Identification and characterization of dominant suppressors of SUVH2 overexpression in *Arabidopsis thaliana*

Dissertation

zur Erlangung des akademischen Grades

Doctor rerum naturalium

(Dr. rer. nat)

Vorgelegt der

Naturwissenschaftliche Fakultät I – Biowissenschaften
Martin-Luther-Universität Halle-Wittenberg

von **Herr Deepak Jain**
geb. am: 14.02.1974 in: Calcutta, Indien

Gutachter:
1.
2.
3.
Halle (Saale), 29.6.10

Tag der mündlichen Prüfung.....

I dedicate this thesis to my parents. They, more then anyone else, have been the best role models I could have ever hoped for. Any accomplishment of mine are due in no small part to their support.

Abbreviations

Amp^r	Ampicillin resistance
bp	base pairs
BAC	Bacterial Artificial Chromosome
BAR	Basta-Resistance
CaMV	Cauliflower Mosaic Virus
DAPI	4', 6-Diamidino-2-phenylindole
DNA	deoxyribonucleic acid
dATP	2´-Desoxy-Adenosintriphosphat
dNTP	deoxynucleoside-triphosphate
EDTA	Ethylendiaminetetraaceticacid
EMS	Ethylmethansulfonat
h	hour
H3K9	Lysine 9 of histone H3
H3K27	Lysine 27 of histone H3
H4K20	Lysine 20 of histone H4
HMTase	Histone methyltransferase
HDAC	Histone deacetylase
kb	kilobases
kDa	kilodalton
LB	Luria Bertani
LRR	Leucine Rich Repeat
MB	Megabases
MBD	methyl-CpG-binding domain
mg	Milligrams
pH	potentia Hydrogenii
PBS	Phosphate Buffered Saline
PCR	polymerase chain reaction
PTGS	Post Transcriptional Gene Silencing
RNA	ribonucleic acid
RNase	ribonuclease
rpm	rounds per minute

SDS	sodiumdodecylsulfate
siRNA	Small interfering RNA
Su(var)	Suppressor of Variegation
T-DNA	Transfer-DNA
TGS	Transcriptional Gene Silencing
Tris	tris(hydroxymethyl)aminomethane
U	unit
WT	Wild type
X-Gal	5-Bromo-4-chloro-3-indoxyl-β-D-galactoside

Table of Contents

Abstract

The aim of this work was to isolate and characterize new suppressors, based upon SUVH2 overexpression phenotype, resulting in tagging of genes involved in chromatin regulation in *Arabidopsis thaliana*. A novel strategy using T-DNA mutagenesis was successfully established by introducing a second T-DNA in the background of SUVH2 overexpression plants to isolate dominant suppressor mutations. The mutant T-DNA was driven by a weak nos promoter to avoid Tran inactivation and silencing of the SUVH2 overexpression construct. *bar* gene, under the control of nopaline synthase (nos) promoter, provided resistance against herbicide spray containing phosphinothricin. Suppressors with a single locus were identified by checking the segregation ratio of progeny in T2 generation as well as by appearence of a single band in southern blot. The first genetic screen was carried out using heterozygous over expression plants, as homozygous line was sterile. Using posttranscriptional gene silencing, the obtained homozygous overexpression plants in F1 generation were used in 2nd screen for T-DNA mutagenesis.

The expression of SUVH2 transgene was verified in the background of the mutant T-DNA using immunocytology and real-time PCR analysis and methylation status was examined using Southern blot analysis. All three tests proved that the expression of the transgene was not influenced by the introduction of another T-DNA used for transformation, therefore ruling out the possibility that the isolated mutations were false positive. Genetic crosses involving SUVH2 overexpression plants with randomly inserted SALK line further validated these results.

Novel mutations showing dominant suppressor effect on SUVH2 overexpression phenotype were identified and the genomic sequences flanking the T-DNA were isolated by I-PCR. Most of the identified loci were novel, not having been identified from the other genetic screens. Based upon the immunocytological data, the mutants were categorized into two classes depending upon the ectopic distribution of H3K9me2 is retained or reverted back to that of wild type in the background of SUVH2 overexpression. The most interesting candidate protein that was found in the genetic screen was mutation in the bromodomain containing protein. It requires immediate further attention. The presence of a nuclear targeting signal, a bromo domain and the presence of a putative RNA binding motif makes this protein a very interesting candidate for further analysis. Genetic crosses using the commercially available SALK line confirmed the suppressor effect and the insertion point of all the isolated mutants.

Taken together this work led to the establishment of a new genetic screen to isolate dominant suppressor mutations, which could elucidate the signal cascade leading to heterochromatin formation in *Arabidopsis thaliana.*

Zusammenfassung

Das Ziel dieser Arbeit war die Isolierung und Charakterisierung neuer Suppressoren, unter SUVH2-Überexpression. Diese Suppressoren sollen zur Identifikation von Genen führen die an der Regulierung der Chromatinstruktur in *Arabidopsis thaliana* beteiligt sind. Eine neuartige Strategie der T-DNA-Mutagenese wurde durch die Einführung einer zweiten T-DNA in die SUVH2 Überexpressions-Pflanzen erfolgreich etabliert. Diese Pflanzen sollten dazu dienen dominante Suppressor-Mutanten zu isolieren. Die mutierte T-DNA stand unter Kontrolle des schwachen nos-Promotors um Trans-Inaktivierung und Silencing des SUVH2 Überexpressions-Konstrukts zu vermeiden. Herbizidresistenz gegen Phosphinothricin wurde durch das *bar* gene unter der Kontrolle des nos-Promotors erzeugt. Die Suppressoren mit einem einzelnen Locus wurden durch Überprüfung des Teilungsverhältnisses der Nachkommen in der T2-Generation und durch das Auftreten einer einzigen Bande im Southern-Blot ermittelt.

Der erste Screen wurde mit heterozygoten Überexpressions-Pflanzen durchgeführt, da die homozygote Linie steril war. Durch die Verwendung von posttranskriptionellem Gen-Silencing konnten homozygote Überexpressionspflanzen in der F1-Generation erzeugt werden, die im zweiten T-DNA Mutagenese-Screen verwendet wurden. Die Expression des SUVH2 Transgens in den Pflanzen die mit mutierter T-DNA behandelt wurden, wurde durch Immunzytologie und real-time PCR-Analyse nachgewiesen. Der Methylierungs-Status wurde mittels Southern-Blot-Analyse überprüft. Alle drei Tests bewiesen, dass die Expression des Transgens nicht durch die Einführung der zusätzlichen T-DNA, die für die Transformation verwendet wurde, beeinflusst wurde. Dies schloss die Möglichkeit aus, dass die isolierten Mutationen falsch positiv waren. Kreuzungen der SUVH2 Überexpressionspflanzen mit einem SALK-Stamm mit zufällig inserierter T-DNA validierten diese Ergebnisse zusätzlich.

Neue Mutationen die einen dominanten Suppressor-Effekt auf die SUVH2 Überexpressionpflanzen haben konnten identifiziert werden und die flankierenden genomischen Sequenzen der T-DNA wurden durch I-PCR isoliert. Die meisten der identifizierten Loci waren bisher unbekannt, da sie nicht von anderen genetischen Screens identifiziert werden konnten. Basierend auf den immunzytologischen Daten wurden die Mutanten in zwei Gruppen eingeteilt. In der ersten wurde die ektopische Verteilung von H3K9me2 beibehalten während sie in der zweiten Gruppe zum wildtyp der SUVH2 Überexpression zurückkehrt. Der interessanteste Kandidat, der durch den

genetischen Screen gefunden wurde, war eine Mutation eines Gens, das für ein bis dato unbekanntes Protein mit Bromodomäne codierte. Dieses Protein verdient weitere Aufmerksamkeit, vor allem da es eine Kernlokalisierungssequenz, eine Bromodomäne und ein mutmaßliches RNA-Bindungsmotiv besitzt. Kreuzungen mit dem kommerziell erhältlichen SALK-Stamm bestätigten die Suppressor-Effekte und die Insertionsstelle aller isolierten Mutanten. Zusammengefasst führte diese Arbeit zur Etablierung eines neuen genetischen Screens um dominante Suppressor-Mutationen zu isolieren die dabei helfen können die Signalkaskade die zur Bildung von Heterochromatin in *Arabidopsis thaliana* führt aufzuklären.

1.Introduction

DNA is the basic genetic material in all living cells. Although the DNA sequence in all cells of an individual is identical, each cell maintains its identity during differentiation, by expressing or repressing specific sets of genes. One of the hot topics, in modern research is to find out how such tight control of gene expression and repression is initiated and maintained. DNA is tightly packed around an octamer of histones (two H2A-H2B dimers and an H3-H4 tetramer) to form a nucleosome. Histones act as a spool around which the DNA winds. This enables tight compaction, required to fit the large amount of DNA in the nuclei of eukaryotic organisms. The lightly packed part of chromatin constitutes the euchromatin, which is rich in genes and under active transcription. In contrast heterochromatin is the more condensed part of chromatin, which is transcriptionally inert. Heterochromatin can be further differentiated into facultative and constitutive heterochromatin depending upon whether the replicating material remains transiently or permanently in condensed form.

In plants, genetic control of histone modifications are much more complex as compared to other eukaryotes. This could partially be attributed to the sessile nature of plants, as they have to cope and adapt to various environmental factors and accordingly maintain a dynamic state of chromatin (Pfluger *et al*., 2007). This assumption is also supported by the fact that most of the genes in *Arabidopsis* are part of multigene families, which allows easy fine-tuning of their expression. For instance there are 47 SET-domain proteins (Baumbusch *et al*., 2001; Ng *et al*., 2007), 12 histone acetyltransferases and 16 histone deacetylases (Probst *et al*., 2004). Single mutations affecting only one of the genes in *Arabidopsis* seldom lead to a severe phenotype owning to functional redundancy. This could be useful to study lethal mutations, which would eliminate the individuals in other model systems (Matzke and Schied 2007).

The modulation of chromatin structure is critical for fine-tuning and differential gene expression. The accessibility and recruitment of the regulatory factors to chromatin determines whether the underlying DNA can either achieve a closed, condensed and transcriptionally inert state or an open, less condensed, transcriptionally active state. Different transcription states can be achieved not only by covalent histone modifications, but also by variation in the local chromatin structure (Quina *et al*., 2006).

The key regulators of chromatin structure, which can potentially modulate gene expression, include the Nucleosome Assembly Proteins (NAP), which are important for incorporation of core histones (Lafos and Schubert, 2009), ATP-dependent chromatin remodelers, which can alter the position and the conformation of nucleosomes and histone variants, which are generally associated with actively transcribed regions as was shown in *Drosophila*. In

Drosophila, histone H3 is replaced by the variant H 3.3, which is found to be abundant in actively transcribed regions (Walter.M, personal communication). The histone variant H2A.Z, which is generally found in the regions of active transcription, specifically associates with regions of low methylation, suggesting a possible protective role in gene silencing (March-Diaz and Reyes, 2009). Modulators like activators or repressors, which bind and mark the chromatin, followed by recruitment of other remodelling factors to nucleate or maintain a particular chromatin state. However, histone modification and DNA methylation are the two most important and conserved covalent modifications for the regulation and proper function of chromatin.

1.1 Epigenetics and its role in plant development

During plant development there is a tight control of gene expression in a time-dependent fashion. The expression patterns are established and maintained during the successive rounds of cell division. PcG proteins play an important role in controlling the expression of homeotic genes. Two groups of PcG proteins form independent protein complexes called PRC1 and PRC2. The PRC2 complex is responsible for the establishment of the repressive H3K27me3 (trimethylation of lysine 27 of histone H3) mark on its target genes and this complex is conserved in plants as well as in animals. However, in contrast to animals, biochemical and interaction studies suggest the presence of three PRC2-like complexes with distinct and overlapping roles in *Arabidopsis* development (Kohler *et al*., 2003; Chanvivattana *et al*., 2004; Wood *et al*., 2006). The key component of the PRC2 complex is Enhancer of Zeste E(Z), which encodes a SET domain-containing histone H3 methyltransferase protein responsible for mono-, di- and trimethylation of H3K27 (H3K27me1 to me3). In *Arabidopsis*, E(Z) proteins are represented by a small gene family consisting of CURLY LEAF (CLF), SWINGER (SWN), and MEDEA (MEA). These genes act in a partially redundant manner in regulating the target genes (Chanvivattana *et al*., 2004; Schubert *et al*., 2006; Wang *et al*., 2006). The PRC1 complex is shown to be responsible for maintenance of silencing initiated by the PRC2 complex. Absence of the PRC1 complex in *Arabidopsis* suggests that PcG regulation is much more dynamic in plants as compared to animals.

Trithorax proteins are localized to the same loci as PcG proteins but act in an antagonistic fashion by stimulating transcription. Trithorax proteins catalyze the trimethylation of H3K4 (Czermin *et al*., 2002). There are five homologs of TRX proteins in *Arabidopsis*, of which ATX-1 and ATX-2 have been studied in detail. Histone methyl transferase activity of ATX-1 specific for H3K4me3 mark has been shown (Alvarez-venegas *et al*., 2003). In contrast to

animals, the major epigenetic changes in plants not only occur at the early embryonic stage but also during post-embryonic developmental transitions thus emphasizing the important role of chromatin remodeling in providing higher flexibility during the entire life cyle of plants.

1.2 DNA methylation and its significance in gene silencing

DNA methylation is a typical epigenetic mark for transcriptional gene silencing in many organisms. It is correlated with silencing of transposons, genomic imprinting and X chromosome inactivation. In mammals DNA methylation is preferentially found at symmetric or CpG methylation. DNMT1 and DNMT3 are the two-methyltransferase enzymes, required for maintenance and de-novo methylation in animals (Margot *et al*., 2003). In *Neurospora*, *dim-2* is the only gene responsible for methyl transferase activity and mutation of this gene leads to complete elimination of DNA methylation (Kouzminova and Selker, 2001).

In *Arabidopsis,* DNA methylation occurs in both symmetrical and asymmetrical motifs, thus leading to methylation at specific sites such as CG, CHG, CHH (where H = C, G, A or T). DRM1, DRM2, MET1, DRD1 and CMT3 are the principal *Arabidopsis* DNA methyltransferase enzymes, which transfer and covalently attach methyl groups onto cytosine residues in the DNA. MET1, a homolog of the mammalian DNA methyltransferase DNMT1, is responsible for CG methylation or symmetric methylation. The maintenance of DNA methylation is primarily carried out by MET1, as it recognizes methylation marks during replication from the hemimethylated parent substrate and transfers them to the daughter strand (Henderson and Jacobsen, 2007; Tariq *et al*., 2003). In the mutant background of *met1*, phenotypic defects appear only after several generations of inbreeding indicating functional redundancy among different methyl transferases (Mathieu *et al*., 2007).

Mutations in MOM1 (Morpheus molecule 1) have been shown to release transcriptional gene silencing (TGS) in the absence of any detectable alteration of DNA methylation (Amedeo *et al*., 2000; Vaillant *et al*., 2006). This might indicate an alternative pathway of silencing, which is independent from DNA methylation. Recently it was suggested that MOM1 might participate in maintaining the silent state of target loci by recognizing specific patterns of cytosines and histone modification controlled by the RNA-directed DNA methylation (RdDM) pathway. It was also postulated that MOM1 maintains H3K9me2 at an intermediate level, allowing chromatin to remain in a more dynamic state (Numa *et al*., 2009; Yokthongwattana *et al*., 2110).

DDM1 (Decrease in DNA methylation) is a member of the SWI/SNF2 family of proteins. In a null-mutant background of *ddm1*, substantial reduction of DNA methylation is achieved,

emphasizing the important role of this chromatin-remodeling enzyme (Law and Jacobsen, 2009). DDM1 binds to the nucleosome and promotes chromatin remodeling in an ATP-dependent manner, indicating its indirect role in DNA methylation (Brzeski and Jerzmanowski, 2003). The loss of functional alleles of *ddm1* leads to transcriptional reactivation of transposons and repeats in heterochromatic regions. This could be attributed to H3K9 redistribution throughout the genome as compared to wild type, where this mark is mainly restricted to heterochromatin, resulting in a general shift from a repressive H3K9 mark to an active H3K4 mark in the heterochromatic region (Gendrel *et al*., 2002). In the null-mutant background of *ddm1,* the phenotypic defects became successively stronger in progressive generations (Kakutani *et al*., 1996). MBD (methyl binding domain protein) proteins associate with both methylated DNA as well as histone deacetylases in plants as well as in animals emphasizing their crucial role in the crosstalk between DNA methylation, histone modification and heterochromatin formation. DDM1 was shown to be responsible for localization of AtMBD proteins at specific nuclear domains (Zemach *et al*., 2005).

Asymmetric methylation is mainly controlled by DRM1 and DRM2 (Domain Rearranged Methyltransferase). They are mammalian homologs of the Dnmt3 group, whose members are responsible for the establishment and maintenance of asymmetric methylation (Henderson and Jacobsen, 2007). siRNA plays an important role in directing these methyl transferases, which are responsible for de novo methylation of the target sequences. RNAi-mediated DNA methylation is crucial for DNA methylation in a symmetric as well as in an asymmetric context (Law and Jacobsen, 2009).

CMT3 (Chromomethylase) is a DNA methyltransferase, which is unique to *Arabidopsis* and plays an important role in maintaining non-CG methylation (Bartee *et al*., 2001; Cao *et al*., 2003; Jackson *et al*., 2002).

In the mutant background of *drd1* (Defective in RNA directed DNA methylation), RNA - induced non-CpG methylation was completely eliminated at the target promoter, implicating involvement of DRD1 in RNA-directed de novo methylation (Kanno *et al*., 2004). DRD1 acts in concert with the RNA polymerase 1V/RDR2/DCL3/AGO4-mediated RNAi pathway, to guide DRM2 to carry out de novo asymmetric DNA methylation (Chan *et al*., 2006; Law and Jacobsen, 2009).

1.3 The role of non-coding RNA in gene silencing in *Arabidopsis thaliana*

Recently, the role of noncoding transcripts in gene silencing and heterochromatin formation has gained considerable importance. In particular, it was shown that noncoding RNA

regulates the expression of the major floral repressor gene FLC. This repression was mediated by localized demethylation of the antisense transcript without affecting the transcription of the sense transcript. Upregulation of noncoding antisense transcripts leads to downregulation of FLC expression. This is the earliest step in silencing prior to Polycomb group proteins (Liu *et al*., 2010; Swiezewski *et al*., 2009). The role of RNA polymerase II in siRNA-mediated silencing was shown by demonstrating that it codes for noncoding transcripts, which acts as a scaffold in the siRNA-mediated pathway. The scaffold also interacts with Argonaute 4 to recruit siRNA to homologous loci resulting in TGS. The Ago4/siRNA/RISC complex could then recruit DNA methyltransferases such as DRM2 for de novo methylation or histone methyltransferase resulting in epigenetic modification. POL II transcripts also recruit Pol 1V and Pol V to different heterochromatic loci resulting in siRNA-mediated TGS. DRD1 is required for the production of Pol II transcripts (Zheng *et al*., 2009). RDM4 is a conserved protein, which interacts with Pol II and Pol V and thus serves as a important component in the RNA-mediated silencing pathway (He *et al*., 2009).

Recently arginine methylation has been shown to play an important role in various processes such as DNA repair, signal transduction, RNA processing, and transcriptional regulation. Arginine methylation is carried out by the enzyme arginine methyltransferase (Bedford and Richard, 2005).

1.4 DNA demethylation in the *Arabidopsis* genome

Methylation status of various genes is dynamically regulated by methylation and demethylation. DEMETER and ROS1 have been studied in detail for their role in active demethylation in *Arabidopsis*. Demethylases remove 5-methyl cytosine residues by their glycosylase and lyase activity. DEMETER is expressed in the central cell of female gametophyte. It is required for the maternal expression of two imprinted genes - MEDEA and FWA. MEDEA and FWA remain silent as a result of DNA methylation. On the other hand, ROS1 actively removes 5-methyl cytosine from the hypermethylated transgene (Agius et al., 2006; Morales-Ruiz et al., 2006). Recently demethylase activity of another protein ROS3 was shown. The binding site of this protein colocalizes with ROS1, indicating that both of them are in the same genetic pathway. ROS3 has a RNA binding motif and therefore might be involved in RNA mediated target sequence recoginition for its demethylase activity (Zheng et al., 2008). In a recent study, it was shown that a putative histone demethylase jmjC domain protein IBM1 (Increase in Bonsai Methylation1) prevents ectopic distribution of non-CG methylation in genic regions of *Arabidopsis*, which is absolutely essential for normal

development. It was also postulated that increase of DNA methylation in the mutant background *of ibm1* could be due to an increase in H3K9 methylation. Interestingly RNAi components involved in DNA methylation were not involved in DNA hypermethylation in the i*bm1* mutant background (Miura et al., 2009; Saze et al., 2008).

1.5 Influence of histone modifications on regulation of gene expression

The N- terminal tails of the histones H3 and H4 histones are protruding out from the nucleosomes and can be post translationally modified at several amino acid positions. The modificaion of the tail includes acetylation, methylation, phosphorylation, ADP-ribosylation and ubiquitination (Strahl and Allis, 2000). Dependent on the type and position of the modification, the chromatin can either be in a transcriptionally active, open conformation or a transcriptionally inactive, closed condensed form. Thus combinations of these modifications constitute an epigenetic histone code. The Histone code is not universal but varies from species to species. The effect of all these modifications are interpreted by the chromatin associated proteins, which actually decide whether a transcriptionally active or silent state has to be maintained at a local level (Jenuwein and Allis, 2001). In contrast to animals, none of these proteins are known in plants.

1.6 Types of histone modifications

1.Histone Methylation: Histones are methylated at specific lysine or arginine residues. Methylation at specific residues marks the chromatin as either transcriptional active or inactive, dependent on the type and place where the modification has taken place. Methylation of histones at H3K9, H3K27 and H4K20 are typical marks of the repressed chromatin state (Fischle *et al*., 2003; Lachner *et al*., 2001), whereas methylation of H3K4 and H3K79 are marks of an active chromatin state (Vakoc *et al*., 2005). Further, a different degree of methylation (mono-, di- and tri –methylation) results in tight control and modulation of this process. These marks are sometimes species-specific, as the mono-, di- and tri-methylation of K9 and K27 of histone H3 (H3K9me1 to me3 and H3K27me1 to me3) are specific marks for heterochromatin in *Drosophila* (Ebert *et al*., 2004). In *Arabidopsis*, H3K9me1 and me2, H3K27me1 and me2 and H4K20me1 are preferentially enriched in heterochromatin regions. In contrast, the histone modification marks H3K9me3, H3K27me3 and H4K20me2 and me3 were found exclusively in euchromatic regions (Naumann *et al*., 2005).

2. Histone acetylation / deacetylation: Another well-studied modification is the acetylation / deacetylation of histones. This process is carried out by histone acetyltransferases (HAT) and histone deacetylase (HDACs). HDAC6 was shown to be involved in maintaining CpG methylation (Probst *et al*., 2004). HAT transfers an acetyl residue from Acetyl-CoA to the amino group of lysine whereas deacetylases reverse this reaction. Acetylation is linked to transcriptional activation as it neutralizes the positive charge, reducing the affinity between histone and the negatively charged DNA. This alters the accessibility for various nucleosomal DNA-binding transcription activators, other chromatin remodeling complexes and chromatin modifying enzymes (Pandey *et al*., 2002).

3. Phosphorylation: Aurora kinases are serine /threonine kinases that are known to be essential for cell proliferation. Three aurora kinases (A, B and C) have been identified, and have been shown to play an important role in chromosomal condensation and cell cycle progression during mitosis and meiosis. These enzymes phosphorylate histone H3 at Ser10 (Kawabe *et al*., 2005; Kurihara *et al*., 2007). Like histone methylation, histone phosphorylation is also a species-specific mark, as seen in *Drosophila*, where H3S10 phosphorylation is a typical euchromatic mark (Ebert *et al*., 2006; Zhu, 2009). However, in *Arabidopsis* it is mostly heterochromatic (Fischer *et al*., 2006).

4.Histone demethylation: Histones can also be demethylated by specific histone demethylases at methyllysine and methylarginine residues. Aminooxidases like Jumonji (Jmjc) domain-containing proteins and LSD1 are responsible for lysine demethylase activity. LSD1 is an amine oxidase that catalyze lysine demethylation in a FAD–dependent (flavin adenine dinucleotide dependant) manner, whereas the hydroxylation reaction carried out by Jumonji (Jmjc) domain-containing proteins, requires the presence of Fe (II) and α-keto glutarate to generate formaldehyde and succinate (Zhu, 2009). Arginine demethylation is carried out by PADI4 (peptidyl arginine deaminase I4). It converts arginine in the histone H3 tail to citrulline (Bedford and Richard, 2005; Tanikawa *et al*., 2009)

1.7 SET domains proteins in *Arabidopsis thaliana*

SET domain proteins have histone methyl transferase activity. These proteins were first described in *Drosophila* and the name SET is derived from suppressor of variegation SU(VAR)3-9, Enhancer of Zeste E(Z) and Trithorax protein. It was shown that interaction of SU(VAR) 3-9 with HP1 is required for the establishment and maintenance of heterochromatin in *Drosophila*. The orthologues of SU(VAR)3-9 are conserved in different organisms (Figure1). In case of fission yeast, the Clr4 methyltransferase complex is responsibe for DNA

methylation and heterochromatin assembly in conjunction with the RNAi pathway (Zhang *et al*., 2008). Similarly, the human SUV39H (the human homolog of SU(VAR)3-9) can partially complement and rescue the silencing defect of SU(VAR)3-9 mutations in *Drosophila* (Schotta *et al*., 2002). It was also shown that SU(VAR)3-9 and HP1 from *Drosophila* specifically associate with the heterochromatin when expressed in *Arabidopsis*, indicating functional conservation of the proteins in different organisms (Naumann *et al*., 2005).

There are 47 SET-domain proteins in Arabodopsis (Baumbusch *et al*., 2001, Qian and Zhou, 2006; Ng *et al*., 2007). The different types and places of modification of lysine residues could possibly explain the existence of such a large family of SET-domain proteins. They are grouped into four classes, Enhancer of Zeste E (Z) homologues, Ash1 homologues, Trithorax (Trx) homologues and Suppression of variegation SU(VAR)3-9 homologues.

Enhancer of Zeste E (Z) homologues in *Arabidopsis* are CURLY LEAF (CLF), SWINGER (SWN) and MEDEA (MEA). CURLY LEAF and SWINGER act in a redundant fashion to control leaf and flower morphology as well as flowering time (Schubert *et al*., 2006). MEDEA (MEA) is an imprinted gene, which represses endosperm development in the absence of fertilization (Arnaud and Feil, 2006).

ASH1 homologues are a group of proteins that are homolog to ASH1 and are characterized by the presence of a cysteine-rich post-SET domain. In *Drosophila,* ASH1 (absent, small, or homeotic disc 1) has been shown to be required for H3K4 methylation (Byrd *et al*., 2003). Consistent with these results, it was hypothesized that these group members are involved in H3K36 methylation in *Arabidopsis* (Zhao *et al*., 2005).

Members of the group of Trithorax homologues were shown to be involved in positive regulation of homeotic gene expression, like their *Drosophila* counterparts. ATX1 was the first protein reported from this group to have H3K4 methyltransferase activity (Santos-Rosa *et al*., 2002). These classes of proteins are characterized by the presence of highly conserved domains such as PWWP, FYRN, FYRC and others.

SU(VAR)3-9 is a highly conserved protein and homologues are present in a wide range of organisms, where they catalyze the formation of the repressive H3K9me2 mark, via their SET domain (Bannister *et al*., 2001). The SET domain of SU(VAR)3-9 is responsible for H3K9me2 activity in chromo centers of *Drosophila,* which facilitates the recruitment of HP1 to SU(VAR)3-9 through its chromodomain (Schotta *et al*., 2002). The interaction between SU(VAR)3-9 and HP1 is an interdependent process (Bannister *et al*., 2001; Rea *et al*., 2000).

In conjugation with SET domain there are many other prominent domains, which are present in SU(VAR)3-9 and its homologues. Chromo domain is commonly found in proteins

associated with remodeling and manipulation of chromatin. Chromo domain containing proteins are involved in binding of methylated histones and also play an important role in the RNA-induced transcriptional silencing complex. SRA (SET and RING associated) and MBD domain containing proteins could directly bind to methylated DNA and therefore could play an important role in crosstalk between histone and DNA methylation. The presence of ankyrin repeat is one of the most common protein-protein interaction motifs found in nature. The additional domains present along with SET domain might provide a hint about their possible functional role of these proteins in different organisms.

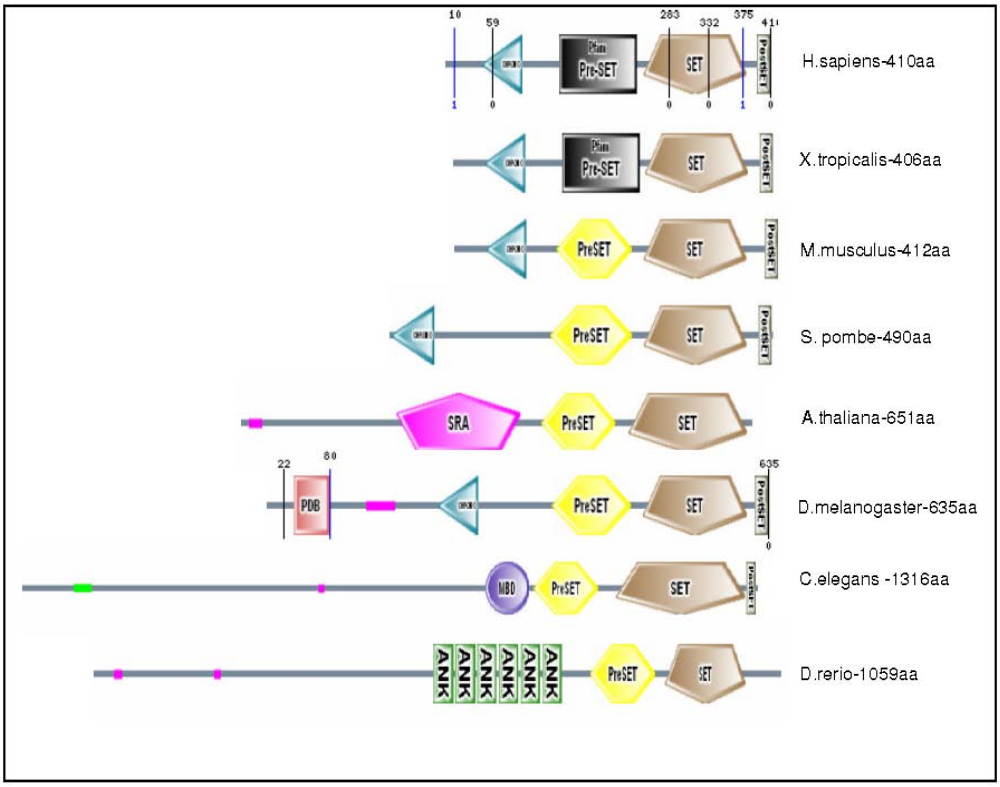


Figure 1.1: Comparision of the homologues of SU(VAR)3-9 protein. Homologues from *H.sapiens, X.tropicalis, M.musculus, S. pombe, A.thaliana , D.melanogaster, C.elegans* and *D.rerio* are shown. The highly conserved PreSET and SET domains, responsible for specific histone methyltransferase activity, are present in all the organisms.

In mammals a similar interaction of DNA methyltransferase Dnmt3a with the mammalian homologue of SU(VAR)3-9, SUV39H1 was shown, also involving HP1ß, the mammalian homologue of HP1 (Fuks *et al.,* 2003). This shows that SU(VAR)3-9-mediated heterochromatin formation is a highly conserved process in different organisms. Comparision of the domain structure in different organisms also revealed that this proteinis highly conserved as seen in Figure1.1.

There are 10 homologues of SU(VAR)3-9 protein in *Arabidopsis*. They all have a typical YDG domain upstream of the pre-SET domain. SUVH4/Kryptonite is one of the well-studied members of this class of proteins. It represses expression of *superman,* a floral homeotic gene through its H3K9 methyltransferase activity. This mutant was isolated in a genetic screen, where, an inverted repeat of the *superman* gene was introduced in a wild type background. This resulted in silencing of the transgene as well as the endogenous locus and formation of an epigenetic allele known as clark kent (*clk–st*) allele. *clk-st* plants showed a defect in the number of floral organs. EMS mutagenesis was performed to score for the suppressor phenotype. It led to identification of *cmt3, suvh4 ago4, drm1* and *drm2* (Jackson *et al*., 2002). *Suvh4* and *cmt3* mutations also suppress the hypermethylation phenotype of *ibm1*-mutation indicating that they work in a common pathway in a complex (Saze *et al*., 2008).

Out of 10 SUVH proteins, SUVH2 has been shown to have profound effect on heterochromatin formation in *Arabidopsis*. This statement is further substantiated by the fact that all the heterochromatin marks were reduced in the null-mutant background of SUVH2. Remarkably a pronounced effect was visibe in H4K20 monomethylation. Accordingly overexpression of SUVH2 leads to ectopic heterochromatization (Naumann *et al*., 2005). Till date, in vivo HMTase activity has been demonstrated for SUVH2, SUVH4, SUVH5 and SUVH6 (Naumann *et al*., 2005 Ebbs and Bender, 2006; Jackson *et al*., 2004). It was also shown that SUVH2 shows a dose-dependent effect and overexpression of SUVH2 leads to a mini plant phenotype. Such pivotal role of SUVH2 in *Arabidopsis* makes this protein an interesting and important candidate to further investigate its role in heterochromatin formation. Genetic interaction-studies with known DNA methyltransferases revealed a direct regulatory link between SUVH2, MET1 and DDM1. Hence a genetic screen was established to find dominant modifiers of the SUVH2 overexpression phenotype. This would in turn help to identify components, which could directly or indirectly interact with SUVH2 in a protein complex, leading to heterochromatin formation in *Arabidopsis.*

Apart from the 10 SUVH proteins, there are four SUVH related proteins, some of which are specifically localized to the nucleolus. SUVR4 was shown to have H3K9me1-activity, involved in control of rRNA expression (Thorstensen *et al*., 2006).

1.8 Role of SET domain genes in flower development in *Arabidopsis thaliana*

Vernalization is the process in which annual plants acquire competence to flower after long cold spells. VRN2 is required for silencing of the FLC locus by the polycomb group of

proteins through acquisition of H3K27 di- and tri-methylation. So, VRN1 is implicated in maintenance of the silent state (Bastow *et al*., 2004; Henderson and Jacobsen, 2007). EMF1 and EMF2 are embryonic flower genes, whose loss of function leads to flowering, bypassing the vegetative shoot growth (Yoshida *et al*., 2001). LHP1 is shown to be responsible for leaf morphology and onset of flowering but does not play any role in heterochromatin formation as its *Drosophila* counterpart (Mylne *et al*., 2006).

1.9 Influence of transcriptional gene silencing (TGS) and post transcriptional gene silencing (PTGS) in regulation of gene expression

Transgene silencing is the major reason for variability in transgene expression and could be attributed to either TGS (transcriptional gene silencing) or PTGS (post-transcriptional gene silencing (Paszkowski and Whitham, 2001). Since this variability in expression does not result from changes in the DNA sequence of the transgene, these changes could be attributed to epigenetic effects.

Transcriptonal gene silencing is characterized by the hyper methylation of the promoter sequence and inhibition of transcription, whereas in post transcriptional gene silencing there could be methylation of the coding region and potential post-transcriptional degradation of the RNA (Mittelsten Scheid *et al*., 1998). When a transgene is present in multiple copies in tandem or inverted repeats in close proximity to each other, transcriptional gene silencing (TGS) is triggered. TGS is mitotically and meiotically stable (Steimer *et al*., 2000). It is characterized by DNA hypermethylation, histone methylation and deacetylation, indicating the formation of a local heterochromatic environment (Hofmann 2004).

Post-transcriptional gene silencing (PTGS) may results from the presence of an antisense transcript, inverted DNA repeats, RNA virus or overexpression of a transgene (De Wilde *et al*., 2000). The silencing effect is much more pronounced in the presence of transgene loci in homozygous state (Vaucheret *et al*., 1998). In a way, PTGS helps to control which genes are active and how much expression is optimum for normal growth. The first step in this process is initiated by Dicer, which cleaves double stranded RNA into shorter fragments of 21-22 nucleotides. These fragmented nucleotides are responsible for mRNA cleavage and epigenetic modification of homologous DNA sequence. Fragmented RNAs are further amplified by the action of SGS2 and SGS3 proteins. SGS2 is a RNA dependent RNA polymerase (RDR6) and SGS3 protein has no functional similarity to any other known protein indicating that it is a plant specific protein required for silencing (Zheng *et al*., 2010). The generated RNA is then incorporated into the RISC complex, pairs with complementary mRNA, and cleaved by

Argonaute, the catalytic component of the RISC complex. PTGS is therefore a natural defense mechanism to prevent viral replication as well as to prevent self-propagation of transposons (Peragine *et al*., 2004).

TGS and PTGS are independent mechanisms, but in a molecular pathway both mechanisms might have common players like DDM1 and MET1 that could methylate target DNA sequence. Both types of silencing leads to the production of sequence specific RNA complementary to the region, which undergoes silencing.

1.10 Available systems for identification of TGS suppressors in *Arabidopsis thaliana*

There are many already established silencing systems to isolate TGS suppressor mutations in *Arabidopsis*. Suppressor mutants of transcriptional gene silencing were isolated for the first time using a transgenic silencing system developed by Mittelsten Scheid *et al*., (1998). In this screening, hygromycin was used as a reporter system in multiple copies and it resulted in a repeat-dependent hypermethylation and silencing of the construct. Due to the repetitive nature of the construct, this line showed stable inactivation of the reporter gene for many generations, resulting in sensitivity to hygromycin. After EMS mutagenesis, mutants were selected based upon reactivation of the hygromycin reporter. The reporter gene could only be transcribed in the background of the mutants defective in transcriptional gene silencing. Eight mutations were isolated and were called som1-8. Later, five of them were found to be *ddm1* alleles. *Ddm1* mutations are strong TGS suppressors in *Arabidopsis*. In a similar screen using T-DNA as a mutagen, MOM1 was isolated (Amedeo *et al*., 2000).

Another screen used additional copies of the CHS gene (chalcone synthase). When these copies were introduced in the wild type ectotype, they resulted in homology-dependent silencing and complete silencing of all homologous copies of the CHS genes (Davies *et al*., 1997). *hog1*-mutation (Homology-dependent gene S- silencing) could release silencing of the homology-dependent silencing state. The HOG1 gene codes for the S-adenosyl–homocysteine hydrolase, that is required for DNA methylation-dependent gene silencing (Rocha *et al*., 2005). In another genetic screen, *sil1* mutation was identified. This mutation is allelic to HDA6 (histone deacetylase 6) and leads to release of transcriptional gene silencing (Probst *et al*., 2004).

One of the well-studied silencing systems is based on the enzymes involved in the tryptophan biosynthetic pathway. The enzyme PAI (phosphoribosyl anthranilate) has a specific

genomic arrangement in the *Arabidopsis* ecotype Wassilewskija (WS). There are three independent loci that can potentially code for this enzyme. However PAI1 and PAI4 are arranged as a tail-to-tail inverted repeat. This complex arrangement leads to methylation of all the loci in an RNA -directed gene-silencing mechanism (Melquist *et al*., 1999). As a result of methylation, the PAI2 locus remains completely silent. Despite hyper methylation in the coding region, the PAI1 gene is expressed leading to the synthesis of an active PAI1 enzyme. Even after *met1* and *ddm1* mutations were introduced in this ecotype, they could only partially demethylate the complex loci indicating a possible role of RNA-directed DNA-methylation (Malagnac *et al*., 2002). In the same genetic screen, second-site mutations could also be isolated, that released silencing from the PAI2 locus in the background of a mutated PAI1 gene and resulted in identification of eleven alleles of CMT3 and seven alleles of SUVH4 (Bartee *et al*., 2001).

In recent studies carried out in our lab, tandem repeats of a 35S-driven luciferase transgene were used, which resulted in complete silencing of the luciferase transgene. EMS mutagenesis was performed in this background, and mutants were selected and identified based upon reactivation of luciferase transgene in the F2 generation. Extensive screening resulted in identification of a large number of transcriptionally silent mutants (Hofmann 2004).

1.11 Aim of the work

Heterochromatin plays an important role in gene regulation, in maintaining the integrity of the chromosome and in silencing of retrotransposons. One way to gain insights into the function of heterochromatin formation in *Arabidopsis* is to generate plants or identify mutants that contain defined phenotypic alterations. Overexpression of the heterochromatin protein SUVH2 results in a miniplant phenotype due to excessive heterochromatization. This provides a useful platform to design and screen for dominant mutations involved in heterochromatin formation based on the loss of the miniplant phenotype. Once such mutations have been identified, their role in heterochromatin formation can be assessed. Therefore detailed analysis of these mutants and corresponding genes would provide better understanding of heterochromatin formation in plants.

The aim of this work was to establish a screening system based on the SUVH2 overexpression phenotype in order to identify dominant suppressor mutations. Such mutations could result in identification of novel unknown suppressors, which could play an important role in the heterochromatic silencing process in *Arabidopsis*. It will also help to identify direct

and indirect interacting proteins of this pathway, which could potentially control the functions of SUVH2. In a previous study carried out in our lab it was shown that SUVH2 functionally interacts with well-known methyltransferase MET1 and DDM1. In this study the screening procedure was fine-tuned in order to obtain seeds from a sterile homozygous overexpression line after PTGS of the SUVH2 transgenes. This helped to screen for mutations in F1 plants at at a faster rate, after T-DNA transformation of insertional mutagen using Basta T-DNA construct. A large number of dominant mutations were isolated, which were initially differentiated into two types based upon H3K9 staining pattern. Type 1 in which ectopic H3K9 distribution was retained and type 2 in which the ectopic H3K9 level was reduce to wild type level in the SUVH2 overexpression background. One of the interesting mutations represents a T-DNA insertion in the gene for global transcription factor 2 (GTE2). The insertion was identified and Southern blot analysis was performed to confirm the presence of a single insert of the mutater T-DNA in the genome. With two independent SALK lines inserted in this gene, the suppressor effect was confirmed by crossing the SALK line to the starting overexpression line and scoring for a dominant suppressor effect in the next generation.

2. Materials and Methods

2.1 Materials used

2.1.1 Chemicals, solvents, enzymes and oligonucleotides

All laboratory chemicals and organic solvents were of analytical grade and obtained from Sigma-Aldrich (Deisenhofen, Germany), Diagonal (Münster, Germany), AppliChem (Darmstadt, Germany), Duchefa (Haarlem, Netherlands), CarlRoth (Karlsruhe, Germany) and Merck (Darmstadt, Germany) unless otherwise stated. Restriction enzymes, T4 DNA Ligase, ribonuclease A and DNA polymerase were used from Boehringer (Mannheim, Germany), MBI Fermentas (St. Leon–Rot, Germany), GibcoBRL (Eggenstein, Germany), Roche Diagnostics (Mannheim, Germany) and Promega (Madison, USA).
Radioactive chemical (α-[^{32}P]-dATP) was obtained from ICN Pharmaceuticals (Irvine, USA). Oligonucleotides were synthesized by MWG-Biotech (Ebersberg, Germany). A list of all the oligonucleotides used in this study is given in appendix 2. Antibodies were obtained from MoBiTec (Göttingen, Germany), Eurogentec (Seraing, Belgium) and Dunn (Germany).

2.1.2 Cultivation conditions of *Arabidopsis thaliana*

For all molecular biological studies, the model organism *Arabidopsis thaliana* ecotype Columbia was used. The genome of this organism is completely sequenced by the *Arabidopsis* Genome sequencing project (*Arabidopsis* Genome Initiative 2000). The soil used for the cultivation was a mixture of ED73 soil and vermiculite in the ratio of 3:1. The cultivation of *Arabidopsis thaliana* was done in a green house under short day and long day conditions. Under short day conditions, 8 hours of light at a temperature of 23 °C and 16 hours of darkness at 20 °C was maintained. The relative humidity was kept constant at 60 %. In order to induce flowering, plants were subjected to long day conditions (16 hours of light and 8 hours of darkness). The axenic cultivation of *Arabidopsis thaliana* was carried out on a modified MS salt medium (Murashige and Skoog, 1962) and to prepare solid medium 8 g/l of bacto-agar was added.

2.1.3 Microorganisms

The bacterial strain *Escherichia coli* DH5α was used for all the culture experiments. The relevant genetic markers on this strain are: F-O- endA1 hsdR17 (rk-mk+) supE44 thi-1 recA1

gyrA96 relA1 Δ(lacZYA-argF) deoR U169 (φ80d lacZΔM15)(Woodcock *et al*., 1989).

2.1.4 Bacterial growth media

LB medium (10 g/l Bacto tryptone, 5 g/l NaCl) was used for the growth of *Escherichia coli*, *Agrobacterium tumefaciens*, and for the selection of clones. To prepare solid medium 15 g/l of Bacto-agar was added. For the plasmid selection, the antibiotics ampicillin (100 mg/l) and kanamycin (50 mg/l) were used. LacZα recombinants and blue-white colonies were selected on plates containing 40 µl X-gal (50 mg/l). *Agrobacterium tumefaciens* GV3101 (Roberts *et al*., 2003) was cultered in the presence of antibiotics gentamycin (10 mg/l) and rifampicin (100 mg/l) were added. All components of the culture medium were purchased from DifcoLab (Detroit, USA).

2.1.5 Vectors

Vector	Features	Reference
pGEM®-3Zf(+)	AMP^r, MCS, LacZα M13-Primer binding site	Cloning vector / Promega (Madison, USA)
pCB302	Kan^r, MCS, T-DNA borders, *bar*	Binary vector / transformation using *Agrobacterium tumefaciens* Xiang *et al*. (1999)
pBI1.4tr-myc	Kan^r, MCS, T-DNA borders, SUVH2	Binary vector / transformation using *Agrobacterium tumefaciens* Nauman *et al*. (2005)

Table 2.1: AMP^r- Ampicillin resisistence, MCS - Multiple cloning site, LacZα - Alpha complementation for selection of clones containing the desired insert, M13 - Primer binding site - to prime DNA synthesis for sequencing reactions and Kan^r- Kanamycin resistence.

2.2 Methods

2.2.1. Isolation of plant total DNA

The extraction of genomic DNA from *Arabidopsis* was carried out according to the protocol of Brandstadter *et al*., (1994). Three to five leaves of *Arabidopsis thaliana* (approximately 1 g) were taken in a 2 ml eppendorf tube, frozen in liquid nitrogen, crushed with a precooled glass tube and then resuspended in 1 ml of extraction buffer (100 mM Tris-HCL, pH 8,5; 50 mM EDTA; 500 mM NaCl; 10 mM β-mercaptoethanol; 1.5 % SDS). The samples were

incubated at 65 °C for 10 minutes. Subsequently, 300 µl of potassium acetate (3 M potassium acetate; 2 M acetic acid) were added to each individual tube and the samples were incubated in ice for at least 10 minutes. Samples were then centrifuged for 15 minutes at 20,000 rpm and the supernatant was collected in a new eppendorf tube. In order to precipitate proteins, samples were subjected to 500 µl phenol-chloroform-isoamyl alcohol mixture (25:24:1) extraction. After vortexing the samples, they were centrifuged at 6,000 rpm to separate the upper aquous phase containing the DNA and the lower organic phase. The supernatant was transferred to a fresh tube and DNA was precipitated using 600 µl of isopropanol, after centrifuging the samples at 15,000 rpm for 10 minutes. The DNA pellet was washed with 70 % ethanol, dried and resuspended in 100 µl TE buffer (1 mM EDTA; 10 mM Tris-HCL, pH 8.0; 100 g/ml RNase A).

2.2.2 Plasmid DNA isolation from *E. coli* and *A. tumefaciens* cells

Plasmid DNA was isolated using Nucleospin Plasmid kit from Macherey-Nagel (ABgene, Germany). Bacterial cultures, obtained from a single colony were grown overnight in LB medium under shaking conditions. The cells were harvested by short centrifugation and were resuspended in 250 µl of solution of A1 (50 mM TrisHCl; 10 mM EDTA, pH 8.0; 100 µg/ml RNAse). Cell lysis was then achieved by mixing with 250 µl of solution A2 (200 mM NaOH; 1 % SDS). A3 buffer (3 M KOAc, pH 5.5) was added for neutralization. The supernatant obtained after centrifugation at 12,000 rpm for 10 minutes was loaded onto an anion exchange column and briefly centrifuged. DNA bound to silica and the flow through was discarded. The column was washed with A4 buffer, containing ethanol. Plasmid DNA was eluted from the column, by adding TE elution buffer followed by centrifuging at high speed (12,000 rpm) for 1 minute. The DNA concentration was measured spectrophotometrically.

2.2.3 Amplification of DNA fragments via polymerase chain reaction (PCR)

Polymerase Chain Reaction was performed using a standard protocol (Seiki, 1990) with Taq-polymerase (GibcoBRL, Germany) or Pfu-polymerase (Stratagene, Germany). A typical temperature profile used during the study was:

a) Denaturation 95 °C (30-45 sec)

b) Annealing 45 - 65 °C (30-45 sec)

c) Extension 70 - 72 °C (1 kb/min)

As PCR progresses, the generated DNA itself is used as a template for the next round of amplification, leading to an exponential increase in product concentration. Hence it is a very sensitive method and requires a very small amount of template. PCR can be extensively modified to perform a wide array of genetic manipulations. A typical reaction mixture with a total volume of 20 µl consists of 2 µl of 10x PCR buffer, 200 µM dNTP mix, forward primer 0.1-0.3 µM, reverse primer 0.1-0.3 µM and 1 U Taq polymerase. The PCR reaction was carried out in a Biometra PCR machine (T-GRADIENT).

2.2.4 Transformation of *E.coli*

Aliquots of 50 µl of the competent *E.Coli* strain *DH5α* (Invitrogen, Karlsruhe, Germany) were thawn on ice. 5 µl of the ligated plasmid was used for transformation in *E.coli* cells. After incubating the bacterial cells with plasmid on ice for 20 minutes, heat shock treatment at 42 °C for 15 sec in a water bath was done. This allows the adhering plasmids to pass through the cell wall. 250 µl of LB medium were added and the cells were incubated at 37 ^{0}C on a shaker for 30-40 minutes. The mixture was then plated on a solid LB agar plate containing appropriate antibiotics to selectively amplify plasmids containing the DNA of interest carrying a resistance marker. The plates were incubated overnight at 37 °C and the obtained colonies were checked for the presence of an insert using PCR.

2.2.5 Agarose gel elctrophoresis

The electrophoretic separation of DNA was carried out in horizontal 1 % agarose gels at a voltage of 1 to 6 V/cm. The agarose gels were made in TAE buffer (40 mM Tris-acetate, pH 7.8; 2 mM EDTA) and agarose and contained 0.01 % ethidium bromide. Prior to loading the samples on the agarose gel, 0.2 volumes of stop buffer (20 mM Tris-HCL, pH 8.0; 120 mM EDTA; 50 % glycerol; 0.1 % (w/v) bromophenol blue) were added. As a size marker 1kb DNA ladder was used for sizing and quantifing of the samples.

2.2.6 Purification, cloning and sequencing of DNA fragments

The desired DNA fragments were amplified by PCR and cloned directly using TA cloning technology. Since many enzymes including Taq DNA polymerase have terminal transferase activity and add a single, 3'-A overhang to each end of the PCR product, the PCR product can

be cloned directly into a linearized cloning vector that has a single 3'-T overhang on each end. Such vectors are called T-vectors. pGEM®-3Zf(+)-Vector (Promega, Germany) was used in all the subcloning experiments.

The PCR product with "A" overhang was mixed with this vector in high proportion. The complementary overhangs of a "T" vector and the PCR product hybridized resulting in ligated plasmid. The ligation was carried out with an insert to vector ratio of 3:1 in the presence of 1 U DNA ligase in a total volume of 10 µl and incubated overnight at 16 °C.

Sequence specific cleavage of double stranded DNA with restriction enzymes was caaried out by incubating the DNA with appropriate restriction enzymes (1 U up to 5 U) and buffer in a total volume of either 20 or 50 µl. Depending upon the enzyme, the reaction mixture was incubated for 2 to 12 hours and the DNA fragments were loaded directly on a 1 % agarose gel. For purification of DNA and removal of the restriction enzymes after digestion, the DNA band was excised from the gel and the fragment was purified using Gel Extraction Kit (Qiagen, Hilden, Germany). If proof reading activity was required, Pfu polymerase was used to amplify the DNA during the PCR reaction. In such cases, however, addition of an "A" overhang was necessary and was done by adding 5 pmol of dATP in the presence of ligase and appropriate buffer, before subjecting the DNA to TA cloning.

Sequencing was carried out according to the Sanger method (1977). For the sequencing PCR reaction, 5 µl of the purified plasmid DNA, which corresponds to 1-1.5 µg, was used as a template. In addition 0.1 µl (10 pmol) of the corresponding primer, 2 µl Big-Dye kit (Applied Biosystems, Darmstadt, Germany) and 1 µl of the corresponding buffer, in a total volume of 10 µl, were used as a reaction mixture in a T-Gradient (Biometra ®, Göttingen) or a T3 Thermocycler (Biometra®, Göttingen). After the PCR was completed, the DNA was precipitated using 1/10 volume of 3 M sodium acetate (pH 4.8) and 3 volumes of 96 % ethanol. The samples were centifuged at 13,000 rpm for 20 minutes. The obtained pellet was washed with 70% ethanol, dried and stored at -20^{0}C. Analysis of the sequencing data was done by ABI 3130x/Genetic Analyzer with the Sequence Analysis Software v5.2 (ABI Applied Biosystems, Darmstadt).

2.2.7 Preparation and transformation of *Agrobacterium tumefaciens*

Competent cells of *A.tumefaciens*, were prepared by inoculating 2-4 ml of overnight culture of *Agrobacterium tumefaciens* GV3101 in 100 ml of LB medium containing 100 µg/ml of rifampicin and 10 µg/ml gentamycin. This bacterial strain harbours a helper Ti plasmid

derived from pTiC58 providing gentamycin resistance and a C58C1 bacterial chromosome carrying rifampicin resistance gene. The culture was further grown for 8-10 hours in a shaker (160 rpm) at 28 °C. The cells were sedimented by centrifuging at 3,000 g for 10 minutes under cold conditions. The obtained pellet was resuspended in 1 ml of fresh LB medium and 200 µl aliquots were frozen in liquid nitrogen and stored at -70 °C.

An aliquot of competent cells was thawed on ice for transformation (Hendrik and Willmitzer, 1998) and 1-5 µg of plasmid DNA were added. They were further incubated on ice, in liquid nitrogen and at 37 °C as a heat shock treatment. Each of these treatments was carried out for five minutes. After addition of 1 ml of LB the medium, the cells were grown for 2-4 hours under shaking conditions (160 rpm) at 28 °C.

For the selection of transformants, the culture was plated on LB agar containing 100 µg/ml of kanamycin, 100 µg/ml of rifampicin and 10 µg/ml gentamycin. The plates were incubated for 2 days at 30 °C in dark.

2.2.8 Transformation of *Arabidopsis thaliana*

The *Agrobacterium* mediated transformation of *Arabidopsis thaliana* was done according to the "flower immersion" method commonly referred to as "floral dip" method (Clough and Bent, 1998). *Agrobacterium* was grown overnight in 50 ml of LB medium containing 100 µg/ml of kanamycin, 100 µg/ml of rifampicin and 10 µg/ml gentamycin. This was used as an inoculum for 500 ml LB medium containing the same concentration of the antibiotics, grown again for another 24 hours at 28 °C. Subsequently, the cultures were centrifuged (20 min, 4,000 g) and the bacterial pellet was dissolved in 100 ml of 5 % sucrose solution containing 0.05 % Silwet L-77. The inflorescence and the leaf rosette of very early flowering plants were dipped in this solution. The plants were then kept in the dark for 24 hours and later transferred to long day conditions.

2.2.9 Crossing of *Arabidopsis thaliana* (flower emasculation and preparation for fertilization)

In the first step, the recipient, designated as female flowering plant, on which crossing has to be performed, was selected. All the flower parts were removed without disturbing the ovary. Nearby flowers, which were either too young or too old, were removed to prevent contamination or self-fertilization. Anthers from the male recipient was taken out and brushed on the prepared ovary of the emasculated plant. This process was repeated on the next day to

increase the success rate of the crossings. A successful crossing is indicated by elongation of the silique within one or two days.

2.2.10 Southern hybridization

The DNA samples (1-3 µg) were digested overnight with an appropriate restriction enzyme. (enzymes that cut only once in pCB 302 vector were chosen for this experiment.) The digested samples were precipitated using 1/10 volume of 3 M NaCl and 3 volumes of 96 % ethanol. The samples were centrifuged at 13,000 rpm for 10 minutes. The obtained pellet was washed with 70 % ethanol, dried and stop buffer was added. These samples were loaded on a 1 % agarose gel, which was run overnight for gradual electrophoretic separation of the DNA fragments.

The DNA on the gel was denatured by placing the gel in a denaturing solution (1 M NaOH and 1.5 M NaCl) for 30 minutes. Later pH was adjusted using neutralization buffer (1 M Tris-HCl, pH 7.4; 1.5 M NaCl) for 60 minutes. DNA from the gel was transferred to an uncharged nylon membrane using 20x SSC (0.3 M sodium citrate buffer, pH 7.4; 3 M NaCl) in a capillary blot fashion. The membrane containing the transferred DNA was UV-crosslinked, in order to covalently attach the DNA to the membrane.

The probe for hybridization was made using Mega Prime DNA-labeling kit (Amersham Pharmacia, Germany) with 50 µCi α-[^{32}P dATP as advised by the manufacturer. Non-incorporated nucleotides were subsequently removed from the membrane-bound DNA by Gel filtration chromatography using ProbeQuant® G-50 Micro Columns (Amersham Pharmacia, Geramany). The amount of incorporated radioactivity was checked using QC4000 XER (Bioscan, USA). The hybridization was carried out using ExpressHyb®-buffer (Clontech, USA) as instructed by the manufacturer. The analysis of the autoradiography was carried out using the Phosphoimager (Molecular Dynamics, Krefeld).

2.2.11 Inverse PCR

Inverse PCR (I-PCR) was used to locate the inserted mutater T-DNA in the genome. As compared to a standard PCR reaction where the two designed primers are directed towards each other leading to amplification of a template, the primers in I-PCR are facing in opposite direction. The primers for this experiment were designed to bind the ends of T-DNA borders. Genomic DNA was prepared from the leaves of *Arabidopsis* carrying the T-DNA. DNA was

dissolved in 100 µl of 10mM Tris pH 8 and was digested by different restriction enzymes (MBI Fermentas, Germany), in presence of respective buffers which were four base pairs (bp) cutters, at 37 °C overnight in a separate reaction. The digested DNA was ligated using T4 DNA ligase (MBI Fermentas, Germany) at 16 °C for 24 hours. The ligated DNA was used as a template (Mathur *et al*., 1998). The fragments containing the junction point between the T-DNA and the chromosome, i.e. the insertion point, would circularize and the primer pairs would then point towards each other. A standard PCR reaction was then used to allow amplification of the specific junction fragments. Products from this PCR were subjected to nested PCR using primers binding downstream of the first set of primers. The fragments were eluted from agarose gel and either subcloned into pGEM-T vector for sequencing or sequenced directly using the forward primer used in the second PCR. Subsequently the adjacent genomic sequences of the DNA were compared with the whole genome of Arabidopsis thaliana in the BLAST (Altschul *et al*., 1997) program of TAIR (http://www.arabidopsis.org/) to identify the insertion position.

2.2.12 Spectrophotometric determination of DNA or RNA concentration

The purity and concentration of DNA or RNA concentration and purity was measured spectrophotometrically at 260-280 nm. Two µl DNA/RNA sample was dissolved in 98 µl sterile water. The ratio between the values obtained at 260 and 280 nm allows an estimation of nucleic acid purity. Pure DNA or RNA preparations have OD_{260}/OD_{280} ratio of 1.8 or 2.0 respectively (Sambrook *et al*., 1989).

2.2.13 Isolation of RNA

Total RNA from leaf material of *Arabidopsis thaliana*, was isolated by the Trizol®- method (GibcoBRL, Germany). Three-five leaves were collected in 2 ml eppendorf tubes, frozen in liquid nitrogen and finely crushed using a precooled glass rod. One ml of Trizol® was added to each tube and vortexed thoroughly. After 5 minutes of incubation at room temperature 240 µl of chloroform were added, mixed and incubated again at room temperature for 5 minutes. In order to separate the aqueous and the organic phase, samples were centrifuged at 6,000 rpm for 10 minutes. The upper aqueous phase containing the RNA was transferred into a fresh eppendorf tube and RNA was precipitated with 1 volume of isopropanol and centrifuged at 15,000 rpm for 10 minutes. The obtained pellet was washed with 70 % ethanol,

dried and dissolved in 20 µl of DEPC water.

2.2.14 RT-PCR analysis

The semi-quantitative RT-PCR experiments were performed with the RT-PCR Kit (Promega, Mannheim) according to the manufacturers specifications. Two µg of RNA was used as starting material for cDNA synthesis by incubating it with10 mM dNTP, 10 pmol random hexamers and 200 U M-MLV reverse transcriptase (M-MLV RT) for 2 hours at 37 °C in a total volume of 25 µl. One µl of the RT reaction was used as a template for the subsequent PCR. The PCR conditions were adjusted to the primers and the size of the amplicon. 18SrDNA was used as a control.

2.2.15 Quantitative real-time PCR

Reverse transcriptase PCR is mainly semi quantitative due to the insensitivity of ethidium bromide and therefore, real-time PCR was used to achieve more reliable information about the total mRNA levels. The real-time PCR was performed in a 96-well plate with the icycler real-Time PCR Optical Detection System controlled by iCycler™ iQ Optical System Software Version 3.0a (Bio-Rad, Germany). A 20 µl reaction was set up, containing cDNA 1x Platinum®SYBR®Green qPCR SuperMix-UDG (Invitrogen, Germany) 250 pmoles of each gene-specific primer (Ay *et al*., 2009). SYBR®green is a fluoroscent dye that binds to double but not to single stranded DNA. The amount of fluorescence emitted, will therefore indicate the amount of double stranded DNA formed in the 96-well plate. By determining the melting point of the amplicon and obtaining a sharp peak in the melting curve the fluorescence from unwanted sources was decreased. To calculate PCR efficiencies, three different cDNA dilutions were applied. To determine and compare the relative gene expression rate of the overexpression of SUVH2 in isolated insertional mutater background relative to the expression in controls, the formula described by Pfaffl, 2001 was used. 18S rRNA was used as a reference gene. Subsequent to the normal quantitative PCR, the determination of a melting curve of the amplified PCR products was carried out. Controls without reverse transcriptase and without template were performed to exclude the amplification of unspecific products.

2.2.16 Immunocytological analyses of H3K9me2

Fresh leaves from *Arabidopsis* ecotype Columbia were used for immunocytological studies. Leaf pieces were fixed in 4 % formaldehyde in PBS on a glass slide, covered with cover slips and squashed. Slides were dipped in liquid nitrogen and were transferred into PBS immediately after removing the cover slip. The slides were pre-incubated at 37 °C in 2 % BSA for 60 minutes and later incubated overnight with H3K9 α-dimethyl antibody (ab 1220, Abcam, USA) prepared in 2% BSA in 1:100 fold dilution. Detection of antibodies was performed by Alexa-488, which, was conjugated to the secondary antibody (Molecular probes, Invitrogen, Germany).
All preparations were counterstained with 4',6-diamidino-2-phenylindole (DAPI, 2 µg/ml) and imaged with a Zeiss Axioskop microscope (Carl Zeiss Inc,) equipped with an epifluorescence attachment.

2.2.17 Mutagenesis of *Arabidopsis thaliana*

A T-DNA mutagenesis approach was used to obtain dominant suppressors of the SUVH2 overexpression phenotype. Since homozygous lines of SUVH2 were almost sterile in the first screen, the heterozygous overexpression line was used. In the transition from the vegetative phase to the inflorescence state, the SUVH2 overexpression line #6 plants were transformed with *Agrobacterium tumefaciens* containing pCB302 resulting in random integration of the T-DNA in the genome of *Arabidopsis*. In the second screen, homozygous overexpression plants that were produced after post-transcriptional gene silencing were used for mutagenesis experiment.

2.2.18 Bioinformatics

For the database search of DNA and proteins blastn and blastp (www.tigrblast.org, www.ncbi.nlm.nih.gov/BLAST) were used. For the analysis of *Arabidopsis* genes, the TAIR database (The *Arabidopsis* Information Resource, www.arabidopsis.org) was used. Sequence analysis was done with the program BIOEDIT (Biological sequence alignment editor for WIN95/98/NT/2K/XP, www.mbio.ncsu.edu/Bioedit/bioedit.html). For the creation of image files, Adobe Photoshop was used. Primers designed with http://frodo.wi.mit.edu/primer3.

3. Results

3.1 Prominent role of SUVH2 in heterochromatin formation in *Arabidopsis thaliana*

SUVH2 was shown to have three important domains, which play a crucial role in heterochromatin formation (Figure 3.1). The N-terminal domain probably helps in substrate recognition. The SRA/YDG [(Tyrosine (Y), Aspartic acid (D), Glycine (G)] domain is responsible for binding to methylated DNA (Naumann *et al*., 2005). Recently it was shown that SUVH2 preferentially binds to methylated CG sites (Johnson *et al*., 2008). Hence SUVH2 could possibly interact with MET1. Through genetic interaction studies, sequence specific methylation of genomic DNA was shown to be dependent on DDM1 and MET1 (Naumann *et al*., 2005). DDM1 was also shown to bind proteins containing a methyl CpG binding domain protein and this interaction might play a role in their localization (Zemach *et al*., 2005).

.

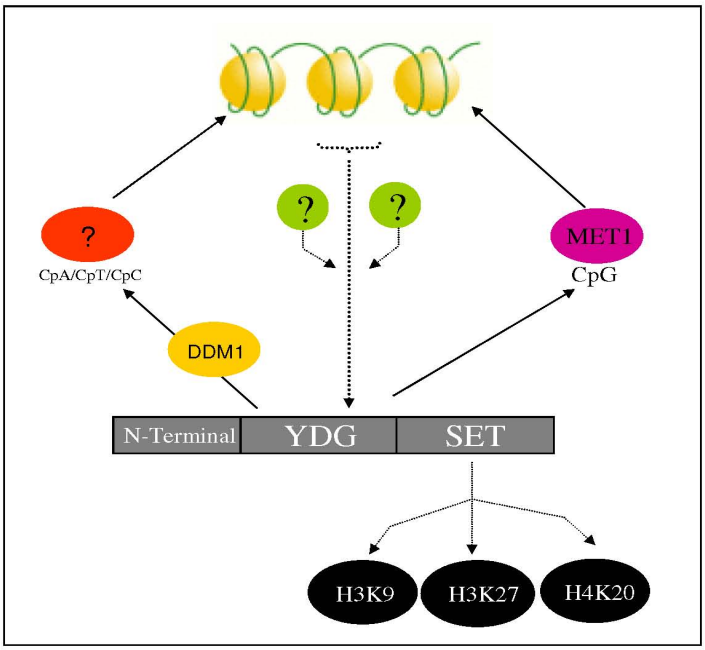


Figure 3.1: Central role of SUVH2 in heterochromatin formation in *Arabidopsis.* SUVH2 with its N-terminal domain and YDG domain can influence DNA methylation of the specific target sequence through interaction with the DNA methyltransferase genes, DDM1 and MET1. The SET domain is responsible for specific lysine methylation on the histones leading to repressed chromatin states.

The SET domain of SUVH2, located at the carboxyl terminus of the protein is responsible for its methyltranferase activity. When SUVH2 was overexpressed, there was a significant increase in all the heterochromatin marks i.e, H3K9me1 and me2, H4K20me1, and mono and

H3K27me1 and me2, compared to wild type. Accordingly all the heterochromatin marks were significantly reduced in the null mutant background thus emphasizeing the importance of SUVH2 in heterochromatin formation in *Arabidopsis thaliana.*

3.2 Overexpression of SUVH2 leads to a mini-plant phenotype

Out of 10 SUVH genes of *Arabidopsis thaliana*, which are homologous to the SU(VAR)3-9 protein, SUVH2 has been shown to have profound influence on heterochromatin formation. Overexpression of SUVH2 with a strong cauliflower mosaic virus (CaMV) 35S promoter led to severe developmental defects as shown in the Figure 3.2. In the heterozygous background of overexpression of SUVH2, the leaves were curly and the plants were smaller as compared to wild type plants [Figure 3.2 (b)]. The effect was much more pronounced in the homozygous background, giving rise to a mini-plant phenotype [Figure 3.2 (c)]. Plants remained quite small and were also sterile (Naumann *et al*., 2005).

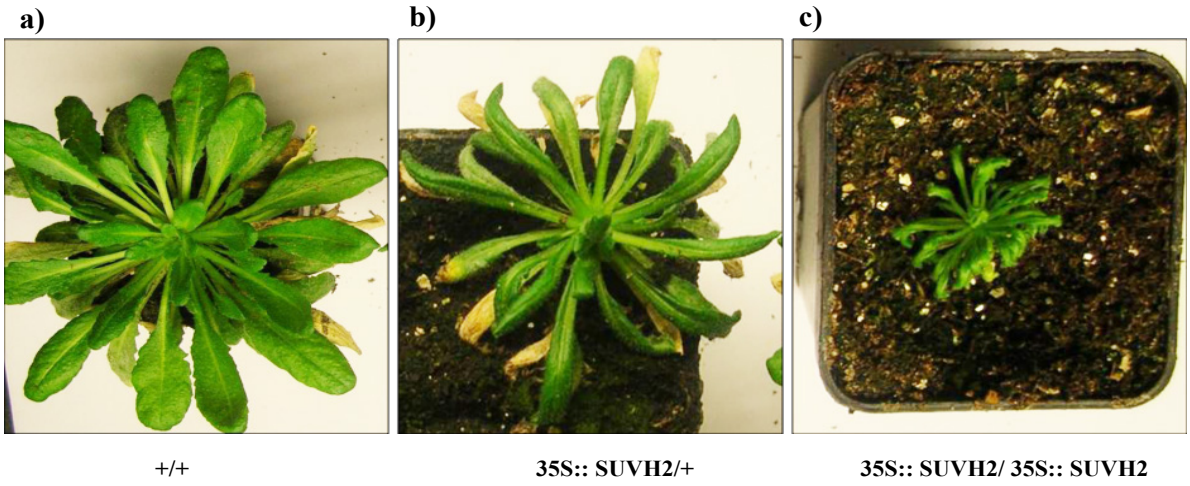


Figure 3.2: Overexpression of the histone methyltransferase SUVH2 protein in *Arabidopsis thaliana* leads to a severe growth defect. (a) In comparison to the wild type (+/+), (b) When SUVH2 is overexpressed with the strong 35S promoter, it results in a curly leaf phenotype in heterozygous (35S:: SUVH2/+) plants; (c) and a mini-plant phenotype in homozygous plants (35S::SUVH2/35S::SUVH2).

The effect of SUVH2 overexpression could be observed in early stages of development. In the overexpression background of SUVH2, the cotyledons were curled compared to that of wild type cotyledons (Figure 3.3), which was a useful criterion for identifying and selecting SUVH2 overexpressing plants during the screening for dominant suppressors.

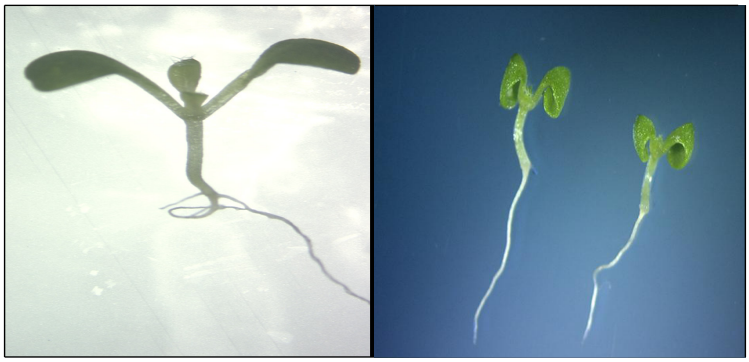


Figure 3.3: SUVH2 dependent phenotypic defects are visible in early stages of growth. Overexpression of SUVH2 has a dramatic effect on the phenotype of *Arabidopsis* at a very early stage of growth. As shown the first cotyledons of the overexpression plants are curled in comparison to the wild type.

3.3 Target and mutant constructs used for establishment of a screening system

For the overexpression construct of SUVH2 in *Arabidopsis thaliana* (ecotype Columbia), CaMV 35S promoter was used. This vector also had a myc tag and a Npt II cassette. Myc is a polypeptide tag derived from the c-myc gene product, and can be very useful in protein localization experiments as antibodies against this tag are commercially available. The Npt II cassette serves as a selection marker and provides resistance against the antibiotic kanamycin [Figure 3.4 (a)]. Transgenic plants containing SUVH2 overexpression constructs were therefore selected based on kanamycin resistance on sugar free MS medium containing 50 mg/lL of the antibiotic (Naumann *et al*., 2005). This T-DNA was designated as target construct as the components involved in heterochromatin formation were intended to be identified based upon the T-DNA containing SUVH2 overexpression construct.

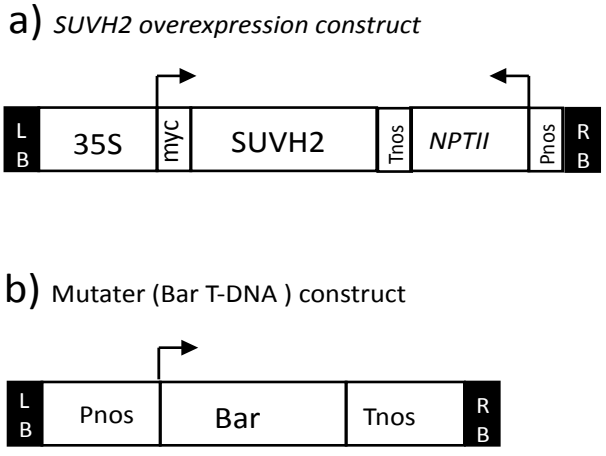


Figure 3.4: Structure of the target (SUVH2 overexpression) and mutater (*bar* gene) T-DNA constructs. a) The transgenic *SUVH2* construct is driven by a strong constitutive CaMV 35S promoter. The NPT II gene is under the control of a weak nos promoter and acts as a selection marker for the identification of transgenic plants. b) The *bar* gene in the mutater T-DNA is under the control of a weak Nopaline synthase (nos) promoter. The *bar* gene provides resistance against the herbicide Basta and thus acts as a selection marker for the identification of transgenic plants. The arrows indicate the respective direction of transcription with in the expression cassette**.** LB and RB represents left and right border of the construct.

For insertional mutagenesis in SUVH2 overexpression plants, another T-DNA of the mini binary vector pCB302 was used (Xiang *et al*., 1999).This T-DNA was referred as "mutater T-DNA". This T-DNA was successfully used, for the identification and isolation of insertional dominant suppressors of the SUVH2 overexpression associated mini plant phenotype. The mutater T-DNA not only disturbs the expression of the gene into which it is inserted but also acts as a marker for subsequent identification of the mutation. For the selection of the transgenic plant, a *bar* gene expression cassette was used [Figure 3.4 (b)]. A weak NOS promoter controls the expression of this gene. The *bar* gene codes for phosphinothricin acetyltransferase (PAT) providing resistance against the herbicide phosphinothricin, commonly referred to as Basta. For the selection of transgenic plants, seedlings were grown on soil till cotyledon stage and then sprayed with 1:5000 dilution solution of the herbicide Basta ® (Hoechst Schering AgrEvo GmbH). The treatment was repeated at intervals of 3-5 days until a sufficient selection of transgenic plants was evident. The plants were maintained at 23 ^{0}C under short-day conditions.

3.4 Principle of T-DNA mutagenesis to isolate dominant suppressor mutants of the SUVH2 overexpression dependent mini plant phenotype

In the process of heterochromatin formation, SUVH2 might interact with another gene here referred to as gene X in a complex pool of proteins, giving rise to curly/mini plant phenotype. (Figure 3.5). In such a background when another T-DNA is transformed, it randomly gets incorporated into the genome. If the introduced T-DNA disturbs the function of the gene X, which interacts with SUVH2, this interaction is lost, due to the insertion of mutater T-DNA leading to a wild type phenotype.

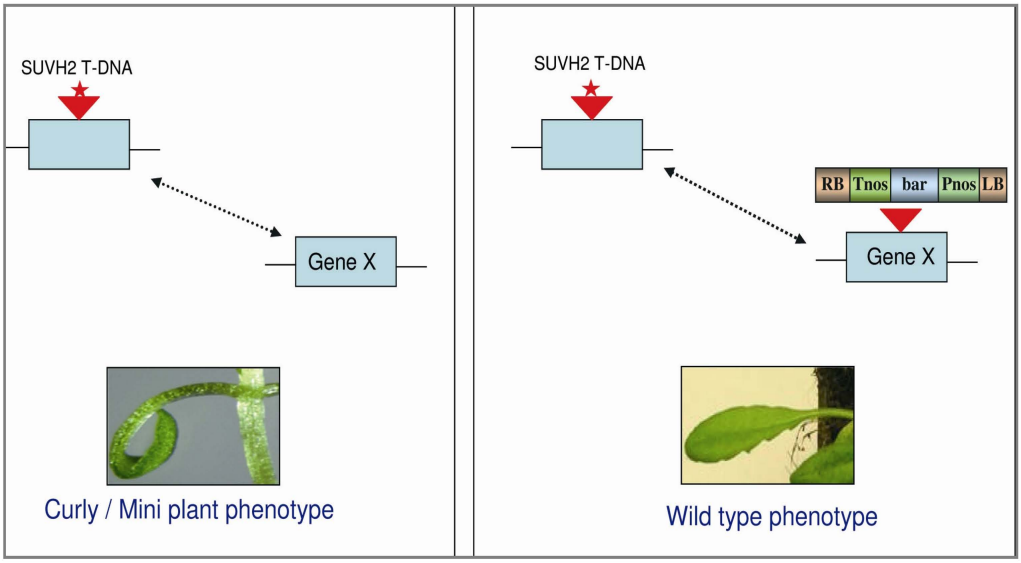


Figure 3.5: Principle of a identification of dominant suppressor. The overexpression of SUVH2 results in ectopic heterochromatization and curly leaf phenotype. When one of the components (encoding gene X) involved in heterochromatin formation is disturbed by insertion of T-DNA, it might result in less heterochromatin formation and accordingly the overexpression phenotype reverts back to wild type.

3.5 Screening of dominant suppressors involved in chromatin regulation

3.5.1 Mutant screen in heterozygous SUVH2 overexpression plants

In the first screen, 35S::mycSUVH2/+ heterozygous plants were used to transform pCB302 plasmids using an *Agrobacterium tumefaciens* mediated transformation. The floral dip method was used and a transformation efficiency of 0.1-1% was achieved. All the transformed plants were transferred to long day conditions. In the next generation, seedlings were grown on soil, and the cotyledons were screened for the phosphinothricin resistance, by challenging them with Basta spray [Figure 3.6 (a)]. Spraying was repeated after 5-6 days in order to confirm the presence of mutater (*bar* T-DNA) construct expressing the *bar* gene and providing resistance.

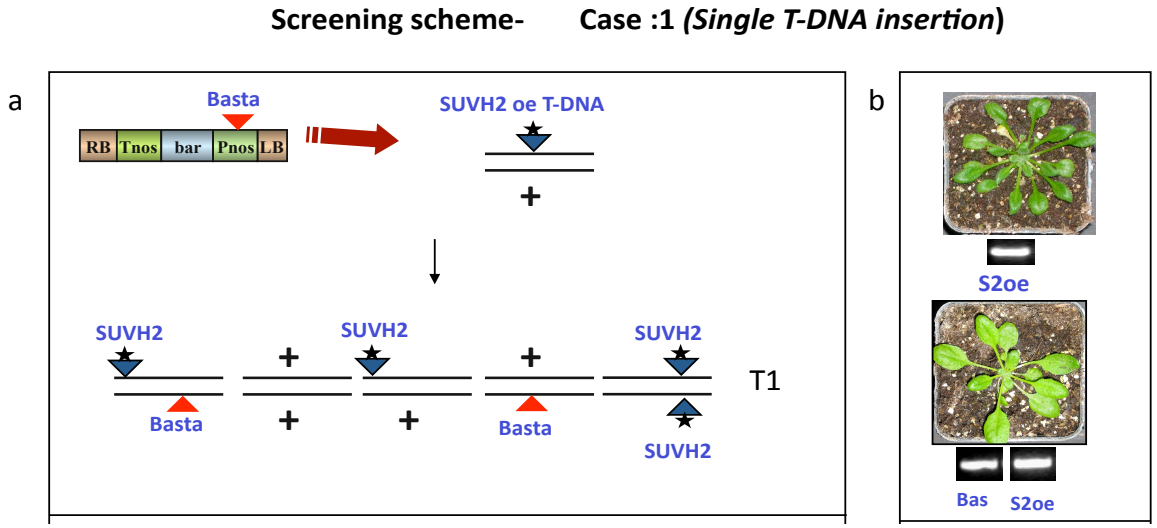


Figure 3.6: Transformation of SUVH2 overexpression plants with a Basta T-DNA construct. a) The pCB302 vector containing the *bar* gene was used as an insertion mutagen in SUVH2 heterozygous plants (abbreviated as S2oe). The Nopaline synthase (nos) promoter was used to drive the expression of the *bar* gene in this construct, conferring phosphinothricin resistance and therefore allowing selection of transgenic plants containing a Basta T-DNA construct. b) The heterozygous SUVH2 overexpression line # 6 was used for the transformation process. In T1 generation plants were selected for the dominant suppressor effect by scoring for wild type phenotype.

In a first screen heterozygous plants were used for transformation as homozygous over expression plants were sterile [upper panel of Figure 3.6 (a) and (b)]. Due to segregation of progeny in the next generation, plants with different genotypes were obtained as shown [lower panel of Figure 3.6 (a)]. Suppressed plants containing both T-DNAs, were subjected to further analysis [lower panel of Figure 3.6 (b)].

In the T1 generation, the plants were screened for dominant suppressors of the SUVH2 over expression phenotype. A loss of phenotype would possibly indicate that a potential suppressor gene, which could interact, with SUVH2 in heterochromatin formation has been disturbed by insertion of the *bar* T-DNA

3.5.2 Analysis of the presence of the SUVH2 transgene in the T1 generation

Due to segregation of the SUVH2 transgene, in the T1 generation, progeny with different genotypes were obtained and all the suppressed plants (no curly leaf phenotype) were checked by PCR analysis to differentiate between dominant suppressors and wild type plants [Figure 3.7 (a) and (b)]. The transgene specific primers, mycSAL and SUVH2574R were used for the detection of the SUVH2 overexpression construct. This was necessary to avoid amplification of sequences from the endogenous SUVH2 locus present in all plants subjected to PCR analysis.

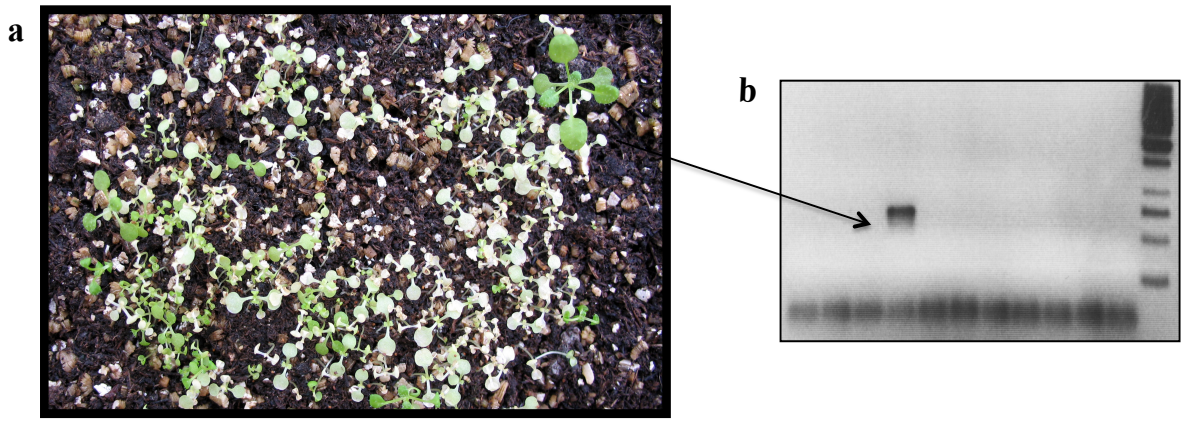


Figure 3.7: PCR analysis of the suppressed Basta resistant plants to confirm the presence of the SUVH2 transgene. In T1 generation, plants with different genotypes containing mutater T-DNA were obtained after Basta treatment (35S::SUVH2 /35S::SUVH2, 35S::SUVH2/+, +/+) [Figure 3.7 (a)]. Curly resistent plants were not processed further and suppressed plants with no curly leaf phenotype were subjected to PCR analysis in order to check for the presence of the SUVH2 transgene [Figure 3.7 (b)].

3.5.3 Selection of suitable independent lines containing single copies of inserted mutater T-DNA

Suppressed SUVH2 transgenic positive plants (T1) were self-pollinated to obtain plants in T2 generation [upper panel of Figure 3.8)]. The pattern of segregation of the Basta resistance selectable marker was followed in succeeding generations. The generations of T-DNA-transformed plants were named T0 (*Agrobacterium*-treated plants), T1 (First generation of Basta-resistant plants), and T2 (first segregating generation). The independent lines, which showed a Mendelian segregation pattern of 3:1 (resistant and loss of phenotype: sensitive and curly), indicated the presence of a single insertion of the mutater T-DNA in T2 generation. All the resistant progeny were completely supressed and the sensitive plants, which died after spray, had a SUVH2 overexpression phenotype [lower panel of Figure 3.8].

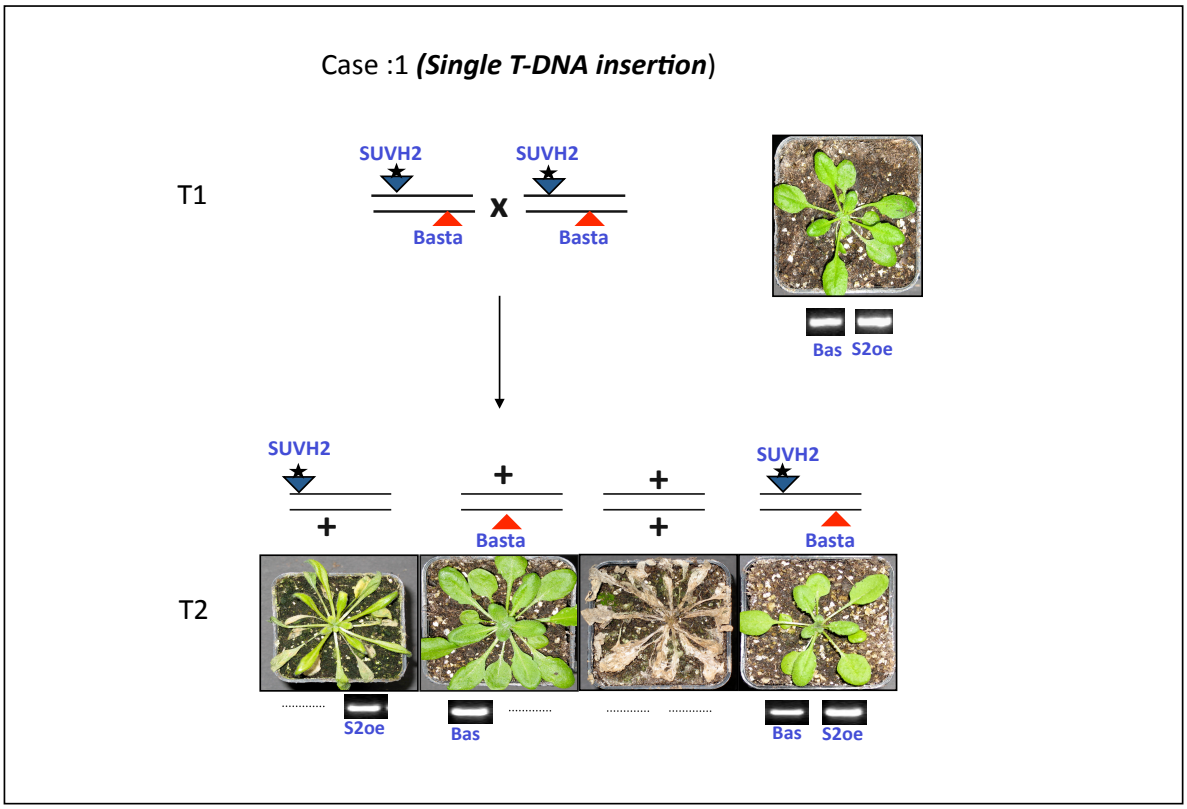


Figure 3.8: Identification of single inserts by segregation of Basta resistance. Progenies of suppressed plants obtained in T1 generation were selfed and the segregation ratio was analyzed in the following generation. A segregation ratio of 3:1(resistant : sensitive) with respect to Basta treatment indicated the presence of single locus, which was later confirmed by Southern blot analysis.

A total of 1956 Basta resistant plants were screened, out of which 692 (T1) plants had wild type phenotype and the remaining (1294) plants had curly phenotype and were not processed further (Table 3.1). From all the 692 transformants, genomic DNA was isolated and was checked for the presence of the SUVH2 transgene by PCR analysis. One hundred and twenty three transformant out of 692 contained the SUVH2 transgene. Fifty-seven independent lines of T1 generation were self pollinated to obtain plants in the next generation and were screened by segregation analysis for the presence of single insert. Five independent lines were chosen for further analysis.

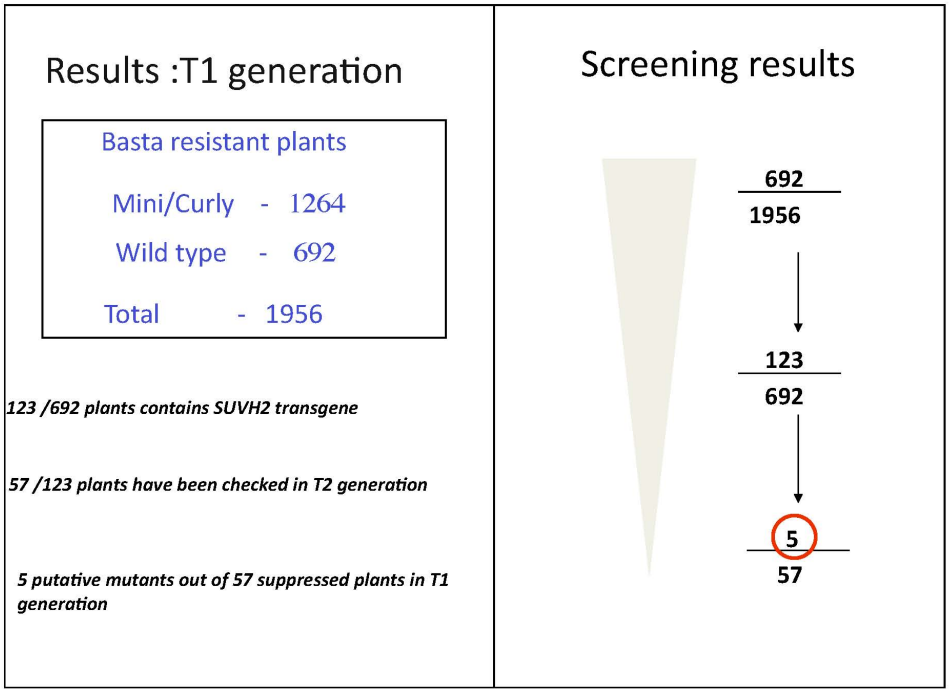


Table 3.1: Summarized data from the 1st screen indicating the mutation frequency obtained from the progeny of Basta resistant transformants. A large number of Basta resistant plants were analyzed. Progeny containing the SUVH2 transgene from the T1 generation were further analyzed by self-pollinating and scoring for suppressor effect in next generation. Five putative suppressors were identified out of 57 SUVH2 transgene containing plants and corresponding genes were identified.

3.5.4 Identification of multiple inserts of mutater T-DNA by a complex segregation pattern

During *Agrobacterium* mediated transformation, the T-DNA could undergo complex rearrangements and may be integrated in multiple copies in the genome. The high segregation ratio after Basta spray could possibly indicate the presence of multiple inserts. This is also evident by the phenotypic appearance of the plants in the T2 generation. As shown [lower panel of Figure 3.9 (b)] apart from appearance of plants which were completely suppressed, multiple inserts could lead to the appearance of curly and Basta resistant plants in heterozygous/ homozygous state. Presence of such siblings indicates the integration of multiple independent T-DNAs in the genome, which did not tag the regions involved in the SUVH2 mediated gene-silencing pathway.

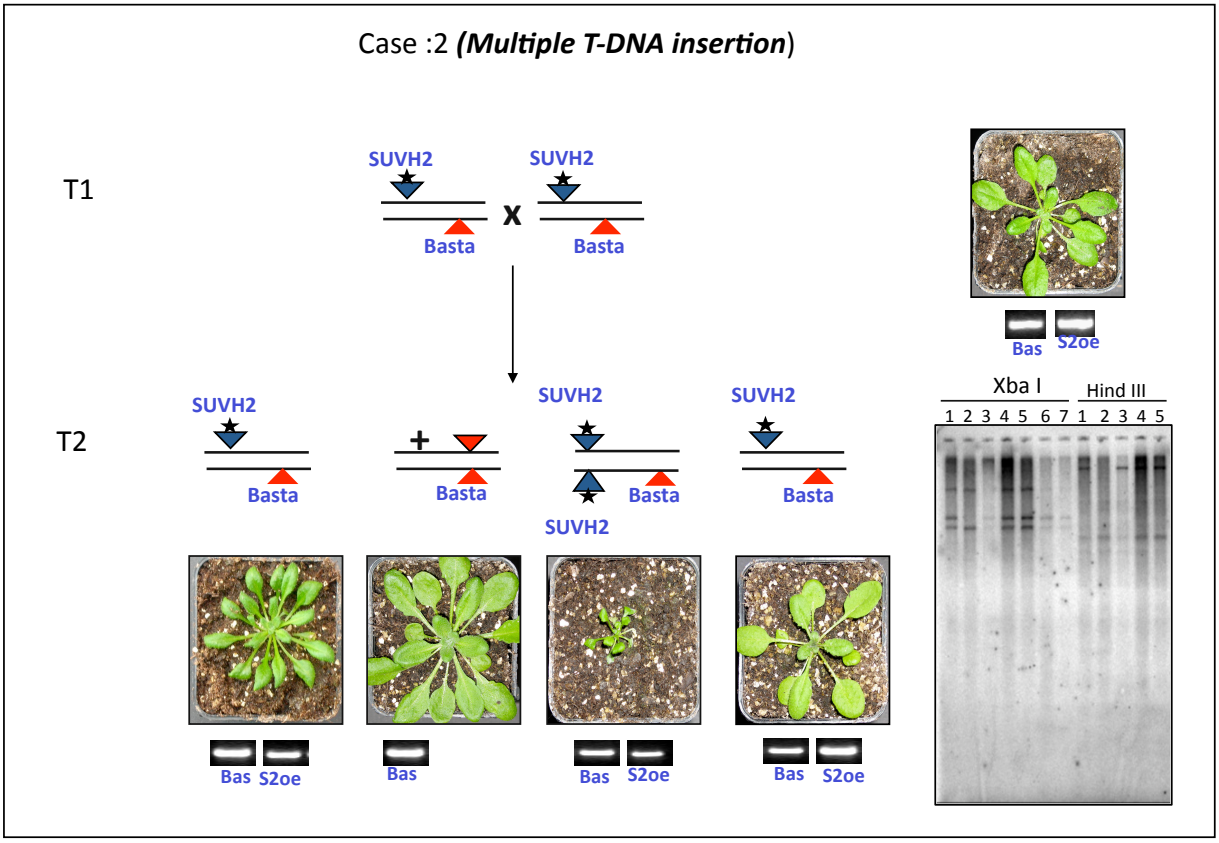


Figure 3.9: Identification of complex and multiple integrations, by phenotypic analysis and segregation of Basta resistance. The obtained progeny of the suppressed plants in T1 generation were self-pollinated and the segregation ratio was analyzed in successive generations. A segregation ratio of 15:1 (resistant: sensitive) indicated the presence of two independent loci. Multiple and complex integrations of mutater T-DNA resulted in different irregular segregation patterns and generally identified by high segregation ratio of resistance: sensitive plants, resulting in appearance of curly resistant progeny in T2 generation.

Therefore based on phenotypic appearance of curly resistant plants, it is possible to differentiate between locus containing single or multiple insert. Only putative single inserts were studied further.

3.5.5 Mutant screen with homozygous SUVH2 overexpression plants using PTGS of the SUVH2 transgene

When the transgene is overexpressed in the early stage of development it might cross a threshold limit of expression leading to specific degradation of the product resulting in posttranscriptional gene silencing (Schubert *et al.*, 2004; Vaucheret *et al.*, 2001). This mechanism was much more pronounced in homozygous SUVH2 overexpression plants [Figure 3.10 (a)] especially under drought conditions [Figure 3.10 (b)]. A new shoot typically appears from the hypocotyl region as shown in Figure 3.10 (c). The post transcriptionally

silenced plant has the same genotype as the parental line but is fertile and therefore can be transformed with *Agrobacterium* containing the mutater DNA. PTGS is lost in the next generation, and all the plants restore the SUVH2 overexpression phenotype [Figure 3.10 (d)]. However in progeny where the mutater T-DNA had disturbed the function of a gene, encoding a functional component of heterochromatin formation, and which could possibly interact with SUVH2, suppressed plants were obtained in T1 generation [Figure 3.10 (e and f)].

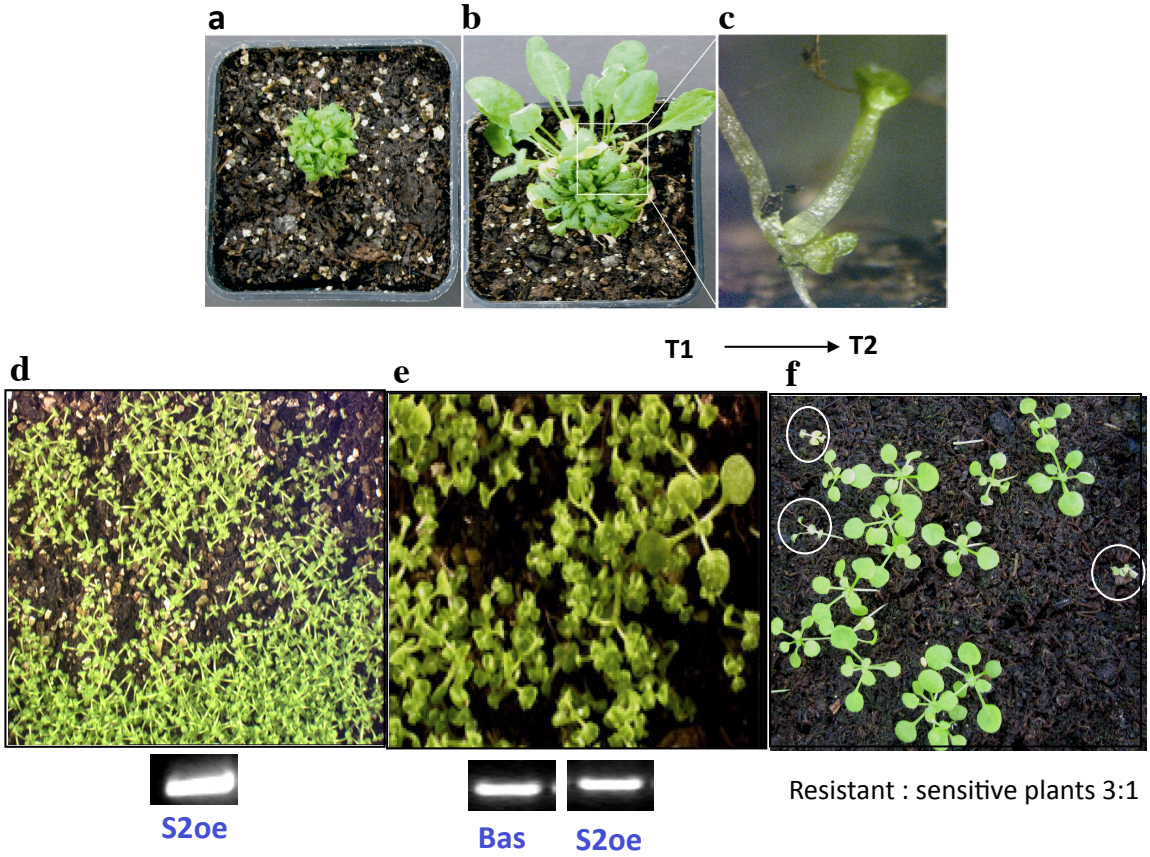


Figure 3.10: The influence of PTGS on fertility of homozygous SUVH2 overexpression plants. (a) Over expressed SUVH2 plants undergoes Post Transcriptional Gene Silencing (PTGS) especially under water deficit stress conditions. (b) This results in appearance of another shoot as shown in the Figure. (c) The shoots, which have undergone PTGS, are fertile, and therefore could be transformed with mutater T-DNA. (d) In the next generation PTGS is lost again resulting in appearance of only curly plants characteristic of homozygous overexpression line. (e and f) However plants, which have undergone PTGS could be transformed with Basta T-DNA construct and screened for resistant suppressed plants, containing single Basta T-DNA insert.

In comparison, all control plants that had not undergone transformation showed the SUVH2 overexpression phenotype. A PCR analysis was performed in order to confirm the presence of SUVH2 transgene. The PCR positive Basta-resistant plants were self-pollinated and the segregation ratio of the progenies were checked in succeeding generation. The segregation

ratio of 3:1 indicates the insertion of a single copy of mutater T-DNA in a particular locus of the genome.

Basta resistant plants	994
Putative mutants	**176** (T1 generation)
Single T-DNA insert	10 (OUT OF 85)
Remaining plants	91 (in progress)

Table 3.2: Summary of the data from 2nd screen indicating the mutation frequency obtained from the progenies of Basta resistant transformant. A suitable population size of the transformant was analyzed. 10 progenies out of 85 suppressed plants appeared to have single inserts as indicated by segregation ratio and were further characterized by Southern blot analysis.

A total of 994 Basta resistant plants were screened, out of which 176 (T1) plants had wild type phenotype and the remaining (818) plants had curly phenotype and were not processed further (Table 3.2). Only a fraction of 85 transformant out of 176 suppressed plants, were analyzed in next generation. In 10 independent lines a segregation ratio of 3:1 was obtained. They were further confirmed by Southern blot analysis for the presence of single inserts.

3.6 DNA methylation status of the SUVH2 transgene using methylation sensitive Southern blot analysis

Occurrence of more then one T-DNA in the transformant might lead to repeat induced gene silencing because of homology between border sequences of the T-DNAs. In addition multiple inserts of mutater T-DNA could also lead to the silencing of the SUVH2 transgene. Repeat induced gene silencing is characterized by DNA hypermethylation of the transgene. Hence in order to check whether the SUVH2 transgenes are methylated and could be in a silenced state, a Southern blot analysis with methylation sensitive enzyme was carried out.

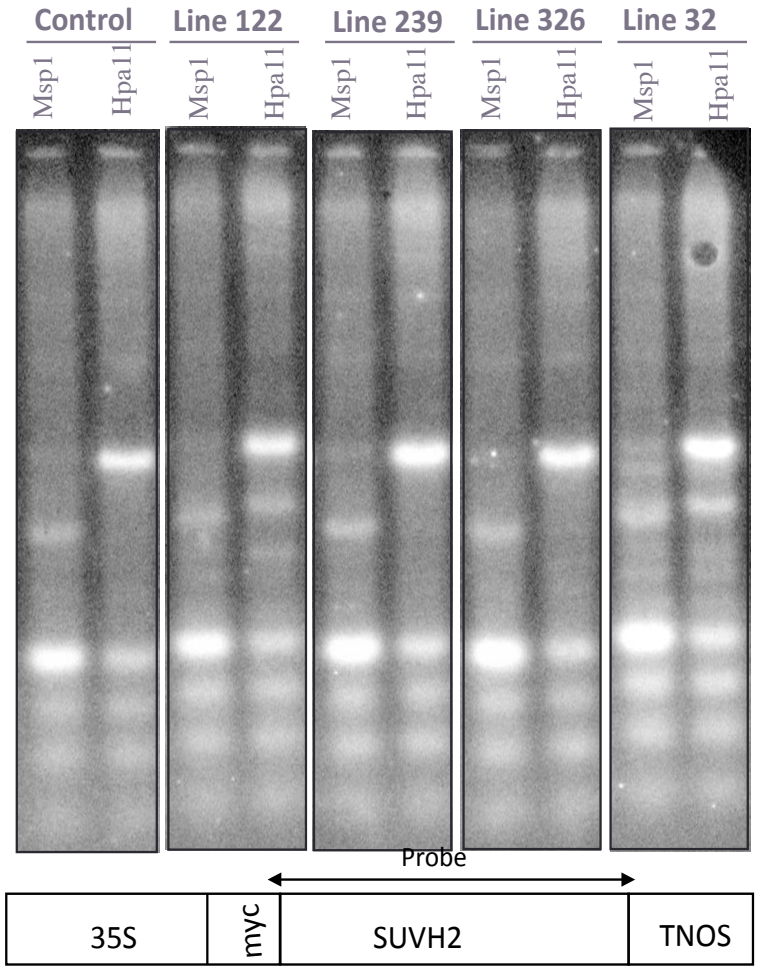


Figure 3.11: Southern blot analysis indicates the absence of methylation in the SUVH2 transgene. Four putative mutants were subjected to DNA blot analysis along with the control. The total DNA was digested with the methylation-sensitive restriction enzymes HpaII (CCGG) and its isoschizomer MspI. After transfer, the blots were probed with ^{32}P-labeled SUVH2 transgene.

Genomic DNA from the putative suppressor mutations was isolated and digested with the methyl-sensitive endonucleases Msp I and Hpa II. The mutant lines 122 and 239 were single inserts, whereas lines 32 and 326 contained multiple copies of the mutater TDNA. Although the enzymes are isoschizomers and therefore have a common recognition sequence (CCGG), they differ in their ability to cleave at 5-methyl-cytosine. When the external C in the CCGG-sequence is methylated, none of the enzymes (Msp I and Hpa II) can cleave the sequence. However, MspI and HpaII differ in their ability to cleave at the second C residue in the recognition site. Hpa II is sensitive to methylation and does not cleave while Msp I can cleave the sequence when the internal C is methylated.

When an unmethyled DNA is cleaved with these two enzymes, due to large number of restriction sites, relatively low molecular weight fragments are observed as compared to methylated DNA. This provides a good hint of the DNA methylation status of the target

locus. When SUVH2 transgene was subjected to restriction by Msp I and Hpa II, as compared to control no higher molecular weight fragments were seen in lines with single or multiple inserts compared to control, indicating that the transgene was not methylated and therefore not silenced (Figure 3.11).

3.7 Real-time analysis of SUVH2 overexpression in control plants

The variability and reproducibility of the SUVH2 transgene expression was checked using quantitative real-time PCR. Many individual control plants of each class were examined as shown in the Figure 3.12. The control includes the parental SUVH2 overexpression lines of heterozygous, homozygous parental and, homozygous F1 plants obtained after PTGS and the plants, which have undergone post transcriptional gene silencing (PTGS) were checked. The results clearly demonstrate that expression of the controls is quite reproducible, and they have minimal intra variability between the samples. Additionaly, the results show that the relative values of heterozygous plants were nearly half of the homozygous plants. The expression profile of homozygous plants obtained after posttranscriptional gene silencing, was much lower then the original homozygous plants (Figure 3.12).

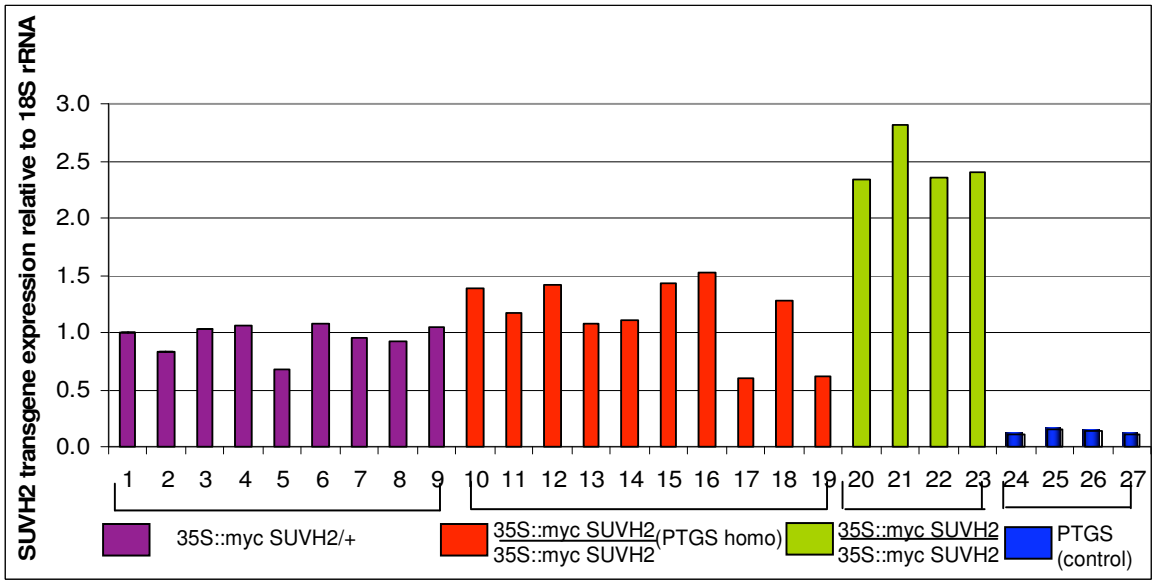


Figure 3.12: Expression analysis of different classes of SUVH2 overexpression lines. Relative levels of expression were determined by real-time PCR using SYBR fluorescent dye. Values were normalized to the expression of 18S rRNA. The graph shows the relative expressions of different lines with respect to the heterozygous overexpression plants. The numerical values on the X-axis represent the number of individual plants analyzed for each class of indicated genotype.

3.8 Real-time data of the SUVH2 overexpression in selected isolated mutants

The relative expression of the SUVH2 transgene from the isolated mutants was measured and compared with the starting overexpression line. As shown in Figure 3.13, the expression pattern of the putative suppressors was not uniform and lower then the starting line. However the expression of the transgene in all the mutants was higher then the post transcriptionally silenced negative control. Lower expression of the SUVH2 transgene could possibly result in identification of false positive suppressors hence only those putative mutants expressing high levels of SUVH2 were analyzed further.

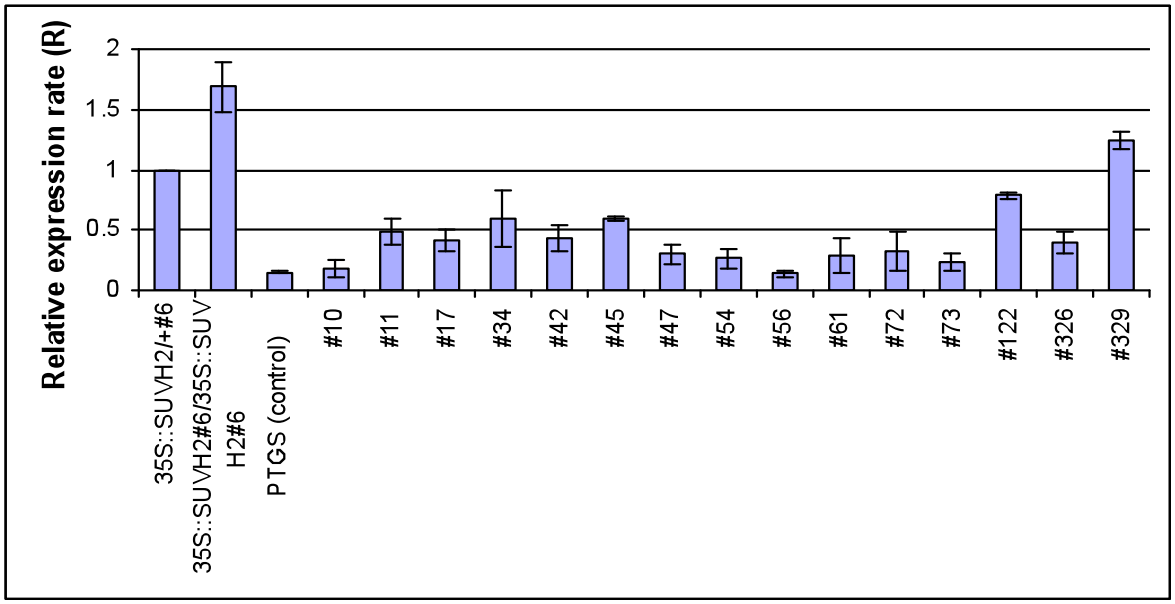


Figure 3.13: Expression analysis of putative dominant suppressors of SUVH2 overexpression. Relative levels of expression were determined by real-time PCR using SYBR fluorescent dye. Values were normalized to the expression of 18S rRNA. The graph shows the relative expression of different independent putative mutants with respect to the heterozygous overexpression line.

3.9 Immunocytological studies of SUVH2 overexpression in the background of the selected mutants

When SUVH2 is overexpressed, it leads to ectopic heterochromatization and severe phenotypic defects. When, interphase nuclei of wild type *Arabidopsis* were stained with H3K9me2 specific antibody, which is a typical heterochromatin mark, the staining localized only in the chromo centers as shown in the Figure 3.14. However in the plants containing SUVH2 overexpression construct, this mark is not only present in the chromocenters but spreads all over the interphase nuclei. When the plants are suppressed post-transcriptionally,

the level of H3K9me2 staining is reduced back to wild type. This is also supported by the real-time expression data of the plants, which have undergone PTGS, as they have a very low level of expression of SUVH2 transgene. However in next generation, PTGS is reset again leading to the appearance of the curly leaf phenotype and to the restoration of expression characteristic of SUVH2 overexpression phenotype.

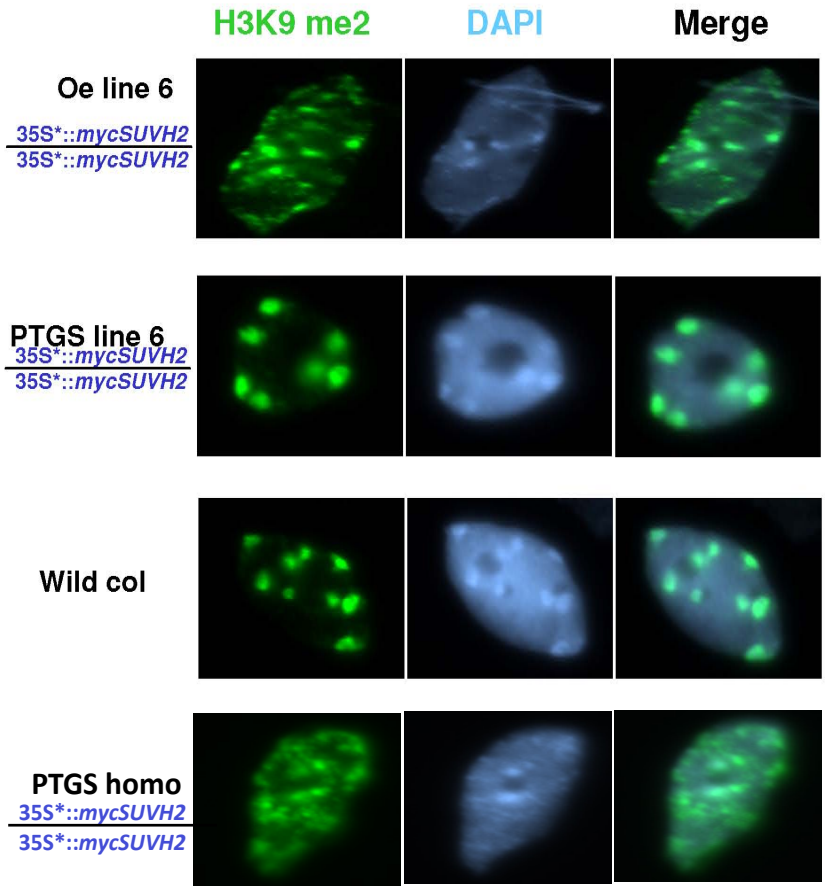


Figure 3.14: Immunocytological analysis of H3K9me2 patterns in wild type, the SUVH2 overexpression line, PTGS line, and the SUVH2 overexpression line after PTGS in next generation. Representative images are shown indicating ectopic distribution of the H3K9me2 mark in SUVH2 overexpression plants in comparison to wild type and PTGS plants, where the signal is restricted only to the chromocenters.

Five of all the isolated mutants were stained with H3K9me2 antibody in the background of SUVH2 overexpression and the mutants were categorized in two groups. The first group retains the ectopic distribution pattern as observed in SUVH2 overexpresion plants, while the second group resembles wild type plants. Surprisingly, when real-time expression data and staining data were examined together, it became evident, that some of the suppressors immunocytologically appeared similar to overexpression lines (in showing ectopic distribution) even though the expression values were lower then the starting overexpressing

lines (Figure 3.27). Therefore the real-time expression data of such mutants were taken as a threshold and all mutants with similar or higher values were considered for further analysis.

3.10 Identification of transgenic *Arabidopsis* plants containing a single insertion of mutater T-DNA using Southern blot analysis

Southern blot analysis was carried out in order to determine the number of independently segregating mutater T-DNA in the insertion lines. Selected transgenic lines whose progeny showed Mendelian characteristics were subjected to Southern blot analysis. *bar* gene was used as a probe for hybridization, in order to determine the number of T-DNA insert. The probe was made by PCR amplification using *bar* gene specific primers. The PCR fragment was run on a 1% agarose gel and subsequently purified.

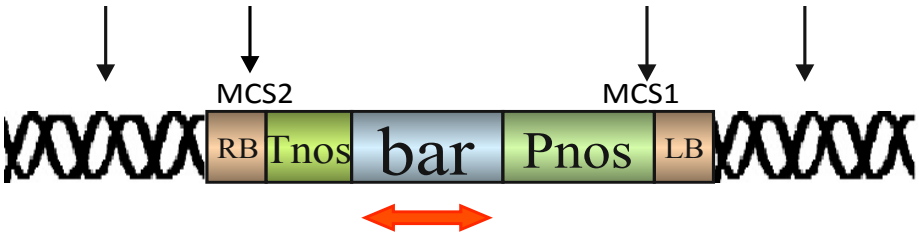


Figure 3.15: Figure indicating the principle of Southern blot analysis. Digestion with HindIII present in MCS I (Mulitiple Cloning Site) cuts the T-DNA only once in the T-DNA and in several places in the genomic region spanning the T-DNA. *bar* gene is used as a probe for hybridization to identify the number of copies of the inserted T-DNA in the genome.

Genomic DNA from individual transgenic lines was digested with restriction enzymes present in either MCS1 (Multiple Cloning Site 1) or MCS2 (Multiple Cloning Site 2). An enzyme was chosen, which cleaves the T-DNA only once as shown in figure 3.15. After hybridization and washing, single bands were detected in most lines thus confirming that these lines contained only a single copy of the T-DNA. Alternatively, appearance of more than one band indicates the presence of multiple copies of the transgene.

3.11 Identification of the genomic DNA sequences flanking the mutater T-DNA using inverse PCR

The insertion point of the mutater T-DNA was determined by Inverse PCR (I-PCR) I-PCR analysis was carried out by designing primers to locate the left border of the mutater T-DNA inserted randomly in the genome of *Arabidopsis thaliana.* After repeated attempts, the sequence derived from the independent lines was of sufficient length and quality to determine the position of integration in the *Arabidopsis* genome. BLAST homology search on the

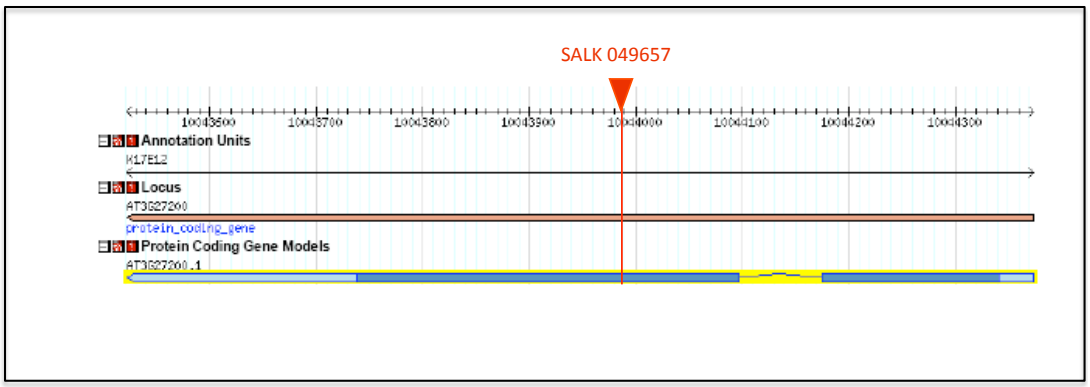
Arabidopsis database (TAIR) was used to exactly determine the position of the obtained sequences on the *Arabidopsis* chromosomes.

3.12 Genetic crosses of 35S::myc SUVH2/SUVH2 with non-specific homozygous SALK line

The decreased levels of SUVH2 expression in the putative mutants as compared to the control lines could be due to various factors. In order to reconfirm that insertion of the respective T-DNA in the putative gene was responsible for loss of the SUVH2 overexpression phenotype, a control cross with three different commercially available SALK lines was performed and analyzed [Figure 3.16 (a, b and c)]. These SALK lines were predicted to have inserts in different locations in the genome. The nonspecific genes were chosen on the criteria that their gene products were potentially targeted to organelles other than the nucleus and thus very unlikely to interact with SUVH2 in heterochromatin formation. The seeds from the respective SALK lines were sown and the plants were molecularly characterized for the presence of T-DNA insertion in the respective genes. Two rounds of PCR were carried out to show that plants were homozygous for the insertion. In the first round of PCR, the T-DNA border-specific primer and the gene specific primer were used. Appearance of bands confirmed the presence of T-DNA in that particular position of the genome. In the next round, gene specific primers were used on same set of genomic DNA probes indicating homozygosity of the T-DNA as no PCR bands were obtained, in contrast to wild type controls.

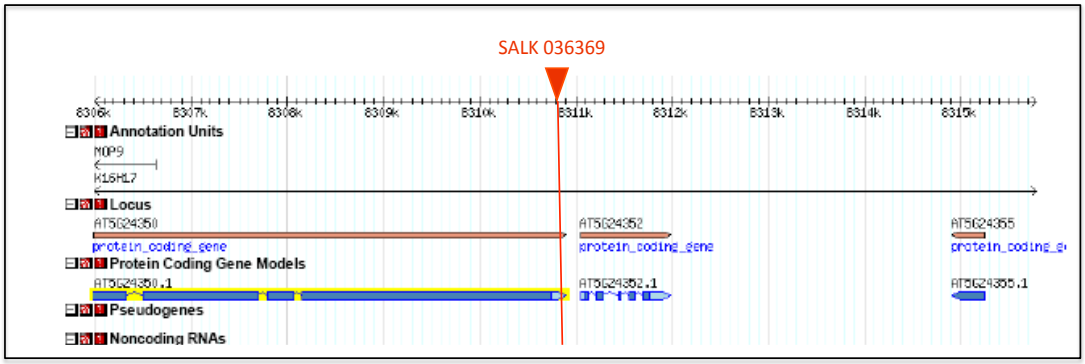
The pollen from homozygous nonspecific control T-DNA (male parent) was used to cross the stigma of the PTGS-derived homozygous overexpression plants (female parent). Crossing results of the next generation were analyzed. All the plants in the F1 generation should either be curly or suppressed, depending on the nature of the (dominant or recessive) suppressor. The SALK lines were only checked for the presence of that T-DNA, which had disturbed the respective genes and not for additional segregating inserts. Even though a strong 35S promoter drives the expression of the T-DNA in the SALK lines and is thus more prone to TGS (Transcriptional Gene Silencing), as the SUVH2 transgene is also driven by a similarly strong promoter, all progeny in F1 generation were curled as shown in Figure 3.16 indicating the functionality of the test system.

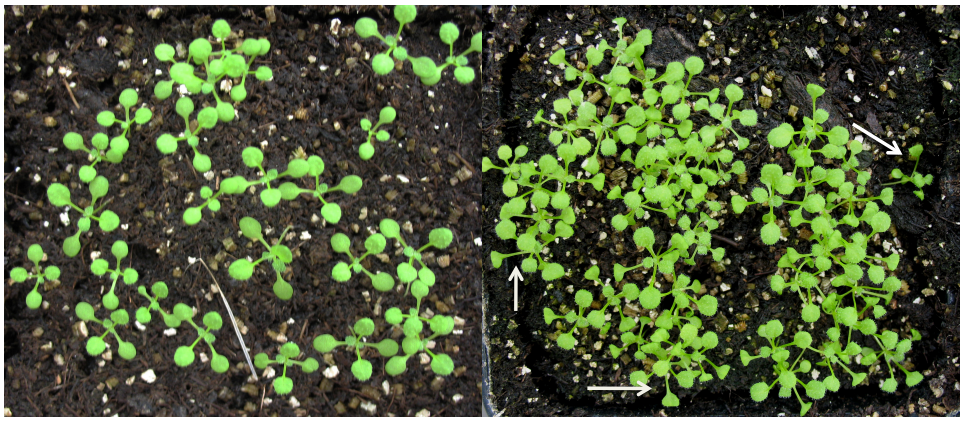
a



Wild Col x SALK049657 (At3g27200) 35S::mycSUVH2#6F1 x SALK049657 (At3g27200)

b





Wild Col x SALK036369 (At5g24350) 35S::mycSUVH2#6F1 x SALK036369 (At5g24350)

c

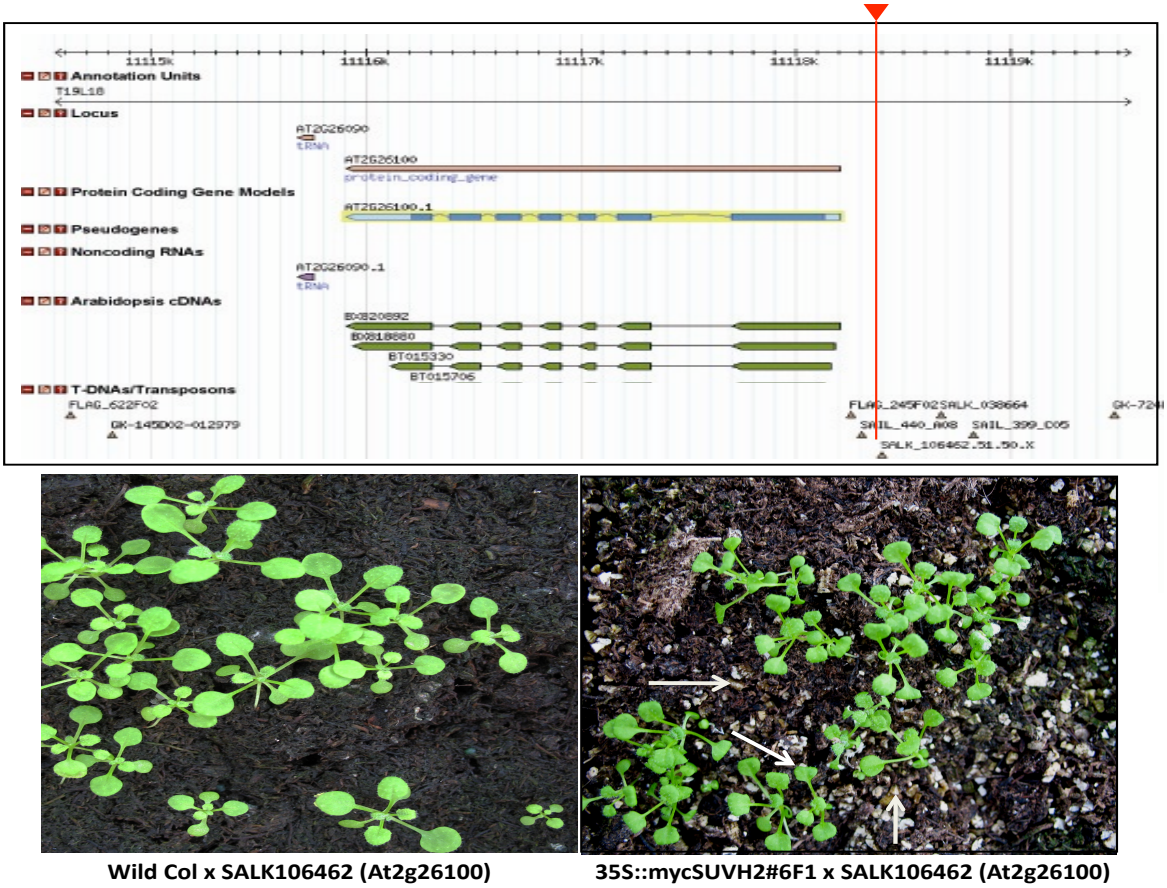


Figure 3.16: Genetic crosses, indicating the functionality of the screen for dominant suppressors. The homozygous starting line was independently crossed to three SALK lines with randomly inserted T-DNAs. (a) In the first SALK line, the T-DNA was inserted in the At5g24350 gene (SALK036369), whose gene product is involved in the secretory pathway, (b) The second had a disturbed At3g27200 gene (SALK049657), whose gene product is involved in aminoacid glycosylation and is targeted to chloroplasts and (c) the third had a disturbed At3g27200 gene (SALK049657), whose product is targeted to plastids and predicted to be involved in electron carrier activity. Curly progenies in F1 generation were obtained in all the crosses clearly demonstrating that transformation of another T-DNA in the background of SUVH2 otverexpression plants did not lead to silencing of the SUVH2 transgene.

The results of the non-specific crossings clearly demonstrate that the suppressed phenotype of the putative mutants obtained during the screening in T1 generation was not strongly affected by TGS of the SUVH2 transgene.

3.13 Identification and characterization of line 122

3.13.1. Identification of the insertion locus in the mutant line 122

PCR amplification and sequencing of the left border of the T-DNA inserted in the mutant line 122 indicated that the insertion had occurred within the promoter region of At3g56370. It is a putative gene located in the T5P19 contig of chromosome 3, as shown in Figure 3.17 (a and

b). The positions of the SALK line T-DNAs are also indicated. The site of Basta T-DNA insertion was precisely mapped on the *Arabidopsis* chromosomes (Appendix 2). Thus, the sequence analysis indicated that the putative gene At3g56370 has been disturbed by the insertion event.

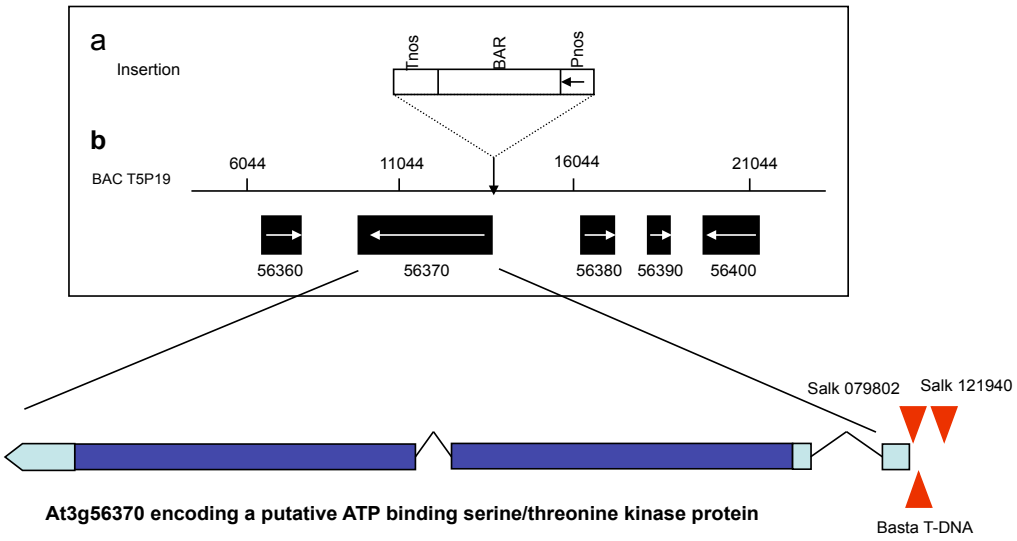


Figure 3.17: Genomic location of the T-DNA responsible for the suppressor effect in line 122. a) The transgenic insert consists of a single copy of the *bar* expression-cassette, driven by weak nos promoter. b) Integration site on chromosome 3. The numbers on top refer to the BAC clone T5P19. The numbers below are the accession numbers of neighboring genes. The position of the T-DNA insertion in the two commercially available SALK lines is also indicated.

According to the TAIR gene prediction, the At3g56370 gene encodes a putative transmembrane receptor protein of 964 amino acids having serine/threonine kinase activity aand therefore might be involved in ATP dependent phosphorylation of the downstream target proteins. The protein also has multiple leucine rich repeats (Figure 3.18), which, are suggested to function in protein-protein interactions (Nam and Li, 2002). Therefore it could be predicted that this protein might be involved as a transmembrane receptor protein in serine/threonine kinase receptor signaling pathway.

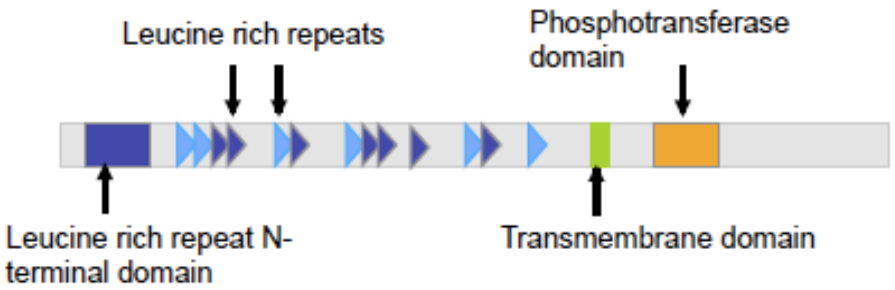


Figure 3.18: Distribution of different domains in the protein structure of At3g56370. Two domains leucine rich repeats and phosphotransferase are quite prominent. The presence of leucine rich repeats which would mediate protein -protein interaction and the phospho transferase domain is responsible for kinase activity.

3.13.2 RT-PCR analysis to check the expression of At3g56370

Plants with homozygous mutation in the At3g56370 gene were identified by PCR, using gene-specific and T-DNA-specific primers as described above. The homozygous plants showed no morphological and developmental alterations apart from the loss of the SUVH2 overexpression phenotype compared to wild type controls.

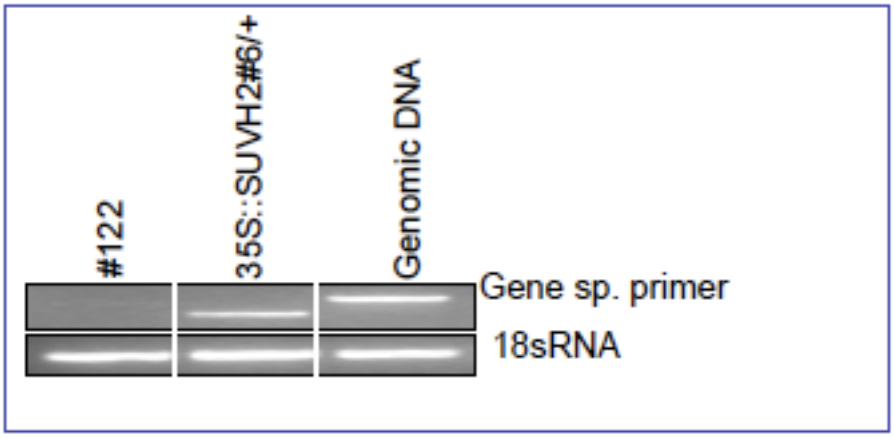


Figure 3.19: RT-PCR analysis indicating loss of function allele of the gene At3g56370. Expression analysis indicating transcript accumulation only in the control using gene specific primer, confirmed that T-DNA insertion resulted in complete disruption of the expression of the kinase gene. 18SrRNA gene (At2g01010) served as a loading control.

The RT-PCR revealed complete absence of the At3g56370 transcript in plants indicating that it is a null alleleas shown in Figure 3.19. The expreesion levels were compared to the levels of 18S rRNA. This reinforced our hypothesis that the down regulation of the expression of this gene could lead to the dominant suppressor effect.

3.13.3 Presence of single insert of mutater T-DNA was confirmed by Southern blot analysis

In T2 generation of line 122, plants showed a segregation ratio of 3:1, indicating the presence of a single locus. However the confirmation of this result was done by Southern blot analysis using *bar* gene as a probe [Figure 3.20(a)]. The genomic DNA was digested with HindIII, as only one restricted site was present in mutater T-DNA. A single band was seen in the

Southern blot confirming the presence of a single copy of inserted T-DNA in the genome. The full length of the gene with 1.2 kb of upstream promoter sequence was cloned in pGEM vector and will be further used for complementation and overexpression studies [Figure 3.20 (b)].

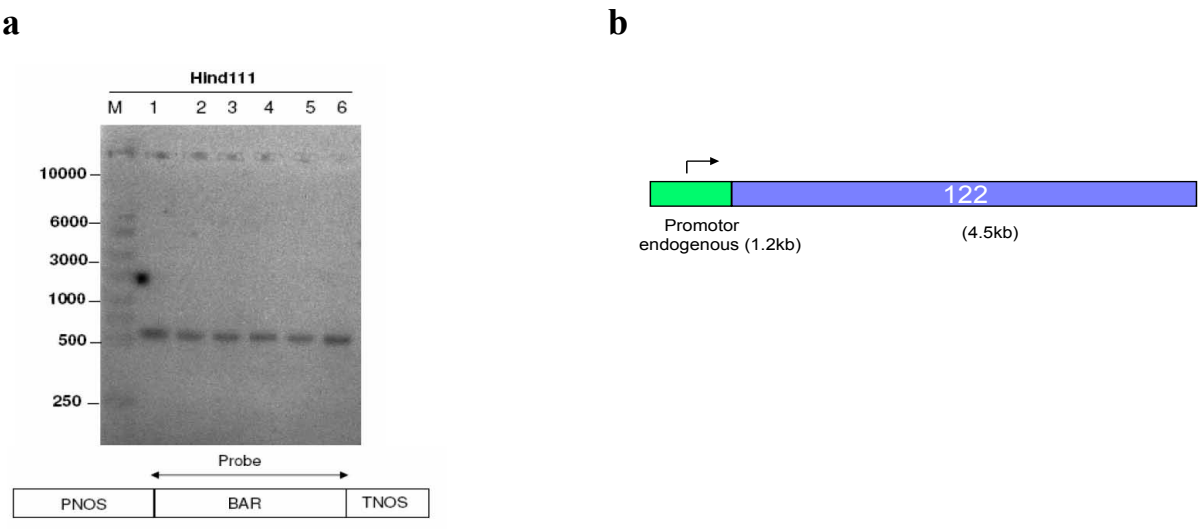


Figure 3.20: Southern blot analysis of mutant line 122 confirming the presence of single insert and cloning the gene for further analysis. a) Six independent progeny from the line 122 were subjected to Southern blot analysis. The genomic DNA was digested with HindIII and after the transfer; the blot was hybridized with ^{32}P labelled *bar* coding sequence. Molecular size markers in kb are indicated on the left. b) The entire coding sequence of the gene along with the endoenous promoter including 1.2 kb upstream genomic region was cloned in pGEM vector using TA cloning for further analysis.

The insertion of a single copy of mutater T-DNA, resulted in complete loss of At3g56370 transcript making this line a promising candidate among other novel dominant suppressor mutation of the SUVH2 overexpression phenotype.

3.13.4 Real-time expression and immunocytological analysis of SUVH2 in the mutant line 122

The presence of another T-DNA could influence the expression of the SUVH2 transgene and hence the expression was quantified by real-time PCR analysis. SUVH2 overexpression in the presence of mutater T-DNA was measured by real-time PCR analysis relative to the starting overexpression line. As shown in Figure 3.21 (a), the relative values are higher than the post-transcriptionally silenced line but lower then the starting line. However, when the interphase nuclei were stained with an H3K9me2-specific antibody, the ectopic H3K9 distribution characteristic of SUVH2 overexpression was also visible in the suppressor line 122 [Figure

3.21 (b)], indicating that the expression levels were sufficient for ectopic H3K9me2 distribution.

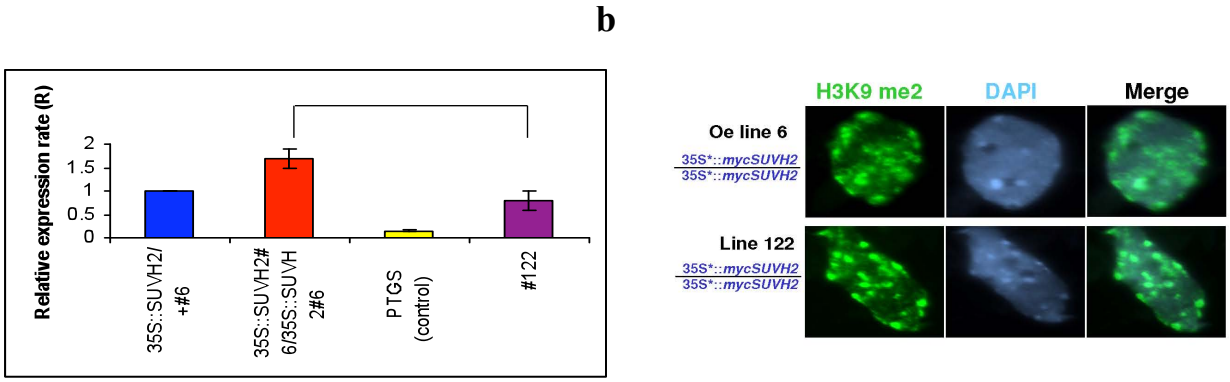


Figure 3.21: H3K9me2 distribution and expression analysis in mutant line 122. a) Relative levels of expression were determined by real-time PCR using SYBR fluorescent dye. The graph represents expression levlels of SUVH2 in homozygous, PTGS (control) and mutant line 122 relative to the heterozygous overexpression line. Values were normalized to the expression of 18S rRNA. b) Representative images indicating ectopic distribution of the H3K9me2 mark in mutant line 122 as well as in SUVH2 overexpression plants are shown.

3.13.5 Trans acting effect of the mutant locus 122 on the SUVH2 overexpression phenotype

The loss of phenotype in SUVH2 overexpression plants could have originated from cis acting mutations within the transgene or from mutations in trans. In order to distinguish between these two possibilities, the mutant locus of 122 was segregated from the *SUVH2* overexpression transgene and retested on the starting overexpression line. The heterozygous nature of the target T-DNA containing the SUVH2 transgene allowed easy separation of the trans acting locus from the SUVH2 overexpression line by segregation analysis as the plants segregated in typical Mendelian ratio of 1:2:1 (35S::SUVH2 /35S::SUVH2, 35S::SUVH2/+, +/+). All the wild type plants without SUVH2 transgene were genotyped again to identify plants containing the homozygous mutater T-DNA. The separated mutant locus was backcrossed to the homozygous SUVH2 overexpression plants and the presence of all suppressed normal plants in next generation confirmed that the trans acting locus was responsible for the dominant suppressor effect (Figure 3.22).

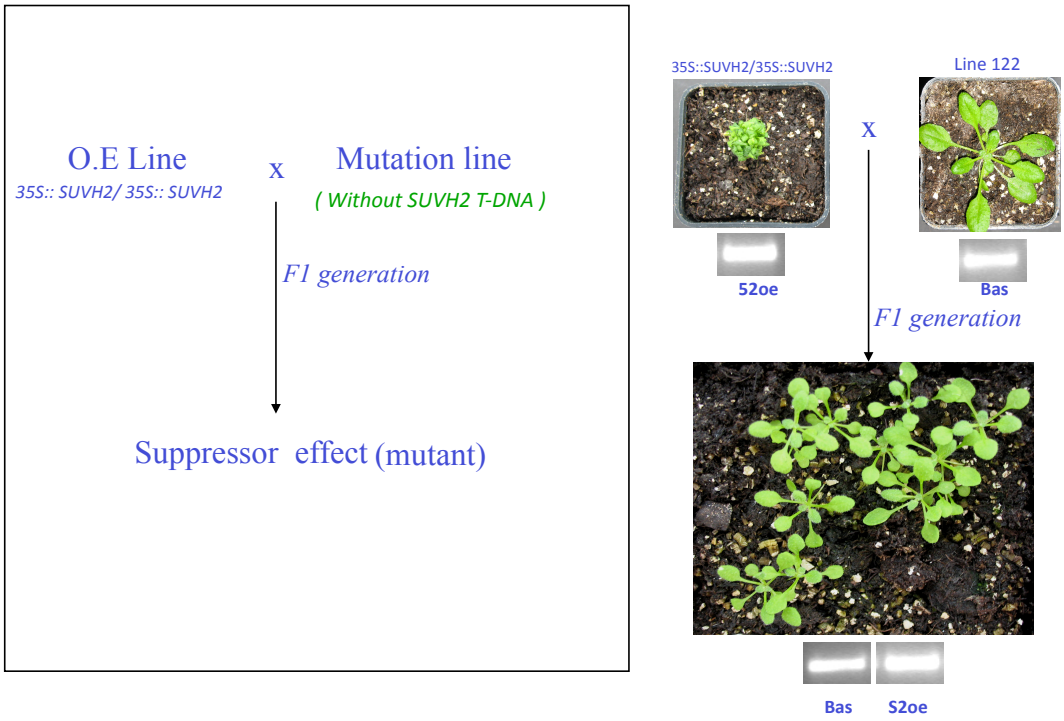


Figure 3.22: Reappearence of suppressor effect in SUVH2 overexpression line backcrossed to line 122 Plants with a homozygous mutant locus of At3g56370 were crossed with the SUVH2 overexpression line. Suppressed progeny were obtained in F1 generation confirming the presence of a dominant effect of the isolated mutant.

3.13.6 T-DNA inserted in the At3g56370 gene is responsible for the dominant suppressor phenotype

To further confirm that the T-DNA insertion in the promoter of the kinase At3g56370 gene is responsible for the dominant suppressor phenotype, two more insertion lines from the SALK collection were analyzed which have a predicted insertion in the promoter region of the At3g56370. The SALK lines had the similar T-DNA insertion as the originally identiied mutant from the genetic screen. Hence these two T-DNA insertions were molecularly analyzed. Homozygous plants of the SALK lines were isolated and their insertion in promoter region was confirmed by PCR analysis. Expression of the kinase gene (At3g56370) was checked in homozygous SALK plants by semi quantitative RT-PCR using gene specific primers. 18SrRNA was used as a loading control. As shown in the Figure 3.23, RT PCR analysis revealed a reduction of gene specific transcripts in the SALK lines as compared to the wild type control.

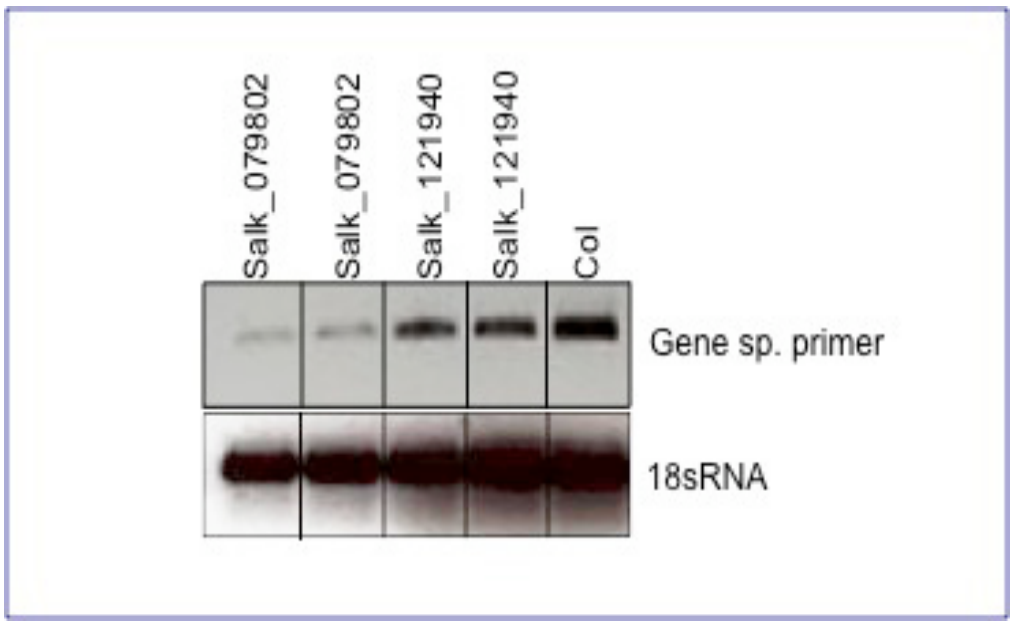
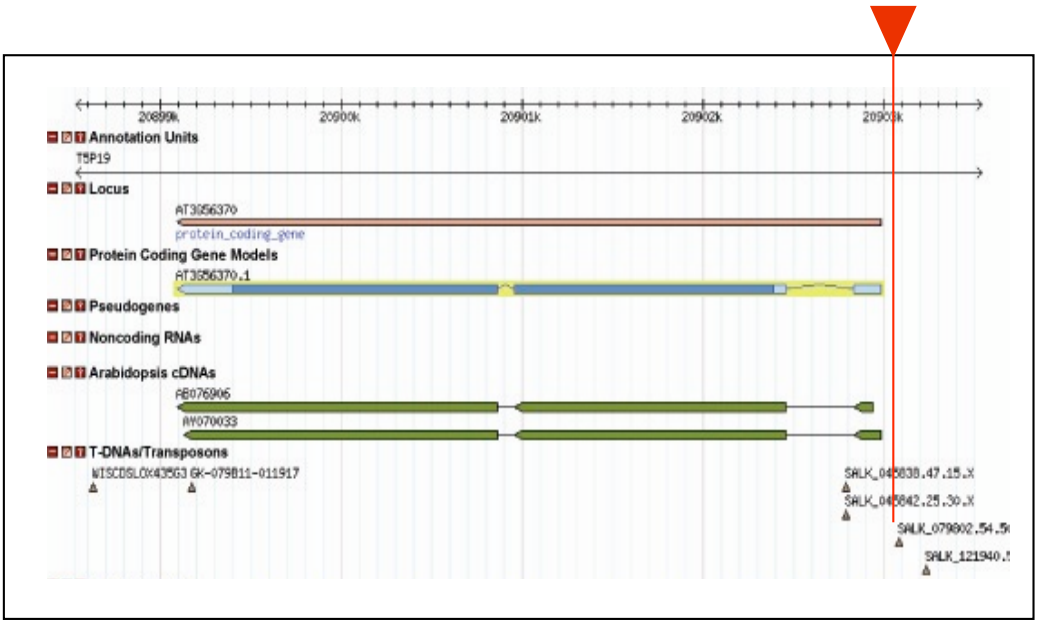


Figure 3.23: RT-PCR analysis indicating the down-regulation of the At3g56370 transcript in the homozygous T-DNA insertion line. Expression analysis indicating low levels of transcript accumulation in the T-DNA insertion SALK lines compared to that of wild type control using gene specific primers. The down regulation of the expression in the T-DNA correlates with the suppressor effect.

In order to attribute the reduction in the transcript of the At3g56370 responsible for the suppressor effect, the pollen from a male homozygous SALK line were used to fertilize the the stigma of homozygous overexpression SUVH2 female, resulting in the generation of F1 seeds. As shown in the Figure and 3.24 and 3.25, all the F1 plants were suppressed, confirming the kinase gene to be a putative suppressor. PCR analysis was done, in order to show the presence of both the T-DNAs in suppressed plants whereas, only curly plants were observed, in the presence of SUVH2 transgene. For the second T-DNA, real-time expression analysis of the suppressed plants indicated that, the expression of the SUVH2 transgene was not reduced even in the presence of CaMV 35S promoter driven SALK T- DNA.



35S::mycSUVH2#6F1 x Wild Col 35S::mycSUVH2#6F1 x SALK079802(At3g56370)

Figure 3.24: The T-DNA inserted in the promoter region of At3g56370 gene is responsible for the suppressor effect. Progeny of homozygous T-DNA line was identified and crossed with the homozygous SUVH2 overexpression plants and scored for suppressor effect in next generation. As compared to the control all suppressed plants, were obtained in next generation.

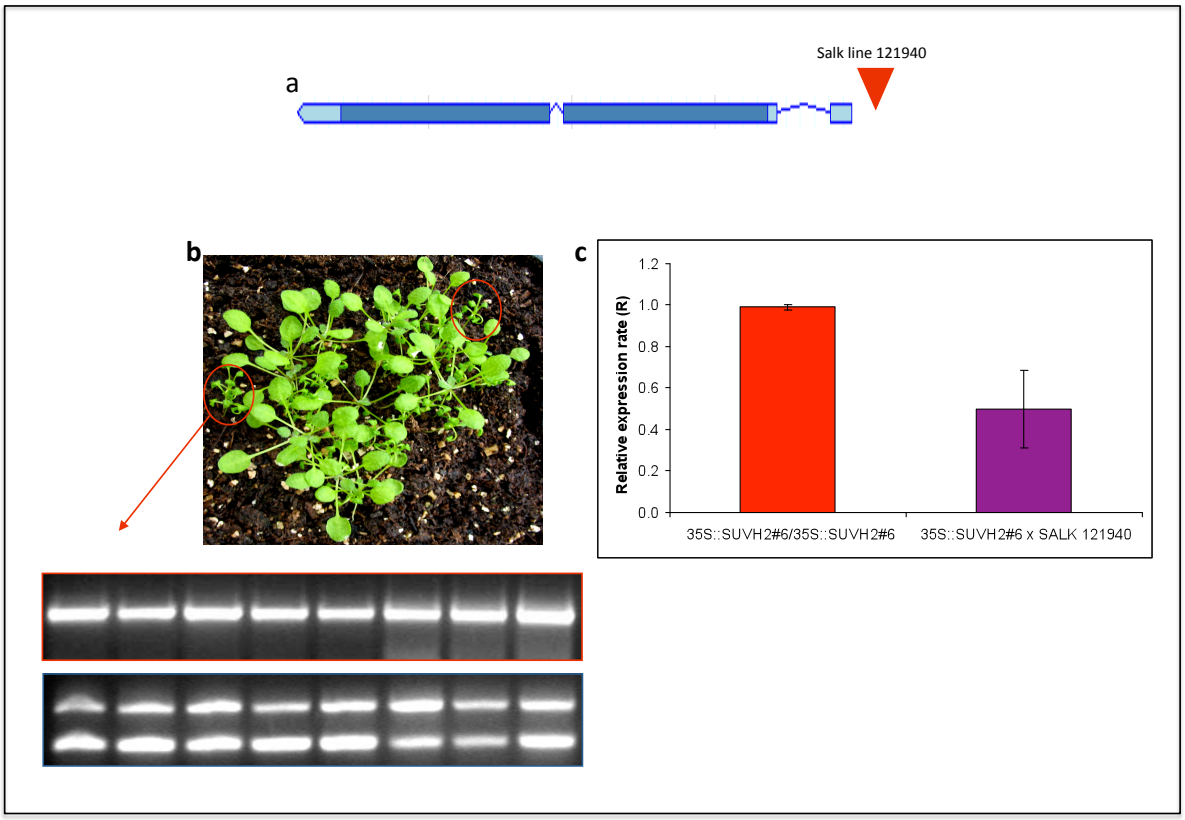


Figure 3.25: The 2nd T-DNA inserted in the promoter region of At3g56370 gene confirmed that this gene is responsible for the dominant suppressor effect. a) Position of the T-DNA insertion in the SALK line 121940. b) Progeny of the homozygous SALK T-DNA line were identified, crossed with the homozygous SUVH2 overexpression plants and scored for suppressor effect in next generation. PCR analysis of the progeny confirmed the presence of only SUVH2 transgene in curly plants while the suppressed plants contain both SUVH2 overexpression construct as well as SALK T-DNA. c) Real-time RT-PCR analysis comparing the expression of SUVH2 overexpression plants with the progeny of the crossed plants that were obtained after crossing the SALK T-DNA line/plants with the PTGS homo plants.

3.14 Identification and characterization of line 326

3.14.1. Identification of the insertion locus in the mutant line 326

The insertion point of the mutater T-DNA in the line 326 was identified by inverse PCR. As shown in the Figure 3.26 (a and b), the insertion was in the first exon of a gene, At2g23690, located in the F27L4 contig of chromosome 2 thus confirming that a null allele was generated by T-DNA insertion. The confirmation that the identified locus was responsible for the suppression effect was achieved by the co segregation analysis of the mutant locus with the suppressor effect. Presence of two-conserved nuclear targeting signal as determined by PSORT program (http://psort.ims.u-tokyo.ac.jp) explicitly point out that, this is a nuclear protein [Figure 3.26 (c)]. However, no conserved domains were found in the protein. The full-length gene with its promoter, located 1.3 kb of upstream was cloned in pGEM vector and

will be further used for complementation and overexpression studies [Figure 3.26 (d)]. The presence of distinct nuclear targeting signal makes this protein an interesting candidate for further analysis.

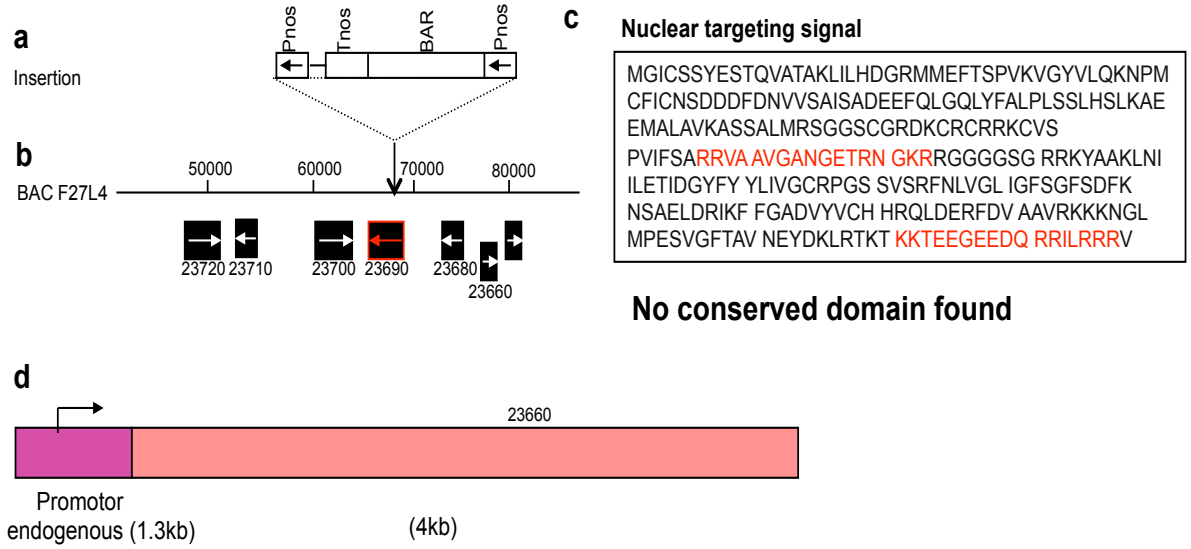


Figure 3.26: Insertion point of the T-DNA in line 326 responsible for SUVH2 overexpression suppressor effect. a) The transgenic insert consisted of a tandem repeat copy of the *bar* expression cassette driven by weak nos promoter. b) Integration site on chromosome 2. Number on top refers to the BAC clone F27L4. Numbers below are the accession numbers of neighboring genes. c) Protein secquence indicating that the conserved basic residues form a nucler-targeting signal. d) The entire coding sequence of the gene along with the endoenous promoter was cloned in pGEM vector using TA cloning, for further analysis.

3.14.2 Real-time expression and immunocytological analysis of SUVH2 in the mutant line 326

The mutant line 326 had more than one inserts. The presence of multiple inserts might induce repeat dependent silencing, could lead to a decreased in the expression of the transgene. Therefore the expression of the SUVH2 transgene was checked both by immunostaining and by quantitative real-time PCR analysis. Real-time PCR analysis showed that, in comparison to the control, the expression of SUVH2 was reduced as shown in the Figure 3.27 (a). This was probably due to TGS (Transcriptional Gene Silencing) as muliple copies of T-DNA often leads to repeat dependent silencing of the transgene. However even in the presence of multiple Basta T-DNA insert, ectopic distribution of H3K9me2 similar to starting SUVH2 overexpression line was observed. This indicated that even though the transgenic expression of SUVH2 was decreased in the mutant line but the level of expression was above the threshold level to show ectopic distribution [Figure 3.27 (b)].

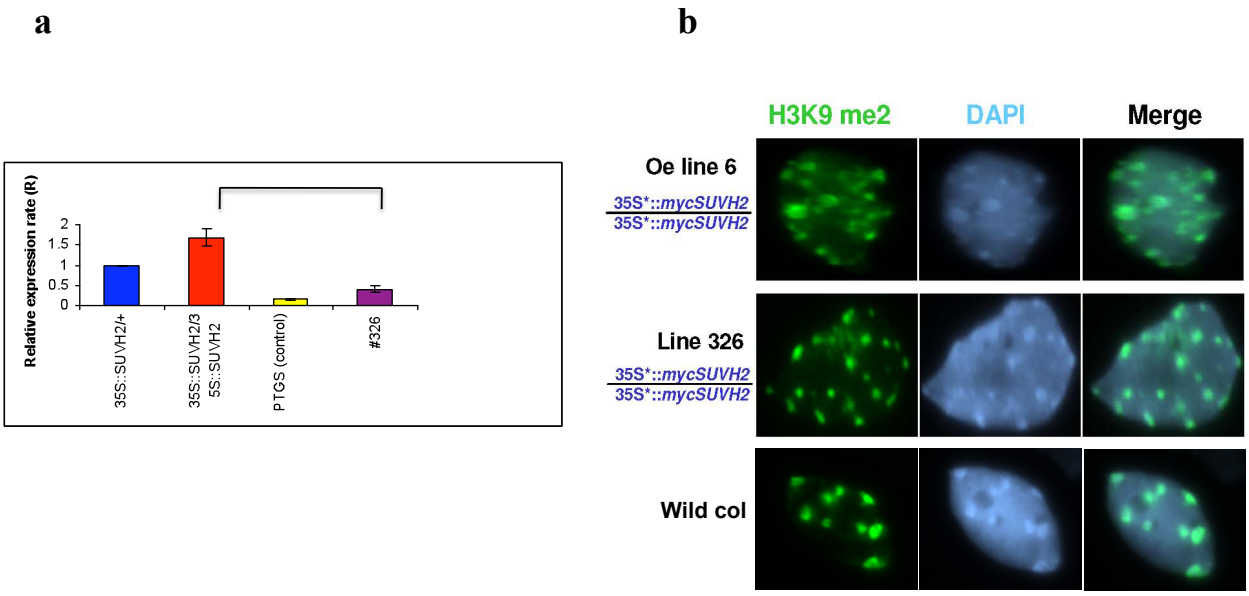


Figure 3.27: Variation in expression analysis and H3K9me2 distribution in mutant line 326 in comparison to SUVH2 overexpression plants. a) Relative levels of expression were determined by real-time PCR using SYBR fluorescent dye. The graph represents expression levlels of SUVH2 in homozygous, PTGS (control) and mutant line 326 relative to the heterozygous overexpression line. Values were normalized to the expression of 18S rRNA. b) Representative images are shown indicating ectopic distribution of H3K9me2 mark in mutant line 326 as well as in control SUVH2 overexpression plants in comparison to wild-type.

3.14.3 Genetic cross of mutant locus with SUVH2 overexpression line

The SUVH2 transgene in insertional line 326 was in homozygous background, and therefore had to be outcross to wild type Columbia to produce heterozygous F1 plants [Figure 3.28(a)]. The obtained plants were selfed to get progeny in which the SUVH2 transgene was removed in succeeding generation. Plants with only the mutant locus, which served as male and was used, to cross female SUVH2 overexpression plants [Figure 3.28(b)]. Completely suppressed plants as shown in Figure 3.28 (c) were obtained indicating that the mutant locus could be a possible modifier of SUVH2 overexpression phenotype. Appearance of curly plants along with the suppressed plants as shown, is indicative of presence of more then one Basta-T-DNA insert also supported by the Southern blot, [Figure 3.28(d)] using *bar* gene as a probe.

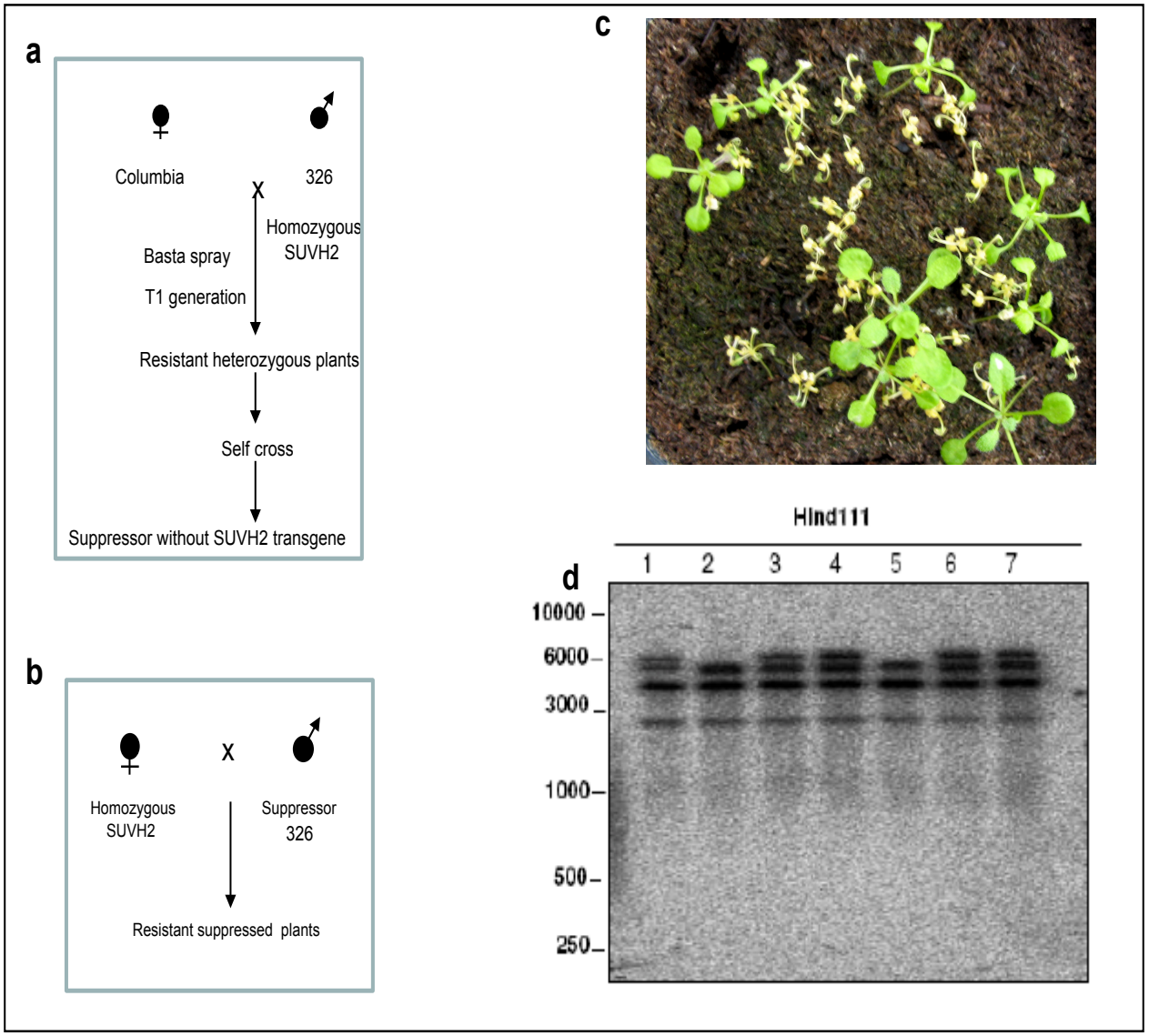


Figure 3.28: Reappearence of suppressor effect in SUVH2 overexpression line back crossed to line 326. a) SUVH2 transgene was segregated out by crossing to wild type and further selfing them in the next generation. b) Homozygous mutant locus without SUVH2 transgene was crossed to SUVH2 overexpression line. c) Suppressed progeny were obtained in F1 generation, confirming the presence of dominant effect of the isolated mutant. d) Seven independent progeny from the line 326 were subjected to Southern blot analysis. The genomic DNA was digested with HindIII and after the ransfer, the blot was hybridized with ^{32}P labelled probe containing the *bar* coding sequence. Molecular size markers in kb are indicated on the left.

3.14.4 Phylogenetic tree analysis indicating At2g23690 and its evolutionary descent

The phylogenetic tree was constructed from the alignment of full-length protein sequences. As seen in the Figure 3.29, the *Arabidopsis* genes cluster in a species-specific distinct clade indicating that this protein is evolutionary conserved among plant species, and has a plant specific functional role in various conserved processes.

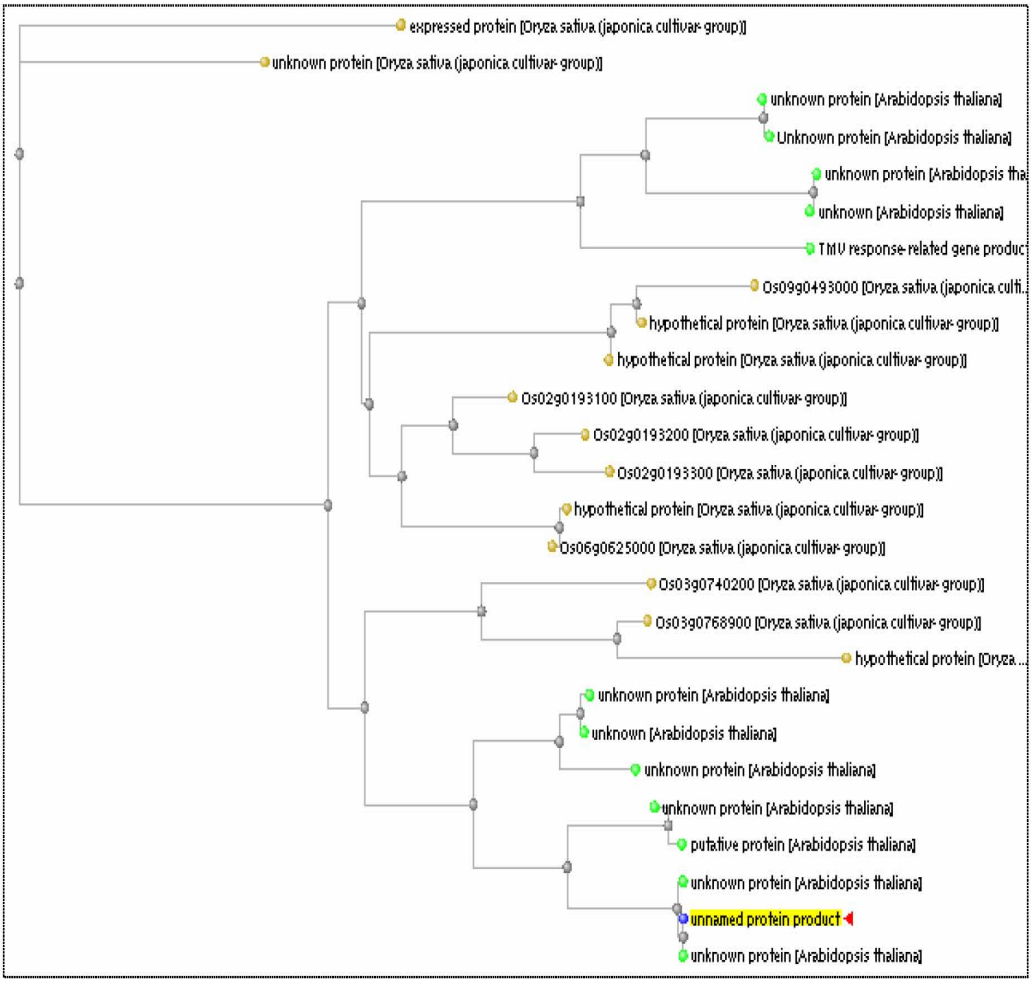


Figure 3.29: Phylogenetic tree showing conserved plant specific protein in different species. Homologous proteins from different plant species were aligned based upon the similarity of the amino acid sequences.

3.14.5 SALK T-DNA inserted in At2g23690 gene also shows a suppressor effect

The T-DNAs from SALK lines were used to reconfirm the suppressor effect. SALK069675C was located in the first exon, similar to the identified Basta mutation in the screen. SALK048198C was inserted in the first intron as indicated in the [Figure 3.30(a)]. The position of T-DNA was confirmed by PCR analysis and homozygous plants were selected for genetic crosses in order to confirm the suppressor effect of this locus [Figure 3.30(b) and (c)]. In the entire crossing scheme, SALK line T-DNA was used as males and were employed to cross female homozygous overexpression line.

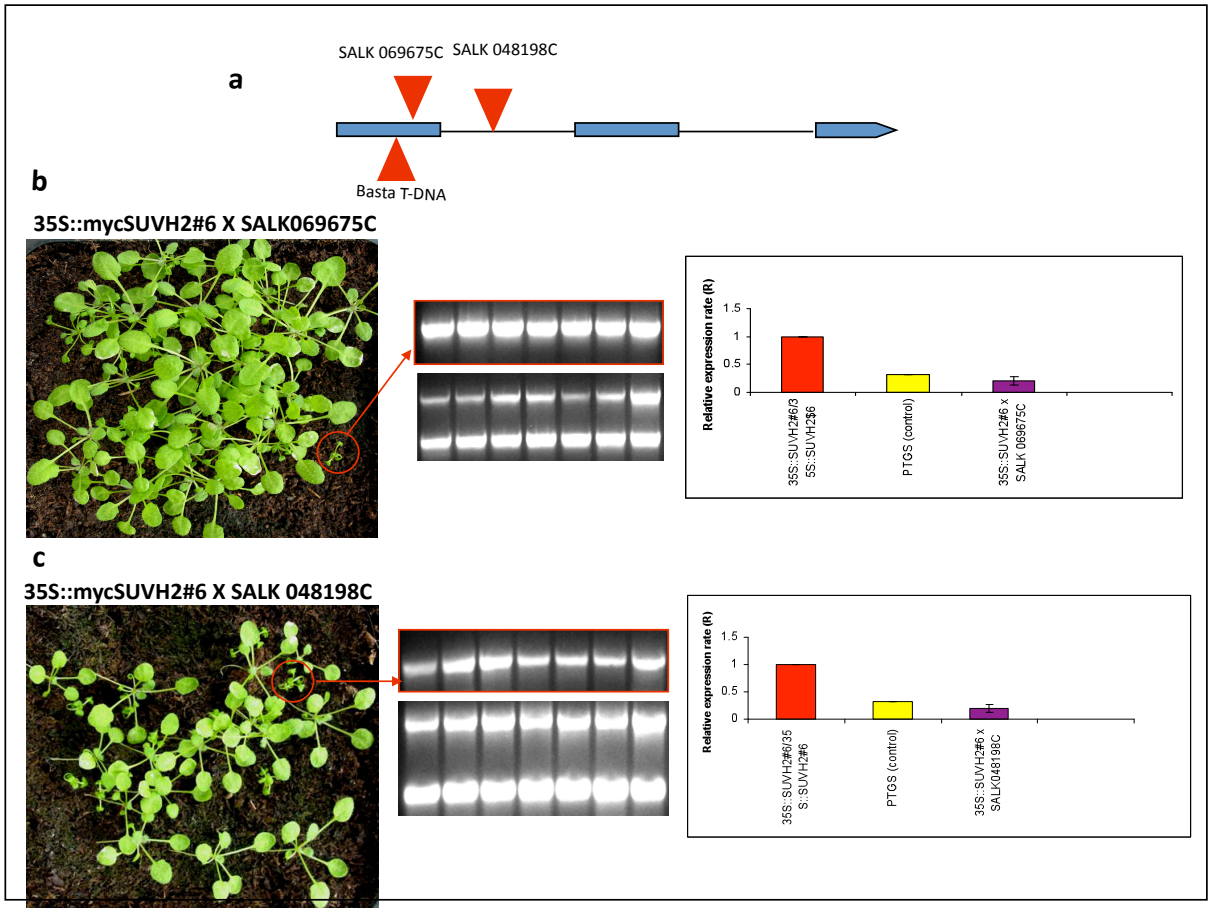


Figure 3.30: The T-DNA inserted in the promoter region of At2g23690 gene is responsible for the suppressor effect. a) Position of the T-DNA insertion in the SALK069675C and SALK048198C. b) Homozygous SALK T-DNA line progeny were identified and crossed with the homozygous SUVH2 overexpression plants and scored for suppressor effect in next generation. In the progeny of crossed plants, the presence of only SUVH2 transgene in curly plants and the SUVH2 overexpression construct as well as SALK T-DNA in suppressed plants was confirmed by PCR. c) Real-time RT PCR analysis comparing the expression of SUVH2 overexpression plants with the progeny of the crossed plants obtained after crossing the SALK T-DNA with the PTGS homo plants.

As shown in the Figure 3.30 (b and c), suppressor effect was only obtained when both the T-DNA were present as indicated by PCR analysis, however in the absence of SALK T-DNA, curly plants were observed. However the expression data indicated the expression levels of the transgene in the progeny was quite low and therefore real-time expression data has to be repeated. The low expression of SUVH2 in these crossed plants could be due to the onset of transcriptional gene silencing (TGS), fluctuation in the SUVH2 expression in different progeny of crossed plants or due to the presence of multiple number of SALK T-DNA insert.

3.15 Identification and characterization of line 329

3.15.1 Identification of the insertion locus of the mutant line 329 encoding a novel bromo domain containing protein (At5g10550)

PCR amplification and sequencing of the left T-DNA insert border indicated that the insertion in line 329 occurred in the promoter region of a putative gene, At5g10550, located in the F12B17 contig of chromosome 5 [Figure 3.31 (a and b)]. The exact insertion point of mutater T-DNA is shown in appendix 2. The points of insertion of the available SALK T-DNAs lines are also indicated. Thus, the sequence analysis indicated that due to the insertion event, possibly a putative gene At5g10550 has been disturbed. Homozygous parents of the insertional mutation At5g10550 gene were isolated by designing and amplifying PCR product using gene specific and T-DNA specific primers. The homozygous plants apart from the loss of SUVH2 overexpression phenotype showed no morphological and developmental alterations in comparison to wild type controls. The RT-PCR revealed significant down regulation of the At5g10550 specific transcript in homozygous mutant plants indicating that it is not a null allele [Figure 3.31 (c)]. Therefore the down regulation of the expression of this gene could lead to dominant suppressor effect in SUVH2 overexpression plants. Gene specific primers were used for the transcript analysis. RT PCR analysis of the 18SrRNA gene At2g01010 showed constitutive expression and served as a loading control.

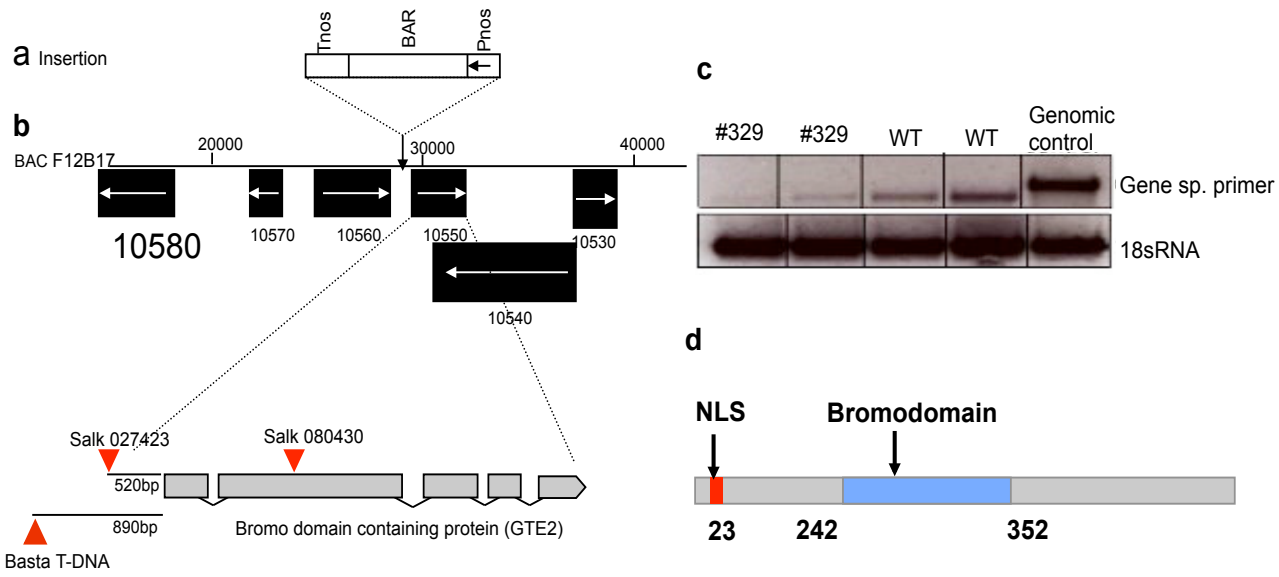


Figure 3.31: Genomic location of the T-DNA responsible for suppressor effect in line 329. a) The transgenic insert consists of a single copy of the *bar* expression cassette driven by weak nos promoter. b) Integration site on chromosome 5. Number on top refers to the BAC clone F12B17. Numbers below are the accession numbers of neighboring genes. The position of the two commercially available SALK line is also indicated. c) Expression analysis indicating less transcript accumulation in the T-DNA insertion line compared to that of wild type

control using gene specific primer. The down regulation of the expression of the protein in the T-DNA lines correlates with the suppressor effect. d) Presence of a bromo domain is quite prominent which recognizes acetylated lysine residues on the N-terminal tails of the histones. A distinct nuclear targeting signal also supports the possible functional role of this protein in chromatin modification.

According to the TAIR gene prediction, the AT5g10550 gene encodes a bromo domain protein of 678 amino acids. This gene is grouped under Global Transcription Factor group E2. The protein has a nuclear targeting signal between amino acid residues 23 and 39, and a characteristic bromo domain at position 242 to 352 [Fig 3.31 (d)]. Therefore it could be predicted that this protein has a functional role in the nucleus and might be directly involved in the SUVH2 mediated silencing pathway.

3.15.2 Southern blot analysis confirming the presence of single copy of mutater T-DNA in line 329.

In T2 generation of the line 329, plants showed a segregation pattern of 3:1 ratio, indicating insertion of a single T-DNA. Further confirmation of this result was done by Southern blot analysis using *bar* gene as a probe as indicated in the Figure 3.32. XbaI and HindIII digestion were used, as both have only one restriction site in the T-DNA. Presence of a single band therefore confirmed that this line contained only one copy of T-DNA. The insertion of single copy of mutater T-DNA resulted in down regulation of transcript of the At5g10550 gene.

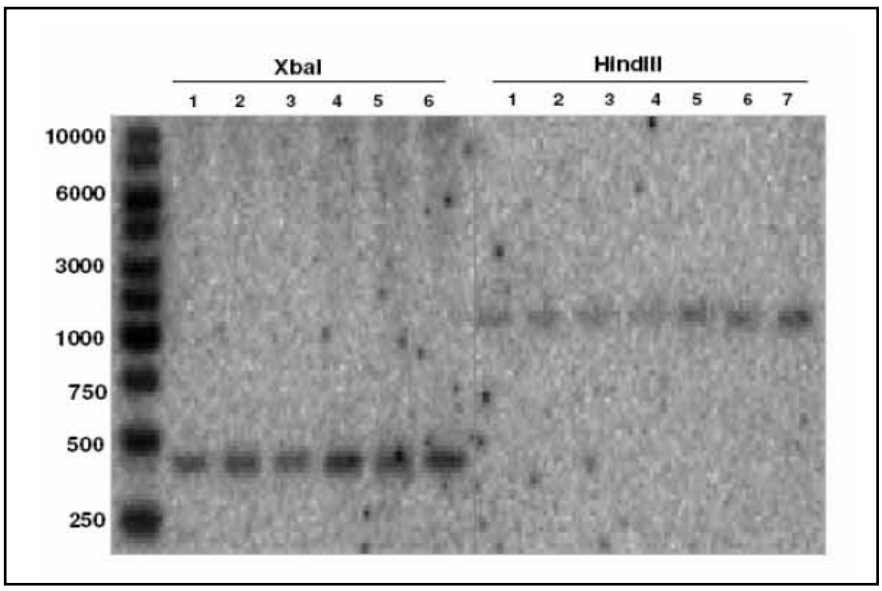


Figure 3.32: Southern blot analysis of mutant line 329 confirming the presence of single insert. Six to seven independent progeny from the line 329 were subjected to Southern blot analysis. The genomic DNA was digested with XbaI and HindIII and after the transfer; the blot was hybridized with ^{32}P labelled *bar* coding sequence. Molecular size markers in kb are indicated on the left.

3.15.3 Real-time expression and immunocytological analysis of SUVH2 in the mutant line 329

The presence of another T-DNA could influence, the expression of SUVH2 transgene and hence the expression was quantified by real-time PCR analysis. Relative to the starting over expression line, the values of SUVH2 overexpression in the presence of mutater T-DNA was measured by real-time PCR analysis. As shown in the Figure 3.33 (a), the relative values are higher then the posttranscriptional silent line and heterozygous plants but lower then the starting homozygous line. In immunocytological analysis, when the interphase nuclei were stained with an H3K9me2 specific antibody, the distribution was restricted to chromocenters similar to the wild type [Figure 3.33 (b)]. Since the expression of the SUVH2 in line 329 was higher then the heterozygous overexpression line which, showed a characteristic ectopic distribution of H3K9me2, suggesting that the gene probably might play an important role in controlling the distribution of H3K9me2 mark.

a **b**

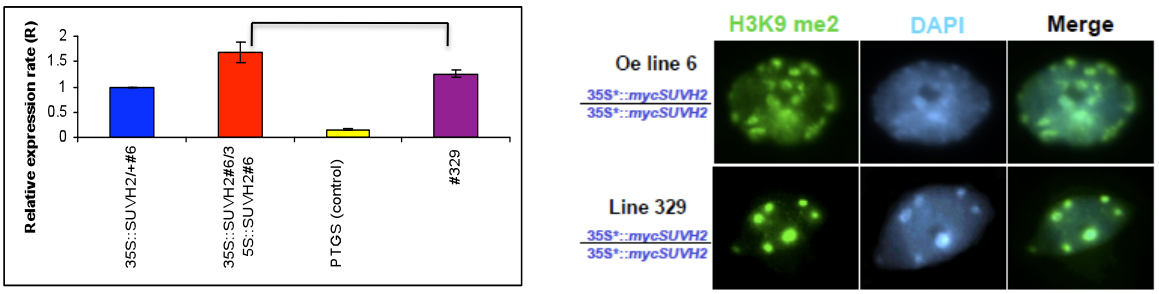


Figure 3.33: Variation in H3K9me2 distribution and expression analysis mutant line 329 in comparison to SUVH2 overexpression plants. a) Relative levels of expression were determined by real-time PCR using SYBR fluorescent dye. The graph represents expression levlels of SUVH2 in homozygous, PTGS (control) and mutant line 329 relative to the heterozygous overexpression line. Values were normalized to the expression of 18S rRNA. b) Representative images are shown indicating distribution of H3K9me2 mark in mutant line 329 as well as in SUVH2 overexpression plants.

3.15.4 Trans acting effect of the mutant locus on SUVH2 overexpression phenotype

To confirm that the suppressor effect is due to the mutater T-DNA, this T-DNA was separated from the T-DNA containing SUVH2 transgene. As both the T-DNAs were in heterozygous state, it allowed easy separation of mutater T-DNA from the SUVH2 overexpression line by segregation analysis. The separated mutant locus was crossed to the homozygous SUVH2

overexpression plants and the presence of all suppressed plants in next generation confirmed that the trans acting locus was responsible for the dominant suppressor effect (Figure 3.34).

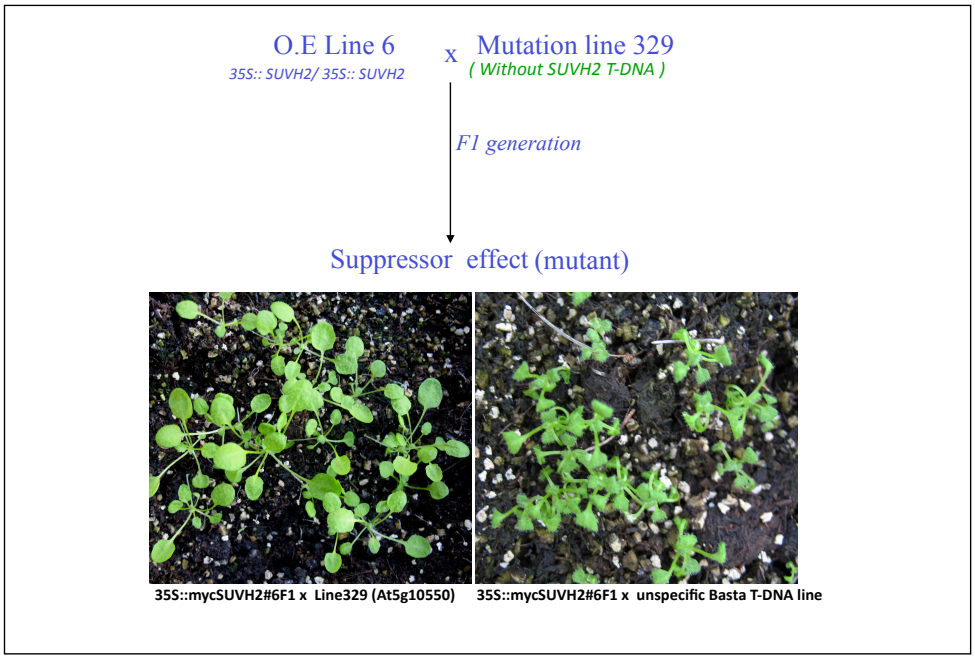
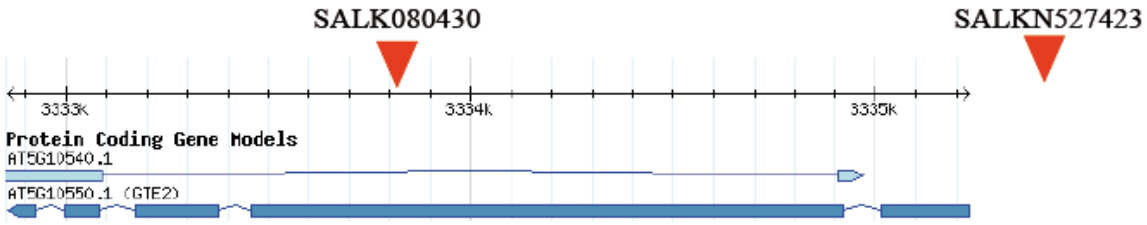


Figure 3.34: Reappearence of suppressor effect in SUVH2 overexpression line back crossed to line 329. Homozygous mutant locus was crossed to SUVH2 overexpression line. Suppressed progeny were obtained in F1 generation confirming the presence of dominant effect of the isolated mutant.

3.15.5 T-DNA inserted in At5g10550 gene is responsible for the dominant suppressor phenotype

To test whether the T-DNA insertion in the promoter of At5g10550 gene coding for a bromo domain containing protein was responsible for the dominant suppressor phenotype, two more insertion lines from the SALK collection were analyzed which have a predicted insertion in the promoter region of the At5g10550 locus. The T-DNA inserted in SALKN527423 line had a similar position in the promoter region, as in the original mutant identified from the genetic screen. The other T-DNA was inserted in the 2nd exon of the gene as shown in the Figure 3.35 (a). Both these insertional lines were molecularly analyzed. Homozygous plants of the SALK lines mutants were isolated and their insertion points were confirmed by PCR analysis. In order to reconfirm the reduction in transcript of the bromodomain-containing gene to the observed suppressor effect, the pollen from homozygous SALK line male parent was used to cross the homozygous overexpression SUVH2 female parent. As shown in the Figure 3.35 (b) all the F1 plants were suppressed, confirming the bromo domain-containing gene is a putative suppressor of SUVH2 overexpression phenotype.

a



b

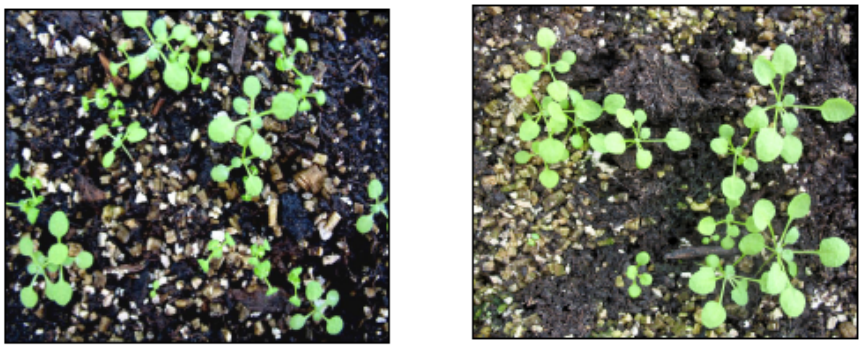


Figure 3.35: The T-DNA inserted in the promoter region of At5g10550 gene is responsible for the suppressor effect. Homozygous T-DNA line progeny were identified and crossed with the homozygous SUVH2 overexpression plants and scored for suppressor effect in next generation.

From the second screen only two putative suppressors were partly characterized with respect to suppressor activity.

3.16 Identification and characterization of line 34

3.16.1. Identification of the insertion locus in the mutant line 34

PCR amplification and sequencing of the left T-DNA insert boundary indicated that the insertion in line 34 occurred 650 bp upstream of the ATG codon in the 5'-transcribed region of a putative gene, At4g38825, located in the T9A14 contig of chromosome 4 (Figure 3.36). The points of insertion of the available SAIL T-DNA lines are also indicated in the Figure 3.36. Thus, the sequence analysis indicated that due to the insertion event, possibly a putative gene At4g38825 has been disturbed.

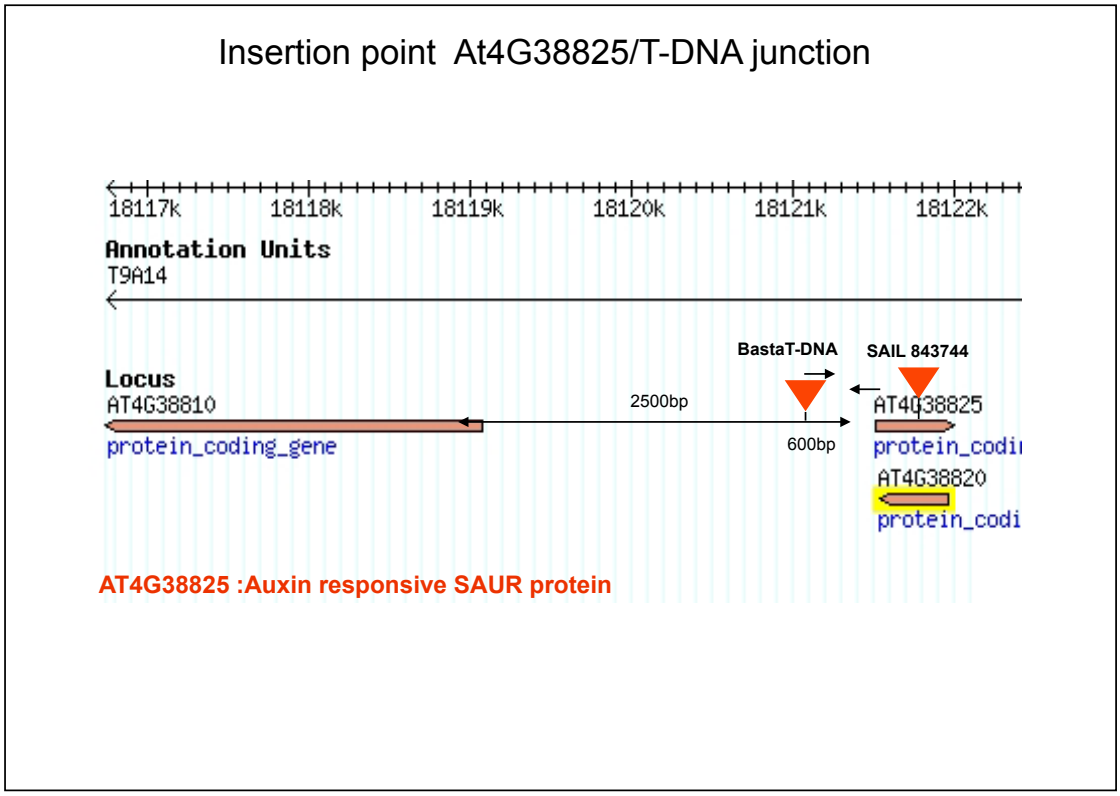


Figure 3.36: Genomic position of the insertion indicating the potential gene disrupted by the T-DNA insertion. The T-DNA was inserted on chromosome 4 on BAC clone T9A14. Auxin responsive SAUR protein was the potential candidate gene responsible for suppressor phenotype.

According to the TAIR gene prediction, the At4g38825 gene encodes a putative SAUR (Small Auxin Up Regulated) protein of 89 amino acids. This family of gene cluster that encodes a group of auxin-regulated RNAs. The promoter of SAUR proteins contains a conserved regulatory motif referred as auxin responsive element implicating in their regulation in response to auxin (Gil *et al*., 1994). They are regulated by auxin at the level of transcription. Proteins from this gene family have no identified functional role. Most of them are intronless indicating that they evolve rapidly either by gene duplication or by reverse transcription.

3.16.2 Southern blot analysis confirming the presence of single copy of mutater T-DNA insert

In T2 generation for line 34, plants showed a segregation pattern of 3:1 ratio, indicating insertion of a single T-DNA. Further confirmation of this result was done by Southern blot analysis using *bar* gene as a probe. Two different enzymes HindIII and BamHI were used for digestion as they have only one restriction site in the MCSI and MCSII in the mutater T-DNA. Detection of single band would therefore confirm that this line contained only one copy

of T-DNA. However Southern blot analysis showed more that one band indicating that the T-DNA might be present as tendem repeats at the site of insertion as shown in Figure 3.37. The exact arrangement of the tandem inserted repeats is required in order to get a clear picture of the architecture of the T-DNA. This knowledge is also important for designing primers for finding out the T-DNA/plant junctions and could be achieved by taking two different restriction enzymes from MCS I and MCS II for Southern blot analysis. Depending on the orientation of the fusion, different combinations are possible resulting in fragments of different size and number. Using Hind III restriction present in MCS I, only one fragment was obtained in Southern blot, indicating that it is a tandem fusion consisting of two T-DNA inserts, and the left borders are free. This hypothesis was confirmed by using BamHI restriction from MCS II, which as expected resulted in appearance of two fragments.

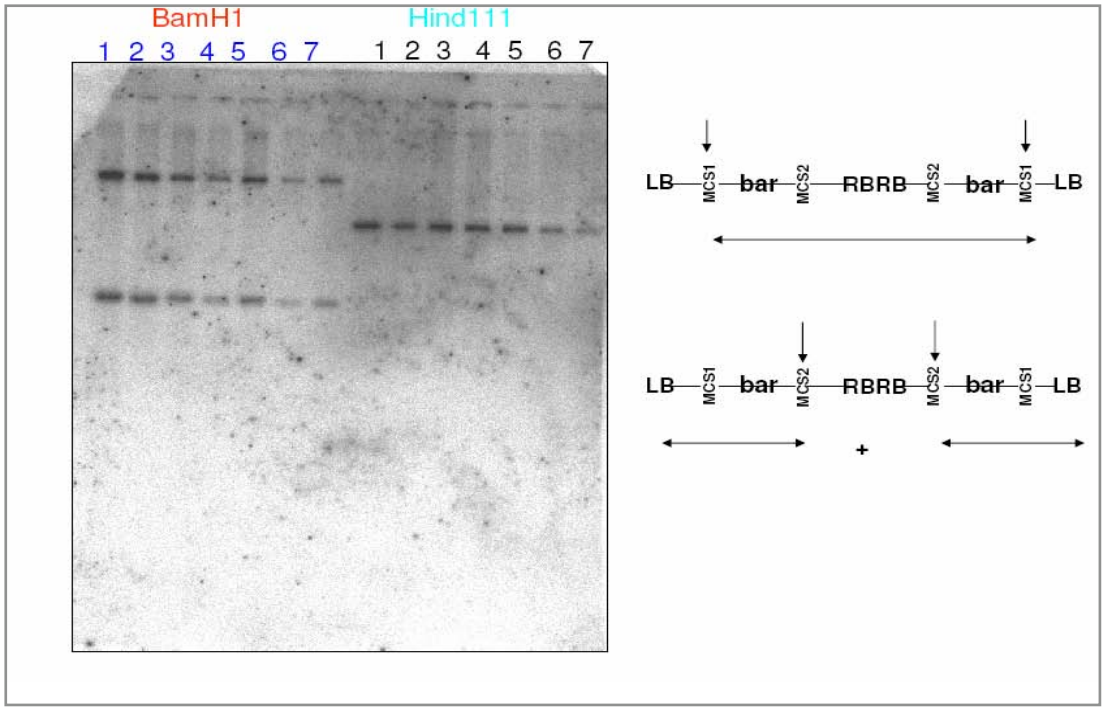


Figure 3.37: Southern blot analysis of mutant line 34 confirming the presence of single insert. Six to seven independent progeny from the line 34 were subjected to Southern blot analysis. The genomic DNA was digested with BamHI and HindIII and after transfer; the blot was hybridized with ^{32}P labelled *bar* coding sequence. Digestion with enzymes present in MCSI and MCSII , indicated that it is a tandem repeat with free left borders as shown .

3.16.3 Real-time expression and immunocytological analysis of SUVH2 in the mutant line 34

Relative to the starting over expression line, the values of SUVH2 overexpression in the presence of mutater T-DNA was measured by real-time PCR analysis. As shown in the Figure

3.38 (a), the relative values are higher then the posttranscriptional silent line but lower then the starting line. However, when the interphase nuclei were stained with H3K9dimethyl antibody, the ectopic H3K9 distribution characteristic of SUVH2 overexpression was also visible in the suppressor line 34, indicating that the expression levels were sufficient to score for suppressor phenotype [Figure 3.38 (b)].

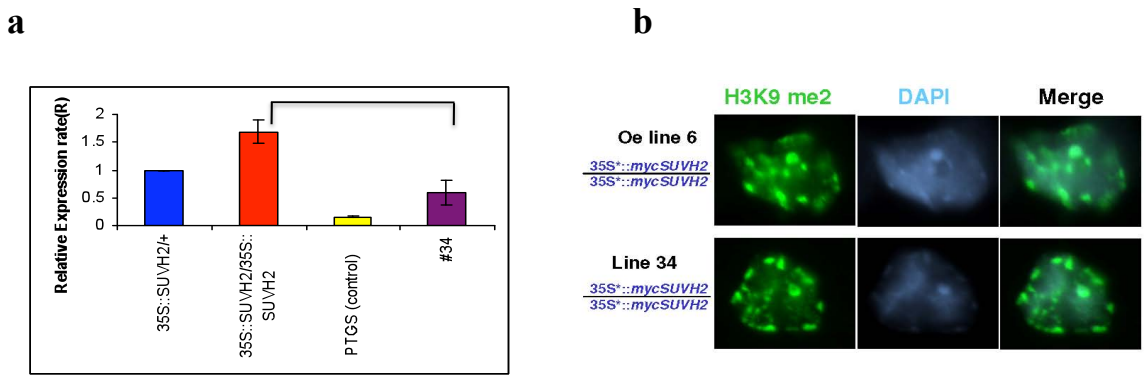


Figure 3.38: Variation in H3K9me2 distribution and expression analysis mutant line 34 in comparison to SUVH2 overexpression plants. a) Relative levels of expression were determined by real-time PCR using SYBR fluorescent dye. The graph represents expression levlels of SUVH2 in homozygous, PTGS (control) and mutant line 34 relative to the heterozygous overexpression line. Values were normalized to the expression of 18S rRNA. b) Representative images are shown indicating ectopic distribution of H3K9me2 mark in mutant line 34 as well as in SUVH2 overexpression plants.

3.17 Identification and characterization of line 11

3.17.1. Southern blot analysis confirming the presence of single T-DNA insertion in the mutant line 11

A segregation ratio of 3:1 in T2 generation of line 11 indicated a insertion of a single T-DNA. This was further confirmed by Southern blot analysis using *bar* gene as a probe. Two different enzymes HindIII and SpeI were used for digestion, as they have only one restriction site in the MCSI and MCSII of the mutater T-DNA. Detection of single band therefore was quite conclusive that this line contains only one copy of T-DNA.

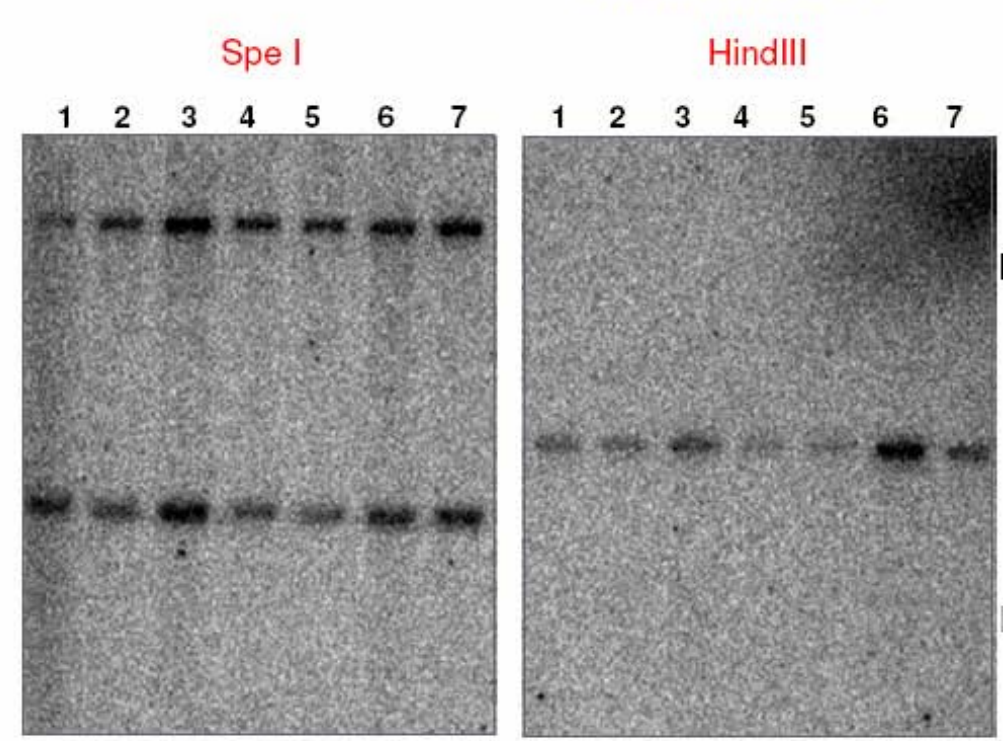
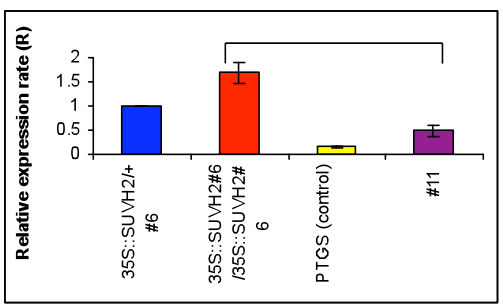


Figure 3.39: Southern blot analysis of mutant line 11 confirming the presence of single insert. Independent progeny from the line 11 were subjected to Southern blot analysis. The genomic DNA was digested with SpeI and HindIII and after transfer; the blot was hybridized with ^{32}P -labelled *bar* coding sequence. Digestion with enzymes present in MCS I and MCS II , indicated that it is a tandem repeat with left borders are free as shown in the Figure 3.39.

3.17.2 Real-time expression and Immunocytological analysis of SUVH2 in the mutant line 11

Relative to the starting over expression line, the values of SUVH2 overexpression in the presence of mutater T-DNA was measured by real-time PCR analysis. As shown in the Figure 3.40 (a), the relative values are higher then the posttranscriptional silent line but lower then the starting line. When the interphase nuclei were stained with H3K9me2 antibody, in suppressor line 11 the distribution was similar to the wild type where, the H3K9me2 distribution was restricted to the chromo centers [3.40 (b)].

a



b

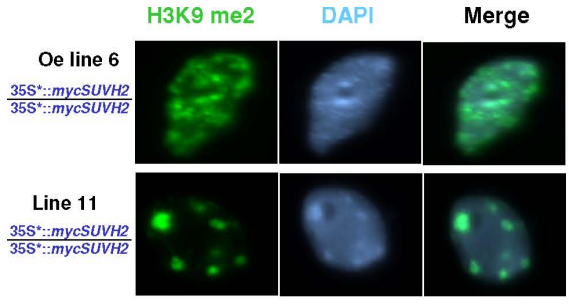


Figure 3.40: Variation in H3K9me2 distribution and expression analysis mutant line 11 in comparision to SUVH2 overexpression plants. a) Relative levels of expression were determined by real-time PCR using SYBR fluorescent dye. The graph represents expression levlels of SUVH2 in homozygous, PTGS (control) and mutant line 11 relative to the heterozygous overexpression line. Values were normalized to the expression of 18S rRNA. b) Representative images are shown indicating distribution of H3K9me2 mark in mutant line 11 as well as in SUVH2 overexpression plants.

A summary of all the mutations and partial characterization with respect to the insertion point, number of copy of insert, and the expression analysis was done as shown in the Table 3.3.

Sl.No	Mutation line	H3K9 staining	Distribution	Real Time	Southern	Insertion point
1	10	X		✓		
2	11	✓	Normal	✓		??
3	17	✓	Normal	✓	>2	>2
4	34	✓	Ectopic	✓	1	Confirm
5	42	✓	Normal	✓	>2	Confirm
6	45	✓	Normal	✓	>2	
7	47	X		✓		
8	54	✓	Normal	✓		
9	56	✓	Normal	✓		
10	61	X		✓		
11	72	X		✓		
12	73	X		✓		
13	122	✓	Ectopic	✓	1	Confirm
14	235	X		X	1	Confirm
15	326	✓	Ectopic	✓	>2	Confirm
16	329	X		X	1	Confirm

Table 3.3: Summary of results for the characterization of all the dominant suppressors of SUVH2 overexpression line found in this study. Isolated suppressor mutations were partially characterized using real-time PCR analysis with respect to expression of SUVH2 transgene and distribution of H3K9me2 staining pattern in the background of SUVH2 overexpression. Number of insertions was confirmed by Southern blot analysis and T-DNA/genomic junction was identified using Inverse PCR.

4. Discussion

The main aim of this work was to establish a screening procedure to genetically decipher the new components of the pathway with respect to SUVH2 mediated heterochromatin formation in *Arabidopsis*. The screen was based upon the already well-characterized histone methyltransferase SUVH2, overexpression of which leads to characteristic mini plant phenotype. A T-DNA mutagenesis approach was used, as this procedure is less tedious compared to classical EMS mutagenesis. Out of 10 SUVH proteins which are homologues of SU(VAR)3-9 protein of *Drosophila* , only SUVH2 has been shown to play a central role in heterochromatin formation by affecting all repressive marks in *Arabidopsis thaliana* (Naumann *et al*., 2005). During this study, a successful attempt has been made to prove the functionality of the test system, and to isolate and identify dominant suppressor mutations of SUVH2 overexpression involved in new chromatin functions in *Arabidopsis thaliana.* This would inturn provide us with insights to genetically decipher the components of the signaling pathway regulating heterochromatin formation in *Arabidopsis.*

From all the tested SUVH proteins only overexpression of SUVH2 leads to severe morphological defect. Under homozygous overexpression condition, plants remain small, and formed small, narrow, curled leaves. It also produced few viable seeds whereas in heterozygous background seeds were normal. Recently, using DNA microarray analysis, it was shown that TFL2 was responsible for enrichment of repressive H3K27me3 marks in about 15% of the target *Arabidopsis* genes (Kotake *et al*., 2003; Turck *et al*., 2007). Such a genome wide effect on heterochromatin formation can also be anticipated with histone methyltransferase SUVH2. As mutation of this protein coordinately reduces not only mono and dimethyl H3K9 but also mono and dimethyl H3K27 and mono methyl H4K20 and suppresses TGS. Overexpression of SUVH2 causes significant enhancement of these marks along with enhancement of transcriptional gene silencing (TGS) of tandem repeats of luciferase construct. Either SUVH2 alone can methylate histones at specific lysine residues in all these repressive marks or in mutant background of *suvh2*, activity of other HMTase is severely affected (Naumann *et al*., 2005). Loss or overexpression of SUVH2 affects both symmetric as well as non-symmetric DNA methylation.

In *Arabidopsis,* 10 genes that are Su(var)3-9 homologs were identified (Baumbusch *et al*., 2001). But only some of these SUVH proteins have been molecularly characterized. The classical heterochromatin silencing pathway used by Su(var)3-9 was also observed in *Arabidopsis*. SUVH4/kryptonite mediated histone H3K9me2 modification provides the

binding site for LHP1 protein. This in turn facilitate binding of the DNA methyltransferase CMT3, thus displaying interplay between histone methylation and DNA methylation (Jackson *et al*., 2002). CMT3 protein is unique to plant kingdom and responsible for CpNpG DNA methylation. This might reflect transcriptional fine tuning present only in the plants in response to different fluctuations in the environment It was also shown that SUVH6 works in concert with SUVH4 in maintaining H3K9me2 mark at transcribed inverted repeats whereas combination of SUVH4 and SUVH5 was shown to control transposon sequences (Ebbs *et al*., 2005; Ebbs and Bender, 2006).

4.1 Establishment of a new screening system for isolation of insertional dominant suppressors of SUVH2 over-expression phenotype

Different approaches have been used to isolate and genetically analyze suppressor mutations. All the previous studies were based upon the reactivation of the typically silent locus after mutagenesis, thus relating the loss of function of a particular gene to the reactivation of the silent locus (Bartee and Bender, 2001; Mittelsten Scheid *et al*., 1998). In the current study, an alternative approach was followed using T-DNA mutagenesis. The use of T-DNA integration has several advantages over conventional EMS mutagenesis. Compared to EMS mutagenesis, which leads to several point mutations in the genome, a low number of insertions per transformants are found in T-DNA mutagenesis. This significantly reduces the additional work required to remove second-site mutations. DNA integration results in relatively stable mutations as compared to transposon-based mutagenesis (Martienssen, 1998). The simple *Agrobacterium* transformation method for *Arabidopsis thaliana* allows high throughput production of T-DNA insertion mutants in this model organism. The identified candidate genes can easily be cloned for further characterization. The overexpression of methyltransferase SUVH2 leads to severe morphological defects and ectopic formation of heterochromatin (Naumann *et al*., 2005).

To further identify and analyze the other component of this pathway and to identify cross-talk points, a large number of *Arabidopsis* mutants needs to be identified using new screening strategies. The suppressor mutations might not only identify interacting protein leading to dissection of the pathway, but also alternative pathways that became activated by the suppressor mutations (Page and Grossniklaus, 2002). The identified components could represent known and unknown factors, which might affect enzymatic activity of SUVH2 or target specific binding of SUVH2. These components could also play an important role in assembly of SUVH2 complex during heterochromatin formation. The principle idea behind

the screen was that, if one of the components involved in heterochromatin formation is reduced in amount due to T-DNA insertion, then relatively less heterochromatin is formed. Accordingly SUVH2 overexpression phenotype is also suppressed. T-DNA mutational analysis could effectively facilitates the identification of important epigenetic functions required for heterochromatin formation. Therefore a novel genetic screen was designed by creating and screening for T-DNA insertions which suppress the mini plant phenotype in the background of SUVH2 overexpression and thus directly implicating the loss of morphological defects to the locus disturbed by the mutater T-DNA insertion.

A constitutive and strong 35S promoter was used to drive the expression of SUVH2 transgene in *Arabidopsis*. For selection of transgenic plants nptII gene under the control of nopaline synthase (nos) promoter was used. In a recent study, it was shown that introduction of another T-DNA driven by a similar strong promoter for insertional mutagenesis could lead to potential homology dependent silencing of the transgene. However such effects were not observed, when weak promoters were used to drive the expression of the transgene (Daxinger *et al*., 2007). In order to prevent trans-inactivation and 35S promoter-driven transgene silencing in the mutater T-DNA, the weak nopaline synthase (nos) promoter was used to derive the expression of *bar* gene. Bar gene codes for phosphinothricin acetyltransferse and provides resistence against the herbicide phosphinothricin (Basta). The integration of the introduced T-DNA during transformation occurs late in flower development after the lineage separation of male and female gametes and hence in T1 generation only heterozygous plants can be produced. Recently it was shown that female gametophyte is the target of T-DNA integration (Page and Grossniklaus, 2002). Since all the Basta sensitive plants die after emergence and only resistant progenies survive, it was confirmed that a single copy of *ba*r gene was sufficient to provide resistance against the concentration of basta spray used in the study.

4.1.1 T-DNA insertional mutagenesis in the heterozygous and homozygous SUVH2 overexpression line

In the first screen, heterozygous SUVH2 overexpression plants were mutagenized with mutater T-DNA construct, as homozygous plants were sterile and the primary transformant were selected based upon the herbicide resistence. However this process was cumbersome as it required extra effort to confirm the presence of SUVH2 transgene by PCR analysis in order to differentiate between dominant suppressors and wild type plants. By utilizing Post Transcriptional Gene Silencing (PTGS), the obtained homozygous SUVH2 overexpression

plants (PTGS-homo) were fertile and could therefore be transformed, resulting in identification of suppressors directly without subjecting them to further molecular analysis.

4.1.2 Identification of single locus plants by segregation ratio and Southern blot analysis

A relatively large number of putative dominant suppressor mutants were obtained in T1 generation. These putative mutants were selfed, and scored for 3:1 (resistant : sensitive) segregation ratio in succeeding generation to obtain single T-DNA locus (Rama Devi *et al*., 2006). However the mutation frequency appeared to be much higher then expected in T1 generation and therefore it was assumed that independent lines from the primary screen also contained a significant proportion of false positive. It could be possible that the suppressor effect is not due to the insertion of the mutater T-DNA but due to the stress environment that resulted in the loss of phenotype. This hypothesis was confirmed when the progeny of T1 plants were analyzed further after selfing them, most of them reverted back to the curly phenotype in the succeeding generation. However in lines containing multiple copies of Basta T-DNA inserts, the onset of Transcription Gene Silencing owning to methylation of the promoter sequence could also results in loss of phenotype. Therefore it was absolutely necessary to measure the expression of transgenic SUVH2 in all the identified mutations and to identify lines containing single copy of Basta T-DNA.

Plant transformation methods, such as *Agrobacterium* mediated transformation do not allow the introduction of a defined number of transgenes into the genome. Mendelian segregation ratio of 3.1 indicates the presence of a single locus whereas a ratio of 15:1 (resistant: sensitive) corresponds to the integration of T-DNA into two independent positions. Even higher ratio of resistant to sensitive plants indicates more then two independent inserts. Although the segregation ratio of 3:1 suggested a single locus for most of the suppressors, this does not necessarily mean the presence of single copy of T-DNA. During the process of transformation, one or multiple intact or rearranged gene copies can integrate at one or multiple unlinked loci, resulting in gene silencing (De Neve *et al*., 1997; Page and Grossniklaus, 2002). Southern blot analysis was done for the isolated line 10, 11, 34, 42, 45, 122, 326, and 329 but only the line 11, 34, 122, and 329 contained a single insert. Tandemly repeated transgenes at the same locus are often silenced in plants as a result of repeat-induced gene silencing (Assaad *et al*., 1993). Independent lines which showed a high resistant: sensitive ratio and complex ratio and a complex segregation pattern indicating the presence of multiple insertions were not analyzed further.

In order to find out whether the expression of the SUVH2 transgene was affected in the presence of mutater T-DNA, Southern blot analysis using methylation sensitive enzyme was carried out (Albini *et al*., 1990). Occurrence of higher molecular weight bands would possibly indicate that the SUVH2 transgene was methylated. However no such higher molecular weight bands were observed as compared to the control. In case of Line 32 and 122, additional lower molecular weight bands were seen but when Line 122 was stained with H3K9me2 antibody, it still showed ectopic distribution of H3K9me2 suggesting that the SUVH2 transgene had not undergone any major change in DNA methylation even after transformation with mutater T-DNA. Bisulfite analysis, which is sensitive to subtle changes in DNA methylation, should be performed in order to determine the degree of DNA methylation. This analysis was further supported by the fact that only plants with curly phenotype were obtained in genetic screen whenever the mutater T-DNA was inserted in the unspecific region of the genome, which had no role in SUVH2 mediated silencing.

4.2 Verification of the SUVH2 transgene in the background of isolated insertion mutations:

4.2.1 Immunocytological analysis of nuclear distribution of SUVH2 overexpression protein

All the mutations were stained with a repressive H3K9me2 mark against the background of SUVH2 overexpression line. On the basis of the staining pattern two classes of mutations could be differentiated. In the first class, mutations displayed the classical ectopic distribution of repressive H3K9me2 mark, similar to the starting SUVH2 overexpression control line. This class of mutations might represents defects in genes, which work in concert with SUVH2, without hindering the function of SUVH2 per se. In the second class of mutations, the staining of H3K9me2 mark reverted back to the wild type condition, where this repressive mark was restricted in chromo centers alone, indicating probably that the target sites are not recognized and hence this class of mutations might be involved in defining the role of SUVH2 in heterochromatin formation. However both types of mutations, identify Arabidopsis genes required for heterochromatin formation. Identification of these mutations would probably results in the convergence of the pathways triggered by distinct genes playing an important role in signal cascade leading to heterochromatin formation.

4.2.2 Quantitative expression analysis of SUVH2 transgene using real-Time PCR

The quantitative overexpression of the SUVH2 transgene in heterozygous, homozygous and homozygous plants obtained upon post-transcriptional gene silencing (PTGS-homo) was monitored by real-time RT-PCR analysis. The variability in expression was checked by quantifying and comparing the expression in many independent lines for each class. The values were quite consistent for each class of overexpression except for the PTGS-homo plants. As expected, the relative expression of the heterozygous line was nearly 50% of that in the homozygous line, thus confirming that the functionality of the measurement. The expression of the PTGS-homo line was very low as compared to original homozygous line and there was a considerable amount of fluctuation in the expression profile of this over expression line. However all the analysed plants of this class had strong curly leaf phenotype characteristic of SUVH2 overexpression indicating that transgene expression of SUVH2 was sufficient to show the mini plant phenotype. The relative expression of different isolated mutants was checked in comparison to the starting overexpression line and only those lines, which had sufficiently high levels of expression, were analyzed further.

The expression profiles are the relative values in comparison to the starting overexpression heterozygous line. However the absolute values are very high as the strong 35S CaMV promoter drives the expression of the transgene. The absolute value of SUVH2 transgene overexpression is 150 times that of endogenous SUVH2 levels. (Ay,N personal communication). Hence a slight decrease in overexpression values could still be useful for identifying dominant suppressors.

Mutants from the second screen were isolated using homozygous SUVH2 overexpression plants that have undergone posttranscriptional gene silencing, as these plants were fertile. However the relative expression values of the PTGS-homo plants in F1 generation, is much lower then the starting homozygote lines. With this background, the relative values of the mutants should be compared to those of the control starting line (PTGS-homo) from which they were derived, instead of the homozygous starting line to show a more accurate profile of the expression.

When a transgene integrates as multiple copies in one or few insertion sites, the dosage effect due to the presence of multiple T-DNA copies results in significant excess of transcript levels. When the level exceeds the transgene specific threshold, silencing is triggered. Therefore the variability in expression pattern is mainly due to the difference in onset of silencing in different plants. The variability in expression could be minimized by proper codon usage and

by making use of a weak promoter to drive the expression of the transgene (Schubert *et al*., 2004). Apart from the strength of the promoter, transcript stability also plays a crucial role in onset of silencing thus supporting the fact that accumulation of transcripts at high concentrations could enhance silencing to optimize the amount of transcript production (Que *et al*., 1997).

For real-time analysis, always three different plants were randomly chosen from each independent line. When a transgene is overexpressed under the control of a strong promoter, there is variability in its expression. Therefore it is also possible that all three assessed plants had low expression values compared to other transformant. This observation is also supported by another study where the expression of the transgene driven by 35S cauliflower mosaic virus promoter was assessed in wild type and posttranscriptional gene silencing (PTGS) mutant background. It was clearly demonstrated that, when compared to wild type, the incidence of highly expressing transformants shifted from from 20% to 100% in *sgs2* and *sgs3* mutant background. Therefore, the PTGS impaired *A.thaliana* mutants hold great promise for obtaining extremely high and uniform transgene expression (Butaye *et al*., 2004).

4.2.3 Confirmation of the screening protocol by crossing with randomly inserted SALK line plants

The cross of homozygous SUVH2 overexpression plants with randomly selected SALK line resulted in reappearance of curly leaf phenotype thus reinforcing our hypothesis that the dominant effect obtained in the screen was not an artifact. The phenotype was weaker when compared to the starting overexpression heterozygous line. This could partially be attributed to the presence of the 35S promoter driving the expression of randomly selected SALK line that has a tendency to undergo trans-inactivation. Appearance of curly plants in these crosses with overexpression line reconfirms our hypothesis that dominant suppressors could be identified, by insertion of second T-DNA into SUVH2 overexpression plants.

4.3 Isolation and characterization of insertional mutations suppressing the SUVH2 overexpression phenotype

In order to check, the functionality of the screening system, three mutants from 1st screen and two mutants from 2nd screen were further characterized.

4.3.1 Insertion of Basta T-DNA in the Line 122 results in complete disruption of the At3g56370

At3g56370 encodes a leucine-rich repeat transmembrane kinase protein. Plant receptor like kinase (RLKs) are one of the largest gene family in *Arabidopsis* and contain more then 600 members, thus accounting for 2.5% of the protein coding genes in *Arabidopsis*. Most of the well characterized RLKs belong to the group of serine/threonine kinase (Shiu and Bleecker, 2001). They participate in various functions in perception and processing of extracellular signals via cell surface receptors and can be grouped into 15 different subfamilies according to their divergent extracellular receptor domains (Shiu and Bleecker, 2001, 2003). This divergence allows them to respond to a wide range of external signals

For line 122, only one insertion site was found by inverse PCR, on chromosome 3 of BAC T5P19. This is in agreement with segregation ratio of 3:1 (Basta resistant: sensitive). This observation was further confirmed by Southern blot analysis in which only one band was visible. The insertion of mutater T-DNA resulted in complete disruption of the At3g56370 gene expression, as indicated by absence of any transcript in RT-PCR. Therefore the identified loss of function allele could be used for further characterization of this protein. In the current study, the kinase gene was only partially characterized. Structural analysis of the protein using a bioinformatic software (smart.embl-heidelberg.de) revealed the presence of a transmemebrane domain and a phosphotransferase domain responsible for serine/threonine kinase activity. These structural features hint at a possible role in signaling pathway. The other prominent feature is the presence of leucine rich repeats which were shown to play an important role in protein-protein interaction (Kobe and Deisenhofer, 1994). In the real-time PCR analysis, the expression of SUVH2 transgene as expected was relatively high in comparison to other isolated mutants. This was in accordance with the previous studies, which pointed out that the expression of a single copy of the transgene was high and stable. Further position effect had minimal influence on the variability of the transgene expression (De Buck *et al*., 2004; Nagaya *et al*., 2005; Schubert *et al*., 2004). In comparison to the starting overexpression line, the expression of SUVH2 transgene in the presence of mutater T-DNA was reduced. However this level of expression was sufficient to show an ectopic distribution of H3K9me2, characteristic of SUVH2 overexpression. Therefore this mutation was a good candidate for further analysis to figure out its potential role in the silencing process. It is likely that this protein is involved in signalling or phosphorylation of other nuclear proteins.

In order to confirm the effect of mutation in trans, crossing of the separated mutated loci with homozygous SUVH2 overexpression line was carried out. This overexpression phenotype was completely suppressed in F1 generation thus reconfirming our hypothesis, that this locus was responsible for the dominant suppressor effect.

4.3.2 Line 326 encodes a plant specific protein of unknown function

A new putative nuclear protein was identified with mutant line 326. However no conserved domain was observed making initial analysis and prediction about the function of this protein a bit more cumbersome then other isolated mutants. Segregation ratio of the progeny in T2 generation suggested the presence of more then two inserts in line 326, and this observation was confirmed by Southern blot analysis. The insertion point of the identified mutater T-DNA always co-segregated with the suppressor effect and Basta resistence. The T-DNA insertion in this line had disturbed the first exon of the gene At2g23690, likely creating a null allele. PSORT program (psort**.**ims.u-tokyo.ac.jp) clearly indicated the presence of two nuclear targeting signals indicating that it is a nuclear protein. Phylogenetic analysis showed that this gene is evolutionary conserved and restricted only to plant species. Taken together, these results show that this protein may have a plant specific role in the nucleus. Therefore this line was chosen for further analysis, inspite of having more then two T-DNA insertions. In the real-time RT-PCR analysis, the expression of transgenic SUVH2 had already reached a certain threshold, sufficient to show an ectopic distribution of H3K9me2, which is characteristic of SUVH2 overexpression. This indicates that enough protein was produced even in presence of multiple inserts, which are generally more prone to silencing. These results are in agreement with findings from previous studies, which suggested that single copy transformants do not always express the highest amount of the transgene. However the variability in transgene expression is relatively low in comparison to multi-copy transformants (De Buck *et al*., 2004).

By crossing the separated mutated loci with homozygous SUVH2 overexpression line, resulted in complete suppression of the overexpression line thus reconfirming our previous results of co-segregation analysis, that this locus in trans was responsible for the suppressor effect.

4.3.3 Characterization of a bromodomain containing protein encoding nuclear/ chromatin function

At5g10550 is predicted to encode bromodomain-containing protein and probably in conjugation with SUVH2 plays an important role in heterochromatin formation. Line 329 was chosen for further analysis, as a segregation ratio of 3:1 was obtained indicating the presence of single insert in T2 generation. The presence of a single copy of mutater T-DNA was further confirmed by presence of a single band in Southern blot analysis. The insertion point of the T-DNA indicated that the bromo domain containing protein At5g10550 present on chromosome 5 on BAC F12B17 could be potentially disturbed by the insertion. Bromo domains are generally found in proteins that regulate chromatin structures and gene expression by binding to acetylated lysine residues on histone 3 and 4 (Dyson *et al*., 2001; Eberharter and Becker, 2002; Zeng and Zhou, 2002). In real-time RT-PCR analysis, the expression of transgenic SUVH2 as expected for a single insert, was higher then the starting heterozygous overexpression line. However, the results of immunocytological studies indicated that the distribution of H3K9me2 was restricted only to the chromo centers similar to wild type. Thus it can be inferred that the mutation resulted in reverting the ectopic distribution of H3K9me2 characteristic of overexpression of SUVH2 to that of the wild type. The insertion point and the suppressor effect on the SUVH2 overexpression phenotype were validated by genetic crosses of the commercially available SALK line with homozygous SUVH2 overexpression line and also by crossing the plants with the suppressor locus containing only mutater T-DNA with the overexpression line. The progenies obtained in T1 generation were all suppressed, thus confirming the dominant effect.

In a previous study, a bromodomain-containing protein from tomato, VIRP1 showed a considerable homology to the protein under investigation not only in the bromodomain but also in the amino- and carboxy-terminal parts of the protein. The carboxy- terminal region of VIRP1 was shown to contain a RNA-binding domain, which was able to interact with RNA in in-vivo conditions (Martinez de Alba *et al*., 2003). Therefore, the presence of a nuclear targeting signal, a bromodomain and potential homology to RNA binding domain of VIRP1 strengthens our hypothesis that this protein might play a role in RNA directed DNA methylation.

4.3.4 The Basta T-DNA insertion site in Line 34 lies in Auxin responsive SAUR protein

The presence of a single T-DNA insert was confirmed by Southern blot analysis for line 34 after obtaining the segregation ratio of 3.1. The insert was located on chromosome 4 of BAC T9A14 and had potentially disturbed the expression of At4g38825 coding for auxin

responsive SAUR protein. These proteins are represented by a large multigene family and are characterized by the presence of short half-lives of 10-50 min. They are regulated by auxin at the level of transcription based on rapid mRNA turnover. This feature makes them a key player in transient regulation, as they can be repressed and induced more rapidly then stable transcripts. The instability of these proteins is mainly due to the presence of a downstream element (DST) found in the 3′ untranslated region (Johnson *et al*., 2000). Auxins are important class of plant harmone essential for coordinating many growth and behavoiural processes in plant life cycle. They can directly stimulate the expression of specific genes by promoting degradation of transcriptional repressor protein. Auxins are essential for cell growth, affecting both cell division and cell elongation during the growth and development of a plant.

Further confirmation of the suppressor effect of this mutant is in progress by crossing the T-DNA homozygotes insertion of this gene obtained from commercially available SAIL line (CS843744) with overexpression line and scoring for dominant suppressor effect in next generation. As pointed out earlier, screening for suppressor mutations could also identify mutations which are components of various other signalling cascades. SUVH2 protein is involved in diverse biological processes as indicated by one of its prominent role in senescence. WRKY53 is a key regulator of leaf senescence in *Arabidopisis.* Overexpression of SUVH2 leads to delay in senescence by inhibiting WRKY53 and senescence associated genes (Ay *et al*., 2009).

Similar to other isolated mutant, the expression of SUVH2 transgene was lower then that of starting overexpression line. However this value was sufficient to show ectopic distribution of H3K9me2, similar to that of the starting overexpression line. Though line 11 and 34 had similar integration pattern as indicated by Southern blot analysis, they differ in their staining pattern. While in line 34, presence of an ectopic distribution was quite evident; in line 11 the staining pattern was similar to that of wild type where H3K9me2 mark was restricted only to the chromo centers. Even though this line was a single tandem insert, the genomic sequences adjacent to the left border could not be amplified after several repeated attempts. One possible reason could be insertion of the mutater T-DNA in the repetitive heterochromatic region of the genome which are difficult to be amplified by I-PCR.

Recently it was also shown that the SRA domain of SUVH2 has a methyl cytosine binding affinity, and preferentially binds to methylated CG residues (Johnson *et al*., 2008). In accordance with these results, SUVH2 was shown to interact with MET1 and DDM1. MET1 is responsible for CG methylation and DDM1, a chromatin-remodeling factor. Mutant

background of DDM1 leads to 70 % loss in DNA methylation (Jeddeloh *et al*., 1999; Teixeira *et al*., 2009). However the obtained suppressor effect on SUVH2 overexpression was recessive in the mutant background of met1 and ddm1, and hda6 allele had no such suppressor effect. In order gain more knowledge of the interacting partners of SUVH2, Transcriptional Gene Silencing (TGS) mutants isolated in the lab by Ingo Hofmann were crossed to homozygous SUVH2 overexpression line, and experiments are underway to score for suppressor effect in F2 generation.

.

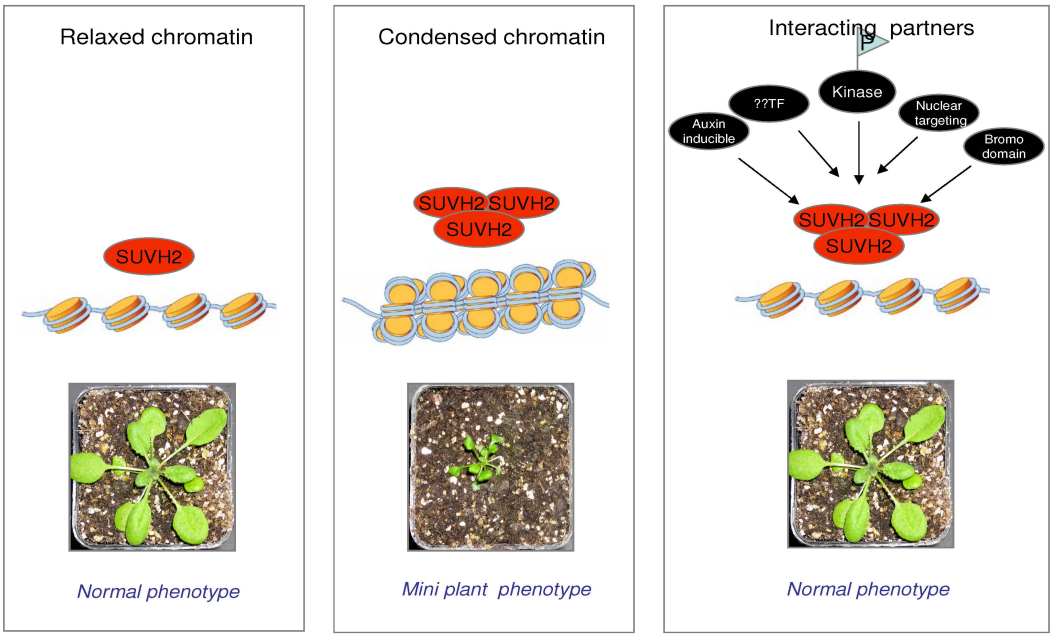


Figure 4.1 Model for the identification of novel suppressors based upon the SUVH2 overexpression phenotype. In the presence of endogenous histone methyl transferase SUVH2, the nucleosomes are relaxed and lead to a wild type phenotype. When SUVH2 is overexpressed with a strong promoter, chromatin becomes condensed resulting in a local heterochromatin environment. This is manifested by severe morphological defects leading to the typical mini plant phenotype. T-DNA mutagenesis in the background of overexpression of SUVH2 plants, leads to disruption of the interaction between SUVH2 and the unknown suppressors, thus reverting the phenotype and nucleosome status back to the wild type level.

Most of the identified loci in this study were novel, not having been identified from the other genetic screens. This might be because the number of identified loci is relatively small or alternatively because the genetic screen is different from former ones and therefore uncovers new classes of mutants. The identified suppressor mutations might not only identify interacting proteins involved in heterochromatin formation but also cross talk points in the various signaling networks initiated by the suppressor mutations. Therefore identified mutations in individual lines might influence several independent signal transduction pathways.

Analysis of all the isolated suppressors revealed that none of the suppressors were essential for survival, as all the homozygous insertion mutants displayed wild type phenotype under normal growth conditions. The other possible reason for wild type phenotype could be the redundancy of the genes as in *Arabidopsis.* Most of the genes are part of gene clusters and hence require all the genes present in the cluster to be disrupted, in order to obtain a phenotype (Krysan *et al*., 1999).

4.4 Outlook

Genetic screen like any other screen has its own limitations. Essential genes whose loss of function mutations leads to lethality cannot be isolated by this method. The presence of multiple inserts and homology between mutater T-DNA and SUVH2 transgene could possibly reduce the expression of the SUVH2 transgene and thus lead to identification of the false positive mutants. This problem could be resolved to a large extent by selecting for single inserts and optimizing the *Agrobacterium* mediated transformation to obtain selectively single inserts. Further variation in the transgene expression driven by a strong promoter in single inserts could be minimized by refining the screening procedure in the background of PTGS impaired mutants.

Similar type of genetic screen using T-DNA mutagenesis to isolate insertional mutations were successfully carried out, utilizing two T-DNA construct (Rama Devi *et al*., 2006) which led to identification of novel loci controlling the expression of stress responsive genes. The results of this screen could also validate our hypothesis of isolating dominant suppressor mutations using T-DNA mutagenesis. Further characterization of the isolated mutations, with respect to their functional role in signaling pathway could provide valuable information about the specific role of the identified genes in structure and dynamics of chromatin regulation.

Appendix 1. List of primers

Primers used for identifying transgenic plants containing SUVH2 and Basta T-DNA
Primers used for inverse PCR

Name	Sequence
MYCSAL	ATAGTCGACATGGGCGGACGCGAACAAAAGTTG
SALK_574R	AAGTACATGATTCTTCATACTCTC
BARF1	GGATCCTCTAGGGGTCATCAGA
BARB596	GATCTCCGACTCTAGGGGGATCT

Primers used for inverse PCR

Name	Sequence
LBLEFT1	AAGTTGTCTAAGCGTCAATTTGTTT
LBLEFT2	ATACAGGCAGCCCATCAGTC
LBRIGHT1	TTTAATGTACTGGGGTGGTTTTG
PCBLBFOR34	CTCGGCACAAAATCACCACTCG
PCBLBREV323	GCTATCTGGACAAGGGAAAACGC

Primers used for genotyping the insertion point of Basta T-DNA in identified putative mutants

Name	Sequence
F27L4 FOR	CAATAAGTTTTCCCACTGTGC
F27L4 REV	CCCAAACCACAGTCCCTAAA
FI7I23NEWF	TAAACGGAGATGTAGTCAAATAAGG
FI7I23NEWR	TATGATCAATCGGTAAAACGATATT
T5P19FOR2	AACGGAGGAAGCAAAACAAA
T5P19REV1	TTCAACGCGATTTCAATCTG
F12B17FOR2	TGCTAGGTCAGCAGTTTCTTGA
F12B17REV1	TCGGAACTTCTTCCACAACA
T9A14FOR139	CAAAAGGCGACTATGCCATC
T9A14REV1030	T TCACATGGCTCTGTCTTGC

Primers used for RT PCR analysis

Name	Sequence
SUVH2HEL6FOR	AATTAATATAATTAGTTGGGACGTG
SUVH2HEL6REV	CTGAGAAATCTGTATGAGTAGTTCA
AT3G56370FOR1	CTGAGCTGCAACATTCTTCG
AT3G56370REV1	CCCACATTGCCTGAAAAACT
AT5G10550RT2FOR1861	CGGTGACAGAAATGGGAAGT
AT5G10550RT2REV2179	CTATCCGCATCCGAATCACT

Primers used for cloning experiments

Name	Sequence
AT3G56370SPEIFOR1	ACTAGTGTCCGGCTCACAC
AT3G563705572REV	GTCCCATGCACATGGTTACA
AT2G23690BAMHFOR1	GGATCCTAATGGCATTGGGTTG
AT2G23690REV	CGGCGCAGAGAGATAAAGAG

Primers used for real-time RT PCR analysis

Name	Sequence
18SF	CTGCCCGTTGCTCTGATGAATCATG
18SR	CAATTAAGACCAGGAGCGTATCG
ACTINF250	ATGGAAAAGATATGGCATCACAC
ACTINR448	TACGACCACTGGCATAGAGAG
SUVH2CTERMINIFOR1	GGTCTATCCTGGTCGGTTCAC
SUVH2CTERMINIREV1	GTAGCAAGCCACGTTCCTCATC

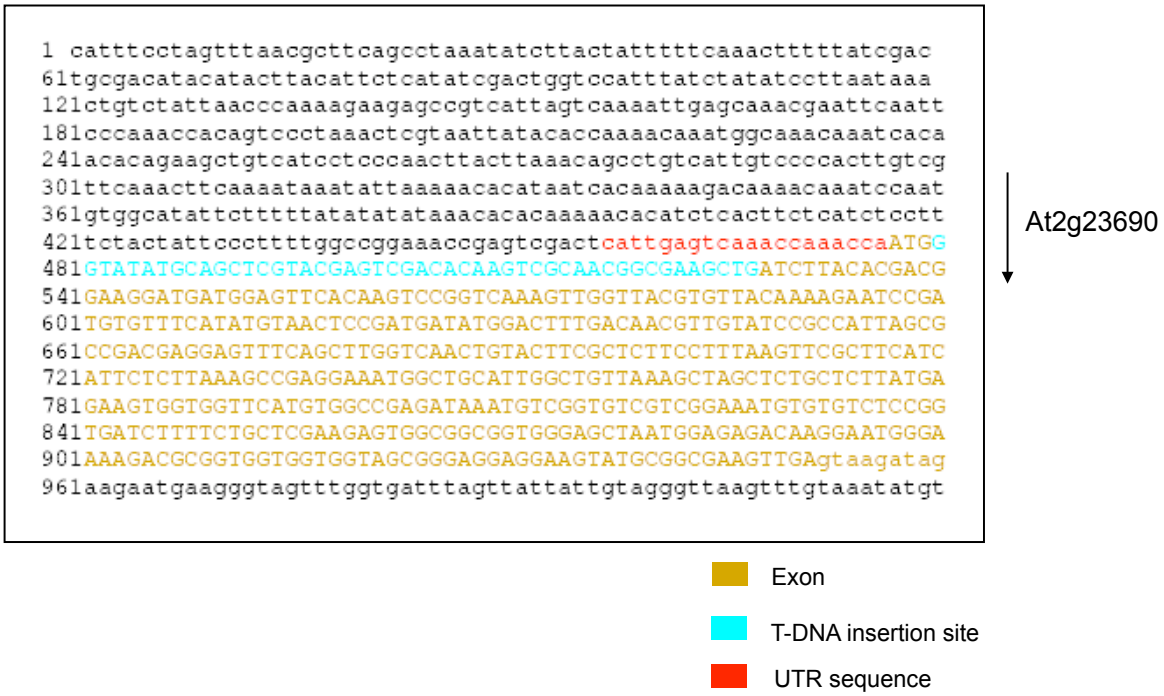
Primers used for genotyping the SALK/SAIL line for the isolated putative mutants

Name	Sequence
SALK121940F	CAACCGCATTTCAATCTGTAGG
SALK121940R	ACGAATCCGACTTGGCTGAC
SALKLBA1RC	TGGTTCACGTAGTGGCGCCATC
SALK048198CFOR	GACACATGGTCCTTTCTTATCC
SALK048198CREV	TGCTCTTATGAGAAGTGGTGG
SALK079802FOR4	TAGTTGCGGAAGACCTCCAT
SALK079802FOR1164	TTTCCAGGACACCATCATCA
SALK080430FOR31	TGGTGAAGCTGATGAAGCAT
SALK080430REV1031	ACCGGAGGCTGATGATAATG
AT3G27200FOR289	CGAACCTGAACCTGAAGGAG
AT3G27200REV1232	TCAAGAGGAAACACCGTTTC
AT5G24350FOR6	CGAATGCAAGTTCCGAGAGT
AT5G24350REV859	ATTGATGGTTTCCGGTGGAA
SAIL843774FOR74	TTCGACACCAAAAGGGTTTC
SAIL84374REV1075	CCTCTTTCTGAACCGGACAC
SALK027423FOR78	TCTGTTGTTGTGTTTGGGGA
SALK027423REV1195	AAAAACAAGGTGAGGCATCG
SALK106462FOR73	CGCATCGATAAACAGAAGCA
SALK106462REV1194	TCCGCACTCATTACCCTACC
SAILLB1	GCCTTTTCAGAAATGGATAAATAGCCTTGCTTCC
SALK069675CFOR	TTTGGATTATTGGTGTGCG
SALK069675CREV	ACAACACAGAAGCTGTCATCC

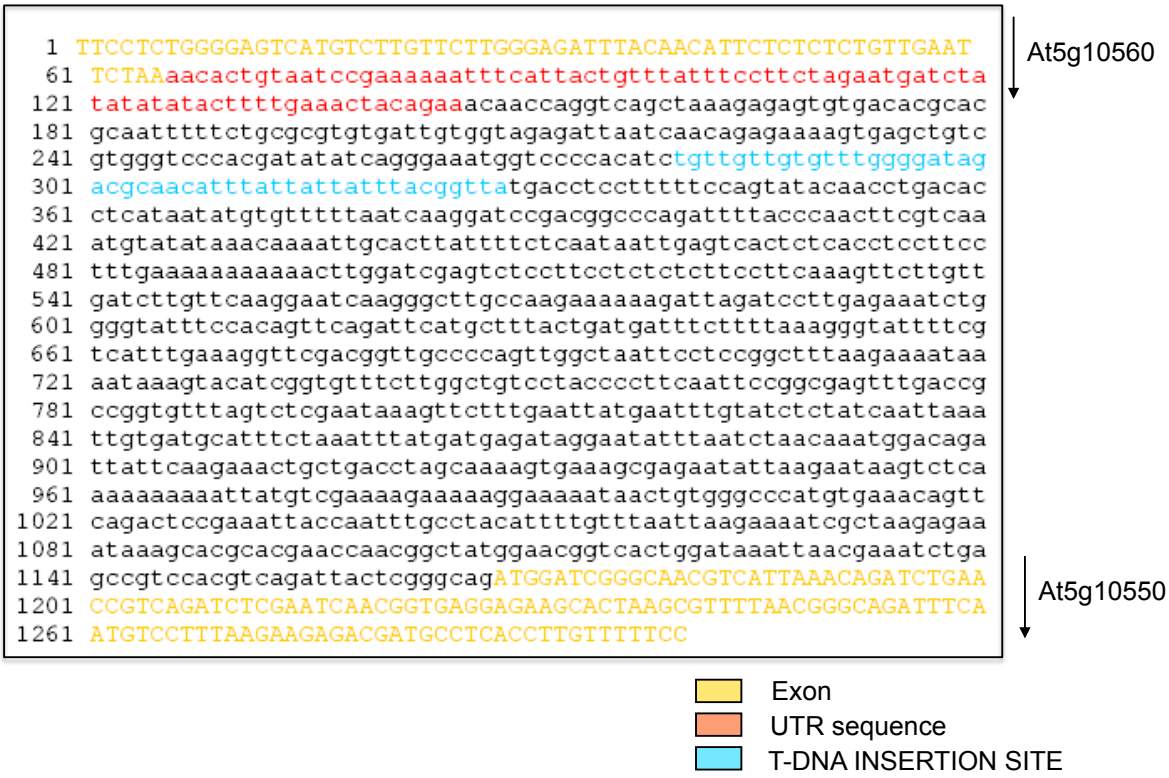
Appendix 2. Insertion sequence of Basta T-DNA in different identified putative mutants

Sequence of the At3G56370/T-DNA junction

Sequence of the At2g23690/T-DNA junction



Sequence of the At5g10550/T-DNA junction



Sequence of the At4G38825/T-DNA junction

```
   1 ccaacttttcttatacgggactggaaatggtttacgcgggccgggatccatacaaaaaat
  61 gttataagtttatattgtatcactgagaattgggcttcgttctgataaattttgaattaa
 121 ggcccaatcacacaagtagaaacgcgctatcggcgataagctgtcgatctcgcaattttc
 181 cccattcgcttatttcttcttcgaactacagaggggatccttgaacgccatgcagttttt
 241 ctgcgaaagagagaaacatcgtgttatttagttggttgcaactagtaacgggatcgctaa
 301 tcccataaagcgccacacagttaaggtgcagagagaagaacgacgtgttgttttaagcgt
 361 gaaaaaacggcgggccctctgtcttttctatttctctcttctcaacacttgcaggacagt
 421 gccctcacggataatatggagggttagtcagatggataatggattatcaattatgtgtag
 481 ccacttggctacttgactcagtgacttagacttttcaggtgctgtgtaggtatataacat
 541 aaaggtcccacgtacatcatctacataccattccaaagagatgtatatgtgccatcgaca
 601 atttcatgtgaagtgcgctcgaagttgttgttagtcttctagtgtcatgaattcttttca
 661 gatataatccatattcaattatcagtttataacatgcaagacagagccatgtgaattctt
 721 cttatgcatgtcccactcttgtataagaagcccacggtttctatcacatgaacgaaattg
 781 gttgtatatatatattgcattgttttcatatcgtgagcataacaacttcttcatctaatc
 841 tgataagctaaaggagacttctaagctgaatctttttccaggatccatcagagcagaaca
 901 aaaaaATGGGAGTGTTCCGAGGTCTTATGGGCGCAAAGAAGATCTTTCAAGGTCGATCAA
 961 TGGCAGCTTCGACACCAAAAGGGTTTCTTGCAGTGTACGTTGGAGAAAGCCAGATGAAAA
1021 GATACATAGTGCCAGTCTCATATTTGAACCAGCCTTCATTTCAAGCTTTGTTGAGTAAAT
1081 CTGAACAAGAGTTTGGGTTT
```

At4g38825

T-DNA insertion site

UTR sequence

Exon

References

Agius, F., Kapoor, A., and Zhu, J.K. (2006). Role of the Arabidopsis DNA glycosylase/lyase ROS1 in active DNA demethylation. Proc Natl Acad Sci U S A 103, 11796-11801.

Albini, A., Toffenetti, J., Zhu, Z., Chader, G.J., and Noonan, D.M. (1990). Hypomethylation of the interphotoreceptor retinoid-binding protein (IRBP) promotor and first exon is linked to expression of the gene. Nucleic Acids Res 18, 5181-5187.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25, 3389-3402.

Alvarez-Venegas, R., Pien, S., Sadder, M., Witmer, X., Grossniklaus, U., and Avramova, Z. (2003). ATX-1, an Arabidopsis homolog of trithorax, activates flower homeotic genes. Curr Biol 13, 627-637.

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.

Arnaud, P., and Feil, R. (2006). MEDEA takes control of its own imprinting. Cell 124, 468-470.

Assaad, F.F., Tucker, K.L., and Signer, E.R. (1993). Epigenetic repeat-induced gene silencing (RIGS) in Arabidopsis. Plant Mol Biol 22, 1067-1085.

Ay, N., Irmler, K., Fischer, A., Uhlemann, R., Reuter, G., and Humbeck, K. (2009). Epigenetic programming via histone methylation at WRKY53 controls leaf senescence in Arabidopsis thaliana. Plant J 58, 333-346.

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

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

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

Baumbusch, L.O., Thorstensen, T., Krauss, V., Fischer, A., Naumann, K., Assalkhou, R., Schulz, I., Reuter, G., and Aalen, R.B. (2001). The Arabidopsis thaliana genome contains at least 29 active genes encoding SET domain proteins that can be assigned to four evolutionarily conserved classes. Nucleic Acids Res 29, 4319-4333.

Bedford, M.T., and Richard, S. (2005). Arginine methylation an emerging regulator of protein function. Mol Cell 18, 263-272.

Brandstadter, J., Rossbach, C., and Theres, K. (1994). The pattern of histone H4 expression in the tomato shoot apex changes during development. Planta 192, 69-74.

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

Butaye, K.M., Goderis, I.J., Wouters, P.F., Pues, J.M., Delaure, S.L., Broekaert, W.F., Depicker, A., Cammue, B.P., and De Bolle, M.F. (2004). Stable high-level transgene expression in Arabidopsis thaliana using gene silencing mutants and matrix attachment regions. Plant J 39, 440-449.

Byrd, K.N., and Shearn, A. (2003). ASH1, a Drosophila trithorax group protein, is required for methylation of lysine 4 residues on histone H3. Proc Natl Acad Sci U S A 100, 11535-11540.

Cao, X., Aufsatz, W., Zilberman, D., Mette, M.F., Huang, M.S., Matzke, M., and Jacobsen, S.E. (2003). Role of the DRM and CMT3 methyltransferases in RNA-directed DNA methylation. Curr Biol 13, 2212-2217.

Chan, S.W., Henderson, I.R., Zhang, X., Shah, G., Chien, J.S., and Jacobsen, S.E. (2006). RNAi, DRD1, and histone methylation actively target developmentally important non-CG DNA methylation in arabidopsis. PLoS Genet 2, e83.

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

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

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

Davies, G.J., Sheikh, M.A., Ratcliffe, O.J., Coupland, G., and Furner, I.J. (1997). Genetics of homology-dependent gene silencing in Arabidopsis; a role for methylation. Plant J 12, 791-804.

Daxinger, L., Hunter, B., Sheikh, M., Jauvion, V., Gasciolli, V., Vaucheret, H., Matzke, M., and Furner, I. (2008). Unexpected silencing effects from T-DNA tags in Arabidopsis. Trends Plant Sci 13, 4-6.

De Buck, S., Windels, P., De Loose, M., and Depicker, A. (2004). Single-copy T-DNAs integrated at different positions in the Arabidopsis genome display uniform and comparable beta-glucuronidase accumulation levels. Cell Mol Life Sci 61, 2632-2645.

De Neve, M., De Buck, S., Jacobs, A., Van Montagu, M., and Depicker, A. (1997). T-DNA integration patterns in co-transformed plant cells suggest that T-DNA repeats originate from

co-integration of separate T-DNAs. Plant J 11, 15-29.

De Wilde, C., Van Houdt, H., De Buck, S., Angenon, G., De Jaeger, G., and Depicker, A. (2000). Plants as bioreactors for protein production: avoiding the problem of transgene silencing. Plant Mol Biol 43, 347-359.

Dyson, M.H., Rose, S., and Mahadevan, L.C. (2001). Acetyllysine-binding and function of bromodomain-containing proteins in chromatin. Front Biosci 6, D853-865.

Ebbs, M.L., Bartee, L., and Bender, J. (2005). H3 lysine 9 methylation is maintained on a transcribed inverted repeat by combined action of SUVH6 and SUVH4 methyltransferases. Mol Cell Biol 25, 10507-10515.

Ebbs, M.L., and Bender, J. (2006). Locus-specific control of DNA methylation by the Arabidopsis SUVH5 histone methyltransferase. Plant Cell 18, 1166-1176.

Eberharter, A., and Becker, P.B. (2002). Histone acetylation: a switch between repressive and permissive chromatin. Second in review series on chromatin dynamics. EMBO Rep 3, 224-229.

Ebert, A., Lein, S., Schotta, G., and Reuter, G. (2006). Histone modification and the control of heterochromatic gene silencing in Drosophila. Chromosome Res 14, 377-392.

Ebert, A., Schotta, G., Lein, S., Kubicek, S., Krauss, V., Jenuwein, T., and Reuter, G. (2004). Su(var) genes regulate the balance between euchromatin and heterochromatin in Drosophila. Genes Dev 18, 2973-2983.

Fischer, A., Hofmann, I., Naumann, K., and Reuter, G. (2006). Heterochromatin proteins and the control of heterochromatic gene silencing in Arabidopsis. J Plant Physiol 163, 358-368.

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

Fuks, F., Hurd, P.J., Deplus, R., and Kouzarides, T. (2003). The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. Nucleic Acids Res 31, 2305-2312.

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.

Gil, P., Liu, Y., Orbovic, V., Verkamp, E., Poff, K.L., and Green, P.J. (1994). Characterization of the auxin-inducible SAUR-AC1 gene for use as a molecular genetic tool in Arabidopsis. Plant Physiol 104, 777-784.

He, X.J., Hsu, Y.F., Zhu, S., Liu, H.L., Pontes, O., Zhu, J., Cui, X., Wang, C.S., and Zhu, J.K. (2009). A conserved transcriptional regulator is required for RNA-directed DNA methylation and plant development. Genes Dev 23, 2717-2722.

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

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.

Jackson, J.P., Lindroth, A.M., Cao, X., and Jacobsen, S.E. (2002). Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. Nature 416, 556-560.

Jeddeloh, J.A., Stokes, T.L., and Richards, E.J. (1999). Maintenance of genomic methylation requires a SWI2/SNF2-like protein. Nat Genet 22, 94-97.

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

Johnson, L.M., Law, J.A., Khattar, A., Henderson, I.R., and Jacobsen, S.E. (2008). SRA-domain proteins required for DRM2-mediated de novo DNA methylation. PLoS Genet 4, e1000280.

Johnson, M.A., Perez-Amador, M.A., Lidder, P., and Green, P.J. (2000). Mutants of Arabidopsis defective in a sequence-specific mRNA degradation pathway. Proc Natl Acad Sci U S A 97, 13991-13996.

Kakutani, T., Jeddeloh, J.A., Flowers, S.K., Munakata, K., and Richards, E.J. (1996). Developmental abnormalities and epimutations associated with DNA hypomethylation mutations. Proc Natl Acad Sci U S A 93, 12406-12411.

Kanno, T., Mette, M.F., Kreil, D.P., Aufsatz, W., Matzke, M., and Matzke, A.J. (2004). Involvement of putative SNF2 chromatin remodeling protein DRD1 in RNA-directed DNA methylation. Curr Biol 14, 801-805.

Kawabe, A., Matsunaga, S., Nakagawa, K., Kurihara, D., Yoneda, A., Hasezawa, S., Uchiyama, S., and Fukui, K. (2005). Characterization of plant Aurora kinases during mitosis. Plant Mol Biol 58, 1-13.

Kobe, B., and Deisenhofer, J. (1994). The leucine-rich repeat: a versatile binding motif. Trends Biochem Sci 19, 415-421.

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

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

Kouzminova, E., and Selker, E.U. (2001). dim-2 encodes a DNA methyltransferase responsible for all known cytosine methylation in Neurospora. EMBO J 20, 4309-4323.
Krysan, P.J., Young, J.C., and Sussman, M.R. (1999). T-DNA as an insertional mutagen in Arabidopsis. Plant Cell 11, 2283-2290.

Kurihara, D., Kawabe, A., Matsunaga, S., Nakagawa, K., Fujimoto, S., Uchiyama, S., and

Fukui, K. (2007). Characterization of a splicing variant of plant Aurora kinase. Plant Cell Physiol 48, 369-374.

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

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

Law, J.A., and Jacobsen, S.E. (2009). Molecular biology. Dynamic DNA methylation. Science 323, 1568-1569.

Liu, F., Marquardt, S., Lister, C., Swiezewski, S., and Dean, C. (2010). Targeted 3' processing of antisense transcripts triggers Arabidopsis FLC chromatin silencing. Science 327, 94-97.

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

March-Diaz, R., and Reyes, J.C. (2009). The beauty of being a variant: H2A.Z and the SWR1 complex in plants. Mol Plant 2, 565-577.

Margot, J.B., Ehrenhofer-Murray, A.E., and Leonhardt, H. (2003). Interactions within the mammalian DNA methyltransferase family. BMC Mol Biol 4, 7.

Martinez de Alba, A.E., Sagesser, R., Tabler, M., and Tsagris, M. (2003). A bromodomain-containing protein from tomato specifically binds potato spindle tuber viroid RNA in vitro and in vivo. J Virol 77, 9685-9694.

Mathieu, O., Reinders, J., Caikovski, M., Smathajitt, C., and Paszkowski, J. (2007). Transgenerational stability of the Arabidopsis epigenome is coordinated by CG methylation. Cell 130, 851-862.

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

Mittelsten Scheid, O., Afsar, K., and Paszkowski, J. (1998). Release of epigenetic gene silencing by trans-acting mutations in Arabidopsis. Proc Natl Acad Sci U S A 95, 632-637.

Miura, A., Nakamura, M., Inagaki, S., Kobayashi, A., Saze, H., and Kakutani, T. (2009). An Arabidopsis jmjC domain protein protects transcribed genes from DNA methylation at CHG sites. EMBO J 28, 1078-1086.

Morales-Ruiz, T., Ortega-Galisteo, A.P., Ponferrada-Marin, M.I., Martinez-Macias, M.I., Ariza, R.R., and Roldan-Arjona, T. (2006). DEMETER and REPRESSOR OF SILENCING 1 encode 5-methylcytosine DNA glycosylases. Proc Natl Acad Sci U S A 103, 6853-6858.

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

Nagaya, S., Kato, K., Ninomiya, Y., Horie, R., Sekine, M., Yoshida, K., and Shinmyo, A. (2005). Expression of randomly integrated single complete copy transgenes does not vary in Arabidopsis thaliana. Plant Cell Physiol 46, 438-444.

Nam, K.H., and Li, J. (2002). BRI1/BAK1, a receptor kinase pair mediating brassinosteroid signaling. Cell 110, 203-212.

Naumann, K., Fischer, A., Hofmann, I., Krauss, V., Phalke, S., Irmler, K., Hause, G., Aurich, A.C., Dorn, R., Jenuwein, T., *et al*. (2005). Pivotal role of AtSUVH2 in heterochromatic histone methylation and gene silencing in Arabidopsis. EMBO J 24, 1418-1429.

Ng, D.W., Wang, T., Chandrasekharan, M.B., Aramayo, R., Kertbundit, S., and Hall, T.C. (2007). Plant SET domain-containing proteins: structure, function and regulation. Biochim Biophys Acta 1769, 316-329.

Ng, R.K., and Gurdon, J.B. (2008). Epigenetic memory of an active gene state depends on histone H3.3 incorporation into chromatin in the absence of transcription. Nat Cell Biol 10, 102-109.

Numa, H., Kim, J.M., Matsui, A., Kurihara, Y., Morosawa, T., Ishida, J., Mochizuki, Y., Kimura, H., Shinozaki, K., Toyoda, T., *et al*. (2009). Transduction of RNA-directed DNA methylation signals to repressive histone marks in Arabidopsis thaliana. EMBO J 29, 352-362.

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

Pandey, R., Muller, A., Napoli, C.A., Selinger, D.A., Pikaard, C.S., Richards, E.J., Bender, J., Mount, D.W., and Jorgensen, R.A. (2002). Analysis of histone acetyltransferase and histone deacetylase families of Arabidopsis thaliana suggests functional diversification of chromatin modification among multicellular eukaryotes. Nucleic Acids Res 30, 5036-5055.

Paszkowski, J., and Whitham, S.A. (2001). Gene silencing and DNA methylation processes. Curr Opin Plant Biol 4, 123-129.

Peragine, A., Yoshikawa, M., Wu, G., Albrecht, H.L., and Poethig, R.S. (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.

Pfaffl, M.W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 29, e45.

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

Probst, A.V., Fagard, M., Proux, F., Mourrain, P., Boutet, S., Earley, K., Lawrence, R.J., Pikaard, C.S., Murfett, J., Furner, I., *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.

Qian, C., and Zhou, M.M. (2006). SET domain protein lysine methyltransferases: Structure,

specificity and catalysis. Cell Mol Life Sci 63, 2755-2763.

Que, Q., Wang, H.Y., English, J.J., and Jorgensen, R.A. (1997). The Frequency and Degree of Cosuppression by Sense Chalcone Synthase Transgenes Are Dependent on Transgene Promoter Strength and Are Reduced by Premature Nonsense Codons in the Transgene Coding Sequence. Plant Cell 9, 1357-1368.

Quina, A.S., Buschbeck, M., and Di Croce, L. (2006). Chromatin structure and epigenetics. Biochem Pharmacol 72, 1563-1569.

Rama Devi, S., Chen, X., Oliver, D.J., and Xiang, C. (2006). A novel high-throughput genetic screen for stress-responsive mutants of Arabidopsis thaliana reveals new loci involving stress responses. Plant J 47, 652-663.

Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D., *et al*. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature 406, 593-599.

Roberts, R.L., Metz, M., Monks, D.E., Mullaney, M.L., Hall, T., and Nester, E.W. (2003). Purine synthesis and increased Agrobacterium tumefaciens transformation of yeast and plants. Proc Natl Acad Sci U S A 100, 6634-6639.

Rocha, P.S., Sheikh, M., Melchiorre, R., Fagard, M., Boutet, S., Loach, R., Moffatt, B., Wagner, C., Vaucheret, H., and Furner, I. (2005). The Arabidopsis HOMOLOGY-DEPENDENT GENE SILENCING1 gene codes for an S-adenosyl-L-homocysteine hydrolase required for DNA methylation-dependent gene silencing. Plant Cell 17, 404-417.

Saleh, A., Alvarez-Venegas, R., Yilmaz, M., Le, O., Hou, G., Sadder, M., Al-Abdallat, A., Xia, Y., Lu, G., Ladunga, I., *et al*. (2008). The highly similar Arabidopsis homologs of trithorax ATX1 and ATX2 encode proteins with divergent biochemical functions. Plant Cell 20, 568-579.

Santos-Rosa, H., Schneider, R., Bannister, A.J., Sherriff, J., Bernstein, B.E., Emre, N.C., Schreiber, S.L., Mellor, J., and Kouzarides, T. (2002). Active genes are tri-methylated at K4 of histone H3. Nature 419, 407-411.

Saze, H., Shiraishi, A., Miura, A., and Kakutani, T. (2008). Control of genic DNA methylation by a jmjC domain-containing protein in Arabidopsis thaliana. Science 319, 462-465.

Schotta, G., Ebert, A., Krauss, V., Fischer, A., Hoffmann, J., Rea, S., Jenuwein, T., Dorn, R., and Reuter, G. (2002). Central role of Drosophila SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing. EMBO J 21, 1121-1131.

Schubert, D., Lechtenberg, B., Forsbach, A., Gils, M., Bahadur, S., and Schmidt, R. (2004). Silencing in Arabidopsis T-DNA transformants: the predominant role of a gene-specific RNA sensing mechanism versus position effects. Plant Cell 16, 2561-2572.

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

Shiu, S.H., and Bleecker, A.B. (2001). Receptor-like kinases from Arabidopsis form a monophyletic gene family related to animal receptor kinases. Proc Natl Acad Sci U S A 98, 10763-10768.

Shiu, S.H., and Bleecker, A.B. (2003). Expansion of the receptor-like kinase/Pelle gene family and receptor-like proteins in Arabidopsis. Plant Physiol 132, 530-543.

Steimer, A., Amedeo, P., Afsar, K., Fransz, P., Mittelsten Scheid, O., and Paszkowski, J. (2000). Endogenous targets of transcriptional gene silencing in Arabidopsis. Plant Cell 12, 1165-1178.

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

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

Tanikawa, C., Ueda, K., Nakagawa, H., Yoshida, N., Nakamura, Y., and Matsuda, K. (2009). Regulation of protein Citrullination through p53/PADI4 network in DNA damage response. Cancer Res 69, 8761-8769.

Tariq, M., Saze, H., Probst, A.V., Lichota, J., Habu, Y., and Paszkowski, J. (2003). Erasure of CpG methylation in Arabidopsis alters patterns of histone H3 methylation in heterochromatin. Proc Natl Acad Sci U S A 100, 8823-8827.

Teixeira, F.K., Heredia, F., Sarazin, A., Roudier, F., Boccara, M., Ciaudo, C., Cruaud, C., Poulain, J., Berdasco, M., Fraga, M.F., *et al*. (2009). A role for RNAi in the selective correction of DNA methylation defects. Science 323, 1600-1604.

Thorstensen, T., Fischer, A., Sandvik, S.V., Johnsen, S.S., Grini, P.E., Reuter, G., and Aalen, R.B. (2006). The Arabidopsis SUVR4 protein is a nucleolar histone methyltransferase with preference for monomethylated H3K9. Nucleic Acids Res 34, 5461-5470.

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

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

Vakoc, C.R., Mandat, S.A., Olenchock, B.A., and Blobel, G.A. (2005). Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. Mol Cell 19, 381-391.

Vaucheret, H., Beclin, C., Elmayan, T., Feuerbach, F., Godon, C., Morel, J.B., Mourrain, P., Palauqui, J.C., and Vernhettes, S. (1998). Transgene-induced gene silencing in plants. Plant J 16, 651-659.

Vaucheret, H., Beclin, C., and Fagard, M. (2001). Post-transcriptional gene silencing in plants. J Cell Sci 114, 3083-3091.

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

Wood, C.C., Robertson, M., Tanner, G., Peacock, W.J., Dennis, E.S., and Helliwell, C.A. (2006). The Arabidopsis thaliana vernalization response requires a polycomb-like protein complex that also includes VERNALIZATION INSENSITIVE 3. Proc Natl Acad Sci U S A 103, 14631-14636.

Woodcock, D.M., Crowther, P.J., Doherty, J., Jefferson, S., DeCruz, E., Noyer-Weidner, M., Smith, S.S., Michael, M.Z., and Graham, M.W. (1989). Quantitative evaluation of Escherichia coli host strains for tolerance to cytosine methylation in plasmid and phage recombinants. Nucleic Acids Res 17, 3469-3478.

Xiang, C., Han, P., Lutziger, I., Wang, K., and Oliver, D.J. (1999). A mini binary vector series for plant transformation. Plant Mol Biol 40, 711-717.

Ye, F., and Signer, E.R. (1996). RIGS (repeat-induced gene silencing) in Arabidopsis is transcriptional and alters chromatin configuration. Proc Natl Acad Sci U S A 93, 10881-10886.

Yokthongwattana, C., Bucher, E., Caikovski, M., Vaillant, I., Nicolet, J., Mittelsten Scheid, O., and Paszkowski, J. (2110). MOM1 and Pol-IV/V interactions regulate the intensity and specificity of transcriptional gene silencing. EMBO J 29, 340-351.

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

Zemach, A., Li, Y., Wayburn, B., Ben-Meir, H., Kiss, V., Avivi, Y., Kalchenko, V., Jacobsen, S.E., and Grafi, G. (2005). DDM1 binds Arabidopsis methyl-CpG binding domain proteins and affects their subnuclear localization. Plant Cell 17, 1549-1558.

Zeng, L., and Zhou, M.M. (2002). Bromodomain: an acetyl-lysine binding domain. FEBS Lett 513, 124-128.

Zhang, K., Mosch, K., Fischle, W., and Grewal, S.I. (2008). Roles of the Clr4 methyltransferase complex in nucleation, spreading and maintenance of heterochromatin. Nat Struct Mol Biol 15, 381-388.

Zhao, Z., Yu, Y., Meyer, D., Wu, C., and Shen, W.H. (2005). Prevention of early flowering by expression of FLOWERING LOCUS C requires methylation of histone H3 K36. Nat Cell Biol 7, 1256-1260.

Zheng, B., Wang, Z., Li, S., Yu, B., Liu, J.Y., and Chen, X. (2009). Intergenic transcription by RNA polymerase II coordinates Pol IV and Pol V in siRNA-directed transcriptional gene silencing in Arabidopsis. Genes Dev 23, 2850-2860.

Zheng, X., Pontes, O., Zhu, J., Miki, D., Zhang, F., Li, W.X., Iida, K., Kapoor, A., Pikaard, C.S., and Zhu, J.K. (2008). ROS3 is an RNA-binding protein required for DNA demethylation in Arabidopsis. Nature 455, 1259-1262.

Zheng, Z., Xing, Y., He, X.J., Li, W., Hu, Y., Yadav, S.K., Oh, J., and Zhu, J.K. (2010). An SGS3-like protein functions in RNA-directed DNA methylation and transcriptional gene silencing in Arabidopsis. Plant J.

Zhu, J.K. (2009). Active DNA demethylation mediated by DNA glycosylases. Annu Rev Genet 43, 143-166.

Acknowledgement

I am deeply indebted to my supervisor Prof. Dr. Gunter Reuter, whose guidance, suggestions and encouragement helped me in all the time of research and writing of this thesis.
This project would never have begun without the assistance of Dr.Ingo Hofmann, who helped me to improve my intellectual and practical skills.
I would like to thank Dr. Andreas Fischer, Kristina Imler and Katrin for their cooperation and good working environment in the lab.
I owe a lot to Sameer Phalke and Amit Sharma for their constant support at every step in my work.
I would like to thank the other members of the research group in the department of Development Genetics for their love ,care and constant support during my good and bad times, making my stay at the Institute a memorable one.
I am bound to thank Martin Zeeb and Jan for careful reading and correcting the draft version of my thesis and providing a lot of detailed and constructive comments for improvement.
I also acknowledge the financial support extended by Prof. Dr. Eckhard Lammert at Institute of Metabolic Physiology Düsseldorf during the final stages of my thesis.
I would like to thank my wife Ruchi, for her persistent encouragement, support and strong faith in my abilities, which had helped me to reach this stage. Also I would like to thank my son Rishabh who did not distract me with his pranks especially during the final stages of my thesis writing, allowing me to quickly wind up my thesis.
I would not be writing this thesis today if it were not for education and encouragement given to me by my parents.
Last but not the least I am grateful to SFB 648 for the financial support.

SELBSTSTÄNDIGKEITSERKLÄRUNG

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe. Desweiteren erkläre ich meine Kenntnisnahme der dem angestrebten Verfahren zugrunde liegenden Promotionsverordnung. Ich habe mich anderwärts nicht um einen Doktorgrad beworben und bin nicht im Besitz eines entsprechenden Doktorgrades.

Halle, den 29.06.10

……..

Deepak Jain

STATEMENT OF AUTHENTICATION

The work presented in this thesis is, to the best of my knowledge and belief, original except as acknowledged in the text. I herby declare that I have not submitted this thesis material, either in full or in part, for a degree at this or any other institution.

Halle, 29.06.10

……..

Deepak Jain

Lebenslauf

Persönliche Daten

Name, Vorname	Jain, Deepak
Anschrift	Thyagarajanagar, Bangalore
Geburtsdatum	14.02.1974
Geburtsort	Calcutta, Indien
Nationalität	Indisch
Familienstand	Verheiratet, 1 Kind

Schulbildung

1980-1986 Besuch der Primary School in Jalgoan, Indien.

1987-1988: Besuch der High School in Tandur (AP), Indien.

1988-1990: Besuch der Higher Secondary School in Hyderabad, Indien.

Wissenschaftlicher Werdegang

1991-1994: Studium an der People Education society, Bangalore (Indien); zur Erlangung des wissenschaftlichen Grades Bachelor of Science.

1995-1996: Studium an der Bangalore University (Indien), zur Erlangung des wissenschaftlichen Grades Master of Science.
Master thesis im Indian institute of Science
IISc Bangalore
Thema: '*Biodegredtion of Endosulfan*.
Betreuer: Dr. P K Shetty

1998-2001: Wissenschaft mitarbeiter in the Jawaharlal Nehru Institute for Advanced Scientific Research, Jakkur Bangalore.

2001-2003: Wissenschaft mitarbeiter in the UFZ (Umweltforschung zentrum) Halle-Leipzig.

2004-2009: Promotionsstudium am Institut für Genetik, Arbeitsgruppe Entwicklungsgenetik der Martin-Luther-Universität Halle-Wittenberg.
Betreuer: Prof. Dr. Gunter Reuter.

2010- Wissenschaft mitarbeiter in the Institute for Metabolic Physiology Düsseldorf.