenes were able to rescue the mutant phenotype.

SCI1 was found to interact with the plant PRC2 members CLF, SWN and MEA. Double mutant analyses revealed a genetic interaction of *SCI1* and *CLF*, as *sci1-1 clf-28* double mutants show a strong enhancement of the *clf* mutant phenotype. SCI1 has a function in PcG target gene regulation, as altered transcript levels of the PcG target genes *FLC*, *SEP3*, *AG* and *FT* were detected in *sci1-1 clf-28* double mutants, which is probably the cause for the enhancement of the phenotype.

FLC expression is strongly repressed in *sci1* single mutants, which results in an early flowering phenotype, and double mutant analyses revealed an epistatic function of FLC and SCI1 in respect to flowering time. Mutant analyses suggest, that SCI1 acts independently of FRI, an activator of *FLC* expression, and of the autonomous pathway component FVE/MSI4, which acts as repressors of *FLC*. Nevertheless it cannot be excluded that SCI1 function in these genetic backgrounds is masked by functional redundancy between SCI1, SCI2 and SCI3.

To further analyze SCI1, putative *in vivo* complex partners were isolated in an immunoprecipitation experiment followed by mass spectrometry. Besides the PRC2 component MSI1, also other proteins involved in chromatin remodeling were found to interact with SCI1. Furthermore the coiled coil protein LINC1 was identified as putative interaction partner of SCI1, and both proteins localize to the nuclear periphery in *Arabidopsis* root cells. *sci1* mutants, similar to *linc1*, show a reduction of nuclear size compared to the wildtype, suggesting that SCI1 might be involved in the regulation of nuclear architecture in *Arabidopsis.*

The amino-terminal PWWP domain of SCI1 was analyzed in respect to histone peptide binding ability. With the exception of phosphorylated H3S28, which inhibits binding of the SCI1 PWWP domain, histone peptides were bound by the SCI1 PWWP domain regardless to other modifications. Since H3S28 is directly adjacent to the PRC2 target site H3K27, SCI1 might be involved in the interplay of these two, probably antagonistic, histone modifications.

In conclusion, SCI1 is involved in the regulation of PcG target genes, especially in the regulation of *FLC*. It is conceivable that SCI1, in concert with SCI2 and SCI3, could restrict PRC2 from binding to chromatin phosphorylated at H3S28, for example during mitosis or in response to abiotic stress. Furthermore SCI1 controls nuclear size, and might have a LINC1 related function in nuclear architecture in *Arabidopsis*.

5.1 ZUSAMMENFASSUNG

Vorausgegangene Studien in *Arabidopsis* über Polycomb-Gruppen (PcG) Proteine und die durch diese vermittelte repressive Histonmodifikation H3K27me3 haben gezeigt, dass nicht alle PcG Proteine aus *Drosophila* homologe Gene in Pflanzen aufweisen. Das kann in den Unterschieden in der Entwicklung und im Lebensstil von Pflanzen und Tieren begründet sein und suggeriert, dass viele pflanzenspezifische Proteine an der PcG vermittelten Genregulation in *Arabidopsis* beteiligt sind.

In dieser Arbeit wurde das bisher nicht beschriebene Protein SCI1 (SWINGER / CURLY LEAF INTERACTOR 1) funktionell charakterisiert. *SCI1* ist Teil einer Genfamilie mit drei Mitgliedern, die während des gesamten Lebenszyklus exprimiert sind. Die Phänotypen der Einzelmutanten der homologen Gene *SCI1*, *SCI2* und *SCI3* unterscheiden sich nur marginal vom Wildtyp, im Gegensatz dazu sind *sci1 sci2 sci3* Mehrfachmutanten im Keimlingsstadium letal. Daraus kann geschlossen werden, dass ein hoher Grad an funktioneller Redundanz zwischen den *SCI* Genen vorliegt.

Im Verlauf dieser Arbeit konnte gezeigt werden, dass SCI1 mit den Histon-Methyltransferase Untereinheiten CLF, SWN und MEA des pflanzlichen POLYCOMB REPRESSIVE COMPLEX 2 (PRC2) interagiert. Zusätzlich zur physikalischen Interaktion ist eine genetische Interaktion zwischen *SCI1* und *CLF* vorhanden, da *sci1 clf* Doppelmutanten eine Verstärkung des *clf*-spezifischen Phänotyps aufweisen. Verantwortlich für diese Verstärkung ist die veränderte Fehlregulation der PcG Zielgene *FLC*, *SEP3*, *AG* und *FT* in *sci1 clf* Doppelmutanten im Vergleich zu *clf* Einzelmutanten.

SCI1 reguliert die Expression des floralen Repressors FLC, und *sci1* Mutanten weisen eine starke Reduktion von *FLC* Transkript auf. Dies resultiert in einem verfrühten Blühzeitpunkt von *sci1* Mutanten im Vergleich zum Wildtyp. Einzelmutanten und Doppelmutanten von *sci1* und *flc* weisen einen identischen Blühzeitpunkt auf, was auf eine epistatische Funktion von *SCI1* und *FLC* hindeutet. Mutantenanalysen haben gezeigt, dass die Regulation der *FLC* Expression durch SCI1 unabhängig ist von FRI, einem Aktivator von *FLC*, sowie von dem Repressor FVE/MSI4. Allerdings könnte der Funktionsverlust von SCI1 auch zum Teil durch Redundanz der homologen Proteine SCI2 und SCI3 ausgeglichen werden.

Um die Funktion von SCI1 weiter zu analysieren, wurden potentielle *in vivo* Komplex-Partner mit Hilfe von Immunoprezipitation von SCI1-GFP und folgender Analyse der Interaktionspartner mit Massenspektroskopie isoliert. Neben der PRC2 Komponente MSI1 konnten noch weitere chromatinmodifizierende Proteine isoliert werden. Zudem wurde mit

LINC1 ein Protein identifiziert, das wie SCI1 eine intrazelluläre Lokalisation an der nucleären Peripherie aufweist.

Durch weiterführende Analysen konnte eine verringerte Zellkern-Größe in *sci1,* vergleichbar mit *linc1* Mutanten, nachgewiesen werden. Dieser Befund deutet darauf hin, dass SCI1, ähnlich wie LINC1, eine Funktion in der strukturellen Organisation des Zellkerns in *Arabidopsis* ausüben könnte.

In dieser Arbeit konnte nachgewiesen werden, dass die PWWP Domäne von SCI1 *in vitro* Histon-Bindungseigenschaften besitzt. Die Bindung an Histon-Peptide wird durch Phosphorylierung an der Position H3S28 inhibiert, während andere Histonmodifikationen keinen Effekt auf die Bindung der SCI1 PWWP Domäne haben. H3S28 grenzt direkt an die durch PcG Proteine tri-methylierte Position H3K27. Da H3S28pho und H3K27me3 antagonistische Auswirkung auf die Genexpression haben, und potentiell die Bindung von Effektor Molekülen konträr beeinflussen, kann angenommen werden, dass SCI1 als Interaktionspartner von PRC2 Komponenten bei der Wechselwirkung dieser Modifikationen eine Rolle spielt.

Zusammenfassend kann man sagen, dass das PWWP-Domänen-Protein SCI1 eine Funktion in der Regulation von *FLC* und anderen PcG Zielgenen aufweist. Es ist denkbar dass SCI1, potentiell redundant mit SCI2 und SCI3, die Bindung von PcG Proteinen an Histone mit der Modifikation H3S28pho verhindert, zum Beispiel während der Zellteilung oder als Reaktion auf abiotischen Stress. Zusätzlich weisen meine Ergebnisse darauf hin, dass SCI1, zusammen mit LINC1, eine Funktion in der strukturellen Organisation des Zellkerns besitzen könnte.

6 LITERATURE

Adrian,J., Farrona,S., Reimer,J.J., Albani,M.C., Coupland,G., and Turck,F. (2010). cis-Regulatory elements and chromatin state coordinately control temporal and spatial expression of FLOWERING LOCUS T in Arabidopsis. Plant Cell *22*, 1425-1440.

Alexandre,C.M. and Hennig,L. (2008). FLC or not FLC: the other side of vernalization. J. Exp. Bot. *59*, 1127-1135.

Alonso,J.M., Stepanova,A.N., Leisse,T.J., Kim,C.J., Chen,H., Shinn,P., Stevenson,D.K., Zimmerman,J., Barajas,P., Cheuk,R., Gadrinab,C., Heller,C., Jeske,A., Koesema,E., Meyers,C.C., Parker,H., Prednis,L., Ansari,Y., Choy,N., Deen,H., Geralt,M., Hazari,N., Hom,E., Karnes,M., Mulholland,C., Ndubaku,R., Schmidt,I., Guzman,P., Aguilar-Henonin,L., Schmid,M., Weigel,D., Carter,D.E., Marchand,T., Risseeuw,E., Brogden,D., Zeko,A., Crosby,W.L., Berry,C.C., and Ecker,J.R. (2003). Genome-wide insertional mutagenesis of Arabidopsis thaliana. Science *301*, 653-657.

Alvarez-Venegas,R. (2010). Regulation by polycomb and trithorax group proteins in Arabidopsis. Arabidopsis. Book. *8*, e0128.

Amasino,R.M. and Michaels,S.D. (2010). The timing of flowering. Plant Physiol *154*, 516-520.

Arib,G. and Akhtar,A. (2011). Multiple facets of nuclear periphery in gene expression control. Curr. Opin. Cell Biol. *23*, 346-353.

Ausin,I., Alonso-Blanco,C., Jarillo,J.A., Ruiz-Garcia,L., and Martinez-Zapater,J.M. (2004). Regulation of flowering time by FVE, a retinoblastoma-associated protein. Nat. Genet. *36*, 162-166.

Ausubel,F.M. (1996). Current protocols in molecular biology. New York: Wiley.

Avery,O.T., Macleod,C.M., and McCarty,M. (1944). Studies on the chemical nature of the substance inducing transformation of pneumococcal types: induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. J. Exp. Med. *79*, 137-158

Avramova,Z. (2009). Evolution and pleiotropy of TRITHORAX function in Arabidopsis. Int. J. Dev. Biol. *53*, 371-381.

Banerjee,T. and Chakravarti,D. (2011). A peek into the complex realm of histone phosphorylation. Mol. Cell Biol. *31*, 4858-4873.

Bannister,A.J., Zegerman,P., Partridge,J.F., Miska,E.A., Thomas,J.O., Allshire,R.C., and Kouzarides,T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature *410*, 120-124.

Bantignies,F. and Cavalli,G. (2011). Polycomb group proteins: repression in 3D. Trends Genet. *27*, 454-464.

Bantignies,F., Roure,V., Comet,I., Leblanc,B., Schuettengruber,B., Bonnet,J., Tixier,V., Mas,A., and Cavalli,G. (2011). Polycomb-dependent regulatory contacts between distant Hox loci in Drosophila. Cell *144*, 214-226.

Barow,M. and Meister,A. (2003). Endopolyploidy in seed plants is differently correlated to systematics, organ, life strategy and genome size. Plant, Cell and Environment *26*, 571-584.

Barrero,J.M., Gonzalez-Bayon,R., del Pozo,J.C., Ponce,M.R., and Micol,J.L. (2007). INCURVATA2 encodes the catalytic subunit of DNA Polymerase alpha and interacts with genes involved in chromatin-mediated cellular memory in Arabidopsis thaliana. Plant Cell *19*, 2822-2838.

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.

Berdasco,M. and Esteller,M. (2010). Aberrant epigenetic landscape in cancer: how cellular identity goes awry. Dev. Cell *19*, 698-711.

Berger,N., Dubreucq,B., Roudier,F., Dubos,C., and Lepiniec,L. (2011). Transcriptional regulation of Arabidopsis LEAFY COTYLEDON2 involves RLE, a cis-element that regulates trimethylation of histone H3 at lysine-27. Plant Cell *23*, 4065-4078.

Bleckmann,A., Weidtkamp-Peters,S., Seidel,C.A., and Simon,R. (2010). Stem cell signaling in Arabidopsis requires CRN to localize CLV2 to the plasma membrane. Plant Physiol *152*, 166-176.

Bratzel,F., Lopez-Torrejon,G., Koch,M., del Pozo,J.C., and Calonje,M. (2010). Keeping cell identity in Arabidopsis requires PRC1 RING-finger homologs that catalyze H2A monoubiquitination. Curr. Biol. *20*, 1853-1859.

Bratzel,F., Yang,C., Angelova,A., Lopez-Torrejon,G., Koch,M., del Pozo,J.C., and Calonje,M. (2012). Regulation of the new Arabidopsis imprinted gene AtBMI1C requires the interplay of different epigenetic mechanisms. Mol. Plant *5*, 260-269.

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.

Brown,J.L., Mucci,D., Whiteley,M., Dirksen,M.L., and Kassis,J.A. (1998). The Drosophila Polycomb group gene pleiohomeotic encodes a DNA binding protein with homology to the transcription factor YY1. Mol. Cell *1*, 1057-1064.

Butenko,Y. and Ohad,N. (2011). Polycomb-group mediated epigenetic mechanisms through plant evolution. Biochim. Biophys. Acta *1809*, 395-406.

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., Tsukada,Y., and Zhang,Y. (2005). Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. Mol. Cell *20*, 845-854.

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.

Cao,R. and Zhang,Y. (2004). The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. Curr. Opin. Genet. Dev. *14*, 155-164.

Carles,C.C. and Fletcher,J.C. (2009). The SAND domain protein ULTRAPETALA1 acts as a trithorax group factor to regulate cell fate in plants. Genes Dev. *23*, 2723-2728.

Chan,C.S., Rastelli,L., and Pirrotta,V. (1994). A Polycomb response element in the Ubx gene that determines an epigenetically inherited state of repression. EMBO J. *13*, 2553-2564.

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.

Chen,D., Molitor,A., Liu,C., and Shen,W.H. (2010). The Arabidopsis PRC1-like ring-finger proteins are necessary for repression of embryonic traits during vegetative growth. Cell Res. *20*, 1332-1344.

Cheutin,T. and Cavalli,G. (2012). Progressive polycomb assembly on H3K27me3 compartments generates polycomb bodies with developmentally regulated motion. PLoS. Genet. *8*, e1002465.

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.

Colon-Carmona,A., You,R., Haimovitch-Gal,T., and Doerner,P. (1999). Technical advance: spatio-temporal analysis of mitotic activity with a labile cyclin-GUS fusion protein. Plant J. *20*, 503-508.

Crevillen,P. and Dean,C. (2011). Regulation of the floral repressor gene FLC: the complexity of transcription in a chromatin context. Curr. Opin. Plant Biol. *14*, 38-44.

Curtis,M.D. and Grossniklaus,U. (2003). A gateway cloning vector set for high-throughput functional analysis of genes in planta. Plant Physiol *133*, 462-469.

Czechowski,T., Stitt,M., Altmann,T., Udvardi,M.K., and Scheible,W.R. (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. Plant Physiol *139*, 5-17.

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 Lucia,F., Crevillen,P., Jones,A.M., Greb,T., and Dean,C. (2008). A PHD-polycomb repressive complex 2 triggers the epigenetic silencing of FLC during vernalization. Proc. Natl. Acad. Sci. U. S. A *105*, 16831-16836.

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.

Dechat,T., Adam,S.A., and Goldman,R.D. (2009). Nuclear lamins and chromatin: when structure meets function. Adv. Enzyme Regul. *49*, 157-166.

Dechat,T., Pfleghaar,K., Sengupta,K., Shimi,T., Shumaker,D.K., Solimando,L., and Goldman,R.D. (2008). Nuclear lamins: major factors in the structural organization and function of the nucleus and chromatin. Genes Dev. *22*, 832-853.

Dellaporta,S.J. and Hicks,J. (1983). A plant DNA minipreparation version II. Plant Molecular Biology Reporter *1*, 19-21.

Deng,W., Ying,H., Helliwell,C.A., Taylor,J.M., Peacock,W.J., and Dennis,E.S. (2011). FLOWERING LOCUS C (FLC) regulates development pathways throughout the life cycle of Arabidopsis. Proc. Natl. Acad. Sci. U. S. A *108*, 6680-6685.

Deng,X., Gu,L., Liu,C., Lu,T., Lu,F., Lu,Z., Cui,P., Pei,Y., Wang,B., Hu,S., and Cao,X. (2010). Arginine methylation mediated by the Arabidopsis homolog of PRMT5 is essential for proper pre-mRNA splicing. Proc. Natl. Acad. Sci. U. S. A *107*, 19114-19119.

Dennis,E.S. and Peacock,W.J. (2007). Epigenetic regulation of flowering. Curr. Opin. Plant Biol. *10*, 520-527.

Dhayalan,A., Rajavelu,A., Rathert,P., Tamas,R., Jurkowska,R.Z., Ragozin,S., and Jeltsch,A. (2010). The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. J. Biol. Chem. *285*, 26114-26120.

Dittmer,T.A. and Richards,E.J. (2008). Role of LINC proteins in plant nuclear morphology. Plant Signal. Behav. *3*, 485-487.

Dittmer,T.A., Stacey,N.J., Sugimoto-Shirasu,K., and Richards,E.J. (2007). LITTLE NUCLEI genes affecting nuclear morphology in Arabidopsis thaliana. Plant Cell *19*, 2793-2803.

Doyle,M.R. and Amasino,R.M. (2009). A single amino acid change in the enhancer of zeste ortholog CURLY LEAF results in vernalization-independent, rapid flowering in Arabidopsis. Plant Physiol *151*, 1688-1697.

Exner,V., Taranto,P., Schonrock,N., Gruissem,W., and Hennig,L. (2006). Chromatin assembly factor CAF-1 is required for cellular differentiation during plant development. Development *133*, 4163-4172.

Farrona,S., Coupland,G., and Turck,F. (2008). The impact of chromatin regulation on the floral transition. Semin. Cell Dev. Biol. *19*, 560-573.

Feng,S., Jacobsen,S.E., and Reik,W. (2010). Epigenetic reprogramming in plant and animal development. Science *330*, 622-627.

Finnegan,E.J. and Dennis,E.S. (2007). Vernalization-induced trimethylation of histone H3 lysine 27 at FLC is not maintained in mitotically quiescent cells. Curr. Biol. *17*, 1978-1983.

Fischle,W., Tseng,B.S., Dormann,H.L., Ueberheide,B.M., Garcia,B.A., Shabanowitz,J., Hunt,D.F., Funabiki,H., and Allis,C.D. (2005). Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. Nature *438*, 1116-1122.

Fischle,W., Wang,Y., and Allis,C.D. (2003a). Binary switches and modification cassettes in histone biology and beyond. Nature *425*, 475-479.

Fischle,W., Wang,Y., and Allis,C.D. (2003b). Histone and chromatin cross-talk. Curr. Opin. Cell Biol. *15*, 172-183.

Fischle,W., Wang,Y., Jacobs,S.A., Kim,Y., Allis,C.D., and Khorasanizadeh,S. (2003c). 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., Kingston,R.E., and Woodcock,C.L. (2004). Chromatin compaction by a polycomb group protein complex. Science *306*, 1574-1577.

Fuchs,J., Demidov,D., Houben,A., and Schubert,I. (2006). Chromosomal histone modification patterns--from conservation to diversity. Trends Plant Sci. *11*, 199-208.

Fujikawa,Y. and Kato,N. (2007). Split luciferase complementation assay to study protein-protein interactions in Arabidopsis protoplasts. Plant J. *52*, 185-195.

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.

Ge,Y.Z., Pu,M.T., Gowher,H., Wu,H.P., Ding,J.P., Jeltsch,A., and Xu,G.L. (2004). Chromatin targeting of de novo DNA methyltransferases by the PWWP domain. J. Biol. Chem. *279*, 25447-25454.

Gehani,S.S., grawal-Singh,S., Dietrich,N., Christophersen,N.S., Helin,K., and Hansen,K. (2010). Polycomb group protein displacement and gene activation through MSK-dependent H3K27me3S28 phosphorylation. Mol. Cell *39*, 886-900.

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.

Gernand,D., Demidov,D., and Houben,A. (2003). The temporal and spatial pattern of histone H3 phosphorylation at serine 28 and serine 10 is similar in plants but differs between mono- and polycentric chromosomes. Cytogenet. Genome Res. *101*, 172-176.

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.

Greb,T., Mylne,J.S., Crevillen,P., Geraldo,N., An,H., Gendall,A.R., and Dean,C. (2007). The PHD finger protein VRN5 functions in the epigenetic silencing of Arabidopsis FLC. Curr. Biol. *17*, 73-78.

Grimaud,C., Negre,N., and Cavalli,G. (2006). From genetics to epigenetics: the tale of Polycomb group and trithorax group genes. Chromosome. Res. *14*, 363-375.

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). Control of reproduction by Polycomb Group complexes in animals and plants. Int. J. Dev. Biol. *49*, 707-716.

Gutierrez,L., Oktaba,K., Scheuermann,J.C., Gambetta,M.C., Ly-Hartig,N., and Muller,J. (2012). The role of the histone H2A ubiquitinase Sce in Polycomb repression. Development *139*, 117-127.

Halfmann,R. and Lindquist,S. (2010). Epigenetics in the extreme: prions and the inheritance of environmentally acquired traits. Science *330*, 629-632.

Hansen,J.C. (2002). Conformational dynamics of the chromatin fiber in solution: determinants, mechanisms, and functions. Annu. Rev. Biophys. Biomol. Struct. *31*, 361-392.

Happel,N. and Doenecke,D. (2009). Histone H1 and its isoforms: contribution to chromatin structure and function. Gene *431*, 1-12.

He,Y., Doyle,M.R., and Amasino,R.M. (2004). PAF1-complex-mediated histone methylation of FLOWERING LOCUS C chromatin is required for the vernalization-responsive, winter-annual habit in Arabidopsis. Genes Dev. *18*, 2774-2784.

He,Y., Michaels,S.D., and Amasino,R.M. (2003). Regulation of flowering time by histone acetylation in Arabidopsis. Science *302*, 1751-1754.

Hekimoglu,B. and Ringrose,L. (2009). Non-coding RNAs in polycomb/trithorax regulation. RNA. Biol. *6*, 129-137.

Hennig,L., Bouveret,R., and Gruissem,W. (2005). MSI1-like proteins: an escort service for chromatin assembly and remodeling complexes. Trends Cell Biol. *15*, 295-302.

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.

Heo,J.B. and Sung,S. (2011). Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. Science *331*, 76-79.

Hirota,T., Lipp,J.J., Toh,B.H., and Peters,J.M. (2005). Histone H3 serine 10 phosphorylation by Aurora B causes HP1 dissociation from heterochromatin. Nature *438*, 1176-1180.

Hofgen,R. and Willmitzer,L. (1988). Storage of competent cells for Agrobacterium transformation. Nucleic Acids Res. *16*, 9877.

Hoppmann,V., Thorstensen,T., Kristiansen,P.E., Veiseth,S.V., Rahman,M.A., Finne,K., Aalen,R.B., and Aasland,R. (2011). The CW domain, a new histone recognition module in chromatin proteins. EMBO J. *30*, 1939-1952.

Horak,J., Grefen,C., Berendzen,K.W., Hahn,A., Stierhof,Y.D., Stadelhofer,B., Stahl,M., Koncz,C., and Harter,K. (2008). The Arabidopsis thaliana response regulator ARR22 is a putative AHP phospho-histidine phosphatase expressed in the chalaza of developing seeds. BMC. Plant Biol. *8*, 77.

Houben,A., Demidov,D., Caperta,A.D., Karimi,R., Agueci,F., and Vlasenko,L. (2007). Phosphorylation of histone H3 in plants--a dynamic affair. Biochim. Biophys. Acta *1769*, 308-315.

Houben,A., Demidov,D., Gernand,D., Meister,A., Leach,C.R., and Schubert,I. (2003). Methylation of histone H3 in euchromatin of plant chromosomes depends on basic nuclear DNA content. Plant J. *33*, 967-973.

Houben,A. and Schubert,I. (2003). DNA and proteins of plant centromeres. Curr. Opin. Plant Biol. *6*, 554-560.

Inoue,H., Nojima,H., and Okayama,H. (1990). High efficiency transformation of Escherichia coli with plasmids. Gene *96*, 23-28.

Jacob,Y., Feng,S., LeBlanc,C.A., Bernatavichute,Y.V., Stroud,H., Cokus,S., Johnson,L.M., Pellegrini,M., Jacobsen,S.E., and Michaels,S.D. (2009). ATXR5 and ATXR6 are H3K27 monomethyltransferases required for chromatin structure and gene silencing. Nat. Struct. Mol. Biol. *16*, 763-768.

Jeon,J. and Kim,J. (2011). FVE, an Arabidopsis homologue of the retinoblastoma-associated protein that regulates flowering time and cold response, binds to chromatin as a large multiprotein complex. Mol. Cells *32*, 227-234.

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.

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.

Kaufmann,K., Muino,J.M., Osteras,M., Farinelli,L., Krajewski,P., and Angenent,G.C. (2010). Chromatin immunoprecipitation (ChIP) of plant transcription factors followed by sequencing (ChIP-SEQ) or hybridization to whole genome arrays (ChIP-CHIP). Nat. Protoc. *5*, 457-472.

Kharchenko,P.V., Alekseyenko,A.A., Schwartz,Y.B., Minoda,A., Riddle,N.C., Ernst,J., Sabo,P.J., Larschan,E., Gorchakov,A.A., Gu,T., Linder-Basso,D., Plachetka,A., Shanower,G., Tolstorukov,M.Y., Luquette,L.J., Xi,R., Jung,Y.L., Park,R.W., Bishop,E.P., Canfield,T.K., Sandstrom,R., Thurman,R.E., MacAlpine,D.M., Stamatoyannopoulos,J.A., Kellis,M., Elgin,S.C., Kuroda,M.I., Pirrotta,V., Karpen,G.H., and Park,P.J. (2011). Comprehensive analysis of the chromatin landscape in Drosophila melanogaster. Nature *471*, 480-485.

Kim,D.H. and Sung,S. (2010). The Plant Homeo Domain finger protein, VIN3-LIKE 2, is necessary for photoperiod-mediated epigenetic regulation of the floral repressor, MAF5. Proc. Natl. Acad. Sci. U. S. A *107*, 17029-17034.

Kim,S., Choi,K., Park,C., Hwang,H.J., and Lee,I. (2006). SUPPRESSOR OF FRIGIDA4, encoding a C2H2-Type zinc finger protein, represses flowering by transcriptional activation of Arabidopsis FLOWERING LOCUS C. Plant Cell *18*, 2985-2998.

Kim,S.Y., He,Y., Jacob,Y., Noh,Y.S., Michaels,S., and Amasino,R. (2005). Establishment of the vernalization-responsive, winter-annual habit in Arabidopsis requires a putative histone H3 methyl transferase. Plant Cell *17*, 3301-3310.

Kim,S.Y. and Michaels,S.D. (2006). SUPPRESSOR OF FRI 4 encodes a nuclear-localized protein that is required for delayed flowering in winter-annual Arabidopsis. Development *133*, 4699-4707.

Kim,S.Y., Zhu,T., and Sung,Z.R. (2010). Epigenetic regulation of gene programs by EMF1 and EMF2 in Arabidopsis. Plant Physiol *152*, 516-528.

Kimura,Y., Kuroda,C., and Masuda,K. (2010). Differential nuclear envelope assembly at the end of mitosis in suspension-cultured Apium graveolens cells. Chromosoma *119*, 195-204.

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. and Muller,J. (2004). The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins. EMBO Rep. *5*, 373-377.

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., 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.

Kooiker,M., Airoldi,C.A., Losa,A., Manzotti,P.S., Finzi,L., Kater,M.M., and Colombo,L. (2005). BASIC PENTACYSTEINE1, a GA binding protein that induces conformational changes in the regulatory region of the homeotic Arabidopsis gene SEEDSTICK. Plant Cell *17*, 722-729.

Koornneef,M., Alonso-Blanco,C., Peeters,A.J., and Soppe,W. (1998). GENETIC CONTROL OF FLOWERING TIME IN ARABIDOPSIS. Annu. Rev. Plant Physiol Plant Mol. Biol. *49*, 345-370.

Koornneef,M., Blankestijn-de Vries,H., Hanhart,C., Soppe,W., and Peeters,T. (1994). The phenotype of some late-flowering mutants is enhanced by a locus on chromosome 5 that is not effective in the Landsberg *erecta* wild-type. The Plant Journal *6(6)*, 911-919.

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.

Kouzarides,T. (2007). Chromatin modifications and their function. Cell *128*, 693-705.

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.

Lachner,M., O'Carroll,D., Rea,S., Mechtler,K., and Jenuwein,T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature *410*, 116-120.

Lafos,M., Kroll,P., Hohenstatt,M.L., Thorpe,F.L., Clarenz,O., and Schubert,D. (2011). Dynamic regulation of H3K27 trimethylation during Arabidopsis differentiation. PLoS. Genet. *7*, e1002040.

Lafos,M. and Schubert,D. (2009). Balance of power--dynamic regulation of chromatin in plant development. Biol. Chem. *390*, 1113-1123.

Lau,P.N. and Cheung,P. (2011). Histone code pathway involving H3 S28 phosphorylation and K27 acetylation activates transcription and antagonizes polycomb silencing. Proc. Natl. Acad. Sci. U. S. A *108*, 2801-2806.

Lee,J.S., Smith,E., and Shilatifard,A. (2010). The language of histone crosstalk. Cell *142*, 682-685.

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,W., Wang,Z., Li,J., Yang,H., Cui,S., Wang,X., and Ma,L. (2011). Overexpression of AtBMI1C, a polycomb group protein gene, accelerates flowering in Arabidopsis. PLoS. One. *6*, e21364.

Libault,M., Tessadori,F., Germann,S., Snijder,B., Fransz,P., and Gaudin,V. (2005). The Arabidopsis LHP1 protein is a component of euchromatin. Planta *222*, 910-925.

Lim,M.H., Kim,J., Kim,Y.S., Chung,K.S., Seo,Y.H., Lee,I., Kim,J., Hong,C.B., Kim,H.J., and Park,C.M. (2004). A new Arabidopsis gene, FLK, encodes an RNA binding protein with K homology motifs and regulates flowering time via FLOWERING LOCUS C. Plant Cell *16*, 731-740.

Lo,S.M. and Francis,N.J. (2010). Inhibition of chromatin remodeling by polycomb group protein posterior sex combs is mechanistically distinct from nucleosome binding. Biochemistry *49*, 9438-9448.

Lopez-Vernaza,M., Yang,S., Muller,R., Thorpe,F., de,L.E., and Goodrich,J. (2012). Antagonistic roles of SEPALLATA3, FT and FLC genes as targets of the polycomb group gene CURLY LEAF. PLoS. One. *7*, e30715.

Lu,F., Cui,X., Zhang,S., Jenuwein,T., and Cao,X. (2011). Arabidopsis REF6 is a histone H3 lysine 27 demethylase. Nat. Genet. *43*, 715-719.

Luger,K., Mader,A.W., Richmond,R.K., Sargent,D.F., and Richmond,T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature *389*, 251-260.

Lukasik,S.M., Cierpicki,T., Borloz,M., Grembecka,J., Everett,A., and Bushweller,J.H. (2006). High resolution structure of the HDGF PWWP domain: a potential DNA binding domain. Protein Sci. *15*, 314-323.

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

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

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.

Mao,Y.S., Zhang,B., and Spector,D.L. (2011). Biogenesis and function of nuclear bodies. Trends Genet. *27*, 295-306.

Margueron,R., Li,G., Sarma,K., Blais,A., Zavadil,J., Woodcock,C.L., Dynlacht,B.D., and Reinberg,D. (2008). Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. Mol. Cell *32*, 503-518.

Margueron,R. and Reinberg,D. (2011). The Polycomb complex PRC2 and its mark in life. Nature *469*, 343-349.

Martin-Trillo,M., Lazaro,A., Poethig,R.S., Gomez-Mena,C., Pineiro,M.A., Martinez-Zapater,J.M., and Jarillo,J.A. (2006). EARLY IN SHORT DAYS 1 (ESD1) encodes ACTIN-RELATED PROTEIN 6 (AtARP6), a putative component of chromatin remodelling complexes that positively regulates FLC accumulation in Arabidopsis. Development *133*, 1241-1252.

Masuda,K., Xu,Z.J., Takahashi,S., Ito,A., Ono,M., Nomura,K., and Inoue,M. (1997). Peripheral framework of carrot cell nucleus contains a novel protein predicted to exhibit a long alpha-helical domain. Exp. Cell Res. *232*, 173-181.

Maurer-Stroh,S., Dickens,N.J., Hughes-Davies,L., Kouzarides,T., Eisenhaber,F., and Ponting,C.P. (2003). The Tudor domain 'Royal Family': Tudor, plant Agenet, Chromo, PWWP and MBT domains. Trends Biochem. Sci. *28*, 69-74.

Meister,R.J., Williams,L.A., Monfared,M.M., Gallagher,T.L., Kraft,E.A., Nelson,C.G., and Gasser,C.S. (2004). Definition and interactions of a positive regulatory element of the Arabidopsis INNER NO OUTER promoter. Plant J. *37*, 426-438.

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., Bezerra,I.C., and Amasino,R.M. (2004). FRIGIDA-related genes are required for the winter-annual habit in Arabidopsis. Proc. Natl. Acad. Sci. U. S. A *101*, 3281-3285.

Michaels,S.D., Ditta,G., Gustafson-Brown,C., Pelaz,S., Yanofsky,M., and Amasino,R.M. (2003). AGL24 acts as a promoter of flowering in Arabidopsis and is positively regulated by vernalization. Plant J. *33*, 867-874.

Mizoguchi,T., Hayashida,N., Yamaguchi-Shinozaki,K., Kamada,H., and Shinozaki,K. (1995). Two genes that encode ribosomal-protein S6 kinase homologs are induced by cold or salinity stress in Arabidopsis thaliana. FEBS Lett. *358*, 199-204.

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.

Muller,J. and Verrijzer,P. (2009). Biochemical mechanisms of gene regulation by polycomb group protein complexes. Curr. Opin. Genet. Dev. *19*, 150-158.

Mylne,J., Greb,T., Lister,C., and Dean,C. (2004). Epigenetic regulation in the control of flowering. Cold Spring Harb. Symp. Quant. Biol. *69*, 457-464.

Mylne,J.S., Barrett,L., Tessadori,F., Mesnage,S., Johnson,L., Bernatavichute,Y.V., Jacobsen,S.E., Fransz,P., and Dean,C. (2006). LHP1, the Arabidopsis homologue of HETEROCHROMATIN PROTEIN1, is required for epigenetic silencing of FLC. Proc. Natl. Acad. Sci. U. S. A *103*, 5012-5017.

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

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

Ooi,S.K., Qiu,C., Bernstein,E., Li,K., Jia,D., Yang,Z., Erdjument-Bromage,H., Tempst,P., Lin,S.P., Allis,C.D., Cheng,X., and Bestor,T.H. (2007). DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature *448*, 714-717.

Papp,B. and Muller,J. (2006). Histone trimethylation and the maintenance of transcriptional ON and OFF states by trxG and PcG proteins. Genes Dev. *20*, 2041-2054.

Pazhouhandeh,M., Molinier,J., Berr,A., and Genschik,P. (2011). MSI4/FVE interacts with CUL4-DDB1 and a PRC2-like complex to control epigenetic regulation of flowering time in Arabidopsis. Proc. Natl. Acad. Sci. U. S. A *108*, 3430-3435.

Perez-Cadahia,B., Drobic,B., and Davie,J.R. (2009). H3 phosphorylation: dual role in mitosis and interphase. Biochem. Cell Biol. *87*, 695-709.

Perry,J. and Zhao,Y. (2003). The CW domain, a structural module shared amongst vertebrates, vertebrate-infecting parasites and higher plants. Trends Biochem. Sci. *28*, 576-580.

Peterson,C.L. and Laniel,M.A. (2004). Histones and histone modifications. Curr. Biol. *14*, R546-R551.

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., Fleury,D., Mylne,J.S., Crevillen,P., Inze,D., Avramova,Z., Dean,C., and Grossniklaus,U. (2008). ARABIDOPSIS TRITHORAX1 dynamically regulates FLOWERING LOCUS C activation via histone 3 lysine 4 trimethylation. Plant Cell *20*, 580-588.

Pirrotta,V. and Li,H.B. (2011). A view of nuclear Polycomb bodies. Curr. Opin. Genet. Dev.

Qiu,C., Sawada,K., Zhang,X., and Cheng,X. (2002). The PWWP domain of mammalian DNA methyltransferase Dnmt3b defines a new family of DNA-binding folds. Nat. Struct. Biol. *9*, 217-224.

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.

Rinn,J.L., Kertesz,M., Wang,J.K., Squazzo,S.L., Xu,X., Brugmann,S.A., Goodnough,L.H., Helms,J.A., Farnham,P.J., Segal,E., and Chang,H.Y. (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell *129*, 1311-1323.

Russo,V.E.A., Martienssen,R.A., and Riggs,A.D. (1996). Epigenetic Mechanisms of Gene Regulation. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Saleh,A., Al-Abdallat,A., Ndamukong,I., Alvarez-Venegas,R., and Avramova,Z. (2007). The Arabidopsis homologs of trithorax (ATX1) and enhancer of zeste (CLF) establish 'bivalent chromatin marks' at the silent AGAMOUS locus. Nucleic Acids Res. *35*, 6290-6296.

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

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.

Sanda,S.L. and Amasino,R.M. (1996). Interaction of FLC and late-flowering mutations in Arabidopsis thaliana. Mol. Gen. Genet. *251*, 69-74.

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

Schatlowski,N., Creasey,K., Goodrich,J., and Schubert,D. (2008). Keeping plants in shape: polycomb-group genes and histone methylation. Semin. Cell Dev. Biol. *19*, 547-553.

Schatlowski,N., Stahl,Y., Hohenstatt,M.L., Goodrich,J., and Schubert,D. (2010). The CURLY LEAF interacting protein BLISTER controls expression of polycomb-group target genes and cellular differentiation of Arabidopsis thaliana. Plant Cell *22*, 2291-2305.

Scheuermann,J.C., de Ayala Alonso,A.G., Oktaba,K., Ly-Hartig,N., McGinty,R.K., Fraterman,S., Wilm,M., Muir,T.W., and Muller,J. (2010). Histone H2A deubiquitinase activity of the Polycomb repressive complex PR-DUB. Nature *465*, 243-247.

Schmitges,F.W., Prusty,A.B., Faty,M., Stutzer,A., Lingaraju,G.M., Aiwazian,J., Sack,R., Hess,D., Li,L., Zhou,S., Bunker,R.D., Wirth,U., Bouwmeester,T., Bauer,A., Ly-Hartig,N., Zhao,K., Chan,H., Gu,J., Gut,H., Fischle,W., Muller,J., and Thoma,N.H. (2011). Histone methylation by PRC2 is inhibited by active chromatin marks. Mol. Cell *42*, 330-341.

Schmitz,R.J., Hong,L., Michaels,S., and Amasino,R.M. (2005). FRIGIDA-ESSENTIAL 1 interacts genetically with FRIGIDA and FRIGIDA-LIKE 1 to promote the winter-annual habit of Arabidopsis thaliana. Development *132*, 5471-5478.

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.

Schubert,I., Dolezel,J., Houben,A., Scherthan,H., and Wanner,G. (1993). Refined examination of plant metaphase chromosome structure at different levels made feasible by new isolation methods. Chromosoma *102*, 96-101.

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

Serebriiskii,I.G. and Golemis,E.A. (2000). Uses of lacZ to study gene function: evaluation of beta-galactosidase assays employed in the yeast two-hybrid system. Anal. Biochem. *285*, 1-15.

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., Conn,A.B., Dennis,E.S., and Peacock,W.J. (2002). Different regulatory regions are required for the vernalization-induced repression of FLOWERING LOCUS C and for the epigenetic maintenance of repression. Plant Cell *14*, 2527-2537.

Shen,X., Liu,Y., Hsu,Y.J., Fujiwara,Y., Kim,J., Mao,X., Yuan,G.C., and Orkin,S.H. (2008). EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. Mol. Cell *32*, 491-502.

Shi,X., Hong,T., Walter,K.L., Ewalt,M., Michishita,E., Hung,T., Carney,D., Pena,P., Lan,F., Kaadige,M.R., Lacoste,N., Cayrou,C., Davrazou,F., Saha,A., Cairns,B.R., Ayer,D.E., Kutateladze,T.G., Shi,Y., Cote,J., Chua,K.F., and Gozani,O. (2006). ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. Nature *442*, 96-99.

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

Simon,J., Chiang,A., Bender,W., Shimell,M.J., and O'Connor,M. (1993). Elements of the Drosophila bithorax complex that mediate repression by Polycomb group products. Dev. Biol. *158*, 131-144.

Simon,J.A. and Tamkun,J.W. (2002). Programming off and on states in chromatin: mechanisms of Polycomb and trithorax group complexes. Curr. Opin. Genet. Dev. *12*, 210-218.

Simpson,G.G. and Dean,C. (2002). Arabidopsis, the Rosetta stone of flowering time? Science *296*, 285-289.

Smaczniak,C., Immink,R.G., Muino,J.M., Blanvillain,R., Busscher,M., Busscher-Lange,J., Dinh,Q.D., Liu,S., Westphal,A.H., Boeren,S., Parcy,F., Xu,L., Carles,C.C., Angenent,G.C., and Kaufmann,K. (2012). Characterization of MADS-domain transcription factor complexes in Arabidopsis flower development. Proc. Natl. Acad. Sci. U. S. A *109*, 1560-1565.

Srikanth,A. and Schmid,M. (2011). Regulation of flowering time: all roads lead to Rome. Cell Mol. Life Sci. *68*, 2013-2037.

Stec,I., Nagl,S.B., van Ommen,G.J., and den Dunnen,J.T. (2000). The PWWP domain: a potential protein-protein interaction domain in nuclear proteins influencing differentiation? FEBS Lett. *473*, 1-5.

Stojic,L., Jasencakova,Z., Prezioso,C., Stutzer,A., Bodega,B., Pasini,D., Klingberg,R., Mozzetta,C., Margueron,R., Puri,P.L., Schwarzer,D., Helin,K., Fischle,W., and Orlando,V. (2011). Chromatin regulated interchange between polycomb repressive complex 2 (PRC2)-Ezh2 and PRC2-Ezh1 complexes controls myogenin activation in skeletal muscle cells. Epigenetics. Chromatin. *4*, 16.

Strahl,B.D. and Allis,C.D. (2000). The language of covalent histone modifications. Nature *403*, 41-45.

Sun,J.M., Chen,H.Y., Espino,P.S., and Davie,J.R. (2007). Phosphorylated serine 28 of histone H3 is associated with destabilized nucleosomes in transcribed chromatin. Nucleic Acids Res. *35*, 6640-6647.

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

Sung,S., He,Y., Eshoo,T.W., Tamada,Y., Johnson,L., Nakahigashi,K., Goto,K., Jacobsen,S.E., and Amasino,R.M. (2006). Epigenetic maintenance of the vernalized state in Arabidopsis thaliana requires LIKE HETEROCHROMATIN PROTEIN 1. Nat. Genet. *38*, 706-710.

Swiezewski,S., Liu,F., Magusin,A., and Dean,C. (2009). Cold-induced silencing by long antisense transcripts of an Arabidopsis Polycomb target. Nature *462*, 799-802.

Tamada,Y., Yun,J.Y., Woo,S.C., and Amasino,R.M. (2009). ARABIDOPSIS TRITHORAX-RELATED7 is required for methylation of lysine 4 of histone H3 and for transcriptional activation of FLOWERING LOCUS C. Plant Cell *21*, 3257-3269.

Till,B.J., Reynolds,S.H., Greene,E.A., Codomo,C.A., Enns,L.C., Johnson,J.E., Burtner,C., Odden,A.R., Young,K., Taylor,N.E., Henikoff,J.G., Comai,L., and Henikoff,S. (2003). Large-scale discovery of induced point mutations with high-throughput TILLING. Genome Res. *13*, 524-530.

Truernit,E., Bauby,H., Dubreucq,B., Grandjean,O., Runions,J., Barthelemy,J., and Palauqui,J.C. (2008). High-resolution whole-mount imaging of three-dimensional tissue organization and gene expression enables the study of Phloem development and structure in Arabidopsis. Plant Cell *20*, 1494-1503.

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.

Turner,B.M. (2002). Cellular memory and the histone code. Cell *111*, 285-291.

vanZanten M., Koini,M.A., Geyer,R., Liu,Y., Brambilla,V., Bartels,D., Koornneef,M., Fransz,P., and Soppe,W.J. (2011). Seed maturation in Arabidopsis thaliana is characterized by nuclear size reduction and increased chromatin condensation. Proc. Natl. Acad. Sci. U. S. A *108*, 20219-20224.

Vezzoli,A., Bonadies,N., Allen,M.D., Freund,S.M., Santiveri,C.M., Kvinlaug,B.T., Huntly,B.J., Gottgens,B., and Bycroft,M. (2010). Molecular basis of histone H3K36me3 recognition by the PWWP domain of Brpf1. Nat. Struct. Mol. Biol. *17*, 617-619.

Wang,J., Kumar,R.M., Biggs,V.J., Lee,H., Chen,Y., Kagey,M.H., Young,R.A., and Abate-Shen,C. (2011). The Msx1 Homeoprotein Recruits Polycomb to the Nuclear Periphery during Development. Dev. Cell *21*, 575-588.

Wang,Y., Fischle,W., Cheung,W., Jacobs,S., Khorasanizadeh,S., and Allis,C.D. (2004). Beyond the double helix: writing and reading the histone code. Novartis. Found. Symp. *259*, 3-17.

Wang,Y., Reddy,B., Thompson,J., Wang,H., Noma,K., Yates,J.R., III, and Jia,S. (2009). Regulation of Set9-mediated H4K20 methylation by a PWWP domain protein. Mol. Cell *33*, 428-437.

Willmann,M.R. and Poethig,R.S. (2011). The effect of the floral repressor FLC on the timing and progression of vegetative phase change in Arabidopsis. Development *138*, 677-685.

Winter,S., Fischle,W., and Seiser,C. (2008). Modulation of 14-3-3 interaction with phosphorylated histone H3 by combinatorial modification patterns. Cell Cycle *7*, 1336-1342.

Wysocka,J., Swigut,T., Xiao,H., Milne,T.A., Kwon,S.Y., Landry,J., Kauer,M., Tackett,A.J., Chait,B.T., Badenhorst,P., Wu,C., and Allis,C.D. (2006). A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. Nature *442*, 86-90.

Xu,L. and Shen,W.H. (2008). Polycomb silencing of KNOX genes confines shoot stem cell niches in Arabidopsis. Curr. Biol. *18*, 1966-1971.

Yan,D., Zhang,Y., Niu,L., Yuan,Y., and Cao,X. (2007). Identification and characterization of two closely related histone H4 arginine 3 methyltransferases in Arabidopsis thaliana. Biochem. J. *408*, 113-121.

Yap,K.L. and Zhou,M.M. (2010). Keeping it in the family: diverse histone recognition by conserved structural folds. Crit Rev. Biochem. Mol. Biol. *45*, 488-505.

Yoo,S.K., Wu,X., Lee,J.S., and Ahn,J.H. (2011). AGAMOUS-LIKE 6 is a floral promoter that negatively regulates the FLC/MAF clade genes and positively regulates FT in Arabidopsis. Plant J. *65*, 62-76.

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.

Zemach,A., Li,Y., Ben-Meir,H., Oliva,M., Mosquna,A., Kiss,V., Avivi,Y., Ohad,N., and Grafi,G. (2006). Different domains control the localization and mobility of LIKE HETEROCHROMATIN PROTEIN1 in Arabidopsis nuclei. Plant Cell *18*, 133-145.

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

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

Zheng,L., Baumann,U., and Reymond,J.L. (2004). An efficient one-step site-directed and site-saturation mutagenesis protocol. Nucleic Acids Res. *32*, e115.

7 APPENDIX

7.1 Abbreviations

aa	amino acid
A. thaliana	*Arabidopsis thaliana*
bp	base pair
°C	degree Celsius
cDNA	copy DNA
Col-0	Ecotype Columbia
C-terminus	Carboxy-terminal end of a protein
dag	days after germination
DAPI	4',6-diamidino-2-phenylindole
DTT	Dithiothreitol
dH2O	destilled water
DNA	Deoxyribonucleic acid
E. coli	*Escherichia coli*
EDTA	Ethylendiamine tetra acidic acid
FRET	Fluorescence resonance energy transfer
g	gram
GAL4-AD	GAL4 activation domain
GAL4-BD	GAL4 DNA binding domain
GFP	Green fluorescing protein
GST	Glutathione-S-Transterase
h	hour
kb	kilobase pair
l	litre
LB	Luria Bertoni broth
LC/MS	light chromatography mass spectrometry
M	molar
mg	milligram
min	minute
ml	millilitre
mM	milli molar
MNE	mean normalized expression
mRNA	messenger RNA
MW	molecular weight
ncRNA	non-coding ribonucleic acid
ng	nanogram
NLS	Nuclear localization signal
nm	nanometer
nmol	nanomolar
N-terminus	amino-terminal end of a protein
PAGE	Polyacrylamid gelectrophoreses
PCR	Polymerase chain reaction
PI	Propidium iodide
PVDF	Polyvinylidenfluoride
RNA	Ribonucleic acid
RT	reverse transcriptase
qRT-PCR	quantitative Real Time-PCR
SDS	Sodiumdodecylsulfate
v/v	volumer per volume
Ws	Ecotype Wassilewskaja
WT	wild type
w/v	weight per volume
Y2H	Yeast two-hybrid
µg	micro gramm
µl	micro litre
µm	micro meter

7.2 List of figures

Figure 1.1: PcG complexes control developmental transitions in *Arabidopsis thaliana* 11

Figure 1.2: The regulation of the switch from vegetative to reproductive development in *Arabidopsis* 15

Figure 3.1: Protein structure of SCI1 and sequence alignment of SCI1, SCI2 and SCI3 39

Figure 3.2: Yeast two hybrid interactions of SCI1 with several PcG proteins 41

Figure 3.3:Intracellular localization of SCI1, SCI1 truncations, CLFΔSET and SWNΔSET . 44

Figure 3.4: Split-luciferase complementation assay with SCI1, CLF and SWN 46

Figure 3.5: Position, phenotype and transcript abundance in *sci1* insertion mutants 50

Figure 3.6: Expression pattern of *SCI1, SCI2* and *SCI3* 51

Figure 3.7: Flowering time analysis of *sci1* and *flc* mutants in LD and SD conditions..... 53

Figure 3.8: Reduced FLC expression in *sci1* mutant background 54

Figure 3.9: Flowering time analysis of *sci1-1 FRI*+ and *sci1-1 fve-3* mutants and expression analysis of *FLC* repressors 56

Figure 3.10: Reduction of root length in *sci1-1* mutants..... 57

Figure 3.11: Genomic loci of *SCI2* and *SCI3* and phenotypes of T-DNA insertion mutants .. 58

Figure 3.12: Phenotypic analysis of *SCI* double and triple mutants 59

Figure 3.13: Scanning electron microscopy analysis of *sci1-1 sci2-2 sci3-2* triple mutants ... 61

Figure 3.14: Phenotypic and flowering time analysis of *sci1 clf-28* double mutants 63

Figure 3.15: Expression analysis of PcG target genes via qRT-PCR 66

Figure 3.16: ChIP analysis of PcG target genes in *sci1-1* mutants 69

Figure 3.17: Global levels of histone modifications in PcG mutants and *sci1-1 clf-28*..... 70

Figure 3.18: SCI1 and its interaction partner LINC1 regulate nuclear size 73

Figure 3.19: Alignment of SCI1 PWWP domain with analyzed domains from mammalian and *S. pombe* proteins 74

Figure 3.20: Histone peptide binding capacity of the SCI1 PWWP domain 76

Figure 4.1: SCI1 regulates PcG target genes in a CLF dependent and independent manner... 90

Figure 4.2: Model for SCI1 function..... 95

7.3 List of tables

Table 2.1: Antibodies and peptides 23

Table 2.2: Oligonucleotides for cloning 24

Table 2.3: Oligonucleotides for genotyping mutant and transgenic plant lines 25

Table 2.4: Oligonucleotides for ChIP analysis 25

Table 2.5: Oligonucleotides for expression analysis 26

Table 2.6: Vectors used for cloning purposes 27

Table 2.7: Created entry clones 27

Table 2.8: Other entry clones used in this study 28

Table 2.9: Created destination vectors 28

Table 2.10: Other vectors used in this study 29

Table 2.11: Microorganisms used in this study 29

Table 2.12: Mutant and transgenic *Arabidopsis thaliana* lines 30

Table 2.13: Excitation and emission spectra 37

Table 3.1: List of selected proteins identified as putative complex partners of SCI1 47

Table 3.2: Analysis of the progeny from *sci1-1 sci2-2/+ sci3-2* plants 60

Table 7.1: List of 50 most enriched putative complex partners of SCI1-GFP 123

Table 7.2: List putative complex partners (# 51-110) of SCI1-GFP 124

7.4 List of putative complex partners of SCI1

#	Ratio	AGI-code	gene names	protein data	uni-prot ID	GO annotation (TAIR)
1	820,30	At2g40930	UBP5	Ubiquitin carboxyl-terminal hydrolase 5	O22207	ubiquitin-specific protease activity
2	82,77	At1g79350	EMB1135		O64516	DNA binding;. embryo development ending in seed dormancy
3	54,23	At4g31880			Q8GLP3	chloroplast; cytosol
4	29,48	At4g15802	HSBP	Heat shock factor binding protein	Q8GW48	response to heat; seed development
5	21,58	At5g03140	SCI1	SCI1	Q0WP05	
6	20,91	At2g17410	ARID3	AT-rich interactive domain (ARID)-containing protein 3	Q940Y3	transcription factor activity
7	18,80	At3g58110			Q9M2K5	unknown biological processes
8	16,23	At5g14170	CHC1	SWI/SNF complex component SNF12 homolog	Q9FMT4	root development
9	15,61	At5g40340		PWWP domain-containing protein	Q9FNE4	nucleolus
10	14,97			DNA-binding protein PcMYB1, putative	Q8LA70	
11	14,85	GFP	GFP	GFP-like fluorescent chromoprotein FP506 (zFP506)	Q9U6Y5	
12	14,16	At1g76510	ARID5	ARID domain-containing protein 5	Q0WNR6	transcription factor activity
13	12,55	At3g43240	ARID4	ARID domain-containing protein 4	Q6NQ79	DNA or RNA binding
14	12,25	At1g18450	ARP4	Actin-related protein 4	Q84M92	chromatin organization
15	12,04	At5g18620		Putative chromatin remodelling complex ATPase chain ISW1	Q94C61	chromatin remodeling
16	11,88	At4g00634		Uncharacterized zinc finger protein	Q2V3L3	nucleic acid binding
17	11,41	At3g06010		Putative transcriptional regulator	Q9SFG5	nucleic acid binding
18	10,96	At5g10710		Centromere protein O	Q8GJP4	chromosome segregation
19	10,33	At2g47620	SWI3A CHB1	SWI/SNF complex subunit SWI3A (ASWI3A)	Q8W4T5	chromatin remodeling
20	9,83	At1g20910	ARID6	ARID domain-containing protein 6	C0SUW7	transcription factor activity
21	8,08	At5g58230	MSI1	WD-40 repeat-containing protein MSI1 (MULTICOPY SUPPRESSOR OF IRA 1)	O22467	chromatin modification
22	7,82	At3g42170	DAYSLEEPER	HAT-like transposase	Q5HT79	DNA or RNA binding
23	7,45	At5g28740		Pre-mRNA-splicing factor SYF1	Q9LKU3	RNA processing
24	6,76	At1g20670		Probable histone H2A.2 (HTA13)	Q9LHQ5	DNA or RNA metabolism |other cellular processes |cell organization and biogenesis
25	6,61	At1g09770	CDC5/MAC1/MYBCD5	Cell division cycle 5-like protein /MOS4-associated complex protein 1 / Atypical R2R3-MYB transcription factor CDC5	P92948	transcription factor activity
26	6,59	At1g28420			Q9SGP0	transcription factor activity
27	6,50	At5g65770	LINC4	LITTLE NUCLEI4 (Putative nuclear matrix constituent protein 1-like protein)	Q9FLH0	chloroplast
28	6,33	At1g32730			Q9LPJ1	
29	6,05	At2g47210		DNA methyltransferase 1-associated protein 1	Q8VZL6	transcription factor activity; DNA or RNA binding
30	5,92	At2g21390		Coatomer subunit alpha-2 (Alpha-coat protein 2)	Q9SJT9	CUL4 RING ubiquitin ligase complex; vesicle-mediated transport
31	5,86	At2g02070		Putative C2H2-type zinc finger protein	Q8RXH0	transcription factor activity
32	5,86	At1g51060		Probable histone H2A.1 (HTA10)	Q9C6S1	DNA or RNA metabolism |other cellular processes |cell organization and biogenesis
33	5,81	At1g68830	STN7	Serine/threonine-protein kinase STN7, chloroplastic (Protein STATE TRANSITION 7)	Q9S713	chloroplast
34	5,36	At5g62690	TUB2	Tubulin beta-2/beta-3 chain	P29512	GTP binding; other biological processes
35	5,33	AtCg00740	rpoA	DNA-directed RNA polymerase subunit alpha (PEP)	P56762	chloroplast
36	5,04	At5g02560		Probable histone H2A.4 (HTA12)	Q9LZ46	DNA or RNA metabolism |other cellular processes |cell organization and biogenesis
37	4,98	At1g79020		Enhancer of polycomb-like protein	Q8GYQ9	transcription factor activity
38	4,94	At5g27670		Probable histone H2A.5 (HTA7)	Q94F49	DNA or RNA metabolism |other cellular processes |cell organization and biogenesis
39	4,83	At1g48610		AT hook motif-containing protein	Q94ALD1	DNA or RNA binding
40	4,57	At5g47690		Sister chromatid cohesion protein PDS5	B3H5K3	nucleus; mitochondrium; unknown biological processes
41	4,55	At4g15900	PRL1/MAC2	Protein pleiotropic regulatory locus 1 (PRL1) / (MAC protein 2)	Q42384	CUL4 RING ubiquitin ligase complex
42	4,46	At1g06190		Rho termination factor	A8MQM6	hydrolase activity |transporter activity
43	4,13	At1g27650		Splicing factor U2af small subunit A (U2 auxiliary factor 35 kDa subunit A)	Q9S709	DNA or RNA binding; photoperiodism, flowering
44	3,95	AtCg00750	rps11	30S ribosomal protein S11, chloroplastic	P56802	plastid small ribosomal subunit
45	3,77	At2g18750		Calmodulin-binding protein (Fragment)	C0SV51	protein binding; unknown biol. Process
46	3,63	At1g67230	LINC1	LITTLE NUCLEI1	Q0WKV7	nucleus
47	3,53	At2g38040	CAC3	Acetyl-coenzyme A carboxylase carboxyltransferase subunit alpha, chloroplastic (ACCase subunit alpha)	Q9LD43	plastid |chloroplast |other cytoplasmic components |other intracellular components
48	3,41	At1g08880		Probable histone H2AXa (HTA5)	O04848	DNA or RNA metabolism |other cellular processes |cell organization and biogenesis
49	3,34	At1g10580		Pre-mRNA-processing factor 17	A4FVN8	nucleotide binding
50	3,33	At4g40030		Histone H3	A8MRL0	DNA or RNA metabolism |other cellular processes |cell organization and biogenesis

Table 7.1: List of 50 most enriched putative complex partners of SCI1-GFP
Putative complex partners are listed according to the ratio and are depicted ranked no., ratio value, AGI code, predicted protein function, uni-prot IDs and gene ontology annotations (GO) retrieved form TAIR and Uni-Prot databases. The bait protein SCI1 is marked in red, GFP in green. Proteins predicted to be involved in chromatin remodeling are shown in blue, histone proteins in yellow and LINC1 in orange. Most proteins are predicted to be nuclear localized.

#	Ratio	AGI-code	gene names	protein data	uni-prot ID	GO annotation (TAIR)
51	3,3256	ATCG00190	rpoB	DNA-directed RNA polymerase subunit beta	P50546	chloroplast/ DNA-directed RNA polymerase activity
52	3,2158	AT3G18790			Q9LS97	cytosol or chloroplast
53	3,2002	ATCG00780	rpl14	50S ribosomal protein L14	P56792	chloroplast, rRNA binding
54	3,1991	AT3G09790	UBQ8	Polyubiquitin 8	Q39256	Ubl conjugation pathway /nucleus /cytoplasm
55	3,1982	AT5G25460			Q94F20	Plant cell wall, plasmodesmi
56	3,0354	AT5G14040	PHT3;1	Belongs to the mitochondrial carrier family	Q9FMU6	mitochondrium / transport
57	2,9957	AT2G23070		Putative casein kinase II catalytic (Alpha) subunit	O64816	protein serine/threonine kinase activity / ATP binding
58	2,937	AT2G34357		Nodulin-like protein	O80784	transmembrane transport
59	2,9056	AT3G08580	AAC1	ADP,ATP carrier protein 1	P31167	mitochondrium / transport
60	2,884	ATCG00160	rps2	30S ribosomal protein S2	P56797	chloroplast
61	2,836	AT1G77180	T14N5.5		O80653	spliceosomal complex / nucleolus
62	2,8155	AT5G19770	TUA3;TUA5	Tubulin alpha-3/alpha-5 chain	P20363	microtubule
63	2,7622	AT5G13490	AAC2	ADP,ATP carrier protein 2	P40941	mitochondrium / transport
64	2,6415	AT5G59870	HTA6	Probable histone H2A7	Q9FJE8	nucleosome assembly / nucleus
65	2,5918	AT5G64260	EXL2		Q9FE06	cell wall
66	2,5854	ATCG00770	rps8	30S ribosomal protein S8	P56801	chloroplast
67	2,5644	AT5G64420		DNA polymerase phi subunit	Q9FGF4	transcription, DNA-dependent
68	2,542	AT5G04990	SUN1		Q9FF75	nuclear envelopme, ER
69	2,5333	ATCG00790	rpl16	50S ribosomal protein L16	P56793	chloroplast, rRNA binding
70	2,4593	AT5G22880	HTB2	Histone H2B.10	Q9FFC0	nucleosome assembly / nucleus
71	2,3638	AT4G17560	AT5G47190	50S ribosomal protein L19-1	Q8W463	chloroplast, rRNA binding
72	2,2968	AT1G09200	HTR2	Histone H3.2	P59226	Chromosome / nucleosome core / nucleus
73	2,255	ATCG00810	rpl22	50S ribosomal protein L22	P56795	chloroplast, rRNA binding
74	2,2525	AT3G57300	INO80	DNA helicase INO80	Q8RXS6	chromatin modification / nucleus
75	2,1633	AT3G57940	UPF0202		Q9M2Q4	ATP binding / nucleus
76	2,1488	AT1G59610	ADL3	Dynamin-2B	Q9LQ55	clathrin-mediated endocytosis
77	2,1449	AT2G28950	ATEXPA6	Expansin-A6	Q38865	plant-type cell wall organization
78	2,0798	ATCG00660	RPL20	50S ribosomal protein L20	P56794	chloroplast, rRNA binding
79	2,04	AT3G05070		Coiled-coil domain-containing protein 12	Q9MAB2	
80	1,9553	AT5G63420	emb2746	EMBRYO DEFECTIVE 2746	Q84W56	regulation of transcription, DNA-dependent
81	1,9512	AT5G51100	FSD2	Superoxide dismutase	Q9LU64	response to UV / chloroplast
82	1,8556	AT1G20960	emb1507		Q8W577	helicase activity
83	1,7726	AT2G46020	BRM	SWI/SNF chromatin remodeling ATPase BRAHMA	Q6EVK6	regulation of gene expression, epigenetic
84	1,7704	AT5G23060	CaS	Calcium sensing receptor	Q9FN48	chloroplast
85	1,7669	AT5G01590			Q7Y1W1	chloroplast envelope
86	1,7023	ATCG00830	rpl2	50S ribosomal protein L2	P56791	chloroplast, translation
87	1,6888	AT4G38780	SUS2	Splicing factor-like protein	Q9T0I6	nuclear mRNA splicing, via spliceosome
88	1,622	AT1G15200			Q9C5G2	
89	1,5335	AT1G23290	RPL27AB	60S ribosomal protein L27a-2	Q9LR33	cytosolic large ribosomal subunit
90	1,49	AT1G78630	emb1473	50S ribosomal protein L13	Q9SYL9	chloroplast, translation
91	1,4418	AT1G04110	SDD1	Subtilisin-like protease SDD1	O64495	regulation of cell proliferation
92	1,4331	AT3G61820		Aspartyl protease-like protein	Q9M356	proteolysis
93	1,3096	AT3G04840		40S ribosomal protein S3a-1	Q9CAV0	cytosolic small ribosomal subunit
94	0,7323	AT5G27120		Probable nucleolar protein 5-1	O04658	ribosome biogenesis / nucleolus
95	0,6416	AT1G29940	NRPA2	DNA-directed RNA polymerase	Q9C8S4	DNA-directed RNA polymerase activity
96	0,6294	AT5G52470	FIB1	rRNA 2'-O-methyltransferase fibrillarin 1	Q9FEF8	nucleolus
97	0,5706	AT1G63810			Q0WVM5	cytosol
98	0,5703	AT3G13940		DNA-directed RNA polymerase	Q9LVK6	DNA binding / DNA-directed RNA polymerase
99	0,5457	AT3G60245		Putative 60S ribosomal protein L37a-1	Q9SRK6	cytosolic large ribosomal subunit
100	0,5438	AT3G05060	NOP5-2	Probable nucleolar protein 5-2	Q9MAB3	ribosome biogenesis / nucleolus
101	0,5438	AT3G11510	RPS14B	40S ribosomal protein S14-2	Q9CAX6	cytosolic small ribosomal subunit
102	0,5265	AT5G08180		H/ACA ribonucleoprotein complex subunit 2-like protein	Q9LEY9	small nucleolar ribonucleoprotein complex
103	0,5055	AT3G49910	RPL26A	60S ribosomal protein L26-1	P51414	cytosolic large ribosomal subunit
104	0,4629	AT3G53740	RPL36A	60S ribosomal protein L36-1	O80929	cytosolic large ribosomal subunit
105	0,457	AT1G60200			Q8VY15	mRNA processing
106	0,3935	AT3G57150	NAP57	H/ACA ribonucleoprotein complex subunit 4	Q9LD90	rRNA processing
107	0,3602	AT2G47610	RPL7AA	60S ribosomal protein L7a-1	P49692	cytosolic large ribosomal subunit
108	0,341	AT3G03920		Putative H/ACA ribonucleoprotein complex subunit 1-like protein 1	Q8VZT0	rRNA processing
109	0,3296	At3g22230		60S ribosomal protein L27-2	Q8LCL3	cytosolic large ribosomal subunit
110	0,0898	AT3G15730	PLDALPHA1	Phospholipase D alpha 1	Q38882	phospholipase D activity

Table 7.2: List of putative complex partners (# 51-110) of SCI1-GFP

Putative complex partners from no. 51 are listed according to the ratio and are depicted ranked no., ratio value, AGI code, predicted protein function, uni-prot IDs and gene ontology annotations (GO) retrieved form TAIR and Uni-Prot databases. Proteins predicted to be involved in chromatin remodeling are shown in blue, histone proteins in yellow

8 EIDESSTATTLICHE ERKLÄRUNG

Eidesstattliche Erklärung zu meiner Dissertation mit dem Titel:

Functional analysis of SCI1 – A PWWP domain protein involved in Polycomb group mediated gene regulation in *Arabidopsis*

Hiermit erkläre ich, dass ich diese Dissertation selbstständig verfasst und keine anderen als die angegebenen Hilfsmittel genutzt habe. Alle wörtlich oder inhaltlich übernommenen Stellen habe ich als solche gekennzeichnet.

Ich versichere außerdem, dass ich die beigefügte Dissertation nur in diesem und keinem anderen Promotionsverfahren eingereicht habe, und dass diesem Promotionsverfahren keine gescheiterten Promotionsverfahren vorausgegangen sind.

Ort, Datum

Unterschrift

9 ACKNOWLEDGEMENTS

Mein besonderer Dank gilt meinem Doktorvater Dr. Daniel Schubert, der es mir ermöglicht hat, an diesem Projekt zu arbeiten. Danke für die vielen hilfreichen Diskussionen, für die Unterstützung und die Begeisterung für mein Projekt. Das alles hat sehr zum Gelingen dieser Arbeit beigetragen.

Bedanken möchte ich mich auch bei Prof. Dr. Rüdiger Simon für die Übernahme des Zweitgutachtens. Danke für die vielen hilfreichen und strukturgebenden Diskussionen während der gesamten Zeit der Doktorarbeit.

Ein herzliches Dankeschön geht an meine Laborkollegen im Schubert-Lab, ganz besonders an Nici, Marcel, Nora, Maurice, Julia und Conny. Danke für die freundschaftliche Arbeitsatmosphäre und die schöne gemeinsame Zeit. Vielen Dank natürlich auch an die ‚Simons': Yvonne, Andrea, Madlen, Helge, Marc, René, Adrian, Julia und Stephan - auf eure Hilfe konnte ich immer zählen, und ohne euch wären die Grillabende nur halb so lustig gewesen.

Ganz besonders danken möchte ich Marcel, Nici, Madlen, Martin und Angelika für das Korrekturlesen dieser Arbeit – danke für eure Unterstützung!

Bei Nici, Andrea und Madlen möchte ich mich für die schöne gemeinsame Zeit in Düsseldorf bedanken. Ihr seid jetzt schon in der großen weiten Welt, aber ich denke, dass unsere Freundschaft auch über die Entfernung standhalten wird. Danke außerdem an Elke, Friederike und Dierk für eure Freundschaft und das Wissen um die Sorgen und Nöte während einer Doktorarbeit – der Austausch mit euch hat mir oft geholfen.

Vielen Dank auch an meine Familie, meine Eltern, die mich immer unterstützt und an mich geglaubt haben. Und natürlich an meine Brüder, die mir mit ihrer Diskussionsbereitschaft über das Leben, das Universum und den ganzen Rest hin und wieder geholfen haben, die Relationen in meinem Leben neu zu ordnen.

Last but not least, vielen Dank Martin, für deine Liebe, Geduld und Lebensfreude, und für die vielen kleinen Aufmunterungen im letzten halben Jahr - es ist schön, dass es dich gibt!