8-2011

Mechanism of Transcriptional Suppression of a Phytochrome A Epiallele in Arabidopsis thaliana

Gulab D. Rangani
University of Arkansas, Fayetteville

Mechanism of Transcriptional Suppression of a Phytochrome A Epiallele in *Arabidopsis thaliana*

Mechanism of Transcriptional Suppression of a Phytochrome A Epiallele in *Arabidopsis thaliana*

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Cell and Molecular Biology

By

Gulabben Rangani
Mumbai University
Bachelor of Science in Microbiology, 1999
Mumbai University
Master of Science in Microbiology, 2001

August 2011
University of Arkansas

ABSTRACT

Cytosine methylation in DNA is an integral part of epigenetically controlled regulatory networks in eukaryotes. Both plants and vertebrates display DNA methylation in the gene coding region; however, its role in gene expression is not well understood. Gene promoter, on the other hand, remains largely unmethylated. Acquisition of methylation in promoter results in transcriptional suppression of the gene. The goal of this research is to study the effect of coding region methylation in gene expression using a unique gene model, *phyA'*. *phyA'* is a transcriptionally suppressed epiallele of the *Arabidopsis thaliana* Phytochrome A gene, which contains methylation in CG sites resident to the exonic region[1]. These exonic methylations confer a strong *phyA* mutant phenotype, characterized by elongated hypocotyls in seedlings grown under continuous far-red light (FRc). Chromatin immunoprecipitation analysis of *phyA'* indicated that the repressive histone mark H3K9me2 is not associated with the *phyA'* locus, and no significant change in the association of euchromatic mark H3K4me3 occurs. Genetic analysis ruled out the involvement of the well-known chromatin modification factors and RNAi genes in *phyA'* silencing, suggesting that *phyA'* silencing is not controlled by the RNA-mediated DNA methylation pathway. To identify novel factors involved in *phyA'* silencing, the forward genetics approach was taken that involved mutagenesis of the *phyA'* epimutant and screening for suppressor mutations. Phenotypic (primary) and molecular (secondary) screening resulted in the isolation of a suppressor line termed as suppressor of *phyA'* silencing 1 (sps1). Genetic and molecular analysis revealed that *sps1* is a second-site mutation that reactivates the *phyA'* locus in spite of *phyA'* hypermethylation. Microarray analysis suggested that targets of SPS1 are the expressed genes. Based on the genetic and

molecular data, it is proposed that the function of SPS1 is to maintain the epigenetic state of the euchromatic loci defined by their methylation state.

[1]Chawla R, Nicholson SJ, Folta KM, and Srivastava V (2007) Transgene-induced silencing of Arabidopsis Phytochrome A gene via exonic methylation. *Plant J.* 52: 1105 - 1118

This dissertation is approved for
Recommendation
to the Graduate Council

Dissertation Director:

_____

Dr. Vibha Srivastava

Dissertation Committee:

_____

Dr. J. Brad Murphy

_____

Dr. Ines Pinto

_____

Dr. Fiona L. Goggin

DISSERTATION DUPLICATION RELEASE

I hereby authorize the University of Arkansas Libraries to duplicate this dissertation when needed for research and/or scholarship.

Agreed_____

Gulab Rangani

Refused_____

Gulab Rangani

DISSERTATION DUPLICATION RELEASE

I hereby authorize the University of Arkansas Libraries to duplicate this dissertation when needed for research and/or scholarship.

Agreed _____
Gulabben Rangani

Refused _____
Gulabben Rangani

TABLE OF CONTENTS

CHAPTER 1

Introduction

Gene expression in eukaryotic cells is regulated by the chromatin structure associated with the gene. This type of regulation supersedes the gene expression determined by the nucleotide sequence, and therefore it is referred to as the *epi*genetic regulation. Epigenetic regulation has an expanding role in a variety of phenomena including the genome defense (silencing of transposons and transgenes), sex chromosome dosage compensation, cellular differentiation and regulation of genes during development (Reviewed by Martiensen and Colot, 2001; Matzke et al., 2000; Matzke and Birchler, 2005; Straub and Becker, 2007; Hsieh and Fischer, 2005). Epigenetic modifications such as methylation of DNA and histones, and acetylation of histones confer specific chromatin structure. Alteration in chromatin structure leads to a change in gene expression, since chromatin structure determines the heterochromatic (silent) and euchromatic (active) state of the genome (Reviewed by Grewal and Jia, 2007).

Heterochromatin mainly consists of repeat DNA sequences. Heterochromatinization inhibits gene transcription leading to gene silencing. Heterochromatin is packaged into tightly arranged nucleosomes with characteristic histone modifications. Methylation of lysine 9 of histone H3 (H3K9) (Richards et al., 2002; Jackson et al., 2004; and Bernatavichute et al., 2008), and deacetylation of histone H4 is often found within heterochromatin (Reviewed by Meyer 2000). In addition, dense methylation in cytosine residues is found in heterochromatic sequences (Reviewed by Tariq and Paszkowski, 2004). In contrast, euchromatin is the gene rich portion of the

genome that is more accessible to gene transcription machinery. Euchromatin is less condensed; and enriched in acetylated histones H3 and H4 (Reviewed by Meyer 2000) as well as methylated histone H3 (H3K4) (Richards et al., 2002; and Zhang et al., 2009).

1.1 DNA Methylation

DNA methylation is an evolutionarily ancient process. In both plants and mammals, hypermethylation of DNA generally correlates with suppression of transcription. DNA methylation is carried out by an enzymatic transfer of a methyl group from the universal methyl donor, S-adenosine-L-methionine (S-AdoMet), onto the cytosine nucleotide located in specific sequence context in DNA. This enzymatic transfer of methyl group is carried out by DNA methyltransferases (DNMTs). Unlike in mammals, where methylation in genome occurs exclusively in CG (symmetric) dinucleotides (Ehrlich et al., 1982), cytosine methylation in plants occurs in both symmetric (CG and CHG) and asymmetric (CHH) sites, where H is A, C or T (Reviewed by Henderson and Jacobsen, 2007).

The understanding of epigenetic gene regulation in plants has been derived by discoveries in model plant, *Arabidopsis thaliana*. *A. thaliana* serves as an ideal model for investigating the role of methylation in gene expression because it provides methodological approaches that are difficult in mammals such as rapid life cycles, easy crosses and mutagenesis. Also, unlike mutations of mammalian DNMTs that cause abnormal development and lethality to embryos, Arabidopsis DNMT mutants have been studied by scoring for morphological defects, without lethal consequences. Many viable Arabidopsis mutants have been generated to study the mechanism of gene silencing

mediated by DNA methylation. For example, the *met1* mutation that eliminates methyl-transferase I activity results in the erasure of genomic CG methylation. Thus, *met1* mutants are extremely useful for studying role of CG methylation in genomes.

Molecular function of DNA methylation in plants was proposed as a defense system to protect genome integrity against endogenous selfish DNA elements (transposable elements or TEs), because most of the methylation in the genome is found in these areas. In Arabidopsis, dense methylation is observed in transposons, retrotransposons, rRNA genes and centromeric repeats in all C contexts, which are normally silent (Reviewed by Matzke et al., 2000; Martiensen and Colot, 2001).

1.2 Plant DNA Methyltransferases

There are at least three types of DNA methyltransferases in plants, METHYLTRANSFERASE 1 (MET1), CHROMOMETHYLASE 3 (CMT3), and DOMAIN REARRANGED METHYLTRANSFEASE (DRMs) responsible for cytosine methylation (Reviewed by Chan et al., 2005).

MET1, a homolog of the mammalian DNMT1, was first isolated from *Arabidopsis thaliana*, and it maintains CG methylation. It is a predominant methyltransferase, and it is responsible for maintenance of global methylation patterns on DNA. *MET1* has no functional redundancy with any other enzyme (Kankel et al., 2003; Saze et al., 2003). It recognizes hemimethylated DNA upon DNA replication, and preferentially methylates unmethylated cytosine residue in CpGs context on newly synthesized strand. Thus, MET1 guarantees the transfer of methylation pattern to next

generation during cell division, which is essential for epigenetic inheritance. Reduction of CG methylation in *met1* mutant was found to be associated with a collection of phenotypic and developmental abnormalities (Kankel et al., 2003; Saze et al., 2003), indicating significant role of MET1 in plant development. MET1 is also required during gametogenesis (Saze et al., 2003) as demonstrated using the following genetic approach: Heterozygous met1-3 (MET1/met1) line was crossed with a transgenic line containing hypermethylated and transcriptionally silenced ß-glucuronidase (GUS) transgene (*MET1/MET1::GUS/GUS*). F1 plants, hemizygous for the *GUS* locus and *met1* locus (*MET1/met1::GUS/-*) did not show release of silencing of *GUS*. But when these F1 plants were backcrossed to the wild type (MET1/MET1), among the heterozygous progeny (*MET1/met1::GUS/-)*, 75% of maternally and 42% of paternally transmitted *GUS* loci were reactivated. Activation of GUS locus through backcrosses indicates that transmission of GUS locus through a *met1* gamete is necessary for the release of gene silencing, indicating a gametopytic role for MET1. MET1 is required for the maintenance of epigenetic information during transition from one generation to the next generation. Difference in the frequency of reactivation of GUS from different parental gametes further confirmed the role of MET1 in gametogenesis (Saze et al., 2003).

A second class of methyltransferases is CMT family. Lindroth et al., (2001) found CMT3 as the main enzyme to be responsible for the maintenance of cytosine methylation at CHG sites by analyzing floral development gene SUPERMAN (*SUP*). *SUP* is densely hypermethylated at non-CG sites (CHG and CHH) (Jacobsen and Meyerowitz 1997). In *cmt3* mutant, *SUP* becomes activated upon losing methylation specifically at CHG sites.

In addition, decreased CHG methylation at centromeric 180 bp repeats, transposable element (Ta3), and retrotransposons (Athila sequences, and Copia-like retrotransposon) was observed. However, *FWA* gene, promoter of which is predominantly methylated at CG sites, was not activated in *cmt3* mutant lines indicating the role of CMT3 in CHG methylation. CMT3 is unique to plant kingdom; no homolog of it has been found in other species. A second gene, KRYPTONITE (KYP), that is required for CHG methylation, encodes a histone methyltransferase protein (Jackson et al., 2002). KYP specifically methylates lysine 9 of histone H3. Cooperative activity of CMT3 and KYP is required to maintain CHG methylation suggesting a self-reinforcing loop mechanism between histone and DNA methylation. The CMT3 chromodomain binds to methylated histone (K9 and K27) (Lindroth et al., 2004), and SRA domain of KYP binds directly to methylated DNA (Johnson et al., 2007).

The third class of methyltransferase genes is composed of two members of DRM family, DRM1 and DRM2. DRM2, the major protein is a homolog of the mammalian DNMT3. *drm1drm2* double mutation did not affect symmetric CG methylation pattern, but showed significant reduction in asymmetric CHH methylations indicating that DRM1 and DRM2 are required for *de novo* establishment of methylation rather than the maintenance of methylation in symmetric sites (Cao and Jacobsen 2002). Role of DRM2 was identified by analyzing *FWA* gene that is silenced in WT plants due to hypermethylation. FWA expression results in late flowering phenotype. By *Agrobacterium*-mediated transformation, Arabidopsis plants were transformed with a copy of *FWA* transgene, where transgene got transcriptionally silenced by *de novo*

methylation. However, when *drm2* mutants were transformed, the *FWA* transgene remained active because, the *de novo* methylation was blocked in *drm2* mutants. The resulting transformants displayed a late flowering phenotype indicative of FWA activity (Cao and Jacobsen 2002; Cao et al., 2003). Null *cmt3* and *drm1drm2* mutants display normal growth and development, even over multiple generations of inbreeding, but *drm1drm2cmt3* triple mutant shows phenotypic aberrations indicating that most of the non-CG methylation is maintained redundantly by DRMs and plant-specific CMT3 (Cao et al., 2003).

1.3 Maintenance of DNA Methylation and *De Novo* DNA Methylation

Once the methylation is carried out at CG and CHG sites, it is maintained by maintenance enzymes MET1 and CMT3 by semi conservative mode of DNA replication; however, methylation at asymmetric sites cannot be maintained by the same mechanism. A persistent RNA signal is required for *de novo* methylation carried out by DRM2 (Reviewed by Chan et al., 2005). Therefore, detection of CHH methylation is indicative of the continuous presence of RNA signal. Most of the non-CG methylation is guided by RNA through a well-known mechanism called RNA directed DNA methylation (RdDM) (Reviewed by Matzke et al., 2001). However, CG methylation at some loci is maintained in the absence of a RNA signal by MET1 (Kanno et al., 2004; and Aufsatz et al., 2004). RdDM is an extensively studied mechanism in the plant kingdom, and so far the only mechanism known for the initiation and establishment of DNA methylation. However, the molecular pathways that control the maintenance of DNA methylation are not clearly understood. RdDM addresses one of the key questions about the sequence specificity of

DNA methylation i.e. specific sequences are targeted by the RNA signal. It was first discovered in viriod infected tobacco plants in 1994 (Wassenegger et al., 1994). RdDM is a nuclear process that is carried out by the componenets of RNAi machinery leading to heterochromatin formation. Small interfering RNA (24 nt siRNA) are the guiding molecules for DNA methylation. DNA with the sequence identity to these guiding siRNA is methylated at its cytosine residues. Recent studies in Arabidopsis have shown that different kinds of 24 nt siRNA (sequence variation) accounts for (a) differences in genomic methylation patterns contributing to natural variation (epigenetic variation), and (b) differential proliferation of transposable elements between closely related ecotypes (Zhai et al., 2008; Hollister et al., 2011). Genomic repetitive sequences such as transposons, retroelements, rRNAs, and centromeric repeats are targets of RdDM. siRNA specific to these targets are called repeat-associated or heterochromatic siRNAs.

DNA-dependent polymerase IV (Pol IV, initially known as Pol IVa) produces long single-stranded RNA transcripts from transposable elements and repetitive sequences (Onodera et al., 2005; Herr et al., 2005; Huettel et al., 2006). Pol IV-generated transcripts move to the nucleolus, where they are converted into dsRNAs by RNA DEPENDENT RNA POLYMERASE 2 (RDR2), and subsequently processed by DICER-LIKE 3 (DCL3) into 24-nt siRNAs (Reviewed by Henderson and Jacobsen, 2007; and Matzke et al., 2009). siRNAs, which confer sequence-specificity are loaded into AGO4 complex to initiate DNA methylation, which in turn imposes transcriptional silencing (Zilberman et al., 2004). Recently, Pol V was found to be associated with intergenic non-coding (IGN) transcripts from several loci (Wierzbicki et al., 2008). The mechanism by

which methylation of the IGN region occurs is not well understood. Nascent Pol V transcripts interact with AGO4 with the help of the adapter protein, SUPRESSOR OF TY INSERTION 5-LIKE (SPT5L), and recruit the silencing machinery to establish DNA methylation (Bies-Etheve et al., 2009; and He et al., 2009). Recently discovered SWI/SNF chromatin remodeling protein, DEFECTIVE IN RNA-DIRECTED DNA METHYLATION 1 (DRD1) is believed to be required for Pol V recruitment to chromatin and for accumulation of IGN transcripts (Kanno et al., 2004; Wierzbicki et al., 2008 and 2009).

DECREASE IN DNA METHYLATION 1 (DDM1) family proteins contribute to the maintenance of DNA methylation. Reduced levels of DNA methylation in all C contexts (Vongs et al., 1993; and Teixeira et al., 2009), transposon reactivation and numerous developmental defects are accumulated over successive generations in *ddm1* mutants (Kakutani et al., 1999; Bartee and Bender 2001). DDM1 gene encodes SWI2/SNF2-like chromatin remodeling factor. *In vitro* analysis showed that purified DDM1 protein redistributes histone octamers on short segment of DNA (Brzeski and Jerzmanowski, 2003). Severe alterations in the distribution of H3K9me2 and decondensation of centromeric heterochromatin was observed in *ddm1* mutants (Gendrel et al., 2002; and Probst et al., 2003), indicating the role of DDM1 in nucleosomes arrangement.

Along with MET1, VARIENT IN METHYLATION (VIM) family proteins play an important role in the maintenance of CG methylation (Woo et al., 2007 and 2008).

Loss of VIM family proteins resulted in the loss of CG methylation in repeat sequences and genic regions. There are five *VIM* genes in *Arabidopsis*, each of which encodes an SRA (SET- and RING-associated) domain methyl cytosine binding protein. VIM1 is a plant homolog of mammalian UHRF1, which is required for the maintenance of CG methylation. Out of five *VIM* genes, *VIM1* was highly expressed in leaves and inflorescence tissue, while *VIM4* and *VIM5* transcripts were not detected. *vim1vim3* mutants displayed more synergetic effect on DNA methylation than *vim1vim2* or either single mutant *vim1*, *vim2* or *vim3*. But *vim1vim2vim3* triple mutant displayed severe hypomethylation compared to *vim1vim3* mutant indicating functional redundancy among VIM proteins (Woo et al., 2008). The precise role of VIM proteins in the maintenance of CG methylation is not well understood. The observations that *vim* mutants phenocopy *met1* mutants indicates VIM proteins are involved in MET1-mediated DNA methylation pathway.

Another class of proteins known as METHYL BINDING DOMAIN (MBD) proteins specifically bind to methylated CG sites. There are twelve MBD proteins in Arabidopsis (Berg et al., 2003; and Grafi et al., 2007). Function of MBD5, MBD6 and MBD7 is suggested as the recruiters of histone deacytelase to methylated DNA (Zemach and Grafi 2003). HISTONE DEACYTELASE 6 (HDA6) is also required for the maintenance of symmetric DNA methylation. In the absence of HDA6, CG and CHG methylation is lost, resulting in the release of transcriptional silencing from several repetitive loci and transgenic loci (Aufsatz et al., 2002; Earley et al., 2010). *hda6* mutants showed decondensation of chromatin around nucleolous organizer regions

(NORs), where rRNA genes are located. Decondensation of genetic loci in the *hda6* mutant is distinct from that seen in *ddm1* mutants. *hda6* mutant displays decondensation of rRNA gene arrays, in contrast to the decondensation of centromeric, pericentromeric repeats and rRNA gene arrays observed in the *ddm1* mutant, indicating the role of HDA6 in structural organization of NORs (Probst et al., 2004). Gene arrays encoding rRNA are arranged in long tandem repeat arrays in hundreds to thousands of copies, but at any one time, only a fraction (less than half of the genes) of rRNA genes are transcribed by POLYMERASE I to provide the sufficient amount of ribosome production. The 18S, 5.8S and 25S genes are clustered in one unit and transcribed together as a 45S RNA gene. Differential transcription of rRNA genes is called dosage control. The effective dosage of their 45S rRNA genes is determined by changes in the density of methylation in the promoter region and specific histone association (Lawrence et al., 2004). Recently a direct role of HDA6 has been demonstrated in rRNA gene dosage control (Earley et al., 2010). 45S rRNA genes were activated in the *hda6* mutant by spurious RNA POLYMERASE II (Pol II)-mediated transcription of intergenic spacer (IGS) located between each unit of the 45S RNA gene. Along with aberrant IGS transcripts, significant hyperacytelation of histone H4, loss of CG and CHG methylations and gain of CHH methylations was observed in the rRNA gene repeats. These observations suggested the function of HDA6 to prevent spurious Pol II transcription in order to control rRNA gene dosage (Earley et al., 2010).

1.4 Global DNA Methylation Pattern in *A. thaliana*

In last few years different groups have attempted to map the distribution of methylation in the entire genome of Arabidopsis using immunoprecipitation and bisulfite sequencing approaches. The latter approach has generated the genome-wide methylation map at single base pair resolution (Zhang et al., 2006; Zilberman et al., 2007; Cokus et al., 2008). Overall levels of CG, CHG and CHH methylation are found to be 24%, 6.7% and 1.7%, respectively in the wild-type Col-0 genome. Genome-wide studies have also revealed an unexpected outcome that approximately one-third (33%) of genes were found to be methylated in their coding region specifically at CG sites (genic CG methylation), without the presence of the corresponding small RNA. Unlike pseudogenes and non-expressed genes, where methylation is evenly distributed across the whole sequence, genic methylation has been found to be distributed away from 5' and 3' end of the genes. This pattern of genic methylation is associated with moderately expressed genes, including the 'housekeeping' genes. Although slight up-regulation of the body-methylated genes was observed in the *met1* mutant background, the precise role of CG methylation in genic region is not well understood.

1.5 Known Epialleles

Most of the understanding regarding the function of DNA methylation comes from the analysis of epigenetically modified alleles or epialleles. Epialleles display altered gene expression as a result of a change in chromatin structure. Altered gene expression may result in a distinct phenotype. The specific epigenetic pattern and transcriptional state of epialleles is heritable. Several naturally occurring or artificially

induced stable epialleles have been found (as a by-product of mutagenesis). *SUPERMAN* (*SUP*) and *FWA* are well studied epialleles that display a distinct phenotype specified by the epigenetic modification. *PAI2* and *BAL2* are naturally occurring epialleles. All of these epialleles undergo transcriptional silencing due to DNA hypermethylation. *SUP* encodes a transcriptional activator required for defining floral whorl boundaries. The *SUP* epiallele (*clark kent* allele) is hypermethylaed in all CG, CHG and CHH sites in the promoter as well as the coding region resulting in the repression of *SUP* transcription (Jacobsen et al., 1997). The WT *FWA* gene consisting of repeat structure is hypermethylated, while the *fwa* epiallele is ectopically expressed due to the hypomethylation in the locus (Soppe et al., 2000). Thus, the *fwa* epiallele is a gain-of-function epimutation that confers a late flowing phenotype. Epiallele of AGAMOUS (AG), a floral regulatory gene, was found in a transgenic line expressing antisense *MET1* transcripts (Jacobsen et al., 2000). In the absence of MET1, the *AG* locus becomes hypermethylated in all C contexts, although the density of methylation is lower than that of the *SUP* locus. Accordingly, the phenotypic stability of AG epiallele was found to be lower than that of *SUP* epiallele in a MET1 antisense line (Jacobsen et al., 2000). Similarly naturally occurring epiallele of Phosphoribosyl Anthranilate Isomerase 2 (PAI2) results from the repeat structures found in the *PAI* locus (Bender and Fink 1995). The *PAI* gene family encodes the enzyme involved in the tryptophan biosynthetic pathway. *Arabidopsis* ecotype Ws has four methylated endogenous *PAI* genes at three unlinked loci: a repeat consisting of *PAI1* and *PAI4*, and singlet *PAI2* and *PAI3* genes. Of these 4 genes, only *PAI1* and *PAI2* encode a functional enzyme, but only *PAI1* is expressed. *PAI1* is transcribed despite dense methylation in both CG and non-CG sites,

whereas endogenous *PAI2* locus remains transcriptionally silent as a result of hypermethylation triggered by *PAI1* and *PAI4* inverted repeat (Bender and Fink 1995; Melquist et al., 1999; and Luff et al., 1999). In addition to these epialleles, some transgene loci also show transcriptional suppression associated with hypermethylation. Genetic analysis of such transgenes has resulted in the identification of DNA hypomethylation mutants, histone modification factors, and other chromatin modification genes. For example, genetic analysis of an Arabidopsis line containing a silent hygromycin-resistance (*HPT*) gene generated *ddm1* and *mom1* mutants (Mittlesten Scheid et al., 1996 and 2003). Thus, natural, induced or transgene epialleles have been used as gene models for understanding epigenetic processes. All of these models have similarities, such as proximity to a repeat structure, DNA methylation in the promoter region, dense chromatin structure, and dependence upon RNA mediated DNA methylation.

Although known epialleles revert back to the WT phenotype at low frequency, they have been successfully used in forward genetic screens to discover factors of the underlying epigenetic pathway. Both *CMT3* and *KYP* were isolated in a suppressor screen for the silenced *SUP* epiallele (Jackson et al., 2002; and Lindroth et al., 2001) and *PAI2* epiallele (Bartee et al., 2001; and Malagnac et al., 2002). Similarly, a genetic screen for mutants that fail to maintain the silent state of the *HPT* transgene resulted in the identification of DDM1 and MOM1 (Mittelsten-Schied et al., 1998; and Amedeo et al., 2000). An Arabidopsis line, consisting of the post-transcriptionally silenced *GUS* gene was mutagenized, and the resulting mutants were screened for the reactivation of

GUS gene. This experiment resulted in the identification of two components of the RNAi pathway: SUPPRESSOR OF GENE SILENCING 3 (SGS3) and SUPPRESSOR OF GENE SILENCING 2 (SGS2) (Elmayan et al., 1998). Similarly, a genetic screen on a complex transgene locus undergoing RNA-directed promoter methylation isolated *DRD1*, a putative SNF2 chromatin remodeling protein (Kanno et al., 2004). Loss of DRD1 reverts silencing induced by the promoter methylation. Consistent with their epigenetic pattern, the genetic screens on *SUP* and *PAI* epialleles isolated factors involved in the RdDM pathway. RdDM is the only pathway known for generating *de novo* methylation, characterized by the presence of dense CHG and CHH methylation. Further, *SUP* and *PAI* genes, as well as transgene models contain methylation throughout the gene sequence covering promoter and coding region. Therefore, the study of known epialleles cannot address (a) the specific role of CG methylation in gene expression, (b) the importance of CG methylation in the exonic region, if any.

1.6 *phyA'* Epiallele

Recently, an epiallele of the Phytochrome A gene (*phyA'*) was isolated. *phyA'* contains an epigenetic pattern similar to that of the body-methylated genes, i.e. occurrence of methylation specifically in CG sites in the coding region (Chawla et al., 2007). *phyA'* is distinct from the previously isolated epiallele as it has following characteristics:

- *phyA'* is hypermethylated at only CG sites in exon 1 and exon 2
- No hypermethylation is detected at CHG or CHH sites in *phyA'*
- No hypermethylation is detected in the promoter or 5'UTR region of *phyA'*

- *phyA'* is transcriptionally suppressed to ~20% of WT level
- Transcriptional silencing of *phyA'* is stable and heritable over many generations. No detectable reversion to the WT phenotype is observed in progeny derived from the selfed or out-crossed parents
- Reduction of *PHYA* transcript in the *phyA'* epimutant confers a distinct phenotype in continuous far-red (FRc) light (2.5 W m^{-2})
- No heterochromatinization is found in the *phyA'* locus as shown by a DNAse I accessibility assay
- In the *met1* mutant background, *phyA'* is up-regulated to a level equivalent to that of WT, restoring the WT phenotype

Thus, *phyA'* is an appropriate model for studying the role of CG methylation in exonic sequences. It is the only known epiallele that displays transcriptional silencing associated strictly with CG methylation without any condensation of the locus. However, the precise mechanism of transcriptional suppression is not known.

Objectives of the Study:

The overall goal of the present work is **to understand the role of DNA methylation, especially CG methylation, in transcriptional regulation using *phyA'* as the gene model.** To achieve this goal the present work is divided into three objectives:

1. To further characterize the *phyA'* locus and the epimutant line and to study the role of already-known epigenetic factors in *phyA'* silencing
2. To identify the factors underlying the epigenetic pathway
3. To characterize a selected *phyA'* suppressor mutation, *sps-1*

CHAPTER 2

Characterization of *phyA'* Epiallele

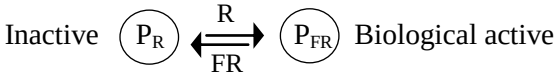
2.1 Introduction

phyA' is an induced epiallele of the *Arabidopsis thaliana* Phytochrome A gene (*PHYA*) found in the phyA-7 and phyA-17 lines. Both phyA-7 and phyA-17 are the non-transgenic progeny of a transgenic line 68lf-3 that carries 3 copies a of *PHYA* transgene construct (Chawla et al., 2007). The *phyA* gene is post-transcriptionally silenced in 68lf-3, a line that accumulates *phyA* siRNAs specific to exon 1 and exon 2 (Nicholson and Srivastava, 2009). 68lf-3 was backcrossed to a WT plant to generate a segregating F2 population. Twenty-two F2 progenies that displayed the *phyA* mutant phenotype were analyzed for the presence of the transgene. Two F2 plants, phyA-7 and phyA-17 showed the absence of transgene while maintaining the mutant phenotype. Lack of the transgene in phyA-7 and phyA-17 lines suggested transgene-induced imprinting of the endogenous *phyA* locus. Sequencing of the *phyA* locus in phyA-17 line showed no change in the DNA sequence compared to the WT (Col-0) *PHYA* locus, indicating the epigenetic modification of the *phyA* gene. The epigenetically modified *phyA* allele is referred to as the *phyA'* epiallele, found in phyA-7 and phyA-17 epimutant lines. DNA methylation analysis of *phyA'* showed hypermethylation only in the symmetric CG sites within the coding region, specifically exon 1 and exon 2. Exonic hypermethylation results in transcriptional silencing of *phyA'*, conferring a strong *phyA* mutant phenotype, characterized by elongated hypocotyls in seedlings grown under continuous Far-Red light (FRc) (Fig. 1).

WT **phyA-211** **phyA-17**

Figure 1: Phenotype of seedlings grown in continuous FR light
WT (Col-0) seedling display short hypocotyls with fully expanded cotyledons, phyA-211 (*phyA* null mutant) seedlings display long hypocotyls and unexpanded cotyledons, and phyA-17 (*phyA'* epimutant line) seedlings display an identical phenotype to that of the null mutant consisting of long hypocotyl and unexpanded cotyledons (Chawla et al., 2007)

PHYA (At1g09570) encodes a Phytochrome A photoreceptor that is the principal mediator of red light (R) and FR induced responses, known as high irradiance responses (HIR). HIR responses include inhibition of hypocotyl elongation, opening of the apical hook, expansion of the cotyledons, and FR light-mediated preconditioned block of greening (Nagatani et al., 1993; and Parks and Quail, 1993). PHYA is also required in very low fluence response (VLFR) like seed germination and gravitational control of hypocotyl growth (Botto et al., 1996). Thus, PHYA plays an important role in germination and seedling establishment (transition from growth in the dark to growth in light) mediated by VLFR and HIR responses. In general, phytochromes exist in two interconvertible forms, P_R and $P_{FR.}$ P_{FR} is the biologically active form, while P_R is the inactive form. Phytochromes are synthesized in P_R form that converts to P_{FR} form (biologically active) upon R absorption. P_{FR} is converted back to P_R form upon FR absorption (as shown below).



PHYA in P_R form is more stable in the cells than PHYA in P_{FR} form, which is photo-labile. PHYA is highly abundant in dark grown seedlings, and upon light exposure, PHYA in P_{FR} form is rapidly degraded. A biological signal generated during P_{FR} to P_R photoconversion was suggested to be responsible for HIR responses (Shinomura et al., 2000).

Homozygous seedlings of the *phyA* null mutant exhibit long hypocotyls and unexpanded cotyledons under FRc light, while heterozygous seedlings exhibit intermediate-length hypocotyls with expanded cotyledons. WT seedlings, on the other hand, exhibit short hypocotyls with expanded cotyledons under FRc light (Fig. 1). Homozygous *phyA* mutants grown under FRc become green upon the transfer to white light, while FRc grown WT and *phyA* heterozygotes bleach and die upon transfer to white light. However, WT seedlings can be rescued in green light on sucrose medium and gradually exposed to white light (Barns et al., 1996). Upon rescue, WT seedlings can grow normally to full maturity. The molecular mechanism of FR blocking of greening is not well understood. Under FRc, protochlorophyllide reductase (POR) activity is suppressed in WT seedlings. As a result, protochlorophyllide (which is a toxic compound) accumulates in FRc grown seedlings, and upon transfer to light causes irreversible photo-bleaching (Barnes et al., 1996). In *phyA* mutants, protochlorophyllide does not accumulate due to a sufficient amount of POR, and become green when transferred to white light (Barnes et al., 1996). WT or heterozygous mutant seedlings exposed to FRc retain the ability to green in the presence of sucrose, because sucrose promotes POR expression by three fold, and provides the required pool of POR enzymes to suppress protochlorophyllide accumulation.

Northern blot analysis and nuclear run-on assays revealed that *phyA'* expression is suppressed to ~80% as compared to the WT level. Phenotypic analysis of progeny derived from selfed plants of epimutant lines (phyA-7 and phyA-17) showed that transcriptional silencing of *phyA'* is maintained in 100% of progeny through successive

generations. Analysis of F_2 progeny derived from an outcross of phyA-17 with Col-0 plants generated a 1:2:1 ratio for long (*phyA*/*phyA*), intermediate (*phyA*/*PHYA*), and short (*PHYA*/*PHYA*) seedlings (Chawla et al., 2007). These experiments indicate that epigenetic modification of *phyA'* is highly stable through meiosis. Demethylation of *phyA'* in a *met1* background resulted in phenotypic reversion of the plants to the WT phenotype accompanied with the restoration of *phyA* expression level to WT level, indicating the pivotal role of the CG methylation in *phyA'* silencing (Chawla et al., 2007). Southern analysis of the revertants revealed the loss of methylation of the exonic CG sites that were uniquely methylated in *phyA'* (Chawla et al., 2007). Thus, transcriptional silencing of *phyA'* is tightly linked with the hypermethylation in its coding region.

Correlation of transcriptional silencing with genic methylation in *phyA'* is surprising. The majority of epialleles including *SUPERMAN*, *AG*, *FWA*, and *PAI2* contain methylation throughout the gene sequence including the promoter and 5' UTR (Jacobsen et al., 1997; Jacobsen et al., 2000; Soppe et al., 2000; Bender and Fink 1995; and Melquist et al., 1999). Therefore, it is impossible to determine whether promoter/enhancer methylation or genic methylation is responsible for the imposed transcriptional silencing. Previously, molecular studies on transgenic lines indicated that transcriptional silencing is mostly correlated with promoter/enhancer methylation (Sidorenko and Peterson, 2001; Stam et al., 2002), while transgenes methylated in their coding regions are transcriptionally active (Reviewed by Fagard and Vaucheret, 2000). These observations suggest that DNA methylation in promoter/enhancer region interferes with the binding of RNA Polymerase II (Pol II) or transcription factors. However, some

recent studies have indicated that transcriptional silencing could also be associated with the methylation of the coding region. Fu et al. (2000) reported the association of transgene coding region (gene-body) methylation at non-symmetrical sites with its transcriptional silencing, although the silenced transgene was not stably inherited through successive generations. Recently, genome-wide methylation mapping studies have shown the presence of cytosine methylation in the transcribed regions of expressed genes (approximately > 20% of expressed genes) (Tran et al., 2005; Zhang et al., 2006; Zilberman et al., 2007; and Aceituno et al., 2008). However, the precise role of gene-body methylation in gene regulation is not well understood. Some studies suggest that gene-body methylation interferes with the progression of Pol II across the coding region, resulting in low transcription rate. Accordingly, three different studies reported lower transcription of densely methylated sequences in plant and mammalian cells (Hohn et al., 1996; Hsieh 1997; Lorincz et al., 2004). Lorincz et al. (2004) further reported the association of heterochromatic histones on the densely methylated (transcriptionally suppressed) gene. Thus, chromatin changes are likely to be associated with transcriptional silencing imposed by gene-body methylation.

Thus, in the present study, fine mapping of DNA methylation in *phyA'* locus was carried out, and the presence of heterochromatic histones was investigated to understand the molecular basis of *phyA'* transcriptional silencing. Several observations indicate that *phyA'* silencing is strictly associated with the maintenance methylation process, and not with *de novo* methylation:

(i) No transgene locus is found in the two epimutant lines, phyA-7 and phyA-17,

(ii) No siRNA associated with the *phyA* locus was found in the two epimutants (Nicholson, unpublished),

(iii) Hypermethylation was found only in CG sites, and not in CHG or CHH contexts. The methylation of all C including CHH sites is characteristic of the *de novo* methylation process.

Thus, the role of RdDM genes were not suspected in *phyA'* silencing. However, the known RdDM or RNAi genes may have additional roles in the epigenetic pathways, which are so far unknown. Therefore, genes in epigenetic pathways were selected to analyze their role in *phyA'* silencing. The selected genes can be categorized into four different groups: RNAi/RdDM, maintenance of DNA methylation, histone methylation, and chromatin modification genes. RNA Directed RNA Polymerase 6 (*RDR6* or *SGS2*), Suppressor of Gene Silencing 3 (*SGS3*), *RDR2*, Argonaute 4 (*AGO4*), and RNA Polymerase IV (*NRPD2A*) are RNAi/RdDM genes. Genes involved in the maintenance of DNA methylation are *CMT3* and *VIM1*. *KYP* is a histone methylation gene is, and *DDM1* and *MOM1* are genes regulating chromatin modifications.

Objectives:

The first objective of the study is to further characterize the *phyA'* locus and the epimutant phyA-17 line and to investigate the role of already-known epigenetic factors in *phyA'* silencing. Following experiments were carried out to address this objective:

1) Bisulfite sequencing of *phyA'* exon 3 and exon 4 to complete the fine mapping of DNA methylation marks in *phyA'* epiallele

2) Chromatin immunoprecipitation (ChIP) analysis to study the association of heterochromatic or euchromatic histone proteins with *phyA'* locus
3) Introduction of mutant alleles of RNAi/RdDM, histone methylation and chromatin modification genes into phyA-17 epimutant line by genetic crosses and analysis of the progeny

2.2 Materials and Methods

2.2a *Bisulfite sequencing of exon 3 and exon 4 in phyA-17 line*

phyA-17 DNA extracted by CTAB buffer (Appendix A-2) was sent to a commercial sequencing service, SeqWright Inc. (Houston, TX) for methylation analysis in exon 3 and exon 4.

2.2b *Chromatin Immunoprecipitation Analysis (ChIP) Assay*

Seeds of phyA-17 and Col-0 (WT) were grown for approximately 3 weeks on soil. Seedlings were used when they were about one inch in height. The protocol for ChIP assay is given in Appendix B. Primers used for amplifying genomic targets and *phyA* gene are given in Appendix C-1 and C-2 respectively. For immunoprecipitation, 10 μg of anti-dimethyl histone H3 (Lys 9; Millipore 07-441), 10 μg of anti-trimethyl histone H3 (Lys 4; Abcam AB8580), and 10 μg of anti-trimethyl histone H3 (Lys 27; Millipore 07-449) was used.

2.2c *Methylation Analysis Of Centromeric Region In phyA-17 Line*

DNA was extracted from pooled seedlings of phyA-17 line, *ddm1* line, and Col-0 using CTAB buffer (Appendix A-2), and digested with *Hpa*II overnight. Southern blot prepared (as given in Appendix A-4) with the digested DNA was hybridized with 180 bp fragment as a radiolabeled probe.

2.2d *Seed stocks used in genetic analysis*

Homozygous or heterozygous mutant lines of each gene were obtained from Arabidopsis Biological Resource Center (ABRC), Ohio State University (Columbus, OH, USA). The list of mutant lines used in the present study is given in Table 1. Each mutant line was crossed with phyA-17 line. Schematic representation of genetic crosses is given in Fig. 2. F1 plants were allowed to self-fertilize to get segregating F2 population. Segregation ratio among F2 population was calculated to investigate the interaction of a particular gene with *phyA'* allele. Approximately 10 - 15 tall F2 plants (homozygous for *phyA'*) were selected and genotyped to find double-mutant plants (homozygous or heterozygous for the mutant allele). Selected F2 plants were allowed to self fertilize to get F3 seeds. Phenotypic analysis under FRc was done on these F3 seedlings to confirm the role of the mutant gene.

Table 1: List of the Selected Genes

Gene	Function	Locus	Allele	Stock Name
VIM1	MET1 accessory facor	AT1G57820	*vim1-2*	CS24737(originally SALK_050903)
CMT3	Methyltransferase mainly responsible for methylation at CHG	AT1G69770	*cmt3-7*	CS6365
DRM1/ DRM2	De novo methyltransferase (methylation in all contexts)	AT5G15380/ AT5G14620	*drm1/drm2-2*	CS6366
KYP	Histone methyltransferase	AT5G13960	*kyp-2*	CS6367
AGO4	Guides methylation at heterochromatic loci	AT2G27040	*ago4-1*	CS6364
RDR2	Involved in RNAi, required for production of endogenous siRNA	AT4G11130	*rdr2-2*	SALK_059661
NRPD2A	Role in RNAi, heterochromatin formation	AT1G63020	*nrpd2A-2*	SALK_046208
RDR6	Generates dsRNA from sense transcripts in PTGS*	AT3G49500	*rdr6-11*	CS24285
SGS3	Involved in PTGS	AT5G23570	*sgs3-11*	CS24289
DDM1	Involved in chromatin remodeling	AT5G66750	*ddm1*	SALK_000590
MOM1	Chromatin modification factor	AT1G08060	*mom1-1*	**Provided by O. Mittelsten Scheid

* Post transcriptional gene silencing
**Dr. Ortrun Mittelsten Scheid (University of Vienna, Austria)

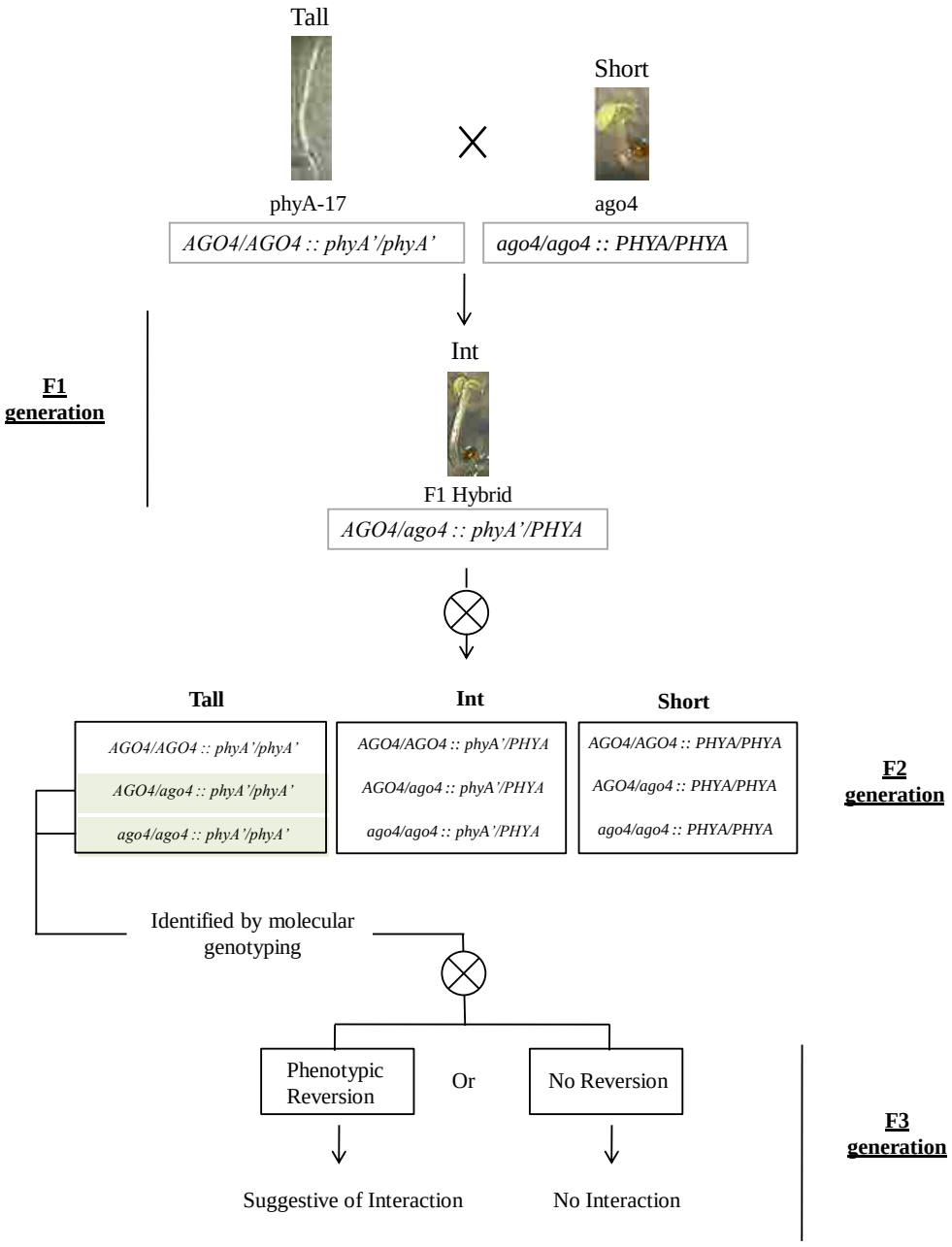


Figure 2: Scheme for genetic crosses. The phyA-17 line was crossed with mutant alleles of selected chromatin modification genes (for example ago4 mutant line). F1 hybrids were self-fertilized to generate segregating F2 population. The F2 individuals containing homozygous *phyA'* locus (tall seedlings) and the mutant allele of the gene under investigation (for example ago4/AGO4 or ago4/ago4) were grown to collected F3 seeds. Phenotypic analysis of F3 seedlings under FRc light was done to find out interaction of the mutant allele with *phyA'* epiallele. Phenotypic reversion in F3 progeny is indicative of the role of the gene in maintaining *phyA'* silencing

2.2e *FRc phenotypic screening and rescuing of revertants*

The following protocol was used for plating and germination of seeds:

1. Seeds of Arabidopsis were surface-sterilized by first soaking in 70% ethanol for 1 min, followed by submerging in 30% bleach and 0.1% SDS for 20 minutes, and rinsing with sterile water twice
2. Seeds were suspended in 0.1% agarose solution and poured on MS media.
3. To induce seed germination, plates were wrapped in aluminum foil and kept for 2 days at 4^{o}C (vernalization) followed by exposure to white light for 14 h and incubation at room temperature for10 h in darkness giving a 24 h light and dark cycle. Plates were then transferred to growth chamber illuminated with continuous far-red light (FRc) (2.5 W m−2) for four days.

After FRc treatment, seedlings were observed in green-safe light (green light filtered through Roscolux™ green filter number 2004). Selected seedlings were transferred to MS media containing 2% sucrose, and kept at RT in dark for 48 h. Plates were then covered with two layers of 3 MM filter papers and exposed to white light. Each layer was removed one by one at 1 h interval. The seedlings were finally exposed to white light till they become green and healthy then transferred to soil.

2.3 Results and Discussion

2.3a *Complete methylation profile of phyA' epiallele*

The methylation profile of the *phyA'* promoter, 5' UTR, exon 1, exon 2 and all introns was already available (Chawla et al., 2007; Nicholson, unpublished). Therefore, bisulfite sequencing of exon 3 and exon 4 was done in the present study to generate the

complete methylation profile of the ~7.0 Kb *phyA'* locus. The data was compared with the published methylation profile of the WT (Col-0) *PHYA* gene using Epi-browser (http://epigenomics.mcdb.ucla.edu/BS-Seq/). This analysis revealed that exon 3 and exon 4 of *phyA'* contain identical methylation marks (CG sites) as the WT allele except for the lack of methylation at three sites within exon 3. Thus, *phyA'* is not hypermethylated in the 3' end of the gene. Representation of the complete methylation profile of *phyA'* compared with Col-0 is given in Fig. 3. In conclusion, hypermethylation in *phyA'* was found only in exon 1 and exon 2, while some demethylations were detected in exon 3.

2.3b *Analysis of histone modification at phyA' locus by the Chromatin Immunoprecipitation (ChIP) assay*

To determine the association of *phyA'* locus with the specific histone proteins that mark active or inactive chromatin, ChIP assays were carried out. Dimethylation of lysine 4 on histone H3 (H3K4me3) is associated with active chromatin (euchromatin), whereas dimethylation of lysine 9 of histone H3 (H3K9me2) is usually associated with inactive chromatin (heterochromatin). Trimethylation of lysine 27 of histone H3 (H3k27me3) is also associated with condensed chromatin; however, it is specifically found on the developmentally regulated genes that turn on or off via chromatin modification (Zhang et al., 2007). Thus, developmentally-regulated genes undergo intermediate level of heterochromatinization characterized by the presence of H3K27me3. Chromatin samples derived from phyA-17 line and Col-0 were immunoprecipitated with antibodies against H3K9me2, H3K4me3, and H3K27me3, the heterochromatic, euchromatic and intermediate heterochromatic marks, respectively. Genomic DNA isolated from the

immunoprecipitated chromatin samples was subjected to PCR using primers corresponding to the *phyA* promoter, 5′ UTR and coding region (Fig. 4). The well-characterized heterochromatin loci (5S rRNA, 180 bp centromeric repeats, AtSN1, Ta2) and euchromatin loci (Actin, PHYB, TUB8, PFK) served as controls (Appendix C).

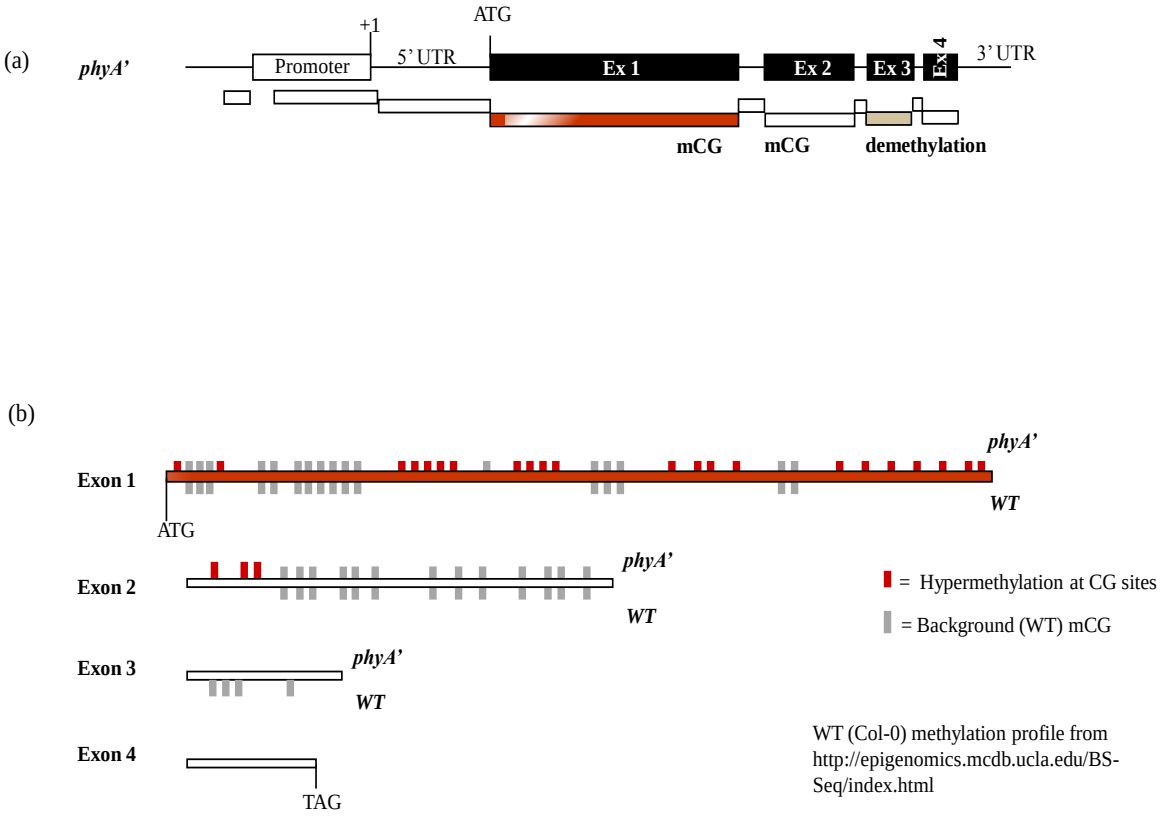


Figure 3: Methylation profile of *phyA'*. (a) Depiction of hypermethylation density (red gradient) in exon 1 and exon2, and demethylation (green gradient) in exon 3 in *phyA'*, (b) Fine mapping of mCG sites in exons based on bisulfite sequencing of *phyA'* (red bar) and wild-type (Col-0) *phyA* gene (gray bar). Red bars represent hypermethylation, and gray bars represent background WT methylations (http://epigenomics.mcdb.ucla.edu/BS-Seq/index.html)

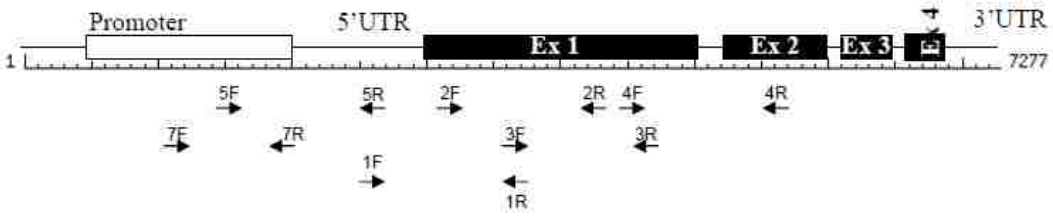


Figure 4: Location of primers used for ChIP analysis across the *phyA*

First, we examined the association of H3K9me2 with *phyA'* locus. As expected the association of H3K9me2 with 5S, 180 bp, Ta2 and AtSN1 was detected, but not with Actin, PHYB, Tub8 and PFK (Fig. 5). However, the binding of H3k9me2 with *phyA'* coding region was also not detected using primers spanning exon 1, exon 2, promoter and 5' UTR. These experiments were repeated 4 times to confirm the findings (Fig. 5). Lack of interaction of H3K9me2 with *phyA'*, especially the exon 1 and exon 2 regions, contradicts the previous report by Chawla et al. (2007), which suggested the weak positive binding of H3K9me2 with the *phyA'* coding region (specifically the hypermethylated region) using ChIP assays. Use of advanced generation seedlings in the present study versus the early generation seedlings in the previous study may have contributed to this discrepancy.

Second, the association of H3K4me3 with *phyA'* was analyzed. ChIP assays with antibodies against H3K4me3 indicated a higher binding of H3K4me3 with the *phyA* locus of phyA-17 line compared to that of Col-0 (Fig. 6). The enrichment of H3K4me3 on phyA-17 locus is in contrast to the other findings, which showed that silenced locus is accompanied by the lower binding of H3K4me3 (Zhang et al., 2009 and Lorincz et al., 2004). Therefore, the significance of abundance of H3K4me3 on phyA-17 locus is not clear. As positive controls, TUB8 and PFK were found to be associated with H3K4me3, whereas the negative control, Ta2 locus, showed much lower association with H3K4me3 (Fig. 6).

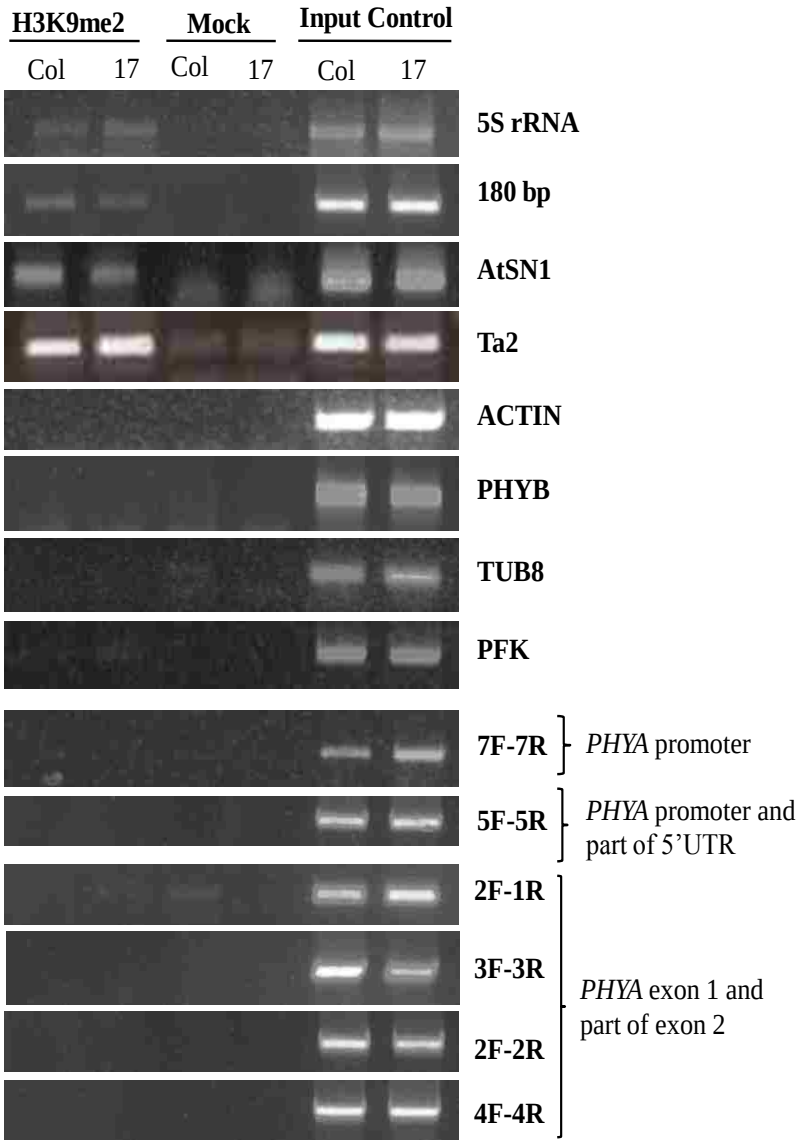


Figure 5: **Chromatin Immunoprecipitation (ChIP) assay using H3K9me2 antibody** Chromatin isolated from Col-0 (WT) and phyA-17 seedlings was subjected to immunoprecipitation with H3K9me2 antibodies, followed by PCR (39 cycles for all reactions) with gene-specific primers (see Fig. 4) along with primers for heterochromatic and euchromatic loci (positive /negative controls). The ChIP assay was also performed without antibody (mock control) and with total DNA from the chromatin sample before immunoprecipitation (input DNA control).

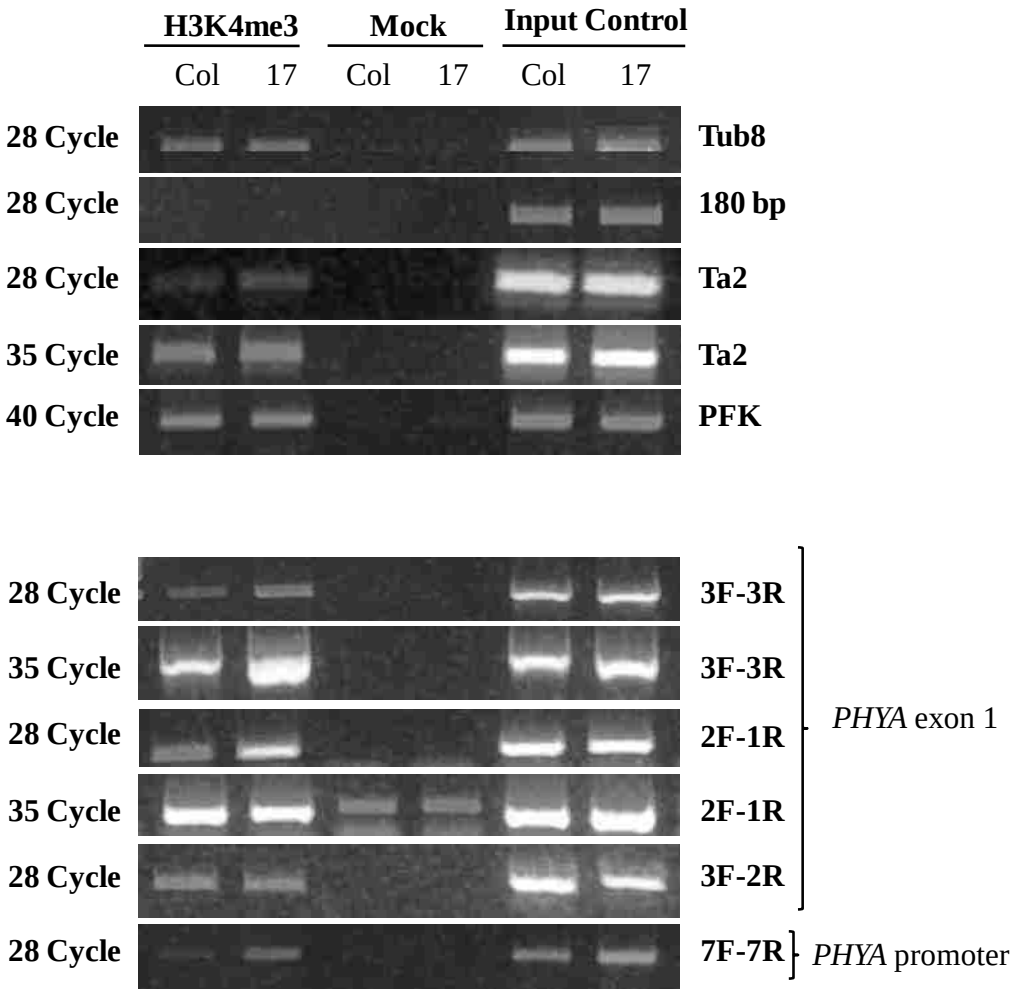


Figure 6: **ChIP assay using H3K4me3 antibody.** Chromatin isolated from Col-0 (WT) and phyA-17 seedlings were immuno-precipitated by H3K4me3 antibodies. Input and no antibody (mock) were also amplified as positive and negative controls, respectively. Immunoprecipitated DNA was amplified by using primers of heterochromatic loci (180 bp, Ta2), euchromatic loci (PFK, Tub8), and phyA gene (see Fig. 4).

Finally, the association of H3K27me3 with *phyA'* was examined, and equal binding of H3K27me3 was found on the promoter region of *phyA* locus in Col-0 and phyA-17 as tested by a single PCR assay (Fig. 7). No further PCR with immune-precipitated chromatin was carried out because the *phyA* locus is not under developmental control. However, a recent study indicated that light mediated induction of the *PHYA* locus involves chromatin modification (Jang et al., 2011). It is possible that H3K27me3 may also be involved in controlling expression of environmentally regulated genes such as light regulated *PHYA* gene.

Together, these results suggest that local chromatin structure of *phyA'* is not subjected to any modification by H3K9me2 and H3K4me3. The specificity of ChIP assay with H3K9me2 and H3K4me3 was verified by binding with known heterochromatic and euchromatic genes, respectively. Lack of the association of H3K9me2 with *phyA'* is consistent with the finding that methylation in *phyA'* is restricted to CG sites. H3K9me2 is critical for transcriptional silencing of heterochromatic region composed of transposons and repeated sequences with dense methylation (Lippman et al., 2004; Bernatavichute et al., 2008). Further, H3K9me2 is correlated with CHG methylation (Bernatavichute et al., 2008), while *phyA'* is strictly maintained by CG methylation. H3K4me3 is generally associated with endogenous genes; however, its presence does not always correlate with active transcription (Alvarez-Venegas et al., 2005; and Zhang et al., 2009). At present, only limited information is available on the histone modifications associated with transcriptionally silent loci located in the euchromatic regions of plant genome (Bernatavichute et al., 2008; Zhang et al., 2007 and 2009).

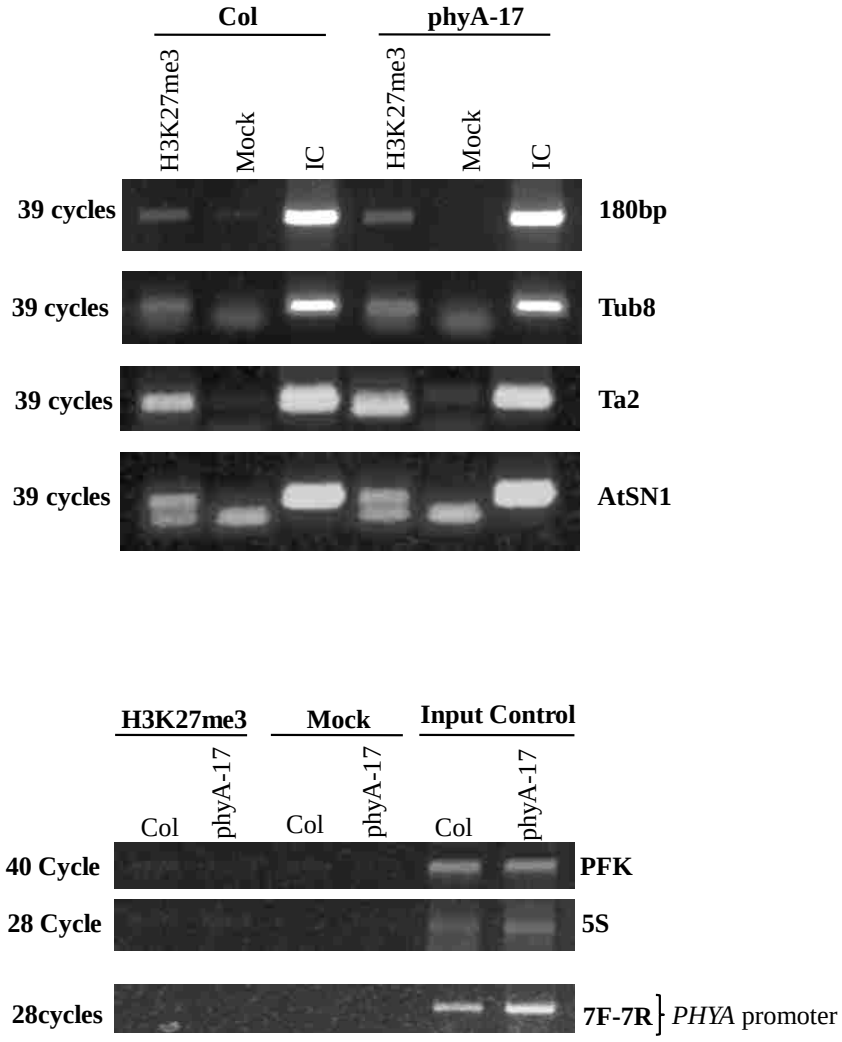


Figure 7: **ChIP assay using H3K27me3.** Chromatin isolated from Col-0 (WT) and phyA-17 seedlings was subjected to immunoprecipitation with H3K27me3 antibodies, followed by PCR with single gene specific primer (see Figure 6) with 180 bp, Ta2, Tub8, AtSN1, PFK and 5S as controls. Input control is chromatin sample before immunoprecipitation, and mock is no antibody control

2.3c *Characterization of the centromeric region in the phyA-17 line*

The phyA-17 line was found to be slightly hypomethylated compared to the WT Col-0 genome. *Hpa*II digested DNA of phyA-17, Col-0 and *ddm1* mutant lines (SALK_000590 was used as a positive control) were hybridized with 180 bp single repeat on a Southern blot. This experiment revealed the presence of a weak ladder of low size DNA fragments compared to WT Col-0 DNA. This is characteristic of demethylation of centromeric repeats in phyA-17 line. The *ddm1* mutant, the positive control, displayed strong bands representing strong demethylation of centromeric repeats (Fig. 8).

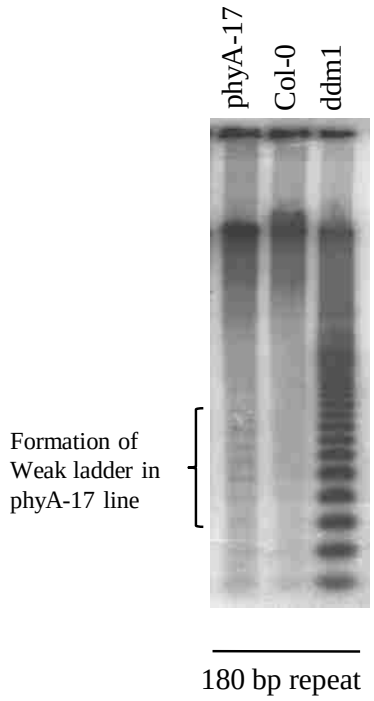


Figure 8: Methylation analysis of centromeric region of phyA-17 line
Southern blot of *Hpa*II digested DNA hybridized with a 180 bp single repeat probe specific to the centromeric region

2.3d *The role of known epigenetic factors in phyA' silencing*

The role of the selected genes of chromatin modification or RNAi pathways in *phyA'* silencing was studied using a genetic approach. Homozygous mutants of the selected genes were obtained from ABRC stock center. Each mutant line was crossed with phyA-17 or phyA-7 line, and the F1 progeny was allowed to self-fertilize to generate F2 populations. Analysis with mutants of *rdr2* and *kyp* were carried out by making reciprocal crosses, while analysis with rest of the mutants were done using one-directional cross (phyA-17 was used as the male parent). The scheme for genetic crosses is given in Fig. 2. A 3:1 ratio for short+intermediate and tall seedlings in the F2 generation would suggest the lack of interaction between the mutant allele and *phyA'*, while a 13:3 ratio for short+intermediate and tall would suggest a direct interaction of the mutant allele with *phyA'*. To confirm the interaction, F3 progeny derived from double-mutant F2 plants were analyzed. For this purpose, tall F2 seedlings (*phyA'/phyA'*) were rescued, and the presence of mutant allele such as *ago4* (heterozygous or homozygous) was determined by PCR or other recommended method. Occurrence of revertant phenotypes (short or intermediate) among the F3 progeny derived from double-mutant F2 parent would confirm the role of the mutant allele in *phyA'* silencing. F2 segregation ratio and FRc phenotype of F3 seedlings derived from each of the double-mutant line (homozygous or heterozygous) is shown in Table 2a and 2b.

Table 2a: Analysis of F2 and F3 derived from the crosses of RNAi/RdDM mutants

Mutant crosses	F2 segregation data			Statistical analysis		Double-mutant F2 plants[A]	F3 phenotype		Interac-tion[B]
	S/I	Tall	Total	3:1 (S/I:T)	13:3 (S/I:T)		Total Tall	Total S/I	
drm1/2 x 17	139	52	191	3:1	-	3 (F2-1, **F2-5, F2-8**)	169	0	No
17 x cmt3	185	55	240	3:1	-	2 (F2-1, F2-2)	143	0	No
ago4 x 17	198	117	412	3:1	-	5 (**F2-B1, F2-B22, F2-B23, F2-B32, F2-37**)	1933	0	No
rdr2 x 17	234	74	308	3:1	-	7 (**F2-2**, F2-3, **F2-4, F2-8, F2-10**, F2-11, **F2-12**)	461	0	No
sgs3 x 17	272	105	377	3:1	-	4 (F2-1, F2-2, F2-8, F-12)	1679	0	No
rdr6 x 17	3016	1010	4026	3:1	-	1 **(F-22)**	489	0	No
kyp x 7	52	18	70	3:1	-	4 (F2-2, **F2-5**, F2-7, F2-9)	1688	0	No
[C]nrpd2A x 7	137	65	202	-	-	1 (F2-4)	250	0	No

[A]Double mutant F2 plants, double-homozygous lines are shown as bold
[B]Interaction of mutant allele with *phyA'*
[C]*NRPD2A* is located on chromosome 1, therefore the segregation ratio indicates genetic linakge with phyA

Table 2b: Analysis of F2 and F3 derived from the crosses of mutants of genes involved in DNA methylation and chromatin remodeling

Mutant crosses	F2 segregation data			Statistical analysis		Double-mutant F2 plants	F3 phenotype		Interaction*
	S/I	Tall	Total	3:1 (S/I:T)	13:3 (S/I:T)		Tall	S/I	
[A]mom1 x 17	343	163	506	-	-	F2-9	1020	-	No
ddm1 x 7	288	84	372	3:1	-	[B]F3-5c	[C]More than 1000	-	No
[A]vim1 x 7	131	76	207	-	-	F2 genotyping is pending			

[*]Interaction of mutant allele with *phyA'*.
[A]*MOM1* and *VIM1* genes are located close to *phyA* locus on chromosome 1, therefore the segregation ratio indicates genetic linkage
[B]For *ddm1* analysis, F3 pooled population was genotyped to identify double mutant plant
[C]FRc phenotype was analyzed in F4 and F5 generation

Both F2 segregation data and F3 FRc phenotyping indicated lack of interaction of following alleles: *kyp, cmt3, drm1/drm2, rdr6, rdr2, ago4, ddm1, nrpd2A,* and *mom1* (Table 2a and 2b). F2 populations derived from the crosses of drm1/drm2, cmt3, rdr2, ago4, kyp, sgs3, rdr6, and ddm1 fit a 3:1 segregation ratio for short+intermediate and tall seedlings, indicating no effect of the mutant allele on *phyA'* silencing. F2 populations derived from the crosses of mom1, nrpd2a and vim1 did not fit a 3:1 ratio and generated more tall seedlings than expected. However, double mutant F2 lines derived from hybrid F1 plants (molecular analysis described below) of each mutant produced only tall F3 (*phyA'/phyA'*) seedlings, indicating that *phyA'* silencing was not released in the mutant background. While multiple F3 populations were analyzed in most cases, only one double-mutant F2 plant was isolated from the mom1 and nrpd2A crosses. Thus, a single F3 population for these crosses was phenotyped that clearly indicated the lack of interaction between *phyA'* and *mom1* and *nrpd2A* alleles. As the *mom1* locus is closely linked to *phyA* locus, recovery of only 1 double-mutant (heterozygous for *mom1*) out of 14 F2 plants was not surprising. However, it is not clear why only one double mutant derived from nrpd2A (heterozygous for *nrpd2A*) out of 12 tall F2 plants was recovered. The explanation may lie in the lack of recombination between the *nrpd2A* and *phyA* locus as both are located on chromosome 1.

As most of these mutant alleles represented null mutants (except ddm1), the role of the associated functional proteins in maintaining *phyA'* silencing was ruled out. These data rely on the fact that F3 lines are derived from a double mutant (homozygous or heterozygous for the mutant allele) F2 line. Therefore, the determination of the presence of mutant allele in each F2 line was an important part of the work. The following

sections describe the F2 genotyping of each mutant cross to ensure the presence of the mutation.

i) Genotyping of *drm1/drm2-2* allele

The *drm1/drm2* mutant was used as the female parent when crossed to phyA-17. Molecular markers used to genotype the presence of *drm1/drm2* mutation were generated using the following three primers in a single PCR reaction: T-DNA, 5' CATTTTATAATAACGCTGCGGACATCTAC-3'; drm1a, 5'-TGCGATTGACAATTTCCAATTTTCTCCAT -3'; and drm1b', 5'-TCTACCACCTCTTCCATACTTGC-3' (Xao X and Jacobsen S.E. 2002). drm1a and drm1b' are gene specific primers. In a PCR reaction containing the three primers, a 1.2 kb band is expected from the WT allele. However, due to the insertion of T-DNA into the locus, ~0.5 kb double band was produced from the *drm1/drm2-2* allele (Fig. 9a). Ten tall F2 seedlings were transferred to soil, and two F2 plants were found to be homozygous and one was heterozygous for the *drm1/drm2* mutation according to the PCR (Fig. 9b). All three F2 plants were grown till maturity to collect F3 seeds. F3 seedlings were screened under FRc light, all of which were found to be tall (Table 2a), indicating no role of DRM1 or DRM2 in maintaining *phyA'* silencing.

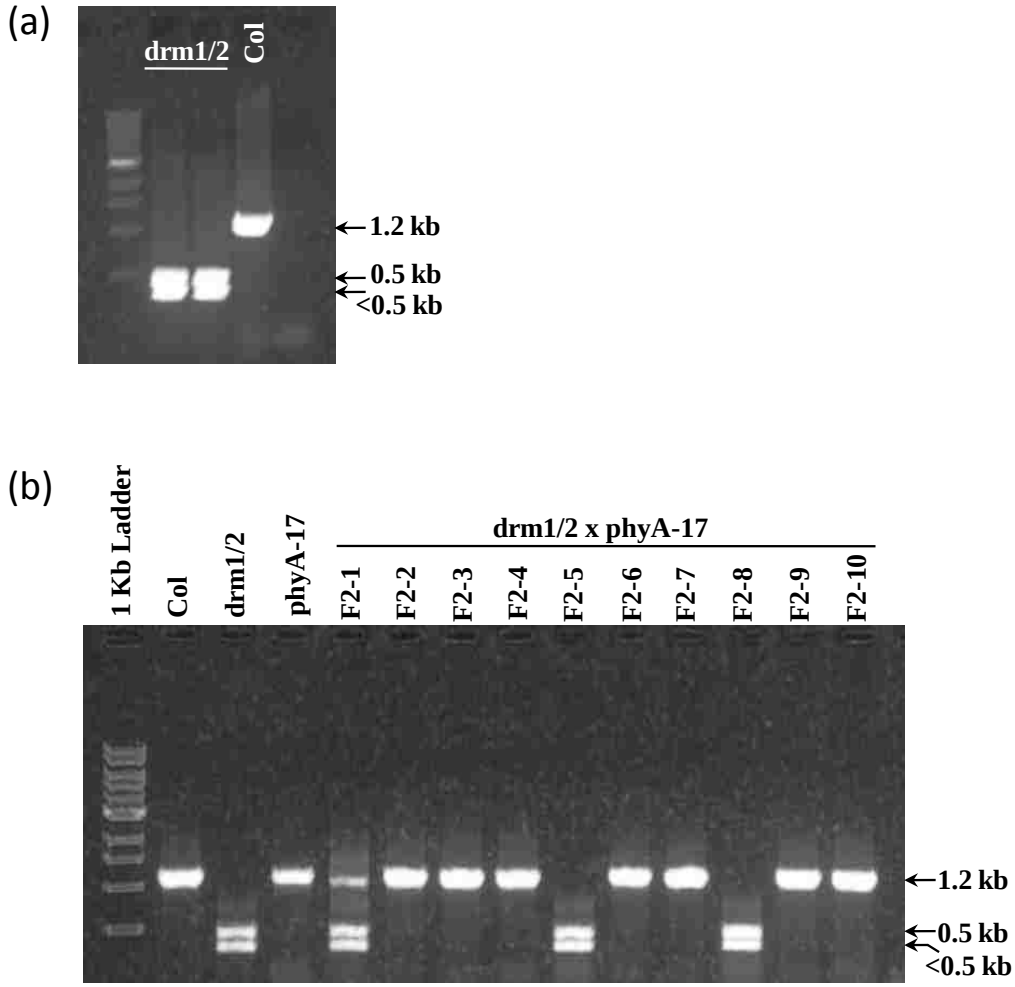


Figure 9: Genotyping of *drm1/drm2-2* allele on 0.8% agarose gel (a) Genomic DNA isolated from two drm1/2 parental line and Col was subjected to PCR using drm1a, drm1b' and T-DNA primers (b) Genomic DNA isolated form Col, drm1/2 parental line, phyA-17 and 10 F2 plants derived from a cross between drm1/2 and phyA-17

ii) Genotyping of *cmt3-7* allele

F1 plant was derived from crossing cmt3 mutant and phyA-17, and its F2 progeny (tall phenotype: *phyA'/phyA'*) were genotyped for the presence of *cmt3-7* allele. *cmt3-7* allele contains a single C/G to T/A transition mutation, which results in the stop codon. Translation of CMT3 stops after 27 amino acids, and thus, *cmt3-7* is likely a null allele (Lindroth et al., 2001). This mutation ('C' to 'T' represented as bold in the sequence) is identified by digesting DNA with *Cac*8I restriction enzyme, as it disrupts *Cac*8I recognition site from G**C**GAGC to G**T**GAGC. Three F2 plants displaying tall phenotype were genotyped by using the primer pair: 5'-TTGACTACCCCGGGAATGAACCCATTTGT-3,' and 5'-GATCTGCAACAAATCTCAGC-3'. The PCR product generated by these primers includes *Cac*8I site. Therefore, when the amplified product of homozygous *cmt3-7* allele is digested with *Cac*8I, an intact band of 735 bp is obtained (Fig. 10a). The PCR product amplified from the WT allele will be sensitive to *Cac*8I digestion, and 522 bp and 213 bp fragments are generated (Fig. 10a). The F1 cross (heterozygous for *cmt3-7* allele) gave all three size fragments; 735 bp, 522 bp, and 213 bp (Fig. 10a). Two F2 plants, F2-1 and F2-2, were found to be heterozygous for *cmt3-7* allele (Fig. 10b). F3 seedlings derived from F2-1 and F2-2 were FRc phenotyped. No reversion was seen among the two F3 populations (Table 2a), indicating no role of CMT3 in *phyA'* silencing.

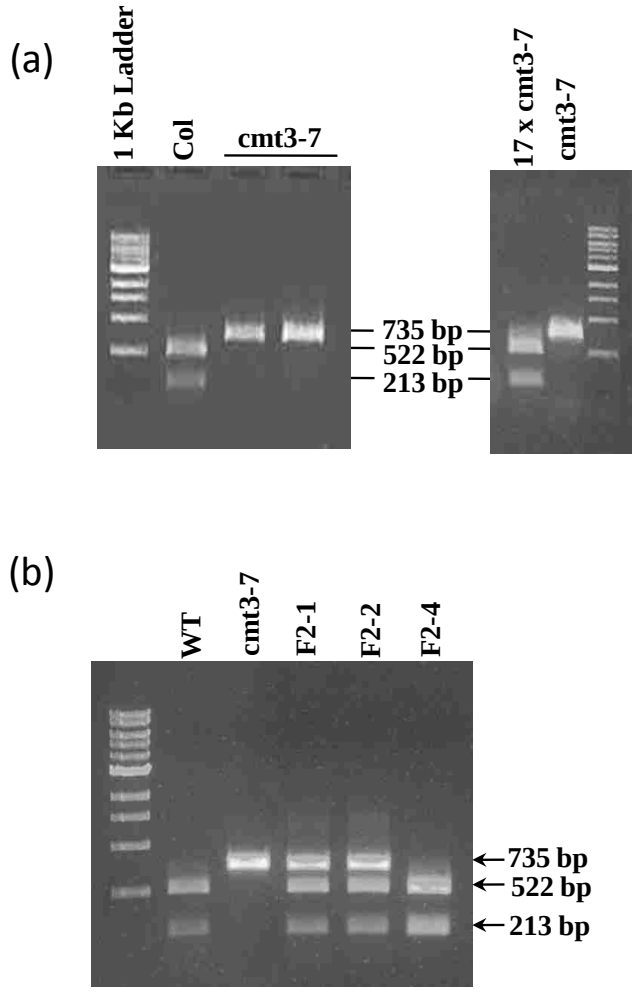


Figure 10: Genotyping of *cmt3-7* allele on 0.8% agarose gel (a) *Cac*8I digestion of PCR product amplified with cmt3R and cmt3F primers of Col, cmt3 parental line and F1 cross (b) *Cac*8I digestion of PCR product of three tall F2 plants derived from a cross between phyA-17 and cmt3-7 line

iii) Genotyping of *ago4-1* allele

Four F1 plants were obtained by crossing *ago4* mutant and phyA-17. F1 plants were genotyped using a combinations of oligonucleotide primers (5'-TGACTGACAGCTGAAAATGGGATGTG GAT-3' and 5'-GCCACTCCCTAGAACTCACCACCTAAGTT-3') and restriction enzyme, *Ava*II. (Zilberman et al., 2003). *ago4-1* allele contains a point mutation, which destroyed a splice acceptor site, and resulted in frame shift. This frame shift caused premature termination after 595 amino acids, and deleted almost the entire conserved domain (PIWI) of AGO protein family. The disrupted splice acceptor site was within the *Ava*II restriction site. Therefore, when the amplified product of homozygous *ago4-1* allele was digested with *Ava*II, it was refractory to cut, and gave an intact band of 1055 bp, while the WT allele gave 615 bp and 440 bp bands (Fig. 11a). As expected, F1 plants (heterozygous for *ago4-1*) gave 1055 bp, 615 bp, and 440 bp bands after digestion with *Ava*II. Twelve F2 plants (displaying tall phenotype) were selected for genotyping to identify double mutant plants. Five F2 plants were found to be homozygous for *ago4-1* as indicated by the presence of 1055 bp band (Fig. 11b). Approximately, 2000 seedlings from double mutant F2 plants were phenotyped under FRc light, and found to be 100% tall (Table 2a), indicating no role for AGO4 in *phyA'* silencing.

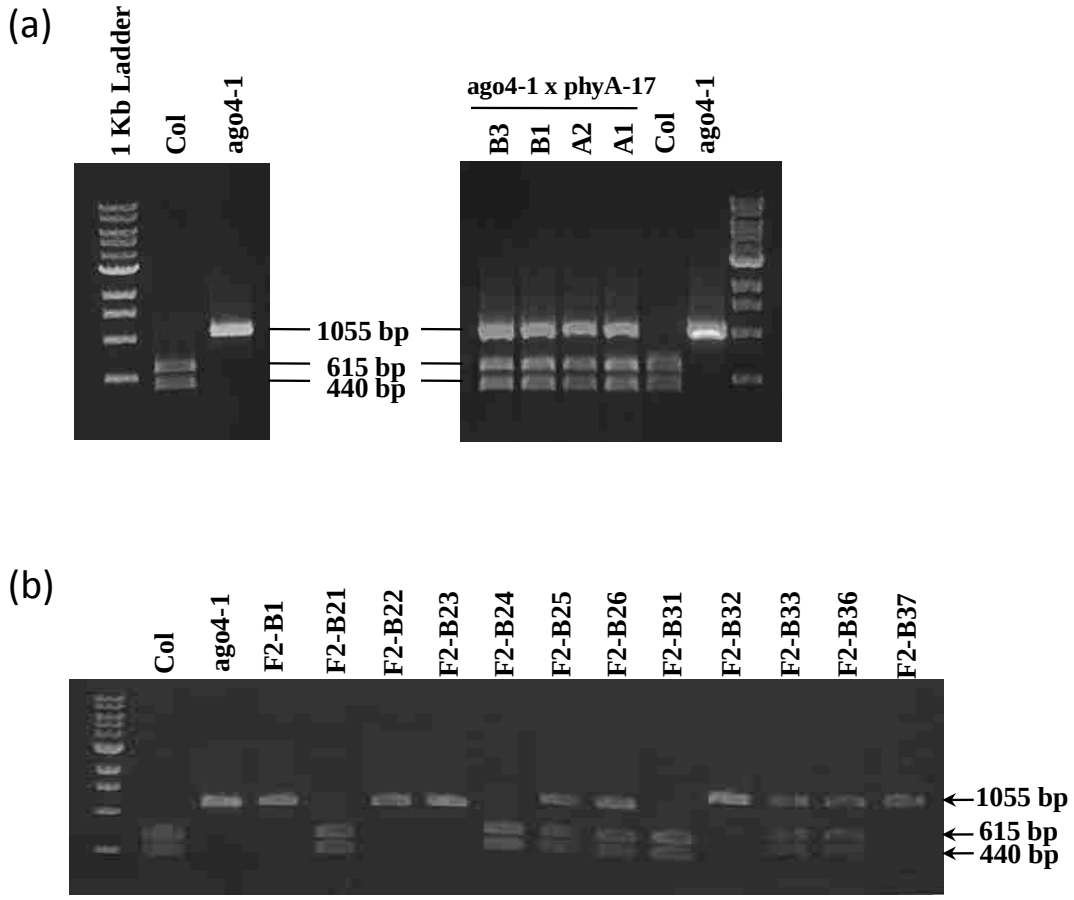


Figure 11: Genotyping of ago4-1 allele on 0.8% agarose gel (a) AvaII digestion of PCR product amplified with ago4F and ago4R primers of Col and F1 plants (b)) AvaII digestion of PCR product of 12 tall F2 plants derived from a cross between ago4-1 and phyA-17 plant

iv) Genotyping *rdr2-2* allele

The *rdr2* mutant was crossed both ways with phyA-17. Genotyping of *rdr2-2* was carried out by PCR using gene-specific primers and T-DNA left border primer. Gene-specific primers are rdr2F, 5'-TCACGCAGAGACAACTTTCG-3'; and rdr2R, 5'-ATTGACCGAGCAGAACATCA -3'; and T-DNA primer is LBb1, 5'-GCGTGGACCGCTTGCTGCAACT- 3'. In a PCR reaction containing gene-specific primers, a 700 bp band would be expected from the WT allele, while no amplification would be expected from *rdr2-2* allele due to T-DNA insertion. However, a weak amplification was observed from the DNA of *rdr2-2* parental line by gene specific primers (Fig. 12a), and ~600 bp fragment was generated by T-DNA primer (Fig. 12b). Ten F2 plants displaying tall phenotype were subjected to genotyping. Double mutant F2 plants were identified based on the amplification of ~600 bp fragment using T-DNA primer due to presence of T-DNA insertion (Fig. 12b). Seven F2 plants were detected as homozygous for *rdr2-2* (Fig. 12a and 12b). F3 seeds derived from each of the seven F2 plants were analyzed under FRc and found to be all tall (Table 2a), indicating no role of RDR2 in *phyA'* silencing.

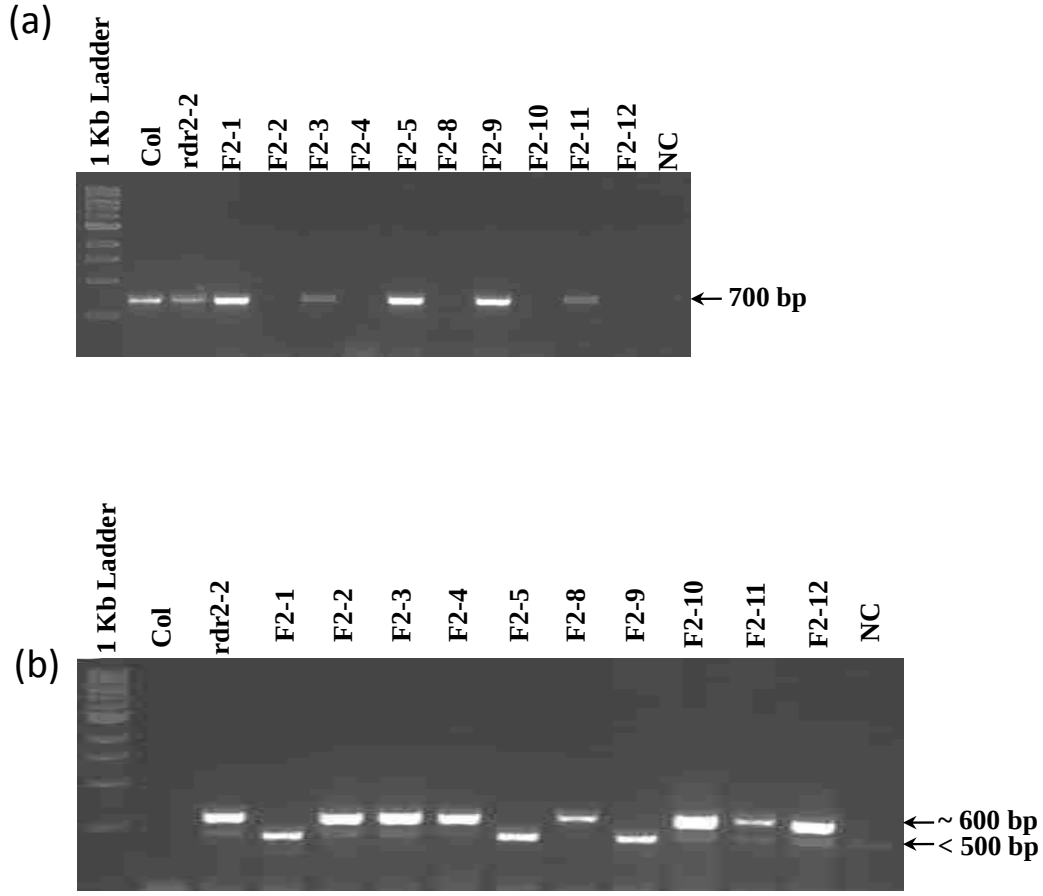


Figure 12: Genotyping of rdr2-2 allele on 0.8% agarose gel Genomic DNA isolated from rdr2-2 parental line, Col and 10 tall F2 plants derived from a cross between rdr2-2 and phyA-17 was subjected to (a) PCR using gene specific primers rdr2R and rdr2F (b) PCR amplification using rdr2F and LBb1 primer. NC represent negative (no DNA) control.

v) Genotyping of *sgs3-11* allele

The *sgs3* mutant was crossed with phyA-17 to get F1 hybrids. Three F1 plants were genotyped to identify the sg*s3-11* allele using a PCR primer pair; 5′-CAAAAAACCTGTGGTGGTCTGCA-3′ and 5′-ACAACCTTGGCACGTTCCTGC-3′. The *sgs3-11* allele was generated when nucleotide at splice site was mutated from G to A (Peragine et al., 2004). In PCR reaction containing above primer pair, *Pst*I site (CTGCAG) is generated in WT, but due to the conversion of G to A nucleotide in mutant line, *Pst*I site is lost in *sgs3-11*. After digestion of PCR product with *Pst*I, bands of 117 bp and 20 bp were generated in WT allele (Fig. 13a). These bands were detected on 15% polyacrylamide gel (TBE-Urea gel, Bio-Rad); however, 20 bp band was not detected (it could be lost during PCR purification step due to its small size). PCR amplification of homozygous *sgs3-11* allele did not work well, as an intact band of 137 bp was not detected in *sgs3-11* mutants (Fig. 13a and 13b). F1 plants (heterozygous for *sgs3-11*) gave two bands; 137 bp and 117 bp (Fig. 13a). Fifteen F2 plants displaying tall phenotype were selected for genotyping to identify double-mutant plants. Considering that amplification of 137 bp band from homozygous *sgs3-11* DNA was difficult, line F2-8 seemed to be a homozygous for *sgs3-11* as it failed to display 117 bp band (Fig. 13b). Approximately, 1500 F3 seedlings derived from F2-8 (presumably homozygous for *sgs3-11*) and other three F2 plants (heterozygous for *sgs3-11*) showed no reversion of seedling phenotype in FRc light (Table 2a and Fig. 13b), indicating no role of SGS3 in *phyA'* silencing.

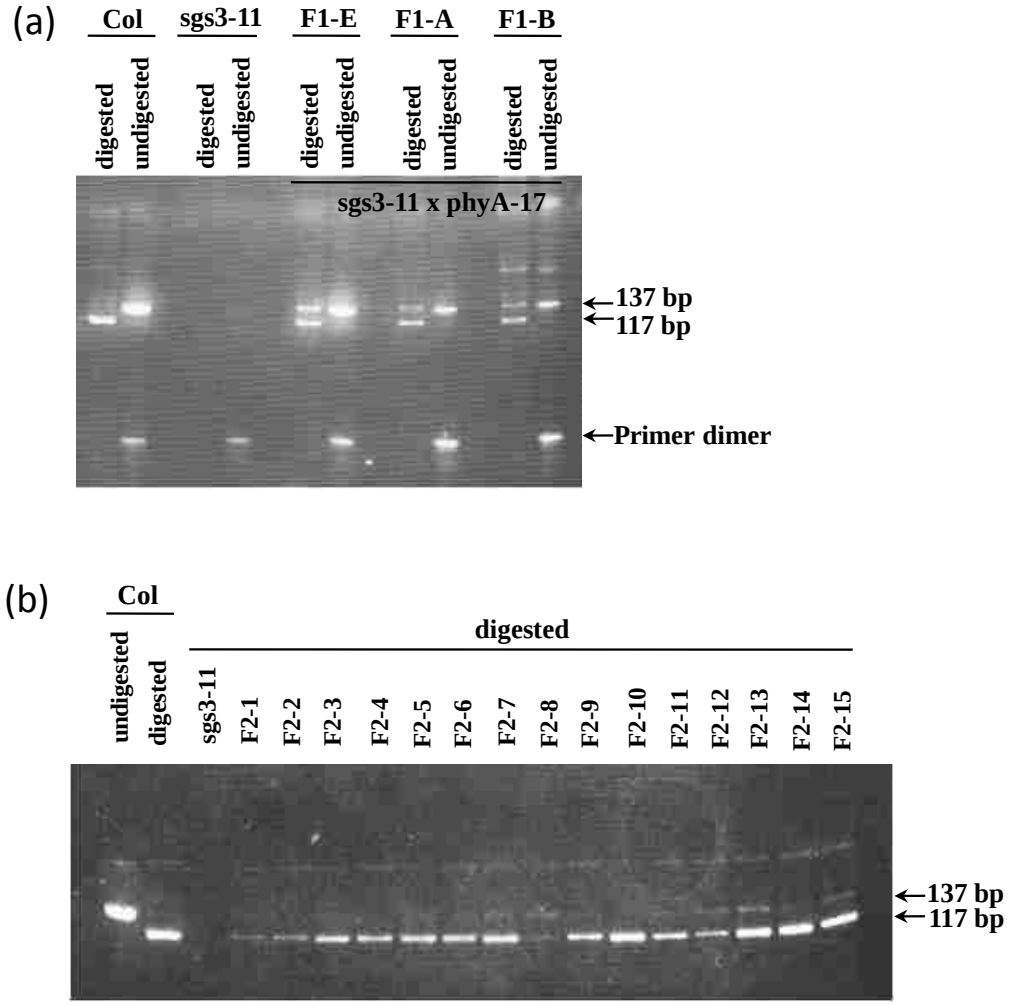


Figure 13: Genotyping of sgs3-11 allele on 15% PAGE (a) PstlI digestion of PCR product amplified with sgs3F and sgs3R primers of Col, parental line sgs3-11 and F1 cross, undigested samples were used as control (b) PstlI digestion of PCR product of 15 tall F2 plants derived from a cross between sgs3-11 and phyA-17 line, undigested samples were used as control

vi) Genotyping of *rdr6-11* allele

The *rdr6* mutant was crossed with phyA-17 to get F1 hybrids. *rdr6-11* contains a spontaneous mutation generating a stop codon. The *rdr6-11* mutation was identified by using amplified product in a PCR reaction using primers 5′-TACTGTCCCTGGCGATCTCT-3′ and 5′-CCACCTCACACGTTCCTCTT-3′ followed by digestion with restriction enzyme *Taq*I (TCGA) (Peragine et al., 2004). The PCR product of the WT allele contains two *Taq*I recognition sites. However, in PCR product amplified from the *rdr6-11* allele, one of the *Taq*I sites falls into the region where mutation occurred, becoming insensitive to *Taq*I digestion. Therefore, the homozygous *rdr6-11* allele gave two bands, 112 bp and 88 bp, whereas the WT allele generated three bands; 97 bp, 15 bp and 88 bp upon digestion with *Taq*I (Fig. 14). The 15 bp fragment was difficult to detect on 15% PAGE. Five F2 plants were genotyped to identify double mutant plants, and F2-22 was identified as homozygous for the *rdr6-11* allele (Fig. 14). No tall seedlings were found in F3 generation derived from F2-22 plant (Table 2a), indicating no role of SGS2 in *phyA'* silencing.

vii) Genotyping of *kyp-2* allele

Three F1 plants derived from a kyp-2 and phyA-7 cross were grown and genotyped as heterozygous for *kyp-2*. A total of 11 tall F2 plants were subjected to genotyping. *kyp-2* allele contains G to A mutation which is a mutations of conserved G residues at the splice acceptor sites corresponding to the last nucleotide of last intron, resulting in a frame shift. Mutation ('g' nucleotide in bold) is present at sequence, tcaattgta**g**/GAGCTCACTT (Jackson et al., 2002), where a/G (in wild-type gene)

represents the intron/exon boundary. Molecular markers used to genotype the *kyp-2* mutation were composed of the combinations of restriction enzyme, *Bgl*II and oligonucleotide primers: JP 1245; 5'-GCAGTGAAGATGAGAATGCGCCAGAGTTC-3', and JP 1246; 5'- CGCTATCAAGCGCA TATCCATAGTCGTAAGTGAGATC-3'. Underlined 'A' nucleotide is a change in base pair in reverse primer, JP 1246 compared to the actual sequence. The amplification product of *kyp-2* allele contains a change in base pair. *kyp-2* mutation together with the mutation introduced by reverse primer creates *Bgl*II site (AGATCT) in *kyp-2* PCR product. Therefore, when the PCR product of *kyp-2* allele is digested with *Bgl*II, two bands 267 bp and 38 bp are expected, while wild-type allele would generate a single band of 304 bp with the same primers. The 267 bp and 304 bp bands were resolved, while 38 bp band was not visible; however, comparison of these bands in separate lanes was difficult (Fig. 15b).

Analysis of F1 plants generated two bands, 304 bp and 267 bp characteristic of WT KYP and *kyp-2* alleles respectively (Fig. 15a). Analysis of 10 F2 plants generated single band (267 bp or 304 bp) from 4 F2 plants, and double bands (267 bp and 304 bp) from the remaining 6 F2 plants (Fig. 15b). Presence of the two bands confirmed the presence of heterozygous *kyp-2* locus. Approximately, 1500 F3 seedlings derived from 4 different double mutant F2 plants heterozygous for *kyp-2* allele were screened under FRc light (Fig. 15b, Table 2a). No reversion was detected in F3 generation indicating KYP is not involved in *phyA'* silencing.

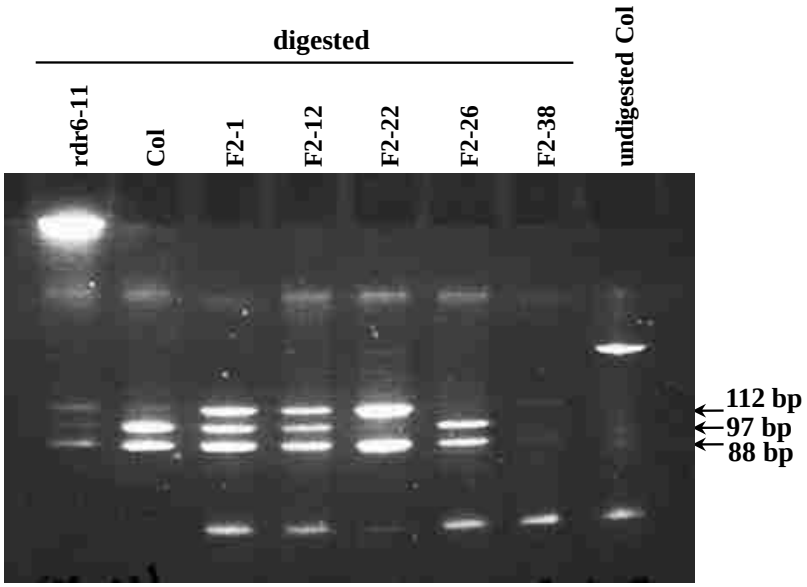


Figure 14: Genotyping of rdr6-11 allele on 15% PAGE TaqI digestion of PCR product amplified with rdr6F and rdr6R primers of Col, parental line rdr6-11 and 5 tall F2 plants derived from a cross between rdr6-11 and phyA-17 line, undigested sample of Col was used as control

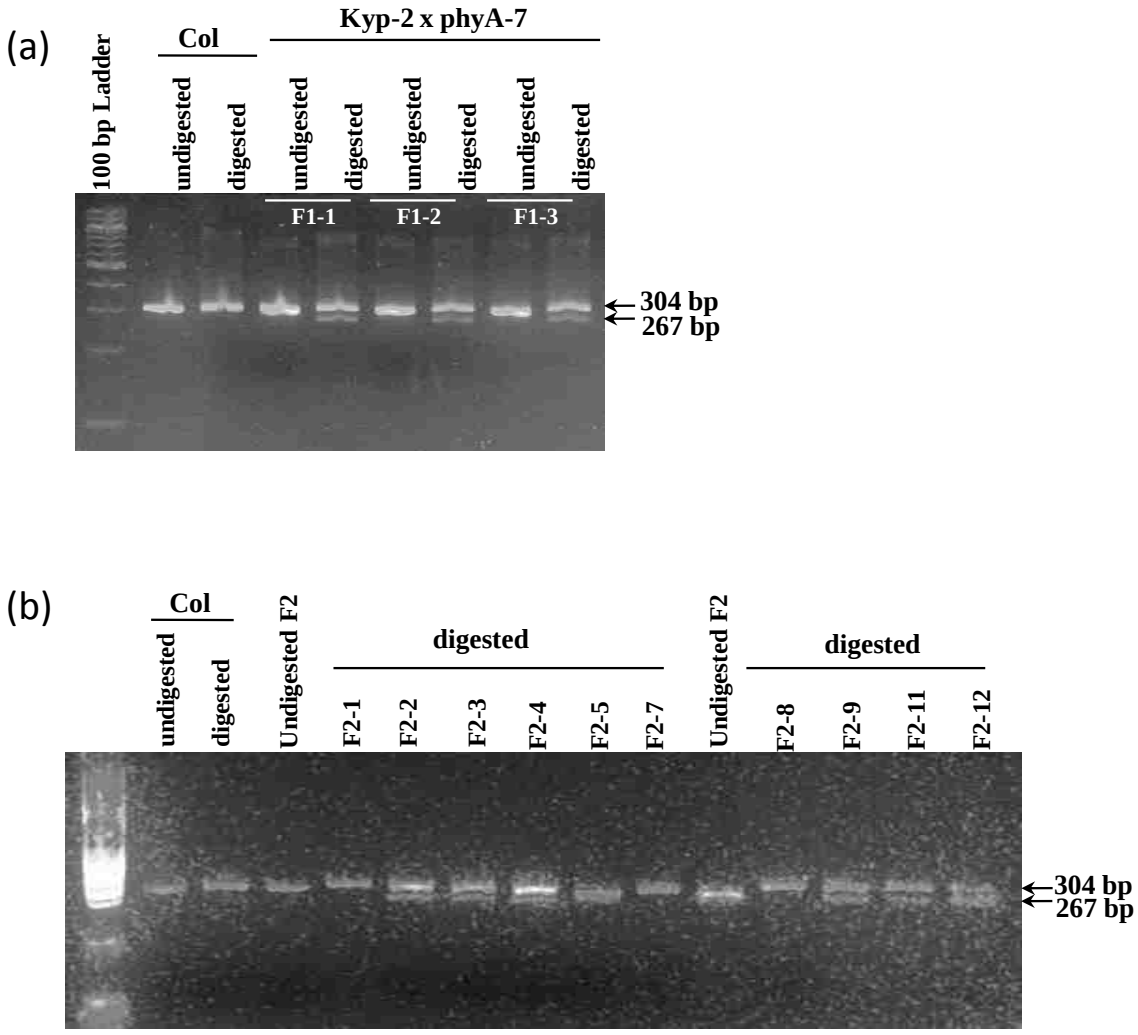


Figure 15: Genotyping of kyp-2 allele on 4% agarose gel (a) BglII digestion of PCR product amplified with kypF and kypR primers of Col and F1 cross, undigested samples were used as control (b)) BglII digestion of PCR product of 10 tall F2 plants derived from a cross between kyp-2 and phyA-7 line, undigested F2 samples were used as control

viii) Genotyping of *nrpd2A-2* allele

Crosses were done between phyA-17 and the *nrpd2A* mutant. Genotyping of *nrpd2A-2* was carried out by PCR using T-DNA left border primer (5′-CGTCCGCAATGTGTTATTAAG-3′), and gene specific primer pair: LP, 5'-TGGAGA TTTTCCACAACCAAG-3' and RP, 5'- CTGGCTTGACCATGAGGCCATG-3'(Onodera et al., 2005). The WT allele showed amplification with only the LP and RP primer pair (Fig. 16a and 16b), while the F1 hybrid (heterozygous for *nrpd2A-2*) showed amplification of a 2 Kb band with the LP and RP primer pair and amplification of a 0.8 Kb size band with the T-DNA and RP primer pair. Thirteen F2 seedlings displaying tall phenotype were selected for genotyping. Only single plant, F2-4 showed amplification with T-DNA primer pair, indicative of the presence of *nrpd2A-2* allele (Fig. 16c and 16d). All F3 seedlings derived from the F2-4 plant showed tall phenotype under FRc light, indicating no role of NRPD2A in *phyA'* silencing (Table 2a).

ix) Genotyping of *mom1-1* allele

Four successful crosses were obtained using *mom1-1* as a female parent and phyA-17 as male parent. All F1 plants were genotyped using PCR primer pairs: pro3-, 5'-CACTTTCCGATTTCGATTCTCG-3' and pro4+, 5'-CATGACTCCCCCAGCCAGTAG-3'; pro5+, 5'-GTGGTTACTGATCAAGTGTCG-3' and barbiE, 5'-GTGAAGGGCAATCAGCTGTTG-3'. Pro4+ is in the genomic sequence, which is deleted in *mom1-1*, while barbiE is present in the T-DNA that caused the deletion. Therefore, pro3- and pro4+ primer pair amplified only the WT allele, and generated a band of 260 bp (Fig. 17b). Pro5+ and barbiE amplified the *mom1-1* allele,

and generated a band of 600 bp (Fig. 17a). Thus, DNA from heterozygous F1 plants gave positive signal with all primers, while Col-0 and phyA-17 line gave amplification with only gene specific primers (Fig. 17a and 17b). To determine the interaction of *mom1-1* with *phyA'*, 13 F2 plants displaying tall phenotype were genotyped. Only one F2 plant, F2-9, was found to be heterozygous for *mom1-1* (Fig. 17c). F2-9 was grown till maturity to collect F3 seeds, and approximately 1000 F3 seedlings were phenotyped under FRc light (Table 2b). None of them were found to be short or intermediate, indicating no role of MOM1 in *phyA'* silencing.

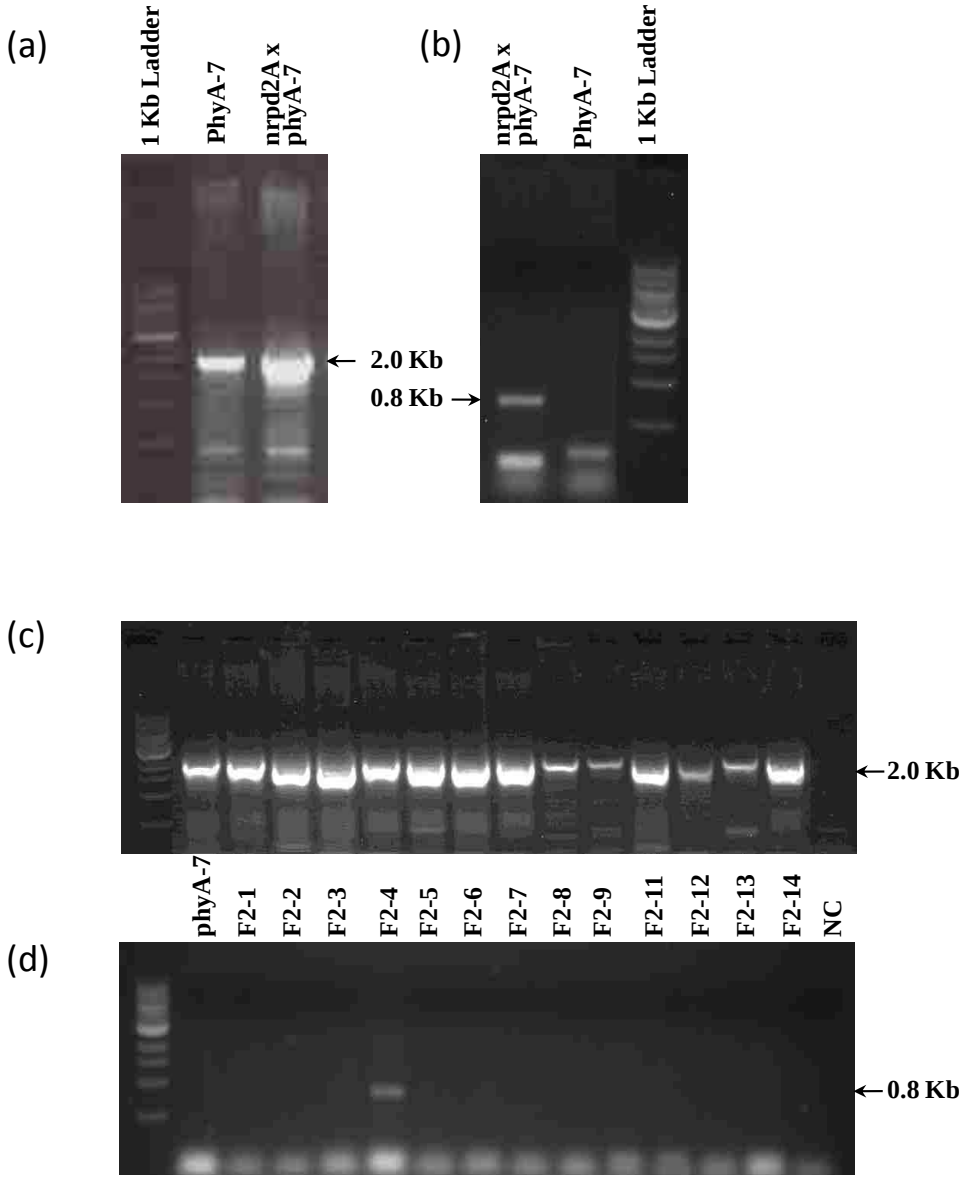


Figure 16: Genotyping of nrpd2A allele on 0.8% agarose gel (a) Genomic DNA isolated from phyA7 and F1 plant was subjected to PCR using RP and LP primers (b) Genomic DNA isolated from phyA7 and F1 plant was subjected to PCR using RP and T-DNA primer (c) Genomic DNA isolated from 13 tall F2 plants was subjected to PCR using RP and LP primers (d) Genomic DNA of 13 tall F2 plants was subjected to PCR using RP and T-DNA primer .

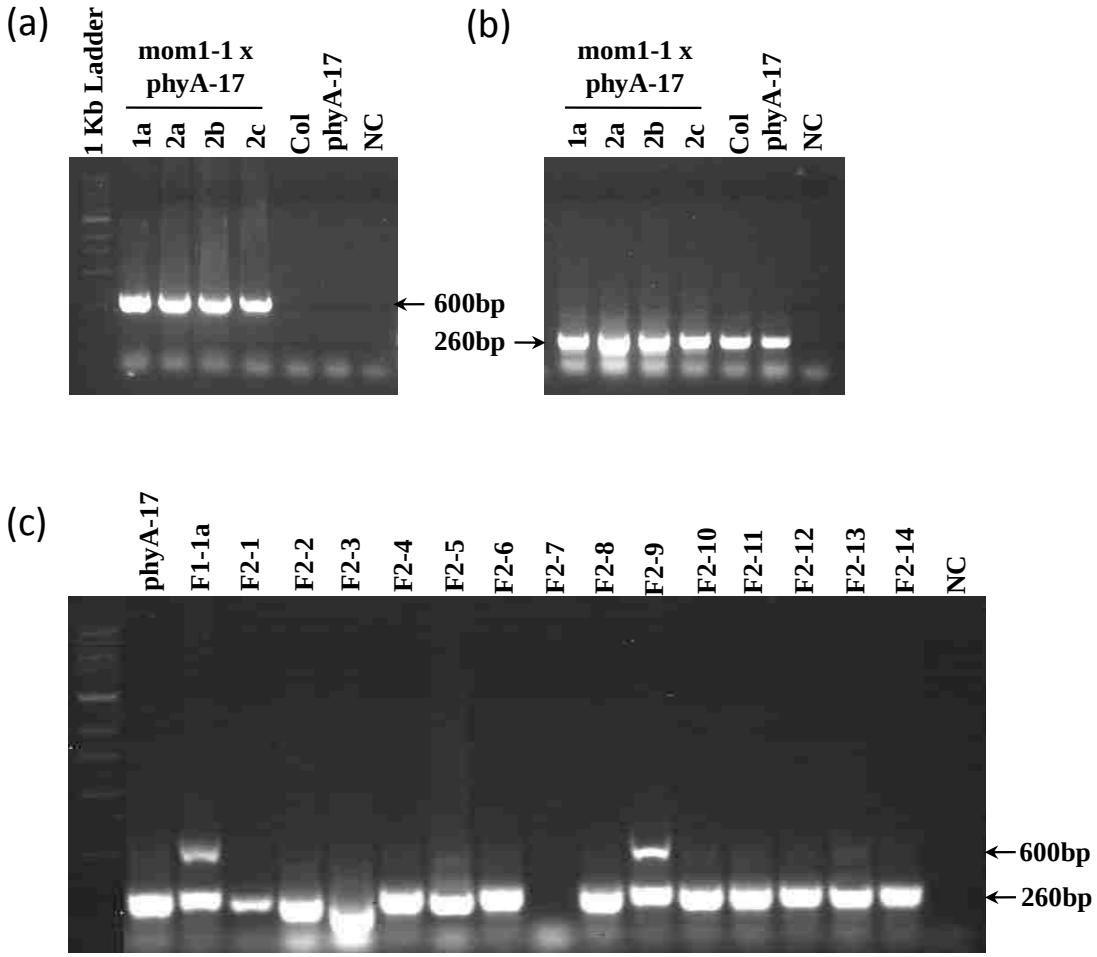
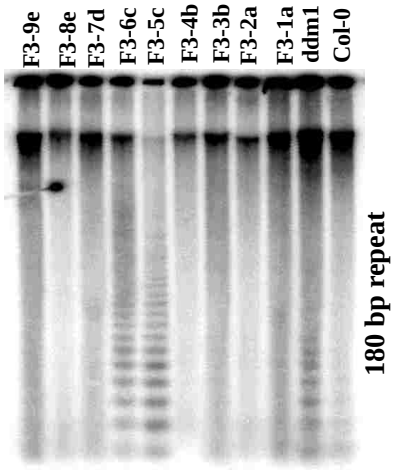


Figure 17: Genotyping of mom1-1 allele on 0.8% agarose gel (a) Genomic DNA isolated from Col, phyA-17 and four F1 plants was subjected to PCR using pro5+ and barbiE primers (b) Genomic DNA isolated from Col, phyA-17 and four F1 plants was subjected to PCR using pro3- and pro4+ primers (c) Genomic DNA isolated from 14 tall F2 plants was subjected to PCR using pro5+,barbiE, pro3- and pro4+ primers. NC represents negative (no DNA) control.

x) Genotyping of *ddm1* allele

Arabidopsis DNA hypomethylation mutant *ddm1* (SALK_000590) was used to study the role of DNA methylation in *phyA'* silencing. This is a T-DNA insertion line, but it is not well characterized. Therefore, presence of a T-DNA insertion in the DDM1 locus was confirmed by standard gene-specific and T-DNA (LB and RB)specific primers (data not shown). Next, the methylation level of three loci, 180 bp centromeric repeat, 5S RNA, Ta2, was analyzed using *Hpa*II digested genomic DNA on a Southern blot. Hypomethylation in each of these loci was found when compared to the WT (Col-0) genomic DNA (Fig. 18), indicating presence of the mutant *ddm1* allele. However, in each case, the hypomethylation in the ddm1 parental line was relatively weak, suggesting that it is a weak *ddm1* allele. Nine F3 populations (displaying tall phenotype) derived from 5 different F1 parents (ddm1 crossed with phyA-7 line) were selected for hypomethylation assay of selected loci. Two F3 populations were found to have strong hypomethylation of centromeric repeats, 5S rDNA and Ta2 loci (Fig. 18a, 18b and 18c) suggesting presence of the *ddm1* allele. Seedlings from the F3-5c line were carried to a successive generation. No reversion to the WT phenotype was detected in F4 and F5 generations (Table 2b). Together, these results indicate that mutation in *ddm1* can specifically relieve silencing of centromeric repeats but not of *phyA'*, indicating no role of DDM1 in maintaining *phyA'* silencing.

(a) **DNA methylation of Centromeric repeats**



(b) **DNA methylation of 5S ribosomal genes** (c) **DNA methylation of Ta2 retrotransposon**

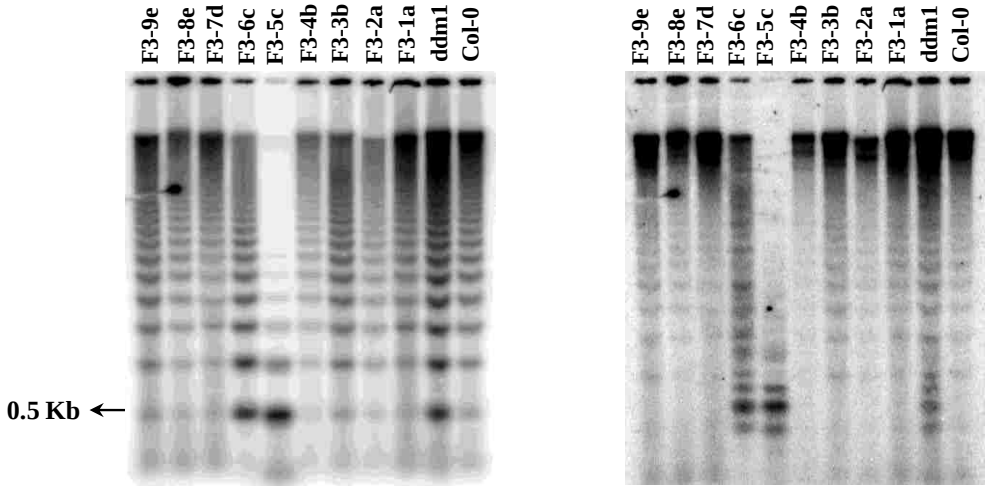


Figure 18: Methylation analysis of genome for characterization of ddm1 mutants by Southern blot analysis (a) Genomic DNAs prepared from light grown pooled seedlings of Col, ddm1 mutant and F3 generation derived from selected tall F2 plants (phyA-7 x ddm1) were digested with HpaII restriction enzyme, and probed with a 180 bp single repeat probe on DNA gel blot (b) The blot shown in Fig 18a was reprobed with 5S rDNA probe (c) The blot shown in Fig 18a was reprobed with Ta2 probe specific to the Ta2 retrotransposon element

2.4 Summary

In summary, *phyA'* exhibits unique qualities compared to previously described epigenetically modified loci. A repressive histone mark, H3K9me2 is absent in *phyA'*, and the known epigenetic regulators are not involved in *phyA'* silencing. These observations are not surprising, as *phyA'* contains a unique methylation profile that is distinct from the previously described epialleles. These observations suggest that *phyA'* silencing is a unique epigenetic phenomenon, which may be controlled by a novel epigenetic mechanism.

CHAPTER 3

Towards Identification of Novel Epigenetic Factors Involved in *phyA'* Silencing

3.1 Introduction

phyA' is an epigenetically modified epiallele of the *Arabidopsis thaliana* Phytochrome A gene. Transcriptional silencing of the *phyA'* epiallele is not maintained by components of an RNA-mediated epigenetic pathway as discussed in the previous chapter. Also, chromatin remodeling factors, DDM1 and MOM1 are not involved in *phyA'* silencing. On the other hand, transcriptional silencing of *phyA'* is tightly associated with the hypermethylation at CG sites, which reside solely in the transcribed region of the gene (Chawla et al., 2007). Previous studies have shown that hypermethylated DNA interacts with additional factors to confer transcriptional silencing (Jeddeloh et al., 1998; Bird 2001; Grafi et al., 2007; Woo et al., 2008; and reviewed by Chan et al., 2005; and Matzke et al., 2009). However, the known factors such as H3K9me2 and H3K4me3 do not play a significant role in *phyA'* silencing. Therefore, we hypothesized that novel factors interact with CG methylation in the coding region to maintain transcriptional silencing. These factors may directly interact with the methylated DNA or may be a part of the complex involved in imposing the transcriptional silencing of *phyA'*. To identify these factors, we took an approach of forward genetic screening of the mutagenized seeds of the phyA-17 line.

Forward genetic screens for the genes involved in the maintenance of transcriptional gene silencing of the well-known epialleles, transgenic loci or endogenous

repeats have identified several components of the epigenetic network in *Arabidopsis* (Jackson et al., 2002; Lindroth et al., 2001; Bartee et al., 2001; Mittelsten-Scheid et al., 1998; Amedeo et al., 2000; Elmayan et al., 1998; Kanno et al., 2004). Forward genetic screens involve random mutagenesis by mutagens, such as chemical reagents, irradiation or T-DNA insertions, and screening for mutants displaying the altered (revertant) phenotype. Success of a forward genetic screenin depends mainly on two factors, a stable genetic background (epigenetic state) and a clear revertant phenotype (Page and Grossniklaus, 2002). The *phyA'* epiallele fulfills both the criteria: silencing of *phyA'* is highly stable (no spontaneous reversion over multiple generations), and release of transcriptional silencing confers a conspicuous phenotype (short seedlings) under FRc light. Seed mutagenesis by the alkylating agent, ethane methyl sulfonate (EMS), is a rapid approach to screen mutants involved in the regulatory pathways. The majority of the time (99%), EMS induces alkylation of nucleotide 'G', which results in formation of O^6-ethylguanine. As O^6-ethylguanine pairs with 'T' instead of 'C', the original G/C pair is subsequently replaced with A/T, resulting in base substitutions (Greene et al., 2003). EMS is known to cause alteration in methylation pattern. It has been shown that CpG sites that are ethylated by EMS have higher affinity for methyltransferase (Farrance and Ivarie 1985; Ivarie and Morris, 1986). EMS modified CpG sites may mimic hemimethylated sites. These sites can be fully methylated in subsequent generation by methyltransferase enzyme and may result in genome-wide hypermethylation. Functionally important mutations, like base substitution from G to A resulting in early stop codon or intronic mutation resulting in truncated proteins or splice variants are considered as significant mutations caused by EMS.

Upon EMS mutagenesis of phyA-17 seeds, individuals that possessed a revertant phenotype (shorter seedling with expanded cotyledons) under FRc light were isolated. The phenotypic reversion is expected to originate from the release of *phyA'* transcriptional silencing. Revertants that carry a second-site mutation, which is responsible for releasing transcriptional silencing and restoring the WT phenotype are the most interesting mutants in the present study. These mutants are referred to as "suppressors of *phyA'* silencing" in the present work.

3.2 Materials and Methods

3.2a *Mutagenesis and Mutant Detection*

Mutagenesis was carried out on homozygous phyA-17 seeds using different strengths of EMS (Table 3). First, small scale screening was carried out consisting of a total of ~4000 seeds. The seed treatments were carried out in two different ways, depending on whether the seeds would be germinated on MS media (for M1 screening) or sown on soil (for collection of M2 seeds). Seed treatment for M1 screening consisted of the following steps: (a) in a 50 ml Falcon tube, 0.05 gm of seeds were soaked in 40 ml of sterile 100 mM phosphate buffer at 4^0C for overnight, (b) phosphate buffer-imbibed seeds were sterilized by replacing the buffer with 40 ml of 6% sodium hypochlorite solution, (c) after 5 min incubation, seeds were thoroughly washed with sterilized water four times, and (d) 40 ml of phosphate buffer was added to sterilized seeds for EMS treatment as described below. Seed treatment for M2 screening was slightly different as surface-sterlization was not necessary. Thus, 0.05 gm of seeds were soaked in 40 ml of

100 mM phosphate buffer overnight at 4^0C followed by replacement with fresh 40 ml 100 mM phosphate buffer.

EMS treatment: Both the tubes (M1 and M2 screening) were treated with EMS in the following way: (1) Under a fume hood, 16 µl of EMS (Sigma) was added to both the tubes containing seeds in the phosphate buffer to get a final concentration of 0.4% EMS, (2) Both the tubes were incubated for 8 hrs at room temperature with gentle agitation, (3) After 8 hrs, seeds were washed thoroughly 20 times with water (40 ml per wash).

For M1 screening, EMS treated seeds were divided equally on 8 MS plates, and germinated under continuous FR (FRc) light. For M2 screening ~25-30 seeds were sown immediately on lightly wetted soil in 8 cm x 5 cm pots. The pots were kept in the dark at 4^0C for 5 days for seed stratification in a growth chamber. After 5 days, growth conditions were changed to 25^0C with a 16/8 light/dark cycle till plants were grown to maturity. All plants (~1300 M1) were allowed to self-pollinate to generate M2 seeds. Each pot (50 in total) was bulk harvested to generate 50 M2 populations.

Large-scale screening consisting of a total of 8000 seeds was carried out by Dr. Jiangqi Wen at Noble Foundation, Ardmore, OK. This screen was divided into two batches (Table 3). The first batch of ~4000 seeds of phyA-17 was treated with 0.25% of EMS for 9 hrs and second batch with 0.25% of EMS for 15 hrs. The resulting 153 M2 populations were screened for the suppressor phenotype.

Table 3 : Design of EMS mutagenesis of phyA-17 seeds					
Name of Screening	**Number of seeds used**	**Screened at stage**	**Concentra-tion of EMS**	**Duration of EMS treatment**	**Number of M2 families collected**
Small scale	~2500	M1	0.4 %	8 hrs	-
	~1350	M2	0.4 %	8 hrs	50
Large Scale	~4000	M2	0.25 %	9 hrs	90
	~4000	M2	0.25 %	15 hrs	63

3.2b *Expression and Methylation Analysis*

M2 seedlings from small scale and the large scale screen were phenotyped under FRc light. Revertants were rescued as described in Chapter 2 (Section 2.2e). The rescued revertants were transferred to soil, and allowed to grow till maturity. For secondary screening, methylation status of the *phyA* gene was carried out on the revertants. Southern analysis was performed as described in Appendix A-4.

3.3 Results and Discussion

3.3a ***Screening of mutants suppressed in phyA' silencing***

Seed mutagenesis was carried out to generate potential mutations that would release transcriptional silencing of *phyA'*. phyA-17 seeds were mutagenized with EMS to induce genome-wide mutations. The mutants displaying the *phyA'* suppressor phenotype (shorter hypocotyls in FRc as compared to phyA-17 seedlings) were identified. Screening and selection of *phyA'* suppressors were based on the following two premises: (i) transcriptional silencing of *phyA'* is extremely stable over generations manifested by the distinct *phyA* mutant phenotype under FRc light (long hypocotyls and unexpanded cotyledons). Thus, recovery of a *phyA'* suppressor phenotype (short or intermediate length hypocotyls and expanded cotyledons) was expected to originate only from the induced mutations, (ii) transcriptional silencing of *phyA'* is released upon demethylation (Chawla et al., 2007). Therefore, the methylation status of *phyA'* was checked in the selected mutants by Southern blot analysis. Only those suppressor mutants were selected for further characterization, in which, full or partial hypermethylation at the *phyA'* locus was maintained. Finally, only those mutants that

maintain the suppressor phenotype in the next generation (M3) were selected for further analysis.

3.3b ***Small Scale Screening***

M1 screening: Design of phyA-17 seed mutagenesis experiment is shown in Table 3. The screening of M1 seed stock was done to determine if unexpectedly high rate of *phyA'* reversions would occur as a result of EMS treatment. Therefore, ~2500 M1 seeds (mutagenized seeds) were directly screened for *phyA'* suppressor phenotype (called as M1 screening). M1 screening can also identify dominant mutations.

In M1 screening, a single mutant, M1-1, was isolated that displayed intermediate length hypocotyl with expanded cotyledons under FRc light, representing a weak *phyA'* suppressor phenotype. M1-1 seedling was rescued and grown further till maturity. Visual phenotyping of M1-1 plant showed slow and weak growth with low fertility. To determine the methylation status of *phyA'*, Southern analysis was performed on DNA isolated from M1-1 plant (Fig. 19a). Southern analysis of *Eco*RI and *Hpa*II double-digested genomic DNA showed that WT *PHYA* gene contains methylation in two *Hpa*II sites: H5 and H8 (see Fig. 19a and 19b). The methylation pattern of the *phyA* locus in the M1-1 line was clearly different (Fig. 19b). The presence of a large ~4 kb band along with lower size bands indicates the presence of a hypermethylated and hypomethylated allele in this line. The progeny analysis confirmed this assumption as the segregation of these bands was seen in M2 progeny. The methylation pattern of three progeny derived from the self-fertilized M1-1 plant is shown in Fig. 19c. The hypermethylated band (~4

kb) and the hypomethylated bands (2.1 kb and 0.6 kb) segregated independently. The intermediate phenotype of M1-1 is consistent with the presence of a heterozygous *phyA* locus consisting of a hypermethylated and a hypomethylated allele. The M3 progeny displayed long seedling phenotype when derived from the parent containing hypermethylated allele, and short seedling phenotype when derived from the M2 parent containing the hypomethylated allele. The M2 plants containing both alleles generated segregating M3 populations consisting of short, intermediate and long phenotype. Low fertility in M1-1 indicates the presence of a second-site mutation. However, co-segregation of hypermethylated *phyA* allele with the long phenotype suggested no role for the second-site mutation in *phyA'* silencing. Therefore, alteration in methylation pattern, most likely due to EMS treatment is the basis of the phenotypic reversion in M1-1. Therefore, no further work was done on M1-1. However, M1 screening established that EMS treatment does not induce unusually high rate of reversions in phyA-17 line, and phyA-17 is an appropriate line for EMS mutagenesis.

M2 Screening: In M2 screening, ~1300 M1 plants were divided into 50 groups (25-30 plants per group), and allowed to set seeds. M2 seeds were harvested from each of the 50 groups in bulk, and designated as fifty independent M2 families. Presence of chlorotic sectors in a few of the M1 plants indicated that M1 seeds were successfully mutagenized. Approximately 500 M2 seeds derived from each of the fifty M2 families were surface sterilized and plated on MS medium to screen under FRc light for the suppressed *phyA'* phenotype. Seedlings with open cotyledons and short or intermediate length hypocotyls were designated as putative suppressors (Fig. 20).

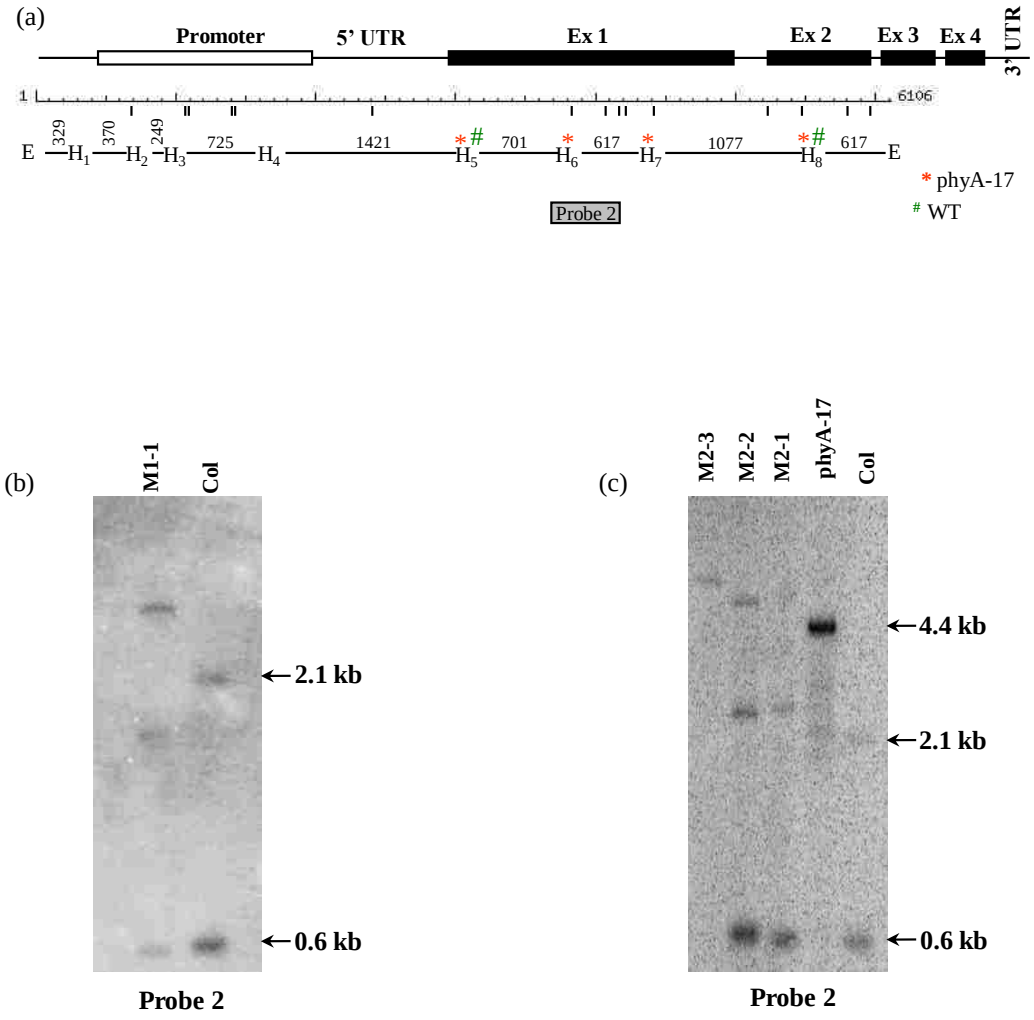


Figure 19: Methylation analysis of M1-1 by Southern blot (a) *Hpa*II map displaying methylated sites in phyA-17 (*), and WT *PHYA* (#) alleles. Fragment length between *Hpa*II (H) sites are given in bp, and fragment used as probe 2 is indicated below the map. Southern analysis of *Eco*RI and *Hpa*II digested genomic DNA of M1-1 plant (a) and M2 progeny (b) along with phyA-17 and Col-0 as controls.

Figure 20: *phyA'* suppressor phenotypes under FRc light
Representative intermediate (Int) and short (S) seedlings isolated in M2 screening

Putative suppressors of *phyA'* silencing were identified in six different M2 families (Table 4, Fig. 20). However, rescued suppressors derived from only one of these families (M2-22) were fertile. M2-22 suppressors displayed both short (total number 6) and intermediate (total number 39) phenotype under FRc light (Fig. 21). A total of 45 suppressors derived from M2-22 were successfully rescued and grown till maturity. No visual phenotypic aberrations were observed among these suppressors.

Southern analysis on twelve M2-22 individuals was done to analyze the methylation of *phyA* locus (Fig. 22). The methylation pattern of *Hpa*II sites in WT *PHYA* locus and *phyA'* epiallele is depicted in Fig. 22a. Of the four *Hpa*II sites in coding region, two are methylated in WT allele, while *phyA'* is methylated in all four. Thus, on a Southern blot, 4.4 kb and 1.7 kb bands are seen in *phyA'* and WT alleles, respectively. The M2-22 lines showed the intermediate bands ranging between 2.2 – 2.5 kb, indicating the presence of methylation over and above the WT level, but lower than that of the *phyA'* epiallele.

Two possibilities exist for the phenotypic reversion in M2-22 lines: (1) EMS-induced demethylation resulted in the release of transcriptional silencing of *phyA'*, (2) the presence of a second-site mutation resulted in the release of *phyA'* silencing. However, *phyA'* demethylation may have occurred independently of the second-site mutation or demethylation could be coupled with the second-site mutation. Therefore, detailed molecular and genetic analysis on M2-22 lines was carried out (described in chapter 4).

Table 4: Summary of M2 screening					
M2 Family	**Number of revertants**	**FRc phenotype**	**Survival on soil**	**Fertility**	**M3 phenotype**
M2-7	6	Short[1] and intermediate[2]	Yes	Non fertile	-
M2-13	4	All intermediate	No	-	-
M2-19	5	Short and intermediate	No	-	-
M2-22	45	Short and intermediate	Yes	Full	All short
M2-29	3	Short and intermediate	No	-	-
M2-35	5	All intermediate	No	-	-

1 - Short hypocotyl and expanded cotyledon
2 - Intermediate length hypocotyl and expanded cotyledon

Figure 21: FRc phenotype of individuals from M2-22 pool along with WT (Col-0) seedlings

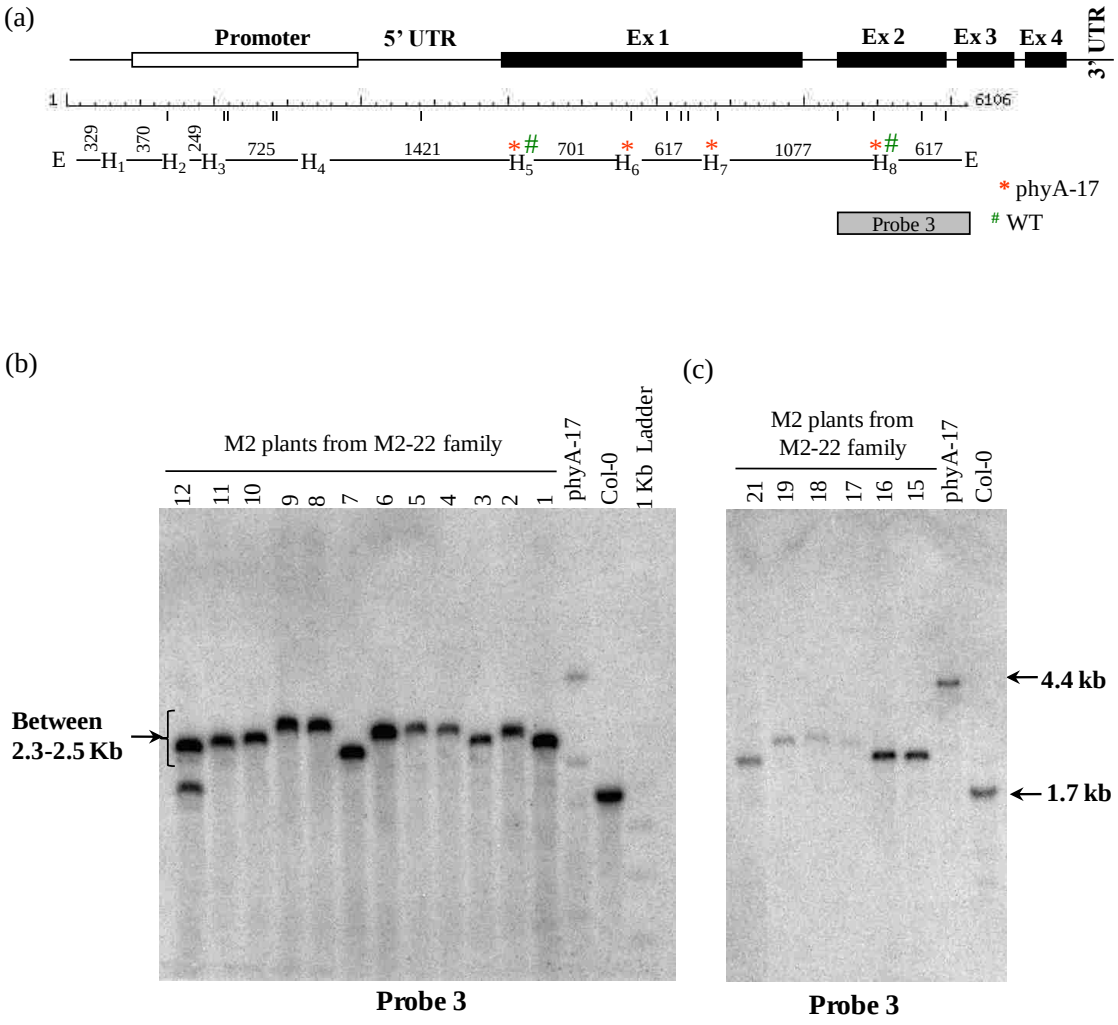


Figure 22: Methylation Analysis of *phyA* locus of M2-22 individuals (a) *Hpa*II map displaying methylated sites in phyA-17 (*) and WT *PHYA* gene (#). Fragment length between *Hpa*II (H) sites are given in bp, and fragment used as probe 3 is indicated below the map. (b) and (c) Southern hybridization of *Eco*RI and *Hpa*II digested genomic DNA with probe 3.

3.3c *Large-scale screening (M2 screening)*

A total of 153 M2 families obtained from the large-scale mutagenesis were screened for *phyA'* suppressor phenotypes under FRc light. Of these, 38 families generated suppressor mutants displaying either short (S) or intermediate (I) phenotype in seedlings grown in FRc light. Attempts to rescue these seedlings were successful on only 24 families (Table 5). The suppressor candidates from the remaining 14 families were lost during the rescue process. Additional attempts to rescue these seedlings should be made to isolate these potential suppressor lines. The majority of families generated a few suppressor lines each; however, some families such as 109 generated more than 15 suppressor lines each (Table 5). As we are most interested in the release of *phyA'* silencing from the hypermethylated *phyA* locus, the next screening consisted of methylation analysis on *Hpa*II Southern blot. So far, suppressor lines derived from 8 different families have been subjected to Southern analysis. Methylation analysis of the *Hpa*II sites in *phyA'* locus on a Southern blot revealed three types of patterns: (i) presence of hypermethylated *phyA'* gene as indicated by a single 4.4 kb or a larger band (e.g. 74a in Fig. 23), (ii) presence of two *phyA* alleles, hypermethylated and hypomethylated as indicated by ~4.4 kb and ~1.7 kb bands (e.g. 109 family, Fig. 24b and 24c), (iii) presence of a single hypomethylated band as indicated by the presence of a single ~1.7 kb band (e.g. 109g; Fig. 25).

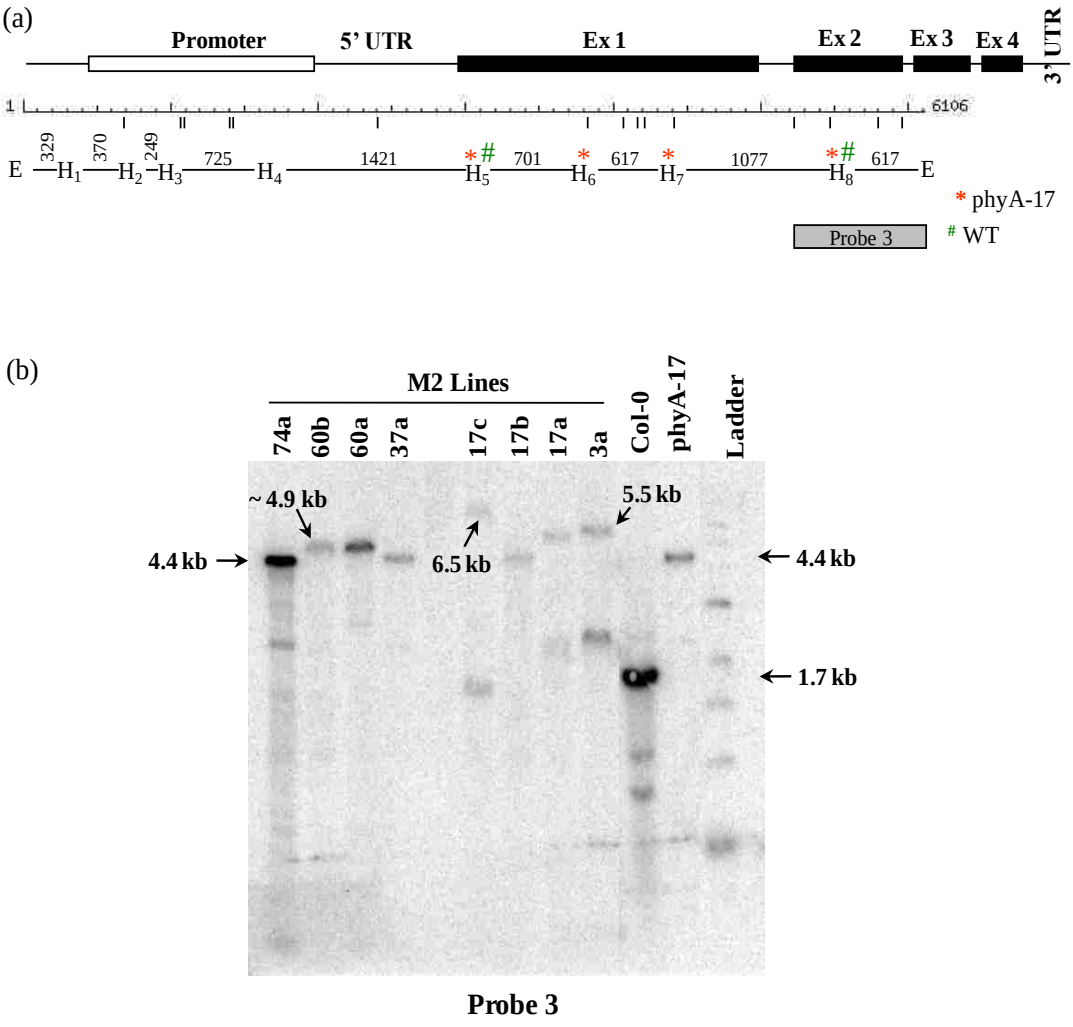


Figure 23: Methylation Analysis of *phyA* gene in suppressor lines identified in large scale screening (a) *Hpa*II map displaying methylated sites in phyA-17 (*) and WT *PHYA* (#) alleles. Fragment length between *Hpa*II (H) sites is given in bp, and fragment used as probe 3 is indicated below the map, (b) Southern analysis of *Eco*RI and *Hpa*II digested genomic DNA of M2 lines using probe3

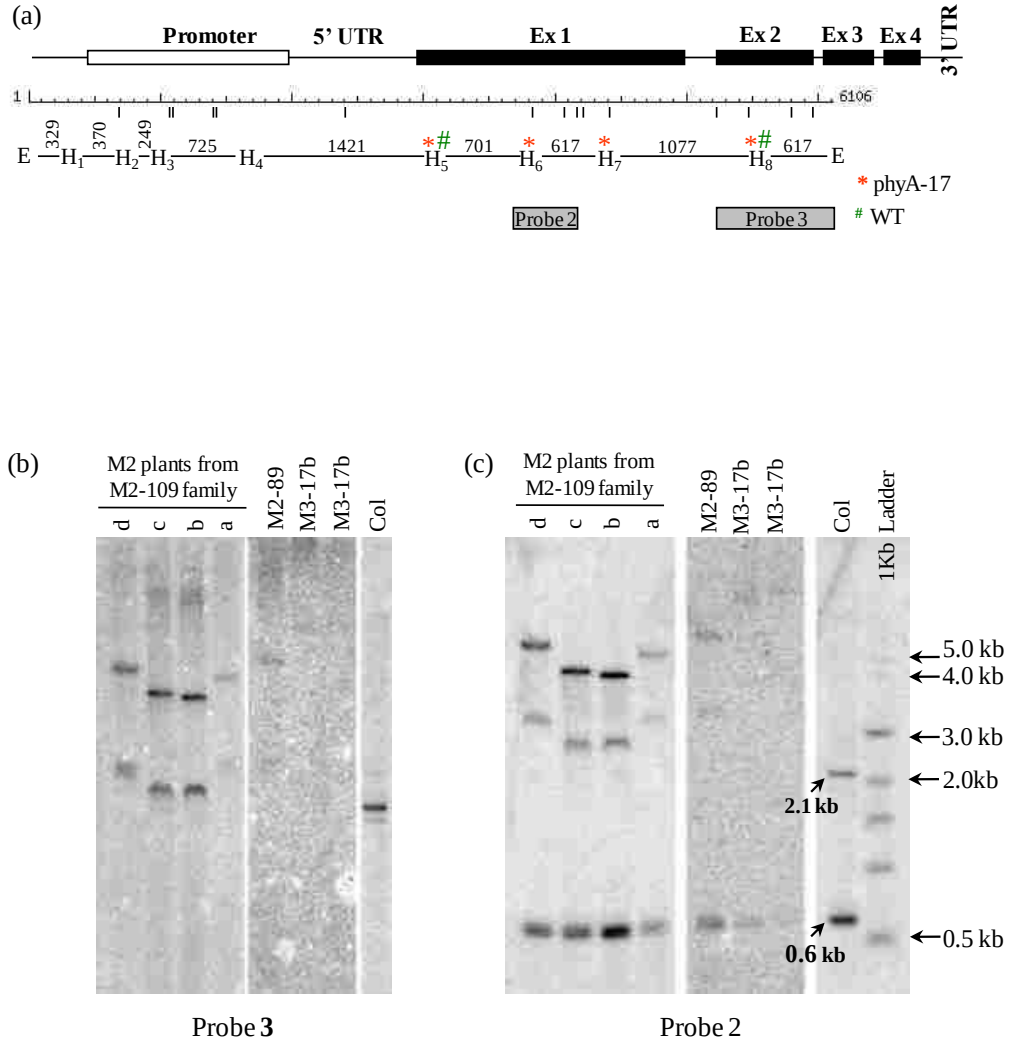


Figure 24: *phyA'* Methylation Analysis of suppressor lines identified in large scale screening (a) *Hpa*II map displaying methylated sites in phyA-17(*) and WT (#). Fragment length between *Hpa*II (H) sites are given in bp, and fragment used as probe 2 and 3 is indicated below the map (b) Southern analysis of *Eco*RI and *Hpa*II digested genomic DNA of M2 lines (M2-109 and M2-89) and M3 plant (M2-17) isolated from large scale screening and hybridized with probe3 (c) the same blot hybridized with probe 2

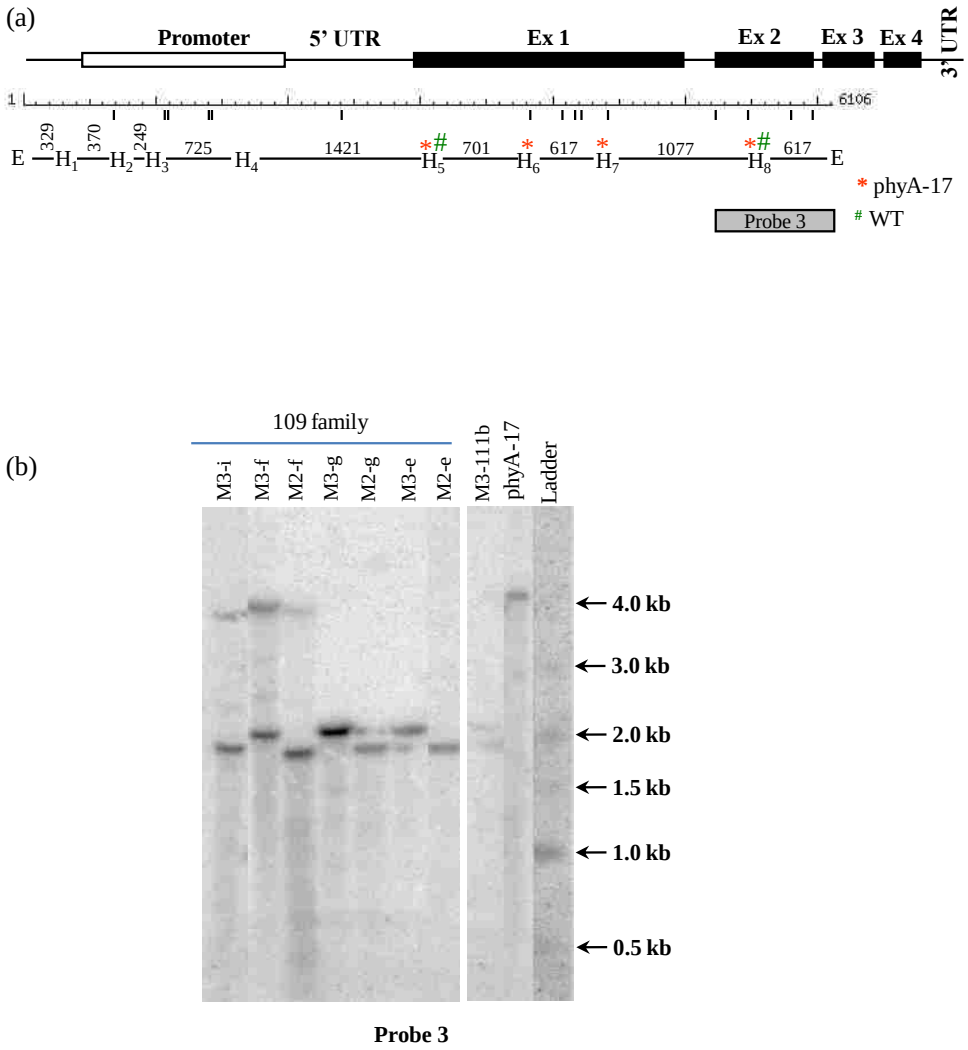


Figure 25: *phyA'* Methylation Analysis of suppressor lines identified in large scale screening (a) *Hpa*II map displaying methylated sites in phyA-17(*) and WT (#). Fragment length between *Hpa*II (H) sites are given in bp, and fragment used as probe 3 is indicated below the map (b) Southern analysis of *Eco*RI and *Hpa*II digested genomic DNA of M2 and M3 plant (M2-109 family) and a single M3 plant from M2-111 family plant isolated in large scale screening using probe2

The suppressor lines containing a hypomethylated *phyA* allele were excluded from future analysis. These lines may also contain second-site mutations that affect *phyA* methylation, such as a mutation in the *MET1* gene. However, analysis of these lines will not generate information on the epigenetic pathway underlying CG methylation mediated transcriptional silencing. Thus, the most interesting mutants are those that display the ~4.4 kb *phyA* band, indicative of fully methylated *phyA* locus. Four families displayed the single hypermethylated *phyA* band, while another four contained both hypermethylated and hypomethylated *phyA* alleles.

Although all suppressor lines displayed *phyA'* suppressor phenotype under FRc, plant phenotype varied among suppressor families. The phenotypic analysis is summarized in Table 5. Some suppressor lines were sterile (e.g. 60a, 17c; Fig. 26a), some lines showed healthy and vigorous growth (e.g. 17a; Fig. 26b), and some lines displayed no phenotype. An aberrant phenotype is indicative of ectopic genomic activity such as activation of transposons or activation of imprinted genes resulting in abnormal development.

M3 progeny derived from the M2 lines were analyzed for the maintenance of *phyA'* suppressor phenotype under FRc. Based on M3 phenotyping, suppressor lines could be divided into the following four categories:

- Stable inheritance of *phyA'* suppressor phenotype: all M3 progeny displayed short phenotype (e.g. 109a, 111b; Table 5)

- Partial inheritance of *phyA'* suppressor phenotype: mostly all M3 progeny displayed short phenotype with a few individuals displaying tall phenotype (e.g. 12 tall seedlings among 141 individuals of 106 suppressor line, and 8 tall seedlings among 138 individuals of 109i suppressor line; Table 5). Occurrence of low *phyA'* phenotype (tall) among the M3 progeny indicates low heritability of the suppressor mutation.
- Rare inheritance: mostly tall M3 seedlings with a few short and intermediate seedlings, indicating very low heritability of the suppressor mutation. Majority of the suppressor lines were found to fall in this category (e.g. 3, 25, 89, 102; Table 5)
- Loss of *phyA'* suppressor phenotype: All M3 progeny displayed tall phenotype. These lines were not selected for the further characterization, although they maintained hypermethylation of *phyA'* in M2 or M3 generation (e.g. 37 and 74; Fig. 23 and Table 5)

Table 5: Suppressors isolated in large scale M2 screening

M2 family	No. of Rever-tants	FRc pheno-type	Plant phenotype	**phyA'* methylation	M3 phenotype
M2-3	3	S	Normal & Fertile	HM-1A (Fig. 23b)	T+S+I
M2-17	17a 17b 17c	I I S	Vigorous, fertile(Fig.26b) fertile sterile (Fig. 26c)	HM-1A (Fig. 23b) HM (Fig. 23b) HM-1A (Fig. 23b)	T+S+I 9(S) + 1(I) + 5(T) -
M2-25	25	I	Weak growth	ND	T + S + I
M2-37	37	I	Normal & fertile	HM (Fig. 23b)	All tall
M2-60	60a 60b	I I	Normal growth, but male sterility (Fig. 26a)	HM (Fig. 23b)	ND
M2-74	74	I	Normal & fertile	HM (Fig. 23b)	All tall
M2-89	89	I	Normal & fertile	HM-1A (Fig. 24b,c)	T + S + I
M2-102	102	S	Normal & fertile	ND	T + S + I
M2-106	106	I	Normal & fertile	ND	129 (S+I) + 12(T)
M2-107	107	I	Normal & fertile	ND	T + S + I
M2-108	108a 108b 108c	I I I	Normal & fertile	ND	- T + S + I -
M2-109	109a 109b 109c 109d 109e 109f 109g 109h 109j 109k 109i	All I	Normal & fertile	HM-1A (Fig. 24b,c) HM-1A (Fig. 24b,c) HM-1A (Fig. 24b,c) HM-1A (Fig. 24b,c) DM (Fig. 25b) HM-1A (Fig. 25b) DM (Fig. 25b) ND ND ND HM-1A (Fig. 25b)	12 (S+I) 7(S)+11(I) +13(T) - - 60 (S+I) T + S + I 61 (S+I) ND 48(S+I) + 8(T) 90(S+I) + 18(T) 130 (S+I) + 8(T)
M2-110	110	I	Normal & fertile	ND-	T+S+I
M2-111	111a 111b 111c	S	Weak & fertile	ND- DM (Fig. 25b) ND	- 3(I) + 12(S) -
M2-113	113a, b	I	Normal & fertile	ND	ND
M2-114	114a-e	I	Normal & fertile	ND	ND
M2-117	117a, b	I	Normal & fertile	ND	ND
M2-131	131a-e	I	Fertile	ND	ND
M2-132	132a, b	I	Fertile	ND	ND
M2-134	134	I	Weak & Fertile	ND	ND
M2-147	147a-e	I	Fertile	ND	ND
M2-149	149	I	Normal & fertile	ND	ND
M2-150	150	I	Normal & fertile	ND	ND
M2-153	153	I	Weak & fertile	ND	ND

* HM-1A: Hypermethylation present in one *phyA* allele; HM: Hypermethylation present in both *phyA* alleles; DM: demethylation, and ND: Not done

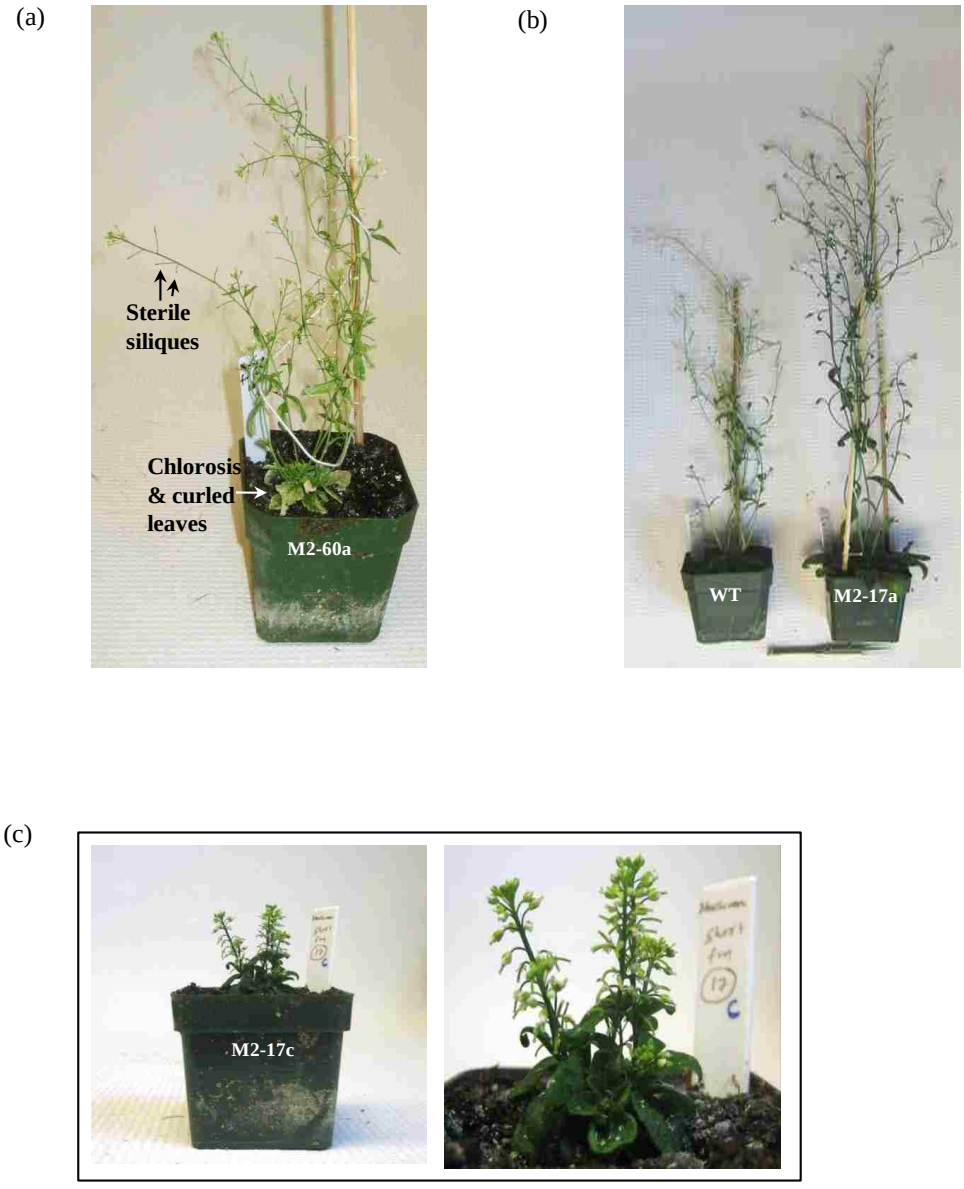


Figure 26: Methylation Analysis of *phyA* gene in suppressor lines identified in large scale screening (a) *Hpa*II map displaying methylated sites in phyA-17 (*) and WT *PHYA* (#) alleles. Fragment length between *Hpa*II (H) sites is given in bp, and fragment used as probe 3 is indicated below the map, (b) Southern analysis of *Eco*RI and *Hpa*II digested genomic DNA of M2 lines using probe3

3.4 Summary

Screening of EMS-mutagenized populations derived from phyA-17 seeds identified several different suppressor lines that are strong candidates for *phyA'* regulators and serve as important genetic resources for unraveling the epigenetic pathway underlying *phyA'* silencing. Future experiments involving map based cloning are likely to identify the underlying mutation. A number of suppressor mutants were found to be sterile, indicating the genes involved in *phyA'* silencing are also involved in gametogenesis and flower development. The screening is still in progress and identification of interesting suppressors is yet to be completed.

CHAPTER 4

Characterization of a *phyA'* suppressor, *sps-1*

4.1 Introduction

To identify the novel factors involved in *phyA'* silencing, we carried out a suppressor screen by mutagenizing the phyA-17 line and isolating mutants suppressed in the *phyA'* phenotype. The expected suppressor phenotype is short hypocotyls in seedlings germinated in FRc light (closer to WT phenotype). We found suppressor mutants in one population, M2-22, which consisted of intermediate or short (WT) seedlings.

Several M2 plants from the M2-22 family displaying a suppressor phenotype were subjected to Southern analysis to determine the methylation pattern of the *phyA* gene. All of them were found to contain hypermethylation compared to the WT *PHYA* gene. One of the individuals, 22-4, was designated as suppressor of *phyA'* silencing 1 (sps-1). The sps-1 line shows strong *phyA'* suppression manifested by a short hypocotyl and well expanded cotyledons in seedlings germinated under FRc for three days (Fig. 27). To further validate the presence of a suppressor mutation in the sps-1 line, the following questions were addressed:

(a) Whether the suppressor mutation is inherited by the progeny? (b) Is the methylation pattern of the *phyA* locus maintained in the subsequent generations? (c) Whether phenotypic reversion in sps-1 is accompanied with the release of *phyA* transcriptional silencing? (d) Is sps-1 a trans-acting mutation?

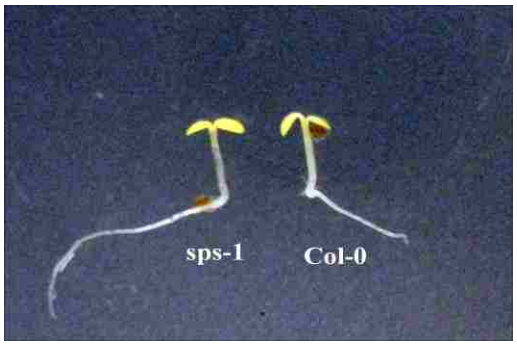


Figure 27: Phenotype of sps-1 seedling alongside wild-type (Col-0) seedling after 3 days of growth in FRc light
sps-1 displays a short hypocotyl and expanded cotyledons, similar to the WT phenotype under FRc

4.2 Materials and Methods

4.2a *RNA isolation, northern hybridization and reverse transcription-PCR (RT-PCR)*

RNA isolation was performed using an RNeasy plant mini kit (Qiagen Inc.) according to the manufacturer's instructions. Total RNA was quantified on Nanodrop (ND-1000). Northern analysis was performed as described in Appendix A-5. Data analysis was performed using ImageQuant (GE Healthsciences Inc.) software. *PHYA* was probed using a probe corresponding to exon 2 and 3. Labeling was conducted with α-P32 dCTP using the Roche Prim-A-Gene kit. Probes were purified using the spin column containing G50 Sephadex beads (Sigma-Aldrich Co.). RT-PCR was carried out using Access RT-PCR kit (Promega Co.) according to the manufacturer's instructions. Primers used in RT-PCR are listed in Appendix C-1.

4.2b *Bisulfite sequencing of sps-1*

Bisulfite sequencing was performed on CTAB-extracted genomic DNA of the sps-1 line. Approximately 2 µg of sps-1 genomic DNA was treated with sodium bisulphite using EpiTect Bisulfite kit (Qiagen Inc), followed by PCR amplification with the primers given in Appendix C-3. For sequencing, the PCR products were gel extracted and sent to the DNA sequencing core facility at University of Arkansas. Overlapping sequences from sps-1 line were assembled and compared with the phyA-17 methylation profile.

4.3 Results and Discussion

4.3a *Phenotyping of sps-1 (suppressor of phyA' silencing 1)*

Several suppressor mutants were identified in an M2-22 family obtained by EMS mutagenesis of phyA-17 seeds (as described in Chapter 3). Twelve individuals were chosen for secondary screening by Southern hybridization to check the *phyA'* methylation status. Later, plant 22-4 was selected for further molecular and genetic analysis, and designated as sps-1. sps-1 displayed a strong *phyA'* suppressor phenotype characterized by short hypocotyls and open cotyledons in FRc-grown seedlings (Fig. 27). No aberrant phenotype was detected in sps-1 plants during the vegetative and reproductive phase of development in growth chamber, indicating sps-1 is healthy and fertile. Further, M3 and M4 seedlings of sps-1 were grown under FRc light to confirm the stability of the suppressor mutation into progeny. M3 and M4 seedlings displayed a consistent short phenotype among the population of approximately 200 seedlings per generation (Fig. 28). Maintenance of a strong suppressor phenotype in large M3 and M4 populations suggests the presence of a homozygous mutation (responsible for the release of *phyA'* silencing) in the sps-1 line.

4.3b *No intragenic mutation detected in phyA'/sps-1*

Complete sequencing of the *phyA* locus in the sps-1 line was performed to look at the possibility of EMS-induced mutations within the *phyA* locus. An approximately 6.1 kb region containing a complete *phyA* coding region and promoter region was sequenced using 22 overlapping primers (Appendix C-2), and compared with the Col-0 *PHYA* sequence (At1g09570). The sequence of the *phyA*/sps-1 gene was identical to that of the

wild-type *PHYA* gene (data not shown). Therefore, the presence of a second-site (extragenic) mutation is most likely responsible for the release of *phyA'* silencing in the sps-1 line. However, since EMS treatment is known to modify the methylation profile of genomic DNA, the next step was to determine the *phyA'*/sps-1 methylation pattern, and rule out the release of *phyA'* silencing due to *phyA'* demethylation.

4.3c *The phyA' gene is hypermethylated in sps-1 line*

The methylation status of the *phyA'* allele in sps-1 line was tested using two approaches: Southern hybridization of genomic DNA treated with methylation-sensitive enzymes, and bisulfite sequencing. Southern analysis was performed on genomic DNA isolated from sps-1 (M3 and M4 plants), phyA-17, and Col-0 (WT). Genomic DNA digested with *Eco*RI and a methylation-sensitive enzyme, *Hha*I or *Hpa*II was blotted and hybridized with *phyA* probes (Fig. 29, 30). Of the seven *Hha*I sites (A1-A7) within the *Eco*RI fragment, two (A1 and A2) are located in the promoter region and the remaining five (A3–A7) in the coding region (Fig. 29a). Southern hybridization patterns of *Hha*I digested DNA suggested that methylation is found only in the coding region. In the wild-type *PHYA* locus, all *Hha*I sites in the coding region, except A6, are methylated as indicated by the presence of 4.3 and 0.8 kb bands, whereas in phyA-17, all sites except A5 are methylated as indicated by 4.8 and 0.3 kb bands. The *phyA'* in sps-1 line, on the other hand, showed a unique pattern defined by methylation in all *Hha*I sites located in the coding region (A3-A7) (Fig. 29b). Thus, on a Southern blot, a single band of 5.2 kb is seen with in sps-1 genomic DNA (Fig. 29b), indicating the presence of methylation over and above the phyA-17 and WT level.

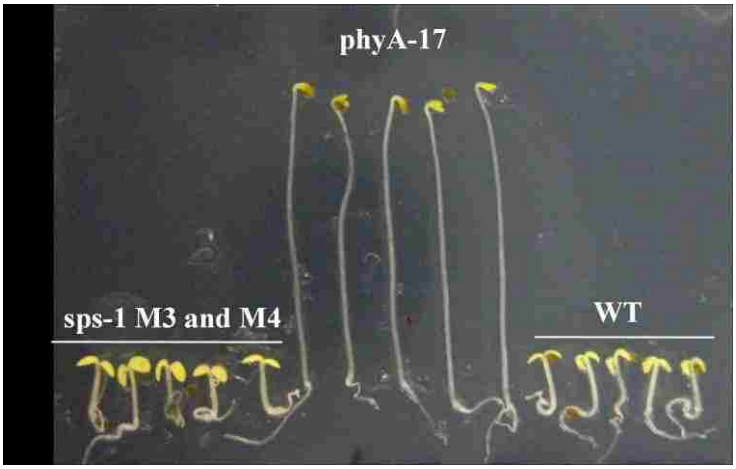


Figure 28: sps-1 maintained FRc phenotype in M3 and M4 generation sps-1 M3 and M4 seedlings displaying short phenotype along with phyA-17 epimutant and WT (Col-0) seedlings

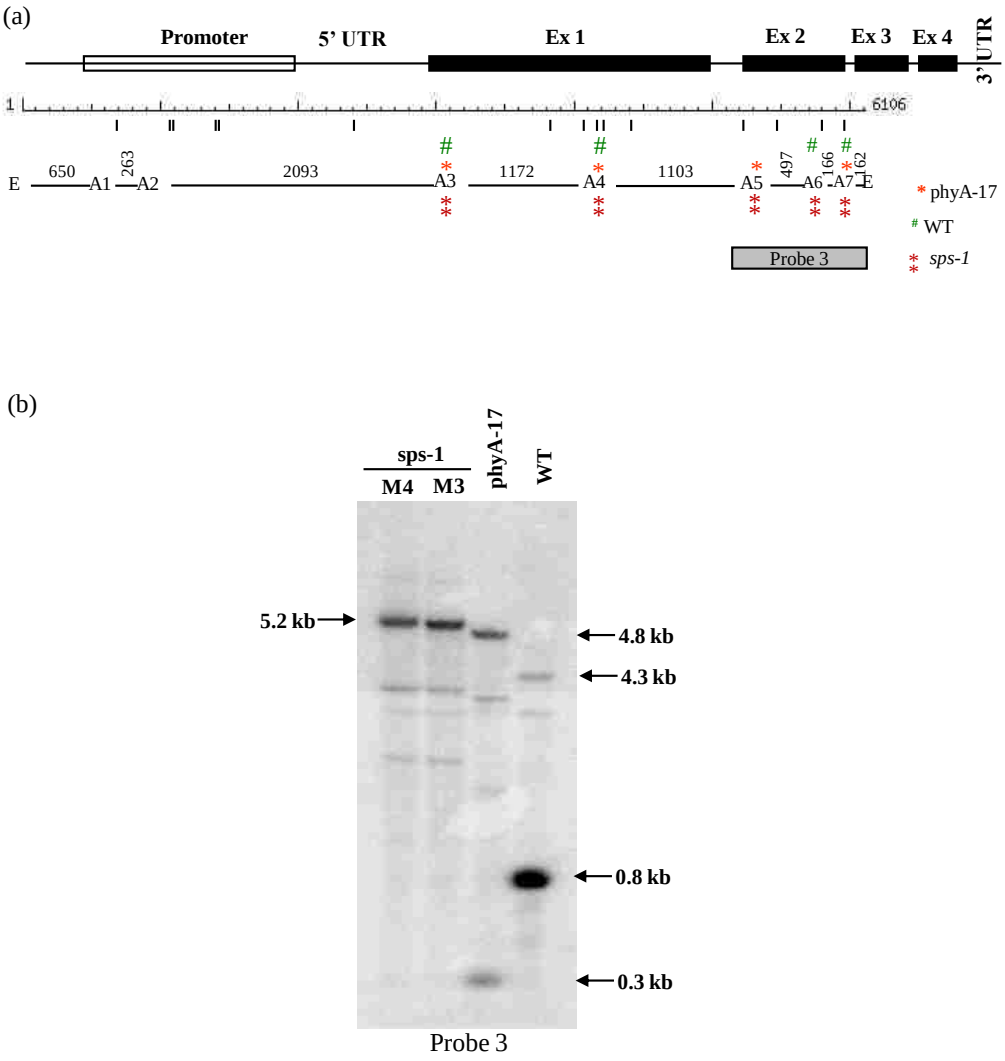


Figure 29: *phyA'* methylation in *sps-1* background based on Southern analysis (a) *Hpa*II and *Hha*1 map displaying methylated sites in phyA-17 (*), WT (#), and sps-1 backgrounds (**). Restriction sites for *Hpa*II (H), *Hha*II (A) and *Eco*RI (E) are shown, and fragment length between sites are given in bp (b) Southern analysis of *Eco*RI and *Hha*II digested genomic DNA of M3 plants of sps-1 and hybridized with probe 3.

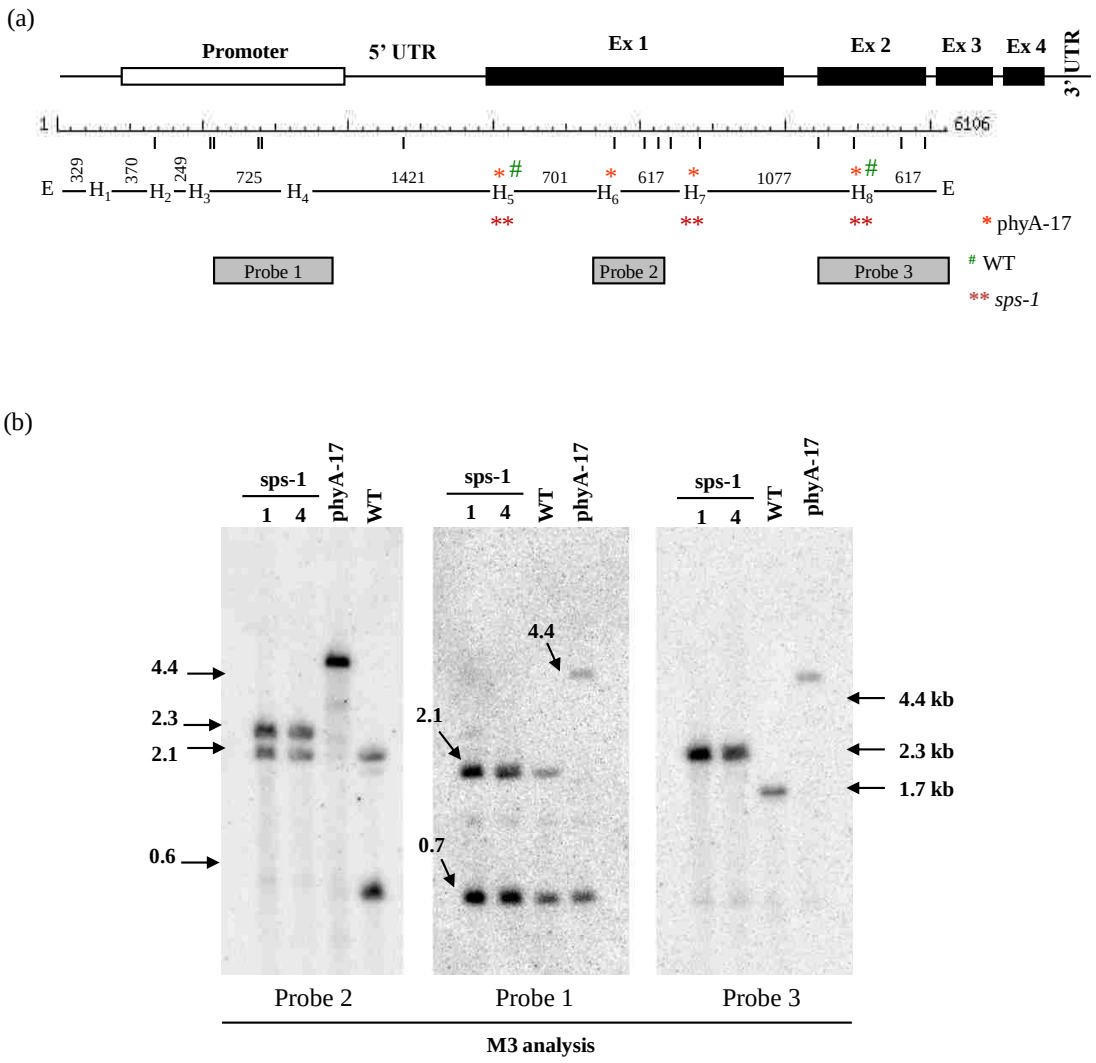
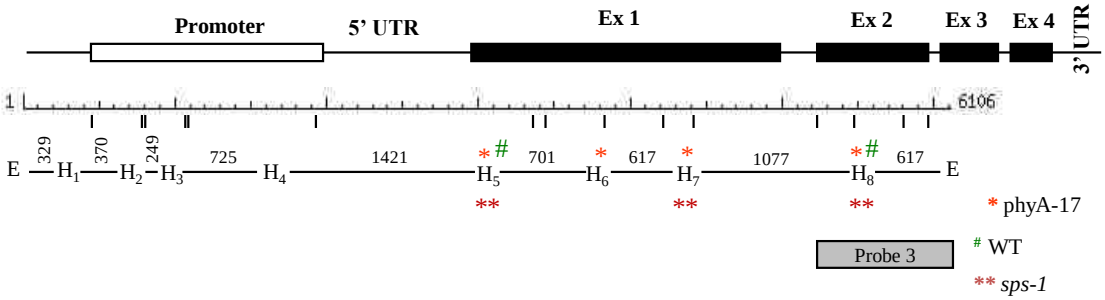


Figure 30: *phyA'* methylation in *sps-1* background based on Southern analysis
(a) *Hpa*II map displaying methylated sites in phyA-17 (*), WT (#), and sps-1 backgrounds (**). Restriction sites for *Hpa*II (H) and *Eco*RI (E) are shown, and fragment length between sites are given in bp. Probes used for Southern analysis are indicated below the map, (b) Southern analysis of *Eco*RI and *Hpa*II digested genomic DNA of M3 plants of sps-1

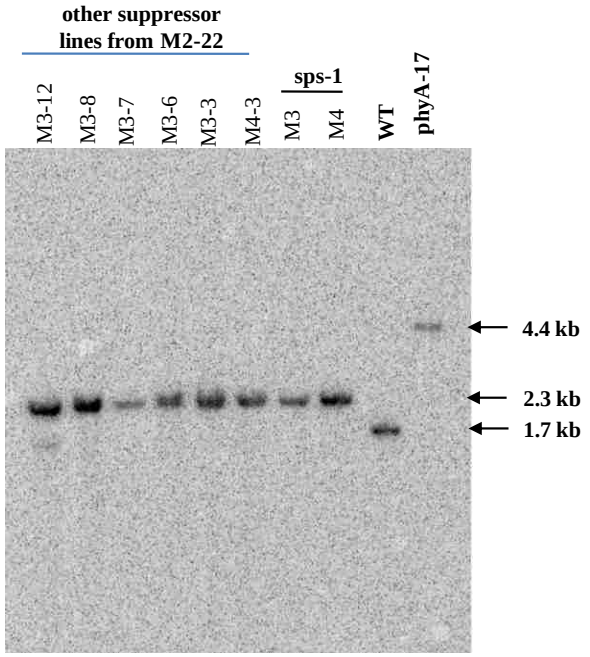
Southern analysis of *Eco*R1 + *Hpa*II digested genomic DNA of sps-1 M3 plants was done to analyze the methylation of *Hpa*II sites in the *phyA* locus using probe 1, probe 2 and probe 3 spanning the *phyA* locus (Fig. 30a). Of the eight *Hpa*II sites (H1-H8) within the *Eco*RI fragment, four sites (H5-H8) that are located in the exonic sequences were found to be methyated in *phyA'*, while the promoter region lacked methylation (Fig. 30a; Chawla et al., 2007). The wild-type *phyA* locus shows methylation of only two *Hpa*II sites, H5 and H8 (Fig. 30b). The *phyA'* /sps-1 again displayed a unique pattern indicative of the loss of methylation in H6 site while retaining methylation at three exonic *Hpa*II sites (H5, H7 and H8). As expected, hybridization with probe 1 generated a 4.4. kb band with phyA-17, but 2.1 and 0.7 kb bands with Col-0 and sps-1, respectively, indicating methylation of all exonic *Hpa*II sites in *phyA'* but lack of H6-methylation in the *phyA* locus found in sps-1 and Col-0. This finding was further confirmed by hybridizing the blot with two additional probes, probe 2 and probe 3 (Fig. 30b). These patterns also indicate that the promoter region in the *phyA* locus in all lines remains unmethylated. As many individuals in the M2-22 pool displayed a *phyA'* suppressor phenotype, Southern hybridization was done on randomly picked suppressor lines and their progeny derived from this pool. All of these lines showed an identical *Hpa*II methylation pattern, suggesting the isogenic nature of these lines. The *Hpa*II Southern analysis with probe 3 on five different suppressor lines (M3 generation) and progeny of a few of these lines is shown in Fig. 31.

Thus, the *phyA'* methylation pattern was maintained in the M4 generation of the sps-1 line (Fig. 31b). In conclusion, the *phyA'* locus in sps-1 line largely maintained the original methylation pattern associated with *phyA'* silencing.

(a)



(b)



Probe 3

Figure 31: *phyA'* methylation in *sps-1* background based on Southern analysis
(a) *Hpa*II map displaying methylated sites in phyA-17 (*), WT (#), and sps-1 backgrounds (**). Restriction sites for *Hpa*II (H) and *Eco*RI (E) are shown, and fragment length between sites are given in bp. Probes used for Southern analysis are indicated below the map. (b) Southern analysis of *Eco*RI and *Hpa*II digested genomic DNA of M3 and M4 progeny of sps-1 line and M3 progeny of other suppressor lines isolated from M2-22

Bisulfite sequencing: To develop the full methylation profile of the *phyA'* in sps-1 line, bisulfite sequencing of the coding region was done and compared to that of phyA-17 and Col-0 *phyA* alleles. Southern analysis of sps-1 revealed that *Hpa*II and *Hha*I sites present in the *phyA* promoter and 5' UTR are not methylated. Also, no cytosine methylation was detected in the 5' regulatory region of the *phyA'* epiallele in the previous study (Chawla et al., 2007). Therefore, only the coding region was selected for bisulfite sequencing. A map of CG methylation in the *phyA'* locus in sps-1 and phyA-17 line is shown in Fig. 32. Consistent with the Southern data, the methylation pattern of *phyA'* remains mostly conserved in the sps-1 line except for loss of H6-methylation. However, bisulfite sequencing revealed gain of methylation in several CG sites located in exon 1-3 (Fig. 32a). As a result, the net methylation level in *phyA'* is even higher in the sps-1 line (Fig. 32b).

Major points of the *phyA'* methylation profile found in sps-1 lines are given below:

(a) The methylation profile of *phyA'* in the sps-1 line is mostly conserved when compared to the phyA-17 allele

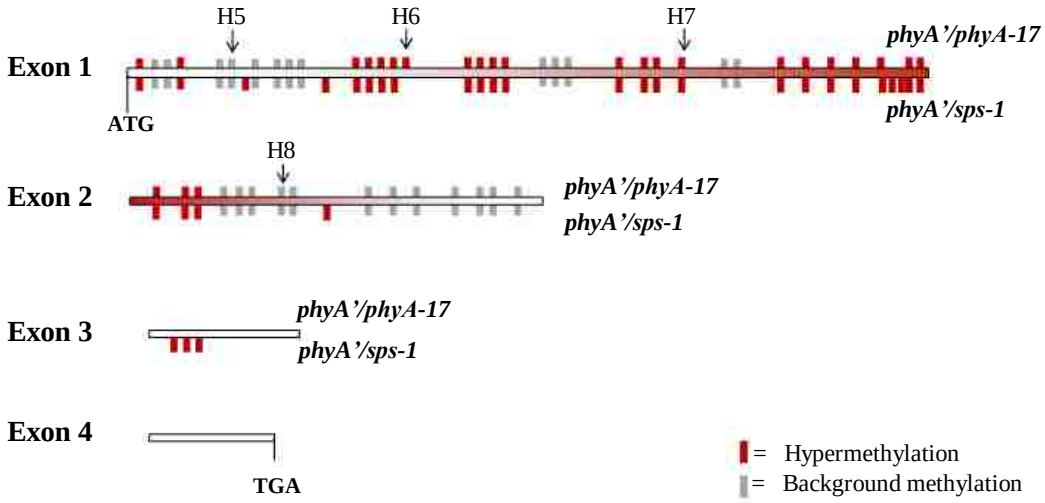
(b) Consistent with Southern data, only a single demethylation (that of H6 site) was detected in sps-1 *phyA'* locus

(c) Gain of methylation was revealed at several CG sites in exon 1, exon 2 and exon 3 (including within A6 site located in exon 2) (Bisulfite sequencing data is given in Appendix D)

(d) Gain of methylation was also confined to only CG sites

(e) No CHG or CHH methylation was detected

(a)



(b)

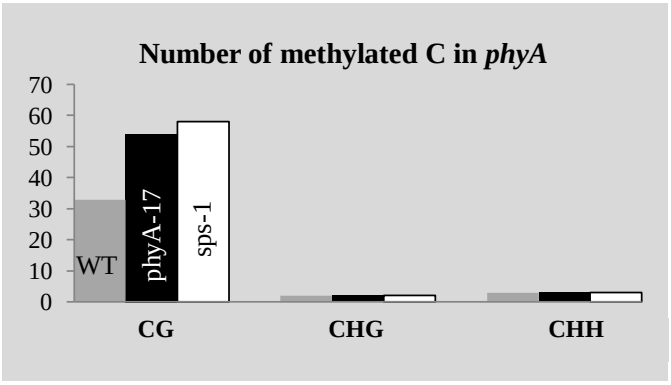


Figure 32: Methylation profile of the *phyA'* allele in the sps-1 line based on bisulfite sequencing (a) Gray bars represent background methylation found in WT (Col-0) allele. Red bars represent hyper-methylations found in phyA-17 line (above bar) or sps-1 line (below bar) backgrounds. *Hpa*II and *Hha*I sites are indicated (b) Graph displays the number of cytosine methylations in *phyA* gene of the three backgrounds: WT, phyA-17, and sps-1. No methylation is found in promoter / UTR / introns of *phyA'* in any of these genetic backgrounds

4.3d *Release of phyA' (transcriptional) silencing in sps-1 background*

To determine if phenotypic reversion of sps-1 seedlings originated from the release of *phyA'* silencing, Northern analysis was performed. The *PHYA* transcript is abundant in dark-grown tissues and is rapidly down-regulated by light; therefore, steady state level of mRNA was measured in dark-grown seedlings. Total RNA from sps-1 M3 and M4 pooled seedlings was hybridized with probe 3 on a Northern blot and compared with that of Col-0 and phyA-17 (Fig. 33). This analysis clearly demonstrated that *phyA'* transcription is reactivated in the *sps-1* background (Fig. 33). Analysis of hybridization intensities using ImageQuant^{TM} software indicated that *phyA'* in sps-1 is expressed at 73 - 133% as compared to the wild-type *PHYA* expression level, while phyA-17 line shows only 15-20% *phyA* expression relative to the wild-type level (Fig. 33). Thus, the phenotypic reversion of sps-1 was accompanied with the release of *phyA'* silencing.

This analysis rules out the presence of the light-signal transduction suppressor *spa1*. SPA1 is a negative regulator of PhyA-specific light signaling, which affects the phenotype without altering *PHYA* transcription rate (Hoecker et al., 1999). As a result, in the absence of SPA1, seedlings display WT phenotype independent of PhyA level. However, WT phenotype of sps-1 under FRc is accompanied with the accumulation of *PHYA* transcripts, eliminating the possibility that *sps-1* is allelic to *spa1*.

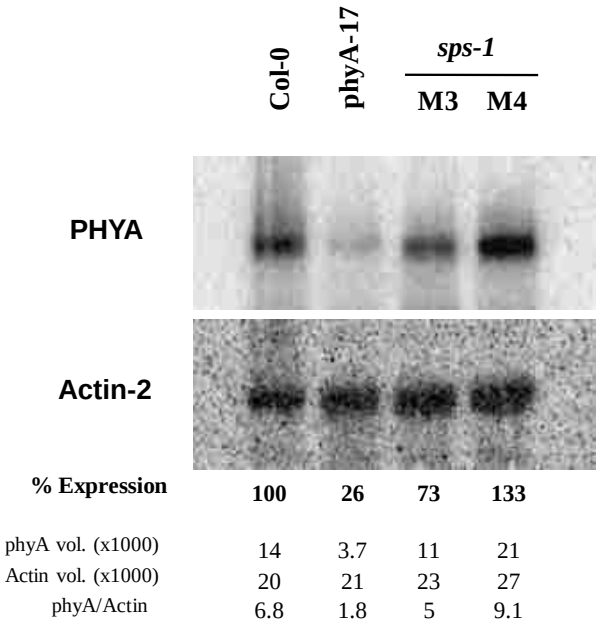


Figure 33: Northern analysis showing reactivation of *phyA* gene in *sps-1* background in M3 and M4 generations Total RNA (5 μg) isolated from the dark grown seedlings of wild type (WT), phyA-17 and sps-1 (M3 and M4 progeny) and hybridized with probe 3 spanning the coding region (top) and Actin, a loading control (bottom). % expression of PHYA in each line is given below each lane

4.3e *Methylation analysis of selected heterochromatin loci*

To study the effect of the *sps-1* mutation on other genetic loci, three well-known heterochromatic and epigenetically controlled loci were chosen: a 180 bp centromeric repeat, 5S RNA, and Ta2. First, the methylation status of these loci was checked on a Southern blot containing *Hpa*II digested genomic DNA from pooled seedlings of sps-1 (M4 generation), Col-0 and a *ddm1* mutant line (as a positive control) (Fig. 34). With the 180 bp single repeat probe, weak ladder formation was observed in the sps-1 line when compared to the WT (Col-0) and ddm1 genomic DNA, indicating the occurrence of hypomethylation in the centromeric repeat sequences (Fig. 34a). However, as a similar level of hypomethylation was also found in the parental line, phyA-17 (Chapter 2, Fig. 8), no effect of the *sps-1* mutation on centromeric repeats occurred. Similarly, the 5S ribosomal repeat and Ta2 loci were also found to be unchanged in the sps-1 line when compared to the WT genome (Fig. 34b and 34c). As expected, all of these loci were found to be hypomethylated in the ddm1 genome.

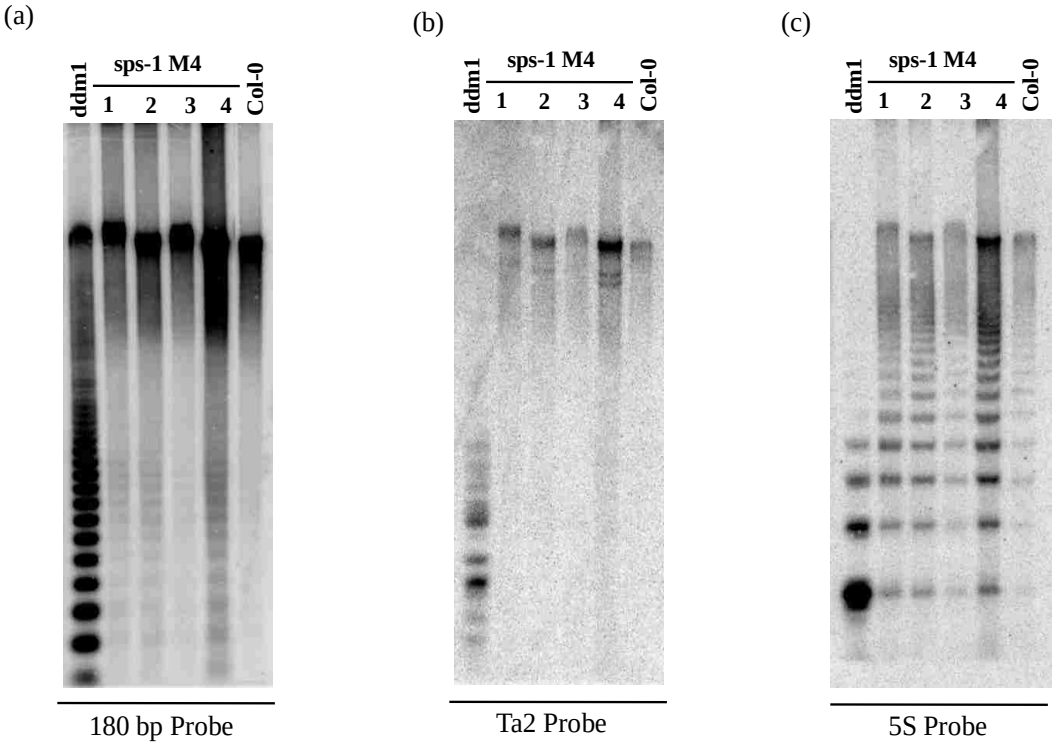


Figure 34: methylation analysis of selected heterochromatic loci in sps-1 by Southern analysis Genomic DNA was isolated from pooled seedlings of the indicated genotype, digested with *Hpa*II and *Eco*R1 and methylation status was determined at three loci; (a) Centromeric repeats using 180bp single repeat probe (b) Retrotransposon, Ta2 using Ta2 probe, and (c) 5S ribosomal genes using 5S probe

4.3f *Evidence for the presence of an extragenic sps-1 mutation*

The genetic (DNA sequencing) and epigenetic (bisulfite sequencing) profile of *phyA'*/sps-1 showed no major change in the *phyA'* locus, and yet transcriptional silencing is released in the sps-1 line. However, demethylation of a single CG site (within the H6 *Hpa*II site) was detected in *phyA'*/sps-1 when compared to phyA-17 line. These observations suggest two possibilities for the release of *phyA'* silencing in sps-1line: (a) the presence of a trans-acting mutation (extragenic) in the sps-1 line is responsible for the release of *phyA'* silencing or (b) loss of H6 methylation alone is responsible (intragenic mutation) for *phyA'* activation. However, the analysis of F2 seedlings derived from an outcross of sps-1 line with Col-0 line and phyA-17 indicated that the *sps-1* mutation lies outside of *phyA* locus as described below.

Crosses of sps-1 with Col-0: Two F1 plants obtained by crossing sps-1 with Col-0 were genotyped as true hybrids by Southern analysis. Genomic DNA from F1 plants was digested with *Eco*R1 and *Hpa*II enzymes, and hybridized with probe 3, which revealed the presence of both *phyA* alleles i.e. *phyA'*/sps-1 and wild-type *PHYA* (Fig. 35b). F1 plants were grown to maturity and allowed to self-fertilize. The resulting F2 progeny were subjected to genotypic and phenotypic analysis. Appearance of the *phyA'* phenotype (long hypocotyl in FRc grown seedlings) in F2 population would indicate the presence of an extragenic mutation in the sps-1 line. Among ~1200 F2 individuals derived from two separate F1 parents, 15 intermediate-length (Int) seedlings were recovered, but no tall seedlings equivalent to phyA-17 were found. Four of these intermediate-length seedlings were rescued successfully and grown till maturity.

Southern hybridization on these F2 plants showed the presence of a 2.3 kb band that is specific to the *phyA'* locus in the *sps-1* background (Fig. 35c). Appearance of a weak *phyA'* phenotype at a low rate in the F2 population indicates that introduction of wild-type SPS-1 does not impose a strong effect on the *phyA'* locus. In other words, the effect of wild-type SPS-1 is not fully heritable. The effect of certain epigenetic factors such as DDM1 is not imposed on their target loci in a single generation (Bartee and Bender 2001), but it may become apparent after a few generations. Therefore, F3 progeny derived from the four intermediate-length F2 plants and 15 randomly picked F2 individuals were phenotyped under FRc light. Nine of these 19 F3 populations displayed partial reversion to intermediate-length seedlings at a high rate (>50%). A representative F3 seedling displaying intermediate-length phenotype is shown in Fig. 36a. A few seedlings from each F3 population were rescued and grown to collect F4 seeds. FRc phenotyping of the F4 population also showed intermediate-length seedlings at a high rate but did not display further increase in the hypocotyl length. This phenotype was maintained at a similar rate in the F5 generation. Presence of a weak *phyA* phenotype among F4 and F5 seedlings was independently verified by Prof. Ute Hoecker (from University of Köln, Germany), an expert in photobiology (Fig. 36b). An increase in the seedling length in the progeny of sps-1 x Col-0 was estimated to be ~2x compared to the parents (Col-0 and sps-1) (Fig. 36a and 36b). As no differences were found in the dark-grown seedlings (Fig. 36b), the increase in seedling length in FRc should be based on *phyA* gene expression or regulation (Park and Quail, 1993; and reviewed by Whitelam and Devlin, 1997). Thus, appearance of intermediate-length seedlings upon backcrossing with wild-type Col-0 is suggestive of *phyA'* suppression.

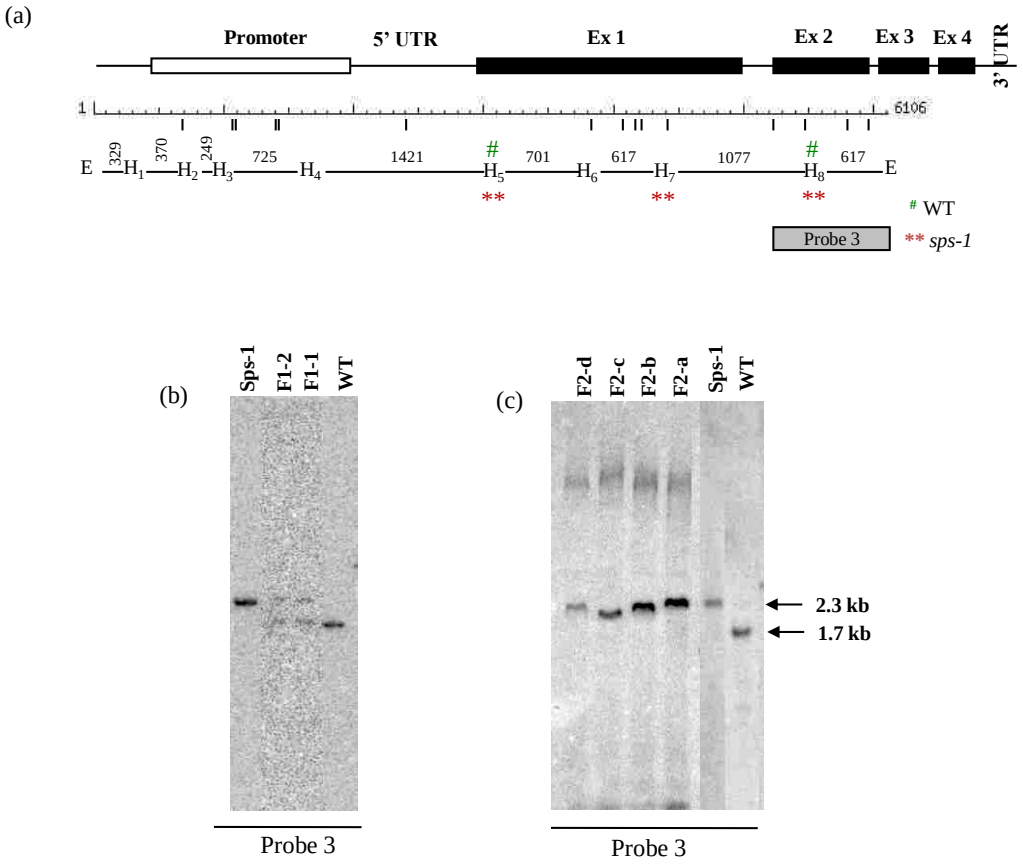
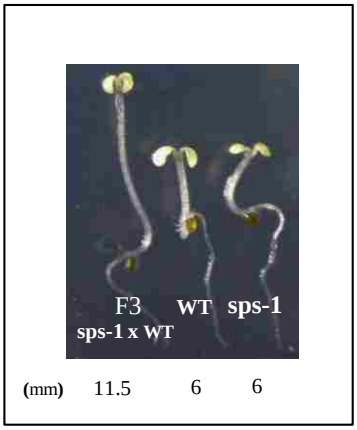


Figure 35: *phyA'* methylation analysis of F1 and F2 progeny derived from Col-0 x sps-1 by Southern analysis (a) *Hpa*II map displaying methylated sites in WT (#), and *sps-1* (**). Southern analysis of *Eco*RI and *Hpa*II digested genomic DNA of (a) two F1 plants (b) Four F2 plants (FRc phenotyped as intermediate), hybridized with probe 3

(a)



(b)

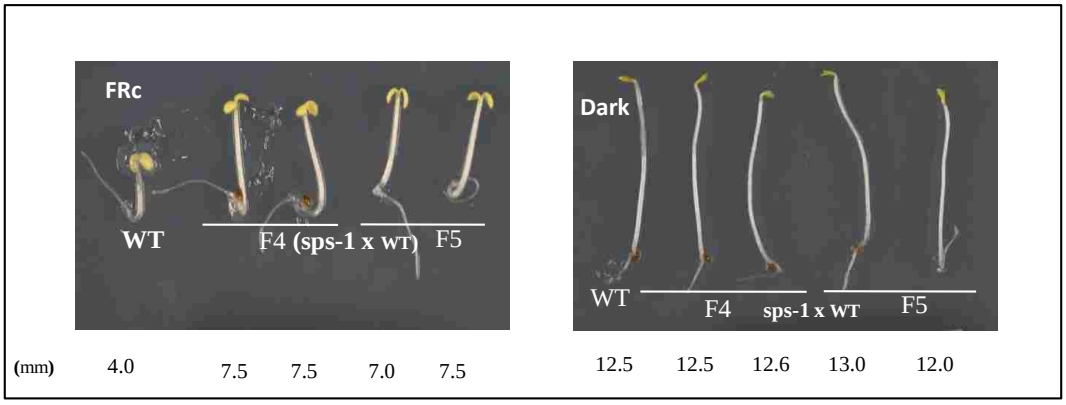
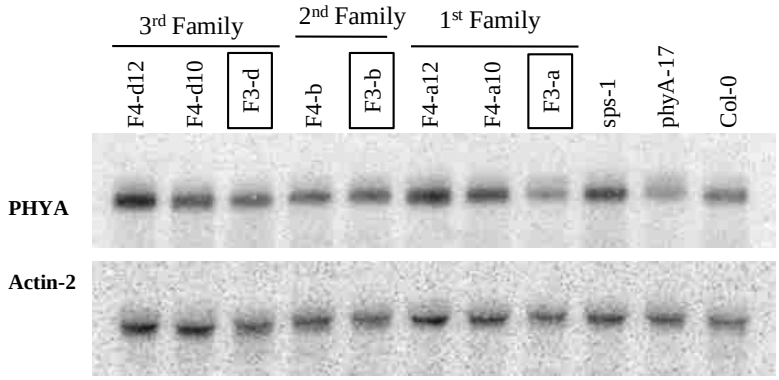


Figure 36: Appearance of weak *phyA* mutant phenotype in F3, F4 and F5 progeny Representative *phyA* mutant phenotype in (a) F3 seedlings derived from F2 plant (WT x sps-1) under FRc light (b) Increased length of F4 seedling compared to WT (Col-0) under FRc, and dark grown seedlings do not display any phenotypic differences ruling out unrelated effects (Photo provided by Prof. Ute Hoecker from University of Köln, Germany)

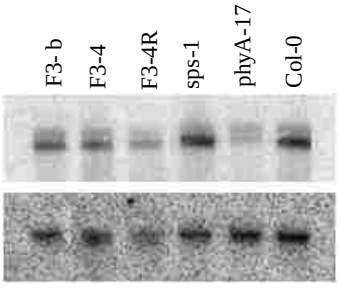
Northern analysis on five F3 families and their F4 derivatives was done to analyze *phyA* expression levels (Fig. 37). Relative expression was calculated by comparing hybridization intensities on ImageQuant. This analysis showed that *phyA'* in most F3 and F4 progeny derived from sps-1 x Col-0 was partially repressed to a level intermediate to phyA-17 and sps-1 lines (Fig. 37a and 37b). PHYA expression in different F3 families and their F4 progeny varied significantly, and some of them expressed at a level equivalent to Col-0. This is quite surprising because the phenotype of the F3 and F4 seedlings was significantly different from that of Col-0, yet expression differences were not consistently observed. However, this analysis showed that while *phyA* expression varied between F3 and F4 families, some level of *phyA* repression occurred when the *phyA'*/sps-1 locus was introduced into the wild-type background. Since pooled seedlings (F3 or F4 generations) were used in Northern analysis, expression variation could be based on the variation within the population. Additional real-time PCR-based analysis using RNA from individual seedlings is required to validate the finding of Northern blot.

(a)



	F4-d12	F4-d10	F3-d	F4-b	F3-b	F4-a12	F4-a10	F3-a	sps-1	phyA-17	Col-0
% Expression	**100**	**59**	**78**	**72**	**87**	**84**	**87**	**50**	**100**	**34**	**84**
PHYA vol. (x1000)	91	53	44	44	46	70	58	24	60	20	32
Acttin vol. (x1000)	28	27	17	18	16	25	20	15	19	17	12
PHYA/Actin	3.2	1.9	2.5	2.3	2.8	2.7	2.8	1.6	3.2	1.1	2.7

(b)



	F3- b	F3-4	F3-4R	sps-1	phyA-17	Col-0
% Expression	**72**	**57**	**43**	**100**	**21**	**79**
PHYA vol. (x1000)	73	63	37	11	30	11
Acttin vol. (x1000)	21	23	19	24	28	30
PHYA/Actin	3.4	2.7	2	4.7	1	3.7

Figure 37: Northern analysis of F3 and F4 progeny derived from Col-0 x sps-1
Two separate northern blots (a) and (b) showing repression (see % expression below each lane) of *phyA* gene in F3 and some of the F4 progenies (WT x sps-1)

Crosses of sps-1 with phyA-17: To study the effect of sps-1 mutation on the naïve *phyA'* locus, sps-1 plants (M2 – M4 generation) were crossed with phyA-17, and 5 F1 hybrids were obtained. Southern blot analysis was done on 4 of these F1 plants that confirmed the presence of two distinct *phyA* alleles (Fig. 38b). Genomic DNA from F1 plants was digested with *Eco*R1 and *Hpa*II enzymes, and hybridized with probe 3, which revealed the presence of 2.3 kb and 4.4 kb bands indicating the presence of two separate *phyA'* alleles originating from phyA-17 and sps-1 lines (Fig. 38b).

The F2 populations derived from the selfed F1 plants were phenotyped in FRc light to study the effect of *sps-1* mutation on the naïve *phyA'* locus characterized by the 4.4 kb *Hpa*II band. Three different kinds of phenotypic segregation were observed in the F2 analysis as given below:

(i) Three populations (F1-1, F1-2 and F1-3) displayed a strong interaction of the *sps-1* mutation and the naïve *phyA'* locus as less than a quarter of seedlings were found to be tall in these populations (Table 6). The F2 population derived from the F1-1 parent generated 27 short, 44 intermediate and 10 tall seedlings in a total of 81 F2 progeny. This segregation ratio is significantly different from that of F2 derived from phyA-17 x Col-0 (WT) crosses. Chawla et al. (2007) reported a 1:2:1 segregation ratio for short, intermediate and tall seedlings among F2 population derived from phyA-17 x Col-0 crosses. Among 307 F2 seedlings grown under FRc, they found 80 tall, 150 intermediate and 77 short seedlings, which fits the 1:2:1 or 3:1 segregation ratio. The F2 progeny of sps-1 x phyA-17 crosses, on the other hand, behaved differently. Similar to F1-1,

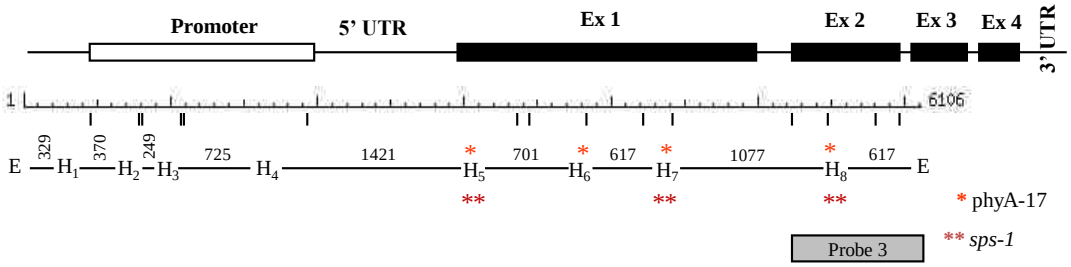
progeny of F1-2 and F1-3 also displayed higher number of short/intermediate seedlings than found in phyA-17 x Col-0 crosses.

(ii) One population (F1-4) consisting of 219 seedlings displayed 178 short and 41 tall seedlings, fitting a ratio of 13:3 for short and tall seedlings. This segregation ratio suggests a direct interaction of *sps1* with *phyA'* locus and phenotypic reversion of the double homozygous F2 individuals.

(iii) One population (F1-5) that was derived from the early generation (M2) sps-1 plant did not show a significant interaction as more than a quarter of seedlings were found to be tall in this population

Distinct segregation ratio in different populations is a characteristic of epigenetic inheritance. The F2 segregation ratios observed in four populations (F1-1 to F1-4) indicated that introduction of *sps-1* mutation releases silencing of the naïve *phyA'* locus. Only one population (F1-5) failed to display interaction between *sps1* and *phyA'*. One explanation could be that this population is derived from the early generation (M2) sps-1 plant, whereas advanced generation (M3 and M4) sps-1 plants were used for generating F1-1 to F1-4 hybrids. These data suggest that the introduction of *sps-1* mutation in phyA-17 line caused phenotypic reversion as a result of the release in *phyA'* silencing.

(a)



(b)

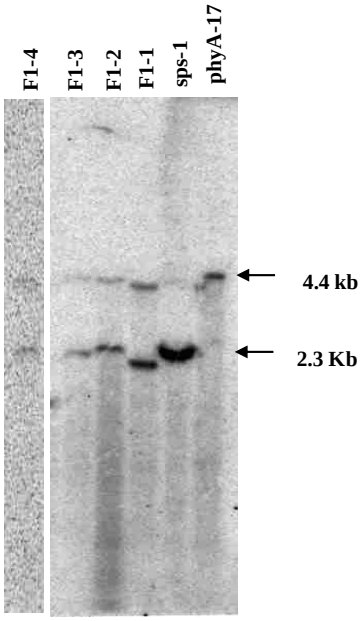


Figure 38: *phyA'* methylation analysis of F1 plants obtained by crossing suppressor lines with phyA-17 by Southern hybridization (a) *Hpa*II map displaying methylated sites in phyA-17 (*), and *sps-1* (**). (b) Southern analysis of *Eco*RI and *Hpa*II digested genomic DNA of four F1 plants hybridized with probe 3.

Table 6: Segregation analysis of F2 (sps-1 x phyA-17)

F1 parent	Crosses	F2 segregation ratio			$X^2_{(3:1)}$	$X^2_{(13:3)}$
		Short+Int (S + I)	Tall	Total		
1	sps-1 (M4) x phyA-17	71 (27 S + 44 I)	10	81	7.1_{NS}	1.7_{NS}
2	phyA-17 x sps-1 (M3)	173	28	201	6.2_{NS}	2.8_{NS}
3	phyA-17 x sps-1 (M3)	29	5	34	1.9_{NS}	0.3_{NS}
4	sps-1 (M3) x phyA-17	178 (71 S + 107 I)	41	219	4.3_{NS}	0.02_{S}
5	sps-1 (M2) x phyA-17	606	235	841	3.9_{NS}	47.4_{NS}

NS : Non significant to accept the hypothesis
S: significant $\chi 2$ value to accept the hypothesis

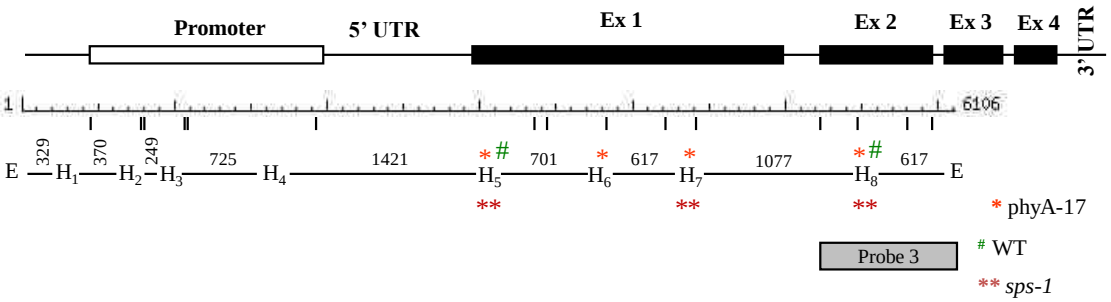
To determine the genotype of F2 revertants, a number of short or intermediate F2 seedlings were rescued and grown to maturity. Genomic DNA isolated from each of the F2 revertants was digested with *Eco*R1 and *Hpa*II enzymes and hybridized with probe 3 (Fig. 39). The hybridization pattern obtained on short F2 revertants revealed:

(i) short F2 revertants do not contain a uniform hybridization pattern. Most of the F2 lines contained a single *phyA'* allele (~2.3 kb band) (sb, sd, sl, sm, sn); however a number of them contained an additional *phyA'* allele (~4.4 kb) (sa, sc, sf, sh) (Fig. 39b),

(ii) One F2 plant (sj) displayed a unique *phyA'* allele that did not match with either of the parental alleles (2.3 kb or 4.4 kb bands) (Fig. 39b).

As the *phyA'* (~4.4 kb) allele in the wild-type background confers a semi-dominant effect on the seedling phenotype (heterozygotes display intermediate-length seedling), the association of 4.4 kb band with short phenotype indicates epigenetic modification of the naïve *phyA'* locus in the *sps-1* background.

(a)



(b)

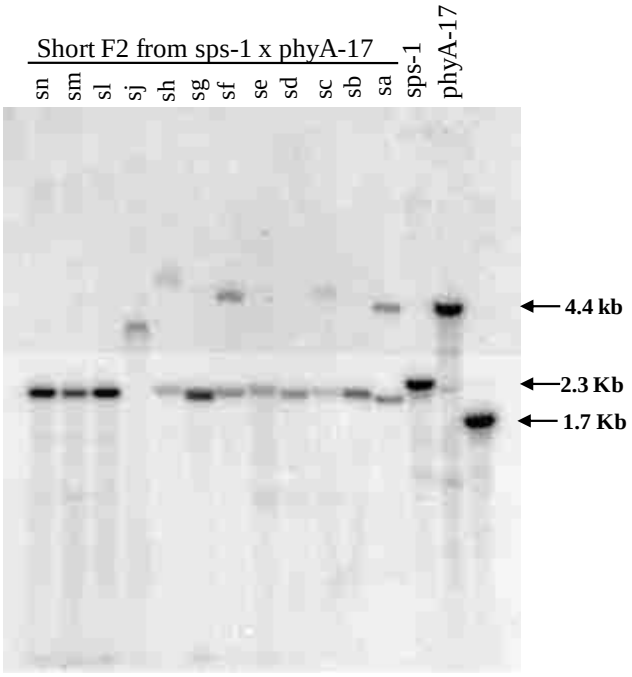


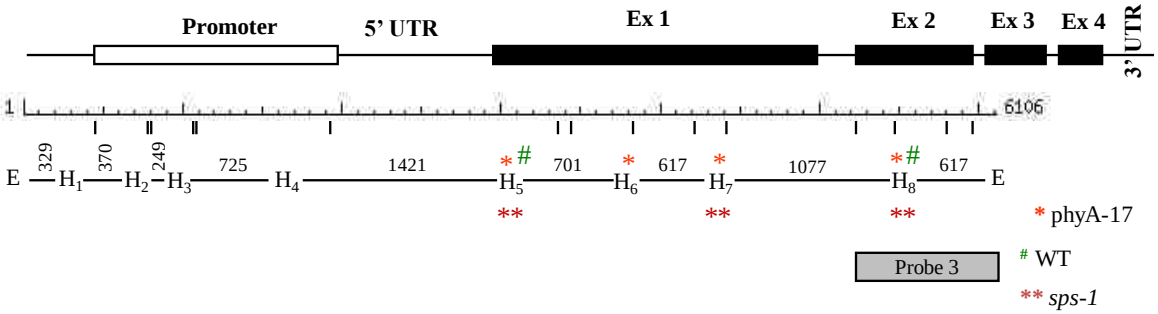
Figure 39: *phyA'* methylation analysis of short F2 plants (sps-1 x phyA-17) by Southern analysis (a) *Hpa*II map displaying methylated sites in WT (#), phySA-17 (*) and *sps-1* (**). (b) Southern analysis of *Eco*RI and *Hpa*II digested genomic DNA of twelve short F2 plants using probe 3.

The hybridization pattern of intermediate F2 revertants also varied (Fig. 40); however, most of them contained the expected two alleles (~4.4 kb and ~2.3 kb bands). Three individuals in this pool (id, if, ii) contained only a single 2.3 kb band, which is characteristic of *phyA'*/sps-1 line that display short phenotype (Fig. 40b).

To study if the two *phyA* alleles (~4.4 kb and 2.3 kb) segregate independently in the next generation, F3 progeny derived from selfed F2 parents was analyzed (Table 7). The F3 progeny derived from the biallelic F2 parents generated a ratio that was significantly different from 1:2:1 for short, intermediate and tall. Each biallelic F2 line generated higher number of short F3 progeny than predicted by Mendelian law.

In summary, genetic and molecular analysis suggest that the naïve *phyA'* locus (4.4 kb band) is modified in the *sps-1* background. This modification involves alteration of the methylation pattern of the locus. Specifically, exonic *Hpa*II site H6 (Fig. 39) is demethylated in the *sps-1* background, which is associated with the release of *phyA'* silencing. Thus, *sps-1* is likely to encode a trans-acting epigenetic factor.

(a)



(b)

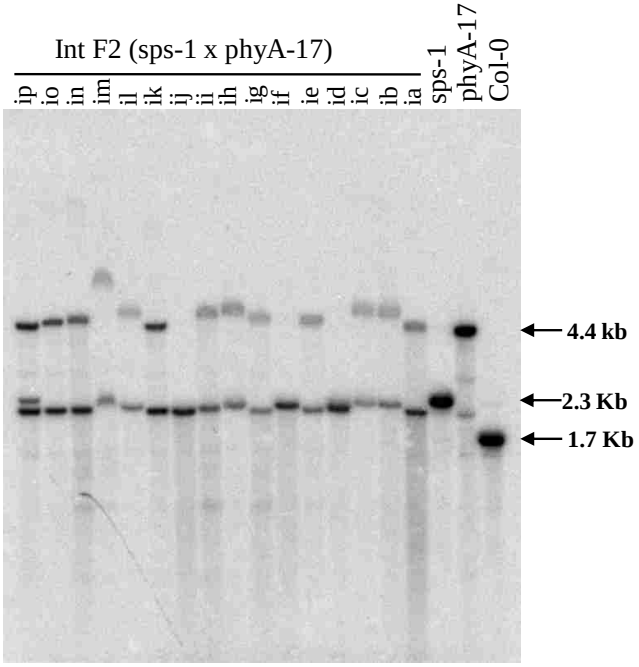


Figure 40: *phyA'* methylation analysis of intermediate F2 plants (sps-1 x phyA-17) by Southern analysis (a) *Hpa*II map displaying methylated sites in WT (#), phyA-17 (*) and sps-1 (**). (b) Southern analysis of *Eco*RI and *Hpa*II digested genomic DNA of fourteen intermediate F2 plants using probe 3.

Table 7: Segregation analysis of F3 from selected F2 (sps-1 x phyA-17)

F2 parent		F3 segregation ratio				$\chi 2_{1:2:1}$
Short/Int	**Genotype**	**Short**	**Int**	**Tall**	**Total**	
Sa	Biallelic	22	10	2	34	$29.25_{(NS)}$
Sb	Monoallelic	67	-	-	67	-
Sf	Biallelic	24	28	14	66	$4.52_{(NS)}$
Sj	Unique monoallelic	29	-	-	29	-
Sm	monoallelic	26	-	-	26	-
Ia	biallelic	34	24	22	80	$16.4_{(NS)}$
Ib	biallelic	80 (s+I)		11	91	$7.25_{(NS)}$
Ie	biallelic	9	6	-	15	$3.75_{(NS)}$
If	monoallelic	35	-	-	35	-
Ii	biallelic	12	2	1	15	$23.46_{(NS)}$
Ik	biallelic	8	8	4	20	$2.45_{(S)}$
Im	biallelic	14	10	5	39	$12.1_{(NS)}$

NS : Non significant $\chi 2$ value to accept the hypothesis
S: significant $\chi 2$ value to accept the hypothesis

4.3g *Effect of the sps-1 mutation on the selected genomic targets*

Genomic loci, other than *phyA*, may also be affected by the *sps-1* mutation. To investigate if targets of DDM1 and MOM1 are up-regulated in the *sps-1* background, transcriptionally Silent Information (*TSI*) element, the 106B long terminal repeat (LTR)-like dispersed repeats, the retrotransposon Ta3, and the 180 bp centromeric repeat were analyzed in the sps-1 line. The 180 bp region was found to be hypomethylated in sps-1 as well as phyA-17, but the methylation level of other loci in sps-1 line was not analyzed. Since, DDM1 and MOM1 regulate their targets in distinct fashion representing separate epigenetic mechanisms (Amedeo et al., 2000), genomic targets of both factors were included. In *ddm1* and *met1* backgrounds, all of the above targets are upregulated, while in *mom1* background only *TSI* is strongly upregulated (Valliant et al., 2006). The expression of the selected targets was analyzed by semi-quantitative reverse transcription–PCR (RT-PCR) using total RNA (Fig. 41). In this assay, ddm1, met1, and phyA-17 lines served as controls.

RT-PCR analysis using primers for each target revealed expression from the respective positive controls (ddm1 and met1). As expected, RT-PCR analysis of the 180 bp repeat revealed expression in sps-1 and phyA-17 lines. Each of these lines displayed hypomethylation of CG sites in 180 bp repeats (Chapter 2, Fig. 8; Southern data for sps-1 hypomethylation is not given). However, the ddm1 line shows much greater hypomethylation of the 180 bp repeat (Chapter 2, Fig. 8). Accordingly, higher expression of the 180 bp transcript was seen in ddm1 compared to sps-1 or phyA-17 (Fig. 41A). The remaining targets, *TSI, 106B* and *Ta3*, were not found to be expressed in the sps-1 line, while they were expressed in ddm1 or met1 (Fig. 41B, C and D).

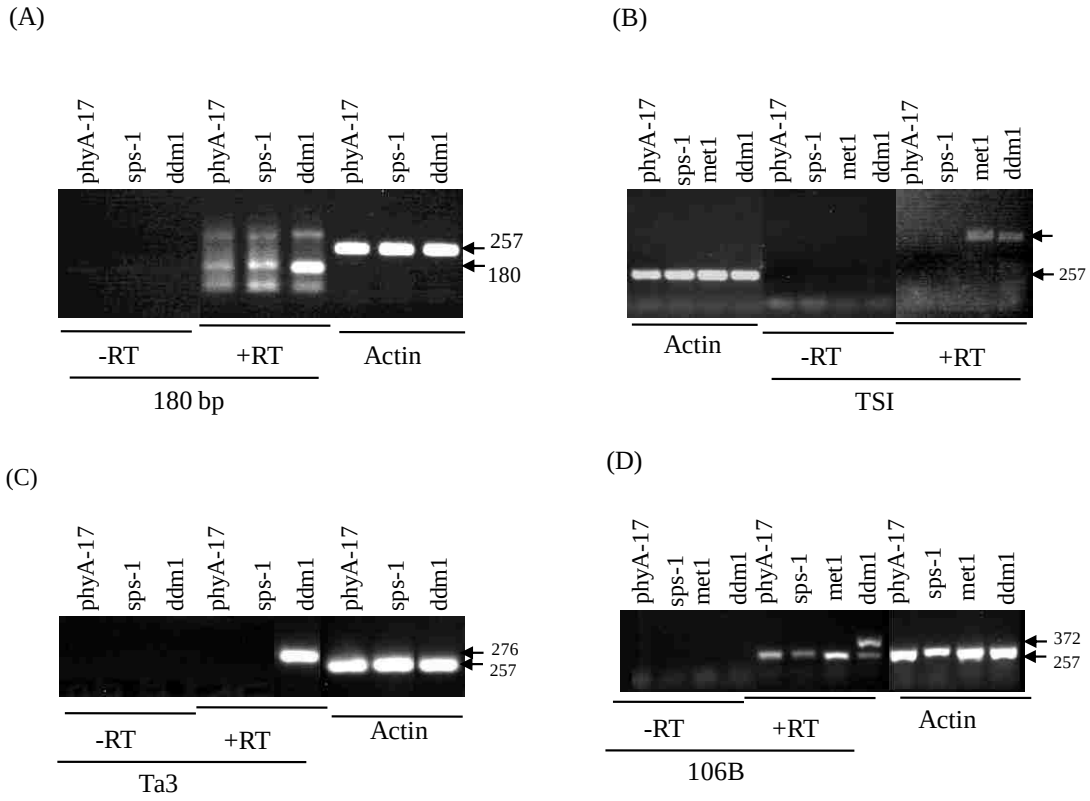


Figure 41: Analysis of selected mom-1 targets in the sps-1 line

Transcriptional analysis of the (A) 180 bp centromeric repeats (B) *TSI* element (C) *Ta3*, a retrotransposon, (D) *106B* LTR-like repeats performed by semi quantitative reverse transcription–PCR (RT–PCR). Expression of *ACTIN 2* was used to normalize the amounts of RNA template. Negative controls lacked reverse transcriptase (no RT). The size of the amplicons is indicated on the right.

This analysis indicates that the underlying epigenetic mechanism of *phyA'* suppression is distinct and does not overlap with that of DDM1 or MOM1. While DDM1 and MOM1 targets include repetitive sequences and heterochromatic regions, targets of *SPS-1* may be limited to euchromatic regions consisting of expressed genes.

4.3h *Candidate gene analysis in the sps-1 line*

Previous analysis showed that *sps-1* is a trans-acting mutation that releases transcriptional silencing of *phyA'* without significantly changing its methylation level. However, demethylation at one of CG site (*Hpa*II site, H6) within *phyA'* coding region is found to be associated with release of *phyA'* silencing in the sps-1 line. On the basis of the phenotype (normal) and demethylation of H6 site, three genes were selected as likely candidates for *sps-1* gene; *MOM1*, *MET1*, and *VIM1*. Following are the reasons for selection of candidate genes: (a) mutation in the *MOM1* gene releases transcription silencing of various genomic loci without significantly changing the methylation pattern of the respective loci (Amedeo et al., 2000; and Mittelsten-Scheid et al., 2002), (b) the phenotype of *mom1* and *vim1* mutant plants are normal and fertile (Amedeo et al., 2000; Woo et al., 2008) (c) transcriptional silencing of *phyA'* epiallele is MET1 dependent (Chawla et al., 2007); therefore, ruling out mutation in *MET1* in sps-1 line was important, (d) VIM1 has a role in maintaining CG methylation in the locus dependent manner (Woo et al., 2007), and plays an important role in the MET1-meditated DNA methylation pathway.

DNA sequencing of the entire coding region (including introns) of the candidate genes in the sps-1 line was carried out, and compared with the parental line, phyA-17 and

Col-0. Overlapping primers were designed for each locus. Location of primers across *VIM1*, *MOM1* and *MET1* is represented in Appendix E. By comparing the sequencing profile of *MOM1*, *MET1* and *VIM1* genes in sps-1 line with that of phyA-17 and Col-0, no mutation was detected in the coding region. Thus, a mutation in these genes is not likely to be associated with *sps-1*.

4.3i *Microarray analysis*

To identify the possible genomic targets of the putative epigenetic factor, SPS-1, we took a global approach based on microarray analysis. Microarray analysis was carried out using mRNA collected from dark grown seedlings. Since PHYA is mostly involved in FR light responses, dark grown tissue is not expected to display differentially expressed *PHYA*-regulated genes. Three different genotypes (all in Col-0 background) were used: (1) *phyA'* suppressor mutant line, sps-1, (2) parent of sps-1, phyA-17, and (3) WT (Col-0). Each genotype (sample) was used in three replicates.

The following steps were carried out by Dr. Mariya Khodakovskaya at the University of Arkansas, Little Rock. Complimentary RNA (cRNA) was synthesized with an Affymetrix IVT Express Kit according to manufacturer's instructions using 500 ng of total RNA for each sample. Following second strand cDNA synthesis, the double stranded cDNA was purified, which served as a template in an in vitro transcription reaction in the presence of a biotinylated nucleotide analog/ribonucleotide mix. The biotinylated cRNA targets were purified and fragmented. Hybridization to the Affymetrix Gene Chips (*Arabidopsis* ATH1 Genome Arrays) was carried out by Expression Analysis Inc. (Durham, NC). The raw fluorescence intensities for all

samples were normalized. Statistical analysis by hierarchical clustering of log (2) fold changes was performed using TM4 Microarray Suite from TIGR (http://www.tm4.org/mev.html). The graphical representation was generated in the form of a heat map (Fig. 42). Columns on the heat map represent samples and rows represent genes, aligned according to the results of the cluster algorithm. The intersection of a gene and sample is colored according to its expression value; red indicates high expression, and green indicates low expression (Fig. 42). Three-way analysis of variance (ANOVA) (p= 0.0001) gave a list of approximately 448 genes which were differentially expressed between these lines with high consistency across replicates (Fig. 42). Important observations based on the transcriptome analysis are given below:

(i) One of the genome-wide effect of EMS is to induce random methylation at CG sites, and thereby alter the gene expression (Farrance and Ivarie 1985; Ivarie and Morris, 1986). Unlike the random effect, which would be uncontrolled and non-specific, transcriptome analysis of sps-1 line indicated altered gene expression at very few and specific loci (euchromatic loci). Three-way ANOVA displayed few genes (~68 out of 448), which were differentially expressed in sps-1 line compared to that of WT and phyA-17 line. When the sps-1 line was compared with that of phyA-17 and WT individually by *T*-test (p= 0.001), even fewer genes were displayed. *T*-test analysis displayed 17 and 15 genes as differentially regulated in the sps-1 line compared to that of phyA-17 and WT, respectively (Appendix F). The gene expression data in the sps-1 line suggests the effect of trans-acting loci, and discount the possibility of an EMS effect.

(ii) Activation of loci residing in the heterochromatin area (e.g. 5S ribosomal DNA arrays, transposons or retrotransposons) was not detected in sps-1 line. Consistent with the analysis of the selected loci (using RT-PCR), microarray analysis also suggested that genomic targets of SPS-1 may be the expressed genes and not the heterochromatic loci or pseudogenes.

(iii) Most of the differentially regulated genes (~68) in sps-1 line were suppressed or activated only up to 1.5 to 2-fold. A few genes (10 - 12) were suppressed to 4 to 6-fold.

(iv) Consistent with the northern analysis, *phyA* expression in the sps1 line was found to be restored to ~1.5-fold compared to the WT level.

(v) By comparing expression profile of phyA-17 line with WT, majority of the genes were found to be down-regulated (total of approximately 200). Only ~100 genes were up-regulated in phyA-17 line.

(vi) As expected, PHYA-regulated genes were not found to be differentially expressed in any of the three genotypes (Appendix G). PHYA-regulated genes are up- or down-regulated upon exposure to FR light as shown in the microarray study carried out using *phyA* null-mutant of *Arabidopsis* seedlings by Tepperman et al., 2001 and Wang et al., 2002.

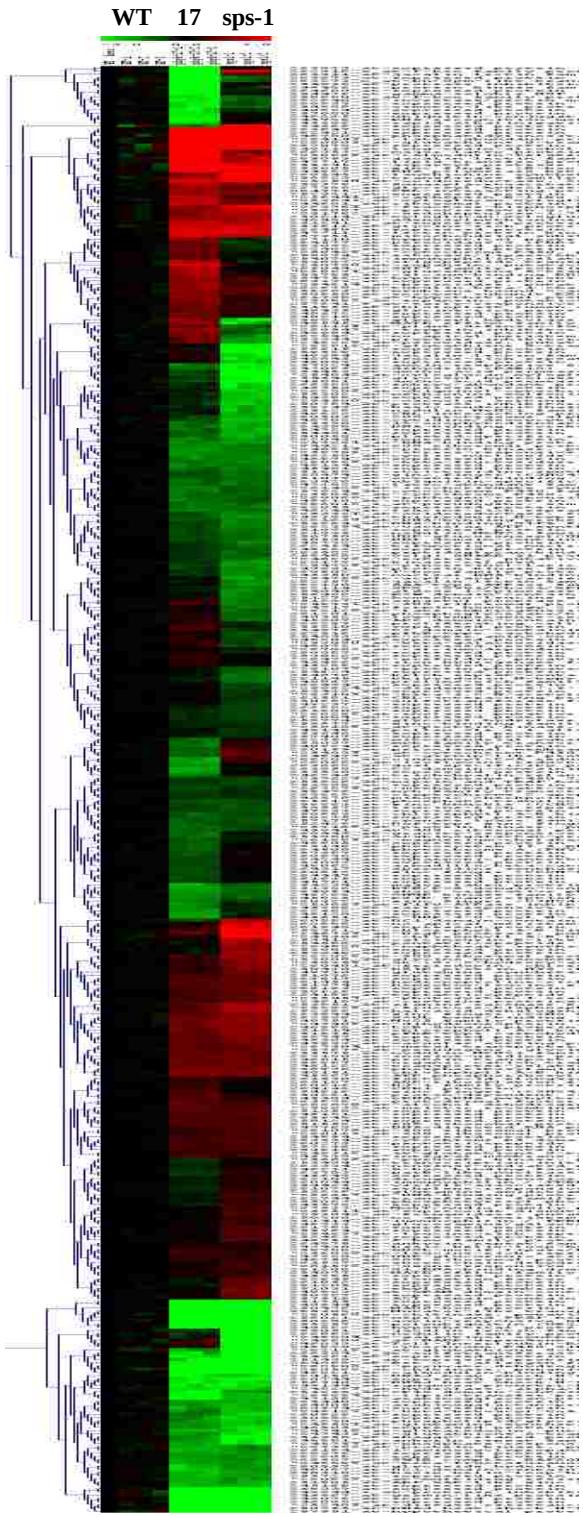


Figure 42: Heat map displaying the expression of 448 transcripts identified by hierarchical clustering . Heat map of three individual hybridizations using total RNA from wild-type, phyA-17 and sps-1

To identify the endogenous targets of SPS-1, we removed those genes from analysis that were expressed at equal levels in phyA-17 and sps-1 lines. The rest were divided into following two groups:

(A) Up-regulated or down-regulated genes in sps-1 line: These genes were expressed at equal levels in WT Col-0 and phyA-17 lines, but at altered level in sps-1 line. A total of 68 genes were found to be uniquely expressed in the sps-1 line (Fig. 43).

(B) The genes that were down or up-regulated in phyA-17 line, but restored to normal or near normal levels in the sps-1 line. A total of 108 genes were found to resume the expression to WT level in the sps-1 line (Fig. 43).

The analysis of expression profiles indicate that a number of genes undergo suppression in *sps-1* background, and the vast majority shows a narrow range (1.5 to 3-fold) of change in expression. Further, a lower range (1.5 to 2-fold) of alteration in gene expression was observed among the activated genes in *sps-1* background. None of the genes showed more than 5-fold change in expression in the *sps-1* background. This moderate change in the expression profile found in sps-1 line suggests that the SPS-1 could be a factor involved in the maintenance of the proper epigenetic state of the expressed genes. The proper epigenetic state imposed by SPS-1 is important for maintaining the proper gene expression in the wild-type background.

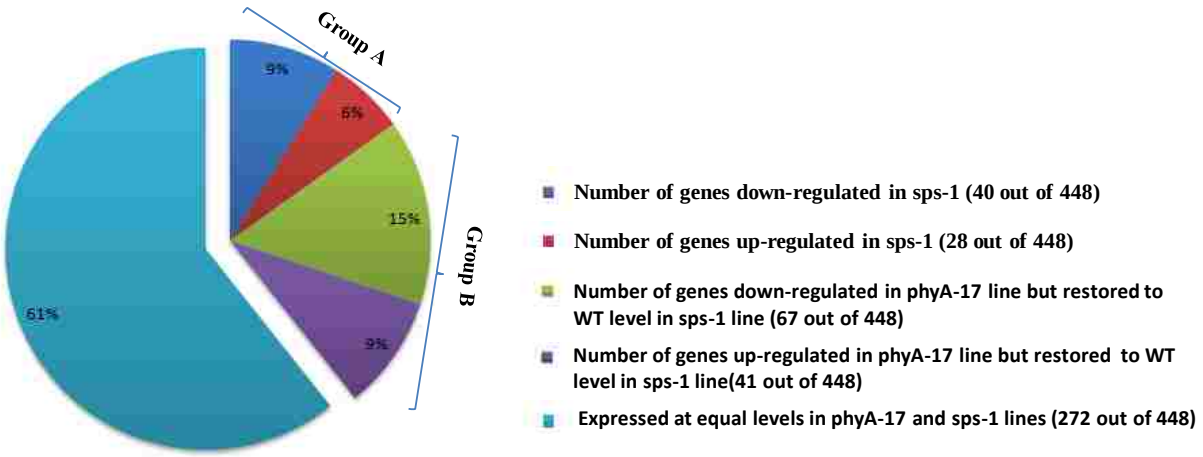


Figure 43: Pie chart representing number of putative targets of SPS-1

As observed in the *phyA'* locus, distribution of methylated sites is altered in the *sps-1* background. Gain of methylation is observed in *phyA'* exon 2 in the *sps1* background; however, it resulted in *phyA* activation. Likewise, other loci in sps1 line could have undergone alteration in the methylation pattern resulting in either hypermethylation or hypomethylation. If hypermethylation occurs in 5' region of the coding region, more likely the gene would be down regulated; however, if hypermethylation occurs in the 3' region, no change would occur (Tran et al., 2006; Zhang et al., 2006; Zilberman et al., 2007). On the other hand, hypomethylation in the 5' region would activate the gene above the WT level (Zhang et al., 2006). Thus, in the absence of SPS-1, genes would either be up or down-regulated depending on the accompanied change in the methylation pattern.

The possibility that SPS-1 could be involved in the maintenance of an epigenetic state is further suggested by an interesting observation that a number of genes suppressed or activated in phyA-17 line are restored to WT expression level in the *sps-1* background. Therefore, we speculate that the function of SPS-1 is to maintain the epigenetic state of these loci defined by its methylation state. For example, hypermethylation of *phyA'* locus results in transcriptional suppression. In the absence of SPS-1, remodeling of the suppressed locus takes place accompanied with the change in the CG methylation pattern. As a result, expression of the *phyA* locus is restored to the WT level. Although, the mechanism of action of SPS-1 is not understood at present, the microarray analysis strongly suggests that the targets of *SPS-1* are transcriptionally active loci.

4.3 Summary

phyA' silencing is maintained by a novel epigenetic pathway, as none of the known epigenetic factors were found to be involved in *phyA'* silencing, except MET1. To identify the factors involved in maintaining the transcriptional silencing of *phyA'*, we performed suppressor screening, and isolated a suppressor line, sps-1, carrying a second site mutation that restored PHYA expression level back to WT. Most of the methylation at *phyA'* locus is maintained in *sps-1* background, however demethylation at one of CG site (H6) and gain of methylation at few CG sites is found to be associated with release of *phyA'* silencing. Although, we cannot confirm the critical importance of H6 methylation in *phyA'* silencing, it is likely that *sps-1*-mediated H6 demethylation is resulting in the release of *phyA'* silencing. SPS-1 is speculated to be a factor that interacts with DNA methylation, which in turn determines the expression state of the gene. To identify the targets of SPS-1, we examined the transcriptome of sps-1 line using the Affymetrix *Arabidopsis* ATH1 genome arrays, and compared to that of WT and phyA-17 line. The resulting gene expression data suggested that the targets of SPS-1 are expressed genes in euchromatin.

CHAPTER 5

References

Aceituno FF, Moseyko N, Rhee SY, Gutiérrez RA. (2008) The rules of gene expression in plants: organ identity and gene body methylation are key factors for regulation of gene expression in Arabidopsis thaliana. *BMC Genomics*; 9: 438

Alvarez-Venegas R, Avramova Z. (2005). Methylation patterns of histone H3 Lys 4, Lys 9 and Lys 27 in transcriptionally active and inactive Arabidopsis genes and in atx1 mutants. *Nucleic Acids Res.* 33: 5199-207

Amedeo, P., Habu, Y., Afsar, K., Mittelsten Scheid, O. *and* Paszkowski, J. (2000) Disruption of the plant gene MOM releases transcriptional silencing of methylated genes. *Nature* 405*:* 203–206.

Aufsatz, W., Mette, M.F., Van Der Winden, J., Matzke, M., and Matzke, A.J. (2002) HDA6, a putative histone deacetylase needed to enhance DNA methylation induced by double-stranded RNA. *EMBO J.* 21: 6832–6841

Aufsatz W, Mette MF, Matzke AJM, Matzke M (2004) The role of MET1 in RNA-directed *de novo* and maintenance methylation of CG dinucleotides. *Plant Mol Biol* 54: 793–804

Barnes, S.A., Nishizawa, N.K., Quaggio, R.B., Whitelam, G.C., and Chua, N.H. (1996). Far-red light blocks greening of Arabidopsis seedlings via a phytochrome A-mediated change in plastid development. *Plant Cell* 8:601–615

Bartee, L., and J. Bender, (2001) Two Arabidopsis methylation-deficiency mutations confer only partial effects on a methylated endogenous gene family. *Nucleic Acids Res*. 29**:** 2127-2134

Bartee L, Malagnac F, Bender J (2001). *Arabidopsis* cmt3 chromomethylase mutations block non-CG methylation and silencing of an endogenous gene. Genes Dev 15: 1753–1758

Bender J, Fink GR. (1995) Epigenetic control of an endogenous gene family is revealed by a novel blue fluorescent mutant of *Arabidopsis. Cell* 83:725–734

Berg, A., Meza, T.J., Mahic, M., Thorstensen, T., Kristiansen, K., and Aalen, R.B. (2003). Ten members of the Arabidopsis gene family encoding methyl-CpG-binding domain proteins are transcriptionally active and at least one, AtMBD11, is crucial for normal development. *Nucleic Acids Res.* **31**: 5291-5304

Bernatavichute YV, Zhang X, Cokus S, Pellegrini M, Jacobsen SE. (2008). Genome-wide association of histone H3 lysine nine methylation with CHG DNA methylation in Arabidopsis thaliana. *PLoS ONE* 3: 3156

Bies-Etheve, N. *et al.* (2009) RNA-directed DNA methylation requires an AGO4-interacting member of the SPT5 elongation factor family. *EMBO Rep 10*: 649-654

Bird A. (2001) Methylation Talk Between Histones and DNA. *Science* 294:2113-2115

Botto J F, Sanchez R A, Whitelam G C, Casal J J. (1996) Phytochrome A mediates the promotion of seed germination by very low fluences of light and canopy shade light in Arabidopsis. *Plant Physiology* 110:439-444

Brzeski, J., and A. Jerzmanowski (2003) Deficient in DNA methylation 1 (DDM1) defines a novel family of chromatin-remodeling factors. *J. Biol. Chem.* 278: 823–828

Cao X, and S.E. Jacobsen (2002a) Role of the *Arabidopsis* DRM methyltransferases in de novo DNA methylation and gene silencing. *Curr. Biol.* **12:** 1138–1144

Cao X, Jacobsen SE. (2002) Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. *Proc Natl Acad Sci U S A*. 99(Suppl 4):16491–16498

Cao X, Aufsatz W, Zilberman D, Mette MF, Huang MS, et al. (2003) Role of the DRM and CMT3 methyltransferases in RNA-directed DNA methylation. *Curr Biol*. 13:2212–2217

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

Chawla R, Nicholson SJ, Folta KM, and Srivastava V (2007) Transgene-induced silencing of Arabidopsis Phytochrome A gene via exonic methylation. *Plant J.* 52: 1105 - 1118

Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, et al. (2008) Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature* 452: 215-219

Earley KW, Pontvianne F, Wierzbicki AT, Blevins T, Tucker S, et al. (2010) Mechanisms of HDA6-mediated rRNA gene silencing: suppression of intergenic Pol II transcription and differential effects on maintenance versus siRNA-directed cytosine methylation. Genes Dev 24: 1119–1132

Ehrlich, M. et al. (1982) Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic Acids Res.* 10: 2709-2721

Elmayan T., Balzergue, S., Beon, F., Bourdon, V., Daubremet, J., Guenet, Y., Mourrain, P., Palauqui, J.C., Vernhettes, S., Vialle, T., et al. (1998) *Arabidopsis* mutants impaired in cosuppression. *Plant Cell* 10**:** 1747-1758

Fagard M, Vaucheret H. (2000) (Trans) gene silencing in plants: how many mechanisms Annual Review of Plant Physiology and Plant Molecular Biology; 51:167-194

Farrance, 1. K., and R. Ivarie. (1985) Ethylation of poly (dC-dG).poly (dC-dG) by ethyl methanesulfonate stimulates the activity of mammalian DNA methyltransferase fn vftro. Proc. Natl. Acad. S c f . USA 82: 1045.

Fu, X., Kohli, A., Twyman, R.M. *and* Christou, P. *(2000)* Alternative silencing effects involve distinct types of non-spreading cytosine methylation at a three-gene single-copy transgenic locus in rice. *Mol. Gen. Genet.* 263: 106-118

Gendrel, A.V., Lippman, Z., Yordan, C., Colot, V. *and* Martienssen, R.A. (2002) Dependence of heterochromatic histone H3 methylation patterns on the *Arabidopsis* gene DDM1. *Science* 297: 1871–1873

Grafi, G., Zemach, A., and Pitto, L. (2007). Methyl-CpG-binding domain (MBD) proteins in plants. Biochim. Biophys. Acta 1769: 287-294

Greene EA, Codomo CA, Taylor NE, Henikoff JG, Till BJ, et al. (2003) Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in Arabidopsis. *Genetics* 164:731-740

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

He, X. J. *et al.* (2009) An effector of RNA-directed DNA methylation in *Arabidopsis* is an ARGONAUTE 4- and RNA-binding protein. *Cell* 137: 498-508

Henderson, I. R. & Jacobsen, S. E. (2007) Epigenetic inheritance in plants. Nature 447, 418–424

Herr, A. J., Jensen, M. B., Dalmay, T. & Baulcombe, D. C. (2005) RNA polymerase IV directs silencing of endogenous DNA. *Science* 308: 118-120

Hohn, T., Corsten, S., Rieke, S., Muller, M. *and* Rothney, H. (1996) Methylation of coding region alone inhibits gene expression in plant protoplasts. *Proc. Natl Acad. Sci. USA*, 93*:* 8334-8339

Hollister JD, Smith LM, Guo YL, Ott F, Weigel D, Gaut BS (2011) Transposable elements and small RNAs contribute to gene expression divergence between Arabidopsis thaliana and Arabidopsis lyrata. *Proc Natl Acad Sci* U S A. Jan 20.

Hsieh, C. L. (1997) Stability of patch methylation and its impact in regions of transcriptional initiation and elongation. *Mol. Cell. Biol.* 17: 5897-5904.

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

Huettel, B. *et al.* (2006) Endogenous targets of RNA-directed DNA methylation and Pol IV in *Arabidopsis*. *EMBO J.* 25: 2828-2836

Ivarie R, and Morris JA. (1986) Activation of a nonexpressed hypoxanthine phosphoribosyltransferase allele in mutant H23 HeLa cells by agents that inhibit DNA methylation. *Mol Cell Biol.* 6: 97-104

Jackson JP, Lindroth AM, Cao X, Jacobsen SE (2002) Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* 416: 556–560

Jackson, J.P., Johnson, L., Jasencakova, Z., Zhang, X., PerezBurgos, L., Singh, P.B., Cheng, X., Schubert, I., Jenuwein, T. *and* Jacobsen, S.E (2004) Dimethylation of histone H3 lysine 9 is a critical mark for DNA methylation and gene silencing in *Arabidopsis thaliana. Chromosoma* 112: 308-315

Jacobsen SE, Meyerowitz EM (1997) Hypermethylated SUPERMAN epigenetic alleles in arabidopsis. *Science.* 277(5329): 1100-1103

Jacobsen SE, Sakai H, Finnegan EJ, Cao X, Meyerowitz EM (2000) Ectopic hypermethylation of flower-specific genes in Arabidopsis. *Current biology* 10(4): 179-186

Jang IC, Chung PJ, Hemmes H, Jung C, Chua NH (2011) Rapid and reversible light-mediated chromatin modifications of Arabidopsis phytochrome a locus *Plant Cell*. 23(2): 459-70

Jeddeloh JA, Stokes TL, Richards EJ. (1999) Maintenance of genomic methylation requires a SWI2/SNF2-like protein. Nat Genet. 22: 94-97

Johnson LM, Bostick M, Zhang X, Kraft E, Henderson I, Callis J, et al. (2007) The SRA methyl-cytosine-binding domain links DNA and histone methylation. *Curr. Biol*. 17(4): 379–384

Kakutani T, Jeddeloh JA, Flowers SK, Munakata K, Richards EJ (1996) Developmental abnormalities and epimutations associated with DNA hypomethylation mutations. *Proc Natl Acad Sci* USA, 93: 12406-12411

Kakutani, T., Munakata, K., Richards, E.J., and Hirochika, H. (1999) Meiotically and mitotically stable inheritance of DNA hypomethylation induced by ddm1 mutation of Arabidopsis thaliana. *Genetics* 151: 831-838

Kankel MW, Ramsey DE, Stokes TL, Flowers SK, Haag JR, Jeddeloh JA, Riddle NC, Verbsky ML, Richards EJ (2003) Arabidopsis MET1 cytosine methyltransferase mutants. *Genetics* 163:1109-1122

Kanno T, Mette MF, Kreil DP, Aufsatz W, Matzke M, Matzke AJM (2004) Involvement of putative SNF2 chromatin remodeling protein DRD1 in RNA-directed DNA methylation. *Curr Biol* 14: 801–805

Lawrence RJ, Earley K, Pontes O, Silva M, Chen ZJ, et al. (2004) A concerted DNA methylation/histone methylation switch regulates rRNA gene dosage control and nucleolar dominance. *Mol Cell.* 13: 599-609

Lindroth AM, Cao X, Jackson JP, Zilberman D, McCallum CM, Henikoff S, Jacobsen SE. (2001) Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation. *Science* 292:2077–2080

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

Lippman Z, Gendrel AV, Black M, Vaughn MW, Dedhia N, et al. (2004). Role of transposable elements in heterochromatin and epigenetic control. *Nature* 430: 471-76

Lorincz, M.C., Dickerson, D.R., Schmitt, M. *and* Groudine, M. (2004) Intragenic DNA methylation alters chromatin structure and elongation efficiency in mammalian cells. *Nat. Struct. Mol. Biol.* 11: 1068-1075

Luff B, Pawlowski L, Bender J. (1999) An inverted repeat triggers cytosine methylation of identical sequences in *Arabidopsis. Mol Cell*. 3: 505-511

Malagnac F, Bartee L, Bender J. (2002) An Arabidopsis SET domain protein required for maintenance but not establishment of DNA methylation. *EMBO J* 21: 6842-6852

Martiensen RA, Colot V (2001) DNA methylation and epigenetic inheritance in plants and filamentous fungi. *Science* 293: 1070–1074

Matzke M., Matzke,A.J.M. and Kooter,J.M. (2001) RNA: guiding gene silencing. Science 293: 1080–1083

Matzke MA, Birchler JA. (2005) RNAi-mediated pathways in the nucleus. *Nat Rev Genet* 6: 24-35

Matzke MA, Mette MF, Matzke AJ. (2000) Transgene silencing by the host genome defense: implications for the evolution of epigenetic control mechanisms in plants and vertebrates *Plant Molecular Biology* 43: 401-415

Matzke, M., Kanno, T., Daxinger, L., Huettel, B., and Matzke, A.J. (2009) RNA-mediated chromatin-based silencing in plants. *Curr. Opin. Cell Biol*. 21: 367-376

Melquist S, Luff B, Bender J. (1999) *Arabidopsis PAI* gene arrangements, cytosine methylation and expression. *Genetics.* 153: 401-413

Meyer, P. (2000) Transcriptional transgene silencing and chromatin components. *Plant Mol. Biol.* **43**:221-234

Mittelsten Scheid, O., Jakovleva, L., Afsar, K., Maluszynska, J. & Paszkowski, J. (1996) A change of ploidy can modify epigenetic silencing. *Proc. Natl. Acad. Sci. USA* 93: 7114-7119

Mittelsten Scheid O., Afsar K., Paszkowski J. (1998). Release of epigenetic gene silencing by trans-acting mutations in Arabidopsis. *Proc. Natl. Acad. Sci. USA* 95: 632–637

Mittelsten Scheid O, Afsar K, Paszkowski J. (2003) Formation of stable epialleles and their paramutation-like interaction in tetraploid Arabidopsis thaliana. *Nat Genet*. 34: 450-454

Nagatani, A., Reed, J. W. *and* Chory, J. (1993) Isolation and initial characterisation of *Arabidopsis* mutants that are deficient in functional phytochrome A. *Plant Physiol.* 102, 269–277.

Nicholson S.J, Srivastava V (2009) Transgene constructs lacking transcription termination signal induce efficient silencing of endogenous targets in Arabidopsis. Mol Genet Genom. 282(3): 319-28

Onodera Y, Haag JR, Ream T, Nunes PC, Pontes O, and Pikaard C (2005) Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. Cell 120: 613-622

Page DR, Grossniklaus U (2002) The art and design of genetic screens: *Arabidopsis thaliana*. *Nat Rev Genet* 3: 124-136

Parks, B.M. *and* Quail, P.H. (1993) *hy8*, a new class of Arabidopsis long hypocotyls mutants deficient in functional phytochrome A. *Plant Cell* **5:** 39–48.

Peragine A, Yoshikawa M, Wu G, Albrecht HL, Poethig RS (2004) SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of trans-acting siRNAs in Arabidopsis. Genes Dev 18: 2368–2379

Probst AV, Fransz PF, Paszkowski J, Mittelstedn Scheid O (2003) Two means of transcriptional reactivation within heterochromatin. Plant J 33: 743–749

Probst AV, Fagard M, Proux F, Mourrain P, Boutet S, et al. (2004) Arabidopsis histone deacetylase HDA6 is required for maintenance of transcriptional gene silencing and determines nuclear organization of rDNA repeats. Plant Cell 16: 1021–1034

Richards E.J and S.C Elgin (2002) Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects *Cell* **108**: 489-500

Saze H, Scheid OM, Paszkowski J (2003) Maintenance of CpG methylation is essential for epigenetic inheritance during plant gametogenesis. *Nat Genet* 34:65–69

Shinomura, T., Uchida, K., and Furuya, M. (2000). Elementary processes of photoperception by phytochrome A for high-irradiance response of hypocotyl elongation in Arabidopsis. *Plant Physiol*. 122: 147–156

Sidorenko, L.V. and Peterson, T. (2001) Transgene-induced silencing identifies sequences involved in the establishment of paramutation of the maize p1 gene. *Plant Cell* 13: 319-335

Soppe WJJ, Jacobsen SE, Alonso-Blanco C, Jackson J, Kakutani T,Koornneef M and Peeters AJM (2000) The late flowering phenotype of fwa mutants is caused by gain of function epi-alleles at a homeodomain gene *Molecular Cell* 6(4): 791-802

Stam, M., Belele, C., Dorweiler, J.E. and Chandler, V.L. (2002) Differential chromatin structure within a tandem array 100 kb upstream of the maize b1 locus is associated with paramutation. *Genes Dev.* 16: 1906–1918

Straub, T., and Becker, P.B. (2007) Dosage compensation: the beginning and end of generalization. *Nat Rev Genet* 8**:** 47-57

Tariq, M. ***and*** **Paszkowski, J**. *(2*004*)* DNA and histone methylation in plants. *Trends Genet*. 20: 244-251

Teixeira FK, Heredia F, Sarazin A, Roudier F, Boccara M, et al. (2009) A Role for RNAi in the Selective Correction of DNA Methylation Defects. *Science* 323: 1600-1604

Tran R, Henikoff J, Zilberman D, Ditt R, Jacobsen S, Henikoff S. (2005) DNA methylation profiling identifies CG methylation clusters in Arabidopsis genes. *Curr Biol.* 15(2):154-159

Vaillant I, Schubert I, Tourmente S, Mathieu O. (2006) MOM1 mediates DNA-methylation-independent silencing of repetitive sequences in Arabidopsis. *EMBO Rep* 7:1273-1278

Vongs A, Kakutani T, Martienssen RA, Richards EJ. (1993) Arabidopsis thaliana DNA methylation mutants. Science 260:1926–1928

Wassenegger M, Heimes S, Riedel L and Sänger HL (1994) RNA-directed de novo methylation of genomic sequences in plants. *Cell* 76:567-76

Wierzbicki, A.T., Haag, J.R. & Pikaard, C.S. (2008) Noncoding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. *Cell* 135: 635-648

Wierzbicki, A. T., Ream, T. S., Haag, J. R. & Pikaard, C. S. (2009) RNA polymerase V transcription guides ARGONAUTE4 to chromatin. *Nature Genet.* 41: 630-634

Whitelam, G.C. and Devlin, P.F. (1997) Roles of different phytochromes in Arabidopsis photomorphogenesis. *Plant Cell Environ*. 20: 752-758

Wang, H.Y., Ma, L.G., Habashi, J., Li, J.M., Zhao, H.Y., and Deng, X.W. (2002). Analysis of far-red light-regulated genome expression profiles of phytochrome A pathway mutants in Arabidopsis. *Plant J.* 32:723-733.

Woo, H. R., Pontes, O., Pikaard, C. S. & Richards, E. J. (2007) VIM1, a methylcytosine-binding protein required for centromeric heterochromatinization. *Genes Dev.* 21: 267-277

Woo, H. R., Dittmer, T. A. & Richards, E. J. (2008) Three SRA-domain methylcytosine-binding proteins cooperate to maintain global CpG methylation and epigenetic silencing in Arabidopsis *PLoS Genet* 15: e1000156

Zemach, A., and Grafi, G. (2003). Characterization of *Arabidopsis thaliana* methyl-CpG-binding domain (MBD) proteins. Plant J. 34**:** 565−572.

Zhai, J., Liu, J., Liu, B., Li, P., Meyers, B.C., Chen, X., and Cao, X. (2008). Small RNA-directed epigenetic natural variation in Arabidopsis thaliana. *PLoS Genet*. 4 e1000056

Zhang, X., Yazaki, J., Sundaresan, A., Cokus, S., Chan, S.W.L., Chen, H., Henderson, I.R., Shinn, P., Pellegrini, M., Jacobsen, S.E., and Ecker, J. (2006). Genome-wide high-resolution mapping and functional analysis of DNA methylation in Arabidopsis. *Cell* 126: 1189-1201

Zhang X, Clarenz O, Cokus S, Bernatavichute YV, Pellegrini M, et al. (2007) Whole-genome analysis of histone H3 lysine 27 trimethylation in *Arabidopsis*. *PLoS Biol* 5:e129

Zhang X, Bernatavichute YV, Cokus S, Pellegrini M, Jacobsen SE. (2009). Genome-wide analysis of mono-, di- and trimethylation of histone H3 lysine 4 in *Arabidopsis thaliana*. *Genome Biol.* 10:R62

Zilberman D, Cao X, Jacobsen SE (2003) ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. *Science* 299:716–719

Zilberman, D. et al. (2004) Role of Arabidopsis ARGONAUTE4 in RNA-directed DNA methylation triggered by inverted repeats. *Curr Biol* 14: 1214-1220

Zilberman, D., Gehring, M., Tran, R.K., Ballinger, T., and Henikoff, S. (2007). Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers interdependence between methylation and transcription. *Nat. Genet*. 39: 61-69

Appendix A

A-1 *Plant growth and condition*

Plants were grown in a mixture of Redi-Earth (Scotts)/vermiculite (50%:50%) in growth chambers [16 hr illumination (fluorescent + incandescent)/day, 85% relative humidity, 22°C].

A-2 *DNA extraction*

CTAB (Cetyl Triethyl Ammonium Bromide) Extraction: DNA was extracted from young flower buds and seedlings using Flower buds were frozen in liquid nitrogen and crushed to powder. The powdered tissues were suspended in CTAB extraction buffer and placed in a water bath maintained at 55˚C for 30-60 minutes. The proteins and lipids present in the samples were removed by adding phenol/chloroform mix (proteins are dissolved in phenol, the lipids are dissolved in chloroform, and DNA is located in the aqueous phase). Supernatant was further mixed with chloroform:Isoamyl alcohol (24:1) to remove any protein contamination. Later the supernatent was precipitated using double the volume of 95% ethanol. The precipitated DNA was washed with 70% ethanol, DNA precipitates were dried properly, and it was dissolved in autoclaved water.

A-3 *PCR clean up kit*

PCR fragments were purified by eluting it from agarose gel using Gene-Clean kit, Qiagene following manufacturer's instructions.

A-4 *Southern blot analysis*

DNA was extracted by the CTAB method from pooled seedlings or flower buds of individual plants. DNA was digested with appropriate restriction endonuclease in the total volume reaction of 200µL. Digested DNA was electrophoresed on 0.8% agarose gel to separate by size, and then transferred to a nylon membrane (Hybond N+®, Amersham Biosciences, UK) by placing it on the top of the gel. Gentle pressure was applied to the gel by placing a stack of paper towels and a weight on top of the membrane and gel for overnight, to ensure good and even contact between gel and membrane. Next day the membrane is dried and exposed to UV to fix DNA on the nylon membrane using the UV Stratalinker® 2400 from Stratagene. The membrane was pre-hybridized at 65◦C using 10 ml of pre-hybridization solution (autoclaved water, 5X SSC, 5X Denhardt's solution and 0.5X SDS solution) in a hybridization bottle (hybond) for at least 2 hours. 100 µl of denatured salmon sperm DNA was added to block non-specific binding during prehybridization. The membrane was then exposed to a P32 labeled probe (a single DNA fragment with a specific sequence whose presence in the target DNA is to be determined)

for overnight at 65◦C in the Hybaid hybridization oven. The radio-labeled probes were synthesized using Random Primed DNA labeling Kit (Roche Diagnostics Corporation, IN, USA) by following the manufacturer's instructions. After hybridization, excess probe is washed from the membrane using wash solutions I (2 X SSC and 0.5 X SDS solutions), II (1 X SSC and 0.25 X SDS solutions) and III (0.5 X SSC and 0.125 X SDS solutions) for 15 minutes each at 65◦C. After washes, the pattern of hybridization is visualized by phosphorimager.

A-5 *Northern blot analysis*

Approximately 5 ug of RNA was loaded onto 1.2% agarose gels. Agarose gels were run in 1x MOPS buffer. Blotting was performed using Amersham Hybond-N+ membranes. Northern prehybridization buffer utilizes Church and Gilbert solution (7% SDS, 250 mM Sodium Phosphate pH 7.2, 1 mM EDTA). Probes were labeled using a-P32 CTP and Roche's Random-Primed DNA Labeling Kit. Hybridization was performed at 42˚C to 65˚C. Radioactive materials were monitored and disposed of in accordance with University protocols.

Appendix B

ChIP Protocol

DAY- 1

On first day seedlings were sliced from bottom (when they were approximately one inch in height), and washed with D/W three times. After thorough washing, excess water was removed with the kitchen paper.

Chromatin Cross-linking: 1.5g of seedlings from both the samples (phyA-17 and Col-0) was placed in 50ml Falcon tube. 37ml of 1% formaldehyde solution was added to crosslink the samples and seedlings were gently submerged in the solution at the bottom of the tube by stuffing it with nylon mesh. Falcon tubes were placed in exsiccator and vacuum was drawn for 10 minutes, 15hg pressure was maintained by opening gauge. After 10 mins vacuum was released slowly, and tubes were shaken to remove air bubbles. Seedlings were appeared translucent at this stage. 2.5ml of 2M glycine was added to quench cross-linking, and vacuum was drawn for 5 mins. Again, vacuum was released slowly and tubes were shaken. Nylon mesh was removed, supernatant was decanted and seedlings were washed with D/W to remove formaldehyde. Water was removed by pressing seedlings between nylon meshes and two kitchen papers over it.

Chromatin Preparation: Seedlings were crushed to powder using precooled mortar and liquid nitrogen. Using a cooled spoon, crushed powder was added to 30 ml of Extraction Buffer 1 stored on ice. Mixture was vortexed (on a VWR Vortex Mixer) at a maximum setting three times for approximately thirty seconds each, and kept at 4°C until solution was homogenous. Homogenized solution was filter through Miracloth into a new, ice-cold 50ml Falcon tube by rigidly pressing the solid material. The extract was again passed through Miracloth, and it was centrifuged in a precooled Beckman rotor at 4000 rpm for 20 minutes at 4°C. Supernatant was discarded and pellet was resuspened in 1ml of Extraction Buffer 2 by pipetting up and down. Solution was transferred to eppendorf tube, and centrifuged at 13000 rpm for 10 minutes. Supernatant was discarded and pellet was resuspended in 300µl of Extraction Buffer 2 by pipetting up and down, care was taken so that bubble should not form. 300µl of Extraction Buffer 3 was taken to fresh Eppendorf tube, and using pipette mix, Extraction Buffer 2 was layered over the Extraction Buffer 3. Tube containing both Extraction Buffer 2 and 3 was centrifuged at 13000 rpm for 1 hr. Pellet was resuspended in 500µl of cold Nuclei Lysis Buffer by pipetting up and down and by vortexing. 10µl was aliquoted to run on an agarose. Each extract was sonicated 6 times for 10 seconds, 40% duty cycle and 20% power. Samples were stored on ice for 1 minute between sonication steps. The cell debris was pelleted by centrifuging at 13000 rpm for 10 minutes twice. Supernatant was added to new eppendorf tube. 10µl was aliquoted from the sonicated extract to run on an agarose gel. Aliquots were run on 1.5% agarose gel. In the sonicated samples, DNA was fragmented between 200-2000bp.

Pre-clearing and immunoprecipitation: 200µl Sonicated chromatin solutions of each sample (Col and phyA-17) was transferred to a new eppendorf tube in three replicates for three antibodies; H3K9me2, K3K4me3, and H3K27me3. 1300µl of CHIP Dilution Buffer

was added per eppendorf tube. Meanwhile, Protein A agarose beads were pre-absorbed with sheared salmon sperm DNA (Upstate, Cat. 16-157) by rinsing 40µl beads with 1ml of CHIP Dilution Buffer for three times. Mixture was centrifuged for 30 seconds at 13000 rpm between the washes to pellet the beads. 1.5ml diluted chromatin solution was pre-cleared by adding to a 40µl of equilibrated beads, rotated at 5000 rpm for 1 hour at 4°C. Beads were separated from the chromatin extracts by centrifuging it for 30 seconds at 13000rpm. Supernatant of the identical genotypes were combined and transferred to a fresh 15ml Falcon tube (4.5ml of Col, and 4.5ml of phyA-17). The beads were subsequently discarded. 60µl of aliquot of pooled chromatin from both the samples was stored at -20°C, which served as an input control. 600µl of pooled chromatin solution from each sample was added to an eppendorf tube with ~10µg of antibody, and mixed up very well. 600µl of chromatin solution from each sample was added to another eppendorf tube without antibody, which serves as mock IP. All samples were immunoprecipitated overnight at 4°C on a rotor at 500rpm speed, they termed as IPs.

DAY-2

Collection, washes and elution of immune complexes: On the second day, fresh CHIP Dilution Buffer was prepared and stored at 4°C. Protein A agarose beads were pre-absorbed with sheared salmon sperm DNA in the same way as previously, by rinsing 40µl of bead three times with 1ml CHIP Dilution Buffer in an eppendorf tube. One aliquot of beads was prepared per IP, and spinned 30 seconds at 13000rpm between the washes to pellet the beads. IPs were added to beads and rotated for 1 hour at 4°C. In the meantime, Elution buffer was prepared, and placed at 65°C. After 1 hour, IPs were removed from rotor, centrifuged at 5000 rpm for 30 seconds to collect beads and discard supernatant. 1ml of Low Salt Wash Buffer was added to beads, rotate for 5 minutes at 4°C, and centrifuge at 5000 rpm for 30 seconds to collect beads and discard supernatant. Performed this step two times. 1ml of High Salt Wash Buffer was added to beads, rotated for 5 minutes at 4°C, and centrifuged at 5000 rpm for 30 seconds to collect beads and discard supernatant. Performed this step two times. 1ml of LiCl Wash Buffer was added to beads, rotated for 5 minutes at 4°C, and centrifuged at 5000 rpm for 30 seconds to collect beads and discard supernatant. Repeated this step two times. 1ml of TE Buffer was added to beads, rotated for 5 minutes at 4°C, and centrifuged at 5000 rpm for 30 seconds to collect beads and discard supernatant. Repeated TE wash and beads were collected. Immune complexes were eluted from the beads by adding 250µl of Elution Buffer. Vortexed briefly to mix and incubated at 65°C for 15 minutes. Samples were centrifuged at 13000 rpm for 30 seconds and transferred supernatant to a fresh Eppendorf tube. Repeated elution and finally combined the two elutes. 460µl of Elution Buffer was added to the 60µl input control aliquoted on Day 1.

Reverse crosslinking: Reverse cross-linking was accomplished by adding 20µl of 5M NaCl to samples and the samples were incubated in a waterbath at 65°C overnight

DAY-3

DNA cleanup: After reverse cross-linking, 10µl of 0.5M EDTA, 20µl 1M Tris-HCl pH6.5 and 1µl of 20mg/ml proteinase K were added to each sample, which was incubated for 1 hour at 45°C. Samples were extracted with 1 volume of phenol-chloroform. Spun in

cooled centrifuge at 13000 rpm for 15 minutes and supernatant was transferred to 2ml reaction tube. The DNA was precipitated with 1/10 volume 3M NaOAc (pH5.2) and 3 volumes absolute ethanol. 4µl of glycogen (Roche, Cat.901393) was added per precipitation, and samples were incubated -20°C for at least 1 hour. Then, the samples were centrifuged at 13000 rpm for 15 minutes and finally washed with 70% ethanol. The pellet was allowed to dry and then resuspended in 50µl 10mM Tris-HCl (pH7.5).

Appendix C

C-1 List of Primers used in ChIP assays, making probes in Southern hybridization and semi quantitative RTPCR

Target	Primers	Sequence (5' to 3')	Methods
180 bp single repeat	180F[1] 180R[1]	GTTATCTGTTCCTAAAAGATAATAGTGTTC ACTCCCACTCATGTATTTCCTATCATAGCG	ChIP assay, Southern analysis and RTPCR
5S rDNA	5S1[1] 5S2[1]	GGATGCGATCATACCAG CGAAAAGGTATCACATGCC	ChIP assay
ATSN1	ATS15[1] ATSN1-F4[1]	ACCAACGTGCTGTTGGCCCAGTGGTAAATC AAAATAAGTGGTGGTTGTACAAGC	ChIP assay
Ta2	Ta2-F[1] Ta2-R[1]	AAACGATGCGTTGGGATAGGTC ATACTCTCCACTTCCCGTTTTTCTTTTTA	ChIP assay, Southern analysis
TUB8	TUB8-F[1] TUB8-R[1]	ATAACCGTTTCAAATTCTCTCTCTC TGCAAATCGTTCTCTCCTTG	ChIP assay
PFK	PFK-F[2] PFK-R[2]	GCCACGAAAACCAAACAGAC CCGGAATTTCGATCAATCCT	ChIP assay
TSI	TSP1F[3] TSP1R[3]	GAACTCATGGATACCCTAAAATAC CTCTACCCTTTGCATTCATGAATC	RTPCR
Ta3	Ta3-F[3] Ta3-R[3]	GATTCTTACTGTAAAGAACATGGCATTGAGAGA TCCAAATTTCCTGAGGTGCTTGTAACC	RTPCR
106B	106B-F[3] 106B-R[3]	TTGATTGATAGATCCCTTCTGGA CGAGGATGGGGTAATTGAGT	RTPCR
Actin	ACT-F[4] ACT-R[4]	- -	Northern analysis and RTPCR

[1] Mathieu, O., Probst, A.V. *and* Paszkowski, J. (2005) Distinct regulation of histone H3 methylation at lysines 27 and 9 by CpG methylation in Arabidopsis. *EMBO J.* 24*:* 2783-2791

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

[3]Vaillant I, Schubert I, Tourmente S, Mathieu O. (2006) MOM1 mediates DNA-methylation-independent silencing of repetitive sequences in Arabidopsis. *EMBO Rep* 7:1273-1278

[5]Sigma (Cat no: C3615)

C-2 List of *phyA* specific primers used in ChIP assays and DNA sequencing of *phyA* Locus in sps-1 line

Primers[1]	**Sequence (5' to 3')**
1F	GCTGTGTTTGAATCTTGAGAGGCTGA
1R	TGATGATGCTCCCTGTAACCCTGTG
2F	CTTCGGTTGTTTACTTGCCTTGGAT
2R	GCAAAATGTTCTCCACCATCTG
3F	GGTTACAGGGAGCATCATCATCGAC
3R	CCACATACCGAATCCCCGAGAGATA
4F	TTGAGCACTGATAGTTTGCATGACG
4R	TGTAATCACCTTCAATCCGCGTAAA
5F	CACAGTCTTAGTCACATCGGTGAAAA
5R	CCTCTCAAGATTCAAACACAGCTACA
6F	AGCTGCGTCTATAAGGAATCGTGTTGG
6R	CGTTTTCTCGGCTTTTTGCGTTTTG
7F	CAAAACGCAAAAAGCCGAGAAAACG
7R	TTGTGCTCCCTAACACAGACGACGA
9F	CGATTGCTTTCCCTGTGGGGACTTT
9R	CAATGCAGAGGCACTAGGAGCAGTGA
11F	TGGTGAAGCTGATGAATGGAGATGTTC
11R	TCAAATCCAAGTTCCAACCCAAGAAATG
12F	GTGCTGAACAATGCTGTGAC
12R	CAAGCCAGACCCCTTTTCT
s	AGTTTGAGATCAAGACATCTG
w	ATACAAAGTGTACAGACATTAC

[1]Chawla R, Nicholson SJ, Folta KM, and Srivastava V (2007) Transgene-induced silencing of Arabidopsis Phytochrome A gene via exonic methylation. *Plant J.* 52: 1105 -1118

C-3 List of primers used to amplify *phyA* gene sequence in sps-1 line after bisulfite conversion

Primers	**Sequence (5' to 3')**
B1F	CAAAaCCTTTCCATACTCCC
B1R	AGTttTtTGAGGGtTtAAGG
B2F	CAGTACATGGCCAACATGG
B2R	CCATGACAGCTGCTGGTGCC
B3F	C TCCATCACATTTCACAA
B3R	ACCTTACCTTGTGTGGCTCCA
B4F	TCCCCATCTCACTTCACCA
B5F	CATGATCCAG
B5R	AAGTGAGTTCCACCTGCAGG
B6F	CATCCTTTTAACaATTTCCAC
B7F	TCCCAAATACTTCTCCTAA
B7R	GATGCAATGTGTTGTTTTTGGT
B8R	TGGTTGGGGTGTGTTTTGTA
B9F	ACTTTCATTCTAaCATACCCTTC
B9R	GAGGAAGTGATTGAtAAAATGtT
B10F	CCATTAAaCTCAATCCTTCTTC
B10R	ATTGAAGGGTATGtTAGAATGAAAGT
B11F	CTCATTTCTCCTATAATTACTTA
B11R	ATTGAGGATATTGGGATTATGA
Lowercase 'a' and 't' represent adenine (A) thymidine (T) corresponding to guanine (G) cytosine (C) in the original (unconverted) sequence	

Appendix D

Bisulfite sequences data of *phyA'*/sps-1 using below primer pairs

B1F-B1R:

NNNNNNNNNNGNNNNNNG**C**GNAGATTATTGNAGATG**C**GAAATTTTATGTTG
ATTTTGAAGAGTTAGGTNNNTTTTTTGATTATTTAATTTTAGTG**C**GTGTTATTG
GTT**C**GGTTGTGGAGAATTAGTTATTAAGGTTTGATAAAGTTATTA**C**GATTTAT
TTTTATTATATATAGAAGGGAAAGTTGATTTAGTTTTTT**C**GGTTGTTTATTTGTT
TTGGATGAGAAGATTTTTAAAGTTATTGTATATAG**T**GAGAATGTATTTGAGTT
GTTGATAATGGTTAGTTATGTAGTTTTTAGTGTTGG**C**GAATATTTTGTTTTAGG
TATTGGGATAGATATAAGGAGTTTTTTTATTGTTTTTAGTGTTTTTGTATTGTA
GAAAGTTTTTGGATTTGGAGATGTTTTTTTTTTGAATTTTATTTTTGTGTATTG
NAGGATTTTTGTAAAGTTTTTTTTATG**C**NATTATTTATAGGGTTATAGGGAGTA
TTATTAT**C**GATTTTGAATTT**T**GNGAAGTTTTATGAAGTTTTTATGANANTTGTT
GGNGTTTTATAATTATATAAGTTNGTNGTNAAAGNAATTATTAGGTNGNAAT
TTTTATTTAN**C**GGGAGTATGGAANGGNNTNGN

B2F-B2R:

TTTATGATNNTTGTTGGTGTTTTATAATTATATAAGTTTGTTGTTAAAGTAATT
ATTAGGTTGTAATTTTTATTTAGCGGGAGTATGGAAAGGTTTTGTGATATAAT
GGTTTAAGAGGTTTTTGAATTTACGGGGTATGATAGGGTGATGGTTTATAAGT
TTTATGAAGATGATTACGGTGAGGTTGTTTTCGAGGTTATAAAATTTGGGTTG
GAGTTTTATTTTGGGTTGTATTATTTTGTTATCGATATTTTTTAAGTAGTTCGT
TTTTTGTTTATGAAGAATAAGGTTTGGATGATAGTTGATTGTAATGTAAAANN
NNTTAGGGTGTTTTAAGATGAAAAGTTTTTTTTTGATTTTATTTTGTGTGGTTT
NNN

B3F-B3R:

NNNNNNNNNTTNNAGTATANGNNTAATATGGATTTAATTGTATTTTTGGTTAT
GG**C**GGTTGNNNNTNA**C**GAGGAAGATGGAGAAGGGGATGTTTTTGATGTTATT

NTNTNNNTTTAAAAGAGAAAGAGATTATGGGGTTTAGTGGTTTGTTATAATA
CGATTTCGAGGTTTGTTTTATTTTTTTTTAGGTATGTTTGTGAGTTTTTAGTTTA
AGTGTTTGTTATATATGTTAATAAGGAGGTGGAATTCGATAATTAGATGGTGG
AGAAGAATATTTTGCGTACGTAGATATTTTTATGCNATATGTTGATGCNTGAT
GTTTTATTGGGTATTGTGTCGTAAAGTTTTAATATAATGGATTTTGTGAAATG
TGATGGAGA

B4F-B4R:
NTATGGGGNNTTAGTGGTTTGTTATAATACGATTTCGAGGTTTGTTTTATTTTT
TTTTAGGTATGTTTGTGAGTTTTTAGTTTAAGTGTTTGTTATATATGTTAATAA
GGAGGTGGAATTCGATAATTAGATGGTGGAGAAGAATATTTTGCGTACGTAG
ATATTTTTATGCGATATGTTGATGCGTGATGTTTTATTGGGTATTGTGTCGTAA
AGTTTTAATATAATGGATTTTGTGAAATGTGATGGAGTAGTTTTTTTGTATAA
AGATAAGATATGGAAATTGGGAATAATTTTAAGTGAGTTTTATTTGTAGGAG
ATAGTTTTATGGTTGTGTGAATATTATATGGATTTAACGGGTTTGAGTATTGA
TAGTTTGTATGACGTCGGGTTTTTTAGGGNNNNANTTTTCGGGGATTCGGTAT
GTGGGATGGTAGTTGNGAGGATATT ATCGAAAGATATGNTNNNNNN

B5F-B5R
NNNNNGNGNGANNTNATATGGATTTAACGGGTTTGAGTATTGATAGTTTGTA
TGACGTCGGGTTTTTTANGGTTTTATTTTTCGGGGATTCGGTATGTGGGATGG
TAGTTGTGAGGATATTATCGAAAGATATGATTTTTTGGTTTCGTTTTTATATCG
TTGGTGAAGTGAGATGGGGAGGTGTGAAGTATGATTTAGA TGATAGGGNN

B6F-B6R
NNNNNNNNANGANNGGGNNGANGTAAGGAGAATGNATTTAAGGTTATTGTT
TAAGGTTTTTTTTGAAGTGGTTAAGATAAGGAGTTTATTTTGGAAGGATTATG
AGATGGATGTTATATATTTTTTGTAATTTATTTTGAGGAATGTTTTTAAGGATA
GTGAAATTATTGATGTGAATATAAAGGTTATTTATT**C**GAAGTTAAATGATTTT
AAAATTGATGGTATATAAGAATTAGAAGTTGTGATTAGTGAGATGGTT**C**GTT
TAATTGAGATTGTTA**C**GGTGTTAATATTGG**C**GGTTGATTTTGATGGATTGGTT
AATGGTTGGAATA**T**GAAAATTGTTGAGTTGATTGGTTTTT**C**GGTTGATGAAGT
AAT**C**GGGAAGTATTTTTTTATATTTGTTGAAGATTTTTTAGTGGAAANNGNNN
AAAGGATGN

B7F-B7R
NNGAAGTTGAGTTTATAATTTTATTTGTTTATTTATGTNNNNNNNNNNGGAGT
ANAATGTTTAGTTTGAGATTAAGATATATTNNNTTNGGGTTGATGTTGGGTTA
ATAAGTTTAGTTGTAAATGTATGCGTAAGTAGAGATTTTTATGAAAATGTGGT
TGGGGTGTGTTTTGTAGTTTATGATTTTATTGGTTANAAGATTGTGATGGATA
AGTTTA**C**G**C**GGATTGAAGGTGATTATAAGGTAATTATTTAAAATTTAAATT**C**
GTTGATTT**C**NNTAATATTTGGTAT**C**GATGAGTTTGGATGGTGTATAGAGTGGA

ATTTAGNAATGTTAAAGTTAAT**C**GGTTTGAAG**C**GAGAGGAAGTGATTGANAA
AATGTTTTTAGGAGAANNATTTGGNAAN

B7F-B8R
NNNNNNNNNNNNNTGTGNNGGNNNNNTTNA**C**G**C**GGATTGAAGGNGATTATA
AGGTAATTATTTAAAATTTAAATT**C**GTTGATTT**C**GTTAATATTTGGTAT**C**GAT
GAGTTTGGATGGTGTATAGAGTGGAATTTAGTAATGTTAAAGTTAAT**C**GGTTT
GAAG**C**GAGAGGAAGTGATTGATAAAATGTTTTTAGGAGAAGTATTTGGGAA

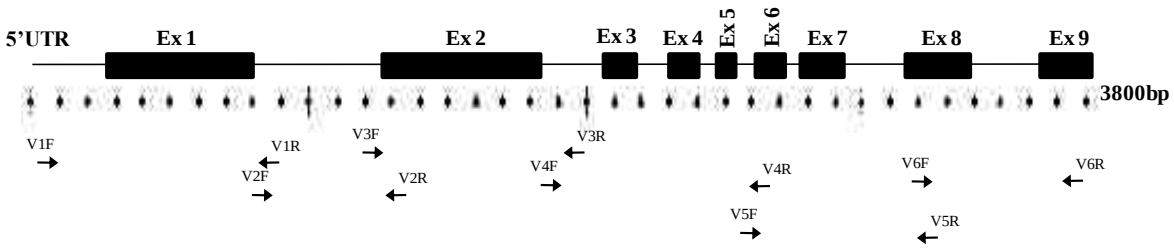
B9F-B9R
NNNNNNNNNNNNNAGANTTATGTTGN**C**GTTTAAAGAATTAAGAAGTTTTTGT
AAATTTTGGGATTGTGTTGAATAATGTTGTGATTAGTTAAGATTTAGAGAAAG
TATTGTTTGTTTTTTTTTATAAGAGGTGGTAAGTATGTGGAGTGTTTGTTGTGTG
TGAGTAAGAAATTGGATAGGGAAGGTGTAGTGATAGGTGTTTTTTGTTTTTTG
TATTTGTTAGTTATGAGTTGTAGTAAG**C**GTTTTATGTTTAA**C**GTTTAGTTGAG
CGAAT**C**GTAGTGAAGAGATTAAAGGTTTTAGTATATATAAAAAGATAGATTA
GGAATTTGTTATTTGGGATTATGTTTATAAGGAAAATGATAGAGGGTATTGA
ATTAGGATTAGAGTAAAGATGGATTTTGTAAATTAG**C**G**C**GTTATGTTAGAAG
TAATTAAGTAAGATTTT**C**GATGATT**C**GGATTTTGAAAGTATTATTGAAGGGTA
TGTTAGAANTGAAAGTN

B10F-B10R
NNNGTNNNNNNGTGTTTTTATNTTGGAATTTGATTGNGTAGATGTTTGGATTT
GGAAATGAAANNNTTTNTTTTAAATGAAGTGTTGATTGTTTTTATAGTTAAGT
NNTGATGAAGAGTAA**C**GGAAAGAGTGTT**C**GGATAATAAATGAGAT**C**GGAGA
AGAAGTAATGTTTGATATTTTGTATGGAGATAGTATTAGGTTTTAATAAGTTT
TGGTAGATTTTATGTTGATGGTTGTAAATTTTATATTATTTGGAGGTTAGTTAA
TTGTTTTAGTTTTTTTGAGGAAGGATTAGTTT**T**GGGTGTTTTGTGTATTTTGTTA
ATTTAGAGATTAGGTAAAATAGATATATTTATAATTAGAATGTATAATATGAT
GTTAAAAATAGTGTATTGAGGATATTGGATTATGAAAATAGGTTAATGTATA
TTGGAGTTGGGATATTTGAGTTTTTATTAAATTAAATGTTTGGGATTGAGGAA
GATGTGTTAGAAGAAGGATTGAGTTTAATGG

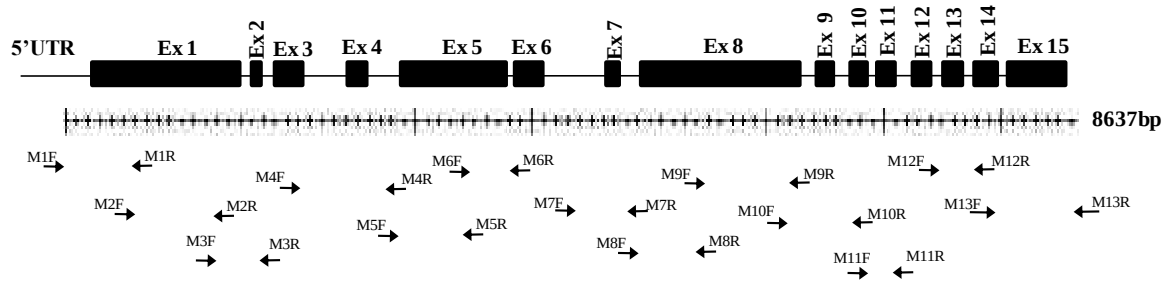
B11F-B11R
NNNNNNNNNNGNNNNNNNGNNNNTTGAGTTTTTATTAAATTAAATGTTTGGG
ATTGAGGAAGATGTGTTANAANAAGGATTGAGTTTAATGGTTAGTTGGAAAT
TGGTGAAGTTGATGAATGGAGATGTTTAGTATTTGAGATAAGTTGGGAAATT
AAGTTTTATTATTATTGTGGAATTTGTTGTAGTAAATAAGTAGTTTTTAAAAG
AAAAGGGGTTTGGTTTGATATAAAATAGTTATTGGTTGTTTTTTGTTTGTAATT
TTTTTTATTGTTTTTGTTTTTGTTTTTAAATTTTAGTAATGATGAAATATTTATT
TATTTATATTTTTTGTTGAATTTTTTTTTGAAGTTGTAAATATGGATGTATATT
TAATTTTTTTTTGAGTGTTTTGGTTTTATGATGTATTATTTAAAATAAGTAATT
ATAGGAGAAATGAGA

Appendix E

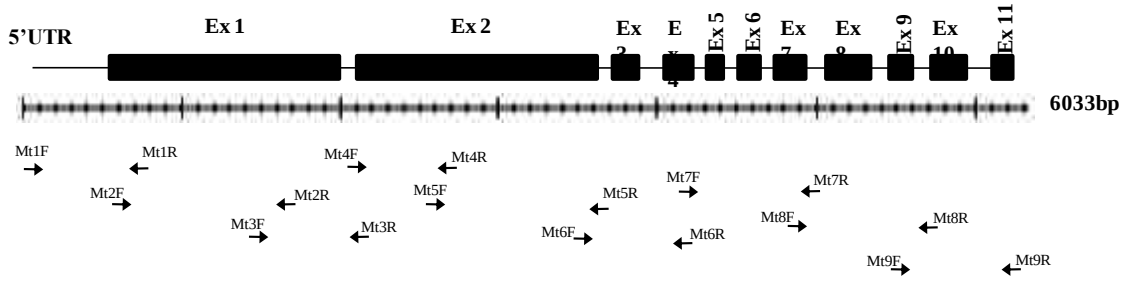
Locus: AT1G57820 (*VIM1*)



Locus: AT1G08060 (*MOM1*)

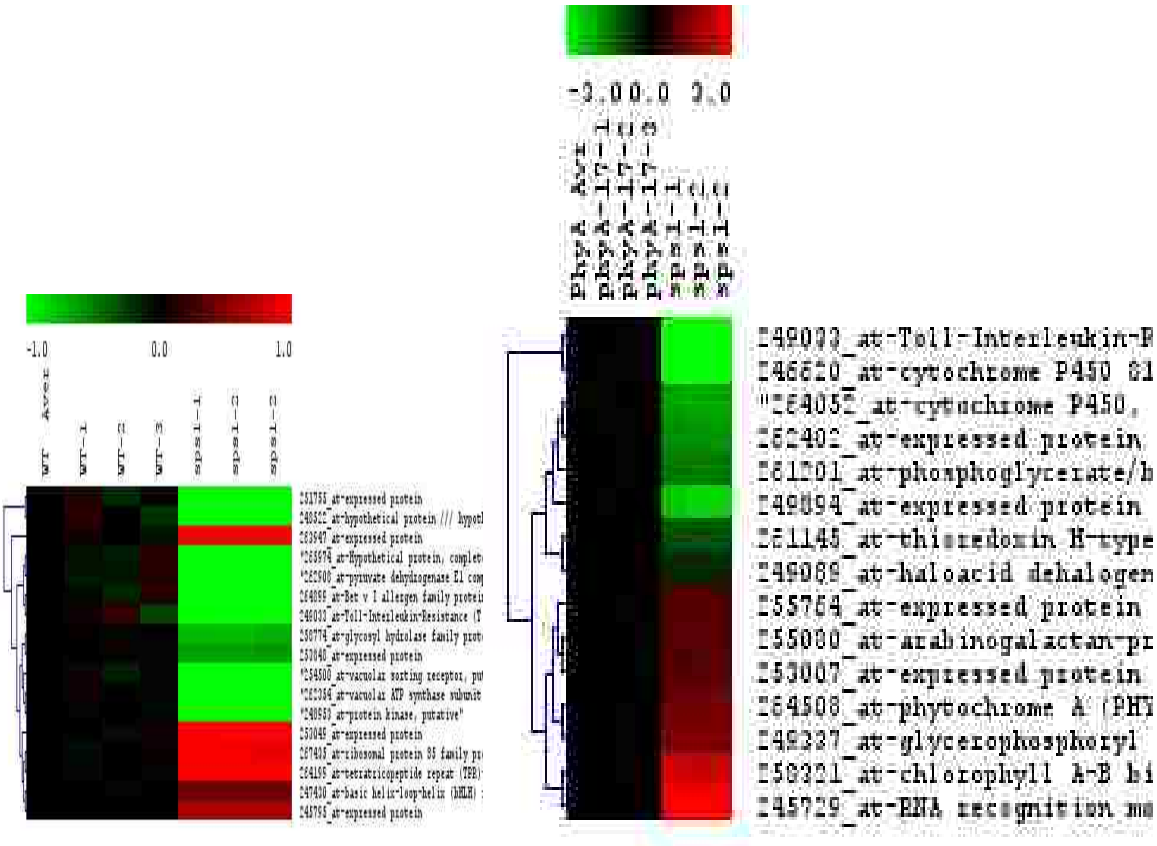


Locus: AT5G49160 (*MET1*)



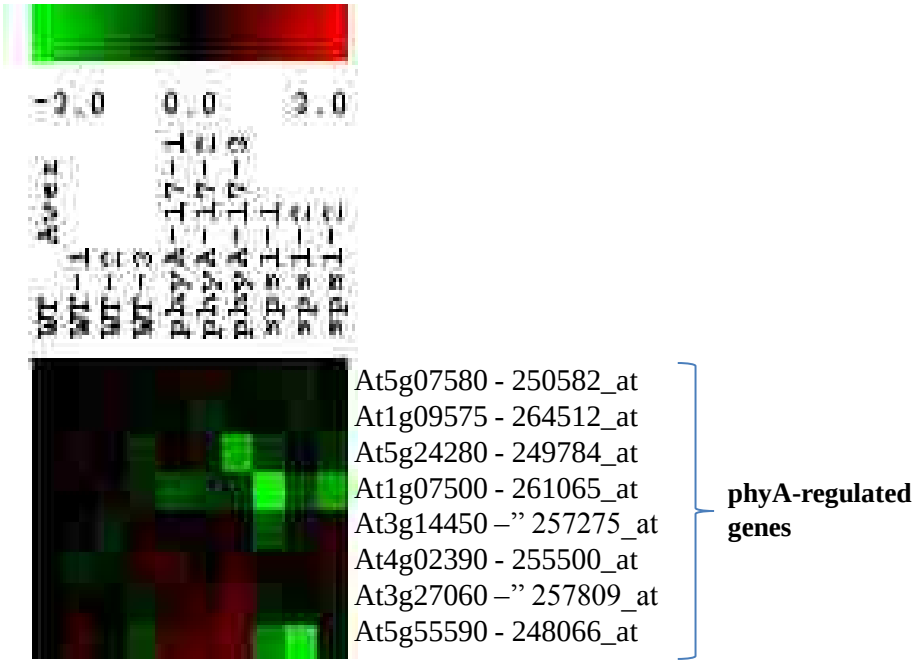
Appendix F

(a) (b)



Appendix F: Hierarchical cluster analysis (heat map) of the microarray data demonstrating the differentially expressed probe sets by three genotype, phyA-17, Col-0 and sps-1. Each sample (genotype) was assayed in triplicate. The bar at the top indicates the magnitude of change in expression. Green indicates induction and red indicates depression of gene expression. A heat map and relational dendrograms showing the average fold-changes in expression in (a) sps-1 line compared with WT and, (b) sps-1 compared with phyA-17

Appendix G



Appendix G: Heat map of the microarray data demonstrating the expression profile of PHYA-regulated genes in three genotypes: phyA-17, Col-0 and sps-1.