universität wien

DISSERTATION

Titel der Dissertation

Maintenance of polyploidy-associated transcriptional gene silencing in *Arabidopsis thaliana*

angestrebter akademischer Grad

Doktor der Naturwissenschaften (Dr. rer.nat.)

Verfasser:	Tuncay Baubec, Mag. rer.nat.
Matrikel-Nummer:	0001447
Dissertationsgebiet (lt. Studienblatt):	A 091 441, Dr.-Studium der Naturwissenschaften Genetik - Mikrobiologie (Stzw)
Betreuerin / Betreuer:	Prof. Dr. Ortrun Mittelsten Scheid
	Gregor Mendel Institut of Molecular Plant Biology
	Austrian Academy of Sciences
	Dr. Bohrgasse 3, 1030 Vienna, Austria

Wien, am 04. Dezember 2008

TABLE OF CONTENTS

1.a. ZUSAMMENFASSUNG **1**

1.b. ABSTRACT **2**

2. INTRODUCTION **3**

2.1. Background **3**

2.1.1. Polyploidy

2.1.2. Epigenetic gene regulation

2.1.3. Epigenetic gene regulation in plants

2.1.4. DNA methylation

2.1.5. Histones and chromatin

2.1.6. Small interfering RNAs

2.2. Hypothesis **13**

2.2.1. Polyploidy associated transcriptional gene silencing (PA-TGS)

2.2.2. A functional genomics approach

2.2.3. Interference with PA-TGS

3. RESULTS **17**

3.1. Forward screen **17**

3.1.1. Introduction

3.1.2. Hygromycin phosphotransferase (HPT) locus

3.1.3. Reactivation of the silent HPT locus

3.1.4. Forward genetics

3.2. Initial characterisation of candidate lines **23**

3.2.1. Hygromycin selection

3.2.2. HPT transgene and transcript analysis

3.3. Detailed characterisation of candidate lines 25

3.3.1. Mutant groups

3.3.2. Putative uptake mutants

3.3.3. HPT transgene mutants

3.3.4. *trans*-acting mutants

3.4. Selected *trans*-mutant candidates – HPT locus analysis 28

3.4.1. HPT expression

3.4.2. DNA methylation at the HPT transgenic locus

3.4.3. Histone modifications at the HPT locus

3.5. Selected *trans*-mutant candidates – global analysis 37

3.5.1. Phenotype analysis

3.5.2. Ploidy analysis

3.5.3. Global 5-methyldeoxycytosine levels

3.5.4. DNA methylation at repeated regions

3.5.5. Chromocenter morphology

3.5.6. Nuclear 5-methyldeoxycytosine localisation

3.5.7. Nuclear histone modification patterns

3.5.8. Transposon reactivation

3.5.9. Transgene reactivation (TS-GUS)

3.5.10. Sensitivity to interference with DNA methylation and histone de-acetylation

3.6. Identification of mutant genes 49

3.6.1. T-DNA insertion analysis

3.6.2. T-DNA linkage analysis

3.6.3. Identification of T-DNA flanking regions

3.6.4. Mutation tracking with additional silent marker gene

3.7. Transcriptome analysis 54

3.8. Identified mutations 62

3.8.1 *HOMOLOGY-DEPENDENT GENE SILENCING* 1, *hog*1-7

3.8.2 *DECREASE IN DNA METHYLATION* 1, *ddm*1-11 to *ddm*1-13

3.9. Mutant overview table 65

4. DISCUSSION **67**

5. MATERIALS AND METHODS **87**

5.1. Plant growth and chemical treatments
5.2. Nucleic acid isolation and gel-blot analysis
5.3. Cation-exchange high pressure liquid chromatography
5.4. Reverse transcription PCR and real time PCR
5.5. *In situ* GUS detection
5.6. Fluorescence *in situ* hybridization (FISH) and immuno-labeling detection
5.7. Bisulfite conversion, sequencing and evaluation
5.8. Chromatin immuno-precipitation
5.9. Microarray expression analysis
5.10. Online databases and tools

6. ACKNOWLEDGMENTS **93**

7. CURRICULUM VITAE **94**

8. LIST OF ABBREVIATIONS **96**

9. REFERENCES **97**

1.a. ZUSAMMENFASSUNG

Die Vervielfachung des gesamten Chromosomensatzes (Polyploidisierung) kommt bei Pflanzen häufig vor und spielt eine entscheidende Rolle in der Evolution und der Pflanzenzüchtung. Der Erfolg polyploider Pflanzen beruht teilweise auf Konsequenzen der Genomverdoppelung: Veränderungen in der Genomorganisation und der epigenetischen Genregulation, die letztendlich zur Diversifizierung polypoider Spezies führen. Deshalb wird vermutet, dass die mit der Polyploidisierung im Zusammenhang stehenden molekularen Veränderungen für das Anpassungspotential der Pflanzen ungemein wichtig sind. Als Modell für die Erforschungen epigenetischer Diversität in polyploiden Pflanzen wurde ein Hygromycinresistenz-Transgen (HPT) in *Arabidopsis thaliana* verwendet, welches durch polyploidie-abhängige epigenetische Mechanismen inaktiviert und in dieser Form an alle nachkommenden Generationen weitergegeben wurde (Mittelsten Scheid *et al.* 2003)

Die vorliegende Arbeit beruht auf Untersuchungen der molekularen Mechanismen der polyploidie-assoziierten transkriptionellen Geninaktivierung (engl.: PA-TGS). Da der inaktive Zustand des Transgens auch nach Auskreuzungen mit diploiden Pflanzen beibehalten wurde, konnte ein genetischer Ansatz mithilfe von Mutagenese und Selektion durchgeführt werden. Zwanzigtausend T-DNA Insertionsmutanten wurden in der M2 Generation auf Hygromycinresistenz selektiert – ein Indiz auf Reaktivierung des HPT Transgens. Einundzwanzig resistente Pflanzen wurden auf genetische und epigenetische Veränderungen am HPT Transgen untersucht und nach ihren Eigenschaften klassifiziert. Zwei Mutanten sind möglicherweise in der Hygromycinaufnahme verändert, sieben Mutanten hatten genetische Veränderungen am HPT Transgen selbst, und zwölf Mutanten betreffen die erwartete *trans*-Regulation. Die letzteren wurden weiterhin auf Veränderungen in Genexpression und Chromatin von hoch-repetitiven genomischen Squenzen, auf veränderte Phänotypen und Expressionsmuster endogener proteinkodierender Gene analysiert. Aufgrund der vorliegenden Ergebnisse konnten diese Mutanten in zwei Gruppen aufgeteilt werden: fünf Mutanten mit starken Veränderungen epigenetischer Merkmale, deregulierter Genexpression und veränderter Morphologie, und sieben Mutanten ohne offensichtliche endogene Veränderungen. Drei Mutanten aus der ersten Gruppe stellten sich als neue Allele des Gens für den „chromatin remodelling“ Faktor *DDM*1 heraus, eine weitere als ein neues Allel des *HOG*1 Gens, dessen Produkt für die Hydrolyse von S-adenosyl-L-homocystein verantwortlich ist. Beide Faktoren sind für die Methylierung sowohl von DNA als auch von Histonen essenziell, entweder durch Regulierung der Chromatinstruktur oder der Biosynthese des Methylgruppen-Donors S-adenosyl-L-methionine. Diese Ergebnisse führen zu der Hypothese, dass PA-TGS durch voneinander unabhängige, aber vereinte Wechselwirkung von DNA- und Histonmethylierung aufrechterhalten wird.

1.b. ABSTRACT

Polyploidisation, the multiplication of the whole chromosome complement, is frequent in higher plants and had a significant role in evolution and breeding. The success of polyploid plants seems to be due to diversification that is least in part generated by major changes in genome organisation and epigenetic regulation associated with the chromosome duplication. Therefore, the molecular events connected with polyploidisation are believed to be very relevant for the adaptation potential in plants. In *Arabidopsis thaliana*, a single copy transgenic hygromycin resistance transgene (HPT) was found to be active or silenced in different lines during tetraploidisation experiments. The gene was maintained in its active or inactive state during subsequent generations and serves as a model for epigenetic diversity in polyploids (Mittelsten Scheid *et al*. 2003)

The aim of this thesis was to investigate the molecular mechanisms involved in maintenance and establishment of polyploidy-associated transcriptional gene silencing (PA-TGS). The silent epigenetic state remained stable in diploid derivatives after backcrosses, so that a forward genetic approach towards identification of *trans*-acting regulators could be performed by screening a collection of insertion mutants. Twenty-thousand T-DNA insertion lines were screened in the M2 generation for restoration of hygromycin resistance, indicating a loss of silencing. Resistant plants were selected and characterized for stability of hygromycin resistance in the progeny and for genetic and epigenetic changes at the HPT transgene. Following a thorough genetic and molecular characterisation, twenty-one hygromycin-resistant mutant lines were classified, including two potential hygromycin uptake mutants, seven mutations within the HPT transgene and twelve *trans*-acting mutants. The *trans*-acting mutants were further analysed for changes in expression and chromatin features at multicopy regions known to be targets of other epigenetic regulatory pathways, for growth and morphological changes and genome-wide changes in expression of endogenous protein-coding genes. Based on the results, the *trans*-acting mutants could be divided in two distinct groups: five mutants showing strong alterations of general epigenetic features and transcriptional regulation, going in line with morphological changes, and seven lines without observable effects on endogenous targets and phenotype. Three *trans*-mutants belonging to the first group were identified as novel mutant alleles of the chromatin remodelling factor *DDM*1 and one as an allele of the gene for the S-adenosyl-homocysteine hydrolase *HOG*1. Both factors are essential for maintenance of histone and DNA methylation, either through changing chromatin structure or regulating the biosynthesis of the methyl-group donor S-adenosyl-methionine. These results, together with the lack of HPT reactivation in the background of numerous other epigenetic mutants, suggest that maintenance of PA-TGS might involve DNA and histone methylation in an independent but combined manner.

2. INTRODUCTION

2.1. Background

2.1.1. Polyploidy

Polyploidisation, the multiplication of the whole chromosome complement, is frequent in higher plants. It contributed significantly to their evolution and adaptation and it turned out to be beneficial for plant breeding (Stebbins 1966). Depending on the origin of the chromosome sets, polyploids can be distinguished into allopolyploids (combination of multiple divergent genomes) and autopolyploids (multiplication of the same genome) that can arise in different ways (Ramsey and Schemske 1998). Autopolyploids occur either by spontaneous or induced lack of cytokinesis, fusion of unreduced gametes or tissue culture regeneration. Allopolyploids are generated through protoplast fusion and plant regeneration or natural or forced hybridisation of two different species and subsequent chromosome doubling. An example of such alloploidisation is the naturally occurring species *Arabidopsis suecica*, which is formed through hybridisation of *Arabidopsis arenosa* and *Arabidopsis thaliana*. In addition to recently produced and enduring polyploidy, ancient polyploidisation events, referred as paleoploidisation, have occurred in many organisms during their evolution and are thought to be the main cause behind gene duplications. Sequence analysis of the *Arabidopsis thaliana* genome revealed that at least two ancient polyploidisation events have occurred (AGI 2000).
Polyploidisation often leads to novel phenotypes not observed in progenitor plants, and these contribute to their success in entering new niches (Osborn *et al.* 2003). Experiments with neopolyploids, artificially produced polyploids, helped to understand the impact of polyploidisation on genome integrity and gene-regulation. They have revealed immediate effects of polyploidisation on phenotypes (e.g. cell or organ size), on gene expression by variation in gene dosage or altered regulatory interactions through increased heterozygosity, changes in epigenetic regulation or genome arrangements (Adams 2007, Bancroft 2001, Osborn 2003, Soltis and Soltis 2000, Wendel 2000) (Figure 1).

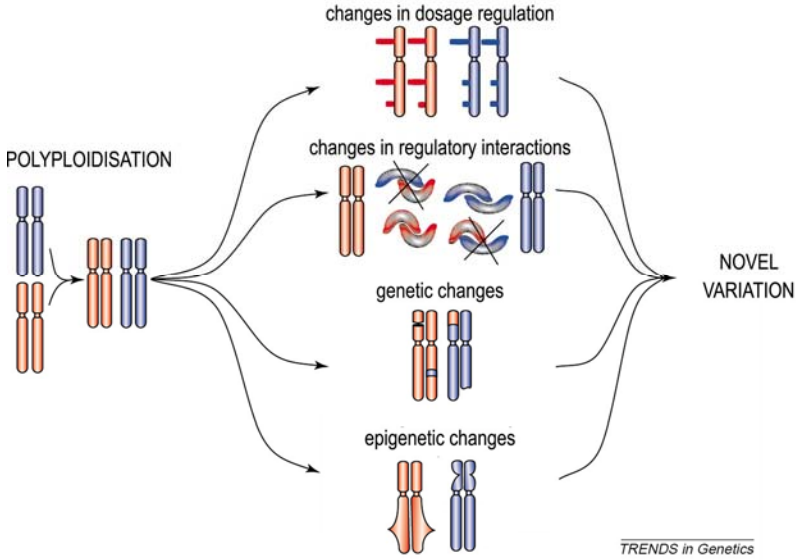

Figure 1. Consequences of polyploidisation.
Polyploidisation leads to novel variation through increased variation from dosage-regulated genes, altered regulatory interactions, genetic and epigenetic changes. (adapted from Osborn *et al*. 2003)

The significant and rapid genetic changes occurring in newly formed neopolyploids are thought to be generated by homeologous recombination, point mutations, gene conversion or transposon reactivation (Pontes *et al.* 2004, Song *et al.* 1995, Zhao *et al.* 1998). Changes in epigenetic gene regulation have also been shown to occur during polyploid formation (Chen and Ni 2006). Nucleolar dominance, a phenomenon in which rDNA loci are silenced in a parent-specific manner, has been well studied in allotetraploids (Pikaard 2000a, Pikaard 2000b). In the allopolyploid *Arabidopsis suecica*, the tandemly arranged 45S rDNA repeats that contribute to the nucleolar organizer region (NOR) are transcriptionally silent at alleles originating from the *thaliana* parent, whereas the *arenosa* alleles remain actively transcribed (Chen *et al.* 1998). This silencing is established during postembryonic development and maintained by epigenetic gene regulation (Pontes *et al.* 2007). Since epigenetic regulation in polyploid plants provides the background for this thesis, its mechanistic principles will be described in more detail in the following.

2.1.2. Epigenetic gene regulation

Epigenetic gene regulation is involved in heritable but dynamic control of gene expression, based on mechanisms found in uni- and multi-cellular eukaryotes ranging from yeast to man (Allis *et al.* 2006). It involves chemical modifications of DNA and associated histones, and structural changes in chromatin organization (Bird 2002, Chan *et al.* 2005, Jenuwein and Allis 2001, Spencer and Davie 1999, Wagner 2003). DNA modification is made up of methylation at the 5th carbon residue of cytosines forming 5-methyldeoxycytosine often referred to as the “fifth base” (Doerfler 2006). Histones are a group of small, basic and highly

conserved proteins that form together with 150bp stretches of DNA the basic sub-units of chromatin, namely the nucleosomes (Chakravarthy *et al.* 2005, Luger *et al.* 1997). The nucleosomes themselves can be subjected to positional alterations, thereby affecting the structure and packaging of chromatin and accessibility of DNA (Becker and Horz 2002, Kornberg and Lorch 1999). These features are not only determined by the density of nucleosomes but also by post-translational modifications of the histone tails protruding from the nucleosomes and interacting with other proteins (Jenuwein and Allis 2001). Once such modifications are established in certain combinations, either by developmental processes or as a response to environmental stimuli, they can dictate the transcriptional activity of single loci or whole chromosomal regions (Gaszner and Felsenfeld 2006, Talbert and Henikoff 2006). This transcriptional regulation is not only directed to endogenous coding genes but it is also actively regulating non-coding genes, transposons, transgenes and viral sequences, therefore contributing to regulation of development and morphology and to genome stability (Allis *et al.* 2006. The epigenetic state of chromosomal regions is evident under cytological analysis, where transcriptionally inert regions are identified as highly compact nuclear compartments surrounded by less compact regions (Heitz 1928). These compact structures are called heterochromatin, whereas the active counterpart is called euchromatin. Heterochromatic regions are known to contain highly repetitive sequences, such as transposons and other duplicated sequences and are enriched in DNA methylation and post-translational modifications of histones, mainly histone H3 methylation at the lysine residue K9 (Grewal and Moazed 2003, Lachner *et al.* 2001). Euchromatin is gene-dense and transcriptionally active with few transposons and repetitive elements. DNA methylation and H3K9 methylation are rarely found in these regions, whereas histone H3/H4-acetylation and methylation of histone H3 at the lysine residue K4 are enriched (Strahl *et al.* 1999, Turner 2000).

2.1.3. Epigenetic gene regulation in plants

Notably, the cytological distinction of hetero- and euchromatin was first described in plants (Heitz 1928). Epigenetic phenomena in plants, such as silencing and reactivation of maize transposons were first described by Barbara McClintock in 1944 (The Dynamic Genome, (Fedoroff and Botstein 1992)). Later, and till today, experiments with the model organisms *Arabidopsis thaliana, Petunia hybrida* and *Nicotiana species* provided numerous major breakthroughs in the epigenetic field. Some milestones are: the discovery of transgene silencing by transcriptional and post-transcriptional gene silencing (Jorgensen 2003, Matzke *et al.* 2004), the involvement of small RNA in RNA interference (RNAi) (Hamilton and

Baulcombe 1999) and recent advances such as whole genome methylation analysis at base-pair resolution (Cokus *et al.* 2008, Lister *et al.* 2008).

2.1.4. DNA methylation

Plants share the epigenetic hallmark of gene silencing by methylation of the 5th carbon residue of cytosines in DNA with some fungi, some insects and mammals. Unlike mammals that predominantly methylate cytosines in a CG sequence context only, plants exhibit methylation at CG, CHG and CHH sites, whereas H is every nucleotide except G (Chan *et al.* 2005). Deep sequencing of bisulfite converted *Arabidopsis* DNA (Cokus *et al.* 2008; Lister *et al.* 2008) have shown that ~5% of all cytosines are 5-methylcytosines, with 55% belonging to CG, 23% to CHG and 22% to GHH sequences. In addition, Cokus *et al.* have calculated the genome-wide levels of methylation in floral tissues to be 24% for CG, 6.7% for CHG and 1.7% for CHH. In order to establish methylation at cytosines, enzymatic activity from *de-novo* methyltransferases is required. In *Arabidopsis* this is achieved by the *DOMAINS REARRANGED METHYLTRANSFERASE* 2 (*DRM*2), a de-novo methyltransferase and *M.m.Dnmt3* homologue. How the methylation is directed to the unmethylated DNA is still not fully uncovered, but short interfering RNAs (siRNAs) are expected to play a crucial role (Cao *et al.* 2003). CG methylation is maintained after replication by the maintenance *METHYLTRANSFERASE* 1 (*MET*1), the pant homologue of *M.m.Dnmt1* (Kankel *et al.* 2003). Mutation in this maintenance methyltransferase results in genome-wide reduction of CG methylation from 56% to 1% (Lister *et al.* 2008) leading to reactivation of many pseudogenes and transposons throughout the genome, suggesting an important role for CG methylation in transcriptional silencing of non-coding genes (Zhang *et al.* 2006). Interestingly, in contrast to the severe loss of CG methylation, an increase in CHG and CHH methylation could be observed in *met*1 mutants that may fulfil redundant functions (Mathieu *et al.* 2007). Similarly, hypermethylation at specific regions, in addition to general reduction of CG methylation, can occur upon interference with MET1 function by antisense approaches (Jacobsen *et al.* 2000). Albeit the most prominent function of CG methylation seems to be maintaining transgenerational silencing of repetitive loci and transposons, another interesting feature is that many coding genes are methylated at CG dinucleotides in their gene-body, without affecting their transcription (Zhang *et al.* 2006) (Zilberman *et al.* 2007). DNA methylation at CHG residues is maintained by *CHROMOMETHYLASE* 3 (*CMT*3), a plant specific methyltransferase that is also thought to be involved in *de-novo* methylation (Lindroth *et al.* 2001). Maintenance of CHH specific DNA methylation after DNA replication is achieved by the *de-novo* methyltransferase *DRM*2 and requires siRNA for targeting (Cao *et al.* 2003).

Mutations in both *drm*2 and *cmt*3 are known to release transcriptional gene silencing mainly from numerous inactive protein-coding genes, suggesting a role of non-CG methylation in tissue-specific gene regulation (Zhang *et al.* 2006).

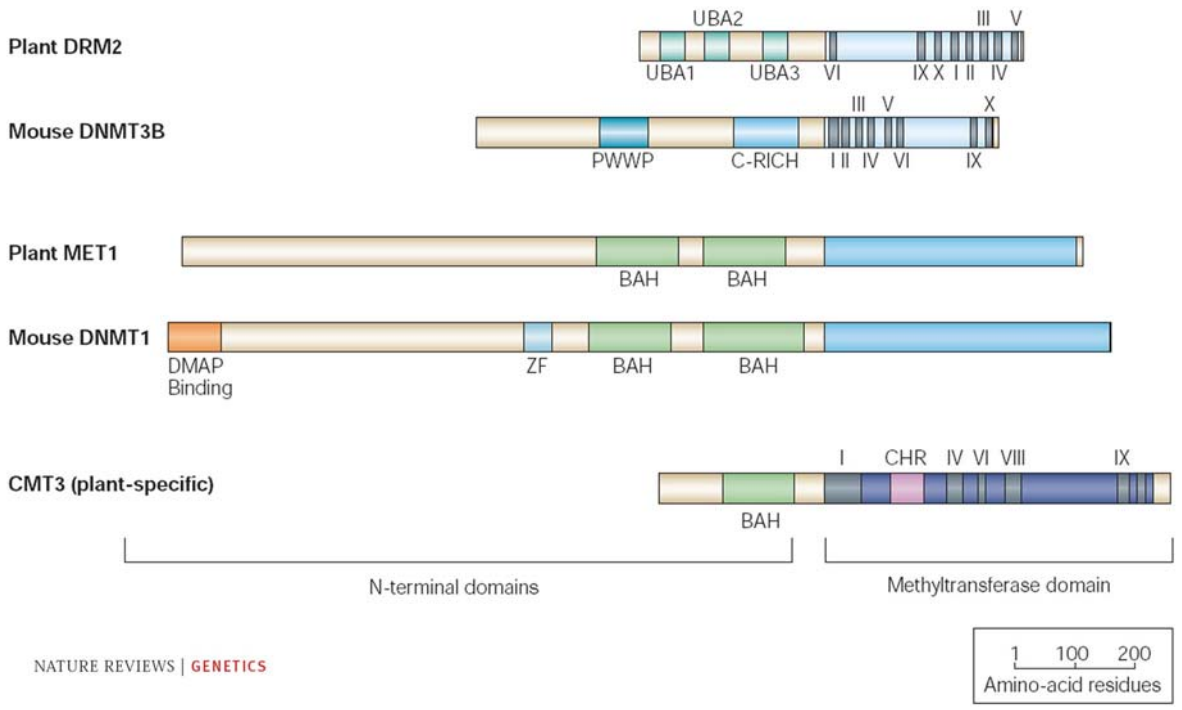

Figure 2. Plant DNA methyltransferases and their mammalian homologues. (Adapted from Chan *et al.* 2005)

In addition to the methyltransferases, the *Arabidopsis* genome also has active DNA demethylases that can remove methylcytosines from DNA through glycosylase activity followed by base excision repair. These glycosylases are DEMETER, that is expressed in the central cell of the female reproductive organ and known to be involved in genomic imprinting (Choi *et al.* 2002), two *DEMETER-LIKE* genes (*DML*2 and *DML*3) that have been shown to be required to maintain appropriate distribution of DNA methylation within the genome and protect gene ends from methylation (Penterman *et al.* 2007a, Penterman *et al.* 2007b) and REPRESSOR OF SILENCING 1 (ROS1) a glycosylase involved in maintaining methylation plasticity in the genome (Gong *et al.* 2002, Zhu *et al.* 2007).

Methyl-binding domain (MBD) proteins, such as the human MECP2, are known to have an important role in the readout of DNA methylation and thus affecting gene-regulation by increasing chromatin compaction (e.g. Rett Syndrome, (Bienvenu and Chelly 2006)). A survey for *Arabidopsis* proteins containing MBD domains resulted in the identification of 12 candidates that can be divided in 7 sub-classes (Zemach and Grafi 2003). However, how and if these MBD proteins are involved in gene regulation is still elusive.

2.1.5. Histones and chromatin

As stated above, posttranslational modifications of histones, especially at the N-terminal residues are substantially involved in transcriptional regulation of underlying DNA sequences (Jenuwein and Allis 2001, Kouzarides 2007). Such modifications are mainly methylation at lysine (K) and arginine (R) residues, acetylation of lysines (K), and phosphorylation at serine (S) and threonine (T) residues (Fuchs *et al.* 2006, Pfluger and Wagner 2007). Depending on the position and their combinations, these modifications provide a readout mechanism - similar to a code (Jenuwein and Allis 2001) - for numerous proteins that are able to bind specifically to the modifications and thereby regulate the chromatin properties and gene activity. The most extensively studied histone modifications associated with transcriptionally active chromatin in many organisms are histone H3K4 methylation and histone H3 and H4 acetylation (Turner 2000, Strahl *et al.* 1999). H3K9 and H3K27 methylation, on the other hand, are the hallmarks of transcriptionally inert chromatin (Kouzarides 2007). More specifically, methylation of H3K9 residues marks constitutive heterochromatin, a very stably maintained and silent entity (Lachner *et al.* 2001), whereas H3K27 methylation is associated with facultative heterochromatin that is able to change its status in a dynamic manner, therefore ensuring rapid response to developmental processes and environmental stimuli (Turck *et al.* 2007). As mentioned before, these marks fulfil the same function in many organisms, ranging from yeast to man including plants. However, some details are different in plants, as indicated by studies in *Arabidopsis.* Lysine residues can be mono-, di- or trimethylated, and this can be differently interpreted during the readout of the genetic code. While H3K9 di- and trimethylation, but not monomethylation are repressive marks in mammalian systems, H3K9 trimethylation is found to be associated with euchromatin in *Arabidopsis*, whereas mono- and dimethylation are heterochromatic marks (Naumann *et al.* 2005).

Although numerous histone-modifying enzymes have been identified in *Arabidopsis* based on sequence signatures in the core domains, their function is still under investigation. In plants these components are found in large gene families, arguing for potentially redundant functions or yet unexplored regulatory mechanisms. The *Arabidopsis* genome encodes 12 putative histone acetyl transferases (HATs) and 18 histone deacteylases (HDACs) – the enzymes that acetylate and deacetylate histonetails (AGI 2000, Pandey *et al.* 2002). Thirty two proteins contain the common SET domain found in histone lysine methyltransferases, and 14 proteins contain a jmjC domain that is involved in removing methylation from histone lysine residues (Saze *et al.* 2008).

Regulation of gene activity by polycomb group (PcG) and trithorax group (trxG) proteins, known from *Drosophila* as an epigenetic mechanism required for cellular memory of gene expression of homeotic genes (Schuettengruber *et al.* 2007), does also occur in plants (Pien *et al.* 2008). These proteins are organised in multimeric complexes and regulate gene expression by changes in chromatin structure. In *Drosophila* the polycomb repressive complex 2 (PRC2) is able to modify histone tails and methylates especially H3K27 and H3K9, whereas PRC1 recognizes these modifications and induces changes in chromatin structure and compaction. While polycomb group proteins suppress gene expression, trithorax proteins are required to maintain an active state. This is accomplished by their involvement in large complexes that act as histone methyltransferases, targeting mostly H3K4, or as ATP-dependent chromatin remodelers (Schuettengruber *et al.* 2007). In *Arabidopsis,* several homologue proteins were found that display similar functions. For example, PRC2 group components such as the WD40 repeat protein FERTILIZATION-INDEPENDENT ENDOSPERM (FIE) (homologous to *Extra sex combs)* or the histone binding domain protein MULTICOPY SUPRESSOR OF IRA 1 (MSI1) (Kohler *et al.* 2003) interact with CURLY LEAF, MEDEA or SWINGER (CLF, MEA, SWN) that belong to the SET domain group proteins homologous to *Enhancer of zeste* or with the zinc finger group proteins FERTILIZATION INDEPENDENT SEED 2, VERNALIZATION or EMBRYONIC FLOWER (FIS2, VRN, EMF2) homologous to *Suppressor of zeste*. Combinations of FIE, MSI1 and the other proteins result in regulation of genomic imprinting (MEA, FIS2), response to vernalization (VRN2, CLF/SWN) and regulation of floral organogenesis (EMF2, CLF/SWN) (Kohler and Villar 2008). No PRC1 component was found in *Arabidopsis*, arguing for a different readout mechanism of the established PRC2 marks (AGI 2000). However, there is growing evidence that the HP1 ortholog LIKE HETEROCHROMATIN PROTEIN 1 (LHP1) may play a role in repressing the genes that have been targeted by the PRC2 complex. *Arabidopsis* LHP1 binding sites overlap with euchromatic H3K27 trimethylation but not with H3K9 mono and dimethylation (Turck *et al.* 2007). As for the trithorax group proteins, five homologues were identified so far in Arabidopsis, namely ARABIDOPSIS TRITHORAX 1 to 5 (ATH1 to ATH5) (Alvarez-Venegas and Avramova 2001, Baumbusch *et al.* 2001). ATH1 has been shown to be implicated in regulation of floral homeotic genes and the floral repressor gene *FLOWERING LOCUS C* (*FLC*) (Pien *et al.* 2008).

The above described, in most cases reversible modifications result in a highly complex but yet very dynamic chromatin structure. This is necessary, considering that the long DNA molecules are densely packaged into a rather small nucleus but have to be still transcribed, replicated and repaired. Several proteins are involved in maintaining and remodelling chromatin in order to ensure compaction and accessibility to the DNA at the same time.

Chromatin assembly factors, such as the *Arabidopsis* FAS1 and FAS2 (Kaya *et al.* 2001, Ono *et al.* 2006), the homologues of the two larger CAF subunits, but also the already mentioned MSI1 protein (Hennig *et al.* 2003) and a rather uncharacterized protein BRU1 (Takeda *et al.* 2004) are involved in ensuring proper chromatin assembly during DNA replication and repair. Mutations in their genes are sensitive to DNA damage, but release also transcriptional gene silencing from different loci, indicating a link to epigenetic regulation. Additionally, a mutation in a gene involved in DNA replication and repair, *REPLICATION PROTEIN* A2 (*RPA*2), releases gene silencing in a DNA methylation-independent manner (Elmayan *et al.* 2005). ATP-dependent chromatin remodelling factors are involved in manipulating the DNA-histone interactions and therefore affect nucleosome positioning. These modifications on the nucleosome level have a major impact on the regulation of the underlying information by making the DNA more or less accessible to the transcription machinery or to histone and DNA modifying enzymes (Wagner 2003). So far, just a handful out of the more than 40 predicted ATP-dependent SWI2/SNF2 chromatin remodelling factors in *Arabidopsis* were functionally linked to regulation of transcription. The *M.m.Lsh1* homologue *DECREASED IN DNA METHYLATION* 1 (*DDM*1, (Jeddeloh *et al.* 1999)) and *DEFECTIVE IN RNA-DIRECTED DNA METHYLATION* 1 (*DRD*1, (Kanno *et al.* 2004)) were the only ones so far shown to be involved in regulating DNA methylation. *DDM*1 maintains transcriptional gene silencing at repetitive loci, mainly through preserving CG and H3K9 dimethylation (Gendrel *et al.* 2002, Mittelsten Scheid *et al.* 1998). *DRD*1, a plant-specific member of the gene family, is involved in the control of non-CG methylation at specific targets (Kanno *et al.* 2004). Another chromatin remodelling factor, *MORPHEUS' MOLECULE* (*MOM*, (Amedeo *et al.* 2000)) is involved in maintaining transcriptional gene silencing at repetitive loci in a DNA methylation-independent manner. All other SWI2/SNF2 factors characterised so far, *SPLAYED* (*SYD*), *PHOTOPERIOD-INDEPENDENT EARLY FLOWERING* 1 (*PIE*1) and *PICKLE* (*PKL*) are so far not connected to epigenetic gene regulation (Wagner 2003).

2.1.6. Small interfering RNAs

Research on epigenetic gene regulation by RNA interference (RNAi) or other processes linked to small RNA molecules has become a major topic over the past years and has demonstrated that it is a major mechanism of gene regulation in many eukaryotic organisms (Bartel 2004, Baulcombe 2004). Major contributions were made by elucidating these pathways in plants. Plant RNA-mediated gene silencing has been shown to be crucial for post-transcriptional and viral-induced gene silencing (PTGS and VIGS) where aberrant RNA

transcripts or viral RNAs that are able to produce double-stranded RNA (dsRNA) are recognized by the RNAi machinery, processed into 21-meres or 24-26-meres which then are able to induce cleavage of the target mRNA and furthermore ensure a systemic response (Voinnet 2005, Yoo *et al.* 2004). Another form of RNAi-mediated gene regulation is promoted by endogenous 21 nucleotide long microRNAs (miRNA) that are cleaved from a precursor pre-miRNA (Han *et al.* 2004, Vazquez *et al.* 2004). The produced miRNAs bind preferentially the 3' UTR of their endogenous target mRNAs, usually with imperfect complementarity. Unlike in animal systems, where the miRNAs mainly inhibit translation of the target mRNA (Pillai *et al.* 2007), plant miRNAs have been shown to induce cleavage of the target, although there is evidence that some miRNAs are also able to induce translational inhibition (Du and Zamore 2005). A similar mode of action has been shown for trans-acting siRNAs (tasiRNA). These 21 nt long RNAs are functioning in a similar way like the siRNAs involved in PTGS namely through cleavage of the target mRNA, but are, like the miRNAs, acting *in trans,* since they are produced from genomic loci that do not overlap with the target genes (Vasquez *et al.* 2004). The pathways described so far have one thing in common. They all act at a post-transcriptional level through targeting mRNAs in the cytoplasm and inhibiting its further processing, either by degradation or blocking translation. This distinguishes them from the RNA directed DNA methylation pathway (RdDM) that guides de novo methylation of DNA at cytosines in all sequence contexts (for review see Matzke *et al.* 2007) and is mainly involved in transcriptional gene-silencing (TGS) of introduced transgenes and various endogenous targets such as transposons, other repeated elements (e.g. rDNA and centromeric repeats, (Vongs *et al.* 1993)) and genes involved in development (e.g. *FWA* (Chan *et al.* 2006)). The methylation is accomplished by the nucleolar production of 24-26 nt long siRNAs that are guiding the DNA methylation machinery to the promoters of the target loci *in cis* or *trans* (Matzke *et al.* 2007).

Although the regulation of gene expression in *Arabidopsis thaliana* through RNAi pathways has been well studied over the past, it can be assumed that not all components of this machinery have been discovered. The known components can be divided into two classes. The first class is responsible for amplification or processing the siRNAs. Prominent members are the RNA-dependent RNA-polymerases RDR2 (TGS) and RDR6 (PTGS and TAS) promoting dsRNA-synthesis from a ssRNA precursor, or the DICER-like enzymes (DCL1 to DCL4), RNase III-like endonucleases that process dsRNA into siRNA. Although the DCL enzyme involved in PTGS is not yet identified (Vaucheret *et al.* 2004), each other pathway seems to have its own DCL. DCL1 for example is involved in miRNA production (Vasquez *et al.* 2004), DCL2 in processing of viral dsRNAs (Xie *et al.* 2004), DCL3 in TGS and DCL4 in the tasiRNA pathway (Allen *et al.* 2005, Gasciolli *et al.* 2005). Another important enzyme involved in the biogenesis or siRNAs is HUA ENHANCER 1 (HEN1), an RNA-specific

methylase that transfers a methyl group to the 3'-end of 21 to 24 nt long siRNAs, thus protecting them from uridylation and further degradation (Li *et al.* 2005). The second group of enzymes that function in the RNAi pathways are interacting with the siRNAs and the respective targets. These are, in the case of PTGS and other mechanisms involving targeting of mRNAs, the RNA-induced silencing complex (RISC) including ARGONAUTE 1 (AGO1), a PAZ domain-containing protein that is able to bind siRNAs and thus induces cleavage of complementary mRNAs (Baumberger and Baulcombe 2005, Carmell *et al.* 2002). A plant-specific DNA-dependent RNA polymerase (Pol IV) that is available in two forms (Pol IVa and IVb), depending on the largest subunit NRPD1a or NRPD1b, has been shown to be involved in RNA-dependent DNA methylation (Herr *et al.* 2005, Onodera *et al.* 2005, Kanno *et al.* 2005). The Pol IVa isoform is involved together with the SNF2-like protein CLASSY1 in the generation of siRNAs precursor transcripts through initially transcribing the target genes (Herr *et al.* 2005, Onodera *et al.* 2005, Smith *et al.* 2007). The second isoform, Pol IVb is involved together with the SNF-like protein DRD1, AGO4 and the *de novo* methylase DRM2 in mediating siRNA-directed DNA methylation at the target genes (Kanno *et al.* 2005a, Kanno *et al.* 2005b).

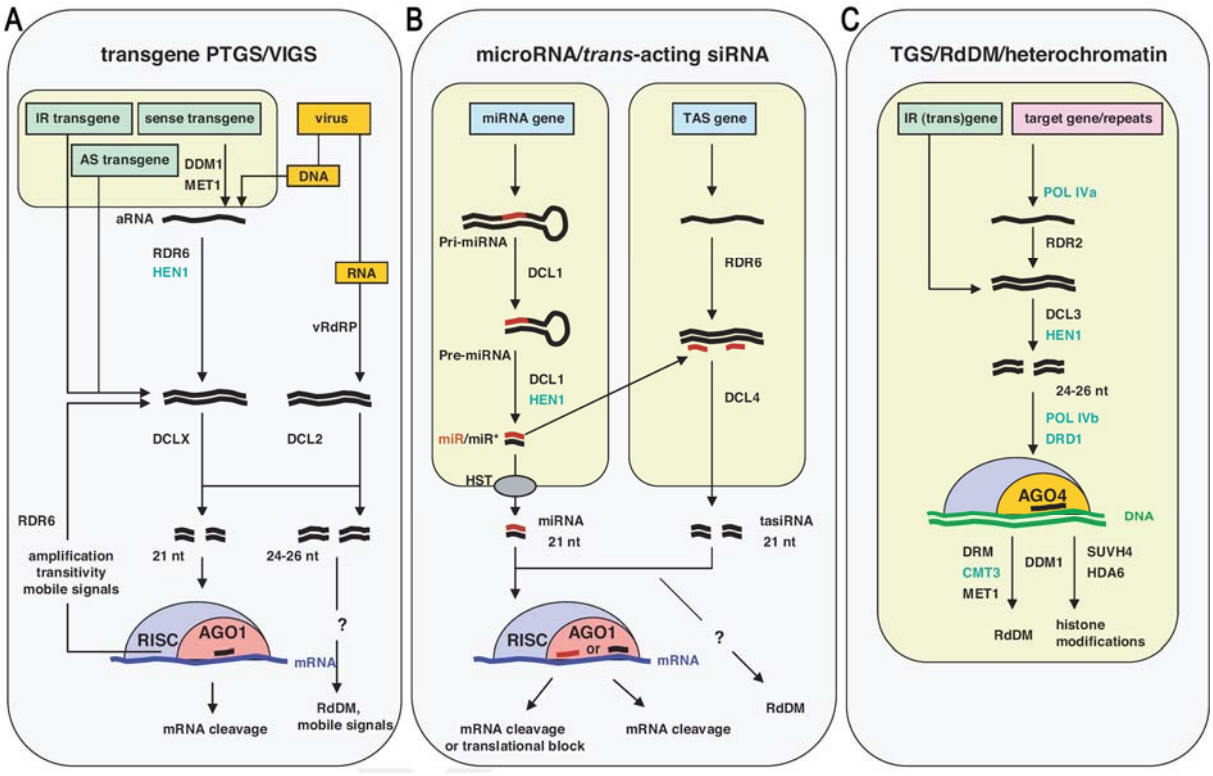

Figure 3. The major siRNA-dependent pathways involved in plant gene regulation. (A) post-transcriptional (PTGS) and virus-induced gene silencing (VIGS), (B) microRNA and trans-acting siRNA pathways and (C) transcriptional gene silencing (TGS) and RNA-dependent DNA methylation (RdDM) pathway (adapted from Allis *et al.* 2006, *Epigenetics*, Chapter 9, Matzke and Mittelsten Scheid, p181)

2.2. Hypothesis

2.2.1. Polyploidy-associated transcriptional gene silencing (PA-TGS)

As mentioned earlier, polyploidisation can lead to drastic effects in genome architecture and gene expression. Such an event occurred during auto-tetraploidisation of an *Arabidopsis thaliana* line carrying a transcriptionally active hygromycin phosphotransferase (HPT) transgene. The hygromycin-resistant line C (Mittelsten Scheid *et al.* 1996) gave rise to several tetraploid lines after plant regeneration from protoplast. Several lines obtained were still hygromycin resistant (referred as C4R) but some lost their ability to survive on hygromycin (referred as C4S) (Figure 4A). Remarkably, the sensitive lines still had the HPT transgene at the same location and genetically unchanged, therefore being isogenic when compared to the active lines. HPT expression analysis showed that the silent lines lost their ability to transcribe the HPT transgene through transcriptional gene silencing since no nascent transcripts could be detected. In addition, DNA methylation-sensitive restriction enzyme digests showed that the HPT transgene in the silent lines accumulated DNA methylation, as observed in other TGS cases (Figure 4B) (Mittelsten Scheid *et al.* 2003). These epialleles, established through polyploidy-associated transcriptional gene silencing (PA-TGS), were faithfully maintained during selfing and remained stable in all analysed generations. Even after reducing the chromosome number by outcrossing to diploid lines and generating diploid derivatives (referred to as C2R or C2S), the epialleles retained their either active or inactive status, inscribed by epigenetic modifications.

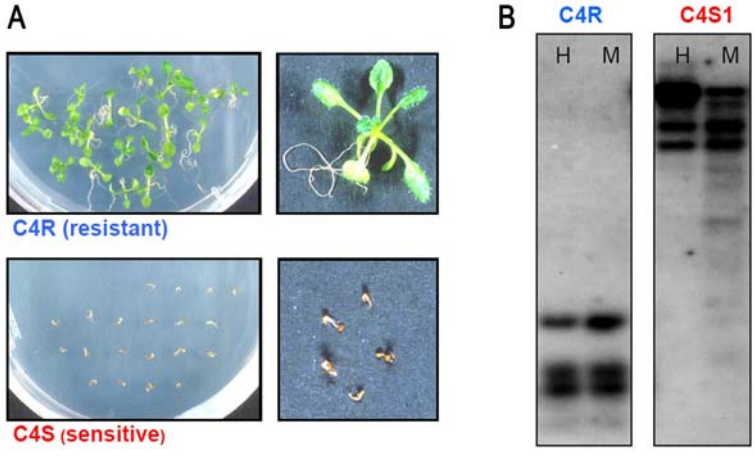

Figure 4. Epigenetic silencing of the single-copy HPT transgene in tetraploid Arabidopsis. **(A)** Hygromycin selection of the active (C4R) and inactive (C4S) tetraploid lines. **(B)** Southern blot analysis with DNA methylation sensitive restriction enzymes detects hypermethylation of the HPT transgene in the inactive tetraploid line. H: *Hpa*II (^{m}CG), M: *Msp*I (^{m}CHG). (Adapted from Mittelsten Scheid *et al*. 2003)

The epialleles had another interesting feature since they behaved paramutagenic upon genetic combination. Crosses between the active tetraploid line C4R and the inactive tetraploid line C4S led to fully hygromycin-resistant F1 plants. However, the F2 plants

obtained after selfing of F1 showed a non-Mendelian segregation of the resistant trait that gave rise to only 0 to 49% resistant plants after 8 weeks on hygromycin, compared to the expected 97.2% resistance, a ratio indeed observed in control crosses between the C4R line and a tetraploid wild type (Mittelsten Scheid *et al.* 2003). This argued for a paramutation-like interaction where the inactive paramutagenic epiallele (S) implied its state on the active paramutable counterpart (R), similar to interaction of certain alleles in maize and tomato (Brink 1973, Stam and Mittelsten Scheid 2005)). Strikingly, this was only observed in tetraploids; crosses between diploid C2R and C2S did not result in paramutation. Such interactions in polyploids were further arguing for the importance of epigenetic changes upon polyploidisation. The molecular mechanism of this polyploidy-associated paramutation is still unknown and differs from the mechanisms found to be crucial for other paramutation systems, where effects are observed already in F1 generations and involve siRNA-mediated silencing (e.g. *b1* and *pl1* paramutation in maize (Alleman *et al.* 2006, Hale *et al.* 2007).

2.2.2 A functional genomics approach

The findings outlined before hypothesize that potential epigenetic regulatory mechanisms may exist in plants that are activated only during or after changes of polyploidy. Such a mechanism might help to regulate and maintain transcriptional regulation form endogenous targets immediately after polyploidisation and could involve yet unknown regulatory components and pathways. In this thesis, the molecular mechanism behind maintenance and establishment of polyploidy-associated transcriptional gene silencing is addressed by a functional genetic approach based on the transcriptional activity of a previously silenced reporter gene.

Based on the Arabidopsis genome sequence (AGI 2000), 25 540 protein coding genes were predicted, including a large number of genes not associated with a known biological function (Yamada *et al.* 2003). The challenge for the scientific community is still to uncover and explore the function of these genes and to connect them to biochemical processes in order to understand how all the genes contribute to morphology, development, metabolism and resistance of the plant. Functional genomics is a widely used approach to study gene function by forward and reverse genetics. The reverse genetic approach makes use of gene identification in other biological model systems and search for available mutant alleles of specific genes in *Arabidopsis*, to analyse their role in the pathway of interest. The forward genetic approach applies random mutagenesis to obtain a certain phenotype (expected or undirected) and aims at identification of the mutated gene and confirmation of its linkage with the effect. Both approaches have advantages and a great potential to uncover gene function,

but also some disadvantages that have to be considered here. The reverse approach, for instance, requires that the gene of interest is available in a non-functional form. This is not trivial since directed gene manipulations like generating knock-outs by gene targeting as used in other systems are technically not possible in higher plants. The situation in *Arabidopsis* is still better than in most plants since there are many randomly generated T-DNA insertion lines with known insertion sites available from stock centres (Alonso *et al.* 2003). It is also possible to use rare, ecotype-specific, polymorphic alleles (e.g. *FWA*) or to generate artificial siRNAs down-regulating the gene of interest. However these are laborious and often unsatisfying approaches. The forward genetic approach has the disadvantage that it usually requires a high number of mutagenized individuals for the screen and a distinct phenotype to identify interesting mutations. Also, gene identification demands mapping of the introduced mutation which may turn out to be difficult, especially when mutagenesis generated multiple mutations or when polymorphisms suitable as genetic markers are missing or are unknown. Additional problems, especially for forward screens that aim to identify epigenetic regulation mechanism, can arise from loss of linkage between phenotype and mutation when the state of the reporter gene becomes autonomous after changing its epigenetic marks. Nevertheless, forward genetic approaches allowed identification of numerous genes involved in epigenetic gene silencing and regulation. Some examples are *DDM*1 and *MET*1, which were found in a forward screen using hypomethylation of endogenous centromeric repeats as readout (Vongs *et al.* 1993). Other screens using reactivation of silent marker genes led to the identification of *CMT*3, *HOG*1, *HDA*6 and many more (Aufsatz *et al.* 2002, Bartee and Bender 2001, Furner *et al.* 1998, Murfett *et al.* 2001). A few genes involved in epigenetic regulation were identified through monitoring silencing of active transgenes (e.g. *ROS*1 (Gong *et al.* 2002)) or through using available insertion mutants of orthologous epigenetic regulators known from other model organisms (*SUVH*2, *HDA*1, *DRM*1/2). For a detailed overview of factors found in mutant screens see Allis *et al.* 2006, *Epigenetics*, Chapter 9, p172-173.

2.2.3 Interference with polyploidy-associated transcriptional gene silencing

Although many factors involved in epigenetic regulation were already identified in reverse and forward screens, bioinformatic analysis of the *Arabidopsis* genome and its comparison to other genomes let it appear likely that there are still numerous components that have not been identified and characterised so far. The Plant Chromatin Database (www.chromdb.org) assigns 517 *Arabidopsis* genes to chromatin regulation, with many of them of unknown or only assumed function and still to be explored.

The initial formation of the silent allele in the freshly generated polyploid line and also the paramutation-like event in the tetraploid hybrids, indicating ongoing establishment of silent alleles, were observed only in tetraploids, arguing that ploidy-dependent epigenetic regulatory mechanisms exist in *Arabidopsis.* Since the link to polyploidy renders classical genetic approaches that are based on recessive mutations difficult, the fact that the silent epiallele was stable also in diploid derivatives was exploited, asking for the role of trans-acting factors required for maintenance of PA-TGS. None of the identified components had been shown to be involved in polyploidy-associated transcriptional gene silencing. Hence, the aim of this thesis was to find the key players involved in maintenance of PA-TGS. This was achieved using a functional genomics approach combining forward genetics (this thesis) and reverse genetics (Milos 2006, and this thesis). C2S1, the diploid derivative of C4S was mutagenised through T-DNA insertion or crossed to numerous TGS mutants affecting DNA methylation, siRNA, histone modification and chromatin-remodelling pathways. The resulting offspring generations were then screened for hygromycin resistance. In addition, interference with DNA methylation and histone acetylation by zebularine (Baubec *et al.* 2008) and trichostatin A or combinations of both was applied to challenge PA-TGS. Preliminary results from the reverse screen had shown that all analysed mutations were not able to reactivate the silent HPT transgene to significant hygromycin resistance (Milos 2006). This was taken as promising evidence that mutants found in the forward screen, based on stringent hygromycin selection, might allow to identify essential regulators of PA-TGS. The resistant lines obtained in the forward genetic approach were characterised for the molecular basis and stability of HPT reactivation and expression of other endogenous targets, for epigenetic properties at the HPT transgene and endogenous loci, and finally mapped for the introduced mutation. It turned out that re-established hygromycin resistance can be caused by interference with the maintenance pathway through introduced mutations either *in trans* or *in cis*. The latter would affect the HPT transgene itself through changing its sequence or genomic context and were expected to identify structural components that are required for susceptibility to PA-TGS. Mutations acting *in trans* would allow identification of the affected genes and suggest a role of their gene products in PA-TGS. These could be challenged subsequently also for a role in the paramutation-like interaction by applying dominant knock-down techniques. In any case, identification of novel *trans*- and *cis*- components would help to understand the mechanism of PA-TGS maintenance at the HPT transgenic locus and furthermore help to identify endogenous targets of PA-TGS and the consequences of PA-TGS interference.

3. RESULTS

3.1. Forward screen

3.1.1. Introduction

A forward genetic screen in the flowering plant *Arabidopsis thaliana* was applied to identify mechanisms involved in epigenetic regulation of gene expression in the context of polyploidy-associated transcriptional gene silencing (PA-TGS). The screen was performed in the diploid transgenic *Arabidopsis* line C2S1 that is a Zürich (wtZH) accession and carries an epigenetically silent, single copy hygromycin-phosphotransferase (HPT) transgene. This HPT transgene was originally active (C-line) and allowed the plants to grow on hygromycin (Mittelsten Scheid *et al.* 1996). During tetraploidisation of the C-line, the transgene became epigenetically silent in some derivatives (C4S1, C4S2) and remained stably in this state in all tested subsequent generations, even after crossing it back to the diploid state (C2S1) (Mittelsten Scheid *et al.* 2003). In addition, the silent epiallele was behaving paramutagenic, i.e. F2 populations from crosses between the tetraploid silent line C4S1 and the tetraploid active line C4R with only 0-50% hygromycin resistance had drastically reduced hygromycin resistance compared to the expected Mendelian ratio of 97.2%, suggesting that the inactive allele silenced the active allele through an unknown mechanism (Mittelsten Scheid *et al.* 2003).

3.1.2. Hygromycin phosphotransferase (HPT) locus

The transgenic locus, referred in this study as HPT locus, is a single insertion in the short arm of chromosome III, within BAC T5N23, at the genomic position 20.306.500, located at an intergenic position 700bp upstream of At3g54810 (*BME*3, BLUE MICROPYLAR END 3, a GATA type zinc finger domain-containing protein) and approximately 5.5kb downstream of At3g54800 (a plextrin homology domain-containing protein). Several *Arabidopsis* genes are regulated by the presence of neighbouring repetitive and transposable elements (Chan et al. 2006). Therefore, the bordering regions were analyzed for the presence of any potential regulatory element that may interact with the expression of the transgenic locus by masking the whole intergenic region between At3g54800 and At3g54810 for repeated sequences and known transposons using RepeatMasker (repeatmasker.genome.washington.edu) and plotting with Pipmaker to visualize the results (Schwartz *et al.* 2000). No known transposons

or repeats could be found in the analysed sequence (Figure R1). Also no significant enrichment in small RNA signatures or DNA methylation indicative of such elements (Llave *et al.* 2002) was found in the Arabidopsis MPSS Plus database (Lu *et al.* 2005) or in the Anno-J methylation database (Lister *et al.* 2008). Therefore no potentially interfering elements could be detected around the insertion locus.

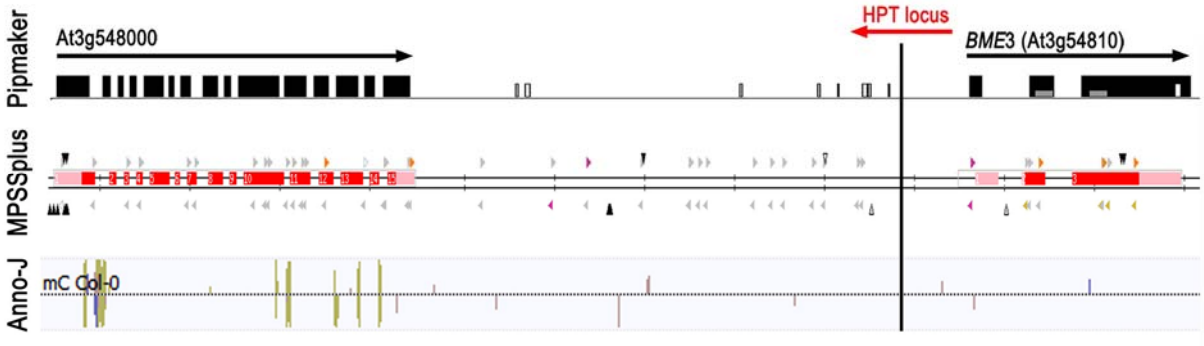

Figure R1: HPT transgene insertion site. The HPT insertion site on chromosome 3 with flanking genes. Multiple database outputs displaying interspersed repeats and transposable elements (Repeatmasker and Pipmaker), RNA signatures (MPSSplus) and DNA methylation (Anno-J). Legends - Pipmaker: small vertical rectangles denote simple repeats and small horizontal rectangles regions with CpG/GpC≥0.75; exons are indicated as black blocks. MPSSplus: coloured triangles show mRNA signatures and black arrowheads unique small RNA signature. Grey triangles indicate lack of significant expression; exons are indicated as red blocks. Anno-J: ^{m}CG is shown in green, while ^{m}CHG and ^{m}CHH are shown in blue and magenta, respectively.

The HPT transgene insertion itself is rearranged insofar as the CaMV35S promoter and some vector sequence are duplicated in tandem downstream of the HPT ORF and the truncated CaMV35S terminator (Figure R2B). Northern blot experiments using HPT-specific probes revealed that the HPT transcript in the active diploid line C2R was longer than expected (Figure R2A). Assuming correct termination, the expected mRNA length should be around 1000 nt in length. However, the observed HPT full length mRNA has approximately 3000 nt, arguing that the transcription is not stopped by the truncated CaMV35S terminator and goes through the second CaMV35S promoter into the plant DNA (Figure R2A and B, green arrow). Additionally, several shorter HPT transcripts at approximately 1700, 1500, 1300, 1000 and 700 nt were observed (Figure R2A). While these smaller transcripts were indeed containing the HPT sequence, an additional transcript could be detected that is presumably originating from the duplicated CaMV35S promoter and does not contain protein-coding information. Like the HPT transcripts, this is only present in the C2R line (Figure R2A) but not in C2S1. The size of this secondary transcript is approximately 800 nt and seems to terminate at approximately the same site in the plant DNA as the main HPT transcript (Figure R2B, red arrow).

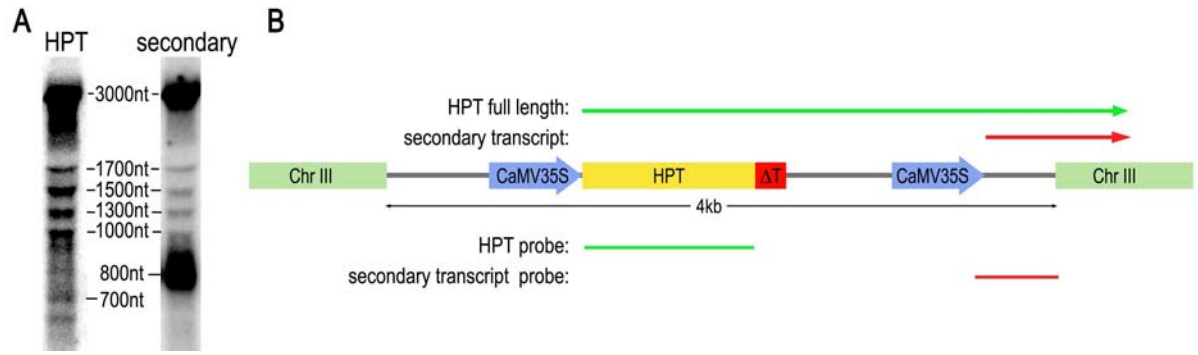

Figure R2: (A) Northern blot analysis of HPT and secondary transcripts in expressing line C2R. **(B)** HPT transgene structure with duplicated CaMV35S promoters and vector sequence, HPT coding region and truncated CaMV35S terminator. Green arrow: full length HPT transcript, red arrow: secondary transcript. Probes used for northern analysis are depicted in the corresponding colours.

3.1.3. Reactivation of the silent HPT locus

Inhibitors of DNA methylation and histone deacetylases have been widely used as potentially reactivating agents for epigenetically silenced endogenes and transgenes (Chang and Pikaard 2005). Therefore the DNA methylation inhibitor zebularine (Zhou *et al.* 2002) and the HDAC inhibitor trichostatin-A (Yoshida *et al.* 1995) were used to test the silencing stability at the HPT locus. Diploid C2S1, C2R, wtZH and line A (containing a multicopy HPT insertion, Mittelsten Scheid et al 1998) seeds were germinated and grown for three weeks on plates containing combinations of hygromycin with zebularine and/or trichostatin-A. Beside the growth retardation that was due to the zebularine treatment and except for the HPT expressing line C2R, no hygromycin resistance could be observed, neither in C2S1 nor in line A (Figure R3). Therefore, the applied drug treatment cannot restore the resistant phenotype.

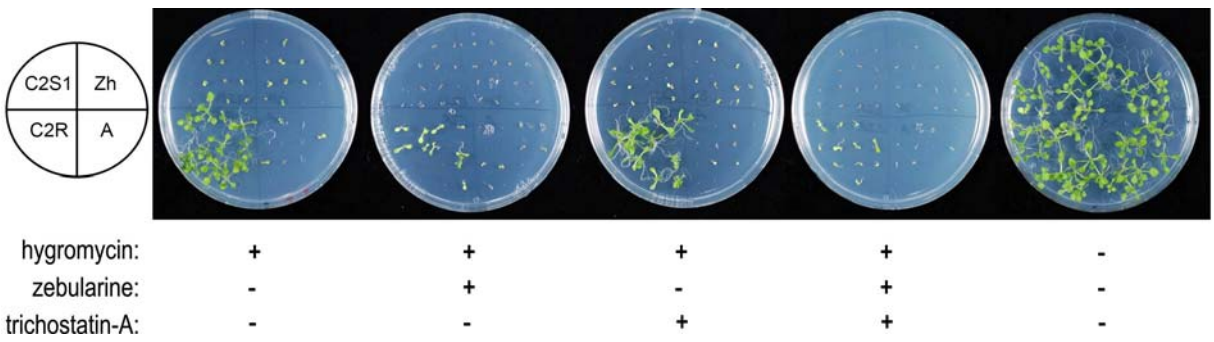

Figure R3: Reactivation test using zebularine and trichostatin-A. C2S1, wtZH, C2R and line A seedlings were tested for hygromycin resistance on plates containing 10 µg/ml hygromycin with or without zebularine (40 µM) and/or trichostatin-A (1.8 µM). A drug-free plate was used as control.

Molecular gene expression analysis of C2S1 seedlings treated with 20, 40 and 80 µM zebularine showed a minimal increase in HPT transcript levels, but obviously not high enough to ensure hygromycin resistance. Namely, HPT transcript abundance in 80 µM zebularine treated C2S1 was 57-fold higher than in mock treated C2S1, but just 0.07-fold compared to the HPT levels in the active line C2R (Figure R4B and C). However, Northern analysis of the secondary transcript revealed strong reactivation after zebularine treatment, suggesting a less stable transcriptional silencing at the duplicated promoter (Figure R4B). DNA methylation analysis of both CaMV35S promoters after zebularine treatment using methylation-sensitive restriction enzymes and subsequent Southern blotting showed increasing hypomethylation at the available CG sites (*Hpa*II) and CHG sites (*Msp*I), with the latter sites being stronger affected by the drug treatment. In both cases, the degree of methylation was unquestionably lower compared to C2S1, but the hypomethylation did not reach C2R levels (Figure R4D). To get a more detailed insight into DNA methylation patterns at the duplicated promoters after zebularine treatment, DNA extracted from treated and untreated seedlings was bisulfite-converted, amplified by PCR and ten clones per promoter and treatment were sequenced and analysed. The observed changes in DNA methylation were remarkably similar between the two promoters. In both cases, total methylation was decreased by 5 and 4% only, whereas sequence-specific methylation was decreased by 9 and 5% at CG, 13 and 10% at CHG and 3% at CHH sites in the main and duplicated promoter, respectively (Figure R4D). This is rather surprising since the duplicated promoter was much stronger reactivated after zebularine treatment.

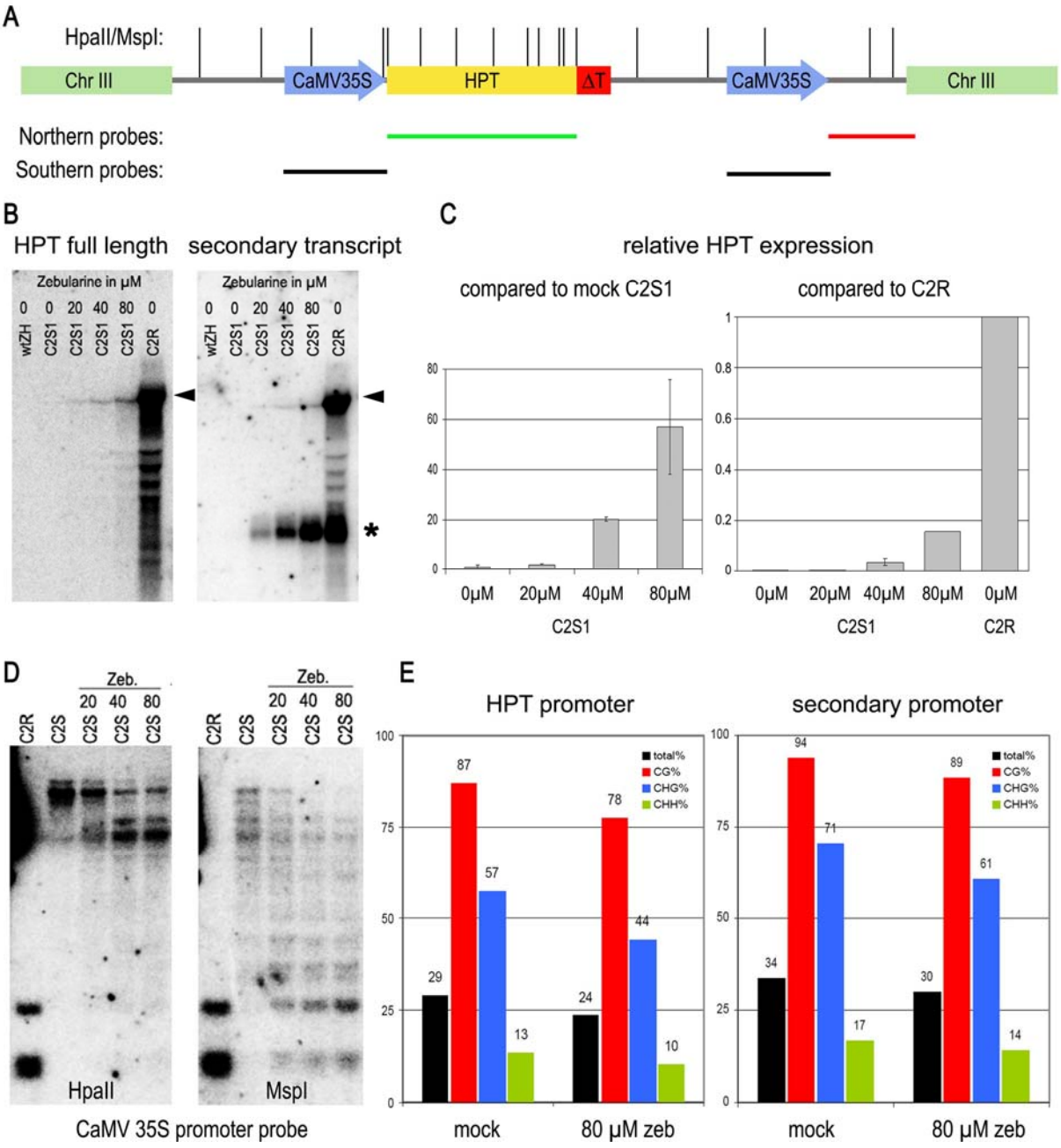

Figure R4: Molecular analysis of the HPT locus after three weeks of zebularine treatment.

(A) HPT locus, denoting *Hpa*II/*Msp*I restriction sites, Northern probes: green for the full length HPT and red for the secondary transcript, and Southern probes for the CaMV35S promoters (black). **(B)** Northern blot analysis for the HPT full length (arrowhead) and secondary transcripts (asterisk) after zebularine treatments compared to mock treated C2S1, wtZH and C2R. **(C)** Real time PCR analysis of HPT mRNA levels normalized to *ACTIN* mRNA and calibrated either to C2S1 or C2R. Note that the comparison to C2R uses log-scale for relative HPT expression. **(D)** DNA methylation Southern analysis with CaMV35S promoter probes on *Hpa*II (^{m}CG) or *Msp*I (^{m}CHG) digested DNA from zebularine- C2S1 or mock-treated C2S1 and C2R plants. **(E)** DNA methylation percentage after bisulfite analysis of the duplicated promoters in mock- and 80 µM zebularine-treated plants. Total DNA methylation: black, ^{m}CG: red, ^{m}CHG: blue and ^{m}CHH: green.

3.1.4. Forward genetics

The molecular mechanism of establishing the silencing during the paramutation-like interaction is difficult to investigate by genetic means, especially since it is limited to the tetraploid background. Therefore, an approach was chosen to identify important components first based on their role in maintaining the HPT silencing, making use of the selectable marker gene. Several *trans*-acting factors were known already to regulate transcriptional gene silencing, and loss-of-function mutants of these were tested whether they would reactivate the silent HPT in a reverse genetic approach. This was done by introducing the silent transgene into several epigenetic mutant backgrounds by crossing and scoring for hygromycin resistance in F2 and later generations. Although representative components of all possible pathways known to be involved in epigenetic regulation were tested, including DNA repair and genome stability mutants, this approach did not result in restoring hygromycin resistance, except for *ddm*1-5 (Mittelsten Scheid et al. 2003, Milos 2007). This mutant is deficient in DDM1 (DECREASED IN DNA METHYLATION 1), an ATP-dependent chromatin remodeler known to be involved in epigenetic regulation (Jeddeloh *et al.* 1999) and could reactivate the HPT transgene, resulting in partial hygromycin resistance, after prolonged inbreeding (Mittelsten Scheid *et al.* 2003, this thesis).

Therefore, a complementing, forward genetic approach was applied to identify the keyplayers involved in this epigenetic regulation. C2S1 plants, carrying the silent HPT transgenic locus, were mutagenized by floral dipping in *Agrobacterium tumefaciens* cultures carrying a T-DNA transformation vector encoding resistance to phosphinotricine (BASTA) and creating insertional mutations randomly in the genome. Therefore, mutagenized M1 seedlings could be easily selected by spraying with BASTA, resulting in lethality of untransformed plants (Figure R5A). M2 seeds (selfed progeny and containing segregating mutant genotypes including homozygotes) from a total number of 20.000 BASTA-resistant M1 plants were then screened on hygromycin for loss of silencing at the HPT transgene resulting in hygromycin resistance. This was done by germinating and growing pooled M2 seeds from 15 individual M1 plants on large Petri dishes with semisolid agar containing 8 µg/ml hygromycin (Figure R5B). The HPT-active and -inactive lines, C2R and C2S1, were used as positive and negative controls, respectively. After two weeks, the plates were screened for resistant plants (Figure R5C), which were transferred to hygromycin-free medium and later to soil. Each individual hygromycin-resistant M2 plant was numbered according to pool number and plant number within the pool, if more than one resistant plant was found.

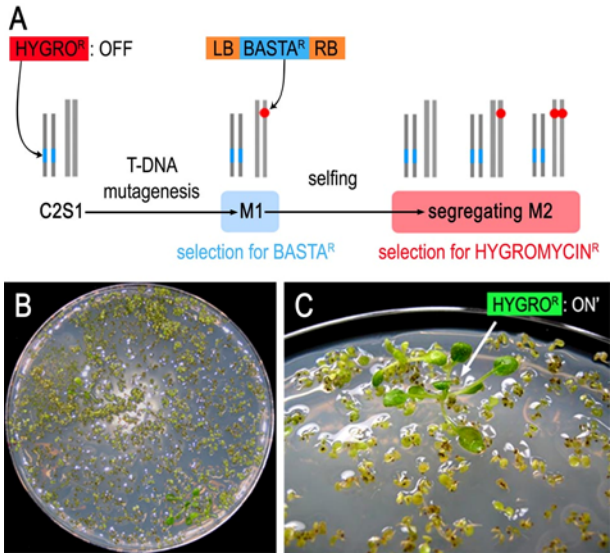

Figure R5: Forward genetic screen
(A) Forward screen procedure: C2S1 plants carrying a silent HPT transgene (blue region) were mutagenized by T-DNA insertion (red dot). Mutagenized M1 plants were screened for BASTA resistance encoded by the T-DNA. M2 seeds from selfed BASTA resistant M1 plants were screened on hygromycin for HPT reactivation. **(B)** Hygromycin screening plate example and **(C)** mutant candidates, hygromycin resistant plants.

3.2. Initial characterisation of candidate lines

3.2.1 Hygromycin selection

During the initial screening procedure, a total number of 34 M2 plants (excluding siblings) were found to be hygromycin-resistant. These plants were subjected to further characterisation in order to proof and analyse the HPT reactivation. One main part of the characterisation was further selection and segregation analysis. For this, M3 selfing populations and F2 populations from crosses between M2 plants and wild type Zürich (wtZH, the ecotype background of C2S1) lines were surface-sterilized, allowed to germinate and grown on selection plates containing 10 µg/ml hygromycin. After 3 weeks, resistance to hygromycin was scored. For the selfed M3 pools, the hygromycin resistance ranged from 0% to 100% (Table R1). Pools with 0% resistance were excluded from further analysis, since the M2 parents were either “false positives” or displayed only transient HPT reactivation. The high rate of M3 populations with resistance ratios below the expected 100% for homozygous recessive mutants suggested that the mutations are either not fully penetrant or dominant and segregating (expected resistance: 25-75%). Hygromycin resistance in F2 after outcrossing was scored from multiple individual populations (in most cases three) and ranged from 0% to 90% (Table R1). Surprisingly, the F2 populations did not show the expected Mendelian segregation ratio of 18.75% (3/16) calculated for recessive mutations in a segregating HPT transgene background. In most cases the ratio was higher than 50%. Hygromycin resistance was surveyed for most lines further during subsequent generations to ensure stability of the transgene reactivation and to monitor segregation, linkage and dominance or recessivity of the introduced mutation, which will be discussed later.

Table R1: Hygromycin resistance in selfed M3 and M2 x wtZH, F2 seedling pools.

Line	M3 Hyg^R[%]	F2 Hyg^R [%]	Line	M3 Hyg^R [%]	F2 Hyg^R [%]
112	4.8	0; 61.2; 52.9	931	0	0; 0; 0
187	4	70; 45.5; 0; 72.1	961	0	nd
243	0	0; 0	964-1	36.7	90; 4.1; 0
262	100	53.9; 81.4	964-2	36.7	0; 70.2; 0
279	81.2	0; 52.8	975	71	0; 0
283	69	77.5; 76	1119	0	nd
298	40	70.7; 62.5	1135	57.1	53.4; 8; 6
348	52.5	73.5; 73.5; 0	1139	98	66; 69; 75
350	85.7	nd	1142	57.8	8.2; 2; 6
409	66.7	70; 73.1; 71.4; 81.3; 67.3; 75	1146	33	76; 70; 66
507	14.3	0; 0	1148	0	nd
511	8	0; 0	1159	0	nd
551	0	62.5; 0; 0; 0	1207	0	nd
552	72.9	0; 38.3	1208	0	nd
555	10.4	0; 0	1213	0	nd
761	53.8	68.8; 62.5; 0; 0	1341	56	nd
774	92.3	71.4; 74.5; 80	1361	77.6	0; 0; 77.3

3.2.2. HPT transgene and transcript analysis

All lines that were still resistant to hygromycin in M3 and/or F2 generations were further analysed for HPT transcription and for the integrity of the transgenic structure to investigate if the resistance was in fact due to transcriptional reactivation of the transgene and to rule out possible rearrangements in the transgenic locus. Northern blot analysis using HPT-specific probes hybridized to total RNA from mutant lines showed that all candidates except 262 and 350 indeed reactivated the HPT transgene, while no transcript could be detected in the parental C2S1 line (Figure R6A). However, the lines 279, 283, 975, 1139 and 1146 expressed HPT transcripts that were shorter than the transcript observed in C2R (Figure R6A).

Southern blot analysis using CaMV35S- and HPT-specific probes against total DNA extracted from mutant lines and fragmented with *EcoR*I or *Hind*III showed that most of the mutants had an intact HPT locus structure, similar to that of C2S1 and C2R (Figure R6C). However, the lines 262, 279, 283, 298, 350, 975, 1139 and 1341 had visible changes in the restriction fragment pattern, indicating rearrangements in the HPT locus structure. The most drastic changes in lines 262 and 350 are concomitant with the lack of any detectable HPT mRNA analysis, while these mutants must have acquired hygromycin resistance differently.

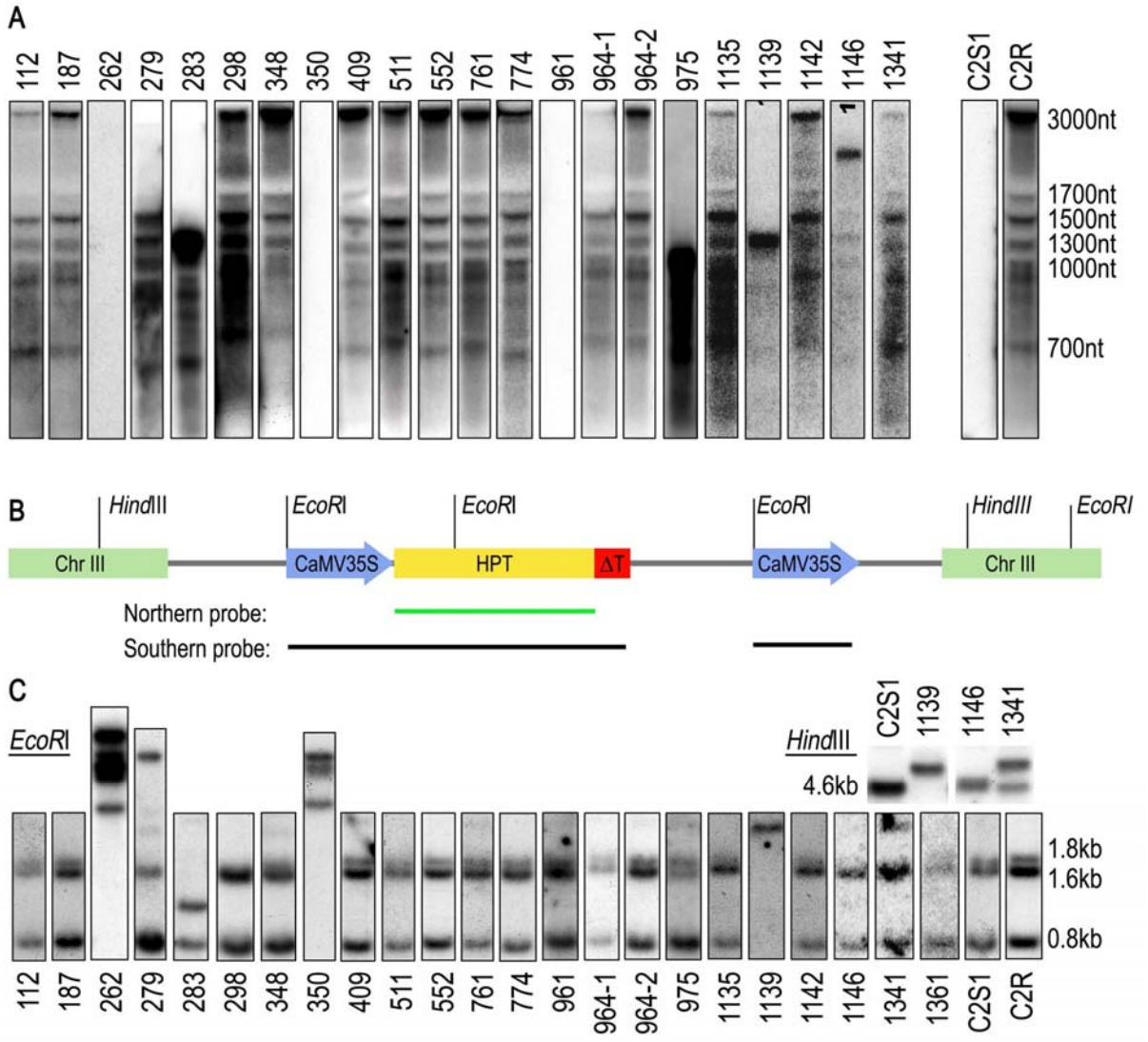

Figure R6: HPT gene and transcript analysis.

(A) Northern blot analysis with probes specific for the full length HPT transcript in all mutants, C2S1 and C2R. **(B)** HPT locus denoting *EcoR*I and *Hind*III restriction sites, Northern probe specific for the HPT transcript in green and Southern probes specific for the CaMV35S promoters and HPT coding region in black. **(C)** HPT locus structure analysis by Southern blot on *EcoR*I- or *Hind*III-digested DNA from mutants, C2S1 and C2R lines.

3.3. Detailed characterization of candidate lines

3.3.1. Mutant groups

Following the data obtained from HPT expression and structure studies, 21 lines were grouped based on their similarities. Three different groups contained potential uptake mutants (2 lines), HPT transgene mutants (7 lines) and *trans*-acting mutants (12 lines).

Figure R7: Mutant groups.
Mutant lines grouped according to similar HPT locus structure and expression. Green: *trans*-acting mutants, blue: HPT transgene mutants and orange: potential uptake mutants.

3.3.2. Putative uptake mutants

The lines 262 and 350 were fully hygromycin-resistant in all tested generations (M2 to M5, Figure R8A). However, no HPT mRNA could be detected by Northern blot or Reverse Transcription PCR (Figure R8B), suggesting a putative mutation in the hygromycin uptake mechanism. Hygromycin-resistant seeds of line 262 were allowed to germinate and grown on kanamycin or gentamycin plates in order to test if the mutation were specific for hygromycin or would also affect the uptake of two related drugs. The 262 seedlings could not survive on kanamycin and gentamycin – like wtZH used as a negative control, while transgenic lines harbouring the appropriate resistance genes survived (Figure R8C). Therefore, the mutation seems to be specific for hygromycin resistance. Although potentially interesting for applications, these lines were not further analysed during this thesis, but are mentioned here for the sake of completeness.

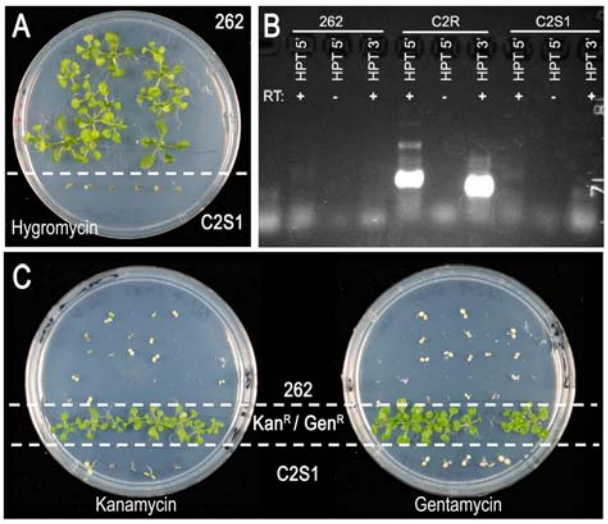

Figure R8: Putative hygromycin uptake mutation. **(A)** Hygromycin selection (10 µg/ml), **(B)** RT-PCR to detect HPT transcripts and **(C)** kanamycin and gentamycin selection with 262, one of the putative hygromycin uptake mutant lines. C2S1 was used as negative control in all experiments, C2R as positive control for the RT-PCR and Kan^R or Gen^R transgenic lines (kindly provided by Dr. Werner Aufsatz) were used as positive controls for the respective selection plates.

3.3.3. HPT transgene mutants

Among the 21 resistant lines were 7 lines that showed either truncated HPT transcripts (975 and 1146), rearranged HPT transgenic structure (298 and 1341), or both (279, 283 and 1139). The results suggest that these lines carry modifications at the HPT locus itself, thereby reactivating the HPT transgene. In addition, the data obtained from hygromycin resistance assays with F2 seeds from wtZH outcrosses show that in most cases, the resistance ratio is either ~75% or 0%, as expected for a hemizygous, dominant HPT-mutation (Table R1). By comparing the truncated HPT transcripts with that of the unchanged HPT locus and by evaluating the southern blot data to reconstruct the locus, it strikes out that all transcripts either terminate or have deletions around the duplicated CaMV35S promoter site, suggesting a role of this region in transcriptional regulation (Figure R9).

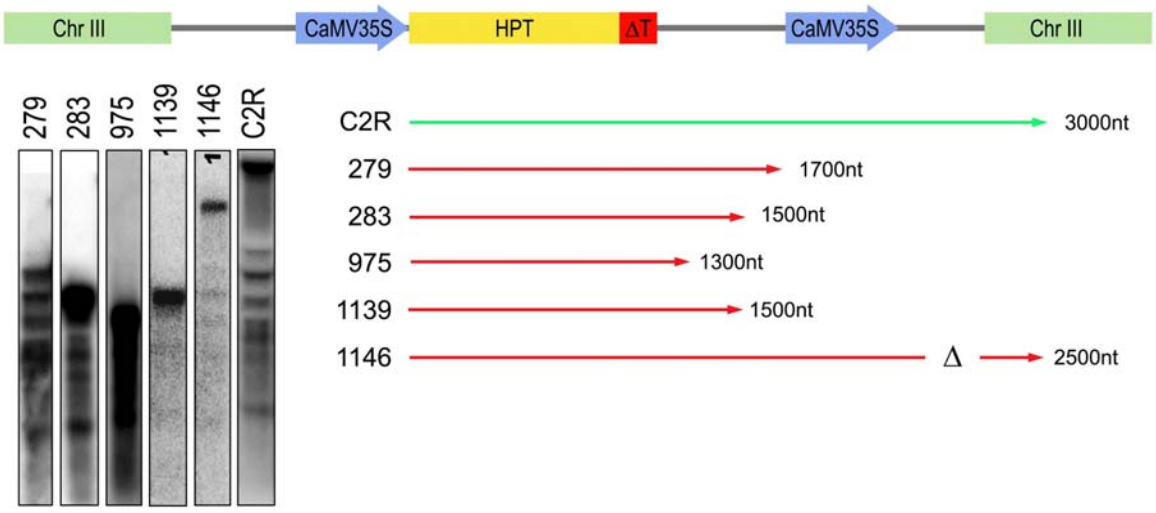

Figure R9: HPT transgene mutations. Northern blot analysis revealed 5 lines with truncated HPT transcripts. The HPT transcript length is plotted from the HPT transcription start site for each mutant (red arrows) to visualize potential breakpoints compared to the full length transcript (green arrow). The estimated deletion in 1146 found by Southern analysis is also shown.

The fact that there were 7 out of 22 mutants (31.8%) occurring in the HPT transgene itself indicates that the screen was saturated. Since these mutations were not obviously coupled to integration events, the *Agrobacterium* transformation must have caused numerous rearrangements and deletions, likely also throughout the genome, which may not be linked to the T-DNA insertion.

3.3.4. *trans*-acting mutants

The remaining 12 lines showed stable expression and full length HPT transcripts resulting in hygromycin resistance observed throughout all generations tested (M1 to M6). In addition, no changes in the HPT transgene structure could be detected. Therefore, these line were chosen for further characterisation and mapping of the introduced mutations. For better comparison between the mutant lines and for genetic homogeneity, seed populations of comparable generations were selected that had ~100% hygromycin resistance. Nine lines were 100% resistant to hygromycin at M4, and 964-1 seeds showed 100% resistance only in M5, while no resistant plants could be found in later generations for 112 and 187 after selfing. These latter lines were therefore not systematically included in all further assays.

3.4. Selected *trans*-mutant candidates - HPT locus analysis

3.4.1. HPT expression

In order to compare the reactivation of the HPT transgene in the *trans*-mutants, its expression was further analysed in three weeks old seedlings by semi-quantitative Northern blot analysis and quantitative reverse-transcription real-time PCR. Using a probe specific to the HPT coding region, the 3kb full length mRNA and the additional shorter transcripts could be detected in all *trans*-mutants, including the positive control C2R, without major differences between the samples (Figure R10A). While the HPT transcripts seemed to be expressed equally, the secondary transcript was strongly reactivated in 348, 552, 774, 1135 and 1142, and was expressed at lower levels in 409, 761, 964-1 and 964-2. This points out that the two CaMV35S promoters are differently affected in the mutant lines (Figure R10B).

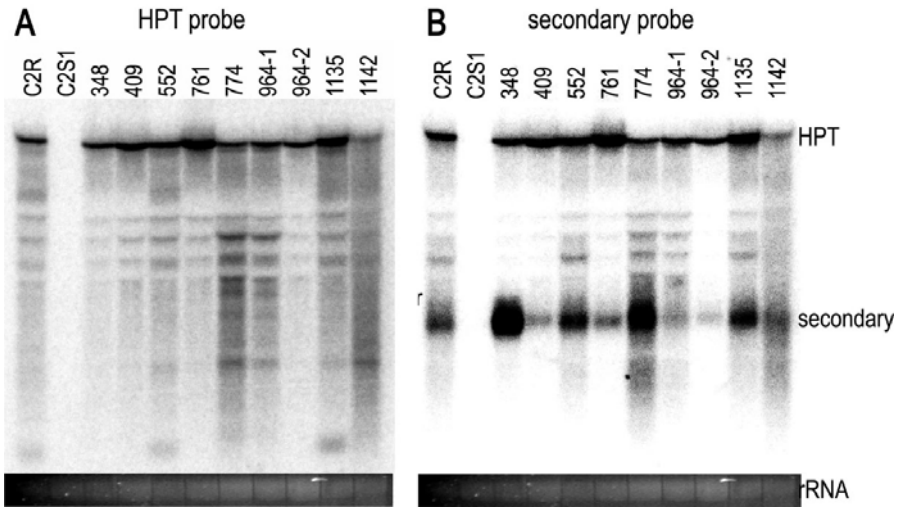

Figure R10: Northern blot analysis of HPT and secondary transcripts. **(A)** HPT transcripts are reactivated in a similar way in all mutant lines, but not in the parental line C2S1. **(B)** The secondary transcript shows uneven reactivation throughout the mutant lines.

While no major differences could be detected between the HPT transcripts in the analysed trans-mutants, real time PCR using cDNA obtained from reverse-transcribed total RNA samples was used to measure the relative abundance of HPT transcripts in the mutant lines compared to the active line C2R, normalized to the endogenous gene *EIF*4A1 (Figure R11). While no transcripts could be detected in wtZH and in the silent line C2S1, HPT was strongly expressed in 552, 774 and 1135 – at similar levels as observed in C2R. HPT expression in 1142 was even 1.5-fold stronger than in the active line. The remaining mutant lines showed weaker expression with 0.7-fold in 761, and 0.1, 0.1 and 0.17 –fold in 348, 409 and 964-2. respectively (Figure R11)

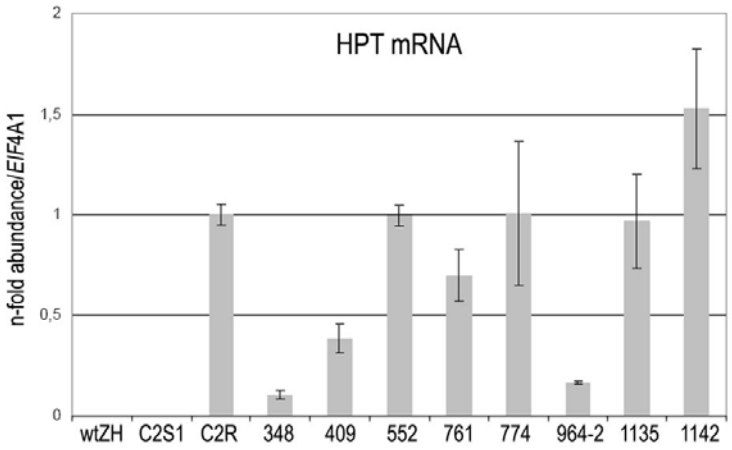

Figure R11: Relative expression of HPT in mutant and control lines compared to the active line C2R using quantitative real-time PCR. *EIF*4A1 was used for normalization. Error bars represent standard deviations from triplicate analysis.

3.4.2. DNA methylation at the HPT transgenic locus

Since the initial transgene silencing was accompanied by DNA methylation, the trans-mutants were analysed for changes in DNA methylation at the HPT locus by Southern blot analysis of individual hygromycin-resistant plants and bisulfite sequencing of pooled, three weeks-old M4 seedlings. Southern blot analysis using the DNA methylation-sensitive restriction enzymes *Hpa*II and *Msp*I and radioactive probes specific to the HPT transgenic locus showed that methylation was changed in all mutants, although to different levels (Figure R12). While the HPT locus was totally hypo-methylated in 187 and 112 (for 112, data not shown), some mutant lines were hypo-methylated to intermediate degrees (e.g. 552, 761, 774, 1135, 1142) and other lines (348, 551, 1341 and 1361) showed both, hypo- and hyper-methylation at the transgenic locus, resembling the additive pattern of C2S1 and C2R (Figure R12).

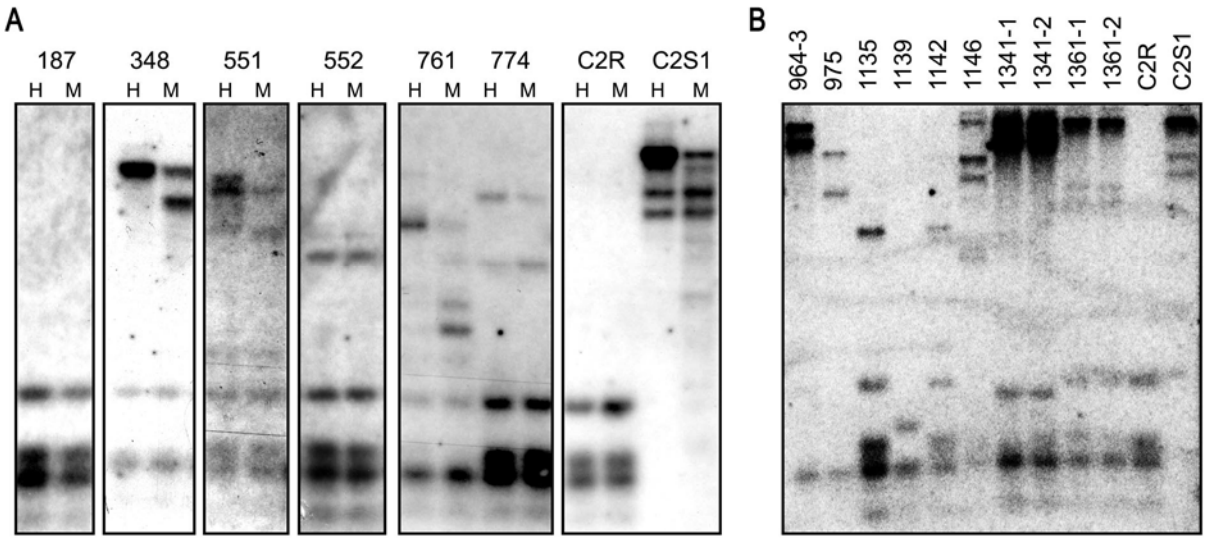

Figure R12. Southern blot analysis of DNA methylation states at the HPT transgene. **(A)** Restriction digests with *Hpa*II (H) or *Msp*I (M). **(B)** Restriction digests with *Hpa*II only.

Because of the different degrees of DNA demethylation observed in the previous results, bisulfite conversion and subsequent sequencing were applied to obtain a more detailed view of the DNA methylation around the duplicated CaMV35S promoters in the mutants and *ddm*1-5 for comparison. For each line, 7 to 10 different bisulfite clones were sequenced and analysed. Total methylation at the main 35S promoter was decreased in almost all mutants, with the most drastic decrease observed in 774 (0%). In 348 and 964-2 however, the total DNA-methylation increased to 49% and 46% respectively, compared to 29% in the C2S1 control line (Figure R13, black bars). CG-specific methylation was affected in all mutants, with most drastic demethylation from 87% in C2S1 to 1% in 774. Strong CG demethylation was also observed in the lines 409 (26%), 552 (18%), 761 (19%), 1135 (7%), 1142 (22%) and *ddm*1-5 (16%), and intermediate CG demethylation in the lines 348 (49%) and 964-2 (55%) (Figure R13, red bars). While CHG methylation levels of C2S1 (57%) were almost retained in 348 (53%), 409 (51%) and 964-2 (55%), no CHG methylation could be detected in 774 (0%). The remaining mutants, 552 (37%), 761 (28%), 1135 (18%), 1142 (32%) and *ddm*1-5 (27%) had a more intermediate effect at CHG sites. Surprisingly, asymmetric DNA methylation at CHH was strongly increased in 348 (49%) and 964-2 (42%), compared to C2S1 (13%). CHH methylation was not drastically changed in the other mutants: some slight increase in 409 (17%), 552 (24%), 761 (21%) and *ddm*1-5 (20%) was observed but not significant, while 774 was free of CHH methylation (0%). 1135 (13%) and 1142 (14%) showed no change in CHH levels (Figure R12, green bars). The duplicated CaMV35S promoter showed decreased methylation levels in all analysed mutants except for 964-2, which retained the C2S1 level. The promoter in all other mutants had reduced total methylation levels ranging from 1% in 774 and 1135 to 18% in *ddm*1-5 (Figure R13, black

bars). CG methylation at the duplicated promoter was strongly decreased in all mutants except 964-2, ranging from 1% in 774 to 33% in 409 (Figure R13, red bars). The same was observed for CHG methylation, where all mutants except 964-2 had strong demethylation ranging from 0% in 1135 to 24% in *ddm*1-5 (Figure R13, blue bars). CHH methylation was not drastically changed in 552 (12%), 761 (12%), 964-2 (22%) and *ddm*1-5 (19%) compared to C2S1 (17%), but was more affected in 348 (9%), 409 (4%), 774 (1%), 1135 (1%) and 1142 (6%) (Figure R13, green bars).

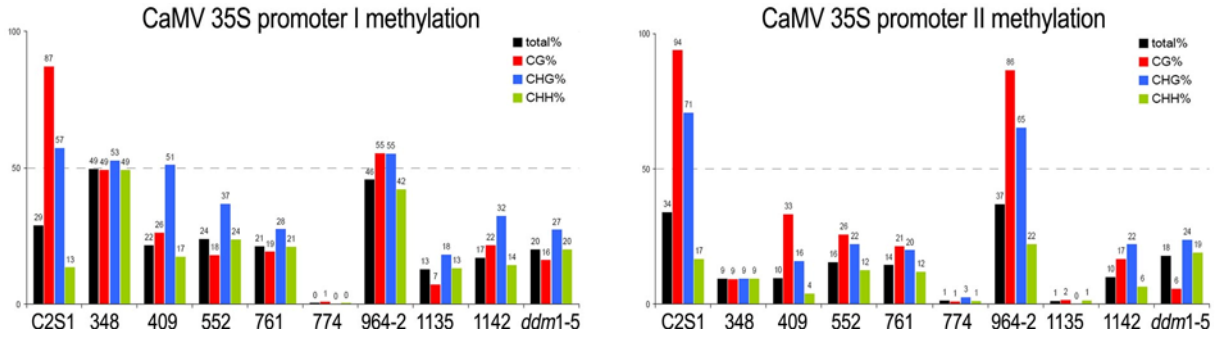

Figure R13: Percentage of methylated cytosines at the duplicated CaMV35S promoters.
Total methylation: black bars, ^{m}CG: red bars, ^{m}CHG: blue bars and ^{m}CHH: green bars.

In addition to the overall quantitative analysis, bisulfite sequencing allows also a more detailed view of the methylation state of each cytosine in the analysed regions, by disclosing the percentage of methylation at any given site and the effect of each mutation on the distribution of 5-methyldeoxycytosines. Although the methylation pattern was similar between the two promoters in the wild type, the mutations displayed diverse effects on the two promoters. While there was a strong bias towards de-methylating the 3' part of the first promoter, the duplicated promoter was demethylated in a more homogenous manner (Figure R14). Additionally, the majority of the mutants displayed increased methylation at the 5' part of the first promoter. This increase was mainly due to a strong increase in CHH methylation, especially in 348 and 964-2. The more prominent loss of CG methylation in the core promoter region was also apparent at the ASF-1 transcription factor binding sites that were significantly methylated in C2S1 (Hetzl *et al.* 2007) but less modified in all *trans*-mutants (Figure R14, asterisk). 774 and 1135 are the only mutants where methylation was strong and equally affected throughout the sequenced region of both promoters, although the first promoter in 1135 retained some residual methylation.

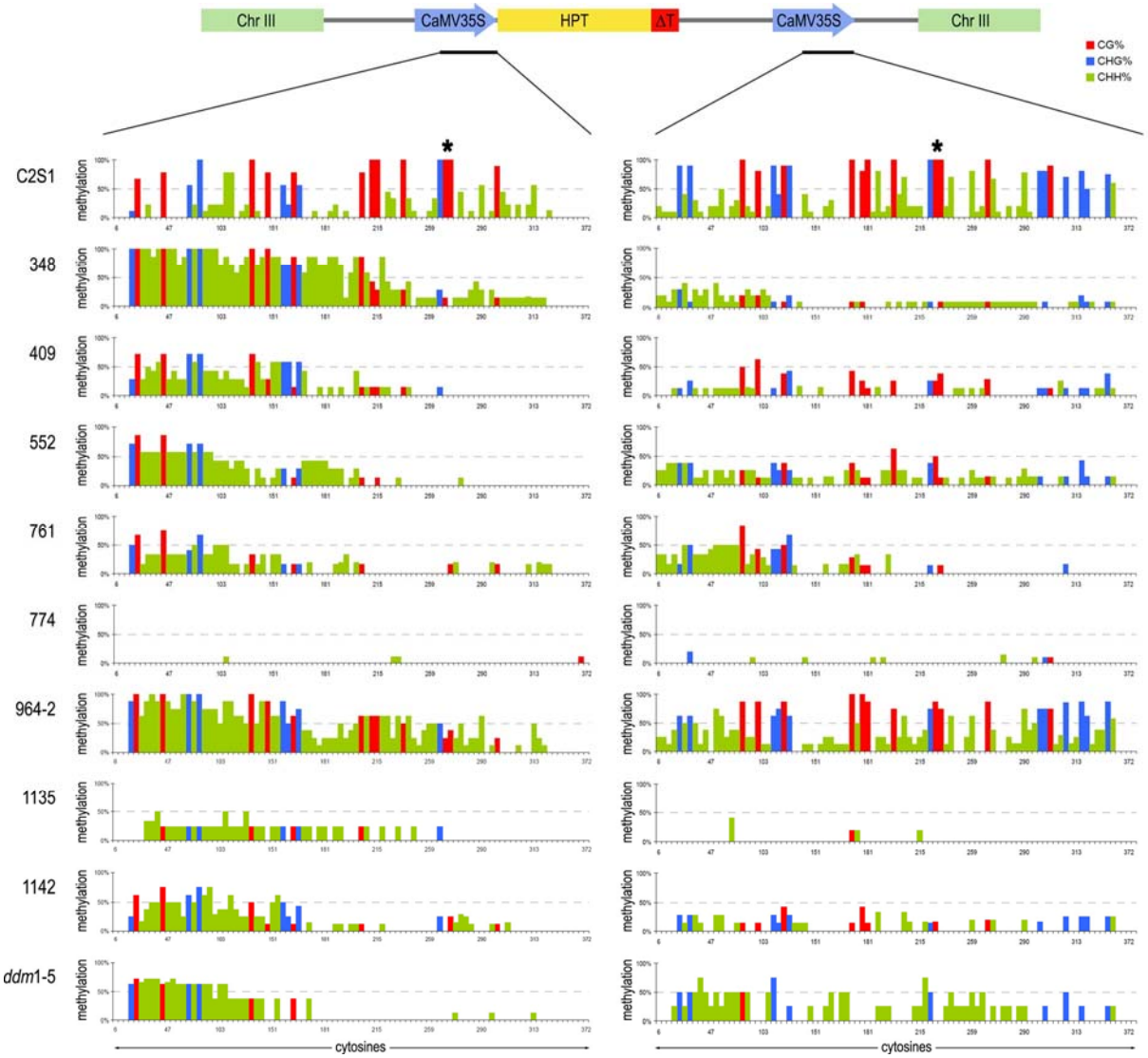

Figure R14: Percentage of methylation at individual cytosine residues.
Clone by clone analysis of each putative methylation site is showing important methylation patterns and changes at the duplicated CaMV35S promoters in C2S1 and mutant lines. Asterisk denotes the ASF-1 binding site.

3.4.3. Histone modifications at the HPT locus

Histone modifications such as methylation and acetylation at certain N-terminal residues are known to be involved in epigenetic regulation of transcription and also to be tightly interwoven with DNA methylation. It has already been shown that the HPT transgene has either heterochromatic or euchromatic marks in the C2S1 or C2R state respectively (Andrea Foerster, unpublished data). For this reason, histone modifications were analysed by chromatin immuno-precipitation (ChIP) using specific antibodies or antisera against the

heterochromatic H3K9me2 and H3K27me2 and euchromatic H3K4me3, H4 acetylation and H3K27me3 methylation-marks (Fuchs *et al.* 2006, Pfluger and Wagner 2007). In addition, antibodies against histone H3, binding independent of modifications, were used to analyse nucleosome occupancy. Selected control sequences and different regions at the HPT transgenic locus (both 35S promoters and the HPT coding region) were analysed in six mutants, C2S1 and C2R by quantitative real time PCR as triplicates and each result plotted as fold-enrichment values over the respective input. Mock-treated samples were used as negative controls (Figure R15). Detailed analysis of the main HPT promoter and the HPT coding region showed that, despite the fact that in all mutants the HPT expression was reactivated, H3K4me3 and H4acetylation levels were not completely regained to the level of C2R, except for 774. The duplicated CaMV35S promoter, however, had an increase in H3K4me3 in almost all mutants, except for 552 and 964-2. Interestingly, H4acetylation was strongly increased in 552 and 964-2, while only minor in the other lines, with exception of 774 that obtained an equal increase in both marks (Figure R15). The inactive mark H3K9me2 was fully removed in 409, 552, 774 and 1135, but 348 and 964-2 still showed residual H3K9 dimethylation at the main 35S promoter and in the HPT coding region, concomitant with results obtained from DNA methylation sequencing. H3K9me2 levels at the duplicated 35S promoter were fairly equally decreased in all mutants, but not fully removed. The heterochromatic H3K27me2 and euchromatic H3K27me3 marks did not give any clear results since the levels of detection were low in almost all analysed lines and regions. However, H3K27me2 could not be detected in C2R, 552, 774 and 1135 in all analysed regions at the HPT locus, where C2S1, 348 and 409 showed increased H3K27 dimethylation. H3K27 trimethylation at 35S promoter could only be detected in the C2S1 line whereas C2R and the mutant lines gave no signals. H3K27 trimethylation at the HPT coding region and the secondary promoter was highly variable in the analysed lines, with the exception of 552 (Figure R15). Analysis of nucleosome occupancy by measuring histone H3 abundance at the three regions showed a high abundance of H3 at the main HPT promoter in the inactive C2S1 line, whereas the mutant lines as well as the active C2R control line had lower levels of nucleosomes at the same region (Figure R15). The prominent difference between active and inactive lines observed at the main promoter was less distinct in downstream regions. Strikingly, H3 abundance was decreased in all analysed regions most strikingly in 1135, but also in 348 and 409 (Figure R15).

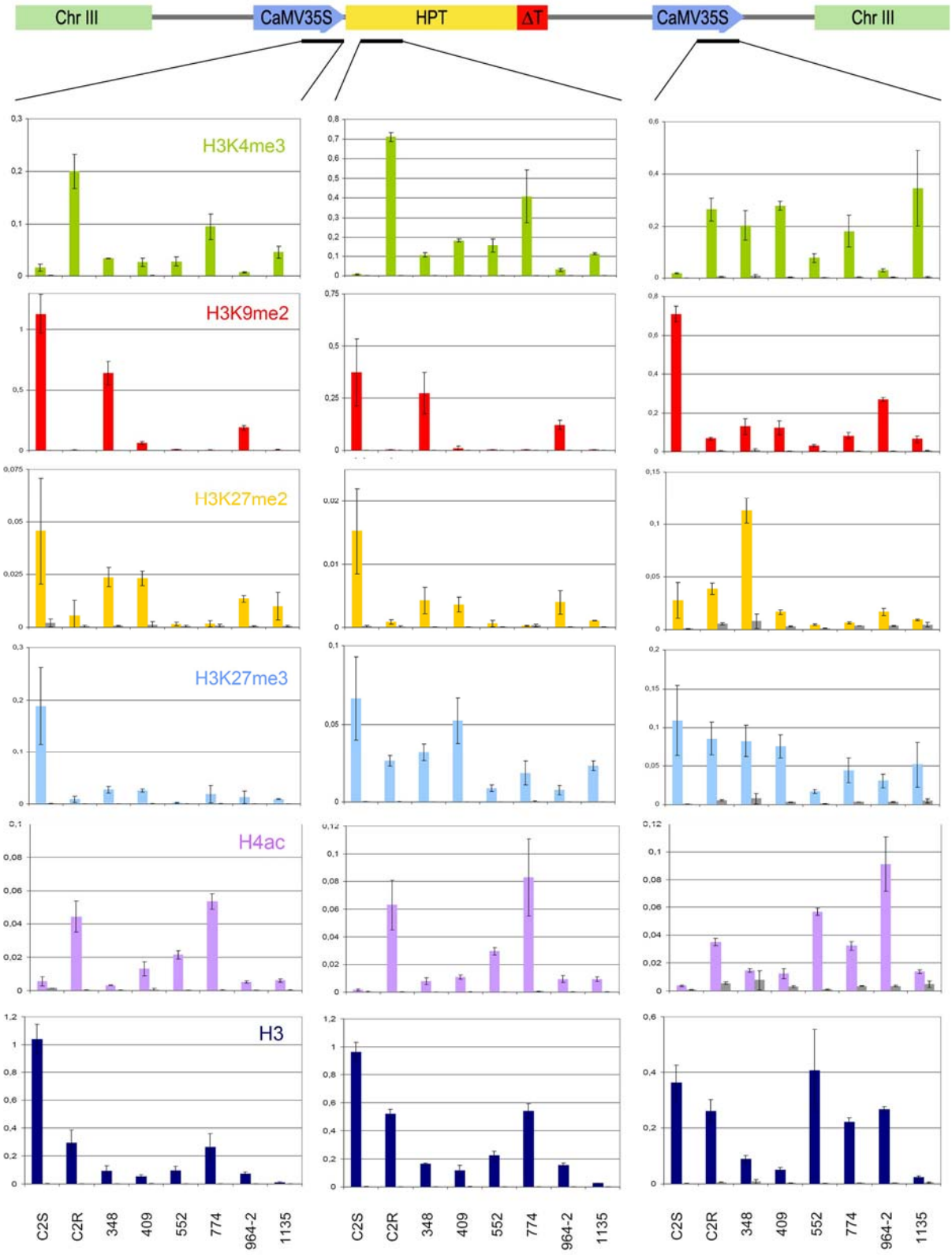

Figure R15: Chromatin immuno-precipitation of histone H3 modifications and abundance at the HPT transgene. Three different regions covering both promoters and the coding region of the HPT gene were analysed for changes in euchromatic marks (H3K4me3, green; H3K27me3, blue; H4acetylation, magenta) and heterochromatic marks (H3K9me2, red; H3K27me2, yellow). Changes in histone H3 occupancy are represented in dark blue. Results are represented as fold enrichment over input with triplicate error bars. Mock results are denoted in grey.

To study the effects of the mutations on chromatin modifications in regions apart from the transgene, the plant region immediately upstream of the HPT insertion and the gene *BME*3, located 2kb upstream of the HPT insertion site, were analyzed in order to determine any spreading effects into the upstream regions. Both regions showed fairly equal H3K4 trimethylation in C2S1 and C2R, but different levels in the mutant lines, depending on the analysed region. While the adjacent upstream region had no increase in H3K4me3 in the mutant lines, the *BME*3 region retained similar K4me3 levels like observed in C2R or C2S1 (Figure R16A and B). H3K9 dimethylation was detected in C2S1 at the HPT adjacent region and also in the *BME*3 gene, albeit at low levels, whereas this mark was diminished from both regions in C2R and the mutant lines (Figure R16A and B). H3K27 dimethylation at the *BME*3 region was detected only at very low levels and therefore not taken into account, while K27 dimethylation in the adjacent upstream region resembled the same pattern observed at the HPT promoter and coding region (Figure R16A and R15, respectively). H3K27 trimethylation was strongly decreased in all mutant lines including the active C2R line at the adjacent region, while showing increased accumulation in C2S1. At *BME*3 however, a strong H3K27me3 decrease could only be detected in 552, 964-2 and 1135 whereas the other lines had a minor decrease compared to the levels detected in C2S1, that were not significantly different from C2R levels (Figure R16A and B).

The third sequence studied was the non-LTR retrotransposon *LINE*1-4, located at a different chromosome It is usually epigenetically silent, therefore enriched in heterochromatic marks and devoid of euchromatic marks (Huettel *et al.* 2006). However, although three mutants (348, 409 and 964-2) showed no H3K4 trimethylation and were still enriched for H3K9 and H3K27 dimethylation as in C2S1 and C2R wild type, the mutants 552, 774 and 1135 showed an inverse pattern, arguing for loss of silencing at endogenous regions in the *Arabidopsis* genome in these mutants (Figure R16C).

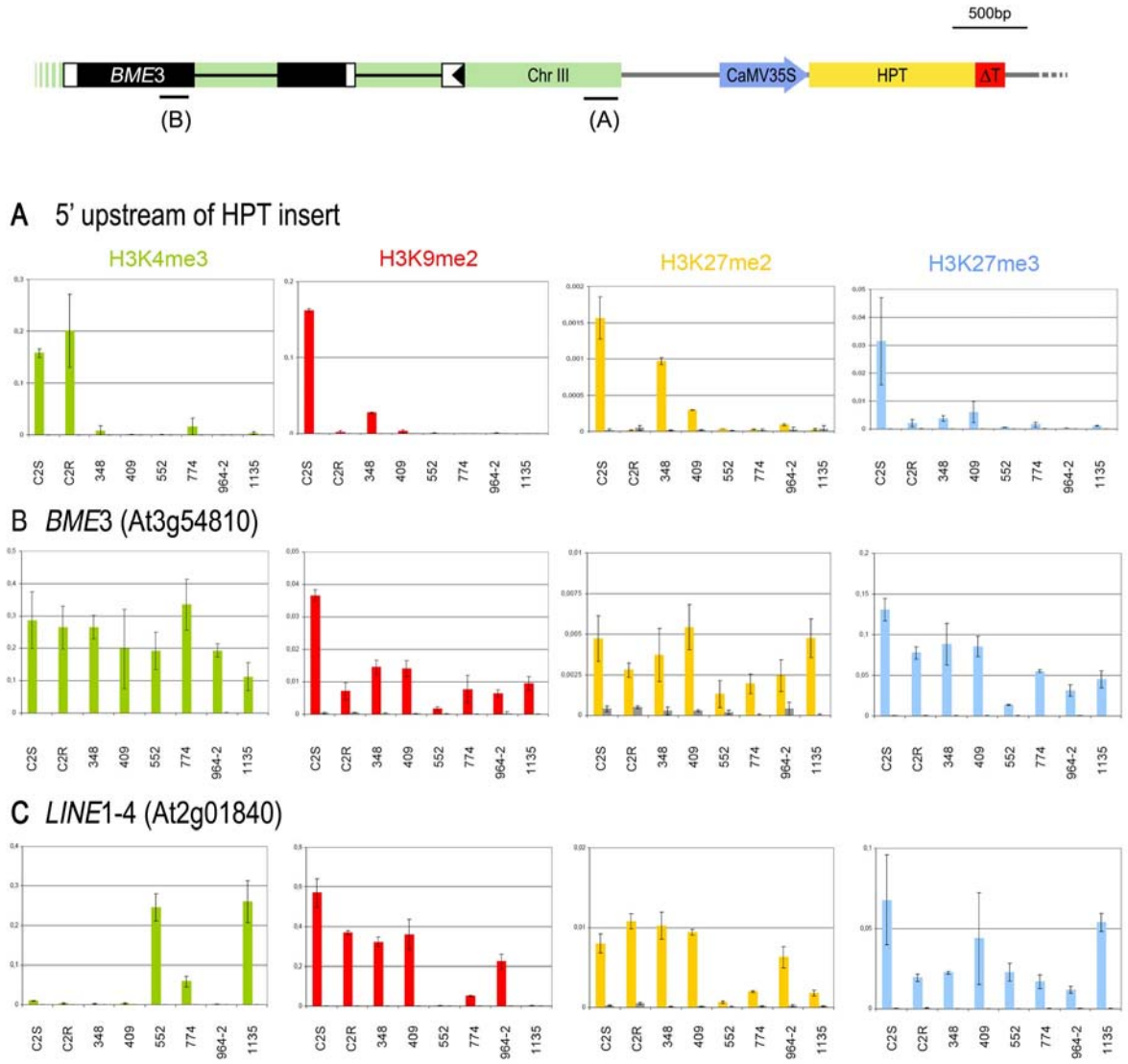

Figure R16: Chromatin immuno-precipitation of histone H3 methylation marks at regions upstream (positions indicated) and outside of the transgene.

(A) Plant genomic region immediately upstream of the HPT insertion, (B) *BME*3, a gene upstream of the HPT locus and (C) *LINE*1-4, a non LTR retrotransposon were analysed for changes in euchromatic marks: H3K4me3 (green), H3K27me3 (blue) and heterochromatic marks: H3K9me2 (red), H3K27me2 (yellow). Results are represented as fold enrichment over input with triplicate error bars. Mock results are denoted in grey.

3.5. Selected *trans*-mutant candidates - global analysis

Besides the detailed analysis of the mutant effects on the HPT transgene, it was of interest to search for additional effects changing the plants in a more global manner.

3.5.1. Phenotype analysis

The most obvious question was to score for changes in growth and morphology caused by the mutations. Therefore, plants were germinated and grown on soil. Growth and morphology were evaluated each two weeks for a period of 6 to 8 weeks in total (Figure R17A and data not shown). C2S1 was used as control. No obvious phenotypic changes could be detected in 348, 409, 761, 964-1 or 964-2 during this analysis. The other tested lines showed more or less drastic differences compared to C2S1. Individuals of mutant 187 were growing slowly and resulted in weakly developed, short plants with increased sterility and only a few siliques. Mutant 774 plants had a prolonged rosette stage with increased leaf size (data not shown), but finally produced normally developed and fertile plants. 552, 1135 and 1142 plants had comparable phenotypes with early bolting and flowering and developed into thin and highly sterile mature plants. 552 had, in addition, an increased flower size compared to all other lines (data not shown). Additional examination of seedlings grown in tissue culture showed that some 964-1 seedlings exhibited a lack of chlorophyll evident as segregating white plantlets in F3 populations after outcrossing to wtZH (Figure R17B), while 774 seeds had a low germination rate (Figure R17C). Seed size analysis showed that all seeds are fairly equal to that of C2S1 except for 552 where the seed size is as big as seeds from tetraploid plants, arguing for increased ploidy in the 552 mutant plants (Figure R17D).

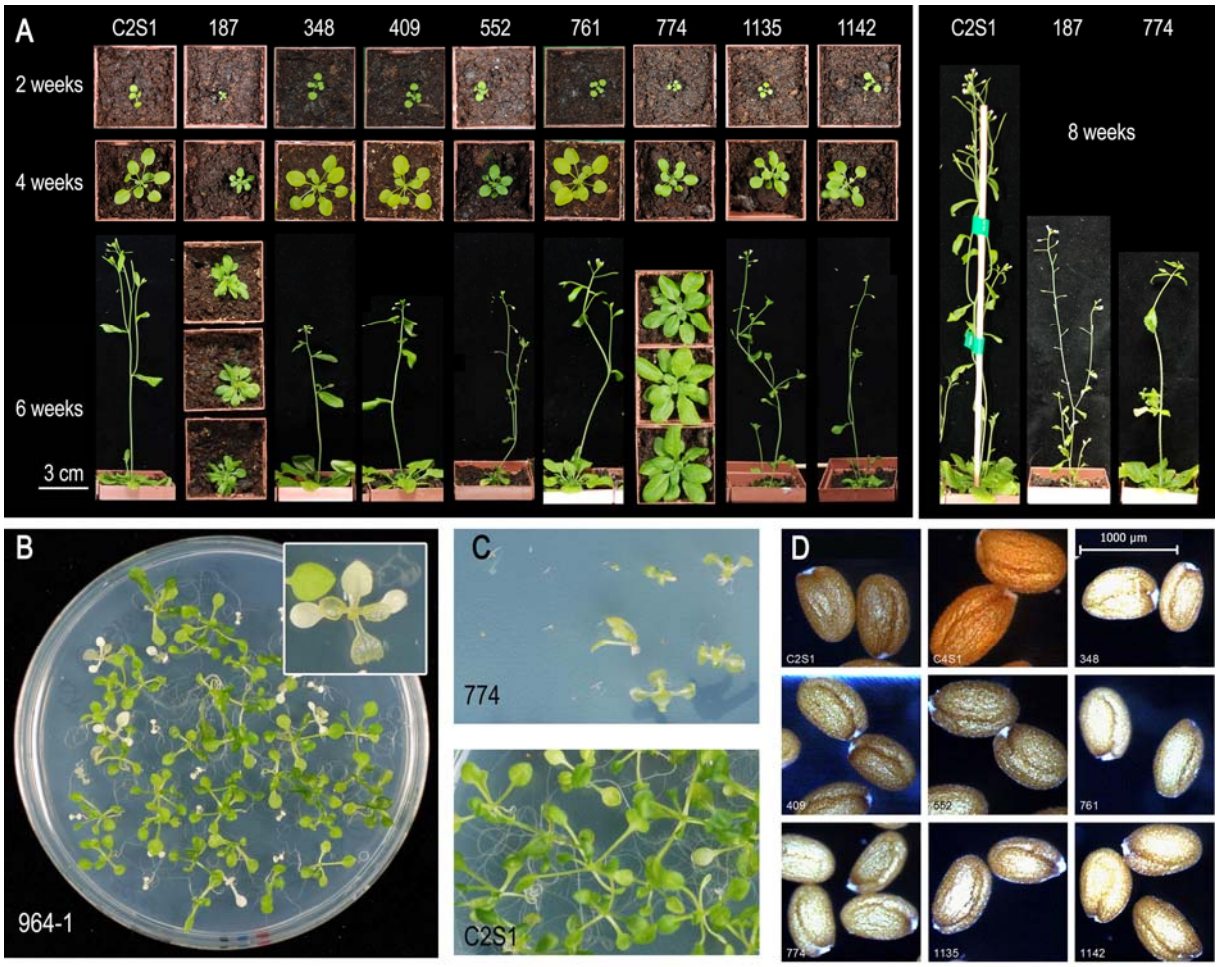

Figure R17: Phenotypic analysis of mutant plants.
(A) Growth and development. Plants were grown for 6 or 8 weeks on soil and pictures were taken every two weeks. (B) Segregating lack-of-chlorophyll phenotype observed in 964-1. **(C)** Low germination rate in 774 **(D)** Seed size comparison between the mutant lines, diploid C2S1 and tetraploid C4S1.

3.5.2. Ploidy analysis

Since 552 plants had increased flower and seed sizes, the ploidy of all mutant lines was tested using flow cytometric analysis of nuclei extracted from rosette leaf tissue. Although *Arabidopsis thaliana* is a diploid plant, additional peaks with more than 2C can be detected. This is due to nuclei after S-phase and to endoreplication that is common in most of the plant tissues. However, the lowest peak is always representing the basic ploidy. Diploid (Zh2n) and tetraploid (Zh4n) plants were used as references and three independent plants were analysed for each mutant line. All analysed plants were diploid except for 552 which was tetraploid, as expected according to the increased seed size (Figure R18).

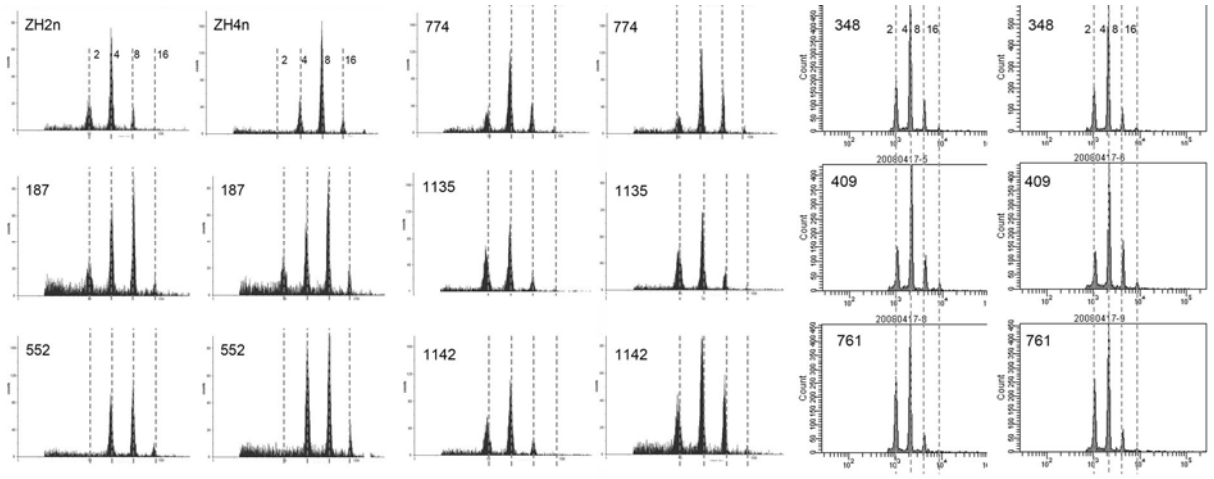

Figure R18: Ploidy analysis

Diagrams with a logarithmic x-axis are shown for two independent plants per line. Peaks represent the number of nuclei with a certain C-value. The lowest peaks represent the basic ploidy of the analysed plant. Higher values result from endoreplication in leaf tissue.

3.5.3. Global 5-methyldeoxycytosine levels

To investigate the effect of the mutations on the global levels of 5-methyldeoxycytidine (5-mdC), mutant plants were compared with wild type C2S1 and wtZH plants, and with plants in which DNA methylation was reduced by genetic means. Mutations in the *ddm*1 gene drastically decrease the level of 5-mdC (Jeddeloh *et al.* 1999, Vongs *et al.* 1993). Plants were germinated and grown for 21 days in tissue culture prior to preparation of genomic DNA. Global 5-mdC levels were analyzed as percentage of 5-mdC in relation to total deoxycytidine (dC) levels using cation exchange HPLC (Rozhon *et al.* 2008). 5mdC levels of 5.91% (±0.43) for 348, 5.49% (±0.15) for 409, 5.50% (±0.23) for 761, 5.51% (±0.38) for 964-1 and 5.67% for 964-2 were similar to C2S1 with 5.92% (±0.5) and wtZH with 5.93% and therefore unchanged (Figure R19). The mutant lines 552, 774, 1135 and 1142 had severely decreased 5-mdC levels, with 1.59% (±0.58), 2.67% (±0.47), 1.74% (±0.13) and 1.82% (±0.15) respectively, suggesting that these mutations are affecting major epigenetic pathways that regulate DNA methylation throughout the genome. Their decrease in global methylation was even more drastic than in the *ddm*1-5 mutant with 4.21% (±0.39) (Figure R19).

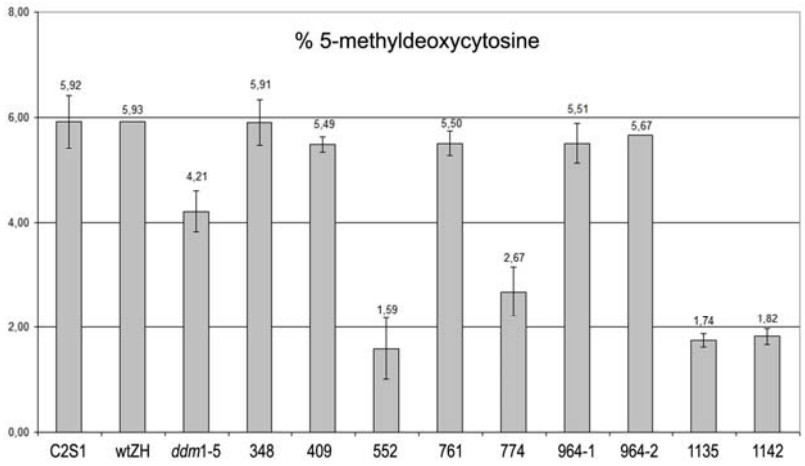

Figure R19: HPLC measurement of global 5-mdC percentage in 3 weeks old seedlings from mutant and different control lines. Error bars denote standard deviations from triplicate analysis.

3.5.4. DNA methylation at repeated regions

The changes in methylation at the HPT transgene and the decrease in global 5-mdC levels in some mutant lines suggested an analysis of repetitive regions that are known to be highly methylated in wild type plants and hypomethylated in DNA methylation-deficient mutants such as *ddm*1 or *met*1 (Vongs *et al.* 1993). Such regions were the Athila-related element TSI (Transcriptionally Silent Information) and the centromeric 180bp repeats. Genomic DNA of the mutants was analysed for changes in DNA methylation by Southern blots after digestion with methylation-sensitive enzymes. TSI and 180bp hypomethylation was evident in several lines: 187, 552, 774, 1135 and 1142 (Figure R20). This is in accordance with the same lines showing either strong decrease in global 5-mdC levels, growth abnormalities or strong hypomethylation of DNA and decreased H3K9me2 at the HPT transgene. The degree of demethylation at the repeats observed in 774 was rather weak, compared to 552, 1135 and 1142 where it was even stronger than in the mutant line *ddm*1-5. All other analysed lines showed no significant demethylation at the analysed regions (Figure R20).

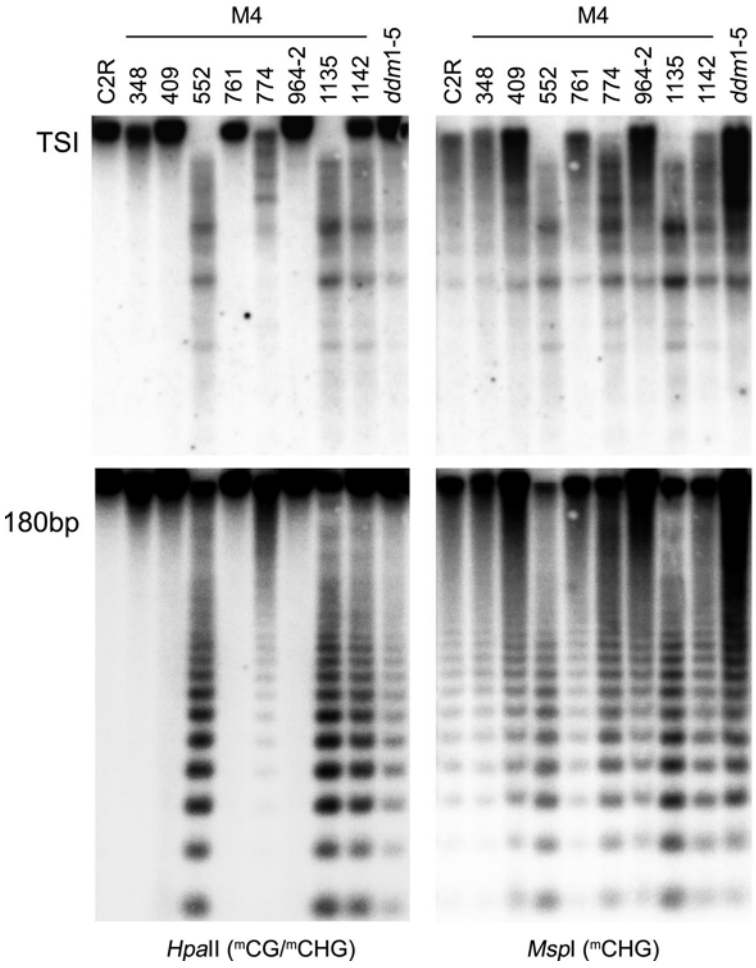

Figure R20. Southern blot analysis of DNA methylation states at endogenous hypermethylated repetitive regions. Restriction digests were done with *Hpa*II (H) or *Msp*I (M). TSI: transcriptionally silent information, Athila-related transposons (Steimer *et al.* 2000). 180bp: centromeric repeats (pAL, Martinez-Zapater *et al.* 1986).

3.5.5. Chromocenter morphology

The compact and intensively stained chromocenters, formed mainly by extreme condensation of the large centromeric and rDNA-coding heterochromatin regions in wild type nuclei, are drastically affected and dissociated upon microscopic examination in numerous epigenetic mutants like *ddm*1 and *suvH*2, with a defect in a H3K9 methyltransferase (Naumann *et al.* 2005). Therefore, nuclei from 3 weeks old seedlings from the *trans*-acting mutants were analysed for loss of chromatin compaction at the 180bp centromeric and the 45S rDNA regions by fluorescence *in situ* hybridisation (FISH). The 180bp centromeric repeats were, as expected from the previous methylation studies, strongly affected in 552, 774, 1135, 1142 and *ddm*1-5. In these lines, just a few nuclei retained intact chromocenters, while the majority had either intermediate or strongly dispersed 180bp chromocenter morphology (Figure R21A). No significant increase in nuclei with dissociated chromocenters was found in 348 or 409. The analysed rDNA regions did not show any obvious changes in any of the analysed mutants and were not further investigated.

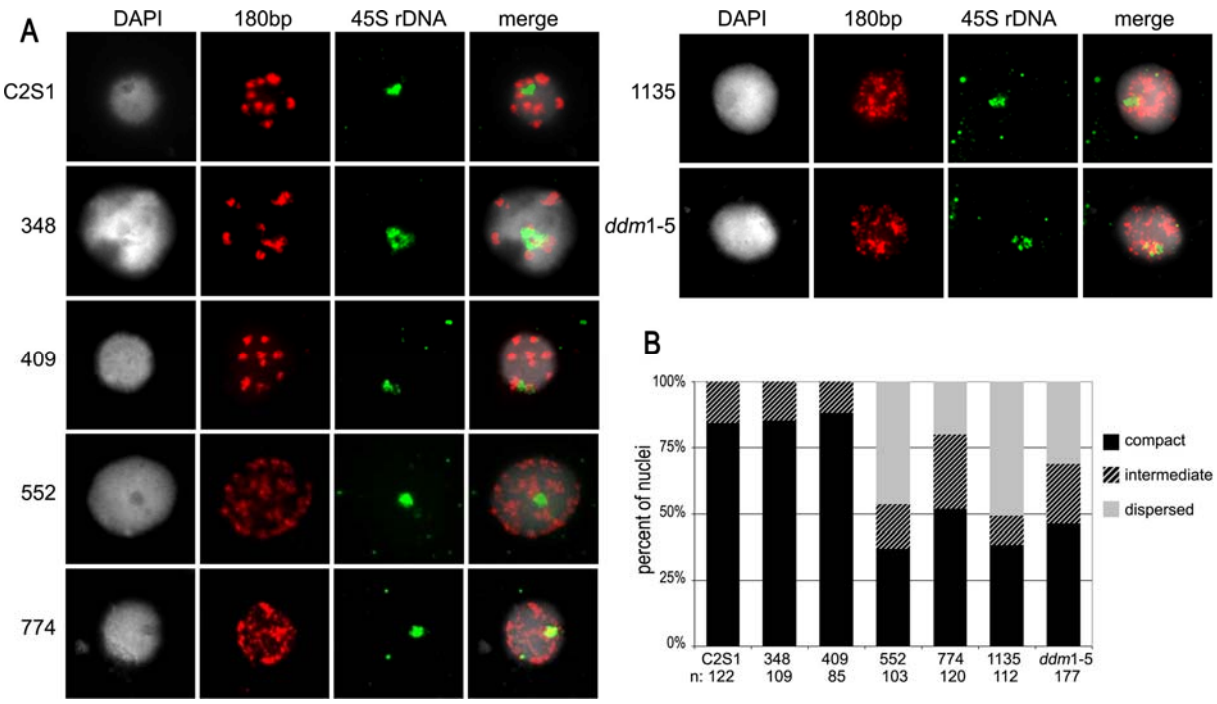

Figure R21: Chromocenter morphology

(A) Fluorescence in situ hybridisation with 180bp (red) and 45S rDNA (green) probes on nuclei extracted from three weeks old seedlings. **(B)** Quantitative evaluation of nuclei with abnormal chromocenters morphology showing either compact, intermediate or dispersed 180bp FISH signals.

3.5.6. Localisation of nuclear 5-methyldeoxycytosine

Since global and locus-specific DNA methylation was affected in several mutants, immuno-fluorescence detection was carried out to analyse the nuclear localisation patterns of 5-mdC in three weeks old seedlings of selected mutants affected in DNA methylation. Nuclei from 761 and 774 showed no changes in 5-mdC localisation (Figure R22A). In both lines the 5-mdC signals were concentrated at the chromocenters, indistinguishable from those in the control line C2S1 and in wild type nuclei analyzed in previous reports (Probst *et al.* 2003) although minor decondensation was observed in 774. Nuclei from 552, 1135 and 1142 showed strong loss of 5-mdC staining from chromocenters (Figure R22A), concomitant with DNA methylation data obtained from HPLC and Southern blot experiments with repetitive probes. The percentage of nuclei with wild type–like 5-mdC signals at chromocenters is shown in Figure R22B.

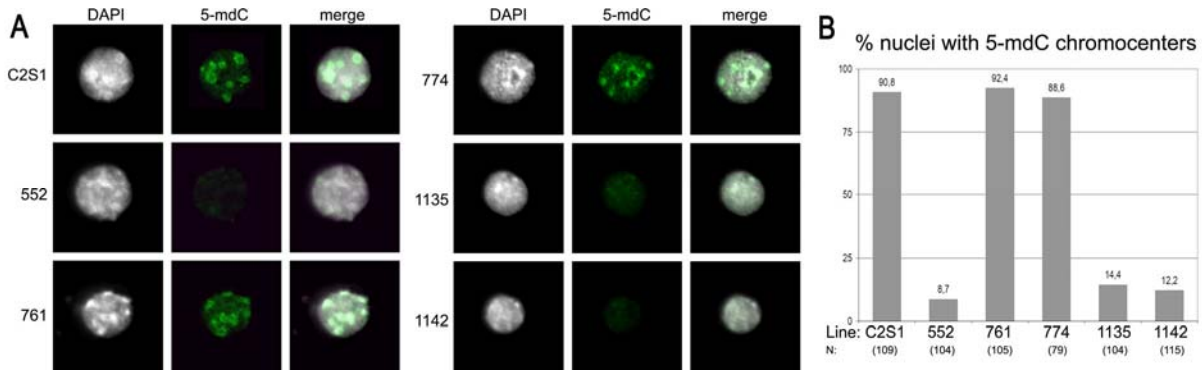

Figure R22: (A) 5-methyldeoxycytosine immuno-fluorescence. **(B)** statistical analysis showing percentage of nuclei with 5-mdC signals at chromocenters. N: number of nuclei analysed.

3.5.7. Nuclear histone modification patterns

After analysing chromocenter morphology and distribution of nuclear 5-mdC in the mutants, patterns of the histone modifications H3K9me2 and H3K27me3 were investigated in mutant nuclei using specific antibodies. H3K9me2 is known to be a heterochromatic mark and associates with DAPI-dense heterochromatic chromocenters, whereas H3K27me3 is found in the surrounding euchromatin (for review Fuchs *et al.* 2006, Pfluger and Wagner 2007). Heterochromatic H3K9me2 was found to co-localize with chromocenters in almost all nuclei analysed from 348, 409 and the parental line C2S1, but this mark was strongly dispersed in 552, 774 and 1142, concomitant with results obtained from 180bp FISH and 5-mdC staining (Figure R23A). The euchromatic H3K27me3 staining could be observed as homogenous staining throughout the nucleus (excluding chromocenters and nucleoli) in all analysed samples from C2S1, 348 and 409. In 552, 774, 1135 and 1142 where most of the chromocenters were dissociated, the H3K27me3 staining was spread throughout the whole nucleus (except the nucleoli) (Figure R23B). Additionally H3K27me3 aggregates that were not evident in C2S1 accumulated in nuclei from 552, 774, 1135 and 1142 (Figure R23B, arrowhead and Table R3).

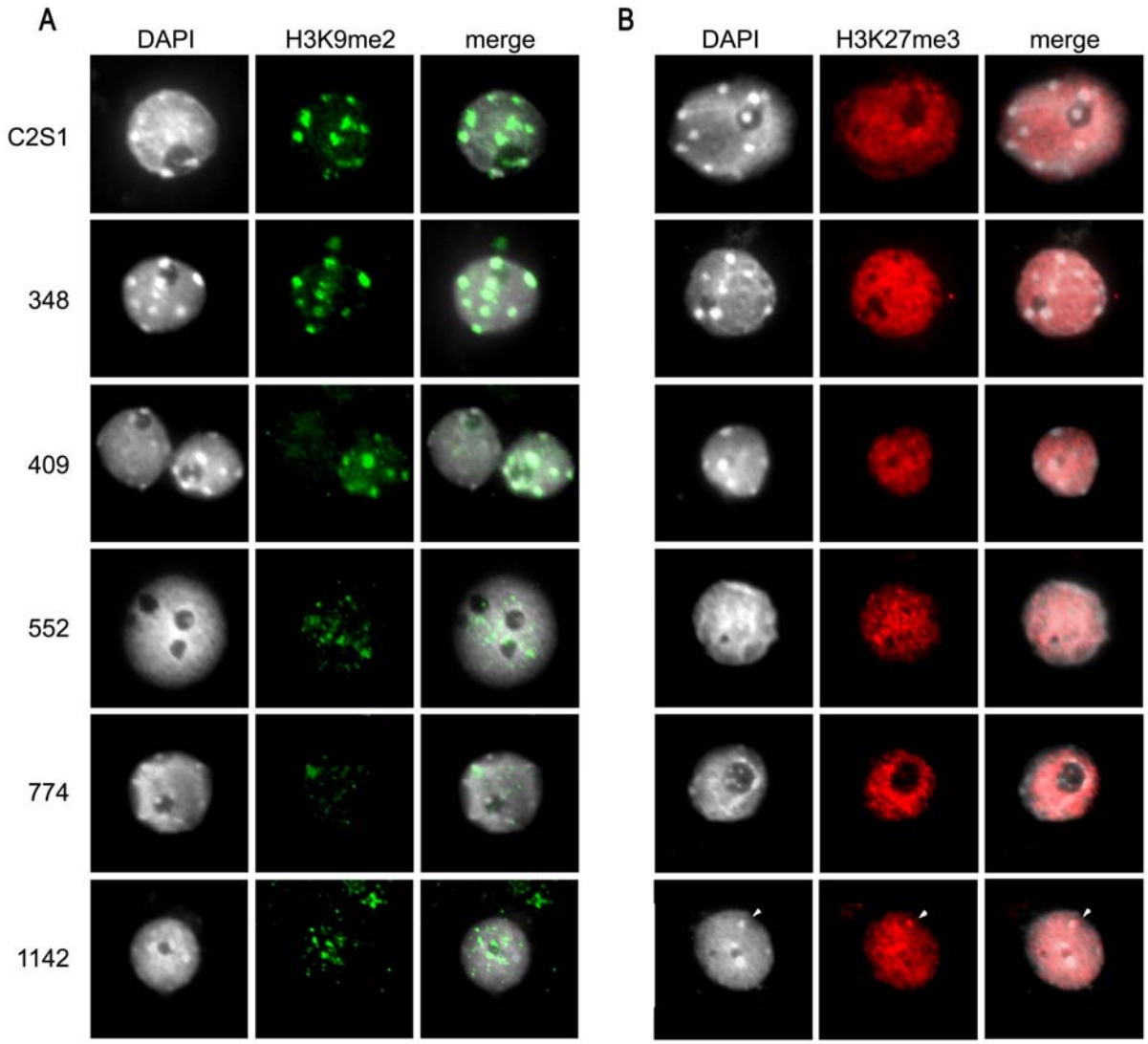

Figure R23: Histone methylation immunofluorescence on nuclei extracted from three weeks old seedlings. **(A)** H3K9 dimethyl staining (green), **(B)** H3K27 trimethyl staining (red).

Table R3: Histone methylation immunofluorescence

	compact H3K9me2	H3K27me3 foci
C2S1	71.3% (n:129)	12.2% (n:181)
348	72.4% (n:116)	n.d.
409	60% (n:135)	n.d.
552	1.6% (n:64)	56.2% (n:178)
774	9.6% (n:114)	32.5% (n:200)
1142	7.5% (n:107)	61.8% (n:207)
1135	n.d.	64.3% (n:258)

3.5.8. Transposon reactivation

Perturbation of DNA and histone methylation leads to upregulation from transcriptionally silent endogenous elements such as transposons and repeats (Hirochika *et al.* 2000), To test for such a trans-acting reactivation effect in the mutants, transcriptional activity of the *Athila*-related TSI repeats was analyzed in total RNA extracted from 3 weeks old *trans*-mutants and control seedlings by Northern blot analysis. Concomitant with previous results, TSI reactivation was observed in 187, 552, 774, 1135, 1142 and *ddm*1-5. No TSI transcripts could be detected in the control lines and the remaining mutant lines (Figure R24A). Since some genes involved in regulation of several Arabidopsis transposons are known (Lippman *et al.* 2004), additional targets were analysed in the *trans*-mutants trying to identify pathways that were deregulated by the mutations. Mutant RNA was subjected to reverse transcription PCR with primers specific to the DNA transposons *MULE* and *CACTA*-like, the non-LTR retrotransposon *LINE*1-4, the retrotransposon *Ta*3 and the LTR element *soloLTR* (Figure R24B). EIF4A was used as reference gene not affected by the mutations. The LTR element *soloLTR* that is regulated by the RNA-dependent DNA methylation (RdDM) pathway was not active in any of the analysed mutants, except for *ago*4 serving as positive control (Huettel *et al.* 2006). The DNA transposons *MULE* and *CACTA*-like and the retrotransposons *LINE*1-4 and *Ta*3 were strongly reactivated in 552, 774, 1135 and 1142 (Figure R24B).

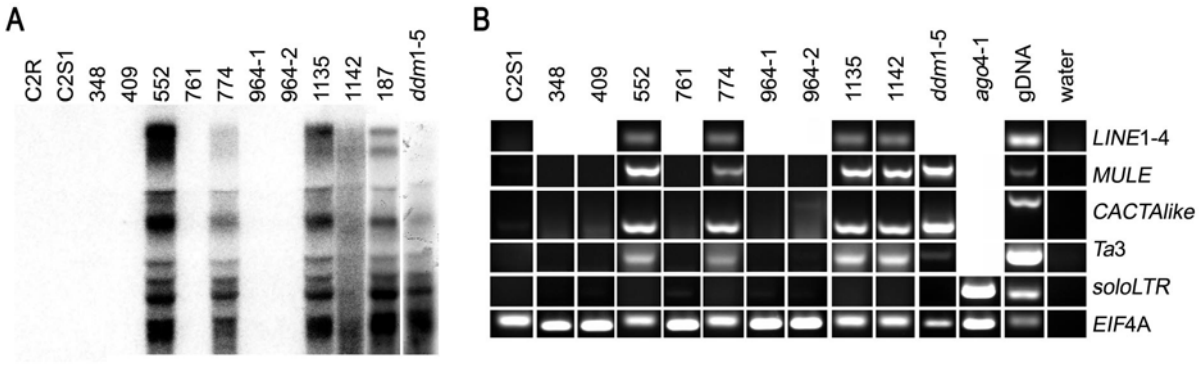

Figure R24: Transposon reactivation.
(A) Northern blot analysis of TSI expression in mutant and control lines. **(B)** Reverse transcription PCR to detect *LINE*1-4, *MULE*, *CACTA*-like, *Ta*3 and *soloLTR* transposons in mutant and control lines. *EIF*4A was used as unaffected reference gene.

3.5.9. Transgene reactivation (TS-GUS)

The reactivation of endogenous transposons, in addition to the HPT reactivation, prompted to analyze the potential of the *trans*-mutants to reactivate another silent, highly repetitive β-

glucuronidase (GUS) transgene. The line carrying this reporter (TS-GUS, L5, former 6b5; (Morel *et al.* 2000)) is reactivated by numerous epigenetic mutants (Elmayan *et al.* 2005) and by treatment with DNA methylation inhibitors (Baubec *et al.* 2008). Mutant plants were crossed to TS-GUS plants and the segregating F2 generations were analysed for GUS reactivation by enzymatic staining of 3 weeks old seedlings. GUS-expressing lines were then allowed to grow on soil and different tissues were analysed for the distribution of GUS staining patterns in adult plants. While no GUS staining except for a few single cells could be detected in the parental TS-GUS line and in F2 plants from crosses to 348 (Figure R25A and B), 60% of the seedlings from crosses to 409 displayed an increased frequency and larger size of GUS-expressing sectors in cotyledons but not in adult tissues (Figure R25C to E). Strong staining was observed in 18.7% of the analysed F2 seedlings from crosses with 187 (Figure R25F). However, an additional 52.1% had intermediate staining affecting only trichomes or leaf margins (Figure R25G and H). The different staining pattern was also observed in leaf tissue from individual adult F2 plants (Figure R25I and J). F2 populations obtained from TS-GUS crosses to 552 resulted in 21.7% fully stained seedlings and adult plants (Figure R25K to M). In the analysis of these crosses, the tetraploidy of 552 needs to be considered. F2 seedlings from TS-GUS crosses to 774 yielded 4.8% positive plants with either full staining of single leaves, marginal staining including trichomes or staining of just one half of the leaf (Figure R25N and O). The crosses with 1135 yielded 24.5% fully stained seedlings and adult leaf tissues (Figure R25P to R) and 22.4% in 1142 (Figure R25S to U).

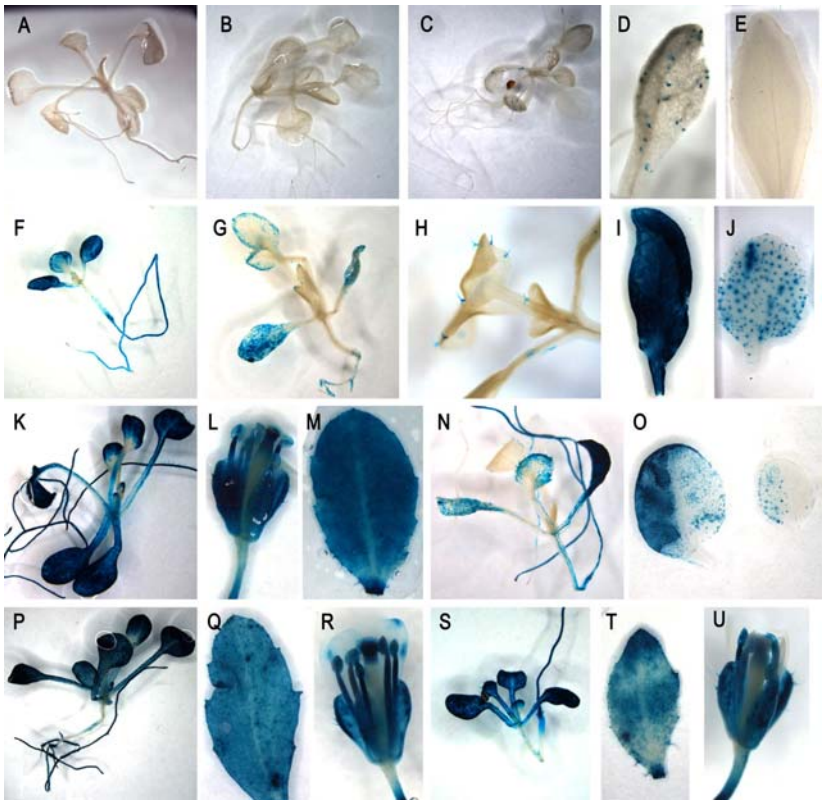

Figure R25: TS-GUS reactivation in F2 seedlings and adult plants from crosses to *trans*-mutants. **(A)** TS-GUS, **(B)** 348 seedling, **(C and D)** 409 seedlings, **(E)** 409 adult leaf, **(F to H)** 187 seedlings and **(I and J**) adult leafs, **(K)** 552 seedling and **(L and M)** adult organs, **(N and O)** 774 seedlings, **(P)** 1135 seedling and **(Q and R)** adult organs, **(S)** 1142 seedling and **(T and U)** adult organs.

F2 plants with expression of the TS-GUS transgene were selfed and the resulting F3 seedling pools were tested for TS-GUS activity (Table R4).

Table R4: TS-GUS reactivation in F2 and F3 populations

	F2 pools		F3 pools
	% GUS positive	n seedlings	GUS activity
187	16.7 + 52.1	78	++
348	0	100	nd
409	60	25	-
552	21.7	115	++
774	4.8	164	+
1135	24.5	188	++
1142	22.4	67	++

3.5.10. Sensitivity to interference with DNA methylation and histone deacetylation

In addition to the DNA methylation inhibitor zebularine, also the histone deacetylase (HDAC) inhibitor trichostatin-A (TSA) induces loss of gene silencing (Chang and Pikaard 2005, Lawrence *et al.* 2004), although hygromycin resistance in C2S1 is obviously not induced as described in 3.1.3. In order to detect additive effects of the mutations and drug-induced decrease in DNA methylation or HDAC activity, seeds from selected *trans*-acting mutant lines, parental line C2S1 and *ddm*1-5 were germinated and grown on 40 µM zebularine and/or 1.8 µM trichostatin-A and compared with mock-treated plants after four weeks. Zebularine treatments induced growth retardation in all treated plants, as previously described (Baubec *et al.* 2008). The growth inhibition was however more pronounced and in most cases lethal in 187 (data not shown), 348, 409, 552, 761 and 774 compared to C2S1 (Figure R26). Seedlings from 964-1 and 964-2 were affected only in some plants, which may be due to segregation. Plants of 1135 and 1142 were similar to *ddm*1-5 mutant plants, with minor developmental effects and viable. Trichostatin-A treatments had no effect in most of the mutants, since all plants developed normally and were even bigger than the mock treated samples. However, 348, 552 and 774 seedlings were strongly affected by this treatment (Figure R26). Combined treatment of zebularine and trichostatin-A was lethal in all cases, with 348, 409, 761 and 774 seedlings being most sensitive.

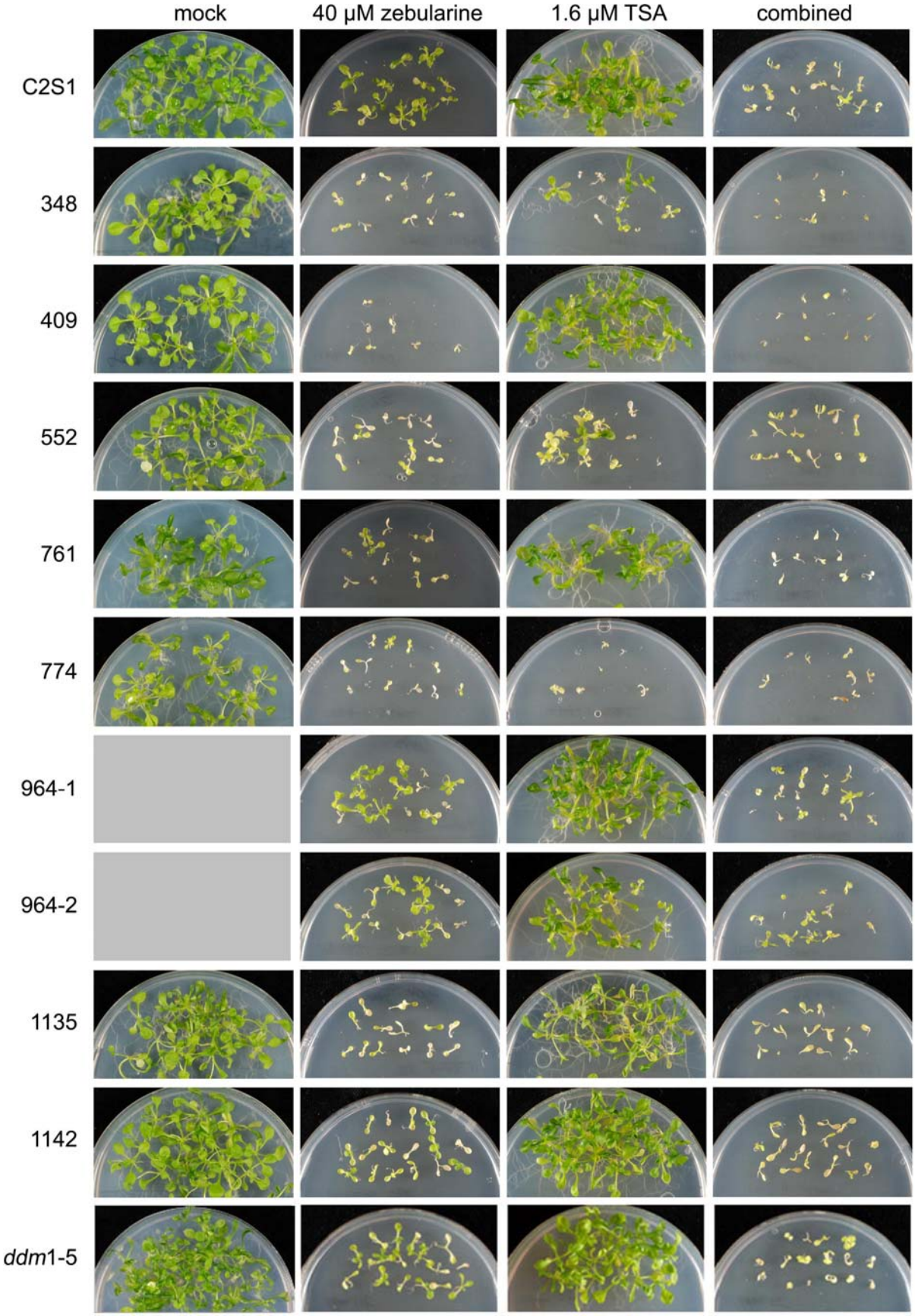

Figure R26: Sensitivity to zebularine and/or trichostatin-A treatments.

Mutant lines were grown for 4 weeks on media containing 40 µM zebularine and/or 1.8 µM trichostatin-A and were compared to treated control lines (C2S1 and *ddm*1-5) and mock-treated sibling-plants.

Upon a repetition of the experiment, lines showing strong growth effects due to zebularine or trichostatin-A treatments were rescued after 14 days and transferred onto drug-free medium to survey recovery. No recovery was observed in 187, 348, 774, 964-1 and 964-2 after zebularine treatment. Seedlings of 409 and 552, however, showed partial recovery after the same treatment. Four lines (348, 409, 552 and 774) could recover from two weeks of trichostatin-A treatment alone or in combination with zebularine. Plants of 552 and 774 recovered partially while 348 and 409 did not (Table R5).

Table R5: recovery on drug-free medium after 14 days treatment

	40 µM Zebularine	1.8 µM Trichostatin-A	combined treatment
C2S1	yes	yes	n.d.
187	no	n.d.	n.d.
348	no	yes	no
409	partial	yes	no
552	partial	yes	partial
774	no	yes	partial
964-1	no	n.d.	n.d.
964-2	no	n.d.	n.d.

3.6. Identification of mutant genes

3.6.1 T-DNA insertion analysis

Since the C2S1 lines were mutagenized using *Agrobacterium tumefaciens*-mediated T-DNA transfer and selected by BASTA treatment, expecting insertional mutants and gene identification via characterization of the flanking sequences (Koncz *et al.* 1989), the number of T-DNA insertions in each mutant was analysed by Southern blot analysis on *Hind*III or *Psi*I fragmented genomic DNA with probes specific to the BASTA resistance gene encoded by the T-DNA (Figure R27A, green bar). While only three lines had a single insertion and thirteen had multiple insertions, five had no detectable T-DNA signals (Table R6). In order to monitor for additional insertions from the binary vector used for the transformation, the membranes were rehybridized with a probe covering the whole 10.9kb T-DNA vector (Figure R27A). Except for some crosshybridisation to the HPT transgene locus caused by a 370bp sequence similarity between the two vectors (Figure R27A, blue bar), numerous additional insertion sites could be detected in eight mutant lines, including three lines that were T-DNA free in the previous hybridization (Figure R27B, arrowheads and Table R6).

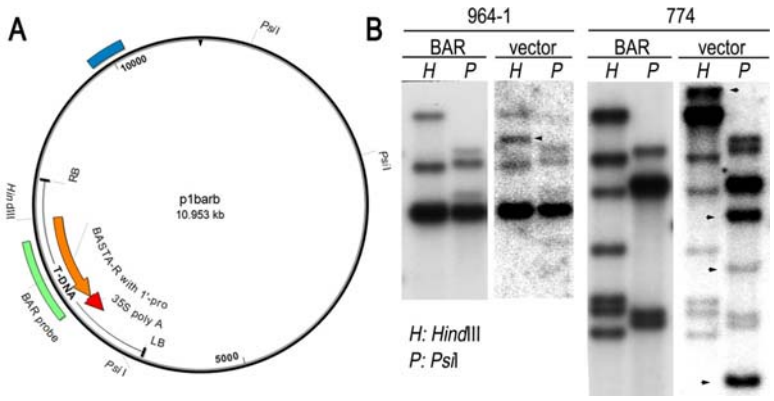

Figure R27: (A) T-DNA construct used for mutagenesis. BASTA resistance gene with 1' promoter and 35S polyA region: orange and red arrows. BAR-specific probe: green bar. HPT transgene cross-hybridisation: blue bar. RB and LB: right and left T-DNA border. (B) Two examples from Southern blot analysis showing results using either BAR-specific probes or whole vector probes. Arrowheads denote binary vector insertions.

Table R6: T-DNA and vector insertions found by Southern blot analysis

Line	*Hind*III (LB)		*Psi*I (RB)		Line	*Hind*III (LB)		*Psi*I (RB)	
	BAR	vector	BAR	vector		BAR	vector	BAR	vector
112	5	5	4	5	761	0	1	0	0
187	nd	nd	7	nd	774	7	8	5	9
262	4	4	3	3	931	1	1	1	1
279	0	1	0	0	964-1	3	4	4	4
283	1	nd	1	nd	964-2	2	3	2	3
298	4	5	4	5	975	4	4	nd	nd
348	1	2	1	1	1135	0	1	0	1
350	2	2	nd	1	1142	3	4	2	3
409	6	6	5	6	1146	2	2	2	4
551	0	0	0	0	1341	0	0	0	0
552	5	6	4	3	1361	nd	nd	5	6

3.6.2. T-DNA linkage analysis

Since in most of the mutants, several T-DNA copies and binary vector sequences were found, the linkage of these inserts to the hygromycin-resistant phenotype had to be evaluated. This was first done by T-DNA segregation analysis in F2 populations from backcrosses between M2 mutant plants and wtZH. Since the T-DNA constructs contain a marker gene that renders the plant resistant to BASTA, segregation behaviour of the T-DNA insertions could be detected by the ratio of BASTA resistance in F2 pools (Table R7). For example, in case of a single T-DNA insertion that is actively expressing the BASTA resistance gene, the expected resistance ratio would be 75%. The same is true for complex insertions containing several T-DNAs, if they do not segregate due to tight linkage. Two unlinked T-DNA insertions expressing the BASTA resistance gene would lead to 94% resistant F2 plants. While the mutant lines 552, 964-1 and 1361 had BASTA resistance of >82% suggesting multiple, segregating T-DNA insertions, no BASTA resistance was observed in 761 and 1135, concomitant with the previous Southern blot results where no T-DNA insert could be detected.

Cosegregation of T-DNA inserts and mutant phenotype, was analyzed by double selection with BASTA and hygromycin (Table R7). If the T-DNA that encodes the BASTA resistance is also the cause for the hygromycin resistance *in trans*, then the percentage of double-resistant plants should equal the percentage of resistant plants after hygromycin selection alone. If the double resistance ratio is significantly lower than the ratio of hygromycin resistance, then the T-DNA insertion is not linked to the hygromycin-resistant phenotype. BASTA-, hygromycin- and double-resistant plants were harvested and their genomic DNA compared by Southern blot for T-DNA and vector insertions, thereby determining which insertion, either T-DNA or binary vector, was linked to the hygromycin phenotype (Table R7). Unfortunately, with the exception of 1142, the T-DNA and binary vector insertions turned out to be unlinked to the mutant phenotype in all other lines.

Table R7: T-DNA segregation and linkage analysis

		Resistance in %			
Line	M2 x wtZH	hyg.	BASTA	double	T-DNA Southern blot[1]
112	F2 #1	0	0	0	unlinked single insertion site.
	F2 #3	61.2	44.4	34.8	
	F2 #4	52.9	69	37.5	
187	F2 #1	67.6	68.8	57.1	unlinked single insertion site
	F2 #2	45.5	0	0	
	F2 #6	0	0	0	
	F2 #7	72.1	62.5	39.3	
348	F2 #1	73.5	70	39.6	unlinked single insertion site
	F2 #4	73.8	77.3	50	
	F2 #5	0	0	0	
	C2 #6[2]	66.7	0	0	
409	F2 #2	70	76.5	31.8	unlinked single insertion site
	F2 #3	73.1	63.5	7.7	
552	F2 #1	0	89.5	0	unlinked, multiple insertions
761	F2 #1	61.8	0	0	no T-DNA
	F2 #2	62.5	0	0	
774	F2 #1a	71.4	74	34	unlinked single insertion site
	F2 #1b	74.5	64.3	52	
	F2 #1c	80	82	52	
964-1	F2 #2	4.1	82	0	unlinked, multiple insertions
	F2 #3	0	94	0	
964-2	F2 #1	0	0	0	unlinked single insertion site
	F2 #2	60.8	82	58	
	F2 #3	0	70	0	
1135	F2 #1	53.4	0	0	no T-DNA
	F2 #2	8	0	0	
	F2 #3	6	0	0	
1142	F2 #1	8.2	71.4	6.3	linked single insert
	F2 #2	2	6.73	0	
	F2 #3	10.2	66	10	
1361	F2 #1	0	83.3	0	unlinked, multiple insertions
	F2 #2	0	0	0	
	F2 #3	77.3	97.6	n.d.	

1: Southern analysis comparing hygromycin-, BASTA- and double-resistant pools
2: F2 population from cross between M2 and C2S1

3.6.3. Identification of T-DNA flanking regions

Since the sequence of the T-DNA used for the mutagenesis was known, this information could be used to specifically amplify and sequence the flanking regions of the insertion site. This was done with thermal asymmetric interlaced (TAIL-) PCR (Liu *et al.* 1995), where in three subsequent PCR reactions the T-DNA flanking region is amplified by a set of walking primers specific to the T-DNA borders and directed outwards, and an excess of degenerated primers that bind randomly in the *Arabidopsis* genome. The amplified PCR products were directly sequenced and the sequences were compared to the *Arabidopsis thaliana* whole genome sequence TAIR7 (http://www.arabidopsis.org) using BLAST algorithms (Table R8).

Table R8: TAIL mapped T-DNA flanking sequences

Line	T-DNA border	genomic location
112	LB	upstream of At2g43610 (putative chitinase)
187	RB	upstream of At1g62300 (*WRKY*6)
	LB	between At2g47580 (*U1A*) and At2g47590 (*PHR*2)
348	RB	5' UTR of At5g02470 (*DPA*)
409	RB	tandem T-DNA, vector sequence
	LB	upstream of At2g37150 (*PER*21), tandem T-DNA
774	RB	4th exon of At1g08140 (*ATCHX*6), vector sequence
	LB	4th exon of At1g08140 (*ATCHX*6)
964-1	LB	between At1g22300 (*GRF*10) and At1g22310 (*MBD*8), tandem T-DNA
964-2	LB	Mutator TN, first intron of At5g46795 (*MSP*2)
1142	LB	7th exon of At5g66750 (*DDM*1), tandem T-DNA

Although many of the T-DNA insertions mapped here were obviously not linked to the hygromycin resistance, a number of genes found by this method were still analyzed for their regulatory role in maintenance of HPT transgene silencing since the linkage analysis could have been obscured by the carryover effect mentioned earlier. The analysis was limited to the most obvious genes, though, such as the spliceosomal factor *U1A* and the photolyase *PHR*2 in 187, the *E2F* binding factor *DP*A in 348, the 14-3-3 gene *GRF*10 and the methyl binding domain gene *MBD*8 in 964-1, the pollen specific gene *MSP*2 in 964-2 and the known ATP-dependent chromatin remodeler *DDM*1 in 1142. This was done by phenocopy and linkage analysis. Additional mutant alleles of the genes identified, available from the NASC Arabidopsis stock centre (http://arabidopsis.info) or other labs, were introduced into the C2S1 line by crossing and F2 generations were analysed for hygromycin resistance in the case of *DPA* and, *DDM*1 (data not shown). Similarly, the original mapped alleles were crossed out from the activated HPT transgene and reintroduced to the parental C2S1 line still carrying a silent HPT transgene (done for *U1A*, *PHR*2, *DPA* and *DDM*1, data not shown). Except for *DDM*1 which was already known to be involved in regulation of HPT transcription (Mittelsten Scheid *et al.* 2003, Milos 2006), none of the analysed alleles could reactivate the naive

introduced C2S1 locus. The mutations in the *GRF*10, *MBD*8 and *MSP*2 genes found in 964-1 and -2 were already shown to be unlinked to the HPT reactivation in previous experiments and therefore not further analysed in this context.

3.6.4. Mutation tracking with additional silent marker gene

Two unexpected results, the segregation of the T-DNA insertions in the mutant lines from the hygromycin resistance phenotype and the autonomous expression of the HPT transgene after reactivation, made conventional mapping approaches futile. Therefore, an additional approach was chosen to confirm and visualize the presence of the mutations. Since the lines 187, 409, 552, 774, 1135 and 1142 were able to reactivate an additional reporter line carrying a silent repetitive GUS transgene (TS-GUS, (L5, Morel *et al,* 2000)), this line was used to track the mutation in single F2 plants from segregating populations. F2 seeds from crosses between hygromycin-resistant mutant plants and TS-GUS were germinated and grown on tissue culture dishes. After two weeks, cotyledons from each single plant were dissected and analysed for GUS reactivation (Figure R28A). Since the GUS reactivation confirmed a new transgene reactivation event and therefore the presence of the mutation in the analysed plants, these were transferred to soil and allowed to develop further. Because, as mentioned in 3.5.9, the GUS staining patterns of 552, 1135 and of some 187 and 774 plants appeared to be similar in the analysed stages and organs to 1142 with the already mapped new ddm1 allele, 187, 552, 774 and 1135 were analysed for mutations in the *DDM*1 gene. Genomic DNA extracted from the GUS-positive plants was digested with *Hind*III and analysed by Southern blotting with a *DDM*1 probe. While 1142 showed rearrangements caused by the mapped T-DNA insertion, no visible changes could be found in 187 (data not shown), 774 and 1135 (Figure R28B), while 552 showed a slight shift in the 3' fragment (Figure R28B, ∆). In order to confirm these results, the *DDM*1 gene in 187, 552, 774 and 1135 was sequenced and compared to wild type *DDM*1. No mutations could be found in the analysed 187 and 774 plants (data not shown), but 552 and 1135 had 38bp and 20bp deletions, respectively, in the *DDM*1 coding region, thereby representing two additional *ddm*1 alleles found in this screen.

To analyse possible linkage of T-DNA or binary vector insertions to the GUS reactivation phenotype, genomic DNA from each GUS-reactivating plant was analysed by Southern blot analysis after *Hind*III digestion. Beside some crosshybridization of the p1barb whole vector probe to the TS-GUS transgene (pNOS region), a clear linkage of a rather complex and multiple T-DNA insertion was found in all analysed 774 plants (Figure R28C, asterisk), contradictory to the results obtained from hygromycin and BASTA resistance in F2

populations from M2 crosses to wtZH plants that suggested non-linkage. This indicates that the HPT locus is indeed becoming autonomous after being reactivated by the mutation and is able to carryover the hygromycin resistance to later generations without the presence of the mutated gene. This could be explained at least in 774 by the complete loss of DNA methylation and H3K9 dimethylation described previously. Other analyzed lines had either no T-DNA insertions (187 and 1135) or had an already mapped insertion (1142).

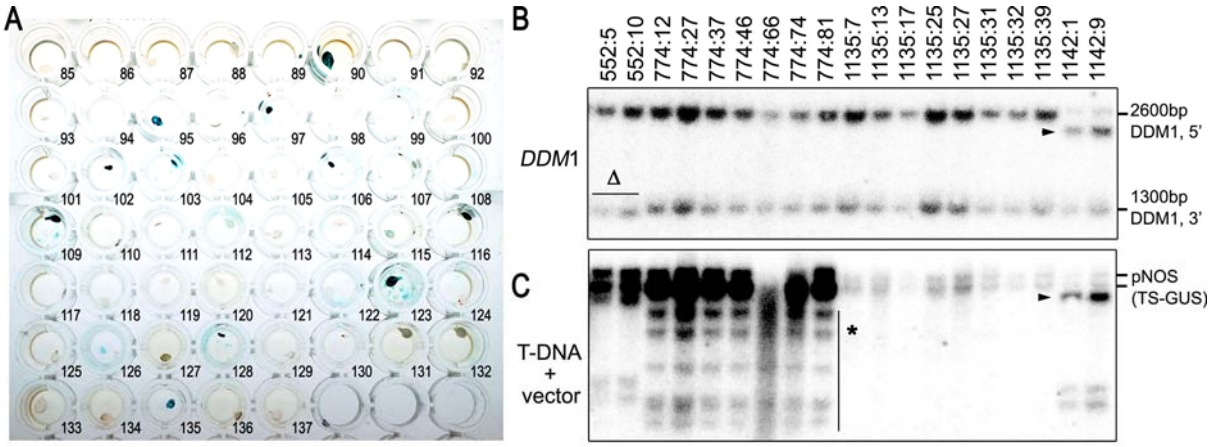

Figure R28: (A) Example for GUS assay on cotyledons from single 187 x TS-GUS F2 plants to monitor new reactivation events in the mutant background. **(B)** Southern blot analysis for *DDM*1 integrity in GUS-positive F2 plants. Δ: deletion in 552, arrowhead: T-DNA insertion in 1142. **(C)** Southern blot analysis for T-DNA and binary insertions in GUS-positive F2 plants. Asterisk: T-DNA insertions in 774, arrowhead: T-DNA insertion in 1142.

3.7. Transcriptome analysis

Expression profiling using microarray analysis is a widely used method to analyse global changes in gene expression. The *Arabidopsis thaliana* Affymetrix ATH1 array contains more than 22500 probesets representing 24000 genes and has been proven to be a useful tool for the *Arabidopsis* community. In this study, the ATH1 array was primarily used to compare gene expression changes between C2S1 and the mutant lines 348, 409, 774 and 1135 in order to gain more insight into global and endogenous changes introduced by the mutations. For this analysis, M4 seeds from mutant lines (known to be 100% resistant to hygromycin) and C2S1 seeds as reference were germinated and grown in triplicates for 17 days on GM media at long day conditions (18h light / 6h dark) and constant 21°C temperature. Total RNA samples extracted from seedling pools were sent to the Nottingham Arabidopsis Stock Centre (NASC) for completing the actual microarray experiments. Resulting raw data were normalized using Robust Multichip Average (RMA, (Irizarry *et al.* 2003)) and were further analysed for log2-fold changes (log2FC) in gene expression using linear modelling (Limma,

(Wettenhall and Smyth 2004)) in Bioconductor/R. Expression analysis of 774 and 1135 resulted in 259 upregulated probe-IDs in 774 and 243 upregulated probe-IDs in 1135, compared to C2S1 (log2FC>1, p-val.<0.05) (Table R9). Three hundred eighteen probe-IDs were found to be downregulated in 774, and 407 probe IDs in 1135 (log2FC<-1, p-val.<0.05) (Table R9). The 259 upregulated probe-IDs in 774 corresponded to 318 gene-IDs, while the gene-IDs contained 154 unique genes, 16 repeated genes and 148 individual transposons (46.5% of total gene-IDs). The 243 upregulated probe-IDs in 1135 corresponded to 349 gene-IDs with 85 unique genes, 16 repeated genes and 248 transposons (71.1% of total gene-IDs) (Table R9, Figure R29A). The increase in gene-IDs over probe-IDs was due to multiple binding of several probes to multiple targets, such as duplicated genes and transposons. The downregulated 318 and 407 probe-IDs in 774 and 1135 were specific to 335 and 420 gene-IDs respectively. Three hundred two unique and 33 repeated genes were found in the downregulated 774 gene set. Three hundred eighty unique and 50 repeated genes were found in the downregulated 1135 set. Both, downregulated 774 and 1135 genesets were free of transposons (Table R9). In total, 141 probe-IDs, 201 gene-IDs, 48 unique genes, 6 repeated genes and 145 transposons (72.1% of the overlapping gene-IDs) were found to be upregulated in both mutant lines. Two hundred forty-three probe-IDs, 258 gene-IDs, 225 unique genes, 33 repeated genes and no transposons were found to be downregulated in both mutant lines (Table R9 and Figure R29A and B).

All up- and downregulated genes in both mutants showed an unbiased and even distribution on the chromosome arms when visualized with the chromosome map tool available at TAIR (www.arabidopsis.org). The upregulated transposons were, as expected from previous publications (Jordan *et al.* 2007) mostly clustered at the pericentromeric regions of chromosome I to III or at the heterochromatic knob on chromosome IV (Figure R29C).

Table R9: Up- and downregulated genes in mutants 774 and 1135

	Upregulated			Downregulated		
	774	1135	Overlap	774	1135	Overlap
Probe-IDs	259	243	141	318	407	243
Gene-IDs	318	349	201	335	420	258
Unique genes	154	85	48	302	380	225
Multiple genes	16	16	6	33	50	33
Transposons	148	248	145	0	0	0

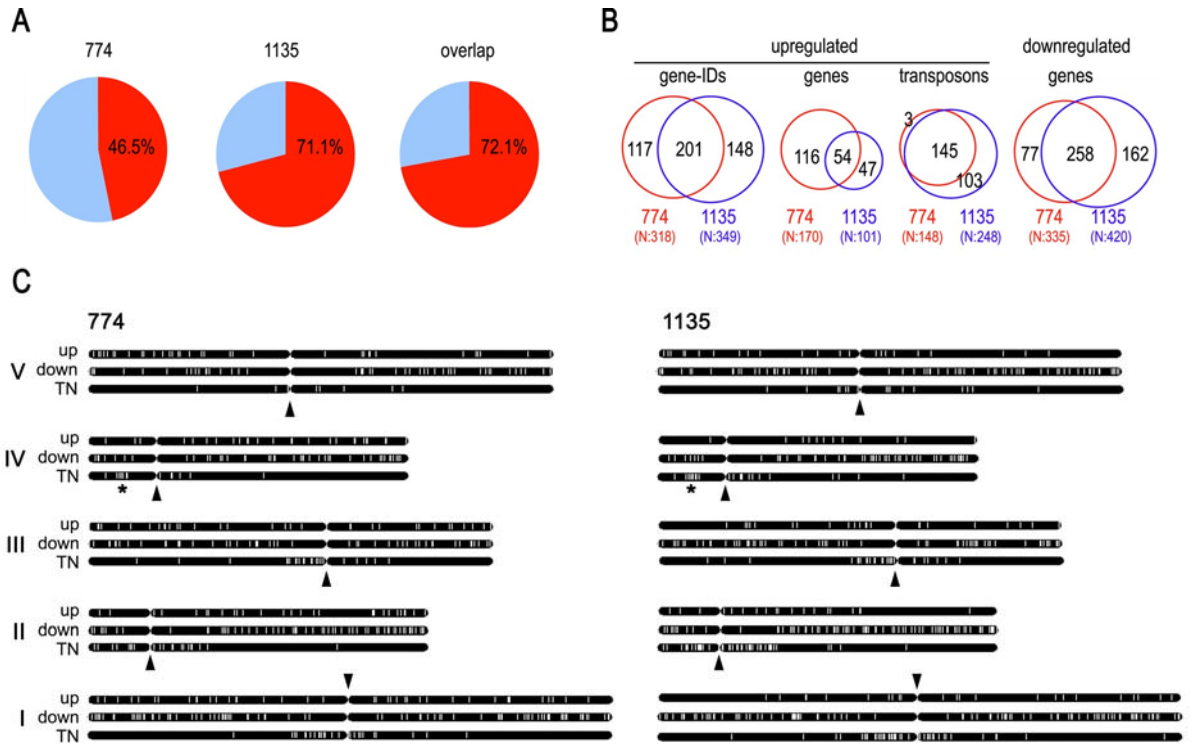

Figure R29: (A) Pie charts showing the percentage of transposons found in the upregulated gene-IDs. **(B)** Venn diagrams showing the overlapping gene-IDs, genes and transposons between the up- and downregulated 774 and 1135 sets. **(C)** Chromosomal positions of the 774, 1135 and overlapping up- and downregulated genes and transposons (arrowheads: centromeres, asterisk: heterochromatic knob on chr. IV).

All probes specific to transposons were removed from the up- and downregulated gene sets in order to have a more specific dataset. This was then used to analyse the gene ontologies (GOs) of the genes showing changes in expression. The datasets were bulk-analysed using the TAIR functional categorization tool GO-SLIM (www.arabidopsis.org) and compared to the whole genome categorization for Cellular Components (CC), Molecular Function (MF) and Biological Process (BP). A significant (chi^2: p-value < 0.05) enrichment of genes in the following GO terms was observed: *other cellular components (CC)* in the up- and downregulated genesets of both mutants including overlaps; *other binding (MF)* and *other enzyme activity (MF)* in the downregulated genesets of both mutants including overlaps; *developmental processes (BP)* and *response to stress (BP)* were enriched in the downregulated genesets of both mutants and overlap, and in the upregulated 1135 geneset (Table R10 and Figure R30). The following GOs were significantly underrepresented in the analysed genesets: *nucleic acid binding (MF)*, *other molecular functions (MF)* and *unknown biological processes (BP)* in the downregulated genesets from both mutants and overlap (Table R10 and Figure R30). Additionally, no genes belonging to the following GO terms were found in the analysed sets: *Golgi apparatus (CC)* in all sets, *extracellular (CC)* in all except 1135 down, *structural molecule activity (MF)* and *receptor binding activity (MF)* in all

upregulated genesets and *DNA or RNA metabolism (BP)* in the 774 and overlapping upregulated genesets (Table R10 and Figure R30).

Table R10: Percentage of GO terms enriched in the up/down regulated genesets

	Total[1]	774		1135		Overlap	
Cellular Component		up	down	up	down	up	down
Unknown cellular components	28,2	29,4	30,4	36,8	29,3	34,0	31,4
Other cellular components	13,8	27,6*	20,4*	21,8*	21,7*	30,0*	21,8*
Other intracellular components	11,2	12,3	11,3	10,3	11,1	10,0	11,8
Other membranes	10,1	4,9	9,4	8,0	9,1	8,0	10,0
Chloroplast	7,7	4,9	6,0	5,7	7,1	8,0	7,0
Other cytoplasmic components	7,6	4,9	5,3	4,6	5,1	4,0	4,8
Nucleus	7,4	4,3	5,0	4,6	4,8	2,0	3,9
Mitochondria	3,8	3,1	3,1	3,4	2,5	2,0	2,2
Plastid	2,8	2,5	2,2	1,1	2,3	2,0	1,7
Ribosome	1,7	1,8	1,9	1,1	2,0	*0,0*	1,3
Plasma membrane	1,3	1,8	1,6	1,1	1,8	*0,0*	1,3
Cytosol	1,2	1,2	1,6	1,1	1,5	*0,0*	1,3
ER	1,0	0,6	1,3	*0,0*	1,0	*0,0*	0,9
Cell wall	0,8	0,6	0,6	*0,0*	0,5	*0,0*	0,4
Extracellular	0,7	*0,0*	*0,0*	*0,0*	0,3	*0,0*	*0,0*
Golgi apparatus	0,7	*0,0*	*0,0*	*0,0*	*0,0*	*0,0*	*0,0*
Molecular Function							
Unknown molecular functions	25,9	26,7	24,2	25,6	22,1	25,0	23,1
Other binding	9,9	14,4	15,2*	13,2	13,9*	16,7	15,0*
Other enzyme activity	8,6	12,3	13,5*	11,6	13,4*	13,3	13,2*
Hydrolase activity	8,3	11,2	9,2	10,7	9,8	11,7	9,6
DNA or RNA binding	7,3	10,2	8,3	9,9	9,3	10,0	9,6
Transferase activity	7,0	9,6	7,8	6,6	8,6	8,3	6,9
Protein binding	6,7	3,7	6,2	5,8	7,0	6,7	6,6
Kinase activity	4,8	3,2	4,7	5,0	4,5	3,3	5,4
Transporter activity	4,8	2,1	3,3	4,1	3,9	3,3	3,6
Transcription factor activity	4,7	2,1	3,1	3,3	3,2	1,7	3,3
Nucleotide binding	3,7	1,6	2,8	1,7	2,3	*0,0*	1,8
Nucleic acid binding	3,6	1,6	0,7*	1,7	0,9*	*0,0*	0,9*
Other molecular functions	2,5	1,1	0,5*	0,8	0,5*	*0,0*	0,6*
Structural molecule activity	1,5	*0,0*	0,2	*0,0*	0,5	*0,0*	0,3
Receptor binding or activity	0,7	*0,0*	0,2	*0,0*	0,2	*0,0*	0,3
Biological Process							
Other metabolic processes	20,9	22,0	21,3	19,1	19,9	23,7	21,2
Other cellular processes	20,8	21,1	19,2	19,1	17,9	22,9	19,0
Unknown biological processes	18,9	15,5	14,2*	15,8	14,1*	13,6	14,4*
Protein metabolism	8,3	6,8	10,0	10,7	10,5*	7,6	10,0
Developmental processes	4,4	6,8	8,4*	8,8*	8,9*	6,8	8,5*
Response to stress	4,3	5,9	7,8*	7,9*	7,7*	5,9	7,5*
Transport	4,0	4,6	4,8	4,2	5,0	5,1	5,2
Response to abiotic or biotic stimulus	3,9	4,6	3,5	4,2	4,0	5,1	3,8
Cell organization and biogenesis	3,5	4,3	3,5	2,8	3,9	3,4	3,5
Transcription	3,4	3,4	2,7	1,9	2,4	2,5	2,9
Other biological processes	3,0	3,1	1,7	1,9	1,8	1,7	1,5
Signal transduction	2,1	1,2	1,4	1,4	1,7	0,8	1,0
Electron transport or energy pathways	1,6	0,6	1,0	1,4	1,6	0,8	1,0
DNA or RNA metabolism	0,7	*0,0*	0,4	0,9	0,3	*0,0*	0,6

1: whole genome classification (TAIR)
*: significantly changed compared to whole genome categorization (chi^2, p-value: <0.05)

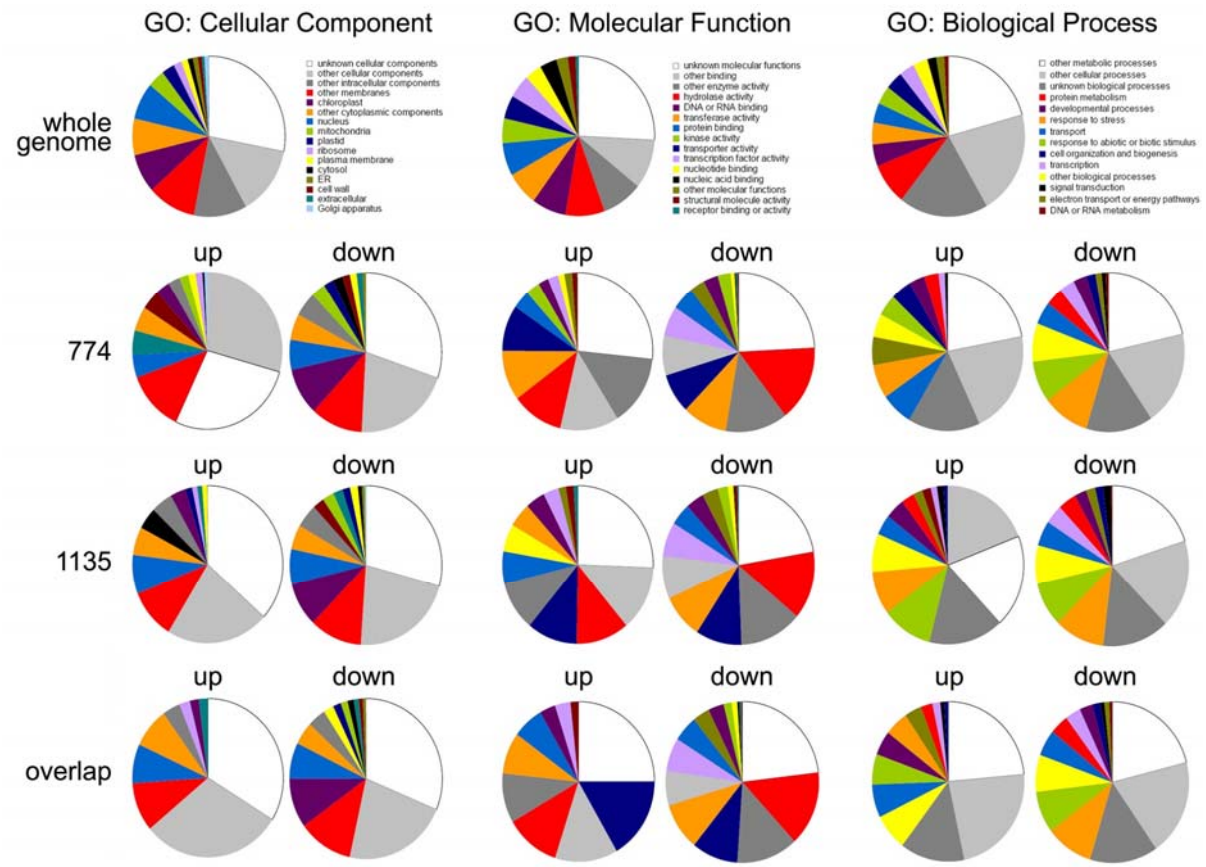

Figure R30: Pie charts showing the distribution of gene ontologies (GOs) for the whole *Arabidopsis* genome as a comparison to the genesets available from up- and downregulated 774, 1135 and overlapping targets.

The up- and downregulated genesets from 774 and 1135 were then further analysed for tissue-specific expression in order to see if the missregulated genes share the same expression patterns. This was achieved using the reference expression database and meta-analysis system Genevestigator (www.genevestigator.ethz.ch (Zimmermann *et al.* 2004)). The available genesets were clustered by Pearson correlation based on their expression profiles in different anatomic tissues (Figure R31). The expression pattern of the analyzed genesets was ranging from ubiquitous to tissuespecific expression. In the upregulated 774 and 1135 sets, many genes were found not to be expressed in anatomic plant tissues, but highly expressed in *in vitro* callus cultures, while no callus-specific genes could be detected in the downregulated sets (Figure R31, a). In addition many genes with tissue specific expression patterns such as seed (b), floral organs (c), adult leafs (d) and roots (e) were found to be upregulated by the 774 mutation (Figure R31). Interestingly, the upregulated 774 and 1135 genesets were not similar in the expression pattern. While the 774 set was enriched with tissuespecific genes and some ubiquitously expressed genes (f), the 1135 set included, beside the mentioned callus-specific and some rootspecific genes a high proportion

of ubiquitously expressed genes (Figure 31). Both downregulated genesets were similar in the observed expression patterns in agreement with the high proportion of overlapping genes. Genes with strong expression in cell suspension (g), senescent leaves (h), roots and pollen (i) were enriched (Figure R31).

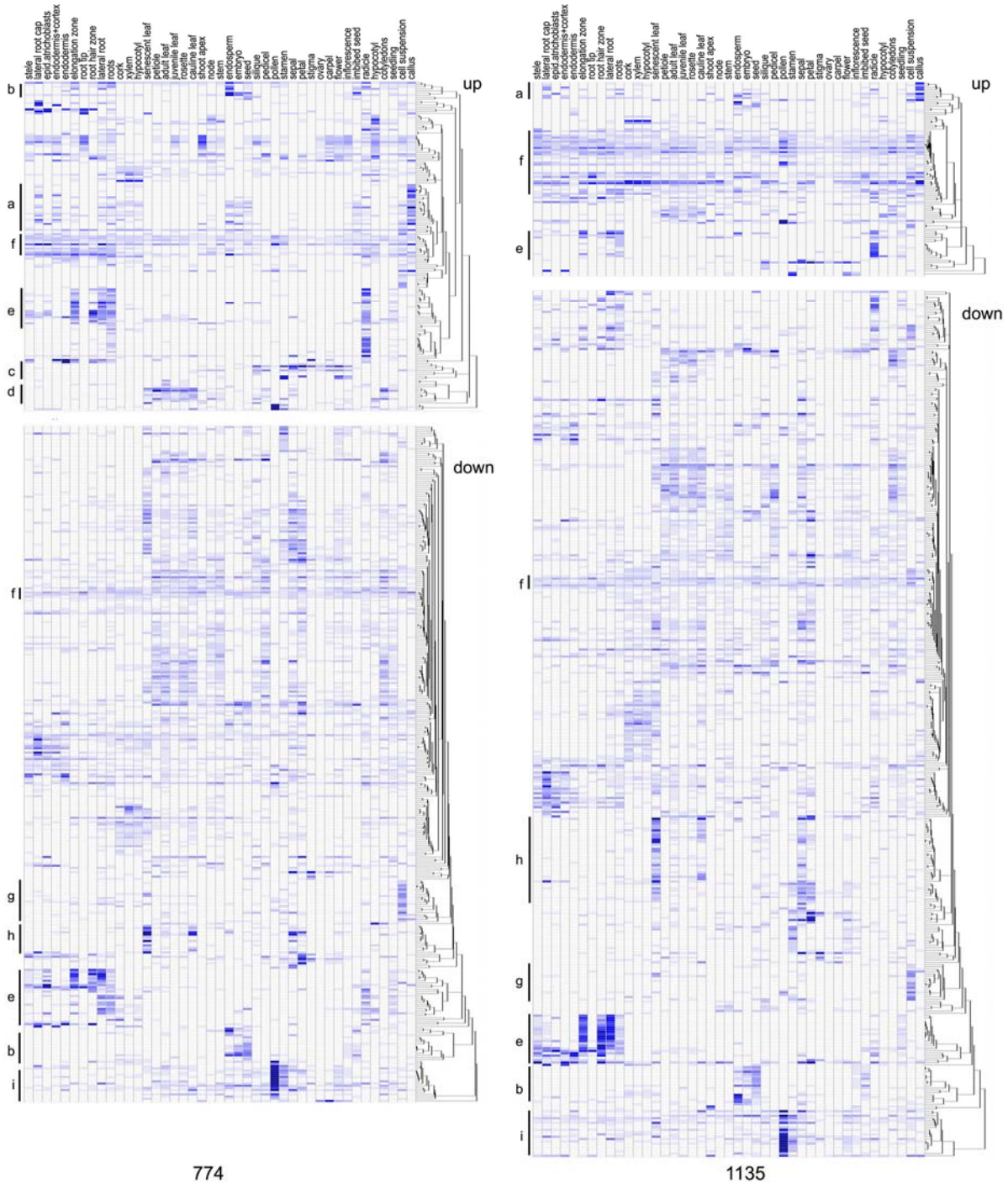

Figure R31: Hierarchical clustering of up- and downregulated 774 and 1135 genes according to tissuespecific gene expression using Pearson coefficient gene correlation in Genevestigator (www.genevestigator.ethz.ch). Top: tissues, right: gene tree. Dark blue indicates high expression; white, no expression. Groups of genes with similar expression patterns - **a:** callus; **b:** seed, **c:** floral organs, **d:** adult leaves, **e:** root tissue, **f:** ubiquitous expression, **g:** cell suspension, **h:** senescent leaf and **i:** pollen.

While analysing the up- and downregulated genesets in 774 and 1135, many genes known to be involved or regulated by epigenetic processes were found to be missexpressed. For example, the DNA N-glycosylase *REPRESSOR OF SILENCING* 1 (*ROS*1, At2g36490), the *FLOWERING LOCUS C* (*FLC, At5g10140)* and the *LINKER HISTONE* 1 variant 3 (*HIS*1-3, At2g18050) were found to be significantly downregulated, while an ATH1-probe specific for the *HISTONE DEACETYLASES* 10 and 17 (*HDA*10, *HDA*17) was significantly upregulated in both mutants. The mutated gene *DECREASE IN DNA METHYLATION* 1 (*DDM*1) was found to be significantly downregulated in 1135 as expected. Mutant 774 showed no change in *DDM*1 expression, although RT-PCR experiments that were done to confirm the ATH1 results even suggested an increase of *DDM*1 transcripts in the 774 background (Figure R32). Two more genes were found to be significantly downregulated in the 774 line, the Jumonji-C domain containing histone demethylase (H3K9) *INCREASE IN BONSAI METHYLATION* 1 (*IBM*1, At3g07610) and the S-homocystein-L-adenosylhomocysteine hydrolase *HOMOLOGY-DEPENDENT GENE SILENCING* 1 (*HOG*1, At4g13940). Both genes have already been shown to be involved in regulation of epigenetic gene silencing (Furner *et al.* 1998, Rocha *et al.* 2005, Saze *et al.* 2008). In order to confirm this result, *IBM*1 and *HOG*1 expression was analysed by RT-PCR and proven to be downregulated in 774, but not in C2S1 or 1135 (Figure R32A). Since both were found to be downregulated in 774, which could be either due to secondary effects caused by an unknown mutation or by a direct mutation and resulting nonsense-mediated mRNA decay, the integrity of both genes was analysed by Southern blot analysis. This was done using genomic DNA obtained from TS-GUS-monitored F2 plants described earlier. While the *IBM*1 locus was unchanged (data not shown), all analysed GUS-positive F2 plants from crosses to 774 were positive for a large insertion at the *HOG*1 gene, therefore suggesting a direct mutation of that gene in the 774 mutant line (Figure R32B).

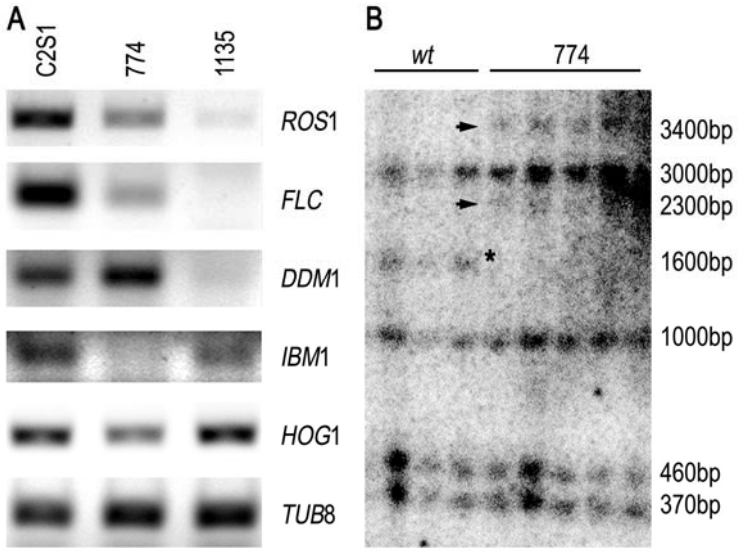

Figure R32: (A) Reverse Transcription PCR validation of downregulated genes identified by ATH1 array expression.analysis **(B)** Southern blot analysis using *Hind*III restriction enzymes showing a change in the *HOG*1 locus in 774 GUS-positive F2 plants. The wild type band at 1600bp (asterisk) is up-shifted to 2300 and 3400bp (arrowheads) in 774.

ATH1 analysis of gene expression in the lines 348 and 409 had a very surprising outcome. In both cases, only one to two genes were significantly (log2FC>1, p-val. <0.05) upregulated and almost no genes were downregulated at all. The single upregulated gene was the gene *BME*3 (At3g54810) neighbouring the HPT insertion. The view that the *BME*3 gene is obviously regulated in an HPT transgene-state dependent manner is further supported by the fact that *BME*3 was also found to be upregulated in 774 (*hog1*-7) and 1135 (*ddm1*-12), together with the reactivated HPT transgene. Interestingly, a second linked gene found to be upregulated in 348 was At3g54730, a gene of unknown function with low transcription. This gene is located ~30kb upstream of the HPT transgene insertion site and tightly regulated by siRNAs and DNA-methylation, since both are highly enriched according to the Anno-J database (Lister *et al*. 2008). This gene has two duplications, one on chromosome II: At2g19850 with 98% sequence similarity to the whole annotated gene, and one on chromosome IV at the annotated 5' part of *OVATE FAMILY PROTEIN* 9 (*OFP*9, At4g04030)**,** with 99% sequence identity (Figure R33). It was not clear whether the observed up-regulation comes from the linked gene At3g54730 or from any of the two unlinked genes since the ATH1-probe was located within the repeated region. Although this probe was not upregulated in 409, there was a clear up-regulation with 5.8 Log2FC (p-val. < 0.05) in 774 and 5.2 Log2FC (p-val. < 0.05) in 1135, suggesting a potential co-regulation with the HPT transgene.

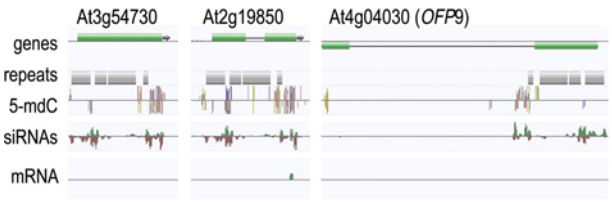

Figure R33. Anno-J output for At3g5730 and its repeated loci At2g19850 and At4g04030. 5-mdC track shows DNA methylation (^{m}CG green; ^{m}CHG blue; ^{m}CHH magenta), siRNA track shows mapped short sequence runs in sense and antisense direction. mRNA track shows expression. (adapted from Anno-J, Lister *et al.* 2008)

To confirm the microarray results of *BME*3 and the At3g54730 repeat, real-time PCR was conducted to quantify the up-regulation of these two targets in the analysed lines including additional mutant and control lines, compared to expression levels in C2S1. *EIF*4A1 was used for normalisation (Figure R34). *BME*3 expression was upregulated in all analysed mutant lines, including the active C2R line. This confirmed the expression-array results for *BME*3 and further showed that this gene is dependent on the HPT transgene activity since no change in expression could be detected in wtZH compared to C2S1 (Figure R34B). *BME*3 was particularly over-expressed in the new *ddm*1 and *hog*1 alleles with over 20-fold expression. Expression of the triple-repeated gene At3g54730 did not show significant changes in expression in the wtZH and C2R control lines and in the 409, 761 and 964-2 mutant lines. Expression was similar to C2S1 levels. However, 5.5-fold up-regulation was

detected in the 348 line and 9- to 16-fold the *ddm*1 lines. Surprisingly, the up-regulation in the *hog*1 line 774 was nearly 60-fold (Figure R34C).

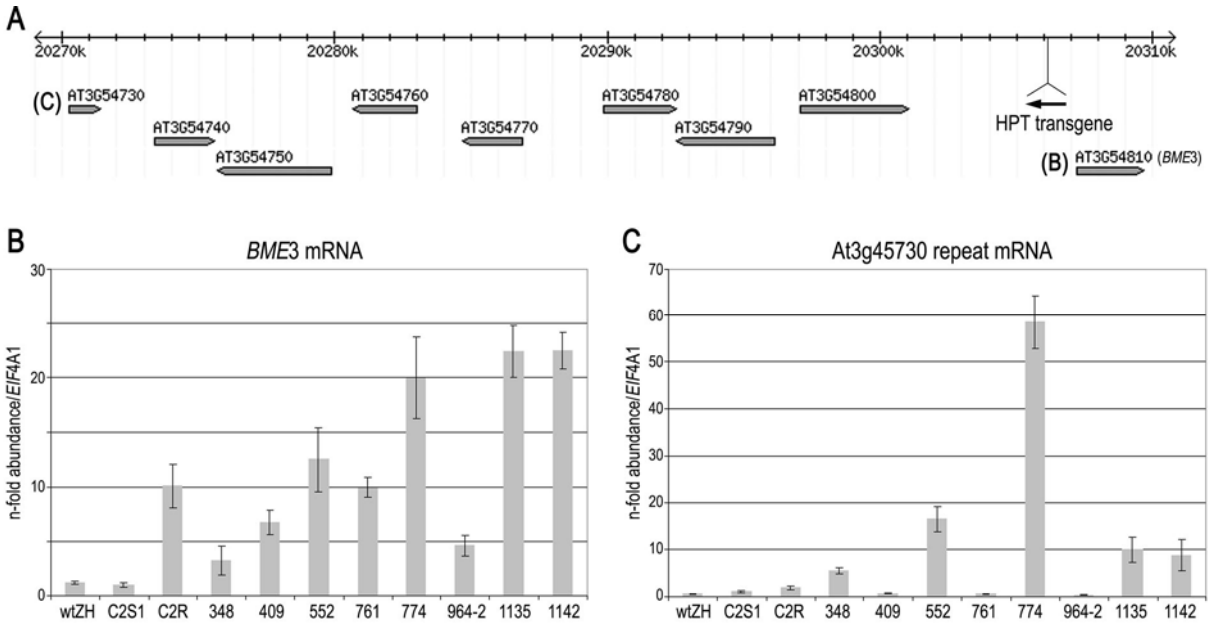

Figure R34: Map illustrating the location of the HPT transgene and the analysed regions. **(A)** Quantitative real-time PCR to measure the fold- expression of **(B)** *BME*3 and **(C)** At3g54730-repeat in forward mutants and control lines compared to C2S1. *EIF*4A1 was used for normalisation. Error bars represent standard deviations from triplicate analysis.

3.8. Identified mutations:

3.8.1 *HOMOLOGY-DEPENDENT GENE SILENCING* 1, *hog*1-7

Since the Southern blot results from mutant 774 obtained from GUS-reactivating F2 plants (Figure R32B) suggested an insertion in the *HOG*1 gene, a PCR approach using combinations of T-DNA- and *HOG*1-specific primers and subsequent sequencing of the obtained products was chosen and indeed revealed a T-DNA integration in the 3' region of the gene. In spite of several efforts to map both insertion borders, only the sequence flanking the T-DNA left border could be mapped so far (Figure R35). The reason is probably a complex insertion, including several T-DNAs and vector sequences, as apparent from the T-DNA Southern in Figure R32B. The sequence of the left border indicated that the T-DNA inserted outside the coding region, in the 3' UTR, therefore very likely not causing a complete loss of function. This is in agreement with the observed zygotic lethality in full loss-of-function mutants (Rocha *et al.* 2005).

Figure R35: Mutant line 774 T-DNA integration site in the *SAHH*1/*HOG*1 gene. Four predicted splicing variants are displayed. Red bar represents the sequenced left border (LB) flanking region. Dashed bar represents the flanking right border region (RB) that failed to be sequenced.

3.8.2 *DECREASE IN DNA METHYLATION* 1, *ddm*1-11 to *ddm*1-13

The three new *ddm*1 alleles identified in this screen were named *ddm*1-11 (552), *ddm*1-12 (1135) and *ddm*1-13 (1142) in continuation of previously identified mutations (Jeddeloh *et al.* 1999, Jordan *et al.* 2007). *ddm*1-11 has a 38 bp deletion in exon V, *ddm*1-12 has a 30 bp deletion in exon XIV and *ddm*1-13 has a T-DNA insertion in exon VII. In contrast to the widely used alleles *ddm*1-2 (point mutation generating a G→A transition in the splice donor site of intron XI, Jeddeloh *et al.* 1999) and *ddm*1-5 (82 bp insert in exon II, Mittelsten Scheid *et al.* 1998), the new mutations are all localized in conserved signature motifs that are characteristic for SWI2/SNF2 family proteins (Bork and Koonin 1993) and affect the domains that are important for ATP-dependent chromatin remodelling, namely SNF2_N and DEAD/DEAH (Figure R36A and C). This may explain why the new alleles caused a stronger reactivation of the HPT than the *ddm*1-5 allele used in the reverse screen and led to hygromycin resistance already in the M2 generations where mutant selection was applied. The *ddm*1-5 allele showed partial reactivation only in F4. The striking difference between the effects of the alleles could be seen from a direct comparison of the *ddm*1-5 F4 seedlings to M4 seedlings obtained from the novel alleles (Figure R36B). In order to confirm the strong reactivating effect of the new *ddm*1 alleles on the silent HPT, the mutations were separated from the original HPT copy upon outcrosses with TS-GUS, identified by GUS-positive F2 plants and genotyped for the loss of the HPT transgene during segregation (expected: 25%). The resulting HPT-negative plants were crossed to C2S1 plants containing the original silent HPT transgene. F2 populations from these backcrosses were screened on hygromycin, and hygromycin resistance was indeed found in F2 generations from in C2S1 crossed to 552, 1135 and 1142, ranging from 4.2% to 18.2% resistant plants. Although resistance was not as strong as in C2R, the plants could develop normally after removal of hygromycin (data not shown). To predict the effects of the mutations on the *DDM*1 protein (comprising 764 amino acids (AA) in the wild type), *in silico* reading frame analysis was conducted. While the *ddm*1-2 mutation generates a splicing variant of 614 AA length and the *ddm*1-5 mutation leads to an early translation stop after 58 AA, the novel alleles introduce a premature stop codon after 653 AA in *ddm*1-11, 260 AA in *ddm*1-12 and 347 AA in *ddm*1-13 (Figure R36C). Interestingly, analysis of the translation initiation sites available at the *DDM*1 gene showed

that there are six putative initiation sites upstream of the sequence encoding the SNF2 domain (Figure R36C, white triangles). Comparison to the consensus translation initiation sequence in *Arabidopsis* (aa(A/G)(A/C)a**AUG**Gcg, (Rangan *et al.* 2008)) suggested that the annotated initiation site for DDM1 has a much lower sequence similarity to the consensus than the downstream initiation sites 3 and 4, suggesting that those could be used as alternative initiation sites to produce a functional protein in *ddm*1-5. Therefore, *ddm*1-5 might still have the potential to produce low amounts of functional DDM1 by using one of the downstream initiation sites, while none of them would circumvent the mutational effects of the new alleles *ddm*1-11 to 13.

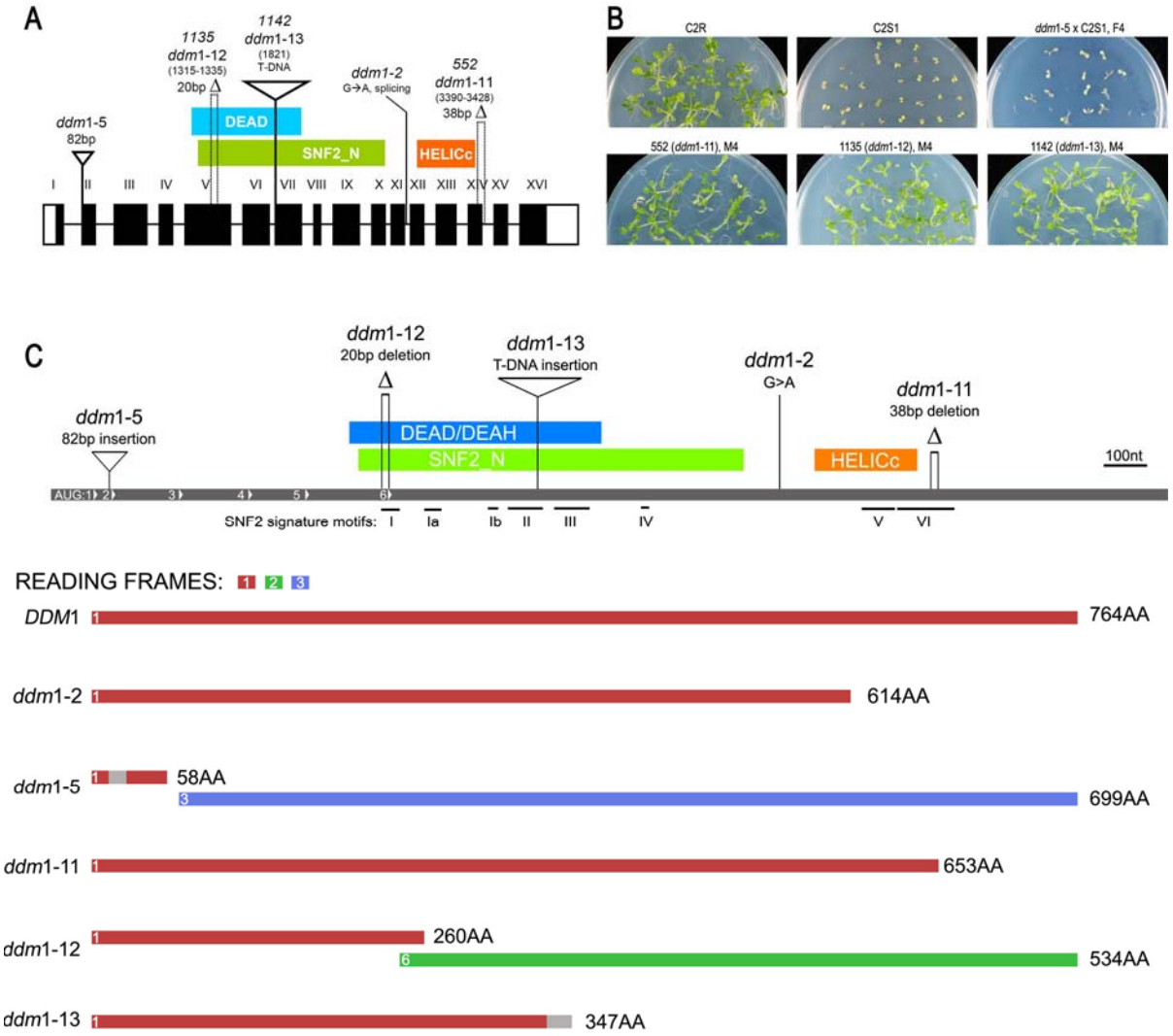

Figure R36: Novel *ddm*1 alleles identified in the screen. (A) *DDM*1 gene region with indicated UTRs (white boxes), exons (filled boxes) and introns (lines). Functional domains are indicated by coloured boxes, while mutations are indicated by insertions or deletions. (B) Allele comparison by hygromycin selection in analogous generations: F4 from crosses between C2S1 x *ddm*1-5 and M4 in the novel alleles. (C) Reading frame analysis in the *ddm*1 alleles. Coding sequence is indicated by the grey bar, mutations and domains are depicted above, conserved SNF2 signatures are shown below. White glyphs indicate potential translation initiation sites in the 5' region. Coding reading frames (in different colours) and encoded protein size are predicted in wt and mutant alleles. Light grey bars indicate non-plant DNA insertions.

3.9. Mutant overview table

The following table summarizes the information available about the different mutant lines that were obtained in the screen. It allows a synoptic view and will serve as a basis for the Discussion.

	Mutants with global effects				
Pool number:	187	552	774	1135	1142
HPT transgene reactivation:					
Hygromycin resistance*	+	+	+	+	+
HPT transcript**	+	+	+	+	+
HPT -fold expression level**		1	1	0.96	1.5
HPT transgene integrity*	+	+	+	+	+
M2 siblings total	1	2	5	12	2
Ploidy	2n	4n	2n	2n	2n
Morphological phenotypic variation:	+	+	+	+	+
Gene affected:		*ddm1*	*hog1*	*ddm1*	*ddm1*
Hygromycin resistance in:					
M3, selfed progeny	4%	73%	96%	78%	71%
Plants tested	25	48	23	46	38
F1, *WtZH* outcross	R	100%	100%	100%	
Plants tested	n.d.	33	31	13	
Transposon reactivation*:					
Transcriptional Silent Information	++	++	++	++	++
LINE1-4 (non LTR retro-TN)	n.d.	+	+	+	+
MULE (TN)	n.d.	++	+	++	++
CACTA like (TN)	n.d.	++	++	++	++
Ta3 (LTR retro-TN)	n.d.	+	+	+	+
SN1 (retro-TN)	n.d.	-	-	-	-
Solo LTR	n.d.	-	-	-	-
Transgene reactivation:					
TS GUS reactivation	+	+	+	+	+
Staining pattern	mixed	total	leaf sector	total	total
F2, TS GUS segregation	42%	25.4%	4.8%	22.6%	22.4%
Plants tested	78	67	164	234	67
DNA methylation:					
% 5-mdC/dC	n.d.	0.92%	2.15%	1.74%	1.82%
HPT transgene hypomethylation*	+++	++	++	++	++
180bp centromeric repeats*	+	++	+	++	++
Changes in nuclear organization*:					
5mdC localization	n.d.	dispersed	wt	dispersed	dispersed
H3K9me2 localization	n.d.	dispersed			
H3K27me3 localization	n.d.	additional foci detected			
180bp centromeric repeats	n.d.	dispersed			
45S rDNA	n.d.	wt	wt	wt	wt
Changes in chromatin structure*:					
H3K9me2	n.d.	lost	lost	lost	
H3K4me3	n.d.	gained	gained	gained	
H3K27me2	n.d.	lost	lost	lost	
H3K27me3	n.d.				
Drug treatments*:					
40 µM Zebularine	++	++	++	+	+
1.6 µM Trichostatin A	n.d.	+	++	-	-
Zebularine and TSA combined	n.d.	-	+	-	-

n.d.: not determined (either not possible or not needed);
*: compared to C2S1; **: compared to C2R

	Mutants without global effects					
Pool number:	348	409	761	964-1	964-2	1361
HPT transgene reactivation:						
Hygromycin resistance*	+	+	+	+	+	+
HPT transcript**	+	+	+	+	+	+
HPT -fold expression level**	0.1	0.4	0.7		0.2	
HPT transgene integrity*	+	+	+	+	+	+
M2 siblings total	1	1	1	5	5	2
Ploidy	2n	2n	2n			
Morphological phenotypic variation:	-	-	-	+	-	-
Gene affected:						
Hygromycin resistance in:						
M3, selfed progeny	53%	67%	53.8	35%	37%	78%
Plants tested	40	36	91	49	49	147
F1, *WtZH* outcross	n.d.	89%	n.d.	48%	47%	100%
Plants tested	n.d.	9	n.d.	25	17	13
Transposon reactivation*:						
Transcriptional Silent Information	-	(+)	+	+	+	
LINE1-4 (non LTR retro-TN)	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
MULE (TN)	-	-	-	-	-	n.d.
CACTA like (TN)	-	+	-	-	-	n.d.
Ta3 (LTR retro-TN)	-	-	-	-	-	n.d.
SN1 (retro-TN)	-	-	-	-	-	n.d.
Solo LTR	-	-	-	-	-	n.d.
Transgene reactivation:						
TS GUS reactivation	-	(+)				
Staining pattern	n.d.	cotyledon spots	n.d.	n.d.	n.d.	n.d.
F2, TS GUS segregation	n.d.	67%	n.d.	n.d.	n.d.	n.d.
Plants tested	50	15				
DNA methylation:						
% 5-mdC/dC	5.61%	5.67%	5.34%	5.89%	5.67%	
HPT transgene hypomethylation*	+	+	+	+	+	+
180bp centromeric repeats*	-	-	-	-	-	
Changes in nuclear organization*:						
5mdC localization	wt	wt	n.d.	n.d.	n.d.	n.d.
H3K9m2 localization	wt	wt	n.d.	n.d.	n.d.	n.d.
H3K27me3 localization	wt	wt	n.d.	n.d.	n.d.	n.d.
180bp centromeric repeats	wt	wt	n.d.	n.d.	n.d.	n.d.
45S rDNA	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Changes in chromatin structure*:						
H3K9me2		partial loss				
H3K4me3		gain				
H3K27me2		partial loss				
H3K27me3						
Drug treatments*:						
40 µM Zebularine	++	++	+	+	+	
1.6 µM Trichostatin A	++	-	-	-	-	
Zebularine and TSA combined	+	+	+	+	-	

n.d.: not determined (either not possible or not needed);
*: compared to C2S1; **: compared to C2R

4. DISCUSSION

4.1. Polyploidy-associated transcriptional gene silencing produces stable epialleles

Polyploidy is a rather common phenomenon in plants and is known to have significantly contributed to plant evolution and speciation (Stebbins 1966). The associated molecular consequences that occur during polyploidisation are still widely unexplored. One possible type of response is based on modified epigenetic regulation (Osborn *et al.* 2003). This includes alteration in chromatin structure and DNA methylation, thereby influencing gene expression from the affected DNA sequence.

In this thesis, such an epigenetic regulation phenomenon was analysed, where a hygromycin phospho-transferase (HPT) transgene was transcriptionally silenced in the course of generating tetraploid *Arabidopsis* plants. Silencing was faithfully maintained throughout subsequent tetraploid generations and even after return into a diploid state (line C2S1) (Mittelsten Scheid *et al.* 2003). Transcriptional inactivation was accompanied by extensive DNA hypermethylation throughout the whole transgene, accumulation of repressive histone marks (H3K9me2, H3K27me2, H3K27me3) and loss of activating histone marks (H3K4me3 and H4acetylation) (Mittelsten Scheid *et al.* 2003, A. Förster and O. Mittelsten Scheid unpublished, this thesis). The transgene consists of a functional single copy insertion on chromosome III, but has been re-arranged during its genomic insertion as described in 3.1.2. In the active C2R line, this rearrangement led to the accumulation of an extended HPT mRNA and an additional, non-coding transcript originating from the duplicated promoter (Figure R2). Both transcripts are not detected in the silent C2S1 line, arguing that both promoters were affected by gene silencing.

Since several examples of transcriptional regulation determined or influenced by neighbouring repetitive and transposable elements have been described in *Arabidopsis* (Lippman *et al.* 2004, Chan *et al.* 2006), extensive analysis of the genomic region surrounding the transgene insertion was performed but showed no evidence for transposable or other repeated elements (Figure R1). This excluded any potential influence from the surrounding region on the epigenetic state of the HPT transgene. In contrast, the insertion of the HPT transgene itself influences the surrounding region. For example, the neighbouring gene *BME*3, 700bp upstream of the transgene insertion site, was found to be influenced by the transcriptional activity of the HPT transgene, probably due to enhancing effects from the CaMV35S promoter (Figure R35B). This effect was also evident from changes in histone marks at the *BME*3 gene (Figure R15). H3K9me2 was detected in the silent line at low

levels, likely indicating spreading of repressive chromatin from the silenced HPT transgene into the coding region of *BME*3 and thereby affecting its expression.

The aim of this thesis was to identify potential regulators that are required to maintain the polyploidy-associated transcriptional gene silencing (PA-TGS) at the HPT transgene. Maintenance of transcriptional gene silencing is known to be suppressed by interfering with DNA methylation, chromatin modifications and siRNA pathways. This is achieved either by loss of function mutations from factors involved in these mechanisms (Jeddeloh *et al.* 1999, Lindroth *et al.* 2001, Cao *et al.* 2003, Kankel *et al.* 2003, Kanno *et al.* 2004) or by pharmacological interference (Chang and Pikaard 2005, Mathieu *et al.* 2003, Baubec *et al.* 2008). A reverse genetic screen to identify the involvement of known epigenetic regulators was done by introducing the silent HPT transgene into various epigenetic mutant lines and screening for hygromycin reactivation (Milos 2006). No hygromycin resistance was observed in all tested mutations except for *ddm*1-5, a chromatin remodelling factor involved in transcriptional gene silencing at repetitive elements (Mittelsten Scheid *et al.* 1998, Jeddeloh *et al.* 1999). However, the observed resistance was only apparent after prolonged inbreeding in the mutant background and was still only partial in the F4 generation (Mittelsten Scheid *et al.* 2003, Milos 2006). Taken together, the results from the reverse genetic screen argue that maintenance of PA-TGS at the HPT transgene is hard to overcome, at least not with the available mutants, since even *ddm*1-5 that has been proven to be a strong silencing suppressor in other experimental systems (Mittelsten Scheid *et al.* 1998, Morel *et al.* 2000, Elmayan *et al.* 2005, Gendrel *et al.* 2002) could not entirely release silencing from the transgene.

To further investigate the stability of transcriptional gene silencing at the HPT locus, zebularine, a DNA demethylating agent (Zhou *et al.* 2002) was applied alone or in combination with the histone deacetylase inhibitor trichostatin-A (Chang and Pikaard 2005). Although it has been reported that zebularine interferes with transcriptional gene silencing in plants (Baubec *et al.* 2008), silencing at the HPT transgene was not affected since the treated plants did not survive on hygromycin (Figure R3). Molecular analysis showed only a minor increase in HPT transcription that was obviously not enough to induce hygromycin resistance. In contrast, the duplicated promoter was strongly reactivated, producing large amounts of the secondary, non-coding transcript (Figure R4B and C). A similar increase of transcriptional activity from the duplicated promoter was observed in numerous mutants analysed in the reverse screen (*ddm*1, *cmt*3, *met*1, *mom*1 and *ago*4, M. Milos and T. Baubec, data not shown). Methylation analysis of the two promoters after zebularine treatment showed that the reduction in DNA methylation was comparable in both promoters (Figure R4E). This might indicate that maintenance at the duplicated CaMV35S promoters is completely different or depends on additional factors. While the duplicated promoter is

obviously regulated by DNA methylation alone and strongly reactivated after zebularine treatment, the promoter driving HPT, just 2kb upstream, maintains its inactive state in spite of being de-methylated to a similar degree. Thus, there could be either DNA methylation-independent or additional redundant regulatory mechanisms, which are hard to overcome and may represent a novel silencing maintenance mechanism set in motion during polyploidisation (Figure D1).

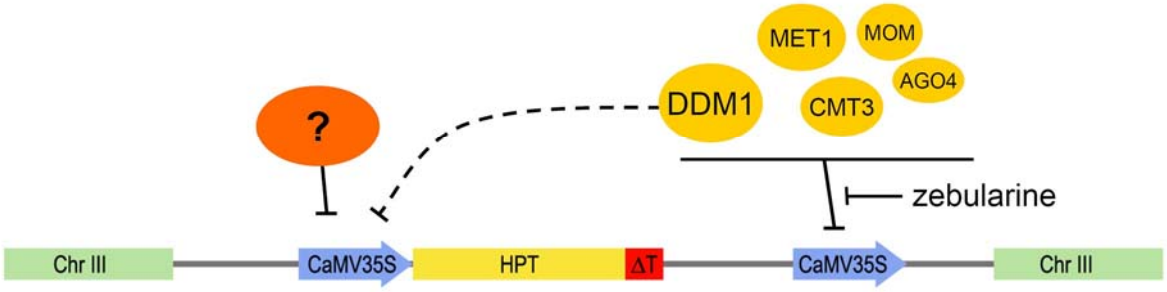

Figure D1: Differential regulation of the duplicated CaMV35S promoters. While release of transcriptional gene silencing from the duplicated promoter is induced in numerous mutant backgrounds and after zebularine treatment, the main promoter is more tightly silenced (reactivation only observed in *ddm*1-5 background after inbreeding), suggesting either unknown or additional redundant layers of regulation.

4.2. A forward genetic screen identified 21 individual hygromycin resistant mutant lines

A forward genetic screen was carried out with the intention to find the regulators involved in maintaining this surprisingly stable polyploidy-associated transcriptional gene silencing (PA-TGS) at the HPT transgene. Genetic analysis in tetraploids is rather complicated, and since the chromatin features at the transgene were faithfully maintained after reduction of ploidy from 4n to 2n, the diploid derivate line C2S1 was mutagenized and 20 000 mutant lines were screened for hygromycin resistance in M2. Several hygromycin-resistant plants were found during the screen. However, not all were able to maintain the resistance since thirteen lines were either fully sensitive or had low resistance ratios (< 5%) in selfed M3 generations obtained from initially hygromycin-resistant M2 plants. This could be explained either by rapid re-silencing of the reactivated transgene or by false-positive M2 plants. The latter case was rather unlikely since all M2 hygromycin-selected plants had strong resistance phenotypes. How and why the transgene was reactivated and then re-silenced in the next generation is still unknown. Interestingly, three lines with hygromycin resistance ratios below 5% in the selfed M3 generations exhibited increased resistance in F2 generations from backcrosses between M2 and wtZH (> 45%, 112, 187 and 551 in Table R1). This could not be explained by loss of plants homozygous for recessive mutations, since M3 plants grown on non-

selective media were developing normal (data not shown). It can be speculated that the out-crossing may have stabilized the transient reactivation of HPT, either by locking in the active state or through depletion of factors involved in triggering re-silencing.

The remaining mutants were able to maintain stable hygromycin resistance in M3 and further selfed generations. However, the resistance ratios in M3 generations did not fit the expected Mendelian segregation ratio calculated for recessive mutants (Figure D2). Assuming that hygromycin resistance in M2 is caused by homozygous recessive mutations, then resistance in selfed M3 seedlings should be 100%. With some exceptions, none of the analysed mutants showed ratios above 90%, suggesting that the recessive mutations were either not fully penetrant and/or would need further inbreeding to display a full effect (Table R1). Another explanation for the non-Mendelian segregation observed in M3 could be due to segregating dominant mutations that would result in ~ 75% resistance if the HPT transgene were homozygous. In addition, hygromycin selection of F2 generations obtained from crosses between M2 mutant lines and wtZH further confirmed the non-Mendelian behaviour of the introduced mutations. The observed ratios were either around 0% or remarkably higher than 18.75%, expected for recessive mutations in a segregating HPT transgene background. This could again indicate dominant behaviour of the introduced mutations (expected resistance: 56%) or due to autonomous behaviour of the HPT transgene itself (Table R1). The latter could have been caused by locus rearrangements releasing silencing or by the failure to regain epigenetic silencing marks after segregating from the responsible mutation. In both cases, the expected resistance ratio would have been 75%. In addition, some F2 populations that were fully sensitive on hygromycin indicated that either the dominant mutation or the autonomous HPT transgene were not passed to the F1 plants during the outcross to wtZH, further strengthening the assumption of a new autonomously active state. Taken together, the observed non-Mendelian segregation apparent in M3 and F2 generations in almost all mutant lines already indicated that the genetic identification of the mutated genes would not be trivial, as will be evidenced later.

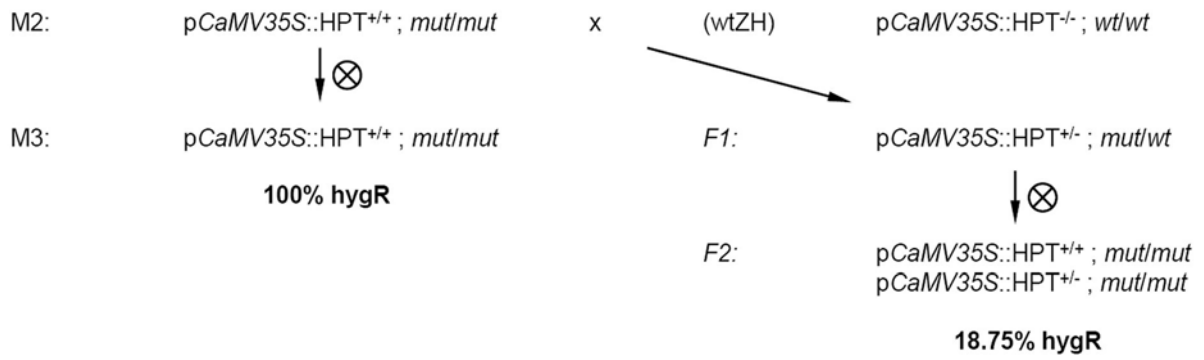

Figure D2: Expected hygromycin resistance in percent calculated for recessive mutations after selfing (M3) and after outcrossing to wtZH (F2). Only genotypes expected to be resistant are shown. The resistance ratios observed from forward mutants are found in Table R1.

4.3. The novel forward mutations result in characteristic changes of structural and epigenetic features at the HPT transgene

The HPT transgene was reactivated in all hygromycin-resistant lines, except for 262 and 350 were no HPT transcripts could be detected (Figure R6A and R8B). Since these two lines were reproducibly shown to be resistant in all tested generations, the lack of HPT transcript suggested alternative resistance mechanisms, e.g. by blocking the uptake of hygromycin. These mutants were not further analysed in the course of this thesis since they would not help to identify trans-acting epigenetic regulators. They might nevertheless be interesting to identify plant-derived antibiotic marker genes for biotechnology purposes (Aufsatz et al, submitted).

Seven mutants had truncated HPT transcripts and rearrangements in the transgene (Figure R6A and C). Interestingly, the rearrangements and deletions affected the 3' end of the transgene by deleting parts of the duplicated 35S promoter and/or adjacent sequences pointing out that this region may be important in regulation of transcriptional gene silencing. They are further characterised by A. Förster (in preparation).

Twelve hygromycin-resistant lines have been found to stably reactivate the HPT transgene in a *trans*-acting manner, i.e. accumulating full length HPT mRNA while maintaining the transgene structure intact. Increased HPT mRNA and secondary transcripts, as seen from Northern blot analysis (Figure R10), were detected in all mutant lines, however with HPT mRNA abundance differing from line to line. This may be due to either partial reactivation after incomplete removal of repressive epigenetic marks or a sign of ongoing re-silencing. Further inbreeding did not result in decreased hygromycin resistance, arguing against re-silencing. Four lines showed highly abundant HPT mRNA levels, comparable to C2R (Figure R11). Interestingly, release of transcriptional gene silencing from the duplicated CaMV35S promoter differed between the mutant lines, as observed from Northern blot analysis (Figure R10B). Due to the HPT transcript that comprises the whole transgenic region, it was not possible to apply PCR-based quantification assays for the secondary transcript. However, the Northern blot analysis suggests that the mutations interfere with silencing at the duplicated promoters in a distinct manner.

Since the initial transcriptional gene silencing at the HPT transgene was associated with increased DNA methylation, a pilot analysis with methylation-sensitive restriction digest indicated that the C2S1-specific hypermethylation at CG and CHG sites was decreased to different levels in the mutant lines (Figure R12). To obtain a more detailed view on the induced changes in DNA methylation, bisulfite sequencing analysis of the two 35S promoters was conducted (Figure R13). Analysis of DNA methylation in a sequence-specific context revealed that CG methylation was significantly decreased in all lines, even in the lines

showing a general methylation increase. CHG methylation, however, was only decreased in some lines, and significant CHH demethylation could only be observed in one mutant (774), while the rest retained or even accumulated CHH methylation (Figure R13). The ^{m}CHH enrichment in 348 and 964-2 also explains the general increase in methylation observed for these mutants since the CHH sites contain 76% of all available cytosines in the analysed sequence. Sequence analysis of the duplicated promoter indicated that total CG and CHG methylation was reduced in all mutant lines except for 964-2 that retained methylation. Asymmetric CHH methylation at the duplicated pCaMV35S showed a minor reduction in five lines (348, 409, 774, 1135 and 1142) but was unchanged in the remaining lines (552, 761, 964-2 and *ddm*1-5) (Figure R13). Methylation analysis at individual cytosines in the analysed promoters showed striking differences in demethylation efficiency. The duplicated promoter was homogenously demethylated in all mutants. Demethylation of the main promoter was biased toward the ASF1 transcription factor binding sites and sequences downstream, as observed in other TGS systems using CaMV35S promoters (Inamdar *et al.* 1991, Linn *et al.* 1990). The 5' end was unchanged or even hypermethylated at most CHH sites in numerous lines (Figure R14). The disparity between the two 35S promoters and between different mutant lines indicated high complexity. While DNA methylation in all sequence contexts and at both promoters was completely removed in 774 and to some degree in 1135, demethylation in the remaining lines was either weaker or inconsistent, depending on the analysed sequence context or promoter. This difference in sequence specificity may come from the fact that sequence-specific DNA methylation in *Arabidopsis* is established and maintained by different enzymes (*DRM*1/2, *MET*1 and *CMT*3, outlined in the introduction). The pathways that regulate the activity and target recognition of these enzymes could be affected either specifically or equally by the mutations. The mutation in 774 for example seems to affect all methylation pathways in a similar manner, while only CG methylation is reduced in 348 and 964-2 and CHH methylation is induced at the first promoter. If this accumulation of CHH methylation in those lines is due to the mutation or an indirect response to loss of CG methylation remains unknown. The latter explanation seems to be more likely since *met*1 mutants induce aberrant CHH methylation at several genomic sites after loss in CG methylation, probably as a rescue mechanism (Mathieu *et al.* 2007, Lister *et al.* 2008). Since CHG and CHH methylation at the first promoter was unchanged or even increased in several mutant lines, one may argue that either methylation at CG sites alone is essential for silencing maintenance at the HPT transgene, or that other - DNA methylation-independent - mechanisms may be involved. Detailed analysis showed that methylation at CG sites was indeed decreased in all lines, especially at the ASF1 binding site. The observed differences between the two promoters may be due to the transcriptional activity of the transgene. Transcription factor binding at the main promoter may be facilitated by loss of

CG methylation at the ASF1 binding site. This in turn would recruit the transcription machinery that elongates the nascent RNA through the HPT coding region and the duplicated promoter until it terminates in the flanking plant genomic region. The ongoing transcription through the duplicated promoter could remove methylation there in general, while inhibiting *de-novo* methylation surrounding the ASF1 binding site in the main promoter, as suggested for active CG rich promoters in mammalian organisms (Bird *et al.* 1981, Macleod *et al.* 1994). The accumulation of CHH methylation at the 5' of the main pCaMV35 could be explained if the transcription machinery occludes the ASF1 site and downstream sequences from CHH *de-novo* methylation, thus redirecting it to the upstream part of the transgene which is lacking transcriptional activity.

Post-translational histone modifications are known to be involved in transcriptional regulation by providing information, similar to, a code, that can be read by the transcription or silencing machinery (Jenuwein and Allis 2001, Kouzarides 2007). Analysis of histone modifications at the HPT transgene revealed that, depending on the epiallelic state, either activating (H3K4me3 and H4acetylation) or inactivating marks (H3K9me2, H3K27me2 and H3K27me3) were detected (A. Förster, unpublished). Concomitant with results obtained from DNA methylation analysis and HPT mRNA quantification, the forward mutants reduced the heterochromatic H3K9me2 and H3K27me2 modifications at all regions analysed, however at different levels (Figure R15). In general, K9me2 and K27me2 levels in 348 and 964-2 were just slightly reduced, whereas no signals could be detected in the remaining lines. H3K27me3, which is known to be associated with tissue-specific genes (Zhang *et al.* 2007), was strongly reduced at the main promoter in all mutant lines, compared to C2S1. The coding region and the duplicated promoter remained either unchanged or decreased only slightly (Figure R15). The activating histone modifications H3K4me3 and histone H4 acetylation at the HPT transgene showed a weak or strong increase depending on the mutant or region analysed. In general, the levels observed in C2R in the HPT promoter and coding region were not reached, even with the strongest increase in 774. This might be due to the fact that in some lines DNA methylation or other repressive histone marks were still remaining and might interfere with complete accumulation of H3K4me3 or H4 acetylation. The duplicated promoter showed alternating enrichment for either H3K4me3 or H4acetylation. The mutants that were enriched in H3K4me3 had lower levels of H4acetylation, and vice versa (Figure R15). This alternation might further reflect the already mentioned residual silencing marks at the duplicated promoter. Nucleosome occupancy and the availability of modified histones at certain regions affect accessibility of the underlying sequence (Wagner 2003). Histone H3 was found to be strongly enriched at the main promoter in the inactive C2S1 line compared to the active C2R line, while less abundant in

the downstream regions. All forward mutant lines had a significant decrease in H3 abundance at the HPT promoter and coding region, strongest in 1135. The observed difference in H3 abundance between the active and inactive lines at the HPT promoter and coding region fits very well to the described transcriptional activity in these lines. The stronger depletion of histone H3 that was observed in some lines could be due to ongoing re-positioning or complete removal of nucleosomes from the analysed regions (all except 774). This in turn could have a big impact on chromatin condensation and might explain transcriptional activity from these lines, even in the presence of heterochromatic marks at the remaining histones (348, 409 and 964-2). Furthermore, the low abundance of nucleosomes at these regions might also explain the low level of activating marks in some mutants.

Analysis of histone modifications upstream of the transgene insertion site showed that the region located immediately upstream was changed in a similar way as reported for the main pCaMV35S, depending on the mutation and analysed modifications (Figure R16A). C2S1-specific H3K9me2 in the gene *BME*3 2kb upstream was reduced in all analysed mutant lines (Figure R16B). This correlated with the transcriptional reactivation of *BME*3 in all mutant lines (Figure R35B). However, it remains unknown if this is solely due to the removal of the H3K9me2 marks or due to the activity of the reactivated 35S promoter in the neighbouring transgene.

4.4. Some, but not all forward mutant lines induce changes at endogenous targets

While the HPT transgene lost transcriptional silencing in the mutants accompanied by more or less drastic loss of repressive epigenetic marks, it was of interest to investigate whether this would occur also at endogenous sequences caused by interference with maintenance of PA-TGS. The first step was to analyse the mutant lines for apparent growth and developmental phenotypes as these may result from changes in gene expression at endogenous loci. Five lines showed strong growth effects (Figure R17A). Interestingly, the remaining lines did not display any obvious phenotypes, besides the mentioned chlorophyll phenotype in one line (964-1) that was possibly due to an unlinked second-site mutation (Figure R17B). In addition, seed size measurements and ploidy analysis revealed that 552 was tetraploid. However, how this mutant line became tetraploid, how this affected gene regulation at the transgene and endogenous targets and whether there is a connection with PA-TGS remains unclear.

In order to link phenotypic observations to the mutations, molecular analysis was done to investigate potentially altered epigenetic properties at specific sites or at a global level. Measurements of global 5-methyldeoxycytosine levels revealed that the mutants with

unaltered phenotypes (348, 409, 761, 964) did not change global DNA methylation, whereas the lines with strong developmental phenotypes (552, 774, 1135 and 1142) had drastically reduced 5-mdC levels (Figure R19). Since DNA methylation in *wt Arabidopsis* occurs mostly at repeated regions, changes in DNA methylation at the pericentromeric TSI and centromeric 180bp repeats were analysed, knowing that these are hypomethylated in mutants deficient in DNA methylation such as *met*1 or *ddm*1 (Mittelsten Scheid *et al.* 1998, Jeddeloh *et al.* 1999). Concomitant with the results obtained from global methylation measurements, strong DNA hypomethylation was observed in the mutants 552, 774, 1135 and 1142, arguing for an impaired DNA methylation machinery in these lines (Figure R20). Loss of DNA methylation at centromeric repeats was shown to drastically influence their nuclear chromocenter morphology by disrupting compact heterochromatic structures (Soppe *et al.* 2002, Probst *et al.* 2003). The same was true for the forward mutants that have been shown to be impaired in maintaining DNA methylation (Figure R21). The nuclear localisation of 5-mdC normally overlapping with dense DAPI-stained chromocenters in the *wild type* was completely erased in these lines. However, in the mutant line 774, 5-mdC staining was still maintained at the chromocenters, although slightly weaker and more de-condensed than in the *wild type* (Figure R22). This is analogous to the moderate demethylation and decondensation of centromeric repeats observed in this line where the effect was not as pronounced as in 552, 1135 or 1142 (Figure R20 and 21). In contrast, demethylation at both 35SCaMV promoters flanking the HPT transgene was most pronounced in 774 (Figure R13 and R14). This could mean that the changes in global DNA methylation measured in these four lines may reflect target-specific demethylation, depending on the mutation.

Histone modifications are known to have distinct nuclear localisation patterns (Fuchs *et al.* 2006, Pflüger and Wagner) and undergo severe changes in diverse mutant backgrounds (Mathieu *et al.* 2003). The same was observed in some of the forward mutants. H3K9me2 signals were completely disassociated in 552, 774, 1135 and 1142 (Figure R23A). Since DNA methylation and H3K9 dimethylation at repetitive loci are known to be tightly interwoven (Bernatavichute *et al.* 2008, Johnson *et al.* 2007, Lindroth *et al.* 2004, Soppe *et al.* 2002), this reduction was not a surprise considering the reduced endogenous DNA methylation. Detailed analysis of histone modifications at a transcriptionally inert retrotransposon located in euchromatin confirmed that H3K9 dimethylation was lost in 552, 774 and 1135 while the activating mark H3K4me3 was enriched (Figure R16C). H3K27 trimethylation is associated with euchromatin and excluded from heterochromatic chromocenters and the nucleolus in wild type nuclei and is thought to be involved in gene regulation of tissue-specific genes (Naumann *et al.* 2005, Turck *et al.* 2007, Zhang *et al.* 2007). Euchromatic staining of H3K27me3 was unchanged in all lines, except that in 552, 774, 1135 and 1142 it extended

into regions previously occupied by heterochromatin. Aggregates of H3K27me3 staining that were not observed in wild type nuclei could be detected at high rates in nuclei from these lines. This aggregation may be due to H3K27me3 taking over a role in silencing at some loci after the severe loss of H3K9me2 marks in 552, 774, 1135 and 1142, as reported for *ddm*1 and *met*1 mutants (Mathieu *et al.* 2003).

Perturbation of DNA methylation and histone modifications by genetic means or by inhibitors is frequently associated with transcriptional reactivation of transposons that are normally tightly silenced and not transcribed in wild type plants, preventing their spreading throughout the plant genome (Zilberman and Henikoff 2004). Therefore, transcriptional reactivation of such transposons can serve as an indicator for interference with epigenetic regulation (Jeddeloh *et al.* 1998) (Jeddeloh *et al.* 1999, Kankel *et al.* 2003). Here, the maintenance of transcriptional gene silencing at transposons, retro-transposons and Athila-related pericentromeric repeats (TSI) was analysed. The lines displaying severe effects on endogenous DNA methylation and histone modification patterns showed strong up-regulation from all analysed transposons (Figure R24), including LINE1-4 that displayed changes in H3K9 di- and H3K4 trimethylation in the same lines (Figure R16C). The fact that no up-regulation could be observed from the soloLTR retrotransposon that is known to be regulated by RNA-dependent DNA methylation (RdDM) (Huettel *et al.* 2006) suggests that all the mutants analysed so far are most likely not affecting the RdDM pathway.
The results obtained from the transposon reactivation experiments stimulated to analyse whether the mutations would reactivate a highly repetitive, silent transgene, encoding a β-glucuronidase transgene (TS-GUS, Morel *et al.* 2000) known to loose silencing in several mutant backgrounds (Elmayan *et al.* 2004) or after treatment with DNA methylation inhibitors (Baubec *et al.* 2008). F2 generations resulting from crosses between mutant lines and the TS-GUS line showed diverse degree and patterns of GUS reactivation (Figure R25). The observed staining patterns might be due to different mode of action or tissue specificity of the mutated genes or due to a spreading effect that counteracts reactivation.

Treatments with the DNA methylation inhibitor zebularine and/or the histone deacetylase inhibitor trichostatin-A have shown to release transcriptional gene silencing from numerous target genes and repeats (Chang and Pikaard 2005). In addition, developmental defects and growth inhibition have been linked to reduced levels of DNA methylation as observed in several DNA methyltransferase mutants and upon pharmacological interference with cytosine methylation (Finnegan *et al.* 1996, Ronemus *et al.* 1996, Xiao *et al.* 2006). This led to the analysis for increased sensitivity to demethylating drugs and/or HDAC inhibition in the mutant lines (Figure R26). While zebularine treatments induced minor growth retardation in

all analysed lines, including controls, some mutants displayed stronger sensitivity and even lethality to prolonged zebularine treatments. Interestingly, only one out of the four lines with lethal outcome after zebularine treatments was genetically impaired in global DNA methylation. The remaining mutant lines with strong changes in global methylation (552, 1135 and 1142) were only slightly more sensitive, similar to *ddm*1-5 that was used as a control, and were able to survive the zebularine treatment. Treatments with trichostatin-A resulted in increased, but recoverable sensitivity in some mutant lines. Combined treatments of TSA and zebularine were deleterious in all analysed lines. The combinatorial and specific effects of zebularine and/or TSA treatments observed in the individual mutants may reflect the role of the mutated gene in regulating DNA methylation and/or histone modifications. It was surprising to see that some mutant lines that had no or minimal changes in DNA methylation at endogenous targets showed an increased sensitivity to the demethylating drug zebularine (348, 409). This could be due to their involvement in redundant regulatory pathways that overlap with DNA methylation. Here the mutation would display its full effect only in a methylation-deficient background. Another possibility is that the mutants are involved in regulatory mechanisms influencing genome integrity, chromatin assembly and DNA repair pathways. This fits well to the fact that zebularine incorporates into DNA during replication and induces covalent bounds with methyltransferases, therefore forming obstacles for any proteins that interact with DNA. Indeed, such a case was observed in prokaryotes where 5-azacytidine, that functions similar to zebularine, led to replication fork stalling (Kuo *et al.* 2007). The generated nucleoprotein complexes would require efficient DNA repair or functional chromatin assembly to be removed or bypassed. In addition, assays for zebularine sensitivity of mutants deficient in DNA repair and chromatin assembly (*fas*1, *bru*1, *ku*70) showed a similarly increased sensitivity as observed in the epigenetic mutant lines (data not shown), further suggesting that these mutants may be affected in genome stability, replication or chromatin assembly. It has been already reported that cytosine methylation inhibitors and HDAC inhibitors have often antagonistic and non-overlapping effects on several genes (Chang and Pikaard 2005), arguing that histone de-acetylation and DNA methylation are not interdependent mechanisms. The outcome of the TSA treatment alone and in combination with zebularine could indicate that most of the mutations analysed here were not affected by HDAC inhibition since they may be involved in regulatory processes at targets that do not require HDAC activity. The few mutants affected by TSA treatments could then be either involved in regulating targets overlapped by HDAC activity in a redundant manner or may be generally enhanced by the lack of histone deacetylation. Combined treatments showed the tremendous impact of DNA methylation and histone deacetylation on plant development, since all lines were severely affected.

4.5. Identification of the mutated genes

Following the results obtained from the analysis of epigenetic changes at endogenous and global targets, the forward mutant lines could be divided into two groups (see overview table 3.9). In both groups, the mutants displayed release of transcriptional gene silencing from the HPT transgene that led to stable hygromycin resistance throughout all analysed generations. However, the two groups differed in transcriptional regulation and chromatin features at endogenous targets. 187, 552, 774, 1135 and 1142 showed reactivation of numerous silent transposons and repetitive regions, a decrease in global 5-mdC levels, changes in nuclear localisation of heterochromatic histone modifications and characteristic growth phenotypes. None of these were observed in the other mutant lines (348, 409, 761, 964-1/2 and 1361) (Overview table 3.9). Assuming that the epigenetic changes and reactivation at the HPT transgene in the first group may be due to the global effects caused by the mutations, the changes observed in the remaining lines may be site-specific only, as a result of HPT reactivation and transcription. However, it still remains unknown for all lines how transcriptional reactivation was initiated. Therefore, the identification of the mutated genes was of high interest in order to shed light on the involved regulatory mechanisms.

Mutagenesis was exerted by transformation with *Agrobacterium tumefaciens* carrying a T-DNA with a BASTA resistance transgene, assuming that this would randomly integrate and allow straightforward identification of the mutated site by thermal asymmetric interlaced (TAIL) PCR that amplifies insertion flanking sites (Liu *et al.* 1995). However, it turned out that this procedure was ineffective in almost all mutant lines, connected with several problems. The first problem encountered was that most of the mutant lines had multiple T-DNA insertions, which were either linked as complex insertions or segregating as independent insertion sites (Figure R27, Table R7). Second, several lines were found to have no T-DNA insertions at all. Therefore the mutations in these lines were either caused by insertions of plasmid sequences outside the T-DNA, which was confirmed by probing with the vector sequence (Figure R27B), or by small deletions and rearrangements in the genome as observed in the mutations within the HPT transgene (Figure R9). These multiple insertions and deletions made the identification of the mutated sites very laborious or almost impossible. A further and even more significant obstacle was the observation that, according to segregation analysis in several outcross generations and comparing observed and expected hygromycin resistance ratios, in almost all mutants the identified T-DNA or vector insertions were unlinked to the hygromycin resistance (Tables R1 and R7). This suggested that the hygromycin resistance was either caused by secondary site mutations such as small deletions or re-arrangements in the genome caused by the T-DNA insertion mechanism, or the expression of the HPT transgene itself became autonomous – once reactivated by the introduced mutation (either T-DNA or vector), it may remain in an active state, without

requiring the mutation to be present. The latter possibility would therefore mask any potential linkage between mutation and hygromycin resistance. The assumption of a “switch and carryover” effect was strongly supported by the high hygromycin resistance ratios observed in F2 generations from crosses between M2 and wild type ZH (Table R7). Another explanation for the observed high resistance ratios in F2 could be a dominant behaviour of the introduced mutation (up to 56.25%) or its tight linkage to the HPT transgene (up to 75%). Regardless of the problems mentioned, TAIL PCR was conducted in some lines and resulted in the identification of numerous targets that had T-DNA insertions (Table R8). Many of those targets were not further analysed since those were either identified as pseudogenes, intergenic hits or involved in biological processes obviously unrelated to gene regulation. However, some genes whose role in gene regulation seemed plausible were analysed for their potential to release transcriptional gene silencing from a naïve HPT transgene. From this analysis a novel *DECREASE IN DNA METHYLATION* 1 (*DDM*1) mutant allele in 1142 was confirmed to be able to reactivate the silent HPT transgene. Due to the described carryover effect observed in most of the mutant lines, linkage to the mutation could not be addressed with the residing hygromycin resistance gene anymore. Therefore, an additional silent marker gene was used to confirm and visualize the presence of the mutations. As mentioned earlier, a transcriptionally silent GUS marker gene was reactivated in several F2 plants from crosses to the mutant lines (Figure R25). Based on the similarities in GUS staining patterns and other endogenous changes observed 1142 that was deficient in *DDM*1, 552 and 1135 could be identified as two novel *ddm*1 mutants. In addition, Southern blot analysis of the T-DNA insertions in the GUS transgene-reactivating lines further suggested HPT-reactivation carryover effects since GUS-positive plants identified in segregating F2 populations showed identical patterns of T-DNA fragments (Figure R28C). The GUS reporter was only reactivated in the mutant lines that showed global effects (187, 552, 774, 1135 and 1142). Therefore, using an additional silent marker gene was helpful for the mutants affecting repetitive sequences but could not be used to identify mutations in the lines that did not display any obvious changes at endogenous targets. Assuming that the mutations generated in the lines lacking endogenous effects may also have multiple and complex insertions, small deletions and carryover leading to masked linkage, their further mapping is impossible with classical methods since the presence of the mutation in a certain plant or generation cannot be validated. Therefore, mutation identification in these lines was abandoned and awaits high-throughput sequencing.

4.6. Transcriptome analysis reveals global changes in gene-expression in some but not all analysed mutant lines

Changes in gene-expression from endogenous genes have been reported to occur in numerous mutant backgrounds affecting transcriptional silencing or chromatin structure (Zhang *et al.* 2006, Jordan *et al.* 2007). In order to analyse if the mutants found in this screen had similar effects on endogenous targets, RNA from two (774, 1135) lines with and two lines (348, 409) without obvious phenotypes and changes in transcriptional gene regulation at transposons were hybridised to Affymetrix ATH1 expression arrays. Concomitant with previous results, significant changes in gene expression at numerous targets including coding genes and transposons could only be detected in 774 and 1135, but not in 348 and 409. In the microarray experiments, 774 could be identified and confirmed as a novel *HOMOLOGY DEPENDENT GENESILENCING* 1 (*HOG*1) mutation through analysis of the down-regulated genes in this mutant (Figure R32 and R35). The majority of up-regulated targets in 774 and 1135 were transposons (46% and 71%) that highly overlapped between the two mutant lines and were mostly clustered at highly repetitive regions such as the peri-centromere or the heterochromatic knob on chromosome IV (Figure R29). This result was in accordance with previous experiments and reports that showed strong up-regulation from transposable elements in these mutant lines and other mutant alleles (Gendrel *et al.* 2002, Jordan *et al.* 2007). The strong overlap in affected transposons indicated that both mutants might be involved in regulation of the same targets; however the ATH1 array does not represent all known transposons and retrotransposons encoded in the *Arabidopsis* genome. Thus target specificity to subclasses of mobile elements and repeats may have been obscured by the small subset of probes available on the array. In contrats to the transposons, the coding genes that showed a significant increase in gene expression were overlapping only to a minor extent between the mutants. This may reflect some biologically meaningful differences in target specificity. Gene ontology analysis of the up- and down-regulated gene sets in 774 and 1135 did not detect specifically enriched or depleted GO terms that would indicate involvement of the mutated genes in characteristic processes or functions. This may be due to either unspecific regulatory functions on genes throughout all GO categories, or to the fact that a high proportion of *Arabidopsis* coding genes are still uncharacterised and thus not assigned to any GO terms. While the up-regulated transposable elements were shown to be clustered at specific chromosomal regions, the up- and down-regulated coding genes were evenly distributed throughout all chromosome arms, therefore contradicting previous reports suggesting an enrichment of up-regulated coding genes from pericentromeric regions in the *hog*1-1 mutant background (Jordan *et al.* 2007).

However this may be due to the different mutant alleles and/or ecotype backgrounds analysed here and previous reports (h*og*1-7 in Zürich vs. *hog*1-1 in Landsberg).
The only gene that was found to be significantly up-regulated in 348 and 409 was *BME*3, which is located 700 bp upstream of the HPT insertion site and has been shown to be over-expressed in all mutant lines reflecting transgene-dependent regulation (Figure R34B). In addition, an unknown, putative transposable element that is found 30 kb downstream of the HPT insertion site and has two more copies on chromosome II and at the heterochromatic knob on chromosome IV was found to be up-regulated in 348. This element showed no transcriptional activity in *wt* plants, in agreement with strong enrichment of DNA methylation and siRNAs according to the AnnoJ database (Lister *et al.* 2008) (Figure R33). How and if this observation is linked to the transcriptional regulation of the HPT transgene is unknown. Transcriptional reactivation from this element may have been spurious and totally unrelated to the HPT reactivation. In addition, due to the high sequence similarity (99%) between the copies at the different locations, the source of transcripts could not be defined. However, expression analysis detected strong up-regulation of this repeat also in the remaining mutant lines (Figure R34C).

4.7. Identified mutations and their involvement in PA-TGS at the HPT transgene

From the twenty one hygromycin-resistant mutant lines identified in the forward screen, nineteen mutant lines were shown to induce loss of silencing maintenance at a hygromycin phospho-transferase (HPT) transgene that was previously silenced during tetraploidisation (Mittelsten Scheid *et al.* 2003). The low number of reactivating mutants generated in this screen suggests that the factors involved in PA-TGS maintenance are either very essential for plant development and therefore causing lethality upon mutation, or redundant, so that their functions can be replaced by similar components. Another very likely explanation is that the regulation of PA-TGS is either tissue-specific or meiosis-dependent, so that potential mutations may not be apparent by the applied screening method. Seven HPT- reactivating mutant lines have been identified as mutations within the HPT transgene that were able to initiate production of truncated HPT transcripts through re-arrangements or deletions in the transgene. These mutants could play a major role in the identification of the sequence elements required for establishing and/or rigorously maintaining transcriptional silencing at the HPT transgene (A. Förster unpublished, and this thesis). How the duplicated region is involved in this regulation mechanism is still unclear. The duplicated promoter or the adjacent sequences could act as a source of yet unidentified short range signals that may spread into the whole transgene and sustain transcriptional inactivity. Deletion of that region in the HPT

mutants might have uncoupled this spreading, resulting in the reactivation of the HPT transgene. Another possibility is that this duplicated region could protect the main promoter from being reactivated. This was evident from zebularine treatments, where the duplicated promoter got strongly reactivated while the main promoter remained silent. The same was observed in the analysed mutant backgrounds from the reverse screen where only the duplicated promoter was reactivated without any effects on the main promoter (T. Baubec, M. Milos and O. Mittelsten Scheid, unpublished). In this case, transcriptional reactivation from the duplicated promoter could lead to increased accumulation of non-coding RNAs that may provide a substrate for re-enforcing or even re-inducing silencing at the main promoter. This hypothesis is currently under investigation (A. Förster and T. Baubec). The relatively high number of mutants affecting the HPT transgene, in relation to the trans-acting mutations, indicates that the mutagenesis and the screen were saturated.

The twelve remaining, HPT-reactivating lines that had no detectable re-arrangements or deletions in the HPT transgene were assumed to be affected in *trans*-acting regulatory mechanisms. Thorough characterisation of these mutant lines revealed that, besides the epigenetic changes at the reporter gene and its stable reactivation in all lines, seven mutations did not show any effects on endogenous targets known to be affected by other epigenetic regulators, while five lines led to reactivation of numerous endogenous transcriptionally inert targets accompanied by global changes in epigenetic features and changes in growth phenotypes. Regrettably, none of the mutations associated with release of HPT silencing without inducing global changes could be identified, due to the reasons outlined before. These seven lines (112, 348, 409, 761, 964-1, 964-2 and 1361) would require further investigation to identify changes at endogenous targets, which were either obscured or too minor to be detected by the selective and stringent analysis applied. These endogenous changes could then be used as additional markers for linkage analysis and identification. In addition, subtle mutations at the HPT transgene or its surrounding sequences, such as minor deletions or re-arrangements missed out during initial analysis, need to be excluded by careful sequence analysis of the transgene and its flanking regions. *Epi*-mutations caused by ectopic DNA hypo –or hypermethylation, similar to those already described in *Arabidopsis*, (Jacobsen et al. 2000) might also have induced release of gene silencing from the HPT transgene. Such ectopic changes in chromatin features at the CaMV35S promoters could have occurred during T-DNA mutagenesis. Induction of salicylic acid (SA)-responsive promoters have been described as an early systemic response to pathogen infection (Ward *et al.* 1991, Wildermuth *et al.* 2001), and since such SA responsive elements were first identified in CaMV35S promoters (Lam *et al.* 1989), it can not be excluded that the HPT promoter could have been reactivated during *Agrobacterium* infection.

Four out of five mutant lines showing global effects could be identified based on available additional markers (TS-GUS) and transcriptome analysis. Three mutant lines carried novel mutations in the already known transcriptional gene silencing effector *DDM*1, which encodes an ATP-dependent SWI2/SNF2 chromatin remodelling factor involved in maintenance of DNA methylation and H3K9 di-methylation at repeats (Jeddeloh *et al.* 1999). However, these novel alleles turned out to have a stronger effect on maintenance of TGS at the HPT transgene since they were able to induce hygromycin resistance in earlier generations than the already known *ddm*1-5 mutation in the reverse screen (Figure R36B). A comparison of changes in global DNA methylation in the novel alleles further confirmed this observation (Figure R19). These differences between the *ddm*1-5 allele, assumed to be a null mutation due to a translational stop codon in the 5' region, and the novel mutations could be due to several factors. Side by side comparison of the mutant alleles indicated that the mutations in the novel alleles occurred at essential domains and conserved motifs (especially 1135 and 1142) required for functional chromatin remodelling. A closer look on *ddm*1-5, with the 82 bp insertion in exon II that produces the preliminary stop-codon, reveals that this could be theoretically either spliced out or bypassed via using one of the translational initiation sites downstream of the mutation, thereby still producing low amounts of functional DDM1 protein (Figure R36A and C). In addition, the increased ploidy observed in 552 (*ddm*1-11) might have intensified the *ddm*1-phenotype, assuming that chromatin regulation in a tetraploid would require double amounts of functional DDM1 protein. Since *ddm*1-5 was already shown to partially interfere with silencing maintenance at the HPT reporter gene (Mittelsten Scheid *et al.* 2003, Milos 2006) the identification of 3 additional *ddm*1 alleles was a functional proof for the forward genetic approach and furthermore indicated mutational saturation.

The other gene involved in maintaining transcriptional gene silencing at the HPT promoter and many other targets in the genome was *HOG*1. Mutant 774 was identified as carrying a novel partial loss-of-function mutant allele: *hog*1-7. *HOG*1 (or SAHH1) encodes an S-adenosyl-L-homocysteine (SAH) hydrolase that is required to hydrolyse the by-product of cellular trans-methylation reactions (Figure D3) and thereby involved in maintaining transcriptional gene silencing at numerous targets (Furner *et al.* 1998, Rocha *et al.* 2005, Jordan *et al.* 2007, this thesis). Hydrolysation of SAH into homocysteine is essential for recycling of the methyl-group donor S-adenosyl-L-methionine (SAM) and also required to prevent inhibition of trans-methylation reactions through increased levels of SAH (Moffatt and Weretilnyk 2001) (Figure D3).

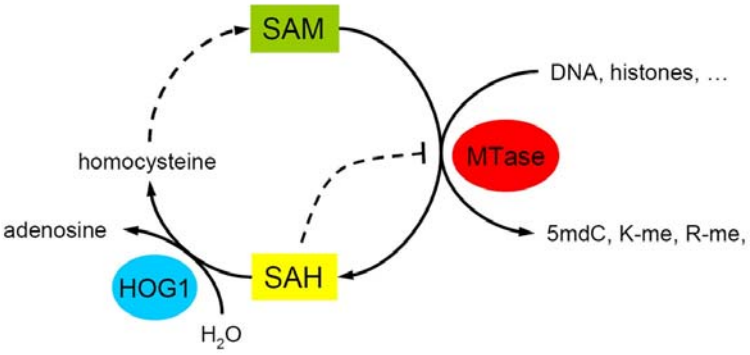

Figure D3: Trans-methylation cycle and HOG1 function. Cellular methylation reactions catalysed by methyl-transferases require (SAM) S-adenosyl-L-methionine as methyl-group donor that is converted into (SAH) S-adenosyl-L-homocystein. The latter is hydrolysed by HOG1 into adenosine and homocysteine to ensure SAM re-synthesis. Increased levels of SAH have been shown to interfere with trans-methylation reactions.

Since all trans-methylation reactions in the cell require SAM as a methyl donor, mutations in the SAH hydrolase would have a global and strong effect on all biochemical processes involving trans-methylation. This is evident from the increased zygotic lethality observed in *hog*1 point mutants, even more pronounced in insertional mutants (Rocha *et al.* 2005). Partial loss of function mutations interfere with DNA methylation, as reported previously (Furner *et al.* 1998, Rocha *et al.* 2005) but likely also with histone methylation, as suggested by Rocha and colleagues. Indeed, cytological changes in H3K9me2 and H3K27me3 methylation were observed in the *hog*1-7 background (Figure R23), and ChIP analysis at single reactivated targets revealed a loss of repressive histone methylation, but also affected H3K4 methylation (Figures R15 and R16).

It is striking that the mutants identified in this screen and able to induce loss of polyploidy-associated transcriptional gene silencing from a HPT transgene were either *ddm*1 or *hog*1. Both were known from previous studies (Jeddeloh *et al.* 1999, Furner *et al.* 1998, Rocha *et al.* 2005) and have been shown here or previously (Gendrel 2002, Habu 2006) to equally affect DNA and histone methylation. This, together with the results obtained from the reverse screen and from zebularine treatments (Milos 2006, this thesis), suggests that maintenance of PA-TGS at the main promoter of the HPT transgene may involve DNA and histone methylation in a combined and redundant manner, therefore requiring mutations that affect both pathways in order to disrupt silencing maintenance. This can occur obviously in at least two different ways: in *ddm*1, the lack of functional protein and its chromatin remodelling function may not attract or not facilitate the accessibility of DNA and histone methyltransferases to the transgene, while reduced functional *HOG*1 would reduce the steady state levels of S-adenosine-L-methionine necessary for both, DNA and histone methylation reactions (Figure D4). In addition, the previously described duplicated region of the transgene may play a crucial role in inducing and/or maintaining stable PA transcriptional gene silencing. A model for the interconnection of these factors regulating the HPT silencing is depicted in Figure D3.

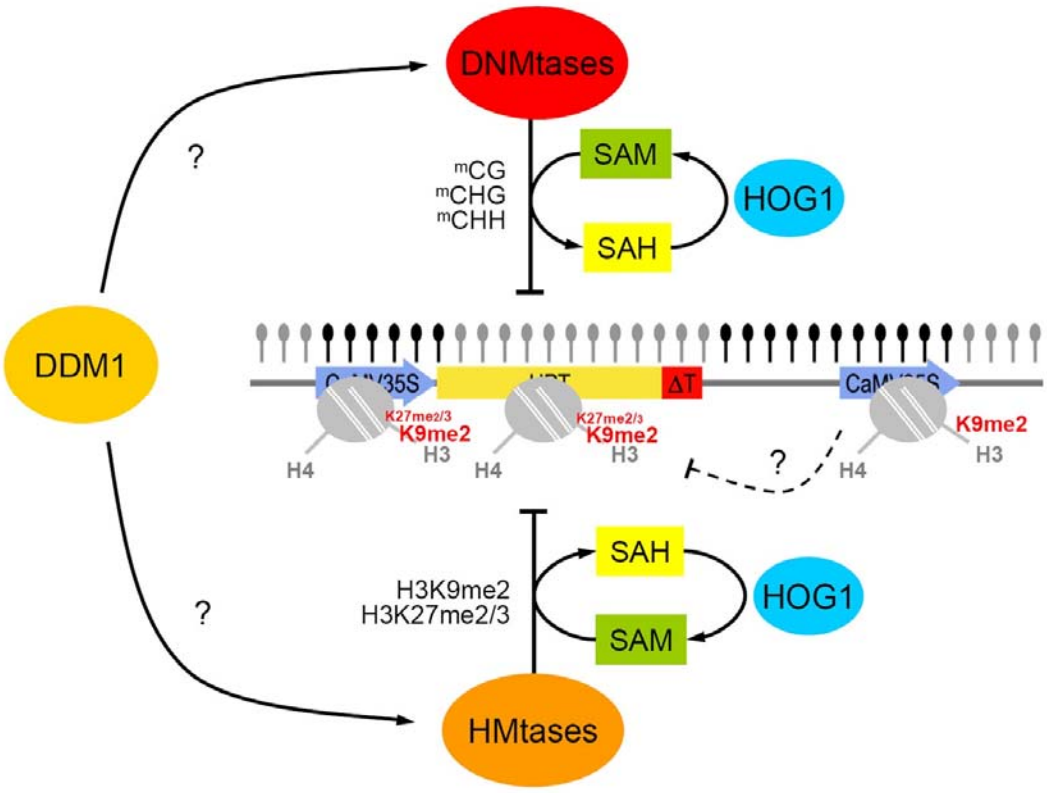

Figure D3: Current model of PA-TGS at the HPT transgene.
Silencing maintenance requires combined action of histone and DNA methyltransferases. Functional DDM1 is either recruiting or facilitating the accessibility of the methyltransferases to the transgene by remodelling the surrounding chromatin. DNA methylation of CG, CHG and CHH, but also histone methylation at K9 and K27 requires S-adenosyl-L-methionine (SAM) as a methyl-group donor, the pool of which in turn is regulated by functional HOG1.

4.8. Concluding remarks and future perspectives

Polyploidy-associated transcriptional silencing induced very stable epialleles that faithfully maintained their transcriptional inactivity during subsequent generations, changes in ploidy, treatments with DNA methylation- or HDAC inhibitors and in the background of numerous TGS mutants. Therefore, maintenance at these epialleles was thought to be regulated by novel mechanisms or by multiple redundant layers of epigenetic regulation. A forward genetic screen finally revealed two factors involved in PA-TGS maintenance. The function of these factors (*DDM*1 and *HOG*1) led to the hypothesis that PA-TGS maintenance requires combined and redundant action of DNA and histone methylation, thus explaining the robust maintenance. To confirm this hypothesis, follow-up experiments are required involving combined interference with DNA and histone methylation inhibitors or double mutant analysis. In addition, the involvement of *DDM*1 and *HOG*1 not only in maintenance but also

in establishing PA-TGS silencing could be challenged, but would require to perform genetic analysis in a polyploid background. Furthermore, the novel alleles could be used to unravel the yet enigmatic molecular function of DDM1 or to analyse genome- and epigenome-wide effects of SAM deprivation. Also, the remaining unidentified mutant lines lacking endogenous effects and the HPT transgene mutants are potentially useful resources to identify novel molecular mechanisms connected with polyploidy-associated gene silencing in *Arabidopsis*.

5. MATERIALS and METHODS

5.1. Plant growth and chemical treatments

Cold-treated seeds were sterilized in 5% sodium hypochlorite and 0.05% Tween-80 for 6 min, washed and air-dried overnight. Sterilized seeds were germinated at 21°C for 3 days in Petri dishes containing agar-solidified germination medium and grown in growth chambers under 16 h light / 8 h dark cycles at 21°C. For hygromycin (CALBIOCHEM), phosphinotrycine (BASTA, Riedel de Haën), trichostatin-A (TSA, SIGMA) and zebularine (SIGMA) treatments, seeds were sown and grown directly on drug-containing plates under the same conditions as described above. Hygromycin (10µg/ml), BASTA (15µg/ml) and zebularine (20, 40 and 80 µM) in aqueous solution or TSA (0.5 µg/ml) dissolved in DMSO were added to the germination medium before solidifying. For recovery, plants were transferred to drug-free growth medium after 14 or 21 days.

5.2. Nucleic acid isolation and gel-blot analysis

Seedlings were harvested as pools of 50-100 plantlets, shock-frozen in liquid nitrogen and homogenized by vortexing for 1 min using 2-3 ceramic spheres of 1 cm diameter. Rosette and stem leafs from 3-5 adult plants were harvested, shock-frozen in liquid nitrogen and homogenized. Homogenized plant tissue was subsequently used for DNA or RNA extraction using Phytopure (Amersham) or RNAeasy (Qiagen) kits, respectively.

For Southern blot analysis, 10 µg of genomic DNA was digested overnight with 1-2 U restriction enzymes. Subsequently, samples were electrophoretically separated on TAE agarose gels, depurinated for 10 min in 250 mM HCl, denaturated for 30 min in denaturation solution containing 0.5 M NaOH and 1.5 M NaCl and neutralized twice in 0.5 M Tris, 1.5 M NaCl and 1 mM EDTA at pH7.2 for 15 min. For Northern blot analysis, 10 µg of total RNA was denatured with 15% glyoxal and DMSO for 1 h at 50°C and separated using 1.4% agarose gels in 10 mM sodium phosphate buffer pH7 in a Sea2000 circular flow electrophoresis chamber (Elchrom Scientific). DNA and RNA gels were blotted onto Hybond N+ (Amersham) membranes over night with 20x SSC, washed and UV-crosslinked using a Stratalinker (Stratagene). Hybridization was performed as described by Church and Gilbert (1984). Radioactive (50 µCi) dCT-α-^{32}P (Amersham or Hartmann Analytic) labeled sequence-specific probes were synthesized from 25 ng of DNA using the Rediprime labeling kit (Amersham) and purified on G50 Probequant (Amersham) columns. Signals were detected with Phosphorimager Screens (Biorad) and scanned with a Molecular Imager FX (Biorad).

5.3. Cation-exchange high pressure liquid chromatography

Total cytosine methylation was determined as described by (Rozhon *et al.* 2008). In short, 5 µg of genomic DNA was digested overnight at 37°C with 0.0025 U nuclease P1 and 0.5 U DNaseI in 20 mM acetic acid, 20 mM glycine, 5 mM $MgCl_2$, 0.5 mM $ZnCl_2$ and 0.2 mM $CaCl_2$, pH5.3 in a total volume of 50 µl. Subsequently, 5 µl of 0.1 M NaOH and 1 U calf intestine alkaline phosphatase were added and the mixture incubated for a further 24 h. Samples were acidified by addition of 44 µl of 12 mM HCl prior to injection into the HPLC system equipped with a 125 x 4 mm Nucleosil 100-10 SA column (Macherey-Nagel) preceded by a Valco 2 µm inline filter. The mobile phase consisted of 60 mM acetic acid and 15% acetonitrile, pH4.8, with a constant flow rate of 1.5 ml/min. UV detection was performed at 277 nm with a bandwidth of 10 nm with a PDA-100 photodiode array detector, and chromatograms were analyzed with Chromeleon 7 (Dionex). All samples were analyzed in technical triplicates and 5-mdC values were expressed as percent of total cytosine.

5.4. Reverse transcription PCR and real time PCR

Prior to reverse transcription, 30 µl RNA solution was treated with 5 U DNase I (MBI Fermentas), 0.4 U Rnasin and 4 µl of 10x DNase I buffer for 40 min at 37°C to remove residual DNA contamination in the RNA samples, extracted with phenol:chloroform (24:1) and subsequently ethanol-precipitated. Reverse transcription was performed on 1 µg of RNA with 0.2 µg random hexamer primers (MBI Fermentas) using 1 U RevertAid M-MuLV-RTase, RNaseH- (MBI Fermentas) at 42°C for 1½ h. The cDNA obtained thereby was used for PCR and real time PCR. Standard PCR was performed with True-Start Taq polymerase (Promega) and the following primers: HPT-qF: 5'-gggtaaatagctgcgccgatggtt-3', HPT-qR: 5'-cacggcgggagatgcaataggtc-3', HM3F: 5'-tcccaatacgaggtcgccaac-3', HPT-RTR: 5'-tatcggc gagtacttctacac-3', LINE1-4F: 5'-cccatggtgaccaagagttt-3', LINE1-4R: 5'-tcaatgtcggagacctcctc-3', Ta3F: 5'-gattcttactgtaaagaacatggcattgagaga-3', Ta3R: 5'-tccaaatttcctgaggtgcttgtaaac-3', LTR625F: 5'-aactaacgtcattacatacacatcttg-3', LTR625R: 5'-aattaggatcttgtttgccagcta-3', CACTA-F: 5'-ggctagctgtccgactcaatgacct-3', CACTA-R: 5'-cagacatcctttccttcagcttagc-3', MULE2-F: 5'-ctgtccgcgagtgtcatcaagtagc-3', MULE2-R: 5'-gatacttgttgacaagtgtttagcaagcc-3', ROS1-F: 5'- agaagaaattcctaccatca-3', ROS1-R: 5'-accgttcttcgaggtaattc-3', FLCqF: 5'-tctgaa ctatggttcacactatgagctac-3', FLCqR: 5'-gctctagtcacggagagggcagtct-3', DDM1ex16F: 5'-tccc tgaacagttaggacacatt-3', DDM1RT: 5'-aaaaacttccgagccctctaata-3', IBM1rtF2: 5'-ccttggaccttc aatcagaagcttgg-3', IBM1rtR2: 5'-tcgctgacattttcaggtgagacgaa-3', HOGex1F2: 5'-ccagatgcag caaaatggaactc-3', HOGex2R2: 5'-ttgttcgtgtcctaacacacatt-3', BME3rtF: 5'-tctcttccaacaccaac

tctgat-3', BME3rtR: 5'-ctcctgaattggctatgagaaga-3', TRINITYqF: 5'-ccctccagacccaccagactatc-3', TRINITYqR: 5'-gctccaaacaagcgagaggaaat-3', EIF4A-F: 5'-atccaagttggtgtgttctcc-3' and EIF4A-R: 5'-gagtgtctcgagcttccactc-3'. Real time PCR analysis was performed with the 2x SensiMix Plus SYBR & Fluorescein Kit (Quantace) protocol using an iQ5 Real-Time-PCR System (BioRad Laboratories). Ct values were analyzed using Excel (Microsoft).

5.5. *In situ* GUS detection

GUS activity was detected by staining in 0.1 M sodium phosphate buffer pH7.0, 10 mM EDTA, 0.1% Triton X-100, 100 µg/ml chloramphenicol, 2 mM potassium ferrocyanide, 2 mM potassium ferricyanide and 0.5 mg/ml X-glucuronide after 30 min vacuum infiltration and overnight incubation at 37°C. Subsequent washes with 70% ethanol at 37°C were performed in order to remove residual chlorophyll. All samples were analyzed using a Leica MZ16FA binocular microscope with a Leica DFC300FX CCD camera. Images were acquired with Leica Application Suite and processed with Adobe Photoshop (Adobe).

5.6. Fluorescence *in situ* hybridization (FISH) and immuno-labeling detection

For the preparation of nuclei, 21 day old plantlets were rinsed in 10 mM Tris buffer pH7.5, fixed by vacuum infiltration in 4% formaldehyde/Tris buffer, rinsed in Tris buffer, chopped in 500 µl Chromosome isolation (CI) buffer (15 mM Tris, 2 mM Na2EDTA, 0.5 mM spermin, 80 mM KCl, 20 mM NaCl, 15 mM beta-mercaptoethanol, 0.1% Triton X-100, pH7.5), and filtered through a 50 µm nylon mesh. Fifty µl nuclei suspension was transferred onto microscope slides using Cytospin (2500 rpm for 10 min). After centrifugation, slides were shortly rinsed in 1x PBS, transferred into 50% glycerol and stored at -20°C until use.

Immuno-localization of methylated cytosine was performed as described (Jasencakova *et al.* 2000) with minor modifications. In brief, slides were treated with pepsin (50 µg/ml in 0.01 M HCl; Roche) at 38°C (1-2 min), post-fixed in 4% formaldehyde/2x SSC, denatured in 70% formamide/2x SSC at 80°C (2 min) and cooled in ice-cold 1x PBS. After blocking (5% BSA, 0.2% Tween 20, 4x SSC) at 37°C (30 min), the slides were incubated with primary monoclonal mouse-anti-5-methylcytosine (1:500, Eurogentec) and secondary goat-anti-mouse-Alexa488 (1:250, Molecular Probes) antibodies.

Immuno-localization of histone H3 modifications was performed as previously described (Jasencankova *et al.* 2000). Slides were fixed in 4% paraformaldehyde/PBS for 20 min and blocked (5% BSA, 0.2% Tween 20, 4x SSC) at 37°C (30 min). The slides were incubated

over night at 4°C with primary antibodies specific to H3K9me2 (1:500, T. Jenuwein 4677) and H3K27me3 (1:500, T. Jenuwein 1199) and secondary goat-anti-rabbit-AF488 (1:400, Molecular Probes).

A biotin-labeled *Arabidopsis* centromeric repeat (pAL, 180 bp) probe for FISH was prepared from genomic DNA by PCR using primers pALU 5'-agtctttggctttgtgtctt-3' and pALR 5'-tggactttggctacaccatg-3'. Slide pretreatment and detection steps were performed as described (Pecinka *et al.* 2004). The probe was detected with subsequent avidin-Texas Red (1:1000, Vector Laboratories), goat-anti-avidin-biotin (1:200, Vector Laboratories) and again avidin-Texas Red (1:1000). The slides were counterstained with DAPI (1 µg/ml in Vectashield (Vector Laboratories)) and analyzed using a Zeiss Axioplan 2 epifluorescence microscope. Monochromatic images were acquired with MetaVue (Universal Imaging) and processed with Adobe Photoshop (Adobe).

5.7. Bisulfite conversion, sequencing and evaluation

After treatment with RNaseA and proteinase K, 1-2 µg of genomic DNA was digested overnight with *Bam*HI (MBI Fermentas). Subsequent bisulfite conversion was carried out using the Epitect Conversion Kit (Qiagen) and controlled for completion as described (Hetzl et al. 2007). Converted DNA was used for PCR amplification with the following primer pairs: pCaMV35S BSA-F: 5'-aattgagatttttttaataaagggtaatat-3', pCaMV35S BSA-X1: 5'-atcccc caaaatccccaaata-3' and pCaMV35S BSA-X2: 5'-ataaaaacccaccacctctac-3'. PCR-amplified DNA was cloned using CloneJet or InsTAclone kits (MBI Fermentas) and transformed into DH5α cells (Invitrogene), sequenced by terminal-labeling using BigDye Terminator v3.1 (Applied Biosystems) and read at vbcgenomics.com. The sequence information obtained was analyzed with CyMATE (www.gmi.oeaw.ac.at/cymate, (Hetzl *et al.* 2007) and Excel (Microsoft).

5.8. Chromatin immuno-precipitation

ChIP was performed as described (www.epigenome-noe.net/researchtools/protocol.php) using 3 week old seedlings. The chromatin was immunoprecipitated with antibodies to histone H3 (Abcam, ab1791), H3K4me3 (Upstate, 07-473), H4K5K8K11K14K17ac (Milipore, P62805), H3K9me2 (T.Jenuwein 4677), H3K27me2 (Upstate, 07-473) and H3K27me3 (T.Jenuwein 1199). Immunoprecipitated DNA was purified using the Quiagen PCR Purification Kit and eluted in 50µl buffer EB. Quantitative real-time PCRs were carried out in

total reaction volume of 25µl and qPCR conditions according to the 2x SensiMix Plus SYBR & Fluorescein Kit (Quantace) protocol using an iQ5 Real-Time-PCR System (BioRad Laboratories). Primer sets used were: HPT-qF: 5'-gggtaaatagctgcgccg atggtt-3', HPT-qR: 5'-cacggcgggagatgcaataggtc-3', P35S-F: 5'-gtgatatctccactgacgtaaggg-3', HUGO: 5'-atatctcattgtcccccgggga-3', Cseq2: 5'-actgacctacagggcagccac-3', CFrevR: 5'-gtc tattcttttactttgcaa-3', POLITAN: 5'-ataatgggaaggtgaaatggca-3', FRankF: 5'-cgacaaacact gattcatcatct-3', FRankR: 5'-agtcaccaccgcacacattgt-3', LINE1-4F: 5'-cccatggtgaccaagagttt-3' and LINE1-4R: 5'-tcaatgtcggagacctcctc-3'. qPCR data were analyzed according to the % of input method described by (Haring *et al.* 2007).

5.9. Microarray expression analysis

Seeds were sawn in a regular matrix on petri dishes containing semisolid germination medium with 0.8% agar and grown for 17 days in growth chambers at 16h light / 8h dark cycles and 21°C. Total RNA was extracted from pooled seedlings using Qiagen RNeasy Plant Mini Kit following standard extraction protocol plus additional RNA cleanup if needed. Three independently prepared RNA samples for each line were sent to the Nottingham Arabidopsis Stock Centre (NASC) for cDNA synthesisand hybridisation to Affymertix ATH1 arrays. The obtained raw data was normalised using Robust Multichip Average (RMA, Irizarry *et al.* 2003) and log2 fold changes were calculated using linear modelling (LIMMA, Wettenhall and Smyth 2004) on the Bioconductor (Reimers and Carey 2006) platform in R. Genes showing significant changes in gene expression (adj.p-val > 0.05, LogFC > or < than 2) were used for further analysis.

5.10. Online databases and tools

Several available online tools and databases were used during this thesis:

DATABASES:
Arabidopsis ensembl: http://atensembl.arabidopsis.info
TAIR: http://www.arabidopsis.org
Arabidopsis MPSS plus (Lu *et al.* 2005): http://mpss.udel.edu/at/
Arabidopsis DNA methylation (Cokus *et al.* 2008): epigenomics.mcdb.ucla.edu/DNAmeth/
AnnoJ (Lister *et al.* 2008): http://neomorph.salk.edu/epigenome.html
Genevestigator (Zimmerman *et al.* 2004): https://www.genevestigator.ethz.ch/gv/index.jsp

TOOLS:

Repeatmasker: http://www.bioinfo.org.cn/relative/RepeatMasker%20at%20EMBL.htm

PipMaker and MultiPipMaker (Schwartz *et al.* 2000): http://pipmaker.bx.psu.edu/pipmaker/

Cymate (Hetzl *et al.* 2007): http://www.cymate.org/

6. ACKNOWLEDGEMENTS

I would like to thank Dr. Ortrun Mittelsten Scheid for her supervision, patience and great support over the past years. Although there were numerous up- and downs connected to this project we managed to make the best out.

I would also like to acknowledge my PhD committee members for their scientific input and encouraging statements: Dr. Karel Riha (GMI, Vienna), Dr. Werner Aufsatz (GMI, Vienna) and Dr. Anton Wutz (IMP, Vienna).

Furthermore, I owe all my colleagues a big “Thank You” for the support, interaction and collaborations during the past years, but also for the involvement in obtaining the results presented in this thesis. Especially Bonnie Wohlrab for her enthusiastic technical assistance, Michael Milos for generating a huge library of reverse genetic mutants, Andrea Förster for the characterization of chromatin features and sequence analysis of the epialleles used in this screen, Huy Dinh and Dr. Marc Rehmsmeier for their great bioinformatics and scientific input in the transcriptome analysis and Dr. Ales Pecinka for helping me with Arabidopsis cytology and for good scientific interaction.

I would also like to thank all collaborators and group-external persons for their help and scientific input: Wilfried Rozhon, Werner Aufsatz, Arndt von Haeseler, Viktor E. Marquez, Jon Reinders, Ingo Hoffman, Kathrin Kiok, James Matthew Watson, Marie Theres Hauser, Lucia Daxinger, Juan Antonio Torres Acosta, Vera Shoft, Marc Berlinger and the people at NASC and VBC genomics.

Finally, I want to thank my family: Heike and Moritz, and my parents for their love, support and never-ending patience.

This work was financially supported by: FWF, OEAW, EPIGENOME and GEN-AU

7. CURRICULUM VITAE

Name:	Tuncay Baubec, Mag. rer.nat.
Date of Birth:	December 26th 1979
Place of Birth:	Alba-Iulia, Romania
Nationality:	Austrian
Children:	Moritz S. Baubec (January 20th 2007)

Academic education:

11.2004 – 12.2008	PhD-student with Prof. Ortrun Mittelsten Scheid, (Vienna Biocenter International PhD programme). Gregor Mendel Institute (GMI) of Molecular Plant Biology Austrian Academy of Sciences (OeAW), Vienna. Title of thesis: *Maintenance of polyploidy-associated transcriptional gene silencing in Arabidopsis thaliana.*
29.10.2004	*Magister rerum naturalis* Master's Degree of Science (with honours), University of Vienna
08.2003 – 10.2004	Diploma student with Prof. Denise P. Barlow. Center of Molecular Medicine (CeMM) Austrian Academy of Sciences (OeAW), Vienna. Title of thesis: *Replication asynchrony and chromatin compaction around the imprinted mouse Igf2r region.*
10.2000 – 10.2004	Undergraduate studies at the University of Vienna. Biology / Genetics and Microbiology
09.1999 – 09.2000	Civil Service, Red Cross Krems a.d. Donau
19.07.1999	Matura exam
09.1991 – 07.1999	Bundesgymnasium, BG Rechte Kremszeille, Krems a.d. Donau

Presentations at meetings and conferences:

Zebularine is a Potent Inhibitor of Plant DNA Methylation
Vienna Plant Meeting, Vienna, October 2008 (Oral presentation)

Zebularine is a Potent Inhibitor of Plant DNA Methylation
Keystone Symposia: Molecular Basis for Chromatin Modifications and Epigenetic Phenomena, Snowmass-CO, April 2008 (Oral and poster presentation)

Trans-acting factors of polyploidy associated gene silencing – Forward and Reverse genetic approaches.
EMBO Conference on Chromatin and Epigenetics. Heidelberg, Mai 2007. (Poster)

Trans-acting factors of polyploidy associated gene silencing – Forward and Reverse genetic approaches.
Vienna Plant Meeting, Vienna, April 2007 (Oral presentation)

Publications:

Baubec T, Pecinka A, Rozhon W and Mittelsten Scheid O
Effective, homogenous and transient interference with cytosine methylation in plant genomic DNA by zebularine treatment.
The Plant Journal, (2008) [in press]

Rozhon W, Baubec T, Mayerhofer J, Mittelsten Scheid O, Jonak C
Rapid quantification of global DNA methylation by isocratic cation exchange high-performance liquid chromatography.
Analytical Biochemistry, 375 (2008), 354-360

Baubec T and Mittelsten Scheid O
Medea in full self-control.
Trends in Plant Sciences, 11 (2006) 469-471.

Baubec T
Analysis of replication and chromatin compaction at the imprinted mouse Igf2r region.
Diploma thesis, University of Vienna, 2004

8. LIST OF ABBREVIATIONS

5-mdC	5-methyldeoxycytosine
bp	basepair
C2R	diploid hygromycin resistant line
C2S1	diploid hygromycin sensitive line
C4R	tetraploid hygromycin resistant line
C4S1	tetraploid hygromycin sensitive line
CaMV	Cauliflower Mosaic Virus
cDNA	complementary DNA
(d)C	(deoxy)cytosine
DDM1	DECREASE IN DNA METHYLATION1
DNA	deoxyribonucleic acid
DNMT	DNA methyltransferase
FISH	fluorescence in situ hybridisation
(d)G	(deoxy)guanosine
GO	gene ontology
GUS	beta-glucoronidase
H	every nucleotide except G
HMT	histone methyltransferase
HOG1	HOMOLOGY DEPENDENT GENE SILENCING1
HPLC	high pressure liquid chromatography
HPT	hygromycin phosphotransferase
kb	kilobase
nt	nucleotide
SAH	S-adenosyl-L-homocysteine
SAM	S-adneosyl-L-methionine
TN	transposon
TSA	trichostatine A
TSI	transcriptionally silent information
wt	wild type
wtZH	Zürich accession
zeb	zebularine

9. REFERENCES

Adams, K.L. (2007) Evolution of duplicate gene expression in polyploid and hybrid plants. *J Hered*, **98**, 136-141.

AGI (2000) Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature*, **408**, 796-815.

Alleman, M., Sidorenko, L., McGinnis, K., Seshadri, V., Dorweiler, J.E., White, J., Sikkink, K. and Chandler, V.L. (2006) An RNA-dependent RNA polymerase is required for paramutation in maize. *Nature*, **442**, 295-298.

Allen, E., Xie, Z., Gustafson, A.M. and Carrington, J.C. (2005) microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell*, **121**, 207-221.

Allis, C.D., Jenuwein, T. and Reinberg, D. (2006) Epigenetics. *Cold Spring Harb Laboratory Press, Cold Spring Harbor, New York.*

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

Alvarez-Venegas, R. and Avramova, Z. (2001) Two Arabidopsis homologs of the animal trithorax genes: a new structural domain is a signature feature of the trithorax gene family. *Gene*, **271**, 215-221.

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

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

Bancroft, I. (2001) Duplicate and diverge: the evolution of plant genome microstructure. *Trends Genet*, **17**, 89-93.

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.

Bartel, D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, **116**, 281-297.

Baubec, T., Pecinka, A., Rozhon, W. and Mittelsten Scheid, O. (2008) Effective, homogeneous and transient interference with cytosine methylation in plant genomic DNA by zebularine. *Plant J.*

Baulcombe, D. (2004) RNA silencing in plants. *Nature*, **431**, 356-363.

Baumberger, N. and Baulcombe, D.C. (2005) Arabidopsis ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs. *Proc Natl Acad Sci U S A*, **102**, 11928-11933.

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.

Becker, P.B. and Horz, W. (2002) ATP-dependent nucleosome remodeling. *Annu Rev Biochem*, **71**, 247-273.

Bernatavichute, Y.V., Zhang, X., Cokus, S., Pellegrini, M. and Jacobsen, S.E. (2008) Genome-wide association of histone H3 lysine nine methylation with CHG DNA methylation in Arabidopsis thaliana. *PLoS ONE*, **3**, e3156.

Bienvenu, T. and Chelly, J. (2006) Molecular genetics of Rett syndrome: when DNA methylation goes unrecognized. *Nat Rev Genet*, **7**, 415-426.

Bird, A. (2002) DNA methylation patterns and epigenetic memory. *Genes Dev*, **16**, 6-21.

Bird, A., Taggart, M. and Macleod, D. (1981) Loss of rDNA methylation accompanies the onset of ribosomal gene activity in early development of X. laevis. *Cell*, **26**, 381-390.

Bork, P. and Koonin, E.V. (1993) An expanding family of helicases within the 'DEAD/H' superfamily. *Nucleic Acids Res*, **21**, 751-752.

Brink, R.A. (1973) Paramutation. *Annu Rev Genet*, **7**, 129-152.

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.

Carmell, M.A., Xuan, Z., Zhang, M.Q. and Hannon, G.J. (2002) The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. *Genes Dev*, **16**, 2733-2742.

Chakravarthy, S., Park, Y.J., Chodaparambil, J., Edayathumangalam, R.S. and Luger, K. (2005) Structure and dynamic properties of nucleosome core particles. *FEBS Lett*, **579**, 895-898.

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

Chan, S.W., Zhang, X., Bernatavichute, Y.V. and Jacobsen, S.E. (2006) Two-step recruitment of RNA-directed DNA methylation to tandem repeats. *PLoS Biol*, **4**, e363.

Chang, S. and Pikaard, C.S. (2005) Transcript profiling in Arabidopsis reveals complex responses to global inhibition of DNA methylation and histone deacetylation. *J Biol Chem*, **280**, 796-804.

Chen, Z.J., Comai, L. and Pikaard, C.S. (1998) Gene dosage and stochastic effects determine the severity and direction of uniparental ribosomal RNA gene silencing (nucleolar dominance) in Arabidopsis allopolyploids. *Proc Natl Acad Sci U S A*, **95**, 14891-14896.

Chen, Z.J. and Ni, Z. (2006) Mechanisms of genomic rearrangements and gene expression changes in plant polyploids. *Bioessays*, **28**, 240-252.

Choi, Y., Gehring, M., Johnson, L., Hannon, M., Harada, J.J., Goldberg, R.B., Jacobsen, S.E. and Fischer, R.L. (2002) DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in arabidopsis. *Cell*, **110**, 33-42.

Cokus, S.J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C.D., Pradhan, S., Nelson, S.F., Pellegrini, M. and Jacobsen, S.E. (2008) Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature*, **452**, 215-219.

Doerfler, W. (2006) The almost-forgotten fifth nucleotide in DNA: an introduction. *Curr Top Microbiol Immunol*, **301**, 3-18.

Du, T. and Zamore, P.D. (2005) microPrimer: the biogenesis and function of microRNA. *Development*, **132**, 4645-4652.

Elmayan, T., Proux, F. and Vaucheret, H. (2005) Arabidopsis RPA2: a genetic link among transcriptional gene silencing, DNA repair, and DNA replication. *Curr Biol*, **15**, 1919-1925.

Fedoroff, N. and Botstein, D. (1992) The Dynamic Genome: Barbara McClintock's Ideas in the Century of Genetics. *CSHL Press*.

Finnegan, E.J., Peacock, W.J. and Dennis, E.S. (1996) Reduced DNA methylation in Arabidopsis thaliana results in abnormal plant development. *Proc Natl Acad Sci U S A*, **93**, 8449-8454.

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

Furner, I.J., Sheikh, M.A. and Collett, C.E. (1998) Gene silencing and homology-dependent gene silencing in Arabidopsis: genetic modifiers and DNA methylation. *Genetics*, **149**, 651-662.

Gasciolli, V., Mallory, A.C., Bartel, D.P. and Vaucheret, H. (2005) Partially redundant functions of Arabidopsis DICER-like enzymes and a role for DCL4 in producing trans-acting siRNAs. *Curr Biol*, **15**, 1494-1500.

Gaszner, M. and Felsenfeld, G. (2006) Insulators: exploiting transcriptional and epigenetic mechanisms. *Nat Rev Genet*, **7**, 703-713.

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.

Gong, Z., Morales-Ruiz, T., Ariza, R.R., Roldan-Arjona, T., David, L. and Zhu, J.K. (2002) ROS1, a repressor of transcriptional gene silencing in Arabidopsis, encodes a DNA glycosylase/lyase. *Cell*, **111**, 803-814.

Grewal, S.I. and Moazed, D. (2003) Heterochromatin and epigenetic control of gene expression. *Science*, **301**, 798-802.

Hale, C.J., Stonaker, J.L., Gross, S.M. and Hollick, J.B. (2007) A novel Snf2 protein maintains trans-generational regulatory states established by paramutation in maize. *PLoS Biol*, **5**, e275.

Hamilton, A.J. and Baulcombe, D.C. (1999) A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science*, **286**, 950-952.

Han, M.H., Goud, S., Song, L. and Fedoroff, N. (2004) The Arabidopsis double-stranded RNA-binding protein HYL1 plays a role in microRNA-mediated gene regulation. *Proc Natl Acad Sci U S A*, **101**, 1093-1098.

Haring, M., Offermann, S., Danker, T., Horst, I., Peterhansel, C. and Stam, M. (2007) Chromatin immunoprecipitation: optimization, quantitative analysis and data normalization. *Plant Methods*, **3**, 11.

Heitz (1928) Das Heterochromatin der Moose. *Jb. Wiss. Bot.*, **69**, 728–818.

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

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

Hetzl, J., Foerster, A.M., Raidl, G. and Mittelsten Scheid, O. (2007) CyMATE: a new tool for methylation analysis of plant genomic DNA after bisulphite sequencing. *Plant J*, **51**, 526-536.

Hirochika, H., Okamoto, H. and Kakutani, T. (2000) Silencing of retrotransposons in arabidopsis and reactivation by the ddm1 mutation. *Plant Cell*, **12**, 357-369.

Huettel, B., Kanno, T., Daxinger, L., Aufsatz, W., Matzke, A.J. and Matzke, M. (2006) Endogenous targets of RNA-directed DNA methylation and Pol IV in Arabidopsis. *Embo J*, **25**, 2828-2836.

Inamdar, N.M., Ehrlich, K.C. and Ehrlich, M. (1991) CpG methylation inhibits binding of several sequence-specific DNA-binding proteins from pea, wheat, soybean and cauliflower. *Plant Mol Biol*, **17**, 111-123.

Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U. and Speed, T.P. (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics*, **4**, 249-264.

Jacobsen, S.E., Sakai, H., Finnegan, E.J., Cao, X. and Meyerowitz, E.M. (2000) Ectopic hypermethylation of flower-specific genes in Arabidopsis. *Curr Biol*, **10**, 179-186.

Jasencakova, Z., Meister, A., Walter, J., Turner, B.M. and Schubert, I. (2000) Histone H4 acetylation of euchromatin and heterochromatin is cell cycle dependent and correlated with replication rather than with transcription. *Plant Cell*, **12**, 2087-2100.

Jeddeloh, J.A., Bender, J. and Richards, E.J. (1998) The DNA methylation locus DDM1 is required for maintenance of gene silencing in Arabidopsis. *Genes Dev*, **12**, 1714-1725.

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., Bostick, M., Zhang, X., Kraft, E., Henderson, I., Callis, J. and Jacobsen, S.E. (2007) The SRA methyl-cytosine-binding domain links DNA and histone methylation. *Curr Biol*, **17**, 379-384.

Jordan, N.D., West, J.P., Bottley, A., Sheikh, M. and Furner, I. (2007) Transcript profiling of the hypomethylated hog1 mutant of Arabidopsis. *Plant Mol Biol*, **65**, 571-586.

Jorgensen, R. (2003) What is the plant cell? *Plant Cell*, **15**, 1497-1498.

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

Kanno, T., Aufsatz, W., Jaligot, E., Mette, M.F., Matzke, M. and Matzke, A.J. (2005a) A SNF2-like protein facilitates dynamic control of DNA methylation. *EMBO Rep*, **6**, 649-655.

Kanno, T., Huettel, B., Mette, M.F., Aufsatz, W., Jaligot, E., Daxinger, L., Kreil, D.P., Matzke, M. and Matzke, A.J. (2005b) Atypical RNA polymerase subunits required for RNA-directed DNA methylation. *Nat Genet*, **37**, 761-765.

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.

Kaya, H., Shibahara, K.I., Taoka, K.I., Iwabuchi, M., Stillman, B. and Araki, T. (2001) FASCIATA genes for chromatin assembly factor-1 in arabidopsis maintain the cellular organization of apical meristems. *Cell*, **104**, 131-142.

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.

Kohler, C. and Villar, C.B. (2008) Programming of gene expression by Polycomb group proteins. *Trends Cell Biol*, **18**, 236-243.

Koncz, C., Martini, N., Mayerhofer, R., Koncz-Kalman, Z., Korber, H., Redei, G.P. and Schell, J. (1989) High-frequency T-DNA-mediated gene tagging in plants. *Proc Natl Acad Sci U S A*, **86**, 8467-8471.

Kornberg, R.D. and Lorch, Y. (1999) Chromatin-modifying and -remodeling complexes. *Curr Opin Genet Dev*, **9**, 148-151.

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

Kuo, H.K., Griffith, J.D. and Kreuzer, K.N. (2007) 5-Azacytidine induced methyltransferase-DNA adducts block DNA replication in vivo. *Cancer Res*, **67**, 8248-8254.

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.

Lam, E., Benfey, P.N., Gilmartin, P.M., Fang, R.X. and Chua, N.H. (1989) Site-specific mutations alter in vitro factor binding and change promoter expression pattern in transgenic plants. *Proc Natl Acad Sci U S A*, **86**, 7890-7894.

Lawrence, R.J., Earley, K., Pontes, O., Silva, M., Chen, Z.J., Neves, N., Viegas, W. and Pikaard, C.S. (2004) A concerted DNA methylation/histone methylation switch regulates rRNA gene dosage control and nucleolar dominance. *Mol Cell*, **13**, 599-609.

Li, J., Yang, Z., Yu, B., Liu, J. and Chen, X. (2005) Methylation protects miRNAs and siRNAs from a 3'-end uridylation activity in Arabidopsis. *Curr Biol*, **15**, 1501-1507.

Lindroth, A.M., Cao, X., Jackson, J.P., Zilberman, D., McCallum, C.M., Henikoff, S. and Jacobsen, S.E. (2001) Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation. *Science*, **292**, 2077-2080.

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

Linn, F., Heidmann, I., Saedler, H. and Meyer, P. (1990) Epigenetic changes in the expression of the maize A1 gene in Petunia hybrida: role of numbers of integrated gene copies and state of methylation. *Mol Gen Genet*, **222**, 329-336.

Lippman, Z., Gendrel, A.V., Black, M., Vaughn, M.W., Dedhia, N., McCombie, W.R., Lavine, K., Mittal, V., May, B., Kasschau, K.D., Carrington, J.C., Doerge, R.W., Colot, V. and Martienssen, R. (2004) Role of transposable elements in heterochromatin and epigenetic control. *Nature*, **430**, 471-476.

Lister, R., O'Malley, R.C., Tonti-Filippini, J., Gregory, B.D., Berry, C.C., Millar, A.H. and Ecker, J.R. (2008) Highly integrated single-base resolution maps of the epigenome in Arabidopsis. *Cell*, **133**, 523-536.

Liu, Y.G., Mitsukawa, N., Oosumi, T. and Whittier, R.F. (1995) Efficient isolation and mapping of Arabidopsis thaliana T-DNA insert junctions by thermal asymmetric interlaced PCR. *Plant J*, **8**, 457-463.

Llave, C., Kasschau, K.D., Rector, M.A. and Carrington, J.C. (2002) Endogenous and silencing-associated small RNAs in plants. *Plant Cell*, **14**, 1605-1619.

Lu, C., Tej, S.S., Luo, S., Haudenschild, C.D., Meyers, B.C. and Green, P.J. (2005) Elucidation of the small RNA component of the transcriptome. *Science*, **309**, 1567-1569.

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

Macleod, D., Charlton, J., Mullins, J. and Bird, A.P. (1994) Sp1 sites in the mouse aprt gene promoter are required to prevent methylation of the CpG island. *Genes Dev*, **8**, 2282-2292.

Mathieu, O., Jasencakova, Z., Vaillant, I., Gendrel, A.V., Colot, V., Schubert, I. and Tourmente, S. (2003) Changes in 5S rDNA chromatin organization and transcription during heterochromatin establishment in Arabidopsis. *Plant Cell*, **15**, 2929-2939.

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.

Matzke, M., Aufsatz, W., Kanno, T., Daxinger, L., Papp, I., Mette, M.F. and Matzke, A.J. (2004) Genetic analysis of RNA-mediated transcriptional gene silencing. *Biochim Biophys Acta*, **1677**, 129-141.

Matzke, M., Kanno, T., Huettel, B., Daxinger, L. and Matzke, A.J. (2007) Targets of RNA-directed DNA methylation. *Curr Opin Plant Biol*, **10**, 512-519.

Milos, M. (2006) A reverse genetic approach to analyse ploidy-associated gene silencing in Arabidopsis thaliana. *University of Vienna*, *Master thesis*.

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.

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

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

Moffatt, B. and Weretilnyk, E.A. (2001) Sustaining S-adenosyl-L-methyonine-dependent methyltransferase activity in plant cells. *Physiol. Plant.*, 435-442.

Morel, J.B., Mourrain, P., Beclin, C. and Vaucheret, H. (2000) DNA methylation and chromatin structure affect transcriptional and post-transcriptional transgene silencing in Arabidopsis. *Curr Biol*, **10**, 1591-1594.

Murfett, J., Wang, X.J., Hagen, G. and Guilfoyle, T.J. (2001) Identification of Arabidopsis histone deacetylase HDA6 mutants that affect transgene expression. *Plant Cell*, **13**, 1047-1061.

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

Ono, T., Kaya, H., Takeda, S., Abe, M., Ogawa, Y., Kato, M., Kakutani, T., Mittelsten Scheid, O., Araki, T. and Shibahara, K. (2006) Chromatin assembly factor 1 ensures the stable maintenance of silent chromatin states in Arabidopsis. *Genes Cells*, **11**, 153-162.

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

Osborn, T.C. (2003) Understanding mechanisms of novel gene expression in polyploids. *Trends Genet*, **19**, 141-147.

Osborn, T.C., Pires, J.C., Birchler, J.A., Auger, D.L., Chen, Z.J., Lee, H.S., Comai, L., Madlung, A., Doerge, R.W., Colot, V. and Martienssen, R.A. (2003) Understanding mechanisms of novel gene expression in polyploids. *Trends Genet*, **19**, 141-147.

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.

Pecinka, A., Schubert, V., Meister, A., Kreth, G., Klatte, M., Lysak, M.A., Fuchs, J. and Schubert, I. (2004) Chromosome territory arrangement and homologous pairing in nuclei of Arabidopsis thaliana are predominantly random except for NOR-bearing chromosomes. *Chromosoma*, **113**, 258-269.

Penterman, J., Uzawa, R. and Fischer, R.L. (2007a) Genetic interactions between DNA demethylation and methylation in Arabidopsis. *Plant Physiol*, **145**, 1549-1557.

Penterman, J., Zilberman, D., Huh, J.H., Ballinger, T., Henikoff, S. and Fischer, R.L. (2007b) DNA demethylation in the Arabidopsis genome. *Proc Natl Acad Sci U S A*, **104**, 6752-6757.

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

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

Pikaard, C.S. (2000a) The epigenetics of nucleolar dominance. *Trends Genet*, **16**, 495-500.

Pikaard, C.S. (2000b) Nucleolar dominance: uniparental gene silencing on a multi-megabase scale in genetic hybrids. *Plant Mol Biol*, **43**, 163-177.

Pillai, R.S., Bhattacharyya, S.N. and Filipowicz, W. (2007) Repression of protein synthesis by miRNAs: how many mechanisms? *Trends Cell Biol*, **17**, 118-126.

Pontes, O., Lawrence, R.J., Silva, M., Preuss, S., Costa-Nunes, P., Earley, K., Neves, N., Viegas, W. and Pikaard, C.S. (2007) Postembryonic establishment of megabase-scale gene silencing in nucleolar dominance. *PLoS ONE*, **2**, e1157.

Pontes, O., Neves, N., Silva, M., Lewis, M.S., Madlung, A., Comai, L., Viegas, W. and Pikaard, C.S. (2004) Chromosomal locus rearrangements are a rapid response to formation of the allotetraploid Arabidopsis suecica genome. *Proc Natl Acad Sci U S A*, **101**, 18240-18245.

Probst, A.V., Fransz, P.F., Paszkowski, J. and Mittelsten Scheid, O. (2003) Two means of transcriptional reactivation within heterochromatin. *Plant J*, **33**, 743-749.

Ramsey, J. and Schemske, D.W. (1998) Pathways, mechanisms and rates of polyploid formation in flowering plants. *Annual Review of Ecology and Systematics*, **29**, 477-501.

Rangan, L., Vogel, C. and Srivastava, A. (2008) Analysis of context sequence surrounding translation initiation site from complete genome of model plants. *Mol Biotechnol*, **39**, 207-213.

Reimers, M. and Carey, V.J. (2006) Bioconductor: an open source framework for bioinformatics and computational biology. *Methods Enzymol*, **411**, 119-134.

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.

Ronemus, M.J., Galbiati, M., Ticknor, C., Chen, J. and Dellaporta, S.L. (1996) Demethylation-induced developmental pleiotropy in Arabidopsis. *Science*, **273**, 654-657.

Rozhon, W., Baubec, T., Mayerhofer, J., Mittelsten Scheid, O. and Jonak, C. (2008) Rapid quantification of global DNA methylation by isocratic cation exchange high-performance liquid chromatography. *Anal Biochem*, **375**, 354-360.

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.

Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B. and Cavalli, G. (2007) Genome regulation by polycomb and trithorax proteins. *Cell*, **128**, 735-745.

Schwartz, S., Zhang, Z., Frazer, K.A., Smit, A., Riemer, C., Bouck, J., Gibbs, R., Hardison, R. and Miller, W. (2000) PipMaker--a web server for aligning two genomic DNA sequences. *Genome Res*, **10**, 577-586.

Smith, L.M., Pontes, O., Searle, I., Yelina, N., Yousafzai, F.K., Herr, A.J., Pikaard, C.S. and Baulcombe, D.C. (2007) An SNF2 protein associated with nuclear RNA silencing and the spread of a silencing signal between cells in Arabidopsis. *Plant Cell*, **19**, 1507-1521.

Soltis, P.S. and Soltis, D.E. (2000) The role of genetic and genomic attributes in the success of polyploids. *Proc Natl Acad Sci U S A*, **97**, 7051-7057.

Song, K., Lu, P., Tang, K. and Osborn, T.C. (1995) Rapid genome change in synthetic polyploids of Brassica and its implications for polyploid evolution. *Proc Natl Acad Sci U S A*, **92**, 7719-7723.

Soppe, W.J., Jasencakova, Z., Houben, A., Kakutani, T., Meister, A., Huang, M.S., Jacobsen, S.E., Schubert, I. and Fransz, P.F. (2002) DNA methylation controls histone H3 lysine 9 methylation and heterochromatin assembly in Arabidopsis. *Embo J*, **21**, 6549-6559.

Spencer, V.A. and Davie, J.R. (1999) Role of covalent modifications of histones in regulating gene expression. *Gene*, **240**, 1-12.

Stam, M. and Mittelsten Scheid, O. (2005) Paramutation: an encounter leaving a lasting impression. *Trends Plant Sci*, **10**, 283-290.

Stebbins, G.L. (1966) Chromosomal Variation and Evolution. *Science*, **152**, 1463-1469.

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., Ohba, R., Cook, R.G. and Allis, C.D. (1999) Methylation of histone H3 at lysine 4 is highly conserved and correlates with transcriptionally active nuclei in Tetrahymena. *Proc Natl Acad Sci U S A*, **96**, 14967-14972.

Takeda, S., Tadele, Z., Hofmann, I., Probst, A.V., Angelis, K.J., Kaya, H., Araki, T., Mengiste, T., Mittelsten Scheid, O., Shibahara, K., Scheel, D. and Paszkowski, J. (2004) BRU1, a novel link between responses to DNA damage and epigenetic gene silencing in Arabidopsis. *Genes Dev*, **18**, 782-793.

Talbert, P.B. and Henikoff, S. (2006) Spreading of silent chromatin: inaction at a distance. *Nat Rev Genet*, **7**, 793-803.

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

Turner, B.M. (2000) Histone acetylation and an epigenetic code. *Bioessays*, **22**, 836-845.

Vaucheret, H., Vazquez, F., Crete, P. and Bartel, D.P. (2004) The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. *Genes Dev*, **18**, 1187-1197.

Vazquez, F., Gasciolli, V., Crete, P. and Vaucheret, H. (2004) The nuclear dsRNA binding protein HYL1 is required for microRNA accumulation and plant development, but not posttranscriptional transgene silencing. *Curr Biol*, **14**, 346-351.

Voinnet, O. (2005) Induction and suppression of RNA silencing: insights from viral infections. *Nat Rev Genet*, **6**, 206-220.

Vongs, A., Kakutani, T., Martienssen, R.A. and Richards, E.J. (1993) Arabidopsis thaliana DNA methylation mutants. *Science*, **260**, 1926-1928.

Wagner, D. (2003) Chromatin regulation of plant development. *Curr Opin Plant Biol*, **6**, 20-28.

Ward, E.R., Uknes, S.J., Williams, S.C., Dincher, S.S., Wiederhold, D.L., Alexander, D.C., Ahl-Goy, P., Metraux, J.P. and Ryals, J.A. (1991) Coordinate Gene Activity in Response to Agents That Induce Systemic Acquired Resistance. *Plant Cell*, **3**, 1085-1094.

Wendel, J.F. (2000) Genome evolution in polyploids. *Plant Mol Biol*, **42**, 225-249.

Wettenhall, J.M. and Smyth, G.K. (2004) limmaGUI: a graphical user interface for linear modeling of microarray data. *Bioinformatics*, **20**, 3705-3706.

Wildermuth, M.C., Dewdney, J., Wu, G. and Ausubel, F.M. (2001) Isochorismate synthase is required to synthesize salicylic acid for plant defence. *Nature*, **414**, 562-565.

Xiao, W., Custard, K.D., Brown, R.C., Lemmon, B.E., Harada, J.J., Goldberg, R.B. and Fischer, R.L. (2006) DNA methylation is critical for Arabidopsis embryogenesis and seed viability. *Plant Cell*, **18**, 805-814.

Xie, Z., Johansen, L.K., Gustafson, A.M., Kasschau, K.D., Lellis, A.D., Zilberman, D., Jacobsen, S.E. and Carrington, J.C. (2004) Genetic and functional diversification of small RNA pathways in plants. *PLoS Biol*, **2**, E104.

Yamada, K., Lim, J., Dale, J.M., Chen, H., Shinn, P., Palm, C.J., Southwick, A.M., Wu, H.C., Kim, C., Nguyen, M., Pham, P., Cheuk, R., Karlin-Newmann, G., Liu, S.X., Lam, B., Sakano, H., Wu, T., Yu, G., Miranda, M., Quach, H.L., Tripp, M., Chang, C.H., Lee, J.M., Toriumi, M., Chan, M.M., Tang, C.C., Onodera, C.S., Deng, J.M., Akiyama, K., Ansari, Y., Arakawa, T., Banh, J., Banno, F., Bowser, L., Brooks, S., Carninci, P., Chao, Q., Choy, N., Enju, A., Goldsmith, A.D., Gurjal, M., Hansen, N.F., Hayashizaki, Y., Johnson-Hopson, C., Hsuan, V.W., Iida, K., Karnes, M., Khan, S., Koesema, E., Ishida, J., Jiang, P.X., Jones, T., Kawai, J., Kamiya, A., Meyers, C., Nakajima, M., Narusaka, M., Seki, M., Sakurai, T., Satou, M., Tamse, R., Vaysberg, M., Wallender, E.K., Wong, C., Yamamura, Y., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R. (2003) Empirical analysis of transcriptional activity in the Arabidopsis genome. *Science*, **302**, 842-846.

Yoo, B.C., Kragler, F., Varkonyi-Gasic, E., Haywood, V., Archer-Evans, S., Lee, Y.M., Lough, T.J. and Lucas, W.J. (2004) A systemic small RNA signaling system in plants. *Plant Cell*, **16**, 1979-2000.

Yoshida, M., Horinouchi, S. and Beppu, T. (1995) Trichostatin A and trapoxin: novel chemical probes for the role of histone acetylation in chromatin structure and function. *Bioessays*, **17**, 423-430.

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

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

Zhao, X.P., Si, Y., Hanson, R.E., Crane, C.F., Price, H.J., Stelly, D.M., Wendel, J.F. and Paterson, A.H. (1998) Dispersed repetitive DNA has spread to new genomes since polyploid formation in cotton. *Genome Res*, **8**, 479-492.

Zhou, L., Cheng, X., Connolly, B.A., Dickman, M.J., Hurd, P.J. and Hornby, D.P. (2002) Zebularine: a novel DNA methylation inhibitor that forms a covalent complex with DNA methyltransferases. *J Mol Biol*, **321**, 591-599.

Zhu, J., Kapoor, A., Sridhar, V.V., Agius, F. and Zhu, J.K. (2007) The DNA glycosylase/lyase ROS1 functions in pruning DNA methylation patterns in Arabidopsis. *Curr Biol*, **17**, 54-59.

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

Zilberman, D. and Henikoff, S. (2004) Silencing of transposons in plant genomes: kick them when they're down. *Genome Biol*, **5**, 249.

Zimmermann, P., Hirsch-Hoffmann, M., Hennig, L. and Gruissem, W. (2004) GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. *Plant Physiol*, **136**, 2621-2632.