# Regulation of Active DNA Demethylation and its Role in Fertility in *Arabidopsis thaliana*

by

Deborah Allison Pohlmann

B.S. Biology (2013)
University of Richmond
Richmond, VA

Submitted to the Department of Biology in Partial Fulfillment
of the Requirements for the Degree of

Doctor of Philosophy in Biology

at the

Massachusetts Institute of Technology
Cambridge, MA

February 2022

© 2022 Massachusetts Institute of Technology. All rights reserved.

Signature of Author ……………………………………………………………………………..
Department of Biology
January 6, 2022

Certified by ……….……………………………………………………………………………..
Mary Gehring
Associate Professor of Biology
Thesis Supervisor

Accepted by ……………………………………………………………………………………..
Amy Keating
Professor of Biology
Co-chair, Graduate Committee

Regulation of Active DNA Demethylation and its Role in Fertility in *Arabidopsis thaliana*

by

Deborah Allison Pohlmann

Submitted to the Department of Biology on January 6th, 2022
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy of Biology

ABSTRACT

The genome of *Arabidopsis thaliana* must maintain control over the balance between *de novo* methylation, maintenance of existing methylation, and active demethylation as it utilizes these processes to facilitate gene expression changes throughout the lifetime of the plant. Expression of the DNA demethylase *REPRESSOR OF SILENCING1* (*ROS1*) is dependent upon this balance as well as perpetuating it, yet many aspects of the regulation of *ROS1* remain unknown. In this work, I show that the downregulation of *ROS1* in mutants of the RNA-directed DNA methylation pathway occurs at the transcriptional level, and is dependent upon an 817-bp region in the proximal promoter region of *ROS1*. The deletion of this region using CRISPR-Cas9 technology resulted in increased expression of *ROS1* in both wildtype and methylation-deficient backgrounds, indicating that this region may be a methylation-sensitive silencer sequence. Additional deletions in the endogenous chromosome identified further regions that contain regulatory elements of *ROS1*. Additionally, I further investigated the results when the balance between methylation and active demethylation is disturbed, by characterizing a quadruple mutant of all four member of the *DEMETER* family of DNA glycosylases in somatic tissues: *dme;ros1;dml2;dml3* (*drdd*). This mutant displays an early flowering phenotype which was linked to downregulation of the floral repressor *FLOWERING LOCUS C,* concurrent with DRDD-dependent hypermethylation in the 5' flanking region. I also characterized a low-penetrance male fertility defect in *drdd* mutants, which I determined is caused by a delay in anther dehiscence that could be a result of altered reactive oxygen species accumulation. This work has led to an increase in our understanding of the mechanisms by which *ROS1* is regulated, and the mechanisms by which active demethylation affect transcription and development of the plant.

Thesis Supervisor: Mary Gehring
Title: Associate Professor of Biology

**ACKNOWLEDGMENTS**

It has been a long journey. A very long journey, in fact.

To those people who support me in my scientific endeavors, who backed me up when my choices were difficult, who helped me to focus and direct my curiosity into a scientific tool, thank you. Mary Gehring, you are a wonderful advisor and you have made my time at MIT enlightening. Thank you for your patience, your willingness to take risks, and your ability to remain calm under every circumstance.

To all of the people who supported me as I spent days in lab vortexing samples every two minutes for a half hour, or taking 4-minutes exposure pictures of endless flats of plants, or extracting DNA from thousands of plants for genotyping, thank you. Labmates former and present, Ben, Satyaki, Daniela, Katherine, Becky, Xiao-Yu, Junsik, Maja, Colette, Rob, Elizabeth, Alex, Lindsey, Aurore, and Eleanor, thank you.

To those who made those tasks worthwhile, by exciting me about basic research and science, by talking over what our results might mean, by providing me with a distracting conversation, a helping hand, or even just a recommendation for a podcast to listen to, thank you. To my longstanding committee members, Jackie and Peter, to Judith for agreeing to be my outside reviewer, to Biograd 2013 and a reprisal of my labmates, thank you.

To those who have shared living spaces with me and put up with my endless talk about getting a cat, and then putting up with my endless talk about the cat I got once I got her, thank you. To Sara, Lindsay, and Eva, the original 245 crew, and to Kasturi and Fiona. Thank you for the food, the talks, the late night gripes, and the endless support.

To those who have learned that they shouldn't ask when this thesis will be done, who put up with my trying to explain advanced scientific concepts to them, who once subscribed to *Nature* magazine because they just wanted to know more about science, despite being business and computer people, thank you. To my family, who have always supported me, thank you.

To those who don't fit into any of these categories, yet remind me that not everything in life needs to be neatly categorized, thank you. For shared meals, for discussions about science and not about science, for trips to the movies and to Broadway shows, and walks on the Esplanade, thank you. To Tej, to my cat Panda, to Curio Coffee and to all others whom I have not specifically mentioned, thank you. This thesis would not exist without you.

## TABLE OF CONTENTS


Thesis Abstract .......... 2

Acknowledgments .......... 3

Chapter One: Introduction .......... 6

Introduction to methylation .......... 7

Establishment and maintenance of DNA methylation .......... 7
RNA-directed DNA Methylation .......... 7
Maintenance of DNA methylation .......... 7

Removal of DNA methylation .......... 9
Discovery and characterization of ROS1 .......... 9
Functionality of the DME family .......... 10
Mechanism of demethylation .......... 11
Demethylation in metazoans .......... 11

Functions of DNA methylation .......... 12
Silencing TEs .......... 12
Imprinting .......... 13
Regulation of transcription .......... 14

Dynamics of DNA Methylation .......... 16
DNA methylation is a highly stable mark .......... 16
DNA methylation is a highly labile mark .......... 17
Balancing methylation and demethylation .......... 18

Summary of thesis .......... 19

References .......... 21

Chapter Two: A DNA demethylase in Arabidopsis is transcriptionally regulated in *cis* by a methylation-sensitive silencer sequence .......... 30

Abstract .......... 31

Introduction .......... 31

Results .......... 33
*ROS1* is regulated in methylation mutants at a transcriptional level .......... 33
4kb of *ROS1* promoter is not sufficient for methylation-sensitive expression .......... 33
Regions 3' of *ROS1* do not have strong effects on *ROS1* transcription .......... 37
The region from -17kb to -2kb of *ROS1* may contain a regulatory element necessary in *cis* for *ROS1* transcription to occur .......... 39
An 817bp region 5' of *ROS1* is a methylation-sensitive repressor of *ROS1* transcription .......... 41

Discussion .......... 43

Methods .......... 50

Acknowledgments .......... 53

References .......... 55

Chapter Three: Active demethylation is required for regulation of *FLC,* correct flowering time, and anther dehiscence ........ 60
Abstract ........ 61
Introduction ........ 61
Results ........ 63
Isolating quadruple demethylase mutants ........ 63
Demethylation is necessary for proper timing of flowering, as regulated by *FLC* ........ 65
Demethylation by DRDD induces timely anther dehiscence and affects male fertility ........ 67
Discussion ........ 69
Methods ........ 72
References ........ 75
Chapter Four: conclusions and future directions ........ 80
Thesis summary ........ 81
Future directions ........ 82
References ........ 86

# Chapter One

# INTRODUCTION

## INTRODUCTION TO METHYLATION

Methylation of the fifth carbon of cytosine is an epigenetic modification found in all vertebrates and flowering plants, as well as numerous fungi, invertebrate, and bacterial species (Goll and Bestor, 2005). Unlike in mammals, where DNA methylation occurs primarily in CG dinucleotides, cytosines in plant genomes can be methylated in all three DNA contexts: CG, CHG, and CHH, where H is any nucleotide besides guanine (Law and Jacobsen, 2010). In Arabidopsis, methylation is enriched over repetitive sequences, including transposable elements (TEs), and heterochromatin, but is also found within approximately one-third of genes (Zhang et al., 2006; Zilberman et al., 2007). By silencing TEs and affecting transcription at a subset of genes, DNA methylation is intrinsic to development, plant defense, genomic imprinting, and responses to abiotic stress (Matzke and Mosher, 2014).

## ESTABLISHMENT AND MAINTENANCE OF DNA METHYLATION

### RNA-directed DNA Methylation

*De novo* methylation in all three sequence contexts is primarily carried out by DOMAINS REARRANGED METHYLTRANSFERASE2 (DRM2) (Law and Jacobsen, 2010) through the RNA-directed DNA Methylation (RdDM) pathway (Matzke and Mosher, 2014). DRM2 is directed to methylate DNA by the actions of two plant-specific polymerases, paralogs of RNA Polymerase II (Ream et al., 2008). RNA Polymerase IV first transcribes a short single-stranded RNA from the loci, which is converted to a double-stranded RNA by RNA-Dependent RNA Polymerase 2 (RDR2) (Haag et al., 2012). This RNA is cut into 24nt small interfering RNAs (siRNAs) by DICER-LIKE3 (Xie et al., 2004), then loaded into ARGONAUTE4 (Zilberman et al., 2003). Through interactions of the AGO-siRNA complex with another local transcript, produced by RNA Polymerase V, DRM2 is directed to add methylation to the locus in question (Wierzbicki et al., 2008; Matze and Mosher, 2014) (Fig. 1).

### Maintenance of DNA methylation

Separate mechanisms have evolved to perform the maintenance of methylation of the three cytosine sequence contexts. METHYLTRANSFERASE1 (MET1), the plant homolog of the mammalian DNA METHYLTRANSFERASE1 (DNMT1), maintains symmetric CG methylation through cell divisions by methylating hemimethylated CG dinucleotides after DNA replication (Saze et al., 2003). CHROMOMETHYLASE3 (CMT3) maintains CHG methylation

(Bartee et al., 2001; Lindroth et al., 2001) in conjunction with a positive feedback loop wherein the presence of CHG methylation promotes recruitment of H3K9 methyltransferases, and H3K9me2 promotes the recruitment of CMT3 (Du et al., 2015). CMT3 is also required to initiate gene body methylation, and plant species that have lost CMT3 have also lost gene body methylation, despite the fact that gene body methylation is almost exclusively in the CG context (Bewick et al., 2016). Finally, CHH methylation at a subset of loci is established and maintained by CHROMOMETHYLASE2 (CMT2), as this context is not symmetric and thus must be re-established *de novo* following DNA replication (He et al., 2021).

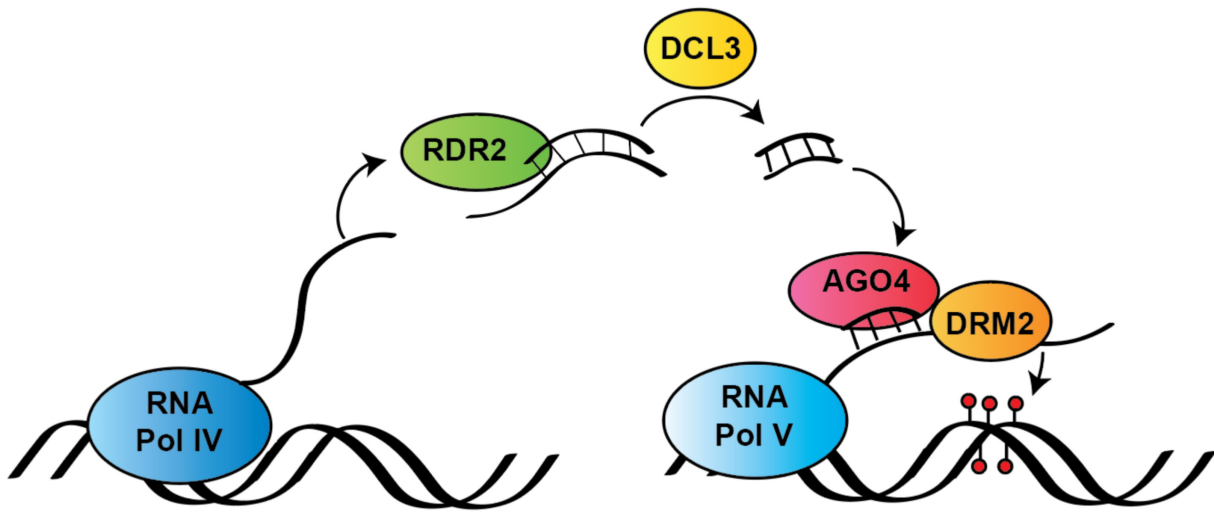


**Figure 1: RNA-directed DNA Methylation (RdDM)**
RdDM is the pathway responsible for the majority of *de novo* methylation in Arabidopsis. This process begins with the transcription of the locus to be methylated by the plant-specific RNA Polymerase IV (dark blue). The resulting transcript is converted into double-stranded RNA by RNA-DEPENDENT RNA POLYMERASE2 (RDR2; green), then processed by DICER-LIKE3 (DCL3; yellow) into 24-nt siRNAs. siRNAs are used as the guide strand in an ARGONAUTE4 complex (AGO4; red), which interacts with DOMAINS REARRANGED METHYLTRANSFERASE2 (DRM2; orange) and a transcript produced by RNA Pol V (light blue) to add methylation to the locus.

## REMOVAL OF DNA METHYLATION

DNA methylation can be lost through passive demethylation when DNA methylation is not reestablished on daughter strands after DNA replication, or DNA methylation can be actively removed by the actions of a family of DNA glycosylases.

### Discovery and characterization of ROS1

*REPRESSOR OF SILENCING1, ROS1*, is the most highly expressed DNA demethylase in the somatic tissues of Arabidopsis (Schmid et al., 2005; Williams et al., 2015). *ROS1* was initially identified by an EMS-mutagenesis screen for anti-silencing mutants, mutants that efficiently silenced a stably-expressed reporter transgene (Gong et al., 2002). It is part of the DME family, which consists of four DNA glycosylases/lyases that remove DNA methylation through base-excision repair: *DEMETER* (*DME*), *ROS1* (or *DEMETER-LIKE1 / DML1*), *DML2*, and *DML3* (Choi et al., 2002; Morales-Ruiz et al., 2006; Gehring et al., 2006; Agius et al., 2006). DME family genes contain a highly conserved helix-hairpin-helix glycosylase domain and two additional well-conserved domains that have no known homology to other proteins (Fig. 2) (Gong et al., 2002; Ponferrada-Marín et al., 2011; Brooks et al., 2014).

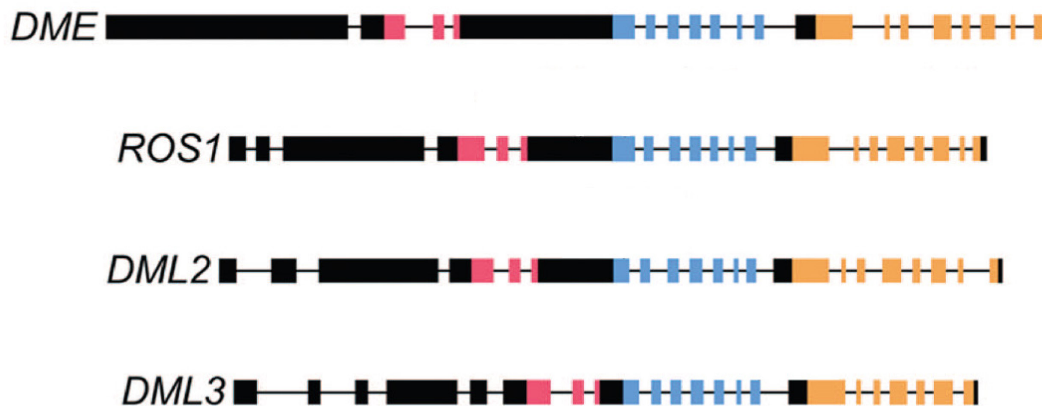


**Figure 2: The DME family of DNA glycosylases/lyases has three conserved domains.** Diagram of gene structure for the DME family, where boxes and lines indicate exons and introns, respectively, and conserved domains are in color. Red = Domain A; Blue = helix-hairpin-helix DNA glycosylase domain; Gold = Domain B. Adapted from Penterman et al. (2007b).

**Functionality of the DME family**

Analysis of DME family mutants has provided insight into the targets of these genes and their role in DNA methylation dynamics in Arabidopsis. While ROS1, DML2, and DML3 all have discrete non-redundant target loci, the majority of their targets are shared to some degree (Penterman et al., 2007b). DML2, for example, specifically demethylates the *SUPERMAN* gene (Penterman et al., 2007b), epialleles of which can result in excess male or female reproductive organs in flowers (Jacobsen and Meyerowitz, 1997, Bondada et al., 2020). Triple mutants of *ros1;dml2;dml3* (*rdd*) are viable, have similar overall genomic methylation levels to wildtype, and display no striking phenotypic differences from wildtype (Penterman et al., 2007b). However, there are hundreds of discrete regions mostly in the chromosome arms where DNA methylation is altered in *rdd* mutants, with approximately 75% of targets being enriched for signifiers of RdDM activity (Penterman et al., 2007b; Lister et al., 2008). Analysis of hypermethylated regions in *rdd* revealed that these proteins preferentially remove methylation from promoters and 3' UTRs, but also act at 5' UTRs and gene bodies (Lister et al., 2008). RDD proteins are also necessary to prevent the spreading of DNA methylation from silenced TEs or from methylated gene body sequences into promoter regions, which could lead to aberrant gene silencing (Penterman et al., 2007b).

These analyses only take into account the lack of activity of these three demethylases, however; DME may be compensating for the loss of the other three demethylases by removing methylation elsewhere in the genome. As the *dme* mutation is lethal if inherited from the mother (Choi et al., 2002), redundancy of the DME family was impossible to investigate until recently, when two labs complemented *dme* mutants in the necessary tissue of the central cell and were able to generate quadruple mutants for all DME genes in somatic tissues (Williams et al., 2021; Zeng et al., 2021). This revealed that hundreds of loci rely on demethylation activity to maintain their wildtype methylation state, including genes necessary for responding to bacterial and fungal pathogens, and genes with differential methylation between reproductive and vegetative tissue (Williams et al., 2021; Zeng et al., 2021). This has led to the conclusion that there is significant redundancy amongst all four DRDD family members, but that DME has certain specific roles that cannot be fulfilled by ROS1, DML2, or DML3.

**Mechanism of demethylation**

The mechanism by which ROS1 performs base excision repair to remove 5-methylcytosine has been well-characterized, though questions still remain. ROS1 binds nonspecifically to both methylated and unmethylated DNA (Ponferrada-Marín et al., 2010) and slides along the DNA strand, flipping out each base as it goes (Ponferrada-Marín et al., 2012, Parilla-Doblas et al., 2013). When it encounters a methylated cytosine, it nicks the DNA backbone to produce a Schiff base intermediate and release β and γ products, leaving an abasic site that can be repaired by DNA polymerase and ligases (Gehring et al., 2006; Jang et al., 2014). Possibly because of this sliding mechanism, ROS1 is a slow-turnover protein that releases DNA after every removal (Ponferrada-Marín et al., 2009). ROS1 target loci are enriched for the chromatin marks H3K18ac and H3K27me3, and ROS1 interacts directly with histone variant H2A.Z to mediate local DNA demethylation (Tang et al., 2016; Nie et al., 2019). Furthermore, ROS1 has been reported to form complexes with several discrete sets of proteins, indicating that our understanding of ROS1 activity is still incomplete as we do not know if these are indeed separate complexes, if they act at discrete loci, or if they are specific to particular conditions (Córdoba-Cañero et al., 2017; Liu et al., 2021).

**Demethylation in metazoans**

While animals do have DNA glycosylases that remove 5-methylcytosine (Jost, 1993), these proteins have been found to have an order of magnitude higher activity removing the thymine of T-G mismatches than 5-methylcytosine (Zhu et al., 2000). This is in contrast to DEMETER and ROS1, which both process 5-methylcytosine faster than thymine in similar mismatches (Morales-Ruiz et al., 2006). Instead, animals utilize the TET enzymes to iteratively oxidize 5-methylcytosine to 5-hydroxymethylcytosine (5-hmC), 5-formylcytosine, and 5-carboxylcytosine, allowing other enzymes or passive demethylation to remove the modified base (Tahiliani et al., 2009; Pastor et al., 2013). It was confirmed that this mechanism is not conserved in Arabidopsis, as 5-hmC is not present in the genome (Erdmann et al., 2015).

TET enzymes serve many and varied roles in metazoan cells. In human pluripotent stem cells active demethylation by TET enzymes was shown to be occurring at approximately 13,000 loci, in balance with *de novo* DNMT3a/b activity (Charlton et al., 2020) and these enzymes are particularly active at enhancer elements (Ginno et al., 2020; Parry et al., 2021). TET activity is in fact necessary for embryonic development (Dawlaty et al., 2014). In the brain and neural tissues,

TET enzymes play a critical role in neural lineage commitment, and the TET product 5-hmC is a stable mark that may promote local transcription (MacArthur and Dawlaty, 2021). Additionally, TET enzymes serve as tumor suppressor genes, hypothesized to utilize vitamin C as a cofactor to directly influence DNA methylation levels (Kohli and Zhang, 2013; Seethy et al., 2021; Brabson et al., 2021). Indeed, 27% of patients with acute myeloid leukemia were found to have mutations in *TET2* (Weissmann et al., 2012), patients with acute lymphoblastic leukemia have low levels of *TET2* (Musialik et al., 2014), and altered *TET* enzyme activity has been linked as well to prostate cancer, gastric cancer, melanoma, and other cancers (Seethy et al., 2021).

While the mechanism of demethylation is not conserved between plants and animals, it is likely that there has been considerable convergent evolution for mechanisms of gene regulation by balancing active methylation, active demethylation, and passive demethylation (Williams and Gehring, 2020). To this end, the study of the regulation of active demethylation in Arabidopsis may lead to important insights into the epigenetic landscape in animals. Additionally, the insertion of a transgene with fused deadCas9:ROS1 was found to be functional in HEK293 cells, allowing for targeted demethylation; TET enzymes were not able to perform in this manner (Devesa-Guerra et al., 2020). Thus, an understanding of the functionality of ROS1 may lead to a better toolkit for experimentation in mammals.

## FUNCTIONS OF DNA METHYLATION

### Silencing TEs

Transposable elements (TEs) are discrete selfish units of DNA that can mobilize within the genome (McClintock, 1950). Translocation of TEs results in mutations that can lead to loss of function of genes or chromosomal breakage (Slotkin and Martienssen, 2007). As 20% of the Arabidopsis genome is composed of TEs, a concerted effort is necessary to silence and maintain silencing of the full complement of TEs. This is accomplished in plants with three mechanisms: siRNA-mediated post-transcriptional silencing, DNA methylation by the RdDM pathway, and histone modifications (Cui and Cao, 2014). Some transcripts of TEs made by RNA Pol II activate the RdDM pathway and are processed into Argonaute complexes; this promotes the activity of the methyltransferase DRM2 at all sequences matching the TE transcript, thus targeting all copies of the TE simultaneously for silencing (Fultz et al., 2015; Rymen et al., 2020). This leads to extensive methylation of TEs within the genome in all three sequence

contexts, and even TEs that overlap genes tend to be methylated in at least one sequence context, CG, CHG, or CHH (Le et al., 2015). Hypomethylation of TEs, as can occur in methylation mutants, can result in increased TE transcription and transposition (Miura et al., 2001; Lister et al., 2008; Slotkin et al., 2009; Mirouze et al., 2009).

Given the compactness of the Arabidopsis genome however, with 44% of genes having a TE within 2kb of their transcribed region, it becomes important that the methylation on the TEs does not expand outside of the element boundaries. Indeed, mutants of the genes required for active demethylation show spreading of methylation from TEs (Penterman et al., 2007b; Le et al., 2014; Tang et al., 2016). Methylation is therefore required to silence TEs, and active demethylation is necessary to prevent ectopic silencing.

**Imprinting**

The phenomenon of imprinting, when allelic expression is altered based on parental origin, occurs in eutherian mammals, some insects, and flowering plants (Gehring, 2013; Batista and Köhler, 2020). Imprinting primarily occurs in the endosperm, a nutritive tissue that is derived from the fertilization of the diploid maternal central cell by the haploid paternal sperm cell. Differential expression is established by epigenetic differences between the parental genomes; in general, the maternal genome loses methylation in all three sequence contexts due to the DNA demethylase DEMETER removing methylation from repeat sequences and transposable elements in the central cell of the female gamete (Choi et al., 2002; Gehring, 2013), while the paternal genome from the sperm cell is hypomethylated solely in the CHH context (Calarco et al., 2012). However, TEs adjacent to genes identified as maternally-expressed in the endosperm were found to be targeted by RdDM and consequently have higher levels of CHH methylation in sperm (Calarco et al., 2012). Thus, between 28% and 54% of maternally expressed genes (MEGs) from Arabidopsis, rice, and maize are associated with a differentially methylated region, such that the hypomethylated maternal allele is expressed and the methylated paternal allele is silenced (Batista and Köhler, 2020). The E(z) homolog *MEDEA* is a maternally-expressed gene in Arabidopsis, and is specifically demethylated in the central cell by DEMETER; this is necessary for the development of the seed, as without maternally-expressed MEDEA, the seed will abort (Gehring et al., 2006). For some MEGs, imprinting can be abolished by the loss of the maintenance methyltransferase MET1 in sperm cells, leading to

hypomethylation of the paternal allele as well and confirming that the methylation differences between the two alleles are causal for imprinting (Kinoshita et al., 2004).

Paternally-expressed imprinted genes (PEGs) are established utilizing different mechanisms from MEGs. The transcription factor *HDG3*, for example, is expressed from the methylated paternal allele and not from the hypomethylated maternal allele; this methylation was determined to be causal as a natural accession that lacked the methylation did not have imprinted expression of *HDG3*, but imprinting could be induced by adding ectopic methylation to the hypomethylated allele (Pignatta et al., 2018).

One mechanism by which this can occur is that the methylation of the paternal genome precludes the activity of the FERTILIZATION INDEPENDENT-Polycomb Repressive Complex 2 (FIS-PRC2) from depositing H3K27me3 at the paternal allele during endosperm development (Gehring et al., 2006; Batista and Köhler, 2020). H3K27me3 induces a transcriptionally-silenced state to the maternal allele, leading to expression occurring solely from the paternal allele. This is the case at *PHERES1* (*PHE1*), a MADS-box transcription factor that is paternally expressed (Köhler et al., 2005; Makarevich et al., 2008). However, only 36% of PEGs in Arabidopsis show H3K27me3 deposition on the maternal allele, indicating that additional regulatory mechanisms exist (Batista and Köhler, 2020). Recent work has indicated that PHE1 may act as a master transcription factor for endosperm development, and that it transcriptionally activates several other genes that have been previously identified as PEGs (Batista et al., 2019). This correlates with the fact that PEGs are generally associated with TEs (Martinez et al., 2018), as the PHE1 binding site has been propagated through the genome by RC/Helitron transposition (Batista et al., 2019). A similar phenomenon has been reported in the crop species *Brassica napus*, where imprinted genes were enriched for the presence of nearby TEs (Rong et al., 2021).

**Regulation of transcription**

While the majority of methylation or demethylation mutants do not incur vast transcriptional dysregulation throughout the genome, there are select loci that are clearly closely regulated by local methylation states (Penterman et al., 2007; Dowen et al., 2012; Williams et al., 2017; Williams et al., 2021). As mentioned previously, imprinted loci are generally governed by methylation state, and interfering with the methylation leads to changes in expression (Pignatta et al., 2018). Methylated epialleles have also been implicated in hybrid incompatibility, petal shape, leaf senescence, vitamin E accumulation in tomatoes, sex determination in melons, and height in

rice (Cubas et al., 1999; Blevins et al., 2017; Srikant and Wibowo, 2021). Methylation by the RdDM pathway even affects apple skin coloration by altering expression of the *MYB1* gene in various apple sports (Jiang et al., 2020).

Genes that are regulated by methylation are more likely to have TEs or TE fragments in close proximity, as is the case for the DNA demethylase *ROS1* (Williams et al., 2015). Indeed, 67% of genes that were strongly downregulated (more than 4-fold) in the triple DNA demethylase mutant *ros1;dml2;dml3* had a TE in their promoter as compared to 30% of randomly selected genes (Le et al., 2014), and ROS1 preferentially targets TEs that are closer to genes (Tang et al., 2016). This is intriguing when considered in conjunction with the evidence that TE insertions can lead to gene diversification and altered transcription, as exemplified by the heat-activated TE *ONSEN* conferring heat-induced transcription to neighboring genes, or *mPING* element insertions in rice leading to cold- and stress-inducible transcription of the genes into which it transposes (Girard and Freeling, 1999; Naito et al., 2009; Ito et al., 2011).

Alternatively, transcription can also be affected by methylation-sensitive transcription factors. The E2F family of transcription factors, for example, have been shown to have different preferences for methylated versus unmethylated DNA *in vitro* (Campanero et al., 2000; O'Malley et al., 2016), and E2F binding sites were identified as being significantly enriched for differential methylation between 1107 accessions of *Arabidopsis thaliana* (Kawakatsu et al., 2017). This differential methylation may lead to E2F factors being more or less likely to bind, altering local transcription. In mammals, the insulator CTCF does not bind methylated DNA (Bell and Felsenfeld, 2000), and several genes require CTCF to bind only the unmethylated allele to establish imprinted expression (Bell and Felsenfeld, 2000; Noordermeer and Feil, 2020). As an example, a SNP at a CTCF binding sites in the 5' UTR of the antiviral protein *IFITM3* was found to alter methylation at the binding site, leading to reduced CTCF binding, lower mRNA levels, and increased risk of severe influenza infections (Allen et al., 2017). Several families of transcription factors show similar sensitivities to methylation in animals, including the bHLH, bZIP, and ETS families (Wang et al., 2019). Similar mechanisms may be in effect in *Arabidopsis*, although it has not been studied to the same extent.

## DYNAMICS OF DNA METHYLATION

### DNA methylation is a highly stable mark

Patterns of DNA methylation remain stable over generational time. The yeast *Cryptococcus neoformans* lost all copies of *de novo* methyltransferases more than 50 million years ago, yet its methylation patterns continue to be faithfully propagated by maintenance methyltransferases (Catania et al., 2019). In plants, gene body methylation patterns have been maintained between orthologs separated by more than 100 million years (Takuno and Gaut, 2013). A study of recombinant inbred lines between two strains of Arabidopsis with different average levels of gene body methylation revealed that after nine generations the methylation status of a region could confidently be used to infer the identity of the parental strain for that region (Picard and Gehring, 2017). This indicates faithful maintenance of regional methylation status, throughout even hybrid and heterozygous states. While there are examples of paramutation, where one epiallele can induce methylation and chromatin state changes in a separate allele in *trans*, this phenomenon appears to be rare (Brink, 1956; Chandler and Stam, 2004; El-Sappah et al., 2021).

Arabidopsis was found to have an epimutation rate of approximately $4.5 \times 10^{-4}$ methylation changes in CG methylation per generation, roughly five orders of magnitude higher than the genetic mutation rate (Schmitz et al., 2011; Weng et al., 2019; Denkena et al., 2021). Methylation of TEs is the most stable, followed by intergenic methylation of gene promoters and downstream regions, with gene body methylation being the most susceptible to change (Becker et al., 2011; Denkena et al., 2021); the genetic mutation rate for these same regions is inverted, with genes accumulating mutations at a much lower rate than TEs (Weng et al. 2019). Indeed, the strongest predictor of instability in CG methylation was found to be an intermediate level of methylation, potentially indicating a region undergoing both *de novo* methylation and active demethylation (Picard and Gehring., 2017). Altogether, gross patterns of DNA methylation are maintained over generational time to the scale of hundreds of millions of years, although individual cytosines may gain or lose methylation. Of particular relevance to this thesis, methylation upstream of *ROS1* orthologs has been maintained for at least 150 million years between *Zea mays* and *Arabidopsis thaliana* (Williams et al., 2015).

**DNA methylation is a highly labile mark**

Paradoxically, DNA methylation is not only stable over millennia, but is also able to vary within a single plant and between plants. In comparisons of 1107 methylomes of Arabidopsis, 78% of methylated cytosines are differentially methylated in at least one accession (Kawakatsu et al., 2016). These methylation differences can lead to variation in gene expression and phenotype; to reiterate an earlier example, the transcription factor *HDG3* is imprinted in the Col-0 ecotype, but not in the Cvi ecotype. The Cvi allele is hypomethylated compared to the Col-0 allele, and this methylation difference was determined to be causal for the difference in expression between the alleles (Pignatta et al., 2018).

Epigenetic changes are also necessary for the proper development of the plant. Loss of DEMETER results in seed abortion due to improper regulation of *MEDEA* in the maternal genome (Choi et al., 2002), and DNA methylation is also necessary for proper embryo development (Xiao et al., 2006; Liang et al., 2021). ROS1 is required for proper cell differentiation in both leaf epidermis and xylem differentiation, and *ros1* mutants have an overproduction of stomata and errors in tracheary development (Yamamuro et al., 2014; Lin et al., 2020). Reproductive tissues of Arabidopsis have increased levels of methylation compared to vegetative tissue (Feng et al., 2020; Williams et al., 2021), and fruit development and ripening in tomato as well as proper pollen tube growth in Arabidopsis require active demethylation from DEMETER and ROS1 or their homologs (Zhong et al., 2013; Liu et al., 2015; Lang et al., 2017; Khouider et al., 2021).

Finally, methylation is required for plants to respond to their environment, especially as they are sessile and cannot remove themselves from hostile climates or events. Many of the genes that regulate DNA methylation or demethylation have altered expression when phytohormone levels are altered, as in stress conditions (Bennett et al., 2021). Rapid active demethylation of defense genes by the DME family upon pathogen detection allows for the transcriptional upregulation of plant disease resistance proteins (Yu et al., 2012; Le et al., 2014; Schumann et al., 2019; Halter et al., 2021). Exposing tomatoes to low temperatures, as in refrigeration after harvest, leads to both substantial methylation changes and a correlated loss of flavor (Zhang et al., 2016). Even exposure to microgravity, as in Arabidopsis plants grown during spaceflight, resulted in altered global methylation patterns, especially near genes associated with RdDM, hormone signaling, and cell-wall modification (Xu et al., 2018).

Exposure to drought, heat, cold, salt, and heavy metals all induce changes in DNA methylation, allowing the plant to regulate its genome in response to stimuli (Chung et al., 2021; Miryeganeh, 2021).

**Balancing methylation and demethylation**

In order to regulate methylation in a dynamic fashion, there must be concurrent lability in the balance of activity of the demethylases and the *de novo* methylation pathways. In accordance with this hypothesis, *ROS1* transcript levels are downregulated in mutants of the RdDM pathway and *met1* (Huettel et al., 2006; Williams et al., 2015). This means that when global methylation levels are low, activity of ROS1 is also lowered. Reduced expression of *ROS1* in RdDM mutants was found to be correlated with proximal promoter methylation upstream of *ROS1,* overlapping a Helitron transposable element. When methylation of the Helitron was lost in RdDM mutants, *ROS1* transcription was lowered; conversely, when methylation was added ectopically in the same mutants *ROS1* transcription was increased (Williams et al., 2015). Furthermore, ROS1 demethylates the same region in the *ROS1* promoter, correlating with an increase in *ROS1* transcript level in the hypermethylated *ros1* mutant (Williams et al., 2015; Córdoba-Cañero et al., 2017). This is indicative of a *ROS1* epigenetic negative feedback loop wherein methylation at the *ROS1* promoter leads to increased production of ROS1 protein, which removes the methylation from the promoter (Fig. 3). On a wider scale, this promotes homeostasis of the global methylation level as the production of ROS1 is proportionate to overall methylation (Williams et al., 2015; Lei et al., 2015).

*ROS1* transcript levels are downregulated to a greater extent in *met1* mutants than in RdDM mutants (Williams et al., 2015). This correlates with the greater methylation difference in *met1* mutants, which lose all but 1% of CG methylation and approximately half of CHH methylation (Lister et al., 2008). However, this downregulation appears to occur via a different mechanism than was described in RdDM mutants. The downregulation of *ROS1* in *met1* mutants is dependent on reduced expression of the histone demethylase *IBM1* (Rigal et al., 2012), while downregulation in RdDM mutants was *IBM1*-independent (Williams et al., 2015). Despite the different mechanism, however, *ROS1* expression levels remain in balance with the overall level of genomic methylation.

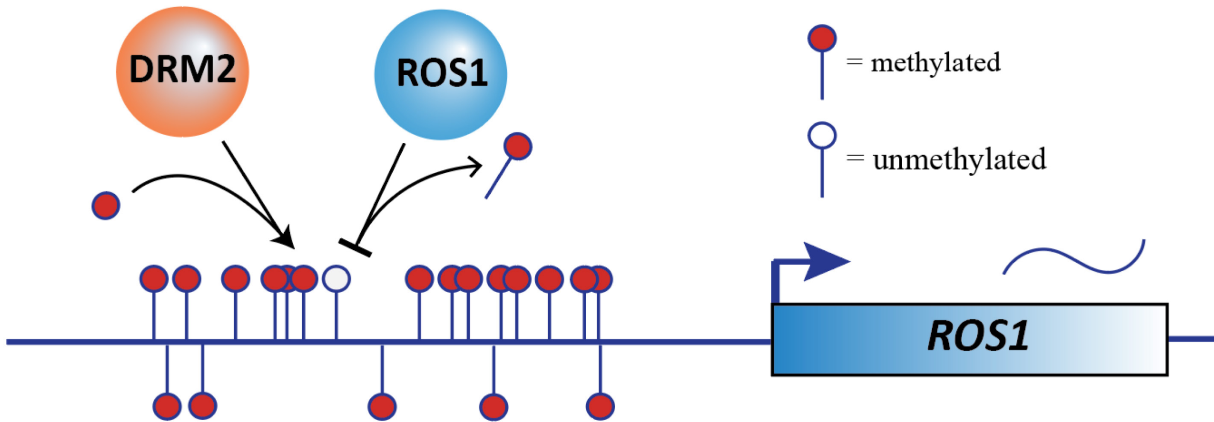


**Figure 3: ROS1 is regulated by a methylation-sensitive negative feedback loop.** Transcription of *ROS1* is increased by the addition of methylation in the proximal promoter region from approximately -300 to -85bp from the transcription start site. This methylation is added by DRM2, the *de novo* methyltransferase of the RdDM pathway (orange). The ROS1 protein (blue) targets this same region, removing methylation via base excision repair and thus lowering *ROS1* transcription levels. This is termed the 'epigenetic rheostat' mechanism of methylation homeostasis (Williams et al., 2015).

**SUMMARY OF THESIS**

This thesis investigates the regulation of active DNA demethylation in the *Arabidopsis thaliana* genome and the function thereof. Transcription of *ROS1* is known to be affected atypically by methylation, in that methylation in the proximal promoter region causes increased transcription rather than gene silencing. The mechanism by which this atypical response is accomplished, however, is unclear. Additionally, while one DNA demethylase is regulated by methylation levels, the other DNA demethylases also play a role in somatic tissue. I also endeavored to investigate the role of DNA demethylation as a whole in the plant.

Chapter Two of the thesis, "A DNA demethylase in Arabidopsis is transcriptionally regulated in cis by a methylation-sensitive silencer sequence", addresses the question of how *ROS1* transcription is regulated. In this chapter, I showed that ROS1 is transcriptionally regulated by sequences outside of its proximal promoter through analysis of transgenes with varying amounts of endogenous sequence. I then used CRISPR-Cas9 technology to create a series of deletions in the endogenous chromosome near *ROS1*. Transcription of *ROS1* was

assayed in siblings with and without these deletions, both in methylation mutant and methylation-typical backgrounds. Deletions 3' of *ROS1* had minimal effects on *ROS1* transcription, but I identified that the region from -17kb to -2kb may be essential for *ROS1* transcription, as well as containing an essential gene. Finally, I showed that an 817bp region within the proximal promoter of *ROS1* is a methylation-sensitive silencer of *ROS1* transcription; without this region, *ROS1* transcription is upregulated in both wildtype and hypomethylated backgrounds. This has led to a new model for *ROS1* regulation by methylation, wherein the repressive effect of this silencer region on *ROS1* transcription is anticorrelated with its level of methylation, resulting in high levels of *ROS1* when the silencer is hypermethylated and low levels when hypomethylated.

Chapter Three, "Active demethylation is required for regulation of *FLC,* correct flowering time, and anther dehiscence", contains our investigation into the role of active demethylation in the adult plant. We generated a transgene to complement a *dme* mutation solely in the central cell, resulting in viable embryos that bypass the *dme* seed abortion phenotype (Choi et al., 2002). This transgene was then used to obtain quadruple mutants of the complete DME family, *dme;ros1;dml2;dml3* (*drdd*). We detected an early flowering phenotype in *drdd* mutant plants and linked this phenotype to a decrease in transcript level of the floral repressor *FLOWERING LOCUS C*, correlated with hypermethylation in the 5' flanking region of *FLC* in *drdd* plants. This is the first time that DNA methylation or active demethylation have been linked to the regulation of *FLC*, a key checkpoint for flowering time in *Arabidopsis*. Additionally, we noticed a fertility defect in *drdd* plants, caused by delays in anther dehiscence, meaning that pollen is being released late from the male organs of the flower. This may be caused by a delay in the accumulation of reactive oxygen species, as reactive oxygen species are necessary for the secondary wall thickening that leads to the anther breaking open and releasing pollen. While recent reports have begun to investigate the role of the RdDM pathway in anther dehiscence, this is the first report of active demethylation being necessary for this process.

Chapter Four, "Conclusions and Future Directions", reviews the findings of this thesis and present lines of inquiry that have been opened and may prove fruitful.

## REFERENCES


Agius, F., Kapoor, A., and Zhu, J. (2006) Role of the Arabidopsis DNA glycosylase/lyase ROS1 in active DNA demethylation. PNAS 103(31):11796-11801.

Allen, E. K., Randolph, A. G., Bhangale, T., Dogra, P., Ohlson, M., Oshansky, C. M., Zamora, A. E., Shannon, J. P., Finkelstein, D., Dressen, A., DeVincenzo, J., Caniza, M., Youngblood, B., Rosenberger, C. M., and Thomas, P. G. (2017) SNP-mediated disruption of CTCF binding at the *IFITM3* promoter is associated with risk of severe influenza in humans. Nature Medicine, 23(8): 975–983. https://doi.org/10.1038/nm.4370

Bartee, L., Malagnac, F., and Bender, J. (2001) *Arabidopsis cmt3* chromomethylase mutations block non-CG methylation and silencing of an endogenous gene. Genes & Development 15: 1753-1758. https://doi.org/10.1101/gad.905701

Batista, R. A., Moreno-Romero, J., Qiu, Y., van Boven, J., Santos-González, J., Figueiredo, D.D., and Köhler, C. (2019) The MADS-box transcription factor PHERES1 controls imprinting in the endosperm by binding to domesticated transposons. eLife 8: e50541. doi: 10.7554/eLife.50541.

Batista, R. A., and Köhler, C. (2020) Genomic imprinting in plants-revisiting existing models. Genes & Development 34(1-2): 24–36. https://doi.org/10.1101/gad.332924.119

Becker, C., Hagmann, J., Müller, J., Koenig, D., Stegle, O., Borgwardt, K., & Weigel, D. (2011). Spontaneous epigenetic variation in the Arabidopsis thaliana methylome. Nature 480(7376): 245–249. https://doi.org/10.1038/nature10555.

Bell, A.C., and Felsenfeld, G. (2000) Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. Nature 405(6785): 482-485. doi: 10.1038/35013100.

Bennett, M., Cleaves, K., and Hewezi, T. (2021) Expression patterns of DNA methylation and demethylation genes during plant development and in response to phytohormones. Int J Mol Sci. 22(18): 9681. doi: 10.3390/ijms22189681.

Bewick, A. J., Ji, L., Niederhuth, C. E., Willing, E. M., Hofmeister, B. T., Shi, X., Wang, L., Lu, Z., Rohr, N. A., Hartwig, B., Kiefer, C., Deal, R. B., Schmutz, J., Grimwood, J., Stroud, H., Jacobsen, S. E., Schneeberger, K., Zhang, X., and Schmitz, R. J. (2016) On the origin and evolutionary consequences of gene body DNA methylation. PNAS 113(32): 9111–9116. https://doi.org/10.1073/pnas.1604666113

Blevins, T., Wang, J., Pflieger, D., Pontvianne, F., and Pikaard, C.S. (2017) Hybrid incompatibility caused by an epiallele. PNAS 114(14): 3702-3707. https://doi.org/10.1073/pnas.1700368114

Bondada, R., Somasundaram, S., Marimuthu, M.P., Badarudeen, M.A., Puthiyaveedu, V.K., and Maruthachalam, R. (2020) Natural epialleles of *Arabidopsis SUPERMAN* display *superwoman* phenotypes. Communications Biology 3: 772. https://doi.org/10.1038/s42003-020-01525-9.

Brabson, J. P., Leesang, T., Mohammad, S., and Cimmino, L. (2021). Epigenetic regulation of genomic stability by vitamin C. Frontiers in Genetics 12: 675780. https://doi.org/10.3389/fgene.2021.675780.

Brink, R.A. (1956) A genetic change associated with the R locus in maize which is directed and potentially reversible. Genetics 41:872. doi:10.2307/2407003.

Brooks, S., Fischer, R., Huh, J., and Eichman, B. (2014) 5-methylcytosine recognition by Arabidopsis thaliana DNA glycosylases DEMETER and DML3. *Biochemistry* 53(15):2525-2532.
Calarco, J. P., Borges, F., Donoghue, M. T., Van Ex, F., Jullien, P. E., Lopes, T., Gardner, R., Berger, F., Feijó, J. A., Becker, J. D., and Martienssen, R. A. (2012) Reprogramming of DNA methylation in pollen guides epigenetic inheritance via small RNA. Cell 151(1): 194–205. https://doi.org/10.1016/j.cell.2012.09.001
Campanero, M.R., Armstrong, M.I., and Flemington, E.K. (2000) CpG methylation as a mechanism for the regulation of E2F activity. PNAS 97(12): 6481-6486. doi: 10.1073/pnas.100340697.
Catania, S., Dumesic, P. A., Pimentel, H., Nasif, A., Stoddard, C. I., Burke, J. E., Diedrich, J. K., Cook, S., Shea, T., Geinger, E., Lintner, R., Yates, J. R., 3rd, Hajkova, P., Narlikar, G. J., Cuomo, C. A., Pritchard, J. K., and Madhani, H. D. (2020) Evolutionary Persistence of DNA Methylation for Millions of Years after Ancient Loss of a De Novo Methyltransferase. Cell, 180(2), 263–277.e20. https://doi.org/10.1016/j.cell.2019.12.012
Chandler, V.L., and Stam, M. (2004) Chromatin conversations: mechanisms and implications of paramutation. Nature Reviews Genetics 5: 532-544. doi: 10.1038/nrg1378.
Charlton, J., Jung, E. J., Mattei, A. L., Bailly, N., Liao, J., Martin, E. J., Giesselmann, P., Brändl, B., Stamenova, E. K., Müller, F. J., Kiskinis, E., Gnirke, A., Smith, Z. D., & Meissner, A. (2020). TETs compete with DNMT3 activity in pluripotent cells at thousands of methylated somatic enhancers. Nature genetics, 52(8), 819–827. https://doi.org/10.1038/s41588-020-0639-9
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.
Chung, S., Kwon, C., and Lee, JH. (2021) Epigenetic control of abiotic stress signaling in plants. Genes & Genomics https://doi.org/10.1007/s13258-021-01163-3.
Córdoba-Cañero, D., Cognat, V., Ariza, R.R., Arjona, T.R., and Molinier, J. (2017) Dual control of ROS1-mediated active DNA demethylation by the DNA DAMAGE BINDING protein 2 (DDB2). Plant Journal 92(6):1170-1181. doi: 10.1111/tpj.13753.
Cubas, P, Vincent, C., and Coen, E. (1999) An epigenetic mutation responsible for natural variation in floral symmetry. Nature 401: 157-161. https://doi.org/10.1038/43657
Cui, X., and Cao, X. (2014) Epigenetic regulation and functional exaptation of transposable elements in higher plants. Current Opinion in Plant Biology 21: 83-88. http://dx.doi.org/10.1016/j.pbi.2014.07.001.
Dawlaty, M. M., Breiling, A., Le, T., Barrasa, M. I., Raddatz, G., Gao, Q., Powell, B. E., Cheng, A. W., Faull, K. F., Lyko, F., and Jaenisch, R. (2014). Loss of Tet enzymes compromises proper differentiation of embryonic stem cells. Developmental Cell 29(1): 102–111. https://doi.org/10.1016/j.devcel.2014.03.003
Dowen, R.H., Pelizzola, M., Schmitz, R.J., Lister, R., Dowen, J.H., Nery, J.R., Dixon, J.E., and Ecker, J.R. (2012) Widespread dynamic DNA methylation in response to biotic stress. PNAS 109(32): E2183-E2191. https://doi.org/10.1073/pnas.1209329109.
Denkena, J., Johannes, F., and Colomé-Tatché, M. (2021) Region-level epimutation rates in *Arabidopsis thaliana*. Heredity 127: 190-202. https://doi.org/10.1038/s41437-021-00441-w.

Devesa-Guerra, I., Morales-Ruiz, T., Pérez-Roldán, J., Parrilla-Doblas, J. T., Dorado-León, M., García-Ortiz, M. V., Ariza, R. R., and Roldán-Arjona, T. (2020) DNA Methylation Editing by CRISPR-guided Excision of 5-Methylcytosine. Journal of Molecular Biology 432(7): 2204–2216. https://doi.org/10.1016/j.jmb.2020.02.007

Du, J., Johnson, L. M., Jacobsen, S. E., and Patel, D. J. (2015). DNA methylation pathways and their crosstalk with histone methylation. Nature Reviews. Molecular cell biology 16(9): 519–532. https://doi.org/10.1038/nrm4043.

El-Sappah, A. H., Yan, K., Huang, Q., Islam, M. M., Li, Q., Wang, Y., Khan, M. S., Zhao, X., Mir, R. R., Li, J., El-Tarabily, K. A., and Abbas, M. (2021) Comprehensive Mechanism of Gene Silencing and Its Role in Plant Growth and Development. Frontiers in Plant Science 12:705249. https://doi.org/10.3389/fpls.2021.705249.

Erdmann, R. M., Souza, A.L., Clish, C.B., and Gehring, M. (2015) 5-Hydroxymethylcytosine is not present in appreciable quantities in *Arabidopsis* DNA. G3 Genes|Genomes|Genetics 5(1): 1-8. doi: 10.1534/g3.114.014670.

Feng, S., Zhong, Z., Wang, M., and Jacobsen, S.E. (2020) Efficient ad accurate determination of genome-wide DNA methylation patterns in *Arabidopsis thaliana* with enzymatic methyl sequencing. Epigenetics & Chromatin 13: 42. https://doi.org/10.1186/s13072-020-00361-9.

Fultz, D., Choudury, S.G., and Slotkin, R.K. (2015) Silencing of active transposable elements in plants. Current Opinion in Plant Biology 27: 67-76. http://dx.doi.org/10.1016/j.pbi.2015.05.027.

Gehring, M., Huh, J.H., Hsieh, T.F., Penterman, J., Choi, Y., Harada, J.J., Goldberg, R.B., and Fischer, R.L. (2006) DEMETER DNA glycosylase establishes MEDEA Polycomb gene self-imprinting by allele-specific demethylation. Cell 124(3): 495-506.

Gehring, M., Reik, W., and Henikoff, S. (2009) DNA demethylation by DNA repair. Trends in Genetics 25(2): 82-90.

Gehring, M. (2013) Genomic imprinting: insights from plants. Annu Rev Genet. 47: 187-208. https://doi.org/10.1146/annurev-genet-110711-155527.

Ginno, P.A., Gaidatzis, D., Feldmann, A. et al. A genome-scale map of DNA methylation turnover identifies site-specific dependencies of DNMT and TET activity. Nat Commun 11, 2680 (2020). https://doi.org/10.1038/s41467-020-16354-x.

Girard, L., and Freeling, M. (1999) Regulatory changes as a consequence of transposon insertion. Developmental Genetics 25(4): 291-296. https://doi.org/10.1002/(SICI)1520-6408(1999)25:4<291::AID-DVG2>3.0.CO;2-5

Goll, M.G., and Bestor, T.H. (2005) Eukaryotic cytosine methyltransferases. Annu. Rev. Biochem. 74: 481-514.

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

Haag, J.R., Ream, T.S., Marasco, M., Nicora, C.D., Norbeck, A.D., Pasa-Tolić, L., and Pikaard, C.S. (2012) In vitro transcription activities of Pol IV, Pol V, and RDR2 reveal coupling of Pol IV and RDR2 for dsRNA synthesis in plant RNA silencing. Molecular Cell 48(5): 811-818. https://doi.org/10.1016/j.molcel.2012.09.027.

Halter, T., Wang, J., Amesefe, D., Lastrucci, E., Charvin, M., Rastogi, M.S., and Navarro, L. (2021) The Arabidopsis active demethylase ROS1 *cis*-regulates defence genes by erasing

DNA methylation at promoter-regulatory regions. eLife 10: e62994. DOI: https://doi.org/10.7554/eLife.62994
He, L., Zhao, C., Zhang, Q., Zinta, G., Wang, D., Lozano-Durán, R., and Zhu, JK. (2021) Pathway conversion enables a double-lock mechanism to maintain DNA methylation and genome stability. PNAS 118(35): e2107320118. https://doi.org/10.1073/pnas.2107320118.
Huettel, B., Kanno, T., Daxinger, L., Aufsatz, W., Matzke, A.J.M., and Matzke, M. (2006). Endogenous targets of RNA-directed DNA methylation and Pol IV in Arabidopsis. EMBO J. 25: 2828–2836.
Ito, H., Gaubert, H., Bucher, E., Mirouze, M., Vaillant, I., & Paszkowski, J. (2011) An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. Nature 472(7341): 115–119. https://doi.org/10.1038/nature09861
Jacobsen, S.E. and Meyerowitz, E.M. (1997) Hypermethylated SUPERMAN epigenetic alleles in *Arabidopsis*. Science 277(5329): 1100-1103. doi: 10.1126/science.277.5329.1100
Jang, H., Shin, H., Eichman, B.F., and Huh, J.H. (2014) Excision of 5-hydroxymethylcytosine by DEMETER family DNA glycosylases. Biochemical and Biophysical Research Communications 446(4): 1067-1072.
Jiang, S., Wang, N., Chen, M., Zhang, R., Sun, Q., Xu, H., Zhang, Z., Wang, Y., Sui, X., Wang, S., Fang, H., Zuo, W., Su, M., Zhang, J., Fei, Z., & Chen, X. (2020). Methylation of *MdMYB1* locus mediated by RdDM pathway regulates anthocyanin biosynthesis in apple. Plant Biotechnology Journal 18(8): 1736–1748. https://doi.org/10.1111/pbi.13337.
Jost, J. (1993) Nuclear extracts of chicken embryos promote an active demethylation of DNA by excision repair of 5-methyldeoxycytidine. PNAS 90(10): 4684-4688.
Kawakatsu, T., Huang, S. C., Jupe, F., Sasaki, E., Schmitz, R. J., Urich, M. A., Castanon, R., Nery, J. R., Barragan, C., He, Y., Chen, H., Dubin, M., Lee, C. R., Wang, C., Bemm, F., Becker, C., O'Neil, R., O'Malley, R. C., Quarless, D. X., 1001 Genomes Consortium, … Ecker, J. R. (2016) Epigenomic Diversity in a Global Collection of *Arabidopsis thaliana* Accessions. Cell 166(2): 492–505. https://doi.org/10.1016/j.cell.2016.06.044
Khouider, S., Borges, F., LeBlanc, C., Ungru, A., Schnittger, A., Martienssen, R., Colot, V., & Bouyer, D. (2021). Male fertility in *Arabidopsis* requires active DNA demethylation of genes that control pollen tube function. Nature Communications 12(1): 410. https://doi.org/10.1038/s41467-020-20606-1.
Kinoshita, T., Miura, A., Choi, Y., Kinoshita, Y., Cao, X., Jacobsen, S. E., Fischer, R. L., & Kakutani, T. (2004) One-way control of FWA imprinting in *Arabidopsis* endosperm by DNA methylation. Science 303(5657): 521–523. https://doi.org/10.1126/science.1089835
Köhler, C., Page, D. R., Gagliardini, V., and Grossniklaus, U. (2005). The *Arabidopsis thaliana* MEDEA Polycomb group protein controls expression of PHERES1 by parental imprinting. Nature Genetics 37(1): 28–30. https://doi.org/10.1038/ng1495
Kohli, R.M., and Zhang, Y. (2013) TET enzymes, TDG and the dynamics of DNA demethylation. Nature 502: 472-479.
Lang, Z., Wang, Y., Tang, K., Tang, D., Datsenka, T., Cheng, J., Zhang, Y., Handa, A. K., and Zhu, J. K. (2017). Critical roles of DNA demethylation in the activation of ripening-induced genes and inhibition of ripening-repressed genes in tomato fruit. PNAS 114(22): E4511–E4519. https://doi.org/10.1073/pnas.1705233114

Law, J.A., and Jacobsen, S.E. (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Natural Reviews Genetics 11: 204-220. https://doi.org/10.1038/nrg2719.

Le, T.N., Schumann, U., Smith, N.A., Tiwari, S., Au, P.C.K., Zhu, Q.H., Taylor, J.M., Kazan, K., Llewellyn, D.J., Zhang, R., Dennis, E.S., and Wang, M.B. (2014) DNA demethylases target promoter transposable elements to positively regulate stress responsive genes in *Arabidopsis*. *Genome Biology* 15: 458-476. https://doi.org/10.1186/s13059-014-0458-3

Lei, M., Zhang, H., Julian, R., Tang, K., Xie, S., and Zhu, JK. (2015) Regulatory link between DNA methylation and active demethylation in *Arabidopsis.* PNAS 112(11): 3553-3557.

Liang, W., Li, J., Sun, L., Liu, Y., Lan, Z., and Qian, W. (2021) Deciphering the synergistic and redundant roles of CG and non-CG DNA methylation in plant development and transposable element silencing. New Phytologist https://doi.org/10.1111/nph.17804.

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(5524): 2077–2080. https://doi.org/10.1126/science.1059745.

Lister, R., O'Malley, R., 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, R., How-Kit, A., Stammitti, L., Teyssier, E., Rolin, D., Mortain-Bertrand, A., Halle, S., Liu, M., Kong, J., Wu, C., Degraeve-Guibault, C., Chapman, N. H., Maucourt, M., Hodgman, T. C., Tost, J., Bouzayen, M., Hong, Y., Seymour, G. B., Giovannoni, J. J., and Gallusci, P. (2015). A DEMETER-like DNA demethylase governs tomato fruit ripening. PNAS 112(34): 10804–10809. https://doi.org/10.1073/pnas.1503362112

Liu, P., Nie, W. F., Xiong, X., Wang, Y., Jiang, Y., Huang, P., Lin, X., Qin, G., Huang, H., Niu, Q., Du, J., Lang, Z., Lozano-Duran, R., and Zhu, J. K. (2021) A novel protein complex that regulates active DNA demethylation in Arabidopsis. Journal of Integrative Plant Biology, 63(4): 772–786. https://doi.org/10.1111/jipb.13045.

MacArthur, I. C., and Dawlaty, M. M. (2021) TET Enzymes and 5-hydroxymethylcytosine in neural progenitor cell biology and neurodevelopment. Frontiers in cell and developmental biology, 9, 645335. https://doi.org/10.3389/fcell.2021.645335.

Makarevich, G., Villar, C. B., Erilova, A., and Köhler, C. (2008). Mechanism of *PHERES1* imprinting in *Arabidopsis*. Journal of Cell Science 121(6): 906–912. https://doi.org/10.1242/jcs.023077

Martinez, G., Wolff, P, Wang, Z., Moreno-Romero, J., Santos-González, J., Conze, L.L., DeFraia, C., Slotkin, R.K., and Köhler, C. (2018) Paternal easiRNAs regulate parental genome dosage in *Arabidopsis*. Nature Genetics 50: 193-198. https://doi.org/10.1038/s41588-017-0033-4.

Matzke, M.A., and Mosher, R.A. (2014) RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. Nature Reviews Genetics 15: 394-408. https://doi.org/10.1038/nrg3683

McClintock, B. (1950) The origin and behavior of mutable loci in maize. PNAS 36(6): 344-355. https://doi.org/10.1073/pnas.36.6.344

Mirouze, M., Reinders, J., Bucher, E., Nishimura, T., Schneeberger, K., Ossowski, S., Cao, J., Weigel, D., Paszkowski, J., and Mathieu, O. (2009) Selective epigenetic control of

retrotransposition in *Arabidopsis*. Nature 461: 427–430. https://doi.org/10.1038/nature08328
Miryeganeh, M. (2021) Plants' epigenetic mechanisms and abiotic stress. Genes 12(8): 1106. https://doi.org/10.3390/genes12081106
Miura, A., Yonebayashi, S., Watanabe, K., Toyama, T., Shimada, H., and Kakutani, T. (2001) Mobilization of transposons by a mutation abolishing full DNA methylation in *Arabidopsis*. Nature 411(6834): 212-214. https://doi.org/10.1038/35075612
Morales-Ruiz, T., Ortega-Galisteo, A., Ponferrada-Marín, M., Ariza, R., and Roldán-Arjona, T. (2006) DEMETER and REPRESSOR OF SILENCING 1 encode 5-methylcytosine DNA glycosylase. PNAS 103(18): 6853-6858.
Musialik, E., Bujko, M., Wypych, A., Matysiak, M., and Siedlecki, J. A. (2014) TET2 promoter DNA methylation and expression analysis in pediatric B-cell acute lymphoblastic leukemia. Hematology Reports, 6(1): 5333.
Naito, K., Zhang, F., Tsukiyama, T., Saito, H., Hancock, C. N., Richardson, A. O., Okumoto, Y., Tanisaka, T., and Wessler, S. R. (2009) Unexpected consequences of a sudden and massive transposon amplification on rice gene expression. Nature 461(7267): 1130–1134. https://doi.org/10.1038/nature08479
Nie, W. F., Lei, M., Zhang, M., Tang, K., Huang, H., Zhang, C., Miki, D., Liu, P., Yang, Y., Wang, X., Zhang, H., Lang, Z., Liu, N., Xu, X., Yelagandula, R., Zhang, H., Wang, Z., Chai, X., Andreucci, A., Yu, J. Q, Berger, F., Lozano-Duran, R., and Zhu, J. K. (2019) Histone acetylation recruits the SWR1 complex to regulate active DNA demethylation in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America, 116(33): 16641–16650. https://doi.org/10.1073/pnas.1906023116.
Noordermeer, D., and Feil, R. (2020) Differential 3D chromatin organization and gene activity in genomic imprinting. Curr Opin Genet Dev. 61: 17-24. doi: 10.1016/j.gde.2020.03.004.
O'Malley, R.C., Huang, S.C., Song, L., Lewsey, M.G., Bartlett, A., Nery, J.R., Galli, M., Gallavotti, A., and Ecker, J.R. (2016) Cistrome and epicistrome features shape the regulatory DNA landscape. Cell 165(5): 1280-1292. DOI:https://doi.org/10.1016/j.cell.2016.04.038
Parilla-Doblas, J.T., Ponferrada-Marín, M.I., Roldán-Arjona, T., and Ariza, R.R. (2013) Early steps of active DNA demethylation initiated by ROS1 glycosylase require three putative helix-invading residues. Nucleic Acids Research 41(18): 8654-8664.
Parry, A., Rulands, S., and Reik, W. (2021) Active turnover of DNA methylation during cell fate decisions. Nature Reviews Genetics 22:59-66.
Pastor, W.A., Arabind, L., and Rao, A. (2013) TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. Nature Revews. Molecular Cell Biology 14(6): 341-356.
Penterman, J., Uzawa, R., and Fischer, R.L. (2007a) Genetic interactions between DNA demethylation and methylation in Arabidopsis. Plant Physiology 145: 1549-1557.
Penterman, J., Zilberman, D., Huh, J.H., Ballinger, T., Henikoff, S., Fischer, R. (2007b) DNA demethylation in the *Arabidopsis* genome. PNAS 104: 6752-6757.
Picard, C.L., and Gehring, M. (2017) Proximal methylation features associated with nonrandom changes in gene body methylation. Genome Biology 18: 73. https://doi.org/10.1186/s13059-017-1206-2.

Pignatta, D., Novitsky, K., Satyaki, P.R.V., and Gehring, M. (2018) A variably imprinted epiallele impacts seed development. PLoS Genetics 14(11): e1007469. https://doi.org/10.1371/journal.pgen.1007469.

Ponferrada-Marín, M.I., Roldán-Arjona, T., and Ariza, R.R. (2009) ROS1 5-methylcytosine DNA glycosylase is a slow-turnover catalyst that initiates DNA demethylation in a distributive fashion. *Nucleic Acids Research* 37(13): 4264-4273.

Ponferrada-Marín, M.I., Martínez-Macías, M.I., Morales-Ruiz, T., Roldán-Arjona, T., and Ariza, R.R. (2010) Methylation-independent DNA binding modulates specificity of Repressor of Silencing 1 (ROS1) and facilitates demethylation in long substrates. Journal of Biological Chemistry 285(30): 23032-23039.

Ponferrada-Marín, M.I., Parilla-Doblas, J.T., Roldán-Arjona, T., and Ariza, R.R. (2011) A discontinuous DNA glycosylase domain in a family of enzymes that excise 5-methylcytosine. Nucleic Acids Research 39(4):1473-1484.

Ponferrada-Marín MI, Roldán-Arjona T, Ariza RR (2012) Demethylation initiated by ROS1 glycosylase involves random sliding along DNA. Nucleic Acids Research 40(22): 11554-11562.

Ream, T.S., Haag, J.R., Wierzbicki, A.T., Nicora, C.D., Norbeck, A.D., Zhu, J.K., Hagen, G., Guilfoyle, T.J., Pasa-Tolić, L., and Pikaard, C.S. (2008) Subunit composition of the RNA-silencing enzymes Pol IV and Pol V reveal their origins as specialized forms of RNA polymerase II. Molecular Cell 33: 192-203. doi: 10.1016/j.molcel.2008.12.015.

Rigal, M., Kevei, Z., Pélissier, T., and Mathieu, O. (2012) DNA methylation in an intron of the IBM1 histone demethylase gene stabilizes chromatin modification patterns. EMBO J 31: 2981-2993. https://doi.org/10.1038/emboj.2012.141.

Rong, H., Yang, W., Zhu, H., Jiang, B., Jiang, J., and Wang, Y. (2021) Genomic imprinted genes in reciprocal hybrid endosperm of Brassica napus. BMC Plant Biology 21(1): 140. https://doi.org/10.1186/s12870-021-02908-8

Rymen, B., Ferrafiat, L., and Blevins, T. (2020) Non-coding RNA polymerases that silence transposable elements and reprogram gene expression in plants. Transcription 11(3-4): 172-191. https://doi.org/10.1080/21541264.2020.1825906

Saze, H., Scheid, O.M., and Paszkowski, J. (2003) Maintenance of CpG methylation is essential for epigenetic inheritance during plant gametogenesis. Nat Genet. 34(1): 65-69. doi: 10.1038/ng1138.

Schmid, M., Davison, T. S., Henz, S. R., Pape, U. J., Demar, M., Vingron, M., Schölkopf, B., Weigel, D., and Lohmann, J. U. (2005). A gene expression map of Arabidopsis thaliana development. Nature genetics, 37(5): 501–506. https://doi.org/10.1038/ng1543.

Schmitz, R. J., Schultz, M. D., Lewsey, M. G., O'Malley, R. C., Urich, M. A., Libiger, O., Schork, N. J., and Ecker, J. R. (2011). Transgenerational epigenetic instability is a source of novel methylation variants. Science 334(6054): 369–373. https://doi.org/10.1126/science.1212959.

Schumann, U., Lee, J.M., Smith, N.A., Zhong, C., Zhu, JK., Dennis, E.S., Millar, A.A., and Wang, MB. (2019) *DEMETER* plays a role in DNA demethylation and disease response in somatic tissues of Arabidopsis. Epigenetics 14(11): 1074-1087. doi: 10.1080/15592294.2019.1631113.

Seethy, A., Pethusamy, K., Chattopadhyay, I., Sah, R., Chopra, A., Dhar, R., and Karmakar, S. (2021) TETology: epigenetic mastermind in action. Applied Biochemistry and Biotechnology 193: 1701-1726.

Slotkin, R.K., and Martienssen, R. (2007) Transposable elements and the epigenetic regulation of the genome. Nature Reviews Genetics 8: 272-285. https://doi.org/10.1038/nrg2072.

Slotkin, R.K., Vaughn, M., Borges, F., Tanurdžić, M., Becker, J.D., Feijó, J.A., and Martienssen, R.A. (2009) Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. Cell 136(3): 461-472. https://doi.org/10.1016/j.cell.2008.12.038

Srikant, T., and Wibowo, A.T. (2021) The underlying nature of epigenetic variation: origin, establishment, and regulatory function of plant epialleles. Int J Mol Sci. 22(16): 8618. doi: 10.3390/ijms22168618.

Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L. M., Liu, D. R., Aravind, L., and Rao, A. (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science 324(5929): 930–935. doi:10.1126/science.1170116.

Takuno, S., and Gaut, B.S. (2013) Gene body methylation is conserved between plant orthologs and is of evolutionary consequence. PNAS 110(5): 1797-1802. https://doi.org/10.1073/pnas.1215380110.

Tang, K., Lang, Z., Zhang, H., and Zhu, JK. (2016) The DNA demethylase ROS1 targets genomic regions with distinct chromatin modifications. Nature Plants 2: 16169. doi: 10.1038/nplants.2016.169

Wang, Y. H., and Warren, J. T., Jr (2010). Mutations in retrotransposon AtCOPIA4 compromises resistance to *Hyaloperonospora parasitica* in *Arabidopsis thaliana*. Genet Mol Biol. 33(1): 135–140. https://doi.org/10.1590/S1415-47572009005000099

Wang, H., Lou, D., and Wang, Z. (2019) Crosstalk of genetic variants, allele-specific DNA methylation, and environmental factors for complex disease risk. Front Genet. 9: 695. doi: 10.3389/fgene.2018.00695.

Weissman, S., Alpermann, T., Grossmann, V., Kowarsch, A., Nadarajah, N., Eder, C., Dicker, F., Fasan, A., Haferlach, C., Haferlach, T., Kern, W., Schnittger, S., and Kohlmann, A. (2012) Landscape of TET2 mutations in acute myeloid leukemia. Leukemia 26(5): 934-942. doi: 10.1038/leu.2011.326.

Weng, ML., Becker, C., Hildebrandt, J., Neumann, M., Rutter, M.T., Shaw, R.G., Weigel, D., and Fenster, C.B. (2019) Fine-grained analysis of spontaneous mutation spectrum and frequency in *Arabidopsis thaliana*. Genetics 211(2): 703-714. https://doi.org/10.1534/genetics.118.301721.

Wierzbicki, A.T., Haag, J.R., Pikaard, C.S. (2008) Noncoding transcription by RNA Polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. Cell 135(4): 635-648. https://doi.org/10.1016/j.cell.2008.09.035.

Williams, B.P., Pignatta, D., Henikoff, S., and Gehring, M. (2015) Methylation-sensitive expression of a DNA demethylase gene serves as an epigenetic rheostat. PLoS Genet. 11(3): e1005142. doi:10.1371/journal.pgen.1005142.

Williams, B.P., and Gehring, M. (2020) Principles of epigenetic homeostasis shared between flowering plants and mammals. Trends Genet 36(10):751-763. https://doi.org/10.1016/j.tig.2020.06.019.

Williams, B.P., Bechen, L.L., Pohlmann, D.A., and Gehring, M. (2021) Somatic DNA demethylation generates tissue-specific methylation states and impacts flowering time. The Plant Cell. koab319. https://doi.org/10.1093/plcell/koab319.

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(4): 805-814. doi:10.1105/tpc.105.038836.

Xie, Z., Johansen, J.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 Biology 2(5): e104. https://doi.org/10.1371/journal.pbio.0020104

Xu, P., Chen, H., Jin, J., and Cai, W. (2018) Single-base resolution methylome analysis shows epigenetic changes in *Arabidopsis* seedlings exposed to microgravity spaceflight conditions on board the SJ-10 recoverable satellite. NPJ Microgravity 4: 12. https://doi.org/10.1038/s41526-018-0046-z

Yamamuro, C., Miki, D., Zheng, Z., Ma, J., Wang, J., Yang, Z., Dong, J., and Zhu, JK. (2014) Overproduction of stomatal lineage cells in *Arabidopsis* mutants defective in active DNA demethylation. Nature Communications 5:4062. DOI: 10.1038/ncomms5062.

Yu, A., Lepère, G., Jay, F., Wang, J., Bapaume, L., Wang, Y., Abraham, A.L., Penterman, J., Fischer, R.L., Voinnet, O., and Navarro, L. (2012) Dynamics and biological relevance of DNA demethylation in *Arabidopsis* antibacterial defense. PNAS 110(6): 2389-2394. doi: 10.1073/pnas.1211757110.

Zeng, W., Huang, H., Lin, X., Zhu, C., Kosami, KI., Huang, C., Zhang, H., Duan, CG., Zhu, JK., and Miki, D. (2021) Roles of DEMETER in regulating DNA methylation in vegetative tissues and pathogen resistance. J Integr Plant Biol. 63(4): 691-706. doi: 10.1111/jipb.13037.

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(6): 1189–1201. https://doi.org/10.1016/j.cell.2006.08.003.

Zhang, B., Tieman, D.M., Jiao, C., Xu, Y., Chen, K., Fei, Z., Giovannoni, J.J., and Klee, H.J. (2016) Chilling-induced tomato flavor loss is associated with altered volatile synthesis and transient changes in DNA methylation. PNAS 113(44): 12580-12585. https://doi.org/10.1073/pnas.1613910113

Zhong, S., Fei, Z., Chen, Y. R., Zheng, Y., Huang, M., Vrebalov, J., McQuinn, R., Gapper, N., Liu, B., Xiang, J., Shao, Y., and Giovannoni, J. J. (2013) Single-base resolution methylomes of tomato fruit development reveal epigenome modifications associated with ripening. Nature Biotechnology 31(2): 154–159. https://doi.org/10.1038/nbt.2462

Zhu, B., Zheng, Y., Hess, D., Angliker, H., Schwarz, S., Siegmann, M., Thiry, S., and Jost, JP. (2000) 5-methylcytosine-DNA glycosylase activity is present in a cloned G/T mismatch DNA glycosylase associated with the chicken embryo DNA demethylation complex. PNAS 97(10): 5135-5139.

Zilberman, D., Cao, X., and Jacobsen, S.E. (2003). *ARGONAUTE4* control of locus-specific siRNA accumulation and DNA and histone methylation. Science 299(5607): 716-719. doi: 10.1126/science.1079695.

Zilberman, D., Gehring, M., Tran, R., Ballinger, T., and Henikoff, S. (2007) Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. Nature Genetics 39(1): 61-69. doi:10.1038/ng1929.

## Chapter Two

## A DNA demethylase in Arabidopsis is transcriptionally regulated in *cis* by a methylation-sensitive silencer sequence

GUS transgenes were designed and generated by B. Williams. ROS1:LUC BAC transgenes were designed, generated, and analyzed by B. Williams. C. Picard designed the ChIP-Seq analysis pipeline. Jing-Ke Weng's lab provided the pU6 and Cas9 constructs. All other experiments were designed and performed by D. Pohlmann.

## ABSTRACT

DNA methylation must be properly regulated for the development of the plant and to respond to environmental perturbations. *ROS1* is the most highly expressed DNA demethylase gene in the somatic tissues of Arabidopsis, and has significant roles in plant development and defense. Prior work has established that *ROS1* expression is regulated atypically by promoter methylation, but the mechanism by which this occurs is unknown. In this research, I show that methylation-sensitive regulation of *ROS1* occurs at a transcriptional level and requires more than 4kb of upstream chromosomal context. I utilize CRISPR-Cas9 to generate several deletions in the endogenous chromosome surrounding *ROS1* and investigate the role of each region in regulating *ROS1* transcription. One deletion resulted in the identification of a methylation-sensitive silencer sequence upstream of *ROS1* that represses *ROS1* transcription but is less repressive as it becomes more methylated. This is an insight into a new mechanism of gene regulation in *Arabidopsis* by DNA methylation.

## INTRODUCTION

Plants must be able to alter their gene expression throughout development and in response to environmental changes. One mechanism by which this is accomplished is DNA methylation, which in plants can occur on any cytosine regardless of sequence context. DNA methylation is an epigenetic mark that is stable over generations, yet labile during development and in various tissues of the plant (Gehring, 2019). Regulation of DNA methylation plays a significant role in pathogen defense (Yu et al., 2012; Le et al., 2014; Zhu et al., 2016; Halter et al., 2021), imprinting (Gehring et al., 2009; Gehring et al., 2011; Hsieh et al., 2011; Park et al., 2016, Gehring and Satyaki, 2017; Pignatta et al., 2018), development (Yamamuro et al., 2014; Lin et al., 2020; Liang et al., 2021) and numerous other pathways. This requires a complex interplay of pathways adding methylation, primarily the RNA-directed DNA Methylation (RdDM) pathway (Matzke and Mosher, 2014), and demethylases removing methylation.

*REPRESSOR OF SILENCING1 (ROS1*) was identified as a DNA glycosylase/lyase that removes DNA methylation through base-excision repair in the somatic tissue of Arabidopsis (Gong et al., 2002; Agius et al., 2006, Morales-Ruiz et al., 2006). Expression of the *ROS1* demethylase was found to be reduced by roughly an order of magnitude in mutants of the RdDM pathway, in which the *ROS1* promoter is hypomethylated (Huettel et al., 2006, Williams et al.,

2015; Lei et al., 2015), suggesting that *ROS1* is itself regulated by DNA methylation. Indeed, it was found that a negative feedback loop exists at *ROS1*: methylation of a 228bp 'rheostat' region approximately 1kb upstream of the transcriptional start site was determined to be both necessary and sufficient to induce *ROS1* expression in an RdDM mutant (Williams et al., 2015; Lei et al., 2015). The introduction into *rdr2* of transgenes that produce double-stranded hairpin RNAs matching this rheostat region bypassed the *rdr2* mutation to ectopically methylate this site; this was causally linked to an increase in *ROS1* transcription (Williams et al., 2015). This resulted in 'Broken Rheostat' plants, in which *ROS1* is expressed at wildtype levels despite their hypomethylated *rdr2* background.

Small RNAs that target *de novo* methylation by RdDM localize to the rheostat region (Pignatta et al., 2014), and this region loses non-CG methylation in RdDM mutants (Williams et al., 2015), indicating that RdDM adds the methylation to induce *ROS1* expression. The ROS1 protein also targets this region for active demethylation (Córdoba-Cañero et al., 2017), thus downregulating its own transcription and modulating *ROS1* expression to maintain epigenetic homeostasis (Williams et al., 2015). Given that methylation of the proximal-promoter region typically results in gene silencing rather than induction, this atypical response to promoter methylation is intriguing.

In this chapter, I investigate the mechanisms by which *ROS1* is regulated. Analysis of transgenes containing *ROS1* promoter sequence is complicated by the fact that endogenous ROS1 protein not only targets its own promoter (Córdoba-Cañero et al., 2017), but also transgenes (Gong et al., 2002). Despite this complication, previous works have still used *ROS1* proximal promoter sequence to driver reporter genes in order to determine *ROS1* expression patterns (Bennett et al., 2021). I show that this does not encompass the full regulation of *ROS1*, as transgenic lines with identical transgenes respond different to being introduced into hypomethylated backgrounds, implying that the greater chromosomal context affects regulation of transcription of *ROS1.* In concordance with that, I use CRISPR-Cas9 technology to create deletions in the endogenous sequence around *ROS1* ranging from 17kb upstream of the gene to 40kb downstream to investigate the role of *cis* regulation on *ROS1* transcription. This has resulted in the identification of several regions that impact *ROS1* transcription.

## RESULTS

### *ROS1* is regulated in methylation mutants at a transcriptional level

*ROS1* has lower steady-state transcript levels in methylation mutants, and methylation of the promoter region is sufficient to upregulate transcript levels (Williams et al., 2015). However, it is unknown whether reduced *ROS1* transcript accumulation reflects transcriptional or post-transcriptional processes. The reduction of transcript levels in methylation mutants could be due to post-transcriptional regulation of the mRNA. In order to determine if lower transcript levels in methylation mutants is due to transcriptional or post-transcriptional regulation, I conducted both ChIP-qPCR and ChIP-Seq for elongating RNA Polymerase II in wildtype, *rdr2* hypomethylated mutants, and *rdr2* mutants with artificially-induced methylation of the *ROS1* promoter, described in Williams et al. (2015) and referred to here as 'Broken Rheostat' (BR) lines. Elongating RNA Pol II with phosphorylated serine 2 accumulates at the transcription termination site (TTS) to allow RNA processing to occur (Kuehner et al., 2011). The ChIP-Seq data showed an increase in Pol II occupancy at the TTS of *ROS1* correlating with the methylation status of the promoter; BR lines had more occupancy than wildtype, which was itself increased over *rdr2* (Fig. 1A-B). ChIP-qPCR for the *ROS1* transcriptional start site and exon 13 also showed decreased Pol II occupancy in *rdr2* compared to wildtype (Fig. 1 C-D). This indicates that the methylation-sensitive expression of *ROS1* reflects transcriptional regulation, rather than post-transcriptional.

### 4kb of *ROS1* promoter is not sufficient for methylation-sensitive expression

To determine if the methylation-sensitive regulation of *ROS1* was linked to a particular developmental time or tissue, I used reporter transgenes to visualize expression from the *ROS1* promoter. A *ROS1*:*LUCIFERASE* fusion was designed and recombineered (Sharan et al., 2009; Zhou et al., 2011) into a bacterial artificial chromosome (BAC) containing 65kb of endogenous Chromosome 2 sequence, including 17kb of sequence upstream of *ROS1* and 41kb of sequence downstream of *ROS1* (Fig. 2A). This transgene was introduced into Col-0 wildtype plants by agrobacterium-mediated floral dipping (Clough and Bent, 1998), then crossed with *rdr2*. *RDR2*$^{+/+}$ (not shown), *RDR2*$^{+/-}$ (Fig. 2B), and *rdr2*$^{-/-}$ (Fig. 2C) plants with the transgene were all evaluated for expression of the luciferase reporter by spraying with the luciferase substrate luciferin and quantifying the light produced from the luciferase reaction by using a highly-sensitive CCD camera. *rdr2* mutants plants produced much less light than the wildtype or heterozygous plants (Fig. 2D), indicating that the transgenic *ROS1*:*LUC* was expressed at much

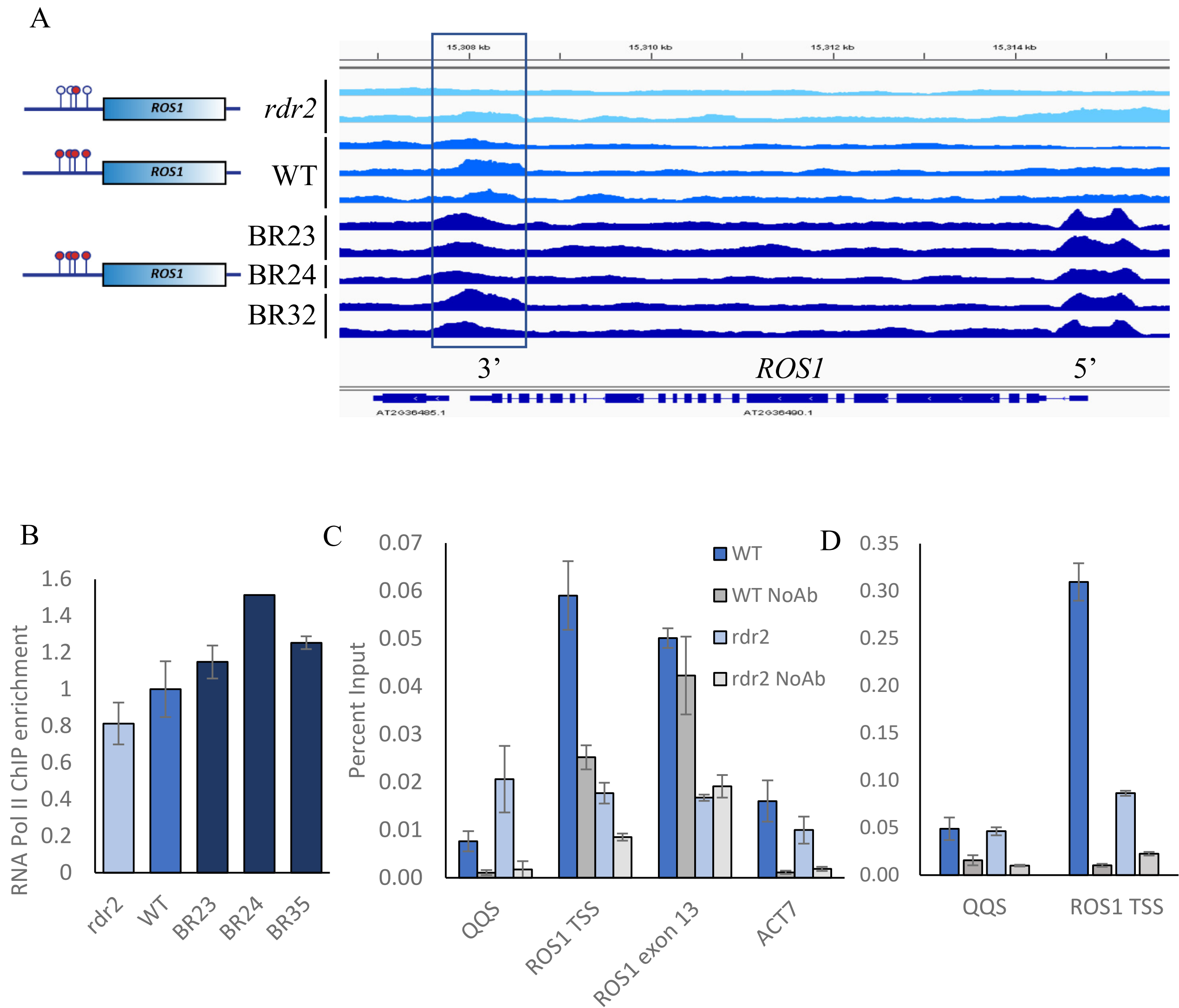


**Figure 1: *ROS1* is transcriptionally silenced in the methylation mutant *rdr2***
A) Browser snapshot of RNA Pol II ChIP-Seq over *ROS1,* with scale of 0-3 set for all tracks. Tracks are colored in rough approximation of *ROS1* promoter methylation level. Peaks in BR lines at the 5' end represent the inverted repeat transgene that matches and drives methylation in this region. B) Average RNA Pol II ChIP enrichment at the *ROS1* TTS ± 500bp, the region boxed in A). Colors are approximations of *ROS1* promoter methylation level. Error bars are standard error of the biological replicates. C-D) Two ChIP-qPCR experiments for Elongating RNA Pol II in both Col (WT) and *rdr2*. *QQS* is upregulated in *rdr2* plants (Kurihara et al., 2008), and *ACT7* serves as a control. NoAb refers to a no-antibody control. Legend in C) also applies to D).

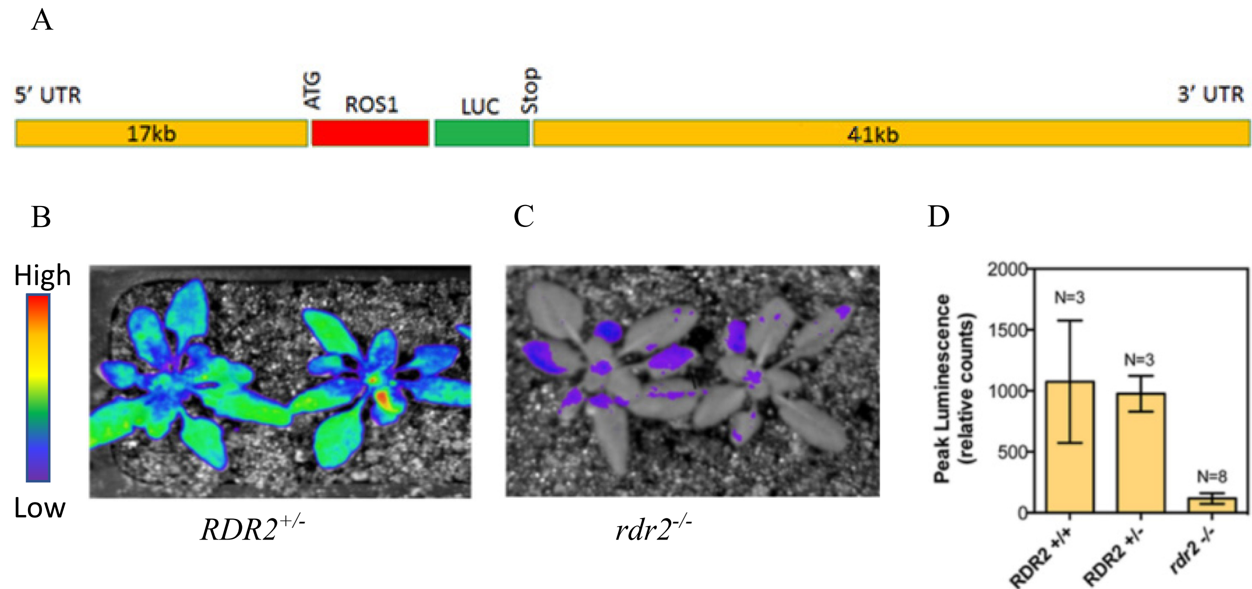


**Figure 2: A transgene containing 65kb of endogenous sequence including *ROS1* is sufficient for methylation-sensitive expression of *ROS1***
A) Schematic of a bacterial artificial transgene containing endogenous sequence from Arabidopsis chromosome 2 and a recombineered *ROS1*:*LUCIFERASE* fusion gene. This transgene contains 17kb of sequence endogenously upstream of the *ROS1* transcriptional start site and 41kb of downstream sequence. B-C) Relative luminescence detected from 3-week old transgenic plants containing the *ROS1:LUC* BAC in both *RDR2*$^{+/-}$ (B) and *rdr2*$^{-/-}$ (C) plants. D) Quantification of maximum luminescence from plants containing the *ROS1*:*LUC* transgene with and without mutation of *RDR2*.

lower levels in the methylation mutant background. The large transgene therefore mimics the endogenous *ROS1* locus with respect to its methylation-sensitive expression.

Smaller transgenes were then assessed. 4kb of sequence upstream of *ROS1* was used to drive expression of the reporter GUS (Fig. 3A-T, Y-Z) and, in a different transgene, a ROS1-GUS fusion protein (Fig. 3U-X). These transgenes were transformed into Col-0 wildtype plants by floral dipping (Clough and Bent, 1998), then crossed into an *rdr2* background so that the same transgene could be assessed in wildtype and *rdr2* siblings descended from the $F_1$. Multiple lines of transgenic plants were followed. While some lines showed a reduction of GUS in an *rdr2* background (compare Fig. 3A-E to F-J), other transgenic lines showed approximately equivalent expression in *rdr2* as in wildtype (compare Fig. 3K-O to P-T). This indicates that the 4kb of upstream sequence is not sufficient to guarantee methylation-sensitive expression or downregulation of ROS1.

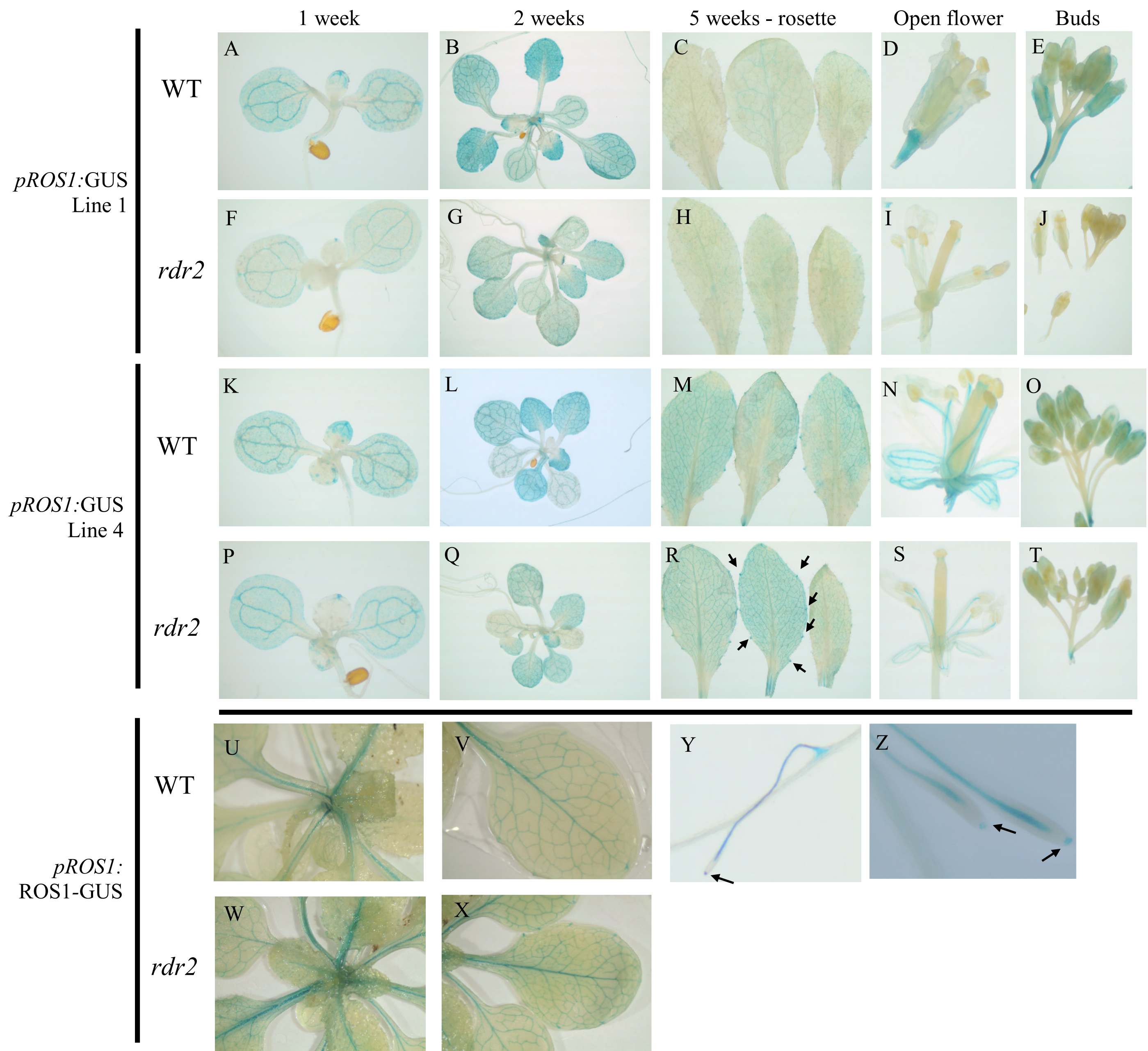


**Figure 3: GUS staining of WT and *rdr2* siblings with *pROS1*:GUS and *pROS1*:ROS1-GUS transgenes in different tissues**

A-T) GUS staining of plants with a GUS transgene driven by 4kb of *ROS1* promoter, in 1 week old seedlings, two week old seedlings, mature rosette leaves at 5 weeks, open flowers, and primary inflorescences (labeled 'Buds'). Arrows in R) point to hydathodes. U-X) GUS staining of plants with a ROS1-GUS fusion protein driven by 4kb of *ROS1* promoter. U, V, W, and X were all collected at 3 weeks and show the vegetative apex (U and W) and rosette leaves (V and X). All plants within a line contain the same single insertion of the transgene. Y-Z) Lateral roots from *pROS1*:GUS Line 4 (Y) and Line 9 (Z) (magnified), with arrows indicating the columella.

Transgenic plants with the same 4kb promoter driving expression of a ROS1-GUS fusion protein also failed to show consistent downregulation of the reporter in an *rdr2* background (Fig. 3U-X). This is further evidence that the regulation is transcriptional in nature, and is not dependent on the sequence of the mRNA or the protein itself.

Comparison of the large transgene (Fig. 2), which is regulated like the endogenous locus in methylation mutant backgrounds, to the transgenes with 4kb of promoter sequence, lines of which are not consistently downregulated in *rdr2* mutants, indicates that sequences outside of the 4kb present in the small transgene influence how *ROS1* is regulated by methylation.

**Regions 3' of *ROS1* do not have strong effects on *ROS1* transcription**

To identify *cis*-elements that regulate *ROS1* in response to global methylation levels, I used CRISPR-Cas9 to create deletions in the regions of the endogenous chromosome contained within the aforementioned BAC transgene, and assayed their effects on *ROS1* transcription. Each deletion line was crossed into two separate RdDM mutants, *drm2* and *rdr2*, and heterozygotes were selfed. The next generation was genotyped to identify siblings that were homozygous with and without the deletion, in plants that were homozygous for the RdDM mutation or wildtype allele (Fig. 4). qRT-PCR was used to assay *ROS1* expression in these siblings. Because the BAC

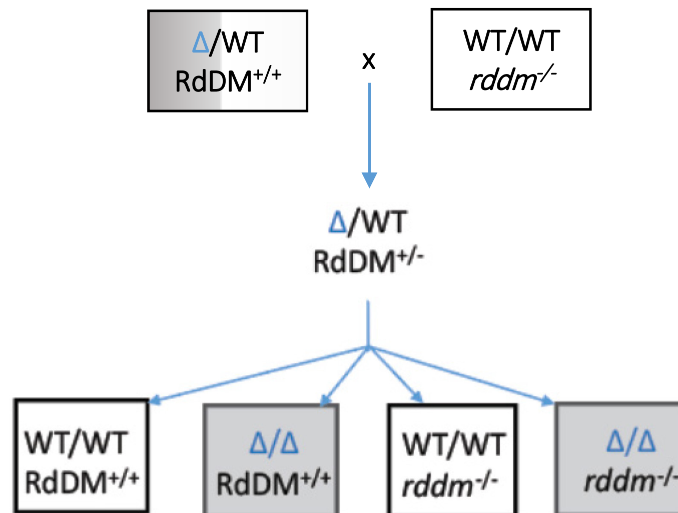


**Figure 4: Crossing scheme for generation of $F_2$ plants used in analyses**
Deletion mutants identified in $T_2$ or later generations were crossed to *drm2* or *rdr2* (mutants in the RdDM pathway, collectively *rddm*$^{-/-}$ here). Heterozygotes of both the deletion and an RdDM mutation were selfed to generate siblings that were homozygous for the deletion and/or the RdDM mutation, or neither, for later analysis.

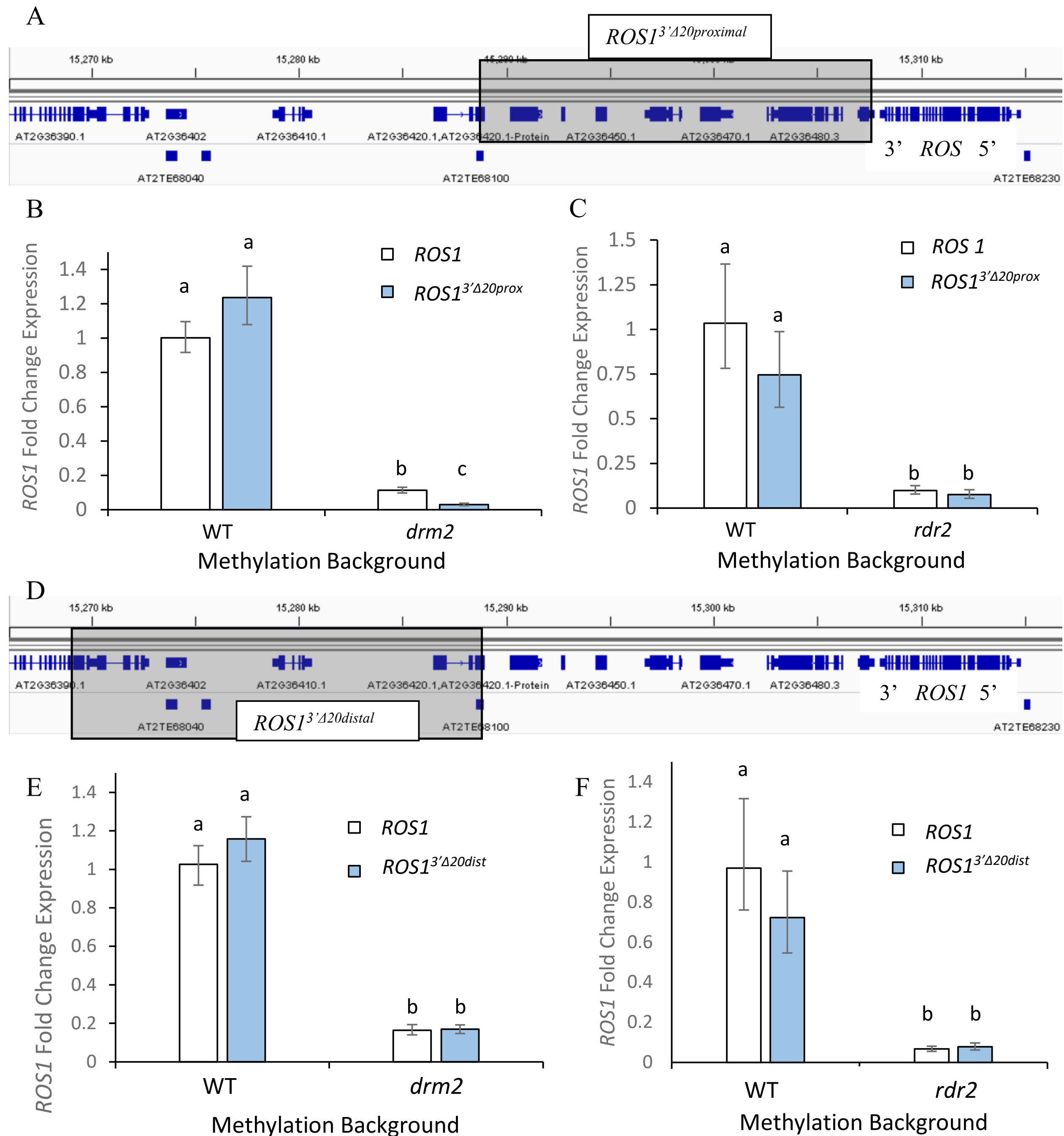


**Figure 5: Deletions 3' of *ROS1* have little effect on *ROS1* transcription**
A) Schematic of a 19.5kb deletion from 7.5kb after the *ROS1* TSS to +28kb. B-C) qPCR results showing fold change of *ROS1* in plants with and without $ROS1^{3'\Delta 20proximal}$ in both wildtype and *drm2* (B) or *rdr2* (C) backgrounds. In B), n=2 for wildtype, n=5 for *drm2* no deletion, n=6 for *drm2* deletion. In C) n=3 for *rdr2* samples, n=4 for WT samples. D) Schematic of a 20kb deletion from +28kb to +48kb. E-F) qPCR results showing fold change of *ROS1* transcripts in plants with and without $ROS1^{3'\Delta 20distal}$ in both wildtype and *drm2* (E) or *rdr2* (F) background. n=2 for *drm2* no deletion in (E) and wildtype no deletion in (F), n=4 for all other genotypes. Error bars represent standard error. Different letters indicate $p<0.01$ by ANOVA.

transgene showed downregulation of *ROS1* in a methylation mutant background, I used this region to set the boundaries for my search for regulatory elements.

To investigate the 40kb of sequence downstream of *ROS1,* given that many *cis*-regulatory elements have been identified downstream of genes in Arabidopsis (Wang and Chekanova, 2019)*,* two adjacent deletions of approximately 20kb each were generated 3' of *ROS1*.

*ROS1*$^{3'\Delta 20proximal}$ removed 20.3kb beginning 551bp after the TTS (Fig. 5A). This deletion did not affect transcription of *ROS1* in wildtype or *rdr2* backgrounds (Fig. 5B-C). However, in a *drm2* background the deletion resulted in significantly further downregulation of *ROS1*, from 9-fold below wildtype levels to 35-fold below levels in wildtype plants without the deletion (43-fold below plants with the deletion in a wildtype background) (ANOVA ($F(3,11) = 58.83$, $p < 0.0001$); Tukey HSD ($p < 0.01$)) (Fig. 5B). This could indicate that a *cis*-regulatory region exists in this region that helps promote *ROS1* expression under low-methylation conditions. However, the fact that this downregulation was not seen in the *rdr2* background is interesting, as *drm2* and *rdr2* are in the same pathway and both result in a hypomethylated state for *ROS1*. DRM2 acts downstream of RDR2, so it is possible that the mechanism for the *drm2*-mediated downregulation may require an RdDM pathway member that acts after RDR2, such as the binding and transcription of RNA Polymerase V.

The second 3' deletion, *ROS1*$^{3'\Delta 20distal}$, begins where 3'Δ20proximal ends, and removes 20.4kb until +48kb (Fig. 5D). This deletion resulted in no transcriptional changes in the *drm2*, *rdr2*, or wildtype backgrounds (Fig. 5E-F). While it is still possible that these regions contain regulatory elements that affect *ROS1* expression in particular tissue types or under particular stress conditions, there is little evidence that they affect *ROS1* expression in response to methylation.

**The region from -17kb to -2kb of *ROS1* may contain a regulatory element necessary in *cis* for *ROS1* transcription to occur**

The deletion of a 15kb region of the endogenous chromosome from -17kb to -2kb from the *ROS1* transcription start site was accomplished with CRISPR technology (Fig. 6A). This deletion was homozygous lethal; 91 descendants from a selfed heterozygote were genotyped and 36 wildtype and 55 plants heterozygous for the deletion were recovered, consistent with Mendelian inheritance if the homozygous deletions were lethal ($X^2$ $(1, N = 91) = 0.86$, $p=0.35$). This lethality may derive from the deletion of *LOS2*/*ENOLASE 2*, contained within this region,

as knockouts of this gene from T-DNA insertions have severe developmental phenotypes and are sterile, with an embryo-lethal phenotype (Eremina et al., 2015).

*ROS1*$^{5'\Delta15}$ heterozygotes did not have any change in *ROS1* expression in the RdDM mutant, but did show a significant, 1.7-fold increase in transcript in a wildtype background

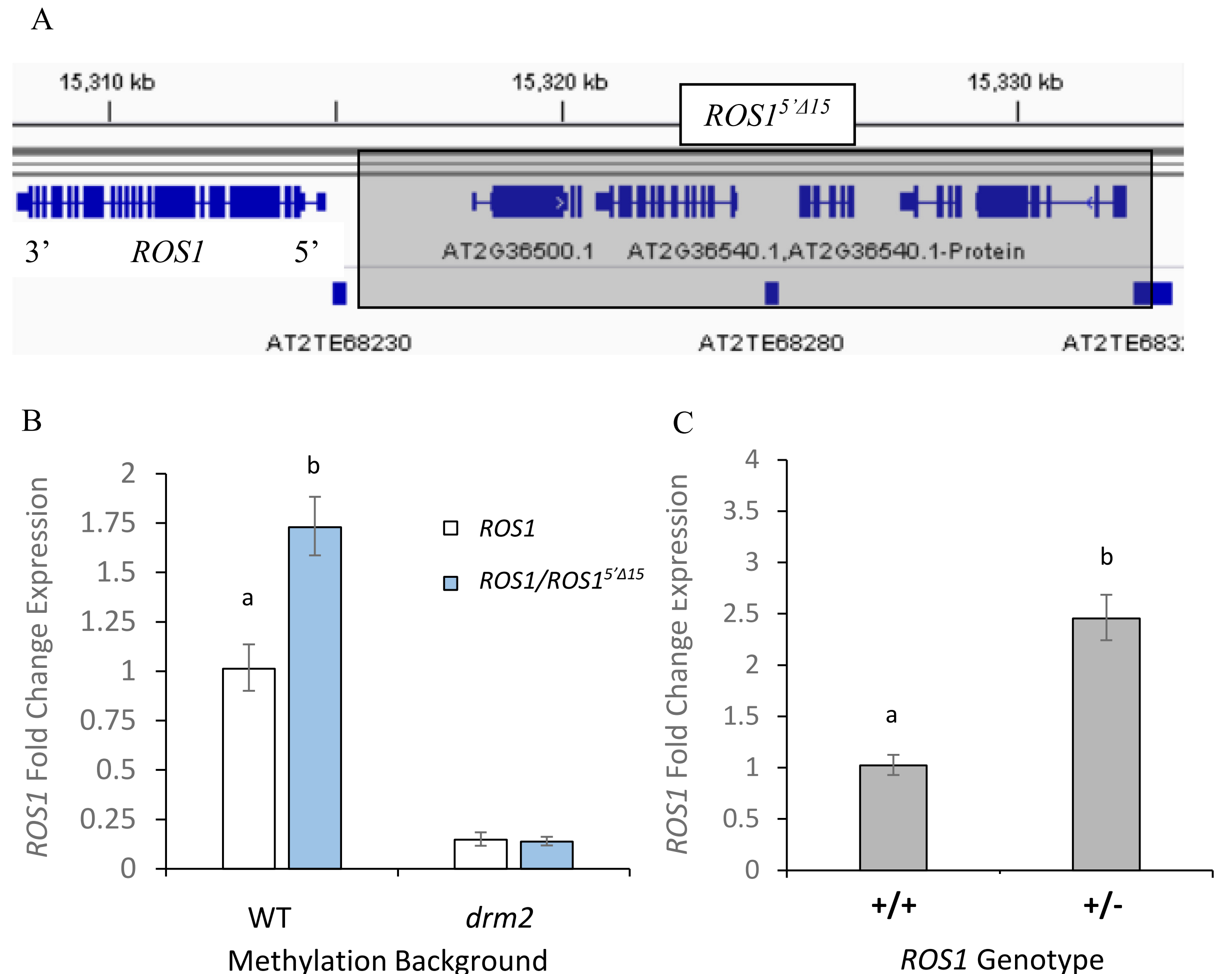


**Figure 6: A 15kb deletion upstream of *ROS1* removes a methylation-dependent repressor of *ROS1* transcription**
A) Schematic detailing the region of Chromosome 2 from the *ROS1* transcription termination site to 20kb upstream, including the *ROS1*$^{5'\Delta15}$ deletion from -17kb to -2kb. Genes and TEs are indicated on separate tracks. B) qPCR results showing fold-change of *ROS1* in wildtype and *drm2 ROS1/ROS1*$^{5'\Delta15}$ heterozygotes versus *ROS1/ROS1* homozygotes. C) qPCR results showing fold-change of *ROS1* in *ROS1/ros1-7* heterozygous plants. n=4 for all genotypes. Error bars indicate standard error of the mean. Different letters indicate $p<0.01$ by Student's T-test.

compared to plants without the deletion (t(6) = 4.69, p = 0.003) (Fig. 6B). While this effect is small, it should be noted that the deletion is not homozygous. A similar increase in transcription was detected in heterozygotes for the point-mutant *ros1-7* (Fig. 6C), which results in a catalytically non-functional enzyme due to an amino acid substitution in the catalytic domain. *ROS1* expression with only one allele producing functional product therefore results in increased transcription. This is consistent with the rheostat feedback model, in that decreases in the amount of functional *ROS1* protein would lead to less demethylation at the *ROS1* locus itself, and the resulting hypermethylation in the proximal promoter of *ROS1* leads to increased *ROS1* transcription (Williams et al., 2015).

**An 817bp region 5' of *ROS1* is a methylation-sensitive repressor of *ROS1* transcription**

I also generated an 817bp deletion from -930bp to -113bp 5' of the *ROS1* transcription start site, designated $ROS1^{5'\Delta 0.8}$ (Fig. 7A). This deletion does not overlap with $ROS1^{5'\Delta 15}$, but does remove almost the entirety of the region described as sufficient to promote *ROS1* expression when methylated in an *rdr2* mutant (Williams et al., 2015). Plants homozygous for this deletion were viable with no visible phenotypes.

The $ROS1^{5'\Delta 0.8}$ deletion resulted in a significant 2 to 2.5-fold increase in levels of *ROS1* in a wildtype background (*drm2:* ANOVA (F(3,9) = 133.33, p < 0.0001), Tukey HSD (p < 0.01); *rdr2*: ANOVA (F(3,11) = 430.35, p < 0.0001), Tukey HSD (p<0.01)) (Fig. 7B-C), indicating that this region functions as a repressor of *ROS1* expression. Despite *ROS1* expression being ten-fold reduced in *drm2* and *rdr2* plants without the deletion, the presence of this deletion in these backgrounds led to *ROS1* levels being elevated even slightly above wildtype levels (Fig. 7B-C). The downregulation of *ROS1* normally seen in methylation mutants is induced by the hypomethylation of this region. However, in the absence of this region, *ROS1* transcription is actually increased.

Increased transcription of *ROS1* does not necessarily indicate increased levels of functional ROS1 protein. *ROS1* mRNA could be post-transcriptionally regulated or produce non-functional products, or there may be additional regulatory measures to maintain ROS1 protein levels at a steady-state. In order to test if the increased levels of *ROS1* transcript correlated with ROS1 function, I assayed expression of *SDC*, a gene that is silenced by methylation in the somatic tissue of the plant (Henderson and Jacobsen, 2008), but also targeted by ROS1 (Williams et al., 2017). In $ROS1^{5'\Delta 0.8}$;*rdr2* double mutants *SDC* expression is significantly

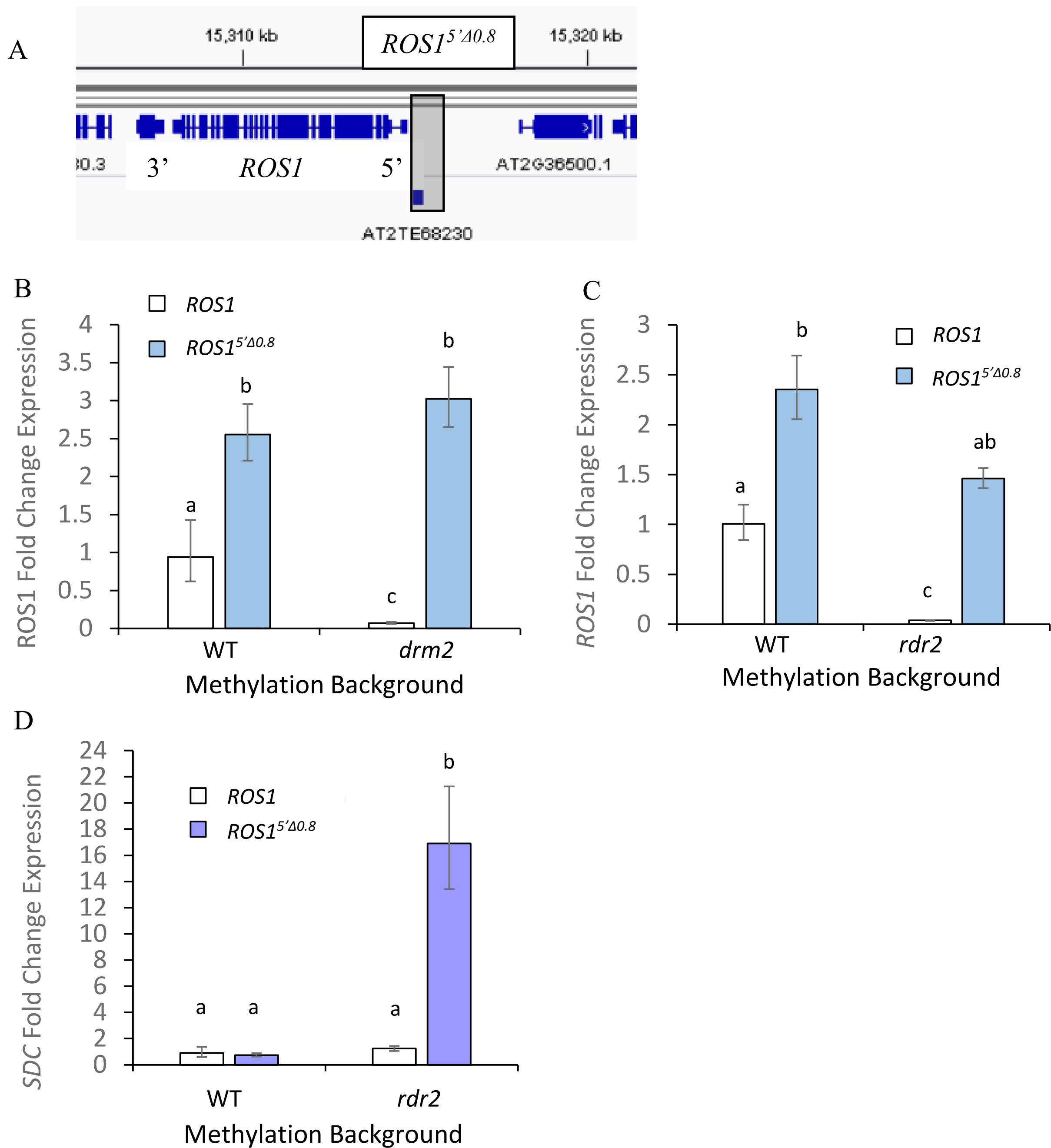


**Figure 7: An 817bp deletion removes a repressor of *ROS1* transcription in methylation-typical and methylation-deficient backgrounds**

A) Gene model of *ROS1* showing the region deleted from -930 to -113bp. B) qPCR results showing fold change of *ROS1* transcription in plants with and without $ROS1^{5'\Delta0.8}$ in both wildtype and *drm2* backgrounds. n=2 for wildtype, no deletion; n=4 for all other genotypes. C-D) qPCR results showing fold change of *ROS1* (C) and *SDC* (D) transcription in plants with and without $ROS1^{5'\Delta0.8}$ in both wildtype and *rdr2* backgrounds. n=3 for wildtype, with deletion; n=4 for all other genotypes. Error bars represent standard error of the mean. Different letters indicate $p<0.01$ by ANOVA.

increased approximately 16-fold over wildtype or *rdr2* levels (ANOVA ($F(3,10) = 26.24$, $p < 0.001$); Tukey HSD ($p < 0.01$)) (Fig. 7D). This indicates that the increased *ROS1* transcription does correlate with increased ROS1 functional activity.

## DISCUSSION

### ***ROS1* is transcriptionally regulated by methylation in its proximal promoter region**

I showed by both ChIP-qPCR and ChIP-Seq that there is a reduction of RNA Polymerase II occupancy at the TTS of *ROS1* in *rdr2*, and an associated increase in occupancy when the methylation upstream of *ROS1* is restored in the BR lines (Fig. 1). This indicates that the mechanism of regulation of *ROS1* in methylation mutants is indeed occurring at the level of transcription, rather than at the mRNA level post-transcriptionally. This conclusion was also upheld by the reporter transgene experiment, in which a ROS1-GUS fusion protein driven by 4kb of endogenous upstream *ROS1* sequence failed to recapitulate the downregulation of *ROS1* in *rdr2*.

Furthermore, even on the transcriptional level the regulation of *ROS1* is not limited to the proximal promoter. GUS expression driven by the 4kb upstream region was variable between lines (Fig. 3). Some transgene insertions were downregulated in *rdr2* mutants (Fig. 3 F-J), while other insertions were expressed near wildtype levels (Fig. 3 P-T), therefore the wider chromosomal context matters for the expression of this transgene. All of the elements necessary for epigenetic or transcriptional regulation of *ROS1* must not be present within the 4kb, or this would recapitulate the endogenous regulation.

This experiment also highlighted several expression patterns of ROS1 that may prove of future interest: GUS expression was high in the hydathodes of the rosette leaves (Fig. 3R, arrows) and in the columella of the lateral root tip (Fig. 3Y-Z). The hydathodes are specialized organs that extrude water from the leaf, and are sources of entry for various pathogens (Cerutti et al., 2019). High levels of *ROS1* expression in this structure is consistent with findings that *ROS1* demethylates several defense genes upon detection of a pathogen (Yu et al., 2012; Halter et al., 2021). In addition, these structures are composed of the ends of the xylem vasculature that transports water throughout the leaf (Yagi et al., 2021). As the lack of *ROS1* causes errors in the development of xylem tracheary elements (Lin et al., 2020), active demethylation may also be necessary for the proper development of the hydathodes and guttation.

The columella of the lateral root tip was found to be the most highly methylated cell types characterized to date, mostly due to hypermethylation in the CHH context in transposable elements, correlated with an increase in RdDM activity (Kawakatsu et al., 2016). An increase in RdDM activity likely correlates with hypermethylation in the *ROS1* promoter region, leading to increased *ROS1* transcription as well. This tissue may prove to be an interesting location to investigate the balance between RdDM and demethylation, as both pathways are at high activity. Additionally, the function of columella cell hypermethylation is still unknown. It was proposed that it may serve to help silence TEs in neighboring cells by producing mobile smRNAs for RdDM, but the concomitant expression of *ROS1* in this case would be counterproductive. Characterizing the role of ROS1 in the columella of the root tip may therefore lead to a better understanding of the role of epigenetic regulation in this cell type.

***ROS1* transcription is not significantly regulated by sequence 3' of the locus**

There appears to be minimal regulation of *ROS1* from any sequence within the 40kb downstream of *ROS1* (Fig. 5), despite a recent finding that many putative enhancer-like elements appear to interact with 3' UTRs in Arabidopsis (Wang and Chekanova, 2019). 3'Δ20proximal did have different results in *rdr2* versus *drm2*, which may indicate an interaction with the latter part of the RdDM pathway, after the role of RDR2.

**Endogenous sequence from -17kb to -2kb of the *ROS1* TSS may be necessary for *ROS1* transcription**

Deletion of 15kb of upstream sequence resulted in a homozygous lethal phenotype, making analysis of this deletion difficult. For future experiments, this deletion could potentially be complemented by a transgene expressing *ENOLASE 2*. The plants heterozygous for the deletion had increased *ROS1* transcript levels, of approximately the same degree as in *ros1-7* heterozygotes (Fig. 6B-C). This could imply that the *ROS1* allele in *cis* with the 5'Δ15 deletion is not producing any mRNA, resulting in upregulation of the wildtype allele in *trans* and the phenocopying of the heterozygous *ROS1/ros1-7.* If this is correct, this indicates that there is a sequence necessary for *ROS1* transcription more than 2kb from the transcription start site.

While this is the parsimonious explanation, it is possible that the increase in transcript level from the *ROS1/ROS1*$^{5'\Delta 15}$ heterozygote could be due to a slight increase in transcription from the deletion allele, and the phenocopying of the *ROS1/ros1-7* heterozygotes may be coincidental. To distinguish between these two possibilities in the heterozygotes, I have crossed

the 5'Δ15 heterozygotes to plants of the L*er* ecotype, which have some SNPs from the Col-0 background in the *ROS1* allele. I have also designed TaqMan primers specific for these SNPs. A TaqMan assay will allow me to determine levels of transcription from the *ROS1*$^{5'\Delta15}$ Col-0 allele versus the wildtype *ROS1* L*er* allele, using wildtype Col-0/L*er* $F_1$ hybrids as a control.

Although this is a large region to scan for potential sites of interest, there is one site close to the *ROS1* locus that deserves additional investigation. High resolution Hi-C data in Col-0 plants identified a DNA loop from within the *ROS1* gene body to a region that overlaps the 3' boundary of *ROS1*$^{5'\Delta15}$ (Liu et al., 2016) (Fig. 10, green). A targeted deletion of this looping site will indicate if this loop affects *ROS1* transcription.

**An 817bp region 5' of *ROS1* contains a methylation-sensitive silencer**

Previous research has shown that *ROS1*$^{5'\Delta0.8}$ encompasses a region that upregulates transcription of *ROS1* in direct correlation with its level of methylation (Williams et al., 2015; Lei et al., 2015). This led to the assumption that methylation was directly inducing transcription by some unknown mechanism. However, if this was true then deleting the region should have resulted in an almost-complete loss of methylation and a consequent reduction of *ROS1* transcription. However, our results from this deletion do not support this hypothesis. Instead, we found strong upregulation of *ROS1* in the absence of this region, even in RdDM mutant backgrounds where *ROS1* transcription is normally 10-fold reduced (Fig. 7). This region is therefore primarily a silencer of *ROS1*. Incorporating our prior knowledge, our new model of regulation proposes that this region becomes less repressive in direct correlation with its level of methylation (Fig. 8). This was previously hypothesized when transgenes of *ROS1* with 1.6kb of endogenous promoter were unable to silence a stably-expressed transgene in the RdDM mutant *nrpe1*, but were able to silence the transgene when the annotated Helitron TE was removed from the promoter (Lei et al., 2015). However, this was considered insufficient evidence for three reasons. First, looking at the silencing of a transgene by ROS1 expressed from a separate transgene allows for the possibility of post-transcriptional regulation of *ROS1* mRNA or protein in the methylation mutant, leading to the loss of function. Second, the insertion of transgenes into the genome by Agrobacterium is random, and insertion location can have strong effects on not only the basal transcription level of a transgene, but also the level of methylation that the

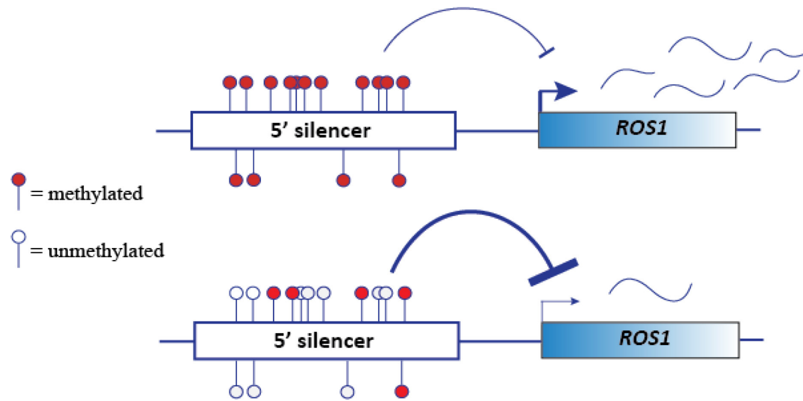


**Figure 8: Model for regulation of *ROS1* by the 5’ silencer**
The 5’ silencer region represses *ROS1* transcription, but this repressive effect is mitigated by methylation of the silencer itself. When this region is methylated (top) there is little repression and abundant transcription of *ROS1*. Demethylation of this region (bottom) leads to an increase in repression and a concomitant reduction of *ROS1* transcription.

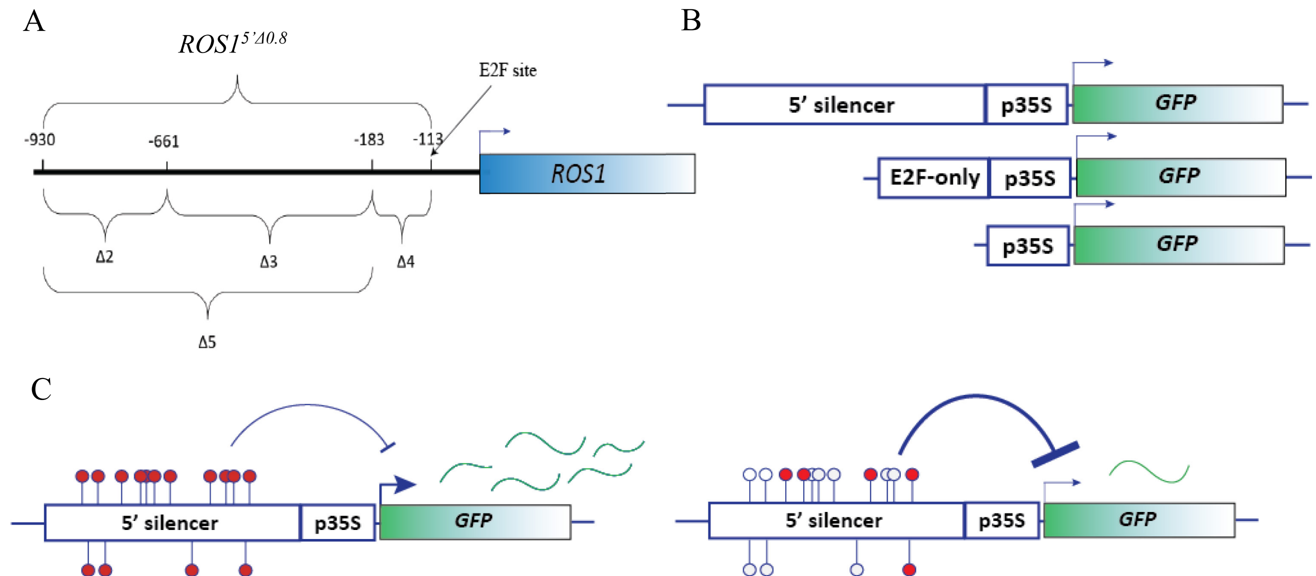


**Figure 9: Future experiments to investigate the 5’ silencer region**
A) Subdeletions of the 5’Δ0.8 region. These CRISPR deletions are currently being generated. The E2F site is from -123 to -116. B) Constructs designed to test the sufficiency of the silencer region. C) Hypothesis for silencer sufficiency: if the silencer is sufficient, then heavy methylation of the silencer (left), as in a *ros1* background, should result in higher expression of *GFP* than the same transgene in a low methylation background like *rdr2* (right).

transgene accumulates. Because of this, directly comparing different transgenes is difficult and subject to error. Third, the ROS1-mediated *ROS1* feedback loop makes the assessment of transgenes difficult, as endogenous ROS1 is able to demethylate transgenes and affect their transcription and transgenic ROS1 will also affect transcription of the endogenous locus. The deletions in the endogenous chromosome are therefore cleaner and more rigorous evidence for this hypothesis.

Previously we had considered the possibility of a methylation-sensitive chromatin loop that was affecting *ROS1* transcription. While high-resolution Hi-C data did not identify a DNA loop within the silencer region, it remains possible that hypomethylation in RdDM mutants allows a novel chromatin loop to form (Liu et al., 2016).

The silencer region (Fig. 10, orange) contains an annotated AtREP5 element, a Helitron TE. Helitrons are class II DNA transposons, and are known to be capable of affecting gene regulation. Paternally-expressed imprinted genes are enriched for neighboring Helitrons, and the epigenetic status of the Helitron typically affects the transcription of the imprinted gene (Kinoshita et al., 2004; Gehring et al., 2009; Wolff et al., 2011; Pignatta et al., 2018). Two Helitrons upstream of the defense gene *RMG1* are differentially methylated; the proximal Helitron is targeted by both RdDM and ROS1, like the *ROS1* Helitron, and active demethylation of this Helitron is required for the induction of *RMG1* in response to pathogens (Halter et al., 2021). Helitrons were also found to be enriched in binding sites for several types of transcription factors, most notably in the MADS-box and E2F families (Batista et al., 2019; Hénaff et al., 2014; Muiño et al., 2016).

The methylated region neighboring *ROS1* also contains a canonical E2F site (Fig. 10, black) (Ramirez-Parra et al., 2003), and ChIP-Seq for E2Fa revealed significant binding at this site (Verkest et al., 2014). However, upregulation of E2Fa did not result in any expression changes at *ROS1* (Vlieghe et al., 2003; Vandepoele et al., 2005). As there are six E2F transcription factors in the Arabidopsis genome, and two DP cofactor genes with which they associate (Vandepoele et al., 2002), the overexpression of E2Fa does not preclude a role for the other E2F factors, especially as they all bind the canonical site. Additionally, with the feedback loop at *ROS1*, even if E2Fa did induce increased transcription of *ROS1*, this would have led to the removal of methylation from the E2F binding site, which may have disrupted E2F binding and returned *ROS1* transcription to a homeostatic level. As E2F transcription factors can act as

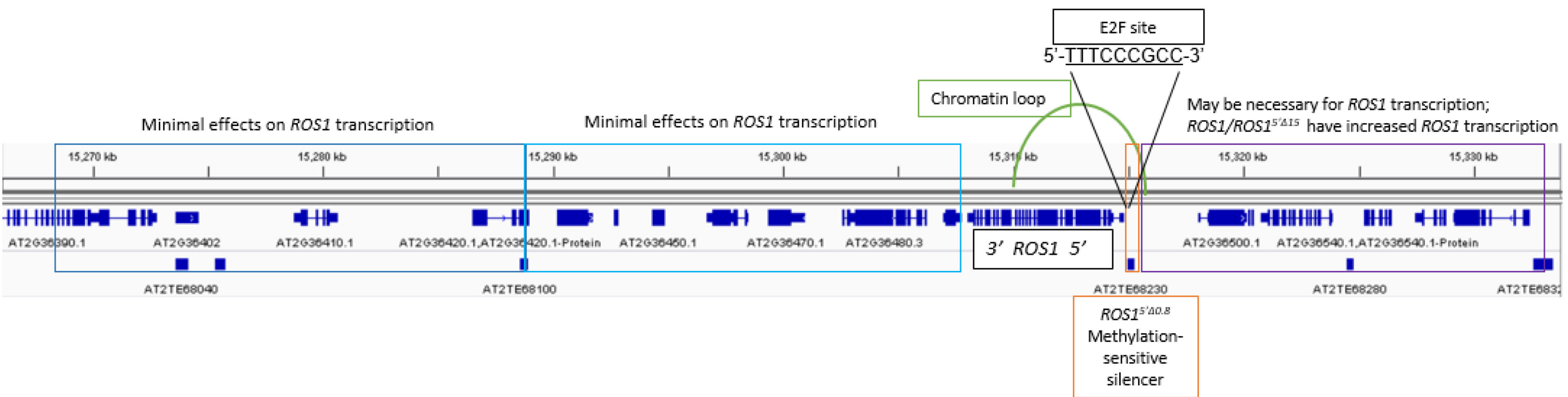


**Figure 10: Summary of *cis*-regulatory elements found in this study and of future interest**
Boxes indicate deletions made and studied in this work. From left to right, deletions are *ROS1*$^{3'\Delta 20distal}$ from Figure 4 (A-C) (dark blue), *ROS1*$^{3'\Delta 20proximal}$ from Figure 4 (D-F) (light blue), *ROS1*$^{5'\Delta 0.8}$ from Figure 6 (orange), and *ROS1*$^{5'\Delta 15}$ from Figure 5 (purple). The chromatin loop (green) is from data from Liu et al., 2016. AT2TE68230 is the Helitron mentioned in this work. The E2F site has been highlighted.

activators or repressors (Trimarchi and Lees, 2002; Ramirez-Parra and Gutierrez, 2007), bind differentially to methylated versus unmethylated DNA (Campanero et al., 2000; O'Malley et al., 2016), and are also known to regulate other factors responsible for DNA methylation (Liu et al., 2017), the hypothesis that *ROS1* transcription is being regulated by the methylation-sensitive binding of an E2F transcription factor is an attractive one.

To investigate this further, I am undertaking two additional experiments: the first will narrow down the region of DNA within the 817bp that has the silencing effect, and the second will determine whether this region is sufficient to function as a methylation-sensitive silencer outside of the endogenous chromosomal context. For the first, I am generating smaller deletions within the 817bp region (Fig. 9A). Deletions $ROS1^{\Delta4}$ and $ROS1^{\Delta5}$ divide the silencer into a 70bp region containing the E2F site ($ROS1^{\Delta4}$) and the remaining 747bp ($ROS1^{\Delta5}$). If E2F binding to this region is responsible for the silencing effect, perhaps by a repressive E2F that preferentially binds non-methylated DNA, then $ROS1^{\Delta4}$ should induce upregulation of *ROS1* while $ROS1^{\Delta5}$ will not. If $ROS1^{\Delta5}$ is found to contain the causal region, I can further narrow down the region of activity by separating the region into $ROS1^{\Delta2}$ and $ROS1^{\Delta3}$.

The second experiment interrogates the ability of the silencing region to act outside of its endogenous context. For this, I have created three constructs that use 113bp of the constitutive *35S* promoter to drive expression of GFP (Fig. 9B); this is greater than the minimal promoter for 35S, and should result in a moderate level of expression. Additionally, the 113bp of *35S* will maintain the endogenous distance of the silencer sequence from the transcriptional start site. The first construct has the full 817bp silencer region before the *35S* promoter; the second only has the 70bp containing the E2F site, as in Δ4; and the third is solely *p35S*, in order to determine the basal level of GFP expression that this construct can generate. These constructs have been transformed into a *ros1* mutant in order to encourage methylation of the transgenes. I will then cross the transgene into less methylated backgrounds, namely Col-0 wildtype and *rdr2*, to create an epiallelic series. I will then assay GFP expression to determine if methylation correlates with reporter expression. If it does, we will have determined that the silencer is modular and can function without additional regions of endogenous sequence in *cis* (Fig. 9C).

The silencer deletion line can also be used to further investigate the role of the ROS1-RdDM feedback loop in Arabidopsis. These plants have high levels of *ROS1* transcription regardless of the genomic methylation status, and may mimic a constitutive ROS1

overexpression line without the need for confounding transgenes. ROS1 expression is no longer coupled to methylation level, and this will allow us to probe other mechanisms by which *ROS1* is regulated. For example, reporter expression driven by 2kb of *ROS1* promoter was upregulated in leaves following treatment with the phytohormones ethylene, auxin, or salicylic acid (Bennett et al., 2021). This deletion line could be used to quickly test whether this upregulation is dependent upon proximal promoter methylation, or relies on more distal regulatory elements. In this way, this deletion line can be used to separate the methylation-sensitive regulation of *ROS1* from all other regulatory mechanisms.

## METHODS

### Chromatin immunoprecipitation

ChIP-qPCR and ChIP-Seq protocols were modified from Saleh et al. (2008) and Lee et al. (2016). Briefly, 3g of 10-day old seedlings were crosslinked in Crosslinking Buffer with 1% paraformaldehyde (Saleh et al., 2008) with vacuum infiltration for 10 minutes. The reaction was stopped with the addition of 2.5M glycine to a final concentration of 125mM, and additional vacuum infiltration for 5 minutes. Cross-linked tissue was rinsed three times with deionized water, blotted dry, then frozen with liquid nitrogen. Frozen tissue was either stored in a -80°C for later use, or ground in a pre-chilled mortar and pestle to a fine powder and transferred to Nuclei Isolation Buffer (Saleh et al., 2008). After resuspension to a homogenous solution, samples were filtered twice through two pieces of Miracloth of each time, then centrifuged at 11,000xg for 30 minutes at 4°C. The resulting pellet was resuspended in Nuclei Lysis Buffer (Saleh et al., 2008) and sonicated in a BioRuptor on Medium for 5 cycles of 30s on, 30s off to produce chromatin from 200-1000bp.

Chromatin was diluted 1:9 in Nuclei Lysis Buffer and pre-cleared with Protein A-conjugated Dynabeads for 90 minutes with rotation at 4°C. After transferring supernatant to fresh tubes, samples were incubated with 5µl antibody ab5095 (abcam, Anti-RNA polymerase II CTD repeat YSPTSPS (phospho S2)) for 5 hours to overnight at 4°C with gentle rotation; a separate tube of chromatin was used for a No Antibody control. Dynabeads were added to each sample and incubated for an additional 4 hours at 4°C with gentle rotation, after which supernatant was discarded. Beads were washed twice each with the following buffers, each with a 5 minute incubation at 4°C with rotation: Low Salt Buffer, High Salt Buffer, LiCl Buffer, TE Buffer

(Saleh et al., 2008). DNA was eluted from the beads in Elution Buffer (Saleh et al., 2008) by incubation at 65°C for 30 minutes, vortexing every 2 minutes. Beads were removed after centrifugation at 16,000xg for 1 minute by transferring the supernatant to a fresh tube. Crosslinking was reversed by incubating samples at 65°C for 6 hours to overnight, but not exceeding 15 hours. A proteinase K digestion was then performed for 1.5hours at 45°C. After a phenol/chloroform/isoamyl alcohol extraction, DNA was precipitated in ethanol after at least 1 hour at -80°C.

Primers used for ChIP-qPCR are listed in Table I.

For ChIP-Seq, libraries were constructed using Swift Accel-NGS 2S Plus DNA Library Kit with 9 cycles of amplification and sequenced on an Illumina 2000 with 40bp single-end protocol at the Whitehead Genome Technology Core. Reads were trimmed with Trim Galore (Babraham Bioinformatics), removing 8bp from the 5' ends of read before mapping to the TAIR 10 genome (Lamesch et al., 2012) using Bowtie2 (Langmead and Salzberg, 2012).

**Construction of transgenes**

For the *ROS1:LUC* BAC transgene, *P. pyralis* Luciferase was isolated by PCR using primers listed in Table I, bounded by homologous sequence for the 3' end of *ROS1*. BAC JAtY62M02 was used for recombineering, following established protocols (Zhou et al., 2011). Col-0 plants were transformed via Agrobacterium-mediated floral dipping (Clough and Bent, 1998), and the transgene crossed into *rdr2*.

For GUS constructs, PCR was used to amplify the 4kb upstream of *ROS1*, the *ROS1* coding sequence, *GUS*, and the *nosT* terminator using primers with designed overhangs. Constructs were initially assembled into a pENTR-TOPO vector via Gibson assembly, then transformed into the binary vector pMDC99 (Curtis and Grossniklaus, 2003). Col-0 plants were transformed via Agrobacterium-mediated floral dipping (Clough and Bent, 1998).

**LUCIFERASE imaging**

3-week old plants were sprayed with 1mM firefly D-luciferin. After >30 minutes had elapsed, luminescence was measured using a Bethold NightOWL II LB 983: plants were placed in complete darkness for 2 minutes to reduce autofluorescence, followed by a 2-minute exposure. Luminescence was quantified using indiGO software.

**GUS staining**

Tissues were collected into ice-cold PBS, then stained with GUS (Jefferson, 1987; 100mM $Na_2HPO_4$, 10mM Potassium ferricyanide, 10mM Potassium ferrocyanide, 10mM EDTA pH8, 0.6% Triton X-100, 2mM X-GlcA): following vacuum infiltration (three times for 3 minutes each) samples were incubated at 37°C for approximately 24 hours. GUS stain was removed, tissues were then fixed in a 3:1 solution of Ethanol : Acetic Acid for 30 minutes at room temperature, and then cleared by incubating in 70% ethanol at 37°C for 2-4 hours. If necessary, samples were further cleared by incubation in 5% (w/v) sodium hydroxide for 2 hours at room temperature. After clearing samples were imaged under a light microscope.

**Generating deletions with CRISPR**

gRNAs were chosen to bound the edges of the desired deletions. If there was a nearby gene, gRNAs were selected from the database provided by Li et al. (2013); if there were no convenient genes, gRNAs were designed by hand or with E-CRISP (Heigwer et al., 2014). gRNA sequences are listed in Table II. For each deletion the two bounding gRNAs were cloned via Gibson Assembly (Gibson et al., 2009) to be under the control of *U6-26* promoters, in the plasmid pJKW0474 containing plant-codon optimized Cas9 under the control of the *CLAVATA3* promoter. pJKW0474 was a gift from Jing-Ke Weng (Addgene plasmid # 107588 ; http://n2t.net/addgene:107588 ; RRID:Addgene_107588). Col-0 plants were transformed with these constructs by Agrobacterium-mediated floral-dipping (Clough and Bent, 1998).

$T_1$ plants were selected by Basta and screened by PCR for deletions, using a 3-primer PCR with a forward and reverse primer outside of the deletion boundaries and a second reverse primer inside the deleted region. If the region is deleted, the inside reverse primer will not be able to anneal. Primers were designed so that the forward primer will create an amplicon of different size based on which reverse primer is used. Plants containing deletions were carried forward by selfing into the next to generation to ensure they were in the germ line. Heterozygous plants with deletions were used as pollen donors to pollinate *drm2-2* and *rdr2-1* plants. The resulting $F_1$ plants were genotyped to confirm the presence of the deletion and the absence of the Cas9 transgene. The $F_1$ plants were selfed. At least 95 $F_2$ plants were genotyped at 14-days old to identify plants that were homozygous with and without the deletion, and with and without the RdDM mutation. More plants were screened if sufficient numbers of those genotypes were not isolated.

**Quantitative real-time PCR**

RNA was isolated from the 5th leaf of 21-day old plants using TRIzol Reagent (Invitrogen) according to manufacturer's instructions. Genomic DNA was removed by treatment with Amplification-grade DNase I (Invitrogen). cDNA was prepared from 500-750ng RNA (standardized within each batch) with Superscript II Reverse Transcriptase (Invitrogen) according to the manufacturer's instructions, with polyadenylated transcripts selected for through use of an oligo-dT primer. Every reaction including one sample that went through the same protocol but without Superscript enzyme to confirm the absence of contaminating genomic DNA in the RNA sample when used in qPCR. qPCR was performed on a StepONE Plus Real-Time PCR system with Fast SYBR-Green PCR master mix (Applied Biosystems). Reactions were normalized to reference gene AT1G58050 (Czechowski et al., 2005). All qPCR reactions were performed with technical triplicates and a minimum of biological triplicates, unless otherwise noted. Cycling conditions were as follows: 95°C for 20s followed by 40 cycles of 95°C for 3s and 60°C for 30s. *ROS1* and *SDC* primers were from Williams et al. (2017) and are listed in Table 1 as *ROS1*_qPCR_ and *SDC*_qPCR_. Relative fold change in expression with respect to the geometric mean of the WT samples was determined using the $\Delta\Delta C_t$ method (Livak and Schmittgen, 2001). Standard error was calculated from combined standard deviations of both technical and biological replicates for each genotype. Significance was determined by one-way ANOVA with post-hoc Tukey HSD or by two-tailed Student's T-test using $\Delta\Delta C_t$ values.

**ACKNOWLEDGMENTS**

Thank you to Colette Picard for her indispensable help with ChIP-Seq analysis. Thank you to Ben Williams for designing and constructing the GUS and LUC transgenes. Thank you to the Bastien Christ and the Jing-Ke Weng lab for providing the backbone for the CRISPR construct and their guidance on creating the construct.

**Table I: Primers used in this study**

| Primer Name | Primer Sequence |
|---|---|
| ACT7_ChIPqF | CGTTTCGCTTTCCTTAGTGTTAGCT |
| ACT7_ChIPqR | AGCGAACGGATCTAGAGACTCACCTTG |
| QQS_ChIPqF | ACGTTGAAAGAAGCTTCAAACC |
| QQS_ChIPqR | TTGCGACACCTGATGTAGAAGT |
| ROS1_TSS_ChIPqF | GAGTCAGAAATGGAGAAACAGAGG |
| ROS1_TSS_ChIPqR | CTTCATGGGTGTCTGAGGAATC |
| ROS1_exon4_ChIPqF | ACTGCTCTCGTTCCTTACACAATG |
| ROS1_exon4_ChIPqR | GGAGTTACAGGCACAATTGCTCC |
| ROS1_exon14_ChIPqF | GTGAACCAATCATCGAAGAGCCT |
| ROS1_exon14_ChIPqR | GTCAGCTATTGATACTTCTGCGGT |
| *ROS1*_qPCR_F | CAGGCTTGCTTTTGGAAAGGGTACG |
| *ROS1*_qPCR_R | GTGCTCTCTCACTCTTAACCATAAGCT |
| *SDC*_qPCR_F | GTAGAAGTCAAGTCCTTGGGAGAT |
| *SDC*_qPCR_R | GAACTCATGAGCCGAAACCGAGA |

**Table II: gRNAs used for CRISPR**

| Deletion identifier | gRNA |
|---|---|
| 5'Δ15a | GGTGACATGTCTATGAGGCT |
| 5'Δ15b | GTGAAGCTTAGGCCTAACTA |
| 5'Δ0.8a | GGTGGTCATAACCTAATGAT |
| 5'Δ0.8b | GCGAAAGTTCGTTTGGTTGG |
| 3'Δ20prox_a | GAAGAAACCCAGGCTCAGCG |
| 3'Δ20prox_b | TCCGGCGCGACGCAGTACTG |
| 3'Δ20distal_a | TCCGGCGCGACGCAGTACTG |
| 3'Δ20distal_b | TACCAGATGGGGTATCCATG |

## REFERENCES


Agius, F., Kapoor, A., and Zhu, J. (2006) Role of the Arabidopsis DNA glycosylase/lyase ROS1 in active DNA demethylation. PNAS 103(31):11796-11801.

Batista, R.A., Moreno-Romero, J., Qiu, Y., van Boven, J., Santos-González, J., Figueiredo, D.D., Köhler, C. (2019) The MADS-box transcription factor PHERES1 controls imprinting in the endosperm by binding to domesticated transposons. eLife 8: e50541. DOI: https://doi.org/10.7554/eLife.50541

Bennett, M., Cleaves, K., and Hewezi, T. (2021) Expression patterns of DNA methylation and demethylation genes during plant development and in response to phytohormones. Int J Mol Sci 22(18): 9681. doi: 10.3390/ijms22189681.

Campanero, M.R., Armstrong, M.I., and Flemington, E.K. (2000) CpG methylation as a mechanism for the regulation of E2F activity. PNAS 97(12): 6481-6486.

Cerutti, A., Jauneau, A., Laufs, P., Leonhardt, N., Schattat, M.H., Berthomé, R., Routaboul, JM., and Noël, L.D. (2019) Mangroves in the leaves: anatomy, physiology, and immunity of epithermal hydathodes. Annu. Rev. Phytopathol. 57: 91-116.

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

Córdoba-Cañero, D., Cognat, V., Ariza, R.R., Arjona, T.R., and Molinier, J. (2017) Dual control of ROS1-mediated active DNA demethylation by the DNA DAMAGE BINDING protein 2 (DDB2). Plant Journal 92(6):1170-1181. doi: 10.1111/tpj.13753.

Curtis, M.D., and Grossniklaus, U. (2003) A Gateway cloning vector set for high-throughput functional analysis of genes in planta. Plant Physiology 133(2): 462-469. https://doi.org/10.1104/pp.103.027979.

Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K., and Scheible, W.-R. (2005). Genome-Wide Identification and Testing of Superior Reference Genes for Transcript Normalization in Arabidopsis. Plant Physiol. 139: 5–17.

Eremina, M., Rozhon, W., Yang, S., and Poppenberger, B. (2015) ENO2 activity is required for the development and reproductive success of plants, and is feedback-repressed by AtMBP-1. The Plant Journal 81(6): 895-906.

Gehring, M. (2019) Epigenetic dynamics during flowering plant reproduction: evidence for reprogramming? New Phytologist Foundation 224(1): 91-96.

Gehring, M., Bubb, K.L., and Henikoff, S. (2009) Extensive demethylation of repetitive elements during seed development underlies gene imprinting. Science 324(5933): 1447-1451. DOI: 10.1126/science.1171609.

Gehring, M., Missirian, V., and Henikoff, S. (2011) Genomic analysis of parent-of-origin allelic expression in *Arabidopsis thaliana* seeds. PLoS One 6(8): e23687.

Gehring, M., and Satyaki, P.V. (2017) Endosperm and imprinting, inextricably linked.

Gibson, D.G., Young, L, Chuang, RY., Venter, J.C., Hutchison III, C.A., and Smith, H.O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. Nature Methods 6: 343-345.

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

Halter, T., Wang, J., Amesefe, D., Lastrucci, E., Charvin, M., Rastogi, M.S., and Navarro, L. (2021) The Arabidopsis active demethylase ROS1 *cis*-regulates defence genes by erasing DNA methylation at promoter-regulatory regions. eLife 10: e62994. DOI: https://doi.org/10.7554/eLife.62994

Heigwer, F. , Kerr, G., and Boutros, M. (2014) E-CRISP: fast CRISPR target site identification. Nat. Methods 11: 122-123.

Hénaff, E., Vives, C., Desvoyes, B., Chaurasia, A., Payet, J., Gutierrez, C., and Casacuberta, J.M. (2014) Extensive amplification of the E2F transcription factor binding sites by transposons during evolution of *Brassica* species. The Plant Journal 77(6): 852-862. https://doi.org/10.1111/tpj.12434.

Henderson, I.R. and Jacobsen, S.E. (2008) Tandem repeats upstream of the *Arabidopsis* endogene *SDC* recruit non-CG DNA methylation and initiate siRNA spreading. Genes & Development 22: 1597-1606.

Hsieh, TF., Shin, J., Uzawa, R., Silva, P., Cohen,S., Bauer, M.J., Hashimoto, M., Kirkbride, R.C., Harada, J.J., Zilberman, D., and Fischer, R.L. (2011) Regulation of imprinted gene expression in Arabidopsis endosperm. Proc Natl Acad Sci U S A 108(5): 1755-1762. doi: 10.1073/pnas.1019273108.

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

Jefferson, R.A. (1987) Assaying chimeric genes in plants: the GUS gene fusion system. Plant Molecular Biology Reporter 5(1): 387-405.

Kawakatsu, T., Stuart, T., Valdes, M., Breakfield, N., Schmitz, R.J., Nery, J.R., Urich, M.A. Han, X., Lister, R., Benfey, P.N., and Ecker, J.R. (2016) Unique cell-type-specific patterns of DNA methylation in the root meristem. Nature Plants 2: 16058. https://doi.org/10.1038/nplants.2016.58.

Kinoshita, T., Miura, A., Choi, Y., Kinoshita, Y., Cao, X., Jacobsen, S.E., Fischer, R.L., and Kakutani, T. (2004) One-way control of *FWA* imprinting in *Arabidopsis* endosperm by DNA methylation. Science 303(5657): 521-523.

Kuehner, J.N., Pearson, E.L., and Moore, C. (2011) Unravelling the means to an end: RNA polymerase II transcription termination. Nat Rev Mol Cell Biol 12: 283-294.

Kurihara, Y., Matsui, A., Kawashima, M., Kaminuma, E., Ishida, J., Morosawa, T., Mochizuki, Y., Kobayashi, N., Toyoda, T., Shinozaki, K., and Seki, M. (2008) Identification of the candidate genes regulated by RNA-directed DNA methylation in *Arabidopsis*. Biochemical and Biophysical Research Communications 376(3): 553-557.

Lamesch, P., Berardini, T. Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., Muller, R., Dreher, K., Alexander, D. L., Garcia-Hernandez, M., Karthikeyan, A. S., Lee, C. H., Nelson, W. D., Ploetz, L., Singh, S., Wensel, A., and Huala, E. (2012) The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. Nucleic Acids Research, 40(Database issue), D1202–D1210. https://doi.org/10.1093/nar/gkr1090.

Langmead, B., and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. Nature Methods. 9: 357-359. https://doi.org/10.1038/nmeth.1923.

Le, T.N., Schumann, U., Smith, N.A., Tiwari, S., Au P.C.K., Zhu, Q.H., Taylor, J.M., Kazan, K., Llewellyn, D.J., Zhang, R., Dennis, E.S., and Wang, M.B. (2014) DNA demethylases target promoter transposable elements to positively regulate stress responsive genes in *Arabidopsis*. Genome Biology 15: 458-476.

Lee, T.I., Johnstone, S.E., and Young, R.A. (2006) Chromatin immunoprecipitation and microarray-based analysis of protein location. Nat Protoc. 1(2): 729-748. doi: 10.1038/nprot.2006.98.

Lei, M., Zhang, H., Julian, R., Tang, K., Xie, S., and Zhu, JK. (2015) Regulatory link between DNA methylation and active demethylation in *Arabidopsis.* PNAS 112(11): 3553-3557.

Li, J.F., Norville, J., Aach, J., McCormack, M., Zhang, D., Bush, J., Church, G.M., and Sheen, J. Multiplex and homologous recombination–mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. Nat Biotechnol 31: 688–691 (2013). https://doi.org/10.1038/nbt.2654

Liang, W., Li, J., Sun, L., Liu, Y., Lan, Z., and Qian, W. (2021) Deciphering the synergistic and redundant roles of CG and non-CG DNA methylation in plant development and TE silencing. New Phytol. 10.1111/nph.17804. Advance online publication. https://doi.org/10.1111/nph.17804

Lin, W., Sun, L., Huang, RZ., Liang, W., Liu, X., He, H., Fukuda, H., He, XQ., and Qian, W. (2020) Active DNA demethylation regulates tracheary element differentiation in *Arabidopsis*. Sci. Adv. 6(26): eaaz2963.

Liu, C., Wang, C., Wang, G., Becker, C., Zaidem, M., and Weigel, D. (2016) Genome-wide analysis of chromatin packing in *Arabidopsis thaliana* at single-gene resolution. *Genome Res.* 26: 1057-1068.

Liu, H. Zhang, H., Dong, Y.X., Hao, Y.J., and Zhang, X.S. (2017) *DNA METHYLTRANSFERASE1*-mediated shoot regeneration is regulated by cytokinin-induced cell cycle in *Arabidopsis.* New Phytologist 217(1): 219-232.

Livak, K.J. and Schmittgen, T.D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2-ddCt Method. Methods 25: 402–408.

Morales-Ruiz, T., Ortega-Galisteo, A., Ponferrada-Marín, M., Ariza, R., and Roldán-Arjona, T. (2006) DEMETER and REPRESSOR OF SILENCING 1 encode 5-methylcytosine DNA glycosylase. PNAS 103(18): 6853-6858.

Muiño, J.M., de Bruijn, S., Pajoro, A., Geuten, K., Vingron, M., Angenent, G.C., and Kaufmann, K. (2016) Evolution of DNA-binding sites of a floral master regulatory transcription factor. Molecular Biology and Evolution 33(1): 185-200. https://doi.org/10.1093/molbev/msv210

O'Malley, R.C., Huang, S.C., Song, L., Lewsey, M.G., Bartlett, A., Nery, J.R., Galli, M., Gallavotti, A., and Ecker, J.R. (2016) Cistrome and epicistrome features shape the regulatory DNA landscape. Cell 165(5): 1280-1292. DOI:https://doi.org/10.1016/j.cell.2016.04.038.

Park, K., Kim, M.Y., Vickers, M., Park, JS., Hyun, Y., Okamoto, T., Zilberman, D., Fischer, R.L., Feng, X., Choi, Y., and Scholten, S. (2016) DNA demethylation is initiated in the central cells of *Arabidopsis* and rice. PNAS 113(52): 15138-15143. https://doi.org/10.1073/pnas.1619047114

Pignatta, D., Erdmann, R.M., Scheer, E., Picard, C.L., Bell, G.W., and Gehring, M. (2014) Natural epigenetic polymorphisms lead to intraspecific variation in Arabidopsis gene imprinting. eLife 3: e03198. doi: 10.7554/eLife.03198

Pignatta, D., Novitzky, K., Satyaki, P.R.V., and Gehring, M. (2018) A variably imprinted epiallele impacts seed development. PLoS Genet. 14(11): e1007469. https://doi.org/10.1371/journal.pgen.1007469

Ramirez-Parra, E., Fründt, C., and Gutierrez, C. (2003) A genome-wide identification of E2F-regulated genes in *Arabidopsis*. The Plant Journal 33(4): 801-811.
Ramirez-Parra, E., and Gutierrez, C. (2007) E2F regulates *FASCIATA1,* a chromatin assembly gene whose loss switches on the endocycle and activates gene expression by changing epigenetic status. Plant Physiology 144(1): 105-120.
Saleh, A., Alvarez-Venegas, R., and Avramova, Z. (2008) An efficient chromatin immunoprecipitation (ChIP) protocol for studying histone modifications in *Arabidopsis* plants. Nature Protocols 3(16): 1018-1025.
Trimarchi, J.M. and Lees, J.A. (2002) Sibling rivalry in the E2F family. Nat Rev Mol Cell Biol 3: 11-20. https://doi.org/10.1038/nrm714.
Vandepoele, K., Raes, J., De Veylder, L, Rouzé, P., Rombauts, S., and Inzé, D. (2002) The Plant Cell 14(4): 903-916.
Vandepoele, K., Vlieghe, K., Florquin, K., Hennig, L., Beemster, G. T., Gruissem, W., Van de Peer, Y., Inzé, D., & De Veylder, L. (2005) Genome-wide identification of potential plant E2F target genes. Plant Physiology, 139(1), 316–328. https://doi.org/10.1104/pp.105.066290
Verkest, A., Abeel, T., Heyndrickx, K. S., Van Leene, J., Lanz, C., Van De Slijke, E., De Winne, N., Eeckhout, D., Persiau, G., Van Breusegem, F., Inzé, D., Vandepoele, K., & De Jaeger, G. (2014) A generic tool for transcription factor target gene discovery in Arabidopsis cell suspension cultures based on tandem chromatin affinity purification. Plant physiology 164(3), 1122–1133. https://doi.org/10.1104/pp.113.229617
Vlieghe, K., Vuylsteke, M., Florquin, K., Rombauts, S., Maes, S., Ormenese, S., Van Hummelen, P., Van de Peer, Y., Inzé, D., and De Veylder, L. (2003) Microarray analysis of *E2Fa-DPa*-overexpressing plants uncovers a cross-talking genetic network between DNA replication and nitrogen assimilation. J Cell Science 116(20): 4249-4259.
Wang, HL. And Chekanova, J.A. (2019) Novel mRNAs 3' end-associated *cis*-regulatory elements with epigenomic signatures of mammalian enhancers in the *Arabidopsis* genome. RNA 25: 1242-1258. Doi:10.1261/rna. 071209.119.
Williams, B.P., Pignatta, D., Henikoff, S., and Gehring, M. (2015) Methylation-sensitive expression of a DNA demethylase gene serves as an epigenetic rheostat. PLoS Genet. 11(3): e1005142. doi:10.1371/journal.pgen.1005142.
Williams, B.P. and Gehring, M. (2017) Stable transgenerational epigenetic inheritance requires a DNA methylation-sensing circuit. Nat Commun. 8: 2124. https://doi.org/10.1038/s41467-017-02219-3.
Williams, B.P. and Gehring, M. (2020) Principles of epigenetic homeostasis shared between flowering plants and mammals. Trends in Genetics 36(10): 751-763.
Wolff, P., Weinhofer, I., Seguin, J., Roszak, P., Beisel, C., Donoghue, M.T.A., Spillane, C., Nordborg, M., Rehmsmeier, M., and Köhler, C. (2011) High-resolution analysis of parent-of-origin allelic expression in the Arabidopsis endosperm. PLoS Genet 7(6): e1002126.
Yagi, H., Nagano, A.J., Kim, J., Tamura, K., Mochizuki, N., Nagatani, A., Matsushita, T., and Shimada, T. (2021) Fluorescent protein-based imaging and tissue-specific RNA-seq analysis of Arabidopsis hydathodes. Journal of Experimental Biology 72(4): 1260-1270. https://doi.org/10.1093/jxb/eraa519.

Yamamuro, C., Miki, D., Zheng, Z., Ma, J., Wang, J., Yang, Z., Dong, J., Zhu, J.K. (2014) Overproduction of stomatal lineage cells in *Arabidopsis* mutants defective in active DNA demethylation. Nature Communications 5: 4062.

Yu, A., Lepère, G., Jay, F., Wang, J., Bapaume, L., Wang, Y., Abraham, A.L., Penterman, J., Fischer, R.L., Voinnet, O., and Navarro, L. (2012) Dynamics and biological relevance of DNA demethylation in *Arabidopsis* antibacterial defense. *PNAS* 110(6): 2389-2394.

Zhu, QH., Shan, WX., Ayliffe, M.A., and Wang, MB. (2016) Epigenetic mechanisms: an emerging player in plant-microbe interactions. MPMI 29: 187-196. https://doi.org/10.1094/MPMI-08-15-0194-FI.

## Chapter Three

## Active demethylation is required for regulation of *FLC,* correct flowering time, and anther dehiscence

Portions of this work have been published as Williams, B.P., Bechen, L.L., Pohlmann, D.A., and Gehring, M. (2021) Somatic DNA demethylation generates tissue-specific methylation states and impacts flowering time. *The Plant Cell*. koab319 doi: 10.1093/plcell/koab319.

Experiments on flowering time, fertility, and anther dehiscence were performed in collaboration with L. Bechen. L. Bechen performed EM-Seq. B. Williams generated the *drdd* mutants and performed RNA-Seq analysis.

## ABSTRACT

Proper regulation of *FLOWERING LOCUS C (FLC*) ensures that flowering plants begin the process of flowering at the right time. While many genetic and epigenetic mechanisms are known to affect *FLC* transcription, no role for DNA methylation or demethylation has been previously described in this regulation. We generated the somatic quadruple mutant *dme;ros1;dml2;dml3* (*drdd*), homozygous for mutations in all of the DNA demethylase genes in Arabidopsis, in order to assess the role of active demethylation in the somatic tissues of the plant. We found that *drdd* plants exhibit early flowering, and link that phenotype to hypermethylation in a region 5' of *FLC*. In addition, we detected a partial male sterility defect in *drdd* caused by delayed dehiscence of the anthers, which may be caused by a lack of reactive oxygen species. This is the first time that DNA demethylation has been implicated in flowering time or anther dehiscence.

## INTRODUCTION

Flowering time and plant fertility are important elements of agriculture. Some cultivars of spinach, for example, are sensitive to conditions that cause them to bolt early, resulting in a loss of yield and quality (Chitwood et al., 2016). In addition, several important agricultural crops such as rice, maize, and cotton report reduced male fertility under high temperature stress, emphasizing the importance of studying these processes for an era of global climate change (Herrero and Johnson, 1980; Rang et al., 2001; Prasad et al., 2006; Snider et al., 2011). A concerted effort has therefore been made to improve our understanding of the genetic and epigenetic regulatory networks that underlie flowering in many different species of flowering plants and the defects in fertility that can occur therein (Abou-Elwafa et al., 2011; Nie et al., 2016; Ou et al., 2017; Avila de Dio et al., 2019; Zhang et al., 2020; Ghorbani et al., 2021).

*FLOWERING LOCUS C* (*FLC*) is a MADS-box transcription factor that represses flowering and must be downregulated before the transition from vegetative to reproductive development can occur (Michaels et al., 1999). The *FLC* locus is regulated by a number of mechanisms through either the autonomous or vernalization flowering time pathways. Before vernalization, *FLC* is upregulated in vegetative tissue by the actions of *FRIGIDA* and a complex of histone acetyltransferases and methyltransferases (Li et al., 2018). Vernalization results in a change in the local chromatin environment mediated by the Polycomb Repressive Complex 2

(Yang et al., 2017; Jiang and Berger, 2017) and the lncRNA *COOLAIR* (Csorba et al. 2014). In the autonomous flowering pathway, in the absence of vernalization, *FLC* is downregulated by a suite of 3' RNA processing proteins, increased transcription of *COOLAIR*, and the loss of H3K4me2 (Wu et al., 2020). However, no role for DNA methylation has been proposed to regulate *FLC* in the autonomous pathway.

Epigenetic changes such as DNA methylation allow plants to modify gene expression throughout development and in response to environmental cues. In plants, cytosine methylation occurs in all three sequence contexts (CG, CHG, and CHH where H is any nucleotide other than G). *De novo* methylation is added through the actions of the RNA-Directed DNA Methylation (RdDM) pathway (Matzke and Mosher, 2014). Methylation can be removed passively, through a failure to renew the methylation on the daughter strand after DNA replication, or actively through the actions of the DEMETER family of Helix-hairpin-helix GPD DNA glycosylase/lyases. This family consists of four members: *DEMETER* (*DME*), which is necessary in the central cell of the female gametophyte, *REPRESSOR OF SILENCING 1* (*ROS1*), *DEMETER-LIKE 2* (*DML2*), and *DEMETER-LIKE 3* (*DML3*). It has not been possible to evaluate the potential roles of DME in vegetative tissues because seeds that inherit a mutant *dme* allele from the maternal parent are aborted (Choi et al., 2002). In contrast, the triple mutant *ros1;dml2;dml3* (*rdd*) is viable and exhibits hypermethylation at the 5' and 3' ends of a subset of genes (Penterman et al., 2007; Lister et al., 2008). There are few reported phenotypes in the *rdd* mutant; the plants are more susceptible to infection by the fungal pathogen *F. oxysporum* (Le et al., 2014), there is an overproduction of stomatal precursor cells due to the hypermethylation and downregulation of a repressor of stomatal formation (Yamamuro et al. 2014), and there are discontinuities in the protoxylem (Lin et al., 2020). However, the role of active demethylation as a whole has been indiscernible because DME might be compensating for the loss of the other three demethylases in the *rdd* mutant.

In this chapter, I describe our bypass of the *dme* seed abortion phenotype by rescuing *DME* transcription solely in the central cell of *dme* mutant plants. We were then able to obtain plants homozygous for mutations in *dme,* including the quadruple mutant *dme;ros1;dml2;dml3* (*drdd*). *drdd* plants are early flowering, which we determine is correlated with hypermethylation and concurrent downregulation of *FLC*. In addition, anther dehiscence is also affected in *drdd*

plants, resulting in a male sterility phenotype. Active demethylation therefore plays a role in the vegetative-to-reproductive transition and in male fertility.

## RESULTS

### Isolating quadruple demethylase mutants

In order to fully understand the role of the DME family of 5-methylcytosine DNA glycosylases in Arabidopsis, we sought to isolate homozygous mutations in all four orthologues in this family – *DME*, *ROS1*, *DML2* and *DML3*. To bypass the *dme* seed abortion phenotype, we created a transgene in which the genomic coding sequence of DME was expressed under the central cell-specific promoter of *AGL61* (Fig. 1A) (Steffen et al., 2008). This transgene was transformed into the previously isolated *rdd* triple mutant (Penterman et al., 2007). Transgenic *rdd* plants were pollinated by heterozygous *dme* mutants to create heterozygous quadruple mutants (Fig. 1B). These heterozygotes were self-fertilized and seed abortion rates were quantified to test the ability of the *pAGL61:DME* transgene to complement the *dme* seed abortion phenotype. Whereas non-transgenic plants harboring the *dme* mutation exhibit 50% seed abortion, multiple transgenic lines expressing *pAGL61:DME* exhibited minimal seed abortion, indistinguishable from *rdd* plants (Fig. 1C). Expression of *DME* in the central cell before fertilization is therefore sufficient to rescue the post-fertilization seed abortion phenotype. After self-fertilizing the heterozygous quadruple mutant, the following four genotypes were isolated over two subsequent generations: homozygous quadruple mutants (hereafter termed *drdd*), homozygous *rdd* mutants, homozygous *dme* mutants, and homozygous wild-type segregants (hereafter termed WT *DRDD)*, which serve as closely-related wildtype controls containing the *pAGL61:DME* transgene for subsequent experiments. All four of these genotypes were determined to be homozygous for a single transgene insertion, and displayed seed abortion rates of <2%, similar to non-transgenic *rdd* (Fig. 1C).

To assay whether the *AGL61* promoter was leaky and causing *DME* expression outside of the central cell, we isolated RNA from adult leaf tissue and used qRT-PCR to evaluate the levels of *DME* transcription. Two primer sets for *DME* transcripts were used: Region 1 primers were designed to span the *dme-2* T-DNA insertion, so are specific to the endogenous wildtype *DME* allele and the transgene, but not to the mutant *dme-2* allele. Region 2 primers target a region downstream of the T-DNA insertion, which therefore should not be found in mRNA from the

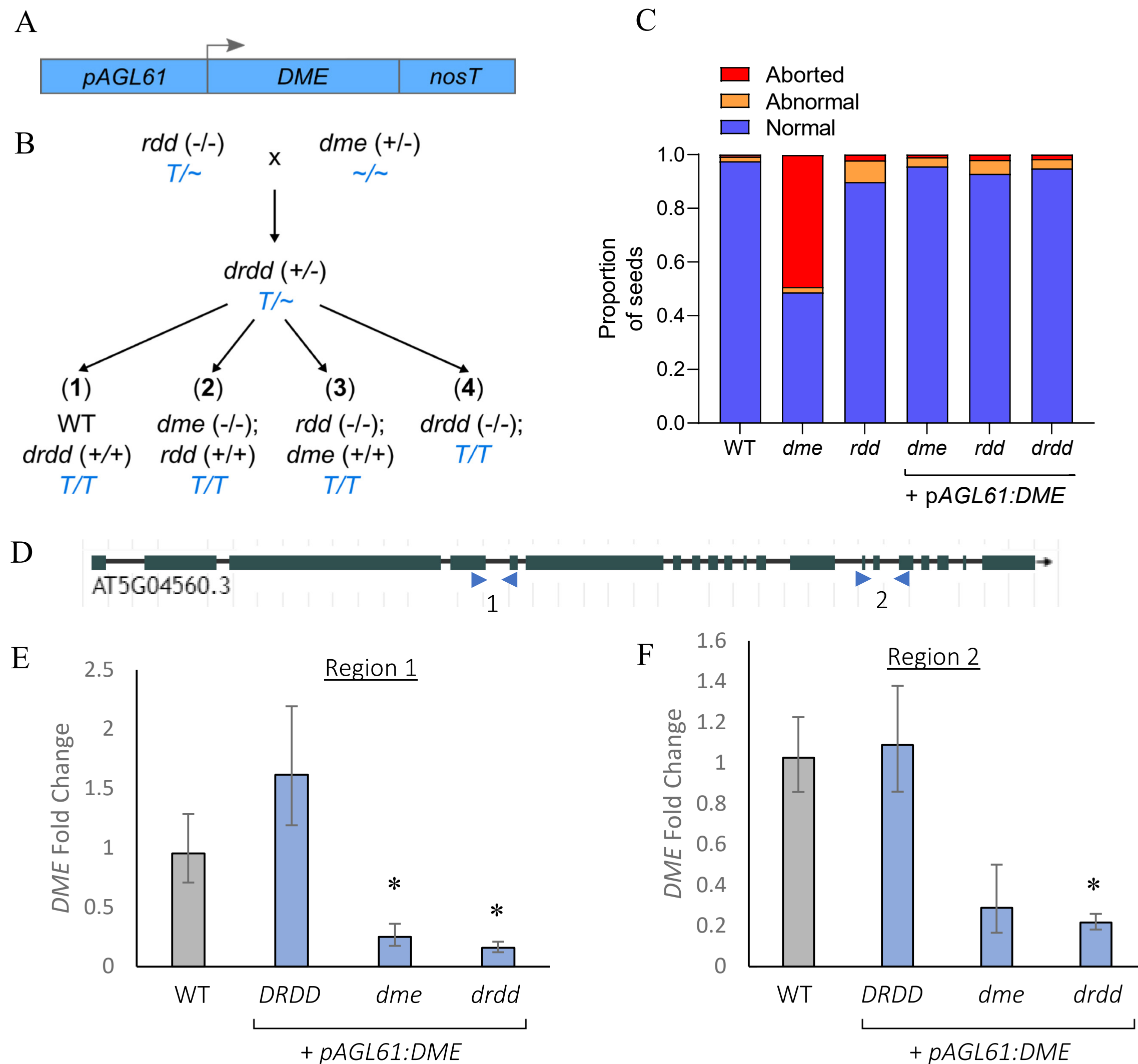


**Figure 1: Generation and evaluation of homozygous *dme* and quadruple *drdd* mutants**
A) Schematic showing the construct designed to express *DME* specifically in central cells using the *AGL61* promoter. B) Schematic showing the segregation of mutant genotypes. Transgenic *rdd* plants were pollinated by heterozygous (+/-) *dme* to generate heterozygous *drdd* F1 progeny. WT, *dme*, *rdd* and *drdd* homozygotes were isolated from two subsequent generations.
C) Proportion of aborted seeds in non-transgenic WT, *dme*, *rdd* plants, as well as *dme*, *rdd* and *drdd* plants homozygous for the *pAGL61:DME* transgene. 250-300 seeds were evaluated for each genotype. D) Schematic of *DME* showing qPCR primer locations. E-F) qPCR results showing fold change of *DME* transcript levels for WT *DRDD* (*DRDD* in legend), WT, *dme*, and *drdd* in 23-day old leaf tissue. All samples with the *pAGL61:DME* transgene are in blue. Error bars represent standard error of the mean. Biological replicates are n=4 for WT *DRDD,* WT, and *dme*, and n=3 for *drdd.* Asterisks indicate a significant difference from WT of $p<0.05$ by Student's T-test. There is no significant difference between WT *DRDD* and WT (Two-tailed Student's T-test: Region 1: $t(3)=-1.32$, $p=.23$; Region 2: $t(3)=-0.33$, $p=0.75$).

*dme-2* allele (Fig. 1D). There was no significant difference in WT *DRDD* samples with the transgene and wildtype samples without, and there was approximately a 5-fold (Fig. 1F) to 10-fold (Fig. 1E) reduction in *DME* levels between WT *DRDD* and *drdd.* Amplification was lower for Region 1 than Region 2 using the same cDNA, likely due to differences in primer efficiencies and bias from the fact that reverse transcription begins at the 3' end (Udvardi et al., 2008). These results suggest that the *pAGL61:DME* transgene is not substantially expressed in leaves and that *drdd pAGL61:DME* have substantially reduced *DME* activity.

**Demethylation is necessary for proper timing of flowering, as regulated by *FLC***

Initial experiments revealed a distinct early flowering phenotype in *drdd* plants compared to WT *DRDD*, *dme*, and *rdd* when all were grown together (Fig. 2A). To quantify this effect, we determined the flowering time for several sets of these plants by counting the number of true leaves that had been produced at the time of bolting, which is when the plants produce a flowering stem. Under both long- and short-day light conditions, *drdd* plants consistently flowered significantly earlier than the other genotypes (Fig. 2B-C). Under short-day conditions (Fig. 2B) the number of mutated demethylase genes correlated with how early the plants flowered, with *drdd* plants flowering the earliest, followed by *rdd, dme*, and WT *DRDD* plants. Under long-day conditions there was no significant difference between *rdd* and *dme* plants, but both flowered slightly earlier than the average WT *DRDD* plant (Fig. 2C).

As flowering time is regulated by *FLOWERING LOCUS C,* we hypothesized that this gene might be dysregulated in the demethylase mutants. RNA-Seq analysis of 3-week old leaf tissue revealed decreased *FLC* transcripts in all three mutants when compared to WT *DRDD*, approximately 2-fold lower in *dme* and *rdd* and 8-fold lower in *drdd* (Fig. 2D). qRT-PCR from 21-day old leaf tissue showed that *FLC* transcription levels varied significantly between biological replicates, but were approximately 20-fold lower on average in *drdd* plants compared to WT *DRDD* (Fig. 2E). While *dme* and *rdd* also showed decreased *FLC* transcript levels, they were not reduced as far as *drdd*, correlating with their flowering after *drdd* plants but before WT *DRDD*. We investigated the methylation status of the 5' flanking region of *FLC* by EM-Seq (Enzymatic Methyl-Seq; Vaisvila et al., 2021) and found hypermethylation in *drdd* in all cytosine sequence contexts 800-2100bp upstream of the transcriptional start site (Fig. 2F). Examination of smallRNA-Seq data from wild-type embryos (Erdmann et al., 2017) also showed

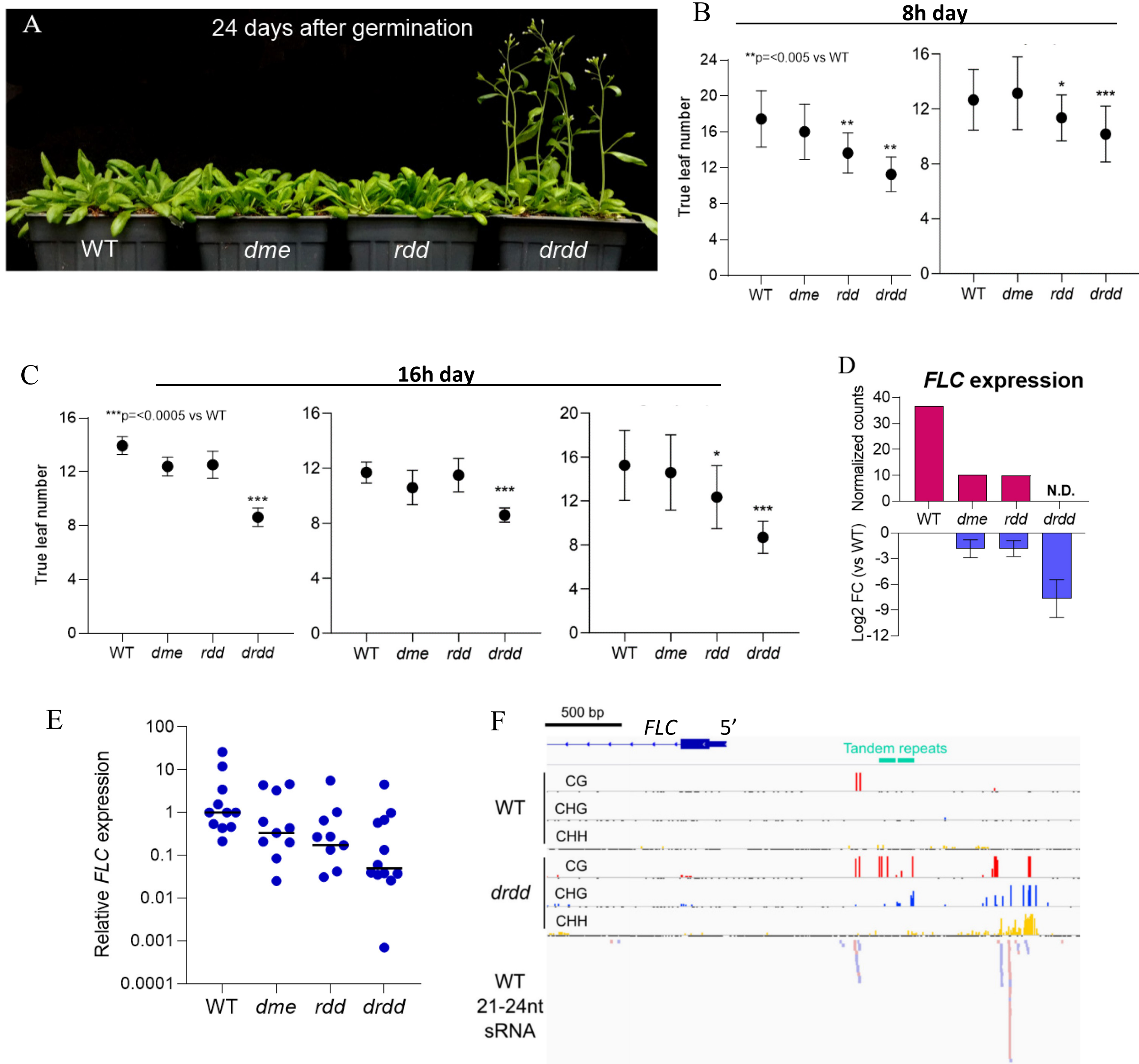


**Figure 2: Loss of DNA demethylases results in hypermethylation and decreased transcription of *FLOWERING LOCUS C,* inducing early flowering.**
A) Early flowering is evident in *drdd* plants compared to WT *(*WT *DRDD*), *dme*, and *rdd* plants all grown in long-day conditions. B-C) Flowering time, as measured by the number of true leaves at the time of bolting, of replicate cohorts of WT (WT *DRDD*), *dme*, *rdd*, and *drdd* plants grown in B) short- (8h) and C) long-day (16h) conditions. Error bars represent standard deviations of biological replicates. p-values are from unpaired t-tests. n>30 for each genotype. D) *FLC* expression and fold-change compared to WT (WT *DRDD*) as measured by RNA-seq. Error bars represent standard error between biological replicates. E) Relative *FLC* expression from long-day plants measured by quantitative real-time PCR. Each point represents an individual biological replicate, horizontal lines denote the median value. F) EM-Seq analysis of methylation levels and small-RNA-Seq coverage of the region 5’ of *FLC*.

a significant number of reads mapping to this area, indicating that this region is likely targeted for methylation by the RdDM pathway in wildtype plants, but any methylation is removed by the 5-methylcytosine DNA glycosylases. This also implies that the RdDM pathway is active at more loci than are methylated in wildtype plants; like *FLC*, there may be a significant number of regions where active demethylation efficiently counters RdDM activity to minimize detectable methylation.

The hypermethylation of *FLC* in *drdd* (Fig. 2F) matches the decrease in *FLC* transcription as detected both in RNA-Seq (Fig. 2D) and by qPCR (Fig. 2E), and links to the early flowering phenotype.

**Demethylation by DRDD induces timely anther dehiscence and affects male fertility**

We also detected a fertility defect in *drdd* plants that was not evident in *rdd* or *dme.* With varying penetrance and expressivity, *drdd* plants exhibit a failure to produce elongated siliques, indicating a lack of developing seeds inside (Fig. 3A-C). Some plants fail to produce any siliques (Fig. 3A), others stop producing siliques after an indeterminate amount of time (Fig. 3B), and still other plants produce some fertile and some infertile bolts (Fig. 3C). We quantified this effect for a cohort of 28 *drdd* plants and found that 71% showed a defect in fertility, including 39% that were completely sterile (Fig. 3D). Sterility can be caused by defects in the male parent, female parent, or both. When pollinated manually by wildtype pollen, *drdd* flowers produced elongated siliques, indicating that female fertility is not compromised (data not shown). Upon examination under magnification, it became evident that sterile *drdd* plants have a failure of anther dehiscence in the first open flower, meaning that pollen is not being released from the anthers, although dehiscence has occurred by the fourth open flower (Fig. 3E). When a flower has aged to become the fourth open flower, however, the pistil has already elongated past the anthers, which may result in a failure of *drdd* flowers to self-fertilize as the pistil never brushes past anthers that are releasing pollen (Fig. 3F). Using a simplified Alexander's stain, which stains non-aborted pollen purple and aborted pollen blue-green (Peterson et al., 2010), we determined that *drdd* pollen showed no viability defect, further indicating that the timing of dehiscence may be the causal problem underlying the *drdd* sterility phenotype.

Anther dehiscence occurs when the anther breaks open to release the pollen that has developed inside. This requires two major processes: cell death of the septum cells that separate

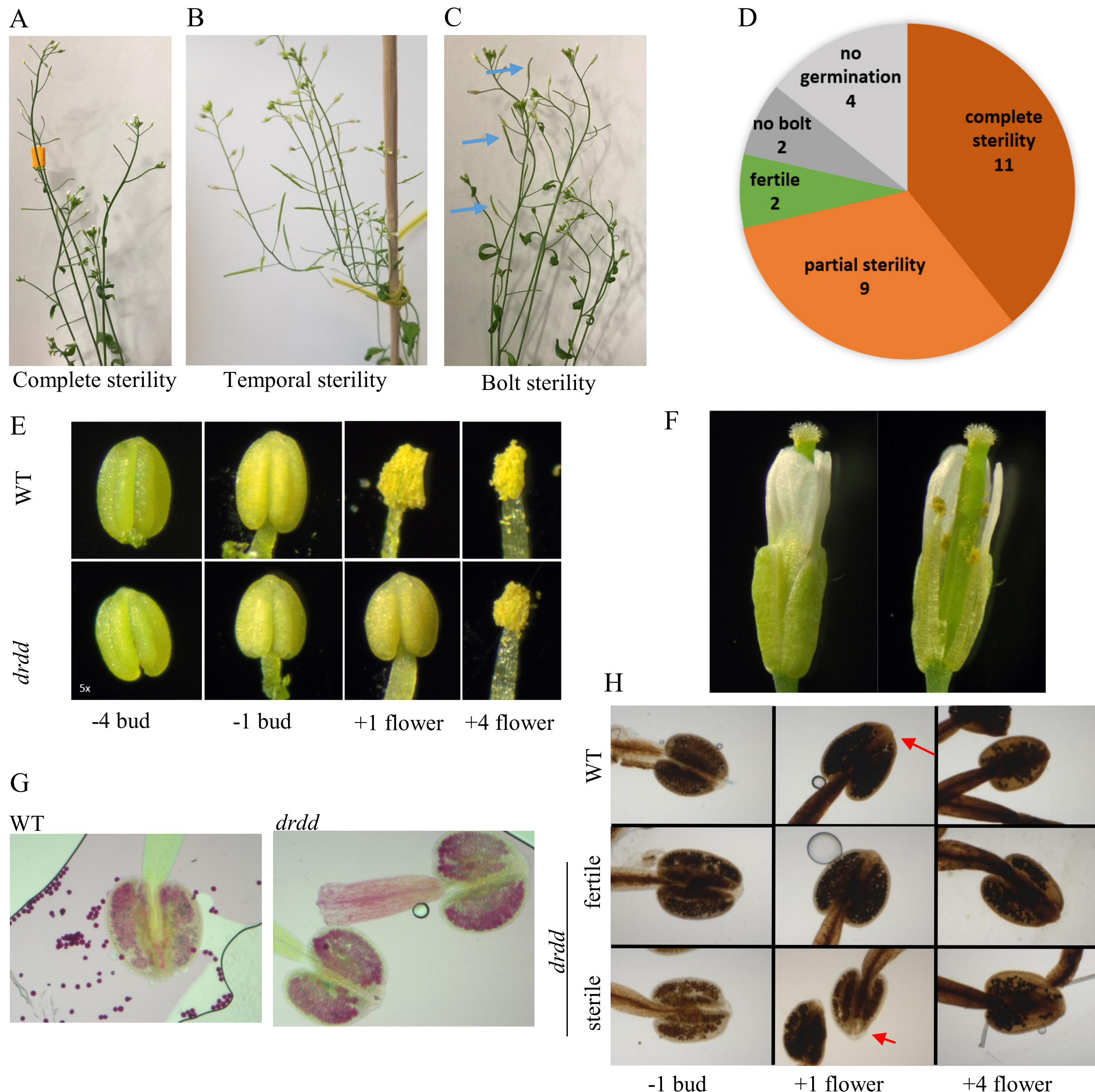


**Figure 3: DRDD are required for proper anther dehiscence and male fertility**
A-C) Variable sterility phenotypes in *drdd* plants, ranging from no fertile siliques (A), to plants that switch from producing fertile siliques to infertile (B), to plants where fertility varies from bolt to bolt within the same plant (C); blue arrows indicate fertile siliques on a single bolt. All plants were grown at the same time under the same conditions. D) Number of 7-week old *drdd* plants from a single flat with each state of fertility. E) Anthers from WT (*DRDD*) and *drdd* from the 4$^{th}$ largest bud, largest bud, first open flower, and 4$^{th}$ open flower at 5x. F) 1.25x magnification of 4$^{th}$ open flower of a sterile *drdd* plant. G) Alexander staining of anthers 1$^{st}$ open flower of WT (*DRDD*) and *drdd*; purple = viable, green = inviable. H) DAB staining of anthers from WT (*DRDD*), fertile *drdd*, and sterile *drdd* inflorescences.

the pollen chambers from each other, and the opening of the anther by a combination of dehydration and secondary cell wall thickening causing physical stress (Wilson et al., 2011).

Previous analyses of mutants with delayed anther dehiscence have linked this phenotype to a failure of secondary wall lignification in the anthers due to a lack of reactive oxygen species (Mitsuda et al., 2005; Dai et al., 2019). To test this hypothesis, we treated anthers of WT *DRDD* and both fertile and sterile *drdd* plants with 3', 3'-diaminobenzidine (DAB), which reacts with $H_2O_2$ to produce a brown precipitate. While pollen grains stain to the same degree in all three genotypes, we found that the tissue of the anther itself, especially at the tip, had less staining in the sterile *drdd* first open flower compared to the anther tissue from bothWT *DRDD* and fertile *drdd* flowers (Fig. 3H, compare red arrows). By the fourth open flower, there is no discernable difference in the intensity of stain between the genotypes. This indicates that the sterile *drdd* anther has a delay in accumulating $H_2O_2$, which correlates with the delay in anther dehiscence.

## DISCUSSION

### Arabidopsis flowering time is regulated by active demethylation of *FLC*

The downregulation of the floral repressor gene *FLC* is a key step for flowering in *Arabidopsis*, and this gene is therefore under complex regulation to ensure that plants transition to the reproductive stage of the life cycle when environmental conditions are optimal. While *FLC* has been reported to be regulated by many epigenetic mechanisms, including histone methylation, antisense lncRNAs, and chromatin looping, neither targeted methylation nor demethylation have been implicated to regulate *FLC* or autonomous flowering (Csorba et al., 2014; Crevillén et al., 2012; Wu et al., 2020). Flowering induced by vernalization involves reduced genomic DNA methylation levels (Finnegan et al., 1998), but no significant changes in methylation at *FLC*. *ddm1* mutants, which have decreased methylation levels, have delayed flowering times in long-day conditions caused by the hypermethylation of tandem repeats in the 5' region of *FWA* (Kakutani et al., 1995; Kakutani, 1997; Soppe et al., 2000; Kinoshita et al., 2006). *FLC,* on the other hand, does not appear to be significantly misregulated in methylation mutants. This correlates with the fact that we found no methylation at *FLC* in Col-0, so mutations in methylation pathways would not confer a significant change at these sites. However, given that we also found that small RNAs were targeted to the 5' region of *FLC* (Fig. 2F), there appears to be a complicated and previously unsuspected interplay between RdDM and

active demethylation at the *FLC* locus, especially as *FLC* was significantly misregulated in *drdd* plants.

To investigate whether the methylation in the 5' flanking region is directly affecting *FLC* transcription, we could ectopically methylate in a Col-0 background the regions we identified as hypermethylated in *drdd* (Fig. 2F). We have previously induced ectopic methylation using a transgene expressing an inverted hairpin construct that will target the RdDM pathway to the desired locus (Williams et al., 2015; Pignatta et al., 2018), and by doing so we can assay the sufficiency of this methylation in repressing *FLC* transcription. Additionally, we should investigate whether or not transcription of *COOLAIR*, an antisense lncRNA at *FLC* that interferes with *FLC* transcription (Wu et al., 2020) is affected in *drdd* as a potential mechanism for this regulation.

Despite the extensive research into regulation of *FLC*, we have shown that additional regulatory mechanisms are at play at this locus, beyond those that have already been described. The DRDD family of protein are actively demethylating the 5' flanking sequence of *FLC,* preventing hypermethylation that is correlated with decreased *FLC* transcription and early flowering.

**DNA demethylation is important for anther dehiscence and male fertility**

We found that *drdd* plants had reduced male fertility, due to a delay in anther dehiscence that correlated with a lack of reactive oxygen species in the anthers (Figure 3). While it has been recently reported that DNA demethylation is required for male fertility in pollen tube formation (Khouider et al., 2021), this is the first evidence for a role of active demethylation in anther dehiscence in Arabidopsis. In cotton, high temperature-induced anther indehiscence was linked to DNA methylation, and to reactive oxygen species accumulation (Ma et al., 2018; Zhang et al., 2020). However, these paper proposed a role for RdDM in that process, rather than a role for active demethylation.

The process of anther dehiscence requires cooperation between several pathways in the maturing anther, any of which could be hindered in *drdd* mutants. Secondary cell wall structures must be thickened through lignification, causing the anther to break open when it undergoes dehydration (Wilson et al., 2011). This lignification requires hydrogen peroxide, and mutants that affect the accumulation of reactive oxygen species (Shin et al., 2019) and/or lignin deposition in secondary cell wall thickening (Mitsuda et al., 2005; Huang et al., 2018; Nguyen et

al., 2019; Zhao et al., 2019; Dai et al., 2019) are known to exhibit delayed anther dehiscence. As *ros1* and *rdd* mutations were found to impair tracheary element differentiation and xylem development in *Arabidopsis* (Lin et al., 2020), it is possible that there is a common failure of lignin deposition in both the xylem and anther endothecium caused by a lack of active demethylation. This could potentially involve *VASCULAR-RELATED RECEPTOR-LIKE KINASE1* (*VRLK1*), overexpression mutants of which show thinner vascular cells with less lignin deposition and anther dehiscence defects (Huang et al., 2018).

In addition to lignification, anther dehiscence also requires jasmonic acid signaling (Stintzi and Browse, 2000; Sanders et al., 2000; Ishiguro et al., 2001; Park et al., 2002; von Malek et al., 2002; Ito et al., 2007; Ogawa et al., 2009; Xiao et al., 2014) and auxin signaling (Cecchetti et al., 2008, 2013; Song et al., 2018). The lack of active demethylation in *drdd* could have altered the expression of genes involved in these pathways.

Another candidate gene is *WRKY27,* one of the 74 WRKY transcription factors in Arabidopsis. Transgenic lines with various degrees of overexpression of *WRKY27* displayed varying levels of sterility caused by anther dehiscence delay, and were unresponsive to the application of exogenous jasmonic or gibberellic acid (Mukhtar et al., 2018). ROS1 was recently described to *cis*-regulate some defense genes through active demethylation in order to promote WRKY transcription factor binding in their promoter-regulatory regions (Halter et al., 2021). In anthers, therefore, it is possible that active demethylation by ROS1 could impact the ability of *WRKY27* to bind to its target genes.

Because our RNA-Seq data is from 3-week old leaf tissue, it is not surprising that it did not reveal a difference in expression of the major genes involved in the process of anther dehiscence. However, this does not preclude a tissue-specific expression difference in anthers for a gene necessary for anther dehiscence. Further RNA-Seq experiments in anthers to compare WT *DRDD* to both fertile and infertile *drdd* anthers would give additional insight as to which pathways are being misregulated. In conjunction with this, we can apply exogenous hydrogen peroxide, jasmonic acid, auxin, or gibberellic acid directly to sterile *drdd* inflorescences and observe if any of these rescue fertility (von Malek et al., 2002; Mukhtar et al., 2018; Dai et al., 2019).

More experiments are necessary to rule out these alternative pathways. As it is, the most parsimonious explanation remains that the lack of active demethylation in *drdd* mutants results

in misregulation of genes involved in reactive oxygen species accumulation, causing slower lignification of the secondary cell wall of the endothecium that then results in anther dehiscence delay. We therefore propose new roles for active demethylation in both the regulation of *FLC* and in male fertility.

## METHODS

### *pAGL61::DME* transgene

The transgene to restore DME expression in central cells was created by amplifying the promoter of *AGL61* (F primer: TCTAGAGGATCCAACCGATTTGACAA, R primer: TGATCGCTAGCTCCTCCTTTTGTA), the full genomic coding sequence (introns included) of *DME* (F primer: ATGAATTCGAGGGCTGATCCG, R primer: TTAGGTTTTGTTGTTCTTCAATTTGCTC) and cloning both fragments into pENTR-TOPO-D via Gibson assembly (overhang sequences are not included in the primers above). The assembled *pAGL61:DME* construct was then transferred to the binary vector pMDC99 (Curtis and Grossniklaus, 2003) using LR clonase.

### Plant material

Triple homozygous mutant *rdd* plants (Penterman et al., 2007) were transformed with *pAGL61:DME* via Agrobacterium floral dipping (Clough and Bent, 1998). Single-insertion transformants were selected and pollinated with *dme-2* (Choi et al., 2002) heterozygote mutant pollen to generate $F_1$ progeny heterozygous for mutations in all four *DRDD* demethylase genes. These quadruple heterozygous plants were self-fertilized, and over two subsequent generations of segregation the following genotypes were isolated, each homozygous for the *pAGL61:DME* transgene: *dme*, *rdd,* and *drdd* and *DRDD* WT segregants (to serve as a closely related WT control for downstream experiments). The selfed progeny of the initial plant of each genotype were used for all downstream experiments. To assess seed abortion, siliques were harvested after drying and seeds examined under a dissecting microscope.

### Flowering time assay

Plants were sown such that every row of the flats contained one plant of each genotype (WT *DRDD*, *dme*, *rdd*, and *drdd*) with the order iterating by one with each successive row. Flats were grown in a Conviron CMP6050 Control System at 22°C and 50% relative humidity, with 16 hours of 120 μMol light and 8 hours of darkness per day (long-day) or 8 hours of light and 16

hours of darkness per day (short-day). Starting at two weeks of age, all plants were visually inspected three times per week. The number of rosette leaves was recorded once a bolt was visible. Populations were compared using a one-way ANOVA with post-hoc Tukey test. The long day flowering time assay was repeated three times independently, and the short-day assay was repeated twice, using >30 biological replicates for each experiment.

**RNA-Seq**

RNA samples were isolated from 3-week old leaf samples using a QIAGEN RNeasy plant mini kit. 400 ng total RNA was used to generate RNA-seq libraries using a QIAGEN QIAseq Stranded mRNA Select Kit, with 13 cycles of amplification. An additional round of purification using QIAseq beads was performed to remove adapter dimers. RNA-seq was performed on an Illumina HiSeq 2500 using a 50 bp single-end protocol at the Whitehead Institute Genome Technology Core. All samples were multiplexed equally in two separate lanes to avoid batch effects in sequencing.

**Quantitative real-time PCR**

RNA was isolated from the 5th leaf of 21-day old plants (FLC qPCR) or 23-day old plants (DME qPCR) using TRIzol Reagent (Invitrogen) according to manufacturer's instructions. Genomic DNA was removed by treatment with Amplification-grade DNase I (Invitrogen). cDNA was prepared from 500-750ng RNA (standardized within each batch) with Superscript II Reverse Transcriptase (Invitrogen) according to the manufacturer's instructions, with polyadenylated transcripts selected for through use of an oligo-dT primer. qPCR was performed on a StepONE Plus Real-Time PCR system with Fast SYBR-Green PCR master mix (Applied Biosystems). *FLC* primer sequences were as previously described (Csorba et al., 2014): FLC_F: AGCCAAGAAGACCGAACTCA, and FLC_R: TTTGTCCAGCAGGTGACATC. *DME* region 1 primers were: DME_reg1F: CGAGGAAGGGCTGATTCCTTCAT, and DME_reg1R: TCCATGGCGAAAAACGTCTATCTC. DME region 2 primers were as previous described (Zhang et al., 2019): qDMEc6F: CATCGTCTCCTTGATGGTATGG, and qDMEc6R: CTTTCCCTCCACACTTCTGTT. Reactions were normalized to reference gene AT1G58050 (Czechowski et al., 2005): AT1G58050_F: CCATTCTACTTTTTGGCGGCT, and AT1G58050_R: TCAATGGTAACTGATCCACTCTGATG. All qPCR reactions were performed with technical triplicates. Cycling conditions were as follows: 95°C for 20s followed by 40 cycles of 95°C for 3s and 60°C for 30s. Relative fold change in expression with respect to

the geometric mean of the WT samples was determined using the $\Delta\Delta C_t$ method (Livak and Schmittgen, 2001). Standard error was calculated from combined standard deviations of both technical and biological replicates for each genotype.

**EM-Seq**

EM-Seq was performed as described (Vaisvila et al., 2021) with the NEBNext® Enzymatic Methyl-Seq Kit according to manufacturer's instructions, with the following specifications: 100ng of input DNA was sheared in 1x TE buffer (130µl total volume) using a Covaris S220 set for 200bp average fragment size (175W, 10% duty factor, 200 cycles/burst, 180 seconds). A 2.5x bead cleanup was performed after shearing, followed by elution in 51µl of $H_2O$. 50µl of this was used as input into the EM-Seq protocol, with 6 cycles of PCR.

**Alexander's staining**

Stain was made as previously described (Peterson et al., 2010). Anthers from the first open flower were dissected and placed in 10µl staining solution on a microscope slide. Slides were incubated in the dark for approximately 3 hours before being imaged on a Zeiss Axio Imager 2 with brightfield optics.

**DAB staining**

Staining was performed as previously described (Daudi and O'Brien 2012) with minor substitutions: DAB tetrahydrochloride was used in place of DAB non-acidified powder and added hydrochloric acid; whole inflorescences were vacuum-infiltrated in a 6-well plate with 5ml of DAB solution for 5 minutes at 10Hg. The plate was placed on an orbital shaker at 90rpm for 4 hours, covered in foil. Tissue was then transferred to the bleaching solution and boiled on a hot plate for 15 minutes, then the bleaching solution was removed and replaced with fresh bleaching solution. Tissue was incubated at room temperature for at least thirty minutes before imaging on a Zeiss Axio Imager 2.

**Accession Numbers**

All high-throughput sequencing data is deposited in NCBI GEO under accession GSE191307.

.

## REFERENCES


Abou-Elwafa, S.F., Büttner, B., Chia, T., Schulze-Buxloh, G., Hohmann, U., Mutasa-Göttgens, E., Jung, C., and Müller, A.E. (2011) Conservation and divergence of autonomous pathway genes in the flowering regulatory network of *Beta vulgaris.* Journal of Experimental Botany 62(10): 3359-3374.

Avila de Dios, E., Delaye, L. and Simpson, J. (2019) Transcriptome analysis of bolting in A. tequilana reveals roles for florigen, MADS, fructans and gibberellins. BMC Genomics 20: 473. https://doi.org/10.1186/s12864-019-5808-9

Cecchetti, V., Altamura, M.M., Falasca, G., Costantino, P., and Cardarelli, M. (2008) Auxin regulates *Arabidopsis* anther dehiscence, pollen maturation, and filament elongation. The Plant Cell 20: 1760-1774.

Cecchetti, V., Altamura, M.M., Brunetti, P., Petrocelli, V., Falasca, G., Ljung, K., Costantino, P., and Cardarelli, M. (2013) Auxin control Arabidopsis anther dehiscence by regulating endothecium lignification and jasmonic acid biosynthesis. The Plant Journal 74: 411-422.

Chitwood, J., Shi, A., Mou, B., Evans, M., Clark, J., Motes, D., Chen, P., and Hensley, D. (2016) Population structure and association analysis of bolting, plant height, and leaf erectness in spinach. HortScience 51(5): 481-486.

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.

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

Crevillén, P., Sonmez, C., Wu, Z., and Dean, C. (2013) A gene loop containing the floral repressor *FLC* is disrupted in the early phase of vernalization. EMBO J. 32: 140-148.

Csorba, T., Questa, J.I., Sun, Q., and Dean, C. (2014). Antisense COOLAIR mediates the coordinated switching of chromatin states at FLC during vernalization. Proc. Natl. Acad. Sci. 111: 16160–16165.

Curtis, M.D. and Grossniklaus, U. (2003). A Gateway Cloning Vector Set for High-Throughput Functional Analysis of Genes in Planta. Plant Physiol. 133: 462–469.

Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K., and Scheible, W.-R. (2005). Genome-Wide Identification and Testing of Superior Reference Genes for Transcript Normalization in Arabidopsis. Plant Physiol. 139: 5–17.

Dai, SY., Hsu, WH., and Yang, CH. (2019) The gene *ANTHER DEHISCENCE REPRESSOR (ADR)* controls male fertility by suppressing the ROS accumulation and anther cell wall thickening in *Arabidopsis*. Scientific Reports 9: 5112.

Daudi, A. and O'Brien, J. A. (2012). Detection of Hydrogen Peroxide by DAB Staining in Arabidopsis Leaves. Bio-protocol 2(18): e263. DOI: 10.21769/BioProtoc.263.

Erdmann, R.M., Satyaki, P.R.V., Klosinska, M., and Gehring, M. (2017) A Small RNA pathway mediates allelic dosage in endosperm. Cell Rep. 21(12): 3364-3372.

Finnegan, E.J., Genger, R.K., Kovac, K., Peacock, W.J., and Dennis, E.S. (1998) DNA methylation and the promotion of flowering by vernalization. Proc. Natl. Acad. Sci. U.S.A. 95(10): 5824-5829.

Ghorbani, F., Abolghasemi, R., Haghighi, M., Etemadi, N., Wang, S., Karimi, M., and Soorni, A. (2021) Global identification of long non-coding RNAs involved in the induction of spinach flowering. BMC Genomics 22: 704.

Halter, T., Wang, J., Amesefe, D., Lastrucci, E., Charvin, M., Rastogi, M.S., and Navarro, L. (2021) The Arabidopsis active demethylase ROS1 *cis*-regulates defence genes by erasing DNA methylation at promoter-regulatory regions. eLife 10: e62994. DOI: https://doi.org/10.7554/eLife.62994

Herrero, M.P. and Johnson, R.R. (1980) High temperature stress and pollen viability of maize. Crop Science 20: 796-800.

Huang, C., Zhang, R., Gui, J., Zhong, Y., and Li, L. (2018) The receptor-like kinase AtVRLK1 regulates secondary cell wall thickening. Plant Physiol. 177(2): 671-683. doi: 10.1104/pp.17.01279.

Ishiguro, S., Kawai-Oda, A., Ueda, J., Nishida, I, and Okada, K. (2001) The *DEFECTIVE IN ANTHER DEHISCENCE1* gene encodes a novel phospholipase A1 catalyzing the initial step of jasmonic acid biosynthesis, which synchronizes pollen maturation, anther dehiscence, and flower opening in Arabidopsis. The Plant Cell 13: 2191-2209.

Ito, T., Ng, KH., Lim, TS., Yu, H., and Meyerowitz, E.M. (2007) The homeotic protein AGAMOUS controls late stamen development by regulating a jasmonate biosynthetic gene in *Arabidopsis*. The Plant Cell 19:3516-3529.

Jiang, D., and Berger, F. (2017) DNA replication-coupled histone modification maintains Polycomb gene silencing in plants. Science 357(6356): 1146-1149.

Kakutani, T., Jeddeloh, J.A., and Richards, E.J. (1995) Characterization of an Arabidopsis thaliana DNA hypomethylation mutant. Nucleic Acids Res. 23(1): 130-137.

Kakutani, T. (1997) Genetic characterization of late-flowering traits induced by DNA hypomethylation mutation in *Arabidopsis thaliana*. The Plant Journal 12(6): 1447-1451.

Khouider, S., Borges, F., LeBlanc, C., Ungru, A., Schnittger, A., Martienssen, R., Colot, V., and Bouyer, D. (2021) Male fertility in *Arabidopsis* requires active DNA demethylation of genes that control pollen tube function. Nat. Commun. 12(1): 410. https://doi.org/10.1038/s41467-020-20606-1

Kinoshita, Y., Saze, H., Kinoshita, T., Miura, A., Soppe, W.J.J., Koornneef, M., and Kakutani, T. (2006) Control of *FWA* gene silencing in *Arabidopsis thaliana* by SINE-related direct repeats. The Plant Journal 49(1): 38-45.

Le, T.N., Schumann, U., Smith, N.A., Tiwari, S., Au, P.C.K., Zhu, Q.H., Taylor, J.M., Kazan, K., Llewellyn, D.J., Zhang, R., Dennis, E.S., and Wang, M.B. (2014) DNA demethylases target promoter transposable elements to positively regulate stress responsive genes in *Arabidopsis*. Genome Biology 15: 458-476.

Li, Z., Jiang, D., and He, Y. (2018) FRIGIDA establishes a local chromosomal environment for *FLOWERING LOCUS C* mRNA production. Nature Plants 4: 836-846.

Lin, W., Sun, L., Huang, RZ., Liang, W., Liu, X., He, H., Fukuda, H., He, XQ., and Qian, W. (2020) Active DNA demethylation regulates tracheary element differentiation in *Arabidopsis*. Sci. Adv. 6(26): eaaz2963.

Lister, R., O'Malley, R., 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.

Livak, K.J. and Schmittgen, T.D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2-ddCt Method. Methods 25: 402–408.

Ma, Y., Min, L., Wang, M., Wang, C., Zhao, Y., Li, Y., Fang, Q., Wu, Y., Xie, S., Ding, Y., Su, X., Hu, Q., Zhang, Q., Li, X., and Zhang, X. (2018) Disrupted genome methylation in

response to high temperature has distinct affects on microspore abortion and anther indehiscence. The Plant Cell. 30(7): 1387-1403.

Matzke, M.S., and Mosher, R.A. (2014) RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. Nature Reviews Genetics 15: 394-48.

Michaels, S.D., and Amasino, R.M. (1999) *FLOWERING LOCUS C* encodes a novel MADS domain protein that acts as a repressor of flowering. The Plant Cell 11(5): 949-956.

Mitsuda, N., Seki, M., Shinozaki, K., and Ohme-Takagi, M. (2005) The NAC transcription factors NST1 and NST2 of *Arabidopsis* regulate secondary wall thickenings and are required for anther dehiscence. *The Plant Cell* 17: 2993-3006.

Mukhtar, M.S., Liu, X. and Somssich, I.E. (2018) Elucidating the role of WRKY27 in male sterility in Arabidopsis, Plant Signaling & Behavior, 13:11, e1363945, DOI: 10.1080/15592324.2017.1363945

Nguyen, H.T.K., Hyoung, S., Kim, H.J., Cho, K.M., and Shin, J.S. (2019) The transcription factor γMYB2 acts as a negative regulator of secondary cell wall thickening in anther and stem. Gene 702: 158-165.

Nie, S., Li, C., Xu, L., Wang, Y., Huang, D., Muleke, E.M., Sun, X., Xie, Y., and Liu, L. (2016) *De novo* transcriptome analysis in radish (*Raphanus sativus* L.) and identification of critical genes involved in bolting and flowering. BMC Genomics 17: 389.

Ogawa, M., Kay, P., Wilson, S., and Swain, S.M. (2009) ARABIDOPSIS DEHISCENCE ZONE POLYGALACTURONASE1 (ADPG1), ADPG2, and QUARTET2 Are Polygalacturonases Required for Cell Separation during Reproductive Development in *Arabidopsis*. The Plant Cell 21(1): 216-233.

Ou, CG., Mao JH., Liu, LJ., Li, CJ., Ren, HF., Zhao, ZW., and Zhuang, FY. (2017) Characterising genes associated with flowering time in carrot (Daucus carota L.) using transcriptome analysis. Plant Biol. (Stuttg.) 19(2): 286-297. doi: 10.1111/plb.12519.

Park, JH., Halitschke, R., Kim, H.B., Baldwin, I.T., Feldmann, K.A., and Feyereisen, R. (2002) A knock-out mutation in allene oxide synthase results in male sterility and defective wound signal transduction in Arabidopsis due to a block in jasmonic acid biosynthesis. Plant J. 31(1): 1-12. doi: 10.1046/j.1365-313x.2002.01328

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

Peterson, R., Slovin, J.P., and Chen, C. (2010). A simplified method for differential staining of aborted and non-aborted pollen grains. International Journal of Plant Biology 1:e13.

Pignatta, D., Novitsky, K., Satyaki, P.R.V., and Gehring, M. (2018) A variably imprinted epiallele impacts seed development. PLoS Genetics 14(11): e1007469. https://doi.org/10.1371/journal.pgen.1007469.

Prasad P.V.V., Boote K.J., Allen L.H. Jr., Sheehy J.E. and Thomas, J.M.G. (2006) Species, ecotype and cultivar differences in spikelet fertility and harvest index of rice in response to high temperature stress. Field Crops Research 95, 398–411.

Rang Z.W., Jagadish S.V.K., Zhou Q.M., Craufurd P.Q. and Heuer, S. (2011) Effect of heat and drought stress on pollen germination and spikelet fertility in rice. Environmental and Experimental Botany 70, 58–65.

Sanders, P.M., Lee, P.Y., Biesgen, C., Boone, J.D., Beals, T.P., Weller, E.W., and Goldberg, R.B. (2000) The Arabidopsis *DELAYED DEHISCENCE1* gene encodes an enzyme in the jasmonic acid synthesis pathway. The Plant Cell 12: 1041-1061.

Shin, J.S., Kim, S.Y., So, W.M., Noh, M., Yoo, K.S., and Shin, J.S. (2019) Lon domain-containing protein 1 represses thioredoxin y2 and regulates ROS levels in *Arabidopsis* chloroplasts. FEBS Letters 594: 986-994.

Snider, J.L., Oosterhuis, D.M., and Kawakami, E.M. (2011) Diurnal pollen tube growth rate is slowed by high temperature in field-grown *Gossypium hirsutum* pistils. Journal of Plant Physiology 168: 441-448.

Song, S., Chen, Y., Liu, L, See, Y.H.B., Mao, C., Gan, Y., and Yu, H. (2018) OsFTIP7 determines auxin-mediated anther dehiscence in rice. Nature Plants 4: 495-504.

Soppe, W.J., Jacobsen, S.E., Alonso-Blanco, C., Jackson, J.P., Kakutani, T., Koornneef, M., and Peeters, A.J. (2000) The late flowering phenotype of fwa mutants is caused by gain-of-function epigenetic alleles of a homeodomain gene. Mol Cell. 6(4): 791-802.

Steffen, J.G., Kang, I.-H., Portereiko, M.F., Lloyd, A., and Drews, G.N. (2008). AGL61 Interacts with AGL80 and Is Required for Central Cell Development in Arabidopsis. Plant Physiol. 148: 259–268.

Stintzi, A., and Browse, J. (2000) The *Arabidopsis* male-sterile mutant, *opr3*, lacks the 12-oxophytodienoic acid reductase required for jasmonate synthesis. PNAS 97(19): 10625-10630.

Udvardi, M.K., Czechowski, T., Scheible, WR. (2008) Eleven golden rules of quantitative RT-PCR. The Plant Cell 20(7): 1736-1737.

Vaisvila, R., Ponnaluri, V.K.C., Sun, Z., Langhorst, B.W., Saleh, L., Guan, S., Dai, N., Campbell, M.A., Sexton, B.S., Marks, K., Samaranayake, M., Samuelson, J.C., Church, H.E., Tamanaha, E., Corrêa Jr., I.R., Pradhan, S., Dimalanta, E.T., Evans Jr., T.C., Williams, L., and Davis, T.B. (2021) Enzymatic methyl sequencing detects DNA methylation at single-base resolution from picograms of DNA. Genome Res. 31(7): 1280-1289.

von Malek, B., van der Graaff, E., Schneitz, K., and Keller, B. (2002) The Arabidopsis male-sterile mutant dde2-2 is defective in the ALLENE OXIDE SYNTHASE gene encoding one of the key enzymes of the jasmonic acid biosynthesis pathway. Planta. 216(1): 187-192. doi: 10.1007/s00425-002-0906-2

Wilson, Z.A., Song, J., Taylor, B., and Yang, C. (2011) The final split: the regulation of anther dehiscence. Journal of Experimental Botany 62(5): 1633-1649.

Wu, Z., Fang, X., Zhu, D., and Dean, C. (2020) Autonomous pathway: *FLOWERING LOCUS C* repression through an antisense-mediated chromatin-silencing mechanism. Plant Physiology 182(1): 27-37.

Xiao, Y., Chen, Y., Charnikhova, T., Mulder, P.P., Heijmans, J., Hoogenboom, A., Agalou, A., Michel, C., Morel, J.B., Dreni, L., Kater, M.M., Bouwmeester, H., Wang, M., Zhu, Z., Ouwerkerk, P.B. (2014) OsJAR1 is required for JA-regulated floret opening and anther dehiscence in rice. Plant Mol Biol. 86(1-2):19-33. doi: 10.1007/s11103-014-0212-y.

Yamamuro, C., Miki, D., Zheng, Z., Ma, J., Wang, J., Yang, Z., Dong, J., and Zhu, J.K. (2014) Overproduction of stomatal lineage cells in *Arabidopsis* mutants defective in active DNA demethylation. Nature Communications 5:4062.

Yang, H., Berry, S., Olsson, T.S.F., Hartley, M., Howard, M., and Dean, C. (2017) Distinct phases of Polycomb silencing to hold epigenetic memory of cold in *Arabidopsis.* Science 357(6356): 1142-1145.

Zhang, C., Hung, YH., Rim, H.J., Zhang, D., Frost, J.M., Jang, H., Liu, F., Xiao, W., Iyer, L.M., Aravind, L., Zhang, X.-Q., Fischer, R.L., Huh, J.H., Hsieh, T.-F. (2019) The catalytic

core of DEMETER guides active DNA demethylation in *Arabidopsis*. PNAS 116(35): 17563-17571.
Zhang, M., Zhang, X., Guo, L., Qi, T., Liu, G., Feng, J., Shahzad, K., Zhang, B., Li, X., Wang, H., Tang, H., Qiao, X., Wu, J., & Xing, C. (2020) Single-base resolution methylome of cotton cytoplasmic male sterility system reveals epigenomic changes in response to high-temperature stress during anther development. Journal of experimental botany, 71(3): 951–969. https://doi.org/10.1093/jxb/erz470
Zhao, SQ., Li, WC., Zhang, Y., Tidy, A.C., and Wilson, Z.A. (2019) Knockdown of Arabidopsis ROOT UVB SENSITIVE4 disrupts anther dehiscence by suppressing secondary thickening in the endothecium. Plant Cell Physiol. 60(10):2293-2306. doi: 10.1093/pcp/pcz127

# Chapter Four

# CONCLUSIONS AND FUTURE DIRECTIONS

**THESIS SUMMARY**

The work presented in this thesis increased our understanding of the mechanism by which transcription of the DNA glycosylase *ROS1* is regulated and the importance of the entire *DEMETER* family of DNA glycosylases as a whole in the proper development of the plant.

I determined that *ROS1* is downregulated at the transcriptional level in response to decreased methylation levels in the proximal promoter. Through a CRISPR-Cas9-mediated deletion of the methylated region at the endogenous locus, I showed that it contains a methylation-sensitive silencer sequence. In the absence of this silencer sequence, transcription of *ROS1* was increased in wildtype, *rdr2*, and *drm2* backgrounds. This led to a new model for regulation at *ROS1* in which this silencer sequence represses *ROS1*, but DNA methylation of the silencer sequence lessens the repressive effect. Thus, the more methylation at the silencer sequence, the less the repressive effect, and the more transcription will occur at *ROS1*. I also determined that the more distal region 5' of *ROS1* contains a regulatory element that is likely necessary for *ROS1* transcription to occur, as deleting this region led to the same level of *ROS1* transcription as is detected in *ROS1*$^{+/-}$ heterozygotes.

Examining the impact of active demethylation as a whole, we generated and characterized a quadruple mutant of *dme;ros1;dml2;dml3* in the somatic tissues of the plant. The *drdd* mutants flowered significantly earlier than wildtype plants and other demethylase mutants. Upon investigation of the floral repressor *FLC*, one of the genes that influences the autonomous flowering pathway, we detected hypermethylation in the 5' flanking sequence. We also determined that transcription of *FLC* was significantly downregulated in *drdd* plants. This suggests that active demethylation is normally required to remove methylation from the *FLC* locus to prevent early silencing of *FLC.* This is the first description of a role for active demethylation in the regulation of *FLC*, a gene that integrates many environmental and genetic inputs to regulate flowering time. Additionally, we investigated a defect in fertility in *drdd* plants and determined that it was caused by a delay in anther dehiscence. This delay was correlated with a concurrent delay in the accumulation of reactive oxygen species, which are necessary for lignification of the secondary cell wall in order for the anther to break open and release pollen. A delay in anther dehiscence can lead to a block to self-fertilization, as the pistil will elongate and brush past the anthers before they release pollen. While active demethylation has been proposed to play a role in male fertility in pollen tube growth (Khouider et al., 2021), there have been no

reports of demethylation being necessary in anther development or dehiscence. This thesis underscores the necessity of understanding the balance between *de novo* methylation and active demethylation in the entirety of the plant.

## FUTURE DIRECTIONS

This thesis has introduced several new lines of inquiry into the regulation of active demethylation and its balance with *de novo* methylation. Now that we have generated the *pAGL61:DME* central cell rescue transgene, we can use this to generate the full complement of mutants of the DRDD genes. In particular, we can compare the four triple mutants (*dme;ros1;dml2, dme;ros1;dml3;, dme;dml2;dml3;* and *ros1;dml2;dml3*) to the quadruple mutant to investigate the sufficiency of each DNA glycosylase. Given the high degree of redundancy of targets of these proteins, identification of hypermethylated loci in single mutants only describes the set of targets for which the single demethylase is necessary. However, by comparing the methylation of the *ddd* mutant genome to the *drdd* genome, for example, we can determine all possible targets of ROS1*,* rather than the set of ROS1-specific targets we can identify from the *ros1* mutant. This investigation will aid the community in our understanding of how *ROS1* targets are defined and prioritized, without the other DNA demethylases acting redundantly.

Additionally, it is becoming clearer that tissue-specific regulation of *ROS1* and active demethylation contributes to several aspects of development*.* Two studies have found phenotypic defects in *rdd*, in the tracheary elements of the venous system and in the stomata of the leaf epidermis (Yamamuro et al., 2014; Lin et al., 2020). Both of these phenotypes were also detected in *ros1*, but to a lesser extent. First, we should investigate the tracheary elements and stomata formation in *drdd* to determine if *DME* can redundantly function at the relevant loci. We should also examine the other structures in which I found high levels of expression of the *pROS1*:GUS transgene, particularly the hydathodes and the root columella.

The hydathodes of the leaf are involved in the release of water from the leaf, and are entry sites for pathogens (van Hulten et al., 2019). The *ROS1* promoter drove high levels of expression in this tissue, which may indicate that there is a constitutive role of ROS1 in plant pathogen defense in this tissue. To test this, we could complement the *ros1* mutant with a *ROS1* transgene expressed solely in the hydathodes and test pathogen susceptibility. If active

demethylation is playing a role in defense in this particular tissue, the increased susceptibility in *ros1* (Yu et al., 2012) will be rescued by the transgene. Alternatively, the role of *ROS1* in the hydathode may be linked to the differentiation of protodermal cells into the stomata, trichomes, and hyathodes (Torii, 2021), especially given the overproduction of stomata in *ros1* mutants (Yamamuro et al., 2014). Yet another hypothesis is that as ROS1 is necessary for the proper development of xylem, ROS1 is thus enriched in the hydathodes which contain the xylem ends (Yagi et al., 2021a). Further experiments are necessary to determine hypothesis or combination of hypotheses is correct.

The columella is a tissue in the root cap that is involved in gravitropism, and has the highest levels of CHH methylation in any tissue studied to date (Kawakatsu et al., 2016). The sedimentation of starch-filled granules within columella cells triggers growth on the side of root on which the granules accumulate (Su et al., 2020). *PIN3* proteins, efflux carriers for the plant hormone auxin, are redistributed in the columella following reorientation of the root, guiding the auxin to direct growth of the root downwards (Su et al., 2020). This redistribution is dependent upon the regulatory kinase *PINOID* (*PID*), which is known to be regulated by DNA methylation and demethylation dynamics (Ariel et al., 2014). Because of this, we should investigate whether gravitropism is disrupted in *ros1*, *rdd*, or *drdd* mutants.

Prior investigations into the role of *ROS1*, *DML2*, and *DML3* in gravitropism indicated that while *ROS1* was induced during this process, *DML2* and *DML3* were repressed (Ariel et al., 2014). This difference between the DNA demethylases implies a unique role for *ROS1* in auxin response, which could have implications outside of the root cap as well. Auxin dynamics have been implicated in the function of the columella (Ariel et al., 2014), as well as the development of hydathodes (Yagi et al., 2021b), stomata (Wei et al., 2021), and anthers (Cecchetti et al., 2008). As several of these structures have developmental roles for DNA demethylation, it would also be instructive to investigate the interactions of *ROS1* and auxin. Auxin was recently found to upregulate a reporter transgene driven by the *ROS1* promoter in young leaves (Bennett et al., 2021). Given the ubiquity of auxin in plant development, there may be a wider role for DNA demethylation than is currently known.

Auxin-induced upregulation of *ROS1* could be tested in the various regulatory deletion mutants that I generated in this thesis. If one of these deletions was found to interfere with auxin-mediated regulation of *ROS1*, this region could then be analyzed for the presence of auxin

response elements, as well as binding sites for the Auxin Response Factor (ARF) transcription factor family.

The regulation of *PID* by DNA demethylation also provides a template for a potential mechanism of regulation of *ROS1*. Methylation silences the nearby lncRNA *APOLO*, allowing formation of a chromatin loop encompassing both *APOLO* and *PID* that represses transcription (Ariel et al., 2014). Demethylation by RDD proteins in response to auxin leads to transcription of *APOLO*, thus breaking the chromatin loop and allowing transcription of *PID* to occur as well (Ariel et al., 2014). In this case, methylation represses a chromatin loop that represses transcription. At *ROS1*, a chromatin loop was detected near *ROS1* in wildtype plants, although not within the silencer region (Liu et al., 2016). It is possible that this chromatin loop could promote *ROS1* transcription, and that silencer methylation could affect this loop as it does the loop at *APOLO*. This would result in methylation-sensitive silencing of *ROS1*. Alternatively, there could be a loop that only forms in the hypomethylated RdDM mutants that is able to repress *ROS1* transcription.

Another hypothesis is that the methylation at *ROS1* could interfere with the transcription of a repressive lncRNA, similar to *COOLAIR* transcription at *FLC*. Transcription of the antisense lncRNA *COOLAIR* directly leads to a reduction of transcription of *FLC* (Swiezewski et al., 2009). A repressive lncRNA could therefore be at play at the *ROS1* locus; transcription of the lncRNA could have a repressive effect on *ROS1*, and methylation could silence this putative repressor, thus also leading to methylation-sensitive silencing.

Finally, methylation levels could be affecting the binding of methylation-sensitive transcription factors. DAP-Seq could therefore be used in both wildtype and RdDM mutants to determine if there is a difference in transcription factor binding near the *ROS1* locus. To differentiate between the previous hypotheses will require additional experiments in the regulation of *ROS1*. Performing targeted 3C at *ROS1* in wildtype and methylation mutant backgrounds will reveal whether a chromatin loop is affected by local methylation levels. We can also delete the looping site identified by Liu et al. (2016) and assay the resulting *ROS1* expression. Furthermore, with the advancements in genomic sequencing, it is feasible to sequence the transcriptome at high enough depth to capture even rare species, such as a lowly-expressed RNA. By comparing the transcriptomes near *ROS1* in both wildtype and methylation mutants, we could determine if there is a methylation-sensitive lncRNA at the *ROS1* locus.

Multiple experiments in this thesis have addressed the importance of the equilibrium between the addition of methylation by the RdDM pathway and its removal by active demethylation. *FLC* is an example of a locus at which the demethylases outcompete RdDM, maintaining a hypomethylated state in the wildtype promoter (Chapter 3, Figure 2D-F). We are able to identify similar loci using the RNA-Seq and EM-Seq data gathered from the *drdd* mutants. On the other hand, *SDC* typifies loci at which RdDM outcompetes demethylation, resulting in heavy methylation and silencing that is only lifted in the *rdr2;ROS1*$^{5'\Delta 0.8}$ double mutant, an RdDM mutant with overexpression of ROS1 (Chapter 2, Figure 7D). Comparing these two sets of loci – those at which demethylases prevent accumulation of methylation by RdDM and those at which RdDM overwhelms demethylation – will give us insight into the establishment, maintenance, and functionality of this epigenetic equilibrium.

## REFERENCES


Ariel, F., Jegu, T., Latrasse, D., Romero-Barrios, N., Christ, A., Benhamed, M., and Crespi, M. (2014) Noncoding transcription by alternative RNA polymerases dynamically regulates an auxin-driven chromatin loop. Mol Cell. 55(3): 383-396. doi: https://doi.org/10.1016/j.molcel.2014.06.011.

Bennett, M., Cleaves, K., and Hewezi, T. (2021) Expression patterns of DNA methylation and demethylation genes during plant development and in response to phytohormones. Int J Mol Sci 22(18): 9681. doi: 10.3390/ijms22189681.

Cecchetti, V., Altamura, M.M., Falasca, G., Costantino, P., and Cardarelli, M. (2008) Auxin regulates *Arabidopsis* anther dehiscence, pollen maturation, and filament elongation. The Plant Cell 20: 1760-1774.

Lin, W., Sun, L., Huang, RZ., Liang, W., Liu, X., He, H., Fukuda, H., He, XQ., and Qian, W. (2020) Active DNA demethylation regulates tracheary element differentiation in *Arabidopsis*. Sci. Adv. 6(26): eaaz2963.

Liu, C., Wang, C., Wang, G., Becker, C., Zaidem, M., and Weigel, D. (2016) Genome-wide analysis of chromatin packing in *Arabidopsis thaliana* at single-gene resolution. *Genome Res.* 26: 1057-1068.

Kawakatsu, T., Stuart, T., Valdes, M., Breakfield, N., Schmitz, R.J., Nery, J.R., Urich, M.A. Han, X., Lister, R., Benfey, P.N., and Ecker, J.R. (2016) Unique cell-type-specific patterns of DNA methylation in the root meristem. Nature Plants 2: 16058. https://doi.org/10.1038/nplants.2016.58.

Khouider, S., Borges, F., LeBlanc, C., Ungru, A., Schnittger, A., Martienssen, R., Colot, V., and Bouyer, D. (2021) Male fertility in *Arabidopsis* requires active DNA demethylation of genes that control pollen tube function. Nat. Commun. 12(1): 410. https://doi.org/10.1038/s41467-020-20606-1.

Swiezewski, S., Liu, F., Magusin, A., and Dean, C. (2009) Cold-induced silencing by long antisense transcripts of an *Arabidopsis* Polycomb target. Nature 462(10): 799-803. doi:10.1038/nature08618

Su, SH., Keith, M.A., and Masson, P.H. (2020) Gravity signaling in flowering plant roots. Plants 9(10): 1290. https://doi.org/10.3390/plants9101290.

Torii, K.U. (2021) Stomatala development in the context of epidermal tissues. Annals of Botany 128(2): 137-148. https://doi.org/10.1093/aob/mcab052.

van Hulten, M., Chatterjee, S., and van den Burg. H. (2019) Infection assay for *Xanthomonas campestris* pv. *campestris* in *Arabidopsis thaliana* mimicking natural entry via hydathodes. Ann Bot. 128(2): 137-148. doi: 10.1093/aob/mcab052.

Wei, H., Jing, Y., Zhang, L., Kong, D. (2021) Phytohormones and their crosstalk in regulating stomatal development and patterning. Journal of Experimental Botany 72(7): 2356-2370. https://doi.org/10.1093/jxb/erab034.

Yagi, H., Nagano, A.J., Kim, J., Tamura, K., Mochizuki, N., Nagatani, A., Matsushita, T., and Shimada, T. (2021a) Fluorescent protein-based imaging and tissue-specific RNA-seq analysis of Arabidopsis hydathodes. Journal of Experimental Biology 72(4): 1260-1270. https://doi.org/10.1093/jxb/eraa519.

Yagi, H., Tamura, K., Matsushita, T., and Shimada, T. (2021b) Spatiotemporal relationship between dynamics and hydathode development in *Arabidopsis* leaf teeth. Plant Signal Behav. doi: 10.1080/15592324.2021.1989216.

Yamamuro, C., Miki, D., Zheng, Z., Ma, J., Wang, J., Yang, Z., Dong, J., Zhu, J.K. (2014) Overproduction of stomatal lineage cells in *Arabidopsis* mutants defective in active DNA demethylation. Nature Communications 5: 4062.

Yu, A., Lepère, G., Jay, F., Wang, J., Bapaume, L., Wang, Y., Abraham, A.L., Penterman, J., Fischer, R.L., Voinnet, O., and Navarro, L. (2012) Dynamics and biological relevance of DNA demethylation in *Arabidopsis* antibacterial defense. PNAS 110(6): 2389-2394.