From Recall to Reset: The Role of DNA (De)methylation in Modulating *Arabidopsis thaliana* Immune Memory

Adam Hannan Parker

A thesis submitted to the University of Sheffield
for the degree of Doctor of Philosophy

School of Biosciences
Faculty of Science
University of Sheffield

Submitted: 25th October 2024

Acknowledgements

First, thanks must go to my primary supervisor, Jurriaan Ton, who has met every idea, piece of data, and question I had during my PhD with utmost enthusiasm, insight, and curiosity. Thanks to Lisa Smith and Steven Rolfe, my co-supervisors, who have provided me with countless creative ideas that have significantly improved the quality of my research. Furthermore, thanks all for providing me opportunities to supervise and teach students, present and share my research, and engage with scientists across both academic and commercial institutions. Without your guidance and encouragement, I would not have had such a fruitful and fulfilling PhD.

I would like to thank the White Rose BBSRC Doctoral Training Partnership in Mechanistic Biology for funding my PhD and offering valuable training and networking opportunities within a supportive community of PhD students across Sheffield, Leeds, and York.

During my PhD, I have been fortunate to mentor and work with talented Summer and Master students. In the second year of my PhD, Kwok Yin Man undertook a Master of Research project, and helped develop a hydroponic system for the early characterisations of *XVE:ROS1-YFP* transformant plants. His exceptional attention to detail and passion for plant epigenetics resulted in a fruitful Master project that has inspired new lines of work in the lab. In my third year, Adam Mitchel was awarded a summer studentship to help characterise a selection of novel epimutagenic lines used during my PhD. Adam quickly embraced and utilised complicated molecular biology techniques and generated a wealth of valuable data. Later, George Cawood joined as an Integrated Master student, where we worked together to expand on the work of Adam Mitchel. George quickly demonstrated his outstanding scientific talent and has since been offered a PhD in the lab. Thank you all for your help and for showing so much interest and enthusiasm in my PhD project. Best of luck in your future endeavours!

Thank you to everyone in the C51/C207 lab group who has been a part of my PhD journey. I am grateful for, though certainly not limited to, the ideas, cakes, coffee, encouragement, and insightful discussions (science or otherwise) we have shared. I would like to specifically thank members who have invested significant time into my PhD project. Firstly, David Pardo and Peijun Zhang for their countless hours of technical support, which made otherwise unwieldy experiments possible. Chia-Nan Tao, for transforming my molecular biology skill set in the lab, generously taking time from her own PhD project to demonstrate techniques, and providing excellent company at scientific conferences. Joost Stassen for equipping me with the computational skills required to complete this PhD. And of course, Sam

Wilkinson, who, perhaps unknowingly, has been an unofficial mentor throughout my PhD, offering extremely valuable (running) advice, project ideas, and enthusiasm throughout.

In the third year of my PhD, I spent 3 months on a professional placement in the Netherlands. Many thanks to Karin Posthuma and Manos Domazakis at Enza Zaden for providing me with this exciting and valuable opportunity. I would like to extend this to the entire Phytopathology team at Enza Zaden for making me feel so welcome a despite my short internship. Likewise, I would like to thank Jake Harris at the University of Cambridge, who hosted me for several weeks in his lab to generate novel epimutagenic Arabidopsis lines, advancing and improving key aspects of my research. Finally, thank you to Helen Hipperson and the Genomics Facility at the University of Sheffield for providing training and access to their Oxford Nanopore Sequencing Technology, which was critical for several of the experiments outlined in this thesis.

The friendships and community I have built during the four years of my PhD in Sheffield have undoubtedly been a key factor in the success of this project. Over the four years I have lived with my closest friends, first the 253 lot (Annie, Alex, Ella, Wes, & Matt T), then the 'Khartoum Network' (Matt T, Matt J, & Danny), and now in a 'cozy communal home' (Jolie & Louis). These homes have given me the warmth, love, laughter, food, and stability that enabled me to navigate the more challenging times in my PhD. Beyond my housemates, I am incredibly fortunate to be part of a wonderful community of friends in Sheffield that have provided an escape from work. Together we have walked in the Peak District, climbed up rocks, swam in freezing water, actively engaged in protests, and, of course, spent plenty of time in warm pubs. Thank you all for providing me with this much-needed balance during my PhD.

I am deeply grateful to my partner Daisy, who has provided me with emotional love and support throughout my PhD. Thank you for all the time we have shared together over the past five years, whether it be on exciting holidays, time outside, or cozy evenings inside. You are a constant source of happiness that has made my life in Sheffield during this PhD infinitely more enriching. Thank you for being there, always.

And most importantly, I would like to thank my family. Mum and Dad, thank you for always being so proud, encouraging, loving, and supportive of everything I do. Thank you for the sacrifices you have made to give me a nurturing upbringing and an education that made this PhD possible in the first place. Liban, Tom, and Safi, thank you for your unconditional love and companionship, offering much-needed breaks from the PhD with concerts, Bananagrams, meals, holidays, and so much more. And finally, thank you to my granddad Chris and great uncle Mark, from whom I presumably (epi)genetically inherited a deep curiosity for science.

Declaration and Acknowledgement of Collaborative Work

The work presented in this thesis is original and has not been submitted for examination at the University of Sheffield or any other institution. The candidate acknowledges the University's Guidance on the Use of Unfair Means (www.sheffield.ac.uk/ssid/unfair-means) and confirms that all work included in this thesis is solely the candidate's own. The candidate acknowledges the use of the generative AI tool ChatGPT for feedback on the phrasing and grammar of text in line with this Guidance. In some Chapters , text and/or figures have been adapted from previous publications in which the candidate was a co-author and contributed to the work during their PhD. Furthermore, in the case of exceptionally large or lengthy experiments, additional co-authors have contributed to aspects of the work, such as in data collection or analysis. Therefore, at the beginning of each Chapter, all co-authors and their contributions are listed, as well as details of previously published works that have been included in the Chapter.

Previously published work that has been included in this thesis is listed below:

Hannan Parker, A.; Wilkinson, S. W. and Ton, J. (2022). Epigenetics: A Catalyst of Plant Immunity against Pathogens. New Phytologist, 233(1): 66–83.

Furci, L.; Pascual-Pardo, D.; Tirot, L.; Zhang, P.; Hannan Parker, A. and Ton, J. (2023). Heritable Induced Resistance in Arabidopsis Thaliana: Tips and Tools to Improve Effect Size and Reproducibility. Plant Direct, 7(8): e523.

Wilkinson, S. W.; Hannan Parker, A.; Muench, A.; Wilson, R. S.; Hooshmand, K.; Henderson, M. A.; Moffat, E. K.; Rocha, P. S. C. F.; Hipperson, H.; Stassen, J. H. M.; López Sánchez, A.; Fomsgaard, I. S.; Krokene, P.; Mageroy, M. H. and Ton, J. (2023). Long-Lasting Memory of Jasmonic Acid-Dependent Immunity Requires DNA Demethylation and ARGONAUTE1. Nature Plants, 9(1): 81–95.

Thesis Summary

Plants can retain memories of past stress events, allowing them to launch an induced resistance (IR) response to defend against recurring attacks by pests and pathogens. In *Arabidopsis thaliana* (Arabidopsis), the DNA demethylase REPRESSOR OF SILENCING 1 (ROS1) is essential for effective IR against diverse stress types. Therefore, in this thesis, I employ a chemically inducible gene construct for controlled activation of ROS1 to better understand the role of DNA (de)methylation in the regulation of immune memory that underpins IR. I demonstrate that transient increases in ROS1 activity alone can establish immune memory against biotrophic pathogens. This was associated with DNA demethylation and reduced small RNA (sRNA) expression in the chromosome arms, along with transcriptional activation of genes involved in salicylic acid-mediated defences and the DNA damage response pathway. Genome-wide identification of genes with promoter regions targeted for ROS1-mediated DNA demethylation uncovered master regulators of both pathways.

Further characterisation of the epigenetic changes associated with this ROS1-driven immune memory revealed a contrasting increase in DNA methylation and sRNA accumulation in (peri)centromeric regions of the genome. Chemically inhibiting DNA methyltransferases that facilitate DNA (re)methylation in these regions prolonged ROS1-driven immune memory and enhanced the associated IR phenotypes against a bacterial pathogen. Thus, regulation of DNA methylation in the (peri)centromeric regions may function to balance the establishment, maintenance, and erasure of immune memory in Arabidopsis.

Finally, I show that altering the developmental stage and frequency of ROS1 induction can cause unique genome-wide DNA (de)methylation patterns and corresponding stress-specific IR phenotypes. Therefore, a single DNA demethylase, ROS1, can generate stress-specific epigenetic immune memory against diverse stress types. Overall, the results and the (epi)genetic tools developed in this thesis lay a solid foundation for further research into the epigenetic control of plant-environment interactions and the future utilisation of immune memory in crop protection schemes.

Table of Contents

List of Figures

List of Supplementary Figures

Supplementary Figure 4.1 Overlaps of differentially expressed loci/genes (DEGs) from whole transcriptome sequencing. 134
Supplementary Figure 4.2 Green leaf area (GLA) of Col-0 and *XVE:ROS1-YFP* lines (L5 and L7) after treating twice E2 or DMSO. 134
Supplementary Figure 4.3 Genome browser snapshots of *cis*-regulated loci. 135
Supplementary Figure 4.4 Characterisation of the protein-coding gene targets of upregulated miRNAs in E2-treated *XVE:ROS1-YFP* lines (L5 and L7). 136
Supplementary Figure 4.5 Verification of antagonistic *cis*-regulation in the promoter of *NPR1* by ROS1 and RdDM. 137

Supplementary Data

Due to the size of the Supplementary Data tables, all files have been provided as separate excel workbooks in a Chapter-specific archive folder. All file names are listed below:

Supplementary Data 3.1 hIR examples in other plants
Supplementary Data 3.2 WGBS sequencing statistics
Supplementary Data 3.3 all identified DMRs

Supplementary Table 4.1 WGBS read stats
Supplementary Table 4.2 RNA-seq read stats
Supplementary Table 4.3 reference genes
Supplementary Table 4.4 DEGs
Supplementary Table 4.5 GO terms DEGs
Supplementary Table 4.6 sRNA-seq read stats
Supplementary Table 4.7 miRNA targets with GO terms
Supplementary Table 4.8 DMRs
Supplementary Table 4.9 DECs
Supplementary Table 4.10 cis regulated loci
Supplementary Table 4.11 TF enrichment PlantRegMap
Supplementary Table 4.12 expression NPR1 components
Supplementary Table 4.13 expression DNA damage

Supplementary Data 5.1 - Nanopore read stats

Supplementary Data 6.1 - Nanopore read stats
Supplementary Data 6.2 - 5mC DMRs
Supplementary Data 6.3 - GO enrichment
Supplementary Data 6.4 - 5mC comparison 2 wk and 4 wk Arabidopsis

Abbreviations and Acronyms

4mC	4-Methylcytosine
5-Aza	5-Azacytidine
5hmC	5-Hydroxymethylcytosine
5mC	5-Methylcytosine
6mA	N6-Methyladenosine
AGO	ARGONAUTE
allC	All cytosines
ANOVA	Analysis of variance
Arabidopsis	*Arabidopsis thaliana*
ATM	ATAXIA-TELANGIECTASIA MUTATED
ATP	adenosine triphosphate
ATR	ATR AND RAD3-RELATED
BABA	β-Aminobutyric
BER	Base excision repair
BF	Bright field
BP	Biological process (in relation to gene ontology)
bp	base pair
cDNA	Complementary DNA derived from RNA
CG/CHG/CHH	Cytosine sequence contexts where H is any base but G
ChIP-seq	Chromatin immunoprecipitation with sequencing
Chr	Chromosome
CLD	Compact letter display
CLSY	CLASSY
CMT	CHROMOMETHYLASE
Co-IP	Co-immunoprecipitation
Col-0	Arabidopsis thaliana ecotype Columbia
CPM	Counts per million
CRC	Chromatin remodelling complex
DAS	Days after sowing
DCL3	DICER-LIKE
DDM1	DECREASED DNA METHYLATION 1
DDR	DNA damage response
DEC	Differentially expressed small RNA cluster
DEG	Differentially express locus/gene
dH2O	distilled water
DME	DEMETER
DML	DEMETER LIKE
DMLs	Differentially methylated locations
DMR	Differentially methylated region
DMSO	Dimethyl sulfoxide
DMTs	DNA methyltransferases
DNA	Deoxyribonucleic acid
dpi	day post infection
DRM2	DOMAINS REARRANGED METHYLTRANSFERASE
DSBs	Double stranded breaks
dsRNA	Double stranded RNAs

E2	17β-estradiol
epiQTL	epigenetic quantitative trait locus
epiRIL	epigenetic recombinant inbred line
ERGs	Environmentally response genes
ETI	Effective triggered immunity
FC	Fold change
FDR	False discovery rate
flg22	Flagelin 22
gDNA	Genomic DNA
GFP	Green fluorescent protein
GLA	Green leaf area
GMO	Genetically modified organism
GO	Gene ontology
H3Kac	Acetylation of a lysine in histone variant H3
H3Kme	Methylation of a lysine residue in histone variant H3
HCA	Hierarchical cluster analysis
h-IR	Heritable induced resistance
h-IS	Heritable induced susceptibility
HMW	High molecular weight
Hpa	*Hyaloperonospora arabidopsidis*
HR	Hypersensitive response
hyper-DMRS	Hypermethylated differentially methylated regions
hypo-DMRs	Hypermethylated differentially methylated regions
IBM	INCREASE IN BONSAI METHYLATION
IR	Induced resistance
IS	Induced susceptibility
ISR	Induced systemic resistance
JA	Jasmonic acid
kb	Kilobase
KYP	KRYPTONITE
L5	*XVE:ROS1-YFP* line 5 (Col-0 background)
L7	*XVE:ROS1-YFP* line 7 (Col-0 background)
lncRNA	Long non-coding RNA
Mb	Megabase
MET	METHYLTRANSFERASE
MF	Molecular function (in relation to gene ontology)
miRNA	MicroRNA
mRNA	Messenger RNA
ncRNA	Non-coding RNA
NHP	N-hydroxypipecolic acid
NLR	Nucleotide binding-Leucine rich repeat
NPR1	NONEXPRESSER OF PR GENES 1
NRPE1	NUCLEAR RNA POLYMERASE D1B subunit of Pol V
nt	Nucleotide
ONT	Oxford Nanopore Technology
p	statistical *p*-value
p_{adj}	Adjusted *p*-value
PAMP	Pathogen associated molecular pattern
Pc	*Plectosphaerella cucumerina*

PCA	Principal component analysis
pMock / pPst	Parental mock/Pst treatment
Pol	RNA POLYMERASE
PR	PATHOGENESIS RELATED
Pst	*Pseudomonas syringae* pv. *tomato* DC3000
Pst-GFP	*Pst* expressing Green Fluorescent Protein
Pst-Lux	*Pst* expressing *LuxCDABE* operon
PTGS	Posttranscriptional gene silencing
PTI	Pattern triggered immunity
PTM	Post translational modification
qPCR	Quantitative polymerase chain reaction
q-value	FDR-corrected *p*-value
R gene	Resistance gene
RdDM	RNA-directed DNA methylation
RH	Relative humidity
RLP	RECEPTOR LIKE PROTEIN
RNA	Ribonucleic acid
ROIs	Regions of interest
ROS	Reactive oxygen species
ROS1	REPRESSOR OF SILENCING 1 / DML1
RT-qPCR	Reverse transcriptase-qPCR
SA	Salicylic acid
SAR	Systemic acquired resistance
seq	Sequencing (of DNA/RNA)
Silwet	Silwet L-77 (LEHLE SEEDS, VIS-30) surfactant
siRNA	Small interfering RNA
sRNA	Small RNA
SSBs	Single stranded DNA breaks
ssRNA	Single stranded RNA
SUVH	SUPPRESSOR OF VARIEGATION 3-9 HOMOLOG
TE	Transposable element
TET	TEN ELEVEN TRANSLOCATION
TF	Transcription factor
TGA	TGACG MOTIF-BINDING FACTOR
TPM	Transcripts per million
TRXh5	THIOREDOXIN H-TYPE 5
TSS	Transcriptional start site
TTS	Transcriptional termination site
WGBS	Whole genome bisulfite sequencing
WRKY	WRKY DNA-BINDING PROTEIN
WT	Wild-type
XVE	Chimeric protein formed of DNA binding domain of LexA (X), the acidic transactivating domain of VP16 (V), and the carboxyl region of the human estrogen receptor (E)
YFP	Yellow fluorescent protein

Chapter 1. General Introduction

Partially adapted from:

Hannan Parker, A.; Wilkinson, S. W. and Ton, J. (2022). Epigenetics: A Catalyst of Plant Immunity against Pathogens. New Phytologist, 233(1): 66–83.

Authors:

Hannan Parker, A[1]; Wilkinson, S. W. [1]; Ton, J[1].

Affiliations:

[1] Plants, Photosynthesis and Soil, School of Biosciences, The University of Sheffield, Western Bank, Sheffield, S10 2TN, UK

Author Contributions:

AHP (the candidate) and JT proposed the original idea and scope of the review. All authors contributed to the writing of the article and/or production of the Figures in the original publication. All material in this Chapter, which is not in the publication, was created by AHP. JT reviewed and provided comments on the Chapter. Funding for the project came from studentships/grants awarded to AHP, SWW, and JT.

1.1 The Impact of Plant Diseases on Society and the Environment

Plant pests and diseases have profound impacts on food security, biodiversity, and economic resources. Globally, 45% of food loss can be contributed to pest and pathogen infestations (FAO, 2022). Examples of historic outbreaks include the spread of the potato pathogen *Phytophthora infestans*, which causes potato late blight, and lead to the Irish Potato Famine in the 1840s, which drove mass emigration from Ireland and the death of approximately 1 million people (Yoshida et al., 2013). Additionally, in the first half of the 20th century, the fungal pathogen *Fusarium oxysporum* f. sp. *cubense* (*Foc*) race 1, which causes Fusarium wilt in banana (also known as Panama disease), entirely collapsed the export trade of 'Gros Michel' bananas, which were the dominant export variety at the time (Ploetz, 2005). As a result, plantations were forced to switch a more resistant variety of banana from the 'Cavendish' subgroup, which comprises over 50% of global banana production today (Lescot, 2020). However, newly emerging causal agents of Fusarium wilt once again threaten the global production of Cavendish bananas (Ploetz, 2006; Roberts et al., 2024). More recently, ash dieback, caused by the fungal pathogen *Hymenoscyphus fraxineus*, has spread rapidly throughout Europe since the 1990s, and has killed up to 85% of European ash (*Fraxinus excelsior*) in some woodlands. In Britain alone, this has an estimated economic cost of £14.8 billion (Coker et al., 2019; Combes et al., 2024). Furthermore, no tree species have been identified which can replace the ecosystem function of European ash, thus posing a severe ecological crisis in Europe (Mitchell et al., 2016).

As human-driven climate change worsens, extreme weather events and shifting temperatures are expected to increase the severity and spread of plant pests and diseases, potentially leading to even greater crop losses (Bebber, 2015; Delgado-Baquerizo et al., 2020; Ristaino et al., 2021; Chaloner et al., 2021; Singh et al., 2023). For instance, between 2018 and 2019, a series of extreme weather events affecting East Africa created favourable conditions for the desert locust (*Schistocerca gregaria*) to swarm. As a result, from 2019 and 2022, Kenya experienced one of its worst desert locust infestations in over 70 years, affecting approximately 20% of the country's land surface and damaging over 190,000 km^2 of cropland (Salih et al., 2020; Kimathi et al., 2020; Liu et al., 2024).

Current management of pests and diseases heavily relies on the use of chemical pesticides. In 2022 alone, 3.7 million tonnes of pesticides were used in agriculture (FAO, 2024), which can have harmful effects on human health and environment (Nicolopoulou-Stamati et al., 2016; Sharma et al., 2019; Tang et al., 2021; Gerken et al., 2024). To reduce the persistent and widespread application of chemical pesticides, and to counteract the emergence and spread of pests and diseases, protection strategies that provide, enhance, or activate internal

plant defence mechanisms need to be developed. The introduction of resistance in crop varieties can be achieved through breeding programmes, genetic engineering, and/or genetic modification (Hickey et al., 2019; Jones et al., 2024; Sundström et al., 2024). However, many genetic resistance breeding initiatives focus on the introgression of single resistance (*R*) genes, which offer a narrow range of protection against selected isolates of biotrophic pathogens. Moreover, in monoculture systems, this genetically determined disease resistance can exert significant selective pressure on diseases to evolve mechanisms that evade these resistance genes (Mundt, 2014; Hickey et al., 2019). Therefore, by advancing our understanding of the plant immune system beyond the genetic basis and functioning of *R* genes, we can develop new strategies that aim to enhance broader internal defence mechanisms of plants. Consequently, these approaches can be integrated with the selection and 'stacking' of conventional *R* genes to develop more resilient and sustainable crops that are better equipped to withstand diverse pests and diseases.

1.2 Overview of the Plant Immune System

1.2.1 Innate Immunity

To survive in their natural environment, plants rely on their immune system for protection against pests and diseases. The plant immune system includes a multifaceted innate component, which is genetically determined, as well as an adaptive component, which is shaped by environmental influences. Unlike animals, where innate and adaptive immunity are governed by distinct mechanisms and specialised cell types, the innate and adaptive components of the plant immune system are tightly intertwined (Jones and Dangl, 2006; Mauch-Mani et al., 2017).

In addition to constitutively active mechanisms, the plant's innate immune system can detect danger signals from the environment, triggering chemical and mechanical defences to mitigate the onset of stress (Wilkinson et al., 2019). These defence-eliciting signals are typically molecules derived from the attacking organism itself (non-self recognition; Bigeard et al., 2015; Boutrot and Zipfel, 2017), or molecules that are produced as a consequence of cellular damage (damaged-self recognition; Li et al., 2020). The resulting pattern-triggered immunity (PTI) protects against the majority of potentially harmful organisms and involves a multitude of signalling mechanisms, including pattern recognition receptors (PRRs; Macho and Zipfel, 2014) and phytohormones (Pieterse et al., 2012). The two most extensively studied phytohormones involved in plant immunity are jasmonic acid (JA) and salicylic acid (SA). JA-dependent defences are mostly effective necrotrophic pathogens and herbivores, whereas

SA-dependent defences facilitate defence against (hemi-)biotrophic pathogens (Glazebrook, 2005; Pieterse et al., 2012).

Specialised attackers have evolved strategies to suppress PTI by deploying effector molecules that supress defence signalling and contribute to pathogen virulence. In response, plants have evolved specific *R* genes that encode R proteins – usually nucleotide-binding, leucine-rich repeat (NLR) proteins (Jones et al., 2016) – that detect immune-suppressive effectors either through direct interaction or by "guarding" internal proteins (Remick et al., 2023; Jones et al., 2024). In the latter case, R proteins detect alterations made by the effector to its target protein, thereby triggering an immune response (Remick et al., 2023). The resulting effector-triggered immunity (ETI) provides strong protection against attackers, often leading to localised cell death in the affected area, known as the hypersensitive response (HR) (Jones and Dangl, 2006). This effectively hinders the invasion of biotrophic attackers that depend on living tissue for survival. However, plant NLR proteins have been shown to have roles in defending against a range of taxonomically diverse organisms, including viruses, bacteria, fungi, oomycetes, parasitic plants, and herbivores (Ngou et al., 2022).

Whilst PTI and ETI have historically been considered distinct components of the innate immune system, two landmark reports using chemically inducible gene constructs and the *Arabidopsis thaliana* – *Pseudomonas syringae* pv. *tomato* DC3000 (Arabidopsis-*Pst*) pathosystem have demonstrated that ETI works, in part, by potentiating PTI (Ngou et al., 2021; Yuan et al., 2021). Two key findings from these independent studies are: (1) ETI cannot be effectively established in Arabidopsis plants with mutations in core PTI receptor genes; and (2) chemically induced expression of the *AvrRpt2* effector alone does not result in HR and triggers weak or absent defence responses, such as ROS production or callose deposition. However, induced *AvrRpt2* significantly enhances PTI responses when plants are simultaneously exposed to the pathogen-associated molecular pattern (PAMP) flagellin 22 (flg22) resulting in HR phenotypes (Ngou et al., 2021; Yuan et al., 2021). Thus, PTI is required for effective ETI responses and is the driving force preventing pathogen invasion in Arabidopsis (Jones et al., 2024). Interestingly, ETI is also required for effective PTI, as Arabidopsis plants with mutations in NLR-encoding genes have weaker PTI responses to flg22 (Tian et al., 2021). Therefore, PTI and ETI are tightly intertwined and involve mutually dependent components in the plant's innate immune system that functions as an essential and immediate defence barrier against harmful organisms.

When an NLR gene fails to recognise an attacker, PTI and/or ETI may be supressed to a lower basal level that is insufficient to prevent disease or infestation. However, this reduced response can still help slow down the colonisation of a wide range of attackers, and

is known as basal or quantitative resistance (Poland et al., 2009; Corwin and Kliebenstein, 2017; Wilkinson et al., 2019).

1.2.2 Induced Resistance: a Form of Plant Adaptive Immunity

Genetically susceptible plants can respond to biotic stress by raising their pre-existing level of innate and/or basal immunity. This acquired/induced resistance (AR / IR) is a form of phenotypic plasticity as it allows the plant to change its defence phenotype in response to specific environmental signals. Accordingly, the mechanisms that facilitate IR are regarded as the adaptive component of the plant immune system (Mauch-Mani et al., 2017; De Kesel et al., 2021). IR is usually systemic and can be triggered by a variety of stimuli, including pathogen-associated molecular patterns (PAMPs), herbivore-induced volatiles from neighbouring plants, root colonization by beneficial microbes, and the application of synthetic chemicals (Prime-A-Plant Group et al., 2006; Heil and Ton, 2008; Pieterse et al., 2014; Yassin et al., 2021). In most cases, IR develops after an initial burst of immune activity resulting in prolonged up-regulation of inducible genes (“induced constitutive response”) and/or increased transcriptional sensitivity of defence genes (“induced primed response”) (Figure 1.1a) (Prime-A-Plant Group et al., 2006; Reimer-Michalski and Conrath, 2016; Wilkinson et al., 2019). Following the recovery from this initial burst of activity, the phase in which plants sustain constitutive defence activation or ‘remember’ to mount a faster and/or stronger response to subsequent stress represents a period of effective immune memory (Figure 1.1b). In contrast, induced resistance (IR) refers to the actual defensive response, such as enhanced resistance, which occurs when the plant encounters a second stress (e.g., pathogen attack) after having established immune memory. Below, I explore the factors and mechanisms that drive the establishment, maintenance, and eventual erasure of immune memory in plants.

While IR can reach full protection if the defence response is activated before effector-driven immune suppression by the attacker, it generally offers partial protection, slowing down the colonisation of pests or pathogens (Ahmed et al., 2011). Well-known examples of IR responses include systemic acquired resistance (SAR), where whole-plant immunity is heightened following a localised pathogen attack, and induced systemic resistance (ISR), where plants become more resistant to pathogens following colonisation of roots by plant-beneficial microbes (Fu and Dong, 2013; Pieterse et al., 2014). However, since IR primarily involves the amplification of PTI responses (Prime-A-Plant Group et al., 2006; Reimer-Michalski and Conrath, 2016; Wilkinson et al., 2019), distinguishing the pathways and mechanisms controlling IR from those of innate immunity is challenging. For instance, IR could be considered a backup mechanism for NLR-mediated ETI, which also enhances PTI to counteract pathogen success (Ngou et al., 2021; Yuan et al., 2021). Despite this, some IR-

specific mechanisms and pathways have been described. For example, Arabidopsis plants infected with pathogens show strong up-regulation of *FLAVIN-DEPENDENT MONOOXYGENASE 1 (FMO1)*, which encodes a protein that catalyses the biochemical conversion of pipecolic acid (Pip) into N-hydroxypipecolic acid (NHP). NHP is a mobile compound which transcriptionally amplifies immune signals and is thus an essential regulator of SAR (Hartmann et al., 2018; Y.-C. Chen et al., 2018; Yildiz et al., 2021). Furthermore, as mentioned above, IR has a distinct temporal aspect that necessitates mechanisms governing immune memory (Figure 1.1b). Therefore, while epigenetic mechanisms contribute to different forms of plant immunity, a growing body of evidence highlights their particular importance and specific roles in immune memory and IR, which will be discussed in more detail below (Luna et al., 2012, 2014; Rasmann et al., 2012; López Sánchez et al., 2016; Wilkinson et al., 2023; Gallusci et al., 2023; Sheikh et al., 2023; Lee et al., 2023; Harris et al., 2023).

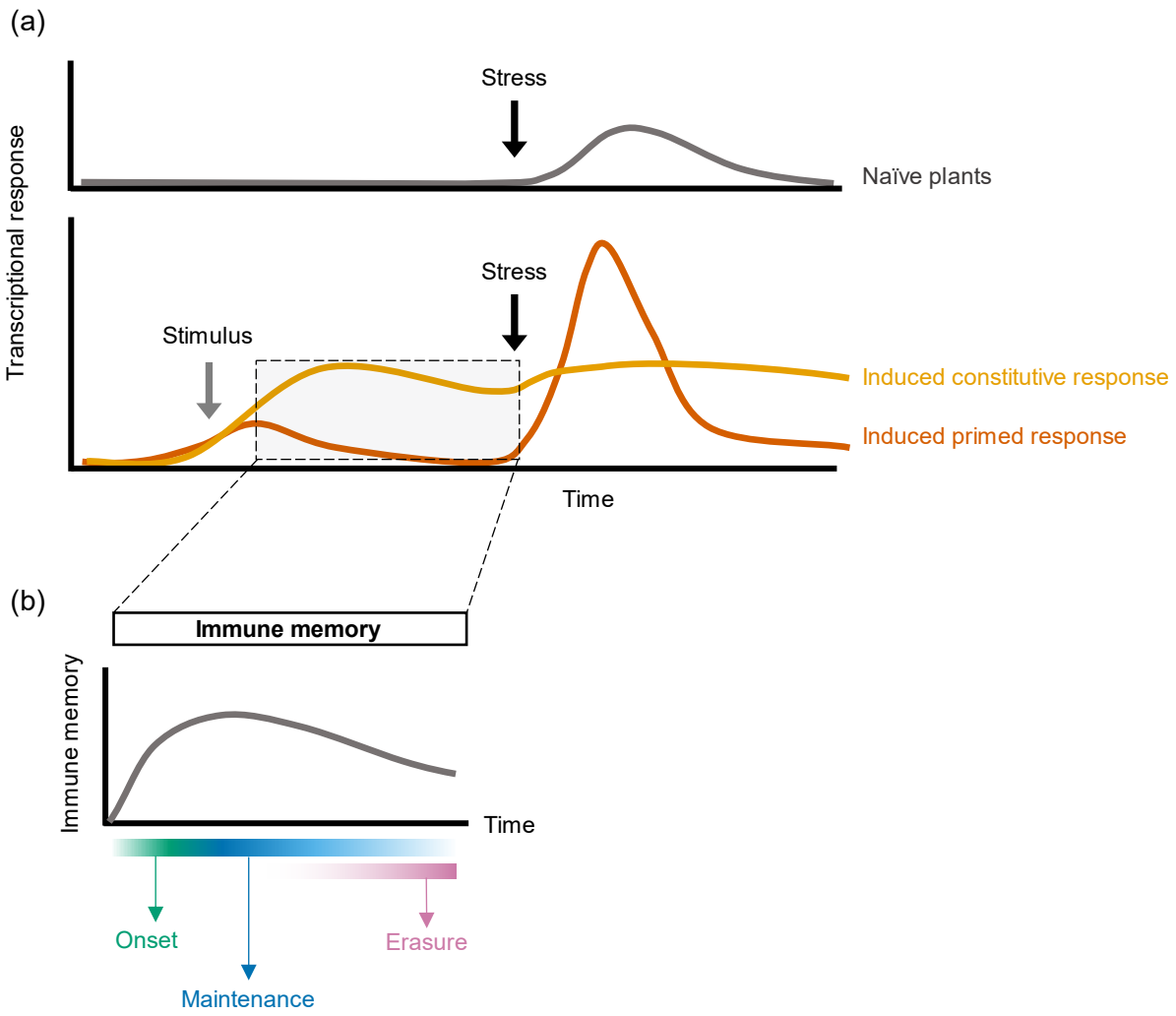


Figure 1.1 Schematic representing the transcriptional responses associated with stress memory. (a) Transcriptional responses of inducible defence-related genes to environmental stress ("stress"; black arrow) in naïve plants (grey line; top graph) and plants that have established stress memory (orange lines; bottom graph). Upon exposure to a stimulus (grey arrow) that mimics or triggers a non-lethal stress response (e.g., localised infection, mild abiotic stress, immune-stimulating chemicals), stress memory genes either show prolonged upregulation following induction (light orange; "induced constitutive response"), or a faster and/or stronger transcriptional induction upon secondary exposure to stress (dark orange; "induced primed response"). 'Induced constitutive' and 'induced primed' transcriptional responses are also referred to as type I and type II transcriptional memory (Bäurle, 2018; Oberkofler et al., 2021). (b) Model representing the distinct phases of immune memory: onset (green), maintenance (maintenance), and eventual erasure (pink). Immune memory is defined as the period between the recovery phase of plants exposed to a memory-inducing stimulus (grey arrow) and subsequent stress (black arrow), during which defence-related genes either maintain prolonged upregulation or retain the capacity to launch a primed transcriptional response to stress.

1.3 Plant Epigenetics: an Overview of the Main Mechanisms

The meaning of the term ‘epigenetics’ has changed since it was first used by Waddington in 1942, who defined epigenetics as the mechanisms by which genes and their products determine phenotypes (Waddington, 1942). Today, epigenetics is more commonly defined as the study of changes in gene function that are mitotically and/or meiotically stable, and that occur independently of changes in DNA sequence (Armstrong, 2014). However, the term is also used to describe chromatin and/or DNA modifications that are not strictly heritable (Crisp et al., 2016; Wilkinson et al., 2019). Thus, I broadly define epigenetic mechanisms here as the cellular processes responsible for these modifications, without a strict requirement of heritability.

1.3.1 Different Forms of DNA Methylation in Plants

DNA methylation at the carbon 5 position (C5) of the pyrimidine ring of the DNA base cytosine (C), resulting in the formation of 5-methylcytosine (5mC), is one of the most extensively studied epigenetic mechanisms in plants. This modification plays a critical role in repressing the activity of transposable elements (TEs), invading viral DNA, and other potentially harmful genetic elements (Erdmann and Picard, 2020). Other forms of DNA methylation include N6-adenine (6mA) and 4-methylcytosine (4mC), which have only recently gained attention in plant research due to technological advancements that allow for improved detection and genome-wide mapping of these modifications (Liang et al., 2018; Zhang et al., 2018; Walker et al., 2021; Kong et al., 2022; Y. Zhang et al., 2023). However, as 5mC is the most abundant and well-characterised form of DNA methylation in plants, I will focus this introduction on 5mC and refer to it below as DNA methylation for simplicity.

1.3.2 RNA Directed DNA methylation (RdDM)

In plants, DNA methylation occurs at three different sequence contexts: CG, CHG, and CHH (H = bases A, T, or C). *De novo* DNA methylation is primarily established at unmethylated sequences by an RNA-dependent DNA methylation (RdDM) pathway, which is mechanistically connected to post-transcriptional gene silencing (PTGS). This pathway involves small RNAs (sRNAs), whose production is dependent on RNA polymerase II (Pol II), and can thus be referred to as Pol II-dependent RdDM (Cuerda-Gil and Slotkin, 2016; Erdmann and Picard, 2020; Sigman et al., 2021). During this process, Pol II actively transcribes a genetic element in the genome, such as a transposable element (TE) or foreign transgene (Panda and Slotkin, 2020). The resulting single-stranded RNA (ssRNA) transcripts are then recognized and converted into double-stranded RNAs (dsRNAs) by RNA-dependent RNA polymerase 6

(RDR6), which are subsequently processed by DICER-LIKE proteins (DCL1-4) to produce 21–24-nucleotide (nt) sRNAs (Cuerda-Gil and Slotkin, 2016; Erdmann and Picard, 2020; Sigman et al., 2021). These sRNAs, primarily those of 21–22 nt in size, are loaded onto ARGONAUTE 1 (AGO1), which reinforces the degradation of TE transcripts and subsequent production sRNAs (Matzke and Mosher, 2014; Cuerda-Gil and Slotkin, 2016). Simultaneously, some 21, 22, or 24-nt sRNAs are loaded onto members of the AGO4-clade protein family (AGO4, AGO6, or AGO9), which, by sequence complementarity, direct RNA polymerase V (Pol V) to Pol II-transcribed genomic regions (Sigman et al., 2021). This recruitment can occur either in *cis* (at the same locus) or in *trans* (at a different locus, such as TEs from the same family). The Pol V recruitment triggers the production of scaffold transcripts that interact with sRNA-loaded AGO4-clade proteins, facilitating the recruitment of DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2) and localised deposition of *de novo* DNA methylation (Erdmann and Picard, 2020; Sigman et al., 2021). Upon activation, Pol V is continually recruited by SUPPRESSOR OF VARIEGATION 3-9 HOMOLOG 2 (SUVH2) and SUVH9, which bind to existing DNA-methylated regions, and so form a self-perpetuating cycle of RdDM (Liu et al., 2014; Johnson et al., 2014). Furthermore, DNA methylation interacts with other histone-modifying enzymes to help establish a transcriptionally repressive heterochromatic environment (see Section 1.3.5 below) (Du et al., 2015; Fang et al., 2022). Once *de novo* RdDM has established, sRNA biogenesis from Pol II diminishes, and Pol IV is recruited to the heterochromatic RdDM regions (Law et al., 2013). Transcripts produced by Pol IV are converted into dsRNAs by RDR2, which are further cleaved by DICER-LIKE 3 (DCL3) into 24-nt sRNAs. Similar to Pol II-dependent RdDM, these 24-nt sRNAs associate with AGO4-clade proteins to recruit Pol V and DRM2, facilitating the deposition of DNA methylation (Matzke and Mosher, 2014; Cuerda-Gil and Slotkin, 2016; Erdmann and Picard, 2020; Sigman et al., 2021). Taken together, Pol II- and Pol IV-dependent RdDM are important mechanisms driving the initiation and reinforcement of DNA methylation in plant genomes, respectively.

1.3.3 DNA Methyltransferases

Plants have evolved distinct DNA methyltransferases that establish and maintain DNA methylation in different cytosine contexts (Chen and Li, 2004; Goll and Bestor, 2005). DRM2, as discussed in the previous paragraph, mediates *de novo* DNA methylation via the RdDM pathway in all sequence contexts, but has a preference for CHH and CHG substrates (Cao and Jacobsen, 2002; Stroud et al., 2013; Fang et al., 2021). The methyltransferase DRM1 is a close paralog to DRM2 but is only expressed at very specific stages of sexual reproduction (Jullien et al., 2012). In contrast, METHYLTRANSFERASE 1 (MET1) and CHROMOMETHYLASE 3 (CMT3) maintain DNA methylation in CG and CHG contexts,

respectively (Jean Finnegan and Dennis, 1993; Lindroth et al., 2001; Kankel et al., 2003; Fang et al., 2022). Whereas CMT2 maintains DNA methylation in CHH contexts, and to a lesser extent CHG contexts (Stroud et al., 2014). In some cases, these methyltransferases also have the capacity to initiate *de novo* DNA methylation at certain genomic loci in an RdDM-independent manner (Zubko et al., 2012; Du et al., 2012; Stroud et al., 2014; Yaari et al., 2019). Whilst CMT2 and RdDM-directed DRM2 activity both contribute to CHH methylation in the genome, CMT2 is mostly restricted to long heterochromatic TEs that reside near the centromeres (pericentromeres), whereas RdDM-directed DRM2 predominantly targets short TEs embedded in the chromosome arms (Zemach et al., 2013; Harris et al., 2024). Mechanistically, the histone variant H1 restricts access of the RNA-directed DNA methylation (RdDM) machinery to pericentromeric regions; however, this restriction can be circumvented by DNA methyltransferases through the action of the nucleosome remodeler DEFECTIVE IN RNA-DIRECTED DNA METHYLATION 1 (DDM1) (Zemach et al., 2013; Harris et al., 2024).

1.3.4 DNA Demethylation by DNA Demethylases

To prevent the spread of DNA methylation from target TEs to neighbouring gene sequences, bifunctional 5-methylcytosine (5mC) DNA glycosylase/lyases in plants catalyse the first steps in a base excision repair pathway which results in active demethylation of cytosines (Zhang et al., 2022). There are four 5mC DNA glycosylase/lyases in Arabidopsis: REPRESSOR OF SILENCING 1 (ROS1), DEMETER (DME), DEMETER-LIKE 2 (DML2), and DML3. Of these, *ROS1* is the most highly expressed DNA demethylase (Zhang et al., 2022) and actively antagonises RdDM at gene-TE boundaries in vegetative tissues (Tang et al., 2016; Schumann et al., 2017; Halter et al., 2021). The *ROS1* gene itself is under positive transcriptional control by RdDM (Lei et al., 2015; Williams et al., 2015), creating a negative feedback loop to ensure tight homeostasis of DNA methylation in the genome, known as the ‘methylstat’ . Unlike the other DNA demethylases in Arabidopsis, DME is essential for the development of gametophytes, and consequently, *dme* mutants have a seed-abortion phenotype (Choi et al., 2002; Schoft et al., 2011; Park et al., 2017). *DME* is also expressed in vegetative tissues but has a smaller influence on genome-wide DNA methylation patterns compared to *ROS1* (Zeng et al., 2021). Likewise, *DML2* and *DML3* are expressed moderately in vegetative tissues, and have a partially redundant role with ROS1, although some studies suggest unique targets for DML2/3 compared to ROS1 (Ortega-Galisteo et al., 2008; Zhang et al., 2022; Kong et al., 2022). Across the plant kingdom, DME and ROS1 are evolutionary conserved, with paralogous genes present in algae through to higher flowering plants (Pei et al., 2019). In contrast, *DML2* is only found in some dicots, whereas *DML3* is limited to dicots and monocots (Pei et al., 2019).

1.3.5 Histone Variants and Chromatin Remodelers

Enzymatic activities that alter chromatin states are considered epigenetic mechanisms, as the resulting changes can influence gene expression and be inherited during cell division (Henikoff et al., 2004; Borg et al., 2021; Jamge et al., 2023). Chromatin density plays a crucial role in transcriptional activity, as it controls the accessibility of DNA to transcription factors. Heterochromatin consists of tightly packed, transcriptionally silent regions of the genome, often associated with repetitive sequences and TEs, whereas euchromatin is less dense and enriched with transcriptionally active genes (Jacob et al., 2009; Ruiz-Velasco and Zaugg, 2017). Chromatin density is dependent on the distribution and composition of nucleosomes (Jamge et al., 2023), each of which consists of two copies of four histone proteins (H2A, H2B, H3, H4), wrapped by ~147 bp of DNA (Luger et al., 1997). These nucleosomes are connected by variable stretches of internucleosomal DNA associated with linker histone H1 (Rutowicz et al., 2019). Chromatin remodelling complexes (CRCs) can break histone-DNA contacts using adenosine triphosphate (ATP) to assemble, edit, and redistribute nucleosomes (Han et al., 2015; Clapier et al., 2017). Eukaryotic CRCs include four major families based on their catalytic ATPase subunit: SWItch/Sucrose Non-Fermentable (SWI/SNF), imitation SWI (ISWI), chromodomain and helicase-like domain (CHD), and Inositol Requiring 80 (INO80) (Bhadouriya et al., 2021).

Histone variants and their post-translational modifications (PTMs) shape the chromatin landscape in plants (Jamge et al., 2023) both directly and indirectly through interactions with other epigenetic machinery such as DNA methyltransferases (see Section 1.3.5 below) (Du et al., 2015). These PTMs include phosphorylation, ubiquitination, SUMOylation, acetylation and methylation, collectively known as the histone code (Bannister and Kouzarides, 2011). Many histone PTMs have been linked to changes in chromatin density. Among these, lysine (K) acetylation (ac) and methylation (me) on histone variant 3 (H3) are commonly used as indicators of chromatin state: H3K9me and H3K27me are markers of heterochromatin, whereas H3K9ac, H3K4me, and H3K36me2/3 are associated with transcriptionally active chromatin (Xiao et al., 2016).

1.3.6 Interactions Between Histone Methylation and DNA Methylation

Histone modifications and DNA methylation share regulatory links that ensure robust chromatin density regulation in plants. For instance, the H3K9 methyltransferases SUVH4 (KRYPTONITE/KYP), SUVH5, and SUVH9 recognize non-CG DNA methylation and catalyse the addition of H3K9me1/2 to nearby histones (Du et al., 2015). Reciprocally, DNA methyltransferases CMT2 and CMT3 bind directly to H3K9me1/2, adding non-CG DNA

methylation to these regions (Stroud et al., 2014; Du et al., 2015; Fang et al., 2022). This mutual interaction creates a reinforcing loop between the two epigenetic marks, and thus, establishing/maintaining a stable heterochromatic state.

Another key player in this co-regulation is the gene *INCREASE IN BONSAI METHYLATION 1* (*IBM1*), which demethylates H3K9me2 in gene regions to protect them from transcriptional silencing (Saze et al., 2008; Miura et al., 2009; Yinwen Zhang et al., 2024). DNA methylation within *IBM1*'s 7th intron, however, is essential for producing the full-length functional transcript (*IBM1-L*). This process is regulated by the INCREASE IN BONSAI METHYLATION2 / ASI1-IMMUNOPRECIPITATED PROTEIN1 / ENHANCED DOWNY MILDEW2 (IBM2/AIPP1/EDM2) protein complex (To et al., 2015; Duan et al., 2017) (Figure 1.2). EDM2 directly binds to H3K9me2 marks found in heterochromatic introns, including the 7th intron of *IBM1*, where it forms a protein complex with the RNA-binding proteins IBM2 and AIPP1 (Saze et al., 2013; Tsuchiya and Eulgem, 2013, 2014; Wang et al., 2013; Duan et al., 2017). Consequently, the IBM2/AIPP1/EDM2 complex prevents access of another RNA binding protein, FPA, which regulates the transcription termination site choices of many mRNAs in Arabidopsis, thus allowing for the long functional transcript of IBM1 (*IBM1-L*) to be transcribed (Duc et al., 2013; Deremetz et al., 2019; Smith, 2019) (Figure 1.2a). As a result, loss of DNA methylation and H3K9me2 in the 7th intron of *IBM1* disrupts the association with the IBM2/AIPP1/EDM2 complex resulting in FPA-dependent premature polyadenylation of *IBM1*, resulting in short non-functional transcripts (*IBM1-S*) (Figure 1.2b) (Saze et al., 2008; Miura et al., 2009; Rigal et al., 2012; Deremetz et al., 2019; Smith, 2019; Yinwen Zhang et al., 2024). This can result in a feedback loop, whereby increased levels of IBM1 protein result in the suppression of *IBM1-L* expression by removing the H3K9me2 marks in this intronic region, leading to more *IBM1-S* expression (Figure 1.2b). Subsequently, there is a decrease in IBM1 protein level and thus, a passive increase in H3K9me2 and DNA methylation marks across the genome, including at this intronic region (Figure 1.2c). This consequently restores functional *IBM1-L* expression and the production of IBM1 protein (Figure 1.2a). Thus, this intronic region acts as a regulatory sensor of heterochromatin by modulating the balance of its own transcript isoforms (Figure 1.2). Furthermore, despite sequence divergence across species, intronic DNA methylation in IBM1 is evolutionarily conserved across flowering plants, underscoring its critical role in regulating chromatin accessibility (Yinwen Zhang et al., 2024).

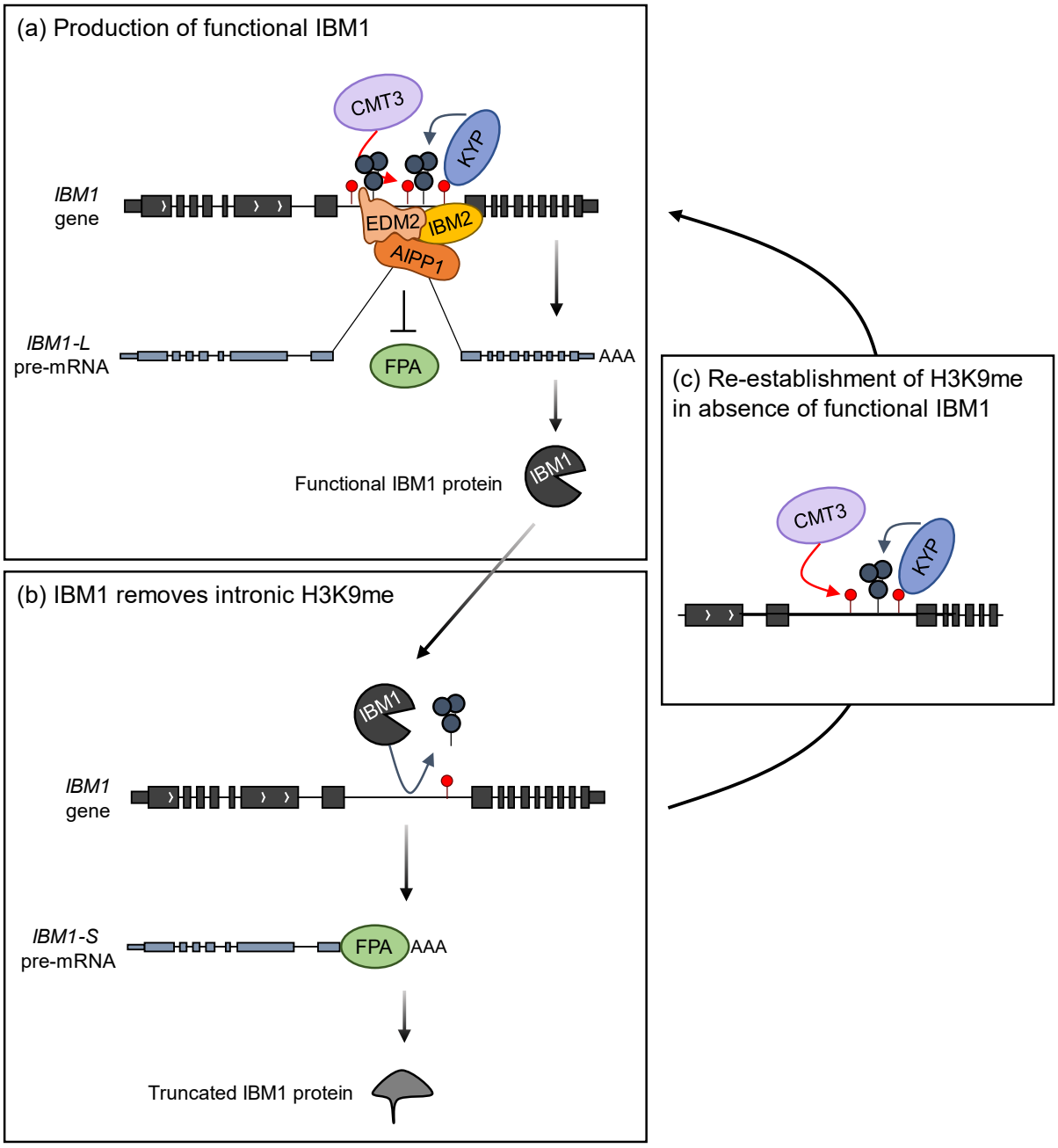


Figure 1.2 Gene model illustrating the transcriptional regulation of *IBM1* by H3K9me2 and DNA methylation.
(a) Recognition of intronic heterochromatin marks (DNA methylation and H3K9me2) in the 7th intron of *IBM1* by the IBM2/AIPP1/EDM2 complex is essential for generating full length IBM1 transcripts (*IBM1-L*) as it repels FPA-dependent premature polyadenylation (Duan et al., 2017; Deremetz et al., 2019). (b) As IBM1 functions as a genic H3K9me2 demethylase, it negatively regulates its own expression by removing the marks required for IBM2/AIPP1/EDM2 complex recruitment (Tsuchiya and Eulgem, 2014). Thus, FPA promotes the premature polyadenylation of *IBM1* resulting in short non-functional *IBM1* transcripts (*IBM1-S*) and a non-functional protein. (c) Consequently, H3K9me marks and DNA methylation are passively restored in IBM1-targeted regions by KYP and CMT3, respectively.

1.3.7 Non-coding RNAs

Non-coding RNA (ncRNA) activities are often considered epigenetic mechanisms because they regulate processes like DNA methylation, histone modifications, and the transcriptional/translational activity of genes (Statello et al., 2021). NcRNAs can be divided between small RNAs (sRNAs; <30 nt), such as miRNAs and siRNAs (Zhan and Meyers, 2023), and long ncRNAs (lncRNAs; >200-nt) (Kapranov et al., 2007). Both siRNAs and miRNAs can coordinate post-transcriptional gene silencing (PTGS) through association with AGO proteins, resulting in mRNA cleavage or translational repression at the endoplasmic reticulum. Furthermore, sRNAs can associate with AGO proteins to direct transcriptional gene silencing (TGS) via RdDM pathways, although this is primarily coordinated by siRNAs and rarely miRNAs (Cuerda-Gil and Slotkin, 2016; Zhan and Meyers, 2023). In plants, long non-coding RNAs (lncRNAs) are primarily transcribed by Pol II and Pol V. Furthermore, Pol IV is known to generate 30-40 nt siRNA precursor transcripts that are occasionally referred to as lncRNAs, although they do not fit the aforementioned description of sRNAs (<30 nt) nor lncRNAs (>200nt) (Wang and Chekanova, 2017; Kornienko et al., 2024). The expression of lncRNAs in Arabidopsis is controlled by other epigenetic marks, with most of the natural variation in lncRNA transcription in natural accessions of Arabidopsis originating from intergenic regions, which may act to reinforce TE (re)silencing (Kornienko et al., 2024). LncRNAs have diverse roles in regulating chromatin density, and gene expression. For example, lncRNAs produced by both Pol II and Pol IV are an important source of 21-24 nt sRNAs that mediate RdDM and transcriptional silencing through interaction with nascent Pol V derived lncRNAs (Cuerda-Gil and Slotkin, 2016; Sigman et al., 2021). However, recent work done in Arabidopsis has also reported that Pol V derived lncRNAs can have a direct influence on defence gene expression independently of RdDM which is associated with positive regulation of basal resistance against both *Botrytis cinerea* and *Spodoptera exigua* (Yuan et al., 2024). This may be driven by the fact that lncRNAs can directly interact with chromatin modifiers, allowing them to manipulate gene expression both in *cis* and in *trans* (Wang and Chekanova, 2017; Statello et al., 2021). Finally, lncRNAs can also stimulate gene expression by acting as a decoy/sponge of miRNAs and siRNAs that drive PTGS and TGS, respectively (Franco-Zorrilla et al., 2007; Wu et al., 2013; Wang and Chekanova, 2017).

1.3.8 The Heritability of Epigenetic Marks in Plants

Changes in epigenetic states, driven by the mechanisms described above, can vary in their stability, enabling phenotypic plasticity in plants across generations. Paramutations offer an example of highly stable epigenetic modifications that are stable across both mitotic and meiotic cell divisions. Paramutations occur when a silent epigenetic state of one allele causes

epigenetic silencing of another allele independently of DNA sequence, resulting in non-Mendelian inheritance of traits. For example, in maize, the expression of a transcription factor gene, *booster 1* (*b1*), which promotes anthocyanin production, is negatively regulated by silent epigenetic marks at a distal region upstream of *b1* (Hollick, 2017). When plants expressing *b1* are crossed with those not expressing *b1*, all heterozygous plants lack expression. Furthermore, even after backcrossing heterozygous F_1 plants with *b1*-expressing parental plants, all F_2 progeny are silenced for *b1* (Chandler, 2007). Genetic screens have demonstrated that establishment of this paramutation is dependent on the presence of functional RdDM machinery, indicating that sRNAs from one allele direct DNA methylation at the other allele in *trans* by sequence complementarity (Arteaga-Vazquez et al., 2010; Hollick, 2017). However, sRNA accumulation alone is not sufficient to cause paramutation at *b1* indicating that other silencing marks are likely involved (Arteaga-Vazquez et al., 2010). More recent work characterising the *sulfuera* (*sulf*) paramutation in tomato demonstrated that paramutation can occur independently of RdDM and instead requires CMT3 and KYP orthologs (Gouil et al., 2016; Martinho et al., 2022). This points towards a mechanism whereby RdDM may be required to initiate paramutation, but other factors such as CMT3/KYP-orthologs are required to drive transcriptional silencing and thus the maintenance/memory phase of paramutation (Martinho et al., 2022).

Maintaining epigenetic marks over a generation, as in the case with paramutation, requires active propagation of silencing epigenetic marks during reproduction. This is also essential to stably repress potentially harmful mobile genetic elements, such as TEs, across generations. During male gametogenesis, gametes are encased by asexual sporophytic cells at various stages of development, which have no genetic contribution to the final sperm cell but are critical for transmitting epigenetic information to the subsequent generation (Chow and Mosher, 2023). In Arabidopsis, the vegetative cell, which encases mature sperm cells, has reduced genome-wide levels of DNA methylation and increased transcriptional activity of TEs, which is caused by the downregulation of *DDM1* and the upregulation of *DME* and *ROS1*. The sRNAs derived from these activated TEs are able to mediate RdDM-dependent silencing of loci in sperm cells, thus ensuring stable repression of TEs in the next generation (Slotkin et al., 2009; Tirot and Jullien, 2022). Likewise, mobile 24 nt sRNAs produced in the sporophytic tapetum cells can promote deposition of DNA methylation at TEs and gene bodies with sequence complementarity (≤3 mismatches) in male meiocytes, resulting in the stable transcriptional silencing of germline TEs (Long et al., 2021). Interestingly, in *Brassica rapa,* a similar mechanism has been described in female ovules. In this case, highly abundant 24 nt sRNAs are produced from ‘siren’ loci, which often contain gene fragments. These sRNAs initiate DNA methylation in *trans* at protein-coding genes with sequence complementarity (≤2

mismatches), which can result in altered expression of the targeted genes (Burgess et al., 2022). However, it is unknown whether this methylation occurs in the germ line or is just restricted to the surrounding maternal somatic cells (Burgess et al., 2022; Chow and Mosher, 2023). Taken together, these studies highlight the intricate mechanisms plants employ to ensure the stable inheritance of DNA methylation and TE silencing during reproduction, thereby safeguarding genome integrity.

In contrast, it can be equally beneficial for plants to reset or 'forget' certain epigenetic states to ensure phenotypic plasticity in changing environments. A notable example of epigenetic reprogramming comes from the floral repressor gene *FLOWERING LOCUS C* (*FLC*). In Arabidopsis, prolonged exposure to cold conditions (e.g., over winter) results in the progressive and stable epigenetic silencing of *FLC* by H3K27me3 (De Lucia et al., 2008; Song et al., 2012). This enables plants to flower more quickly when floral-stimulating conditions, such as longer daylight hours in spring, arise (De Lucia et al., 2008; Song et al., 2012). However, during embryogenesis, the epigenetic silencing of *FLC* is reset by active de-methylation of H3K27, which allows for the subsequent generation to detect and respond to seasonal changes before flowering (Crevillén et al., 2014).

1.4 Epigenetics and Plant Immunity: an Emerging Field With Cross-Disciplinary Implications

1.4.1 Historical Overview

Throughout the late 1990s and early 2000s, independent research groups reported that mutations in epigenetic regulatory machinery affect plant disease resistance, while others reported that biotic stress exposure alters DNA methylation and histone PTMs (reviewed by Bruce et al., 2007; van den Burg and Takken, 2009; Alvarez et al., 2010). In subsequent years, more evidence emerged about the causal mechanisms linking epigenetic mechanisms to plant immunity. López et al (2011) demonstrated that Arabidopsis RdDM mutants express increased basal resistance against the bacterial pathogen *Pseudomonas syringae* pv. *tomato* DC3000 (*Pst*), which was associated with priming of SA-dependent defence genes. In support, Yu et al (2013) reported that the *ros1-4* mutant shows increased susceptibility to *Pst* and that flg22-induced PTI is linked to the transcriptional suppression of genes involved in RdDM. Furthermore, the increased accessibility of next-generation DNA sequencing enabled Dowen et al (2012) to perform genome-wide analyses of DNA methylation and gene transcription in Arabidopsis plants treated with SA and *Pst*. Their study revealed that immune-related changes in DNA methylation mostly occur at TEs and are linked to the production of 21-nt siRNAs. Together, these studies pointed to a model whereby the repression of DNA methylation at TEs during the expression of plant innate immunity leads to the priming of defence genes. This was consistent with results reported by Jaskiewicz et al. (2011), who found that SAR in Arabidopsis is associated with euchromatic histone modifications and priming of transcription factors associated with SA-dependent defences. On the other hand, DNA methylation and H3K9me2 in the first intron of *RECOGNITION OF PERONOSPORA PARASITICA 7* (*RPP7*) enables full-length transcription of this *R*-gene and effective ETI against an oomycete pathogen (Tsuchiya and Eulgem, 2013; Lei et al., 2014). This illustrates that heterochromatin and DNA methylation can also act as a positive regulator of plant immunity. Finally, four independent research groups between 2010 and 2012 reported that progeny from plants exposed to pathogens, herbivores, and/or IR-eliciting agents, develop heritable induced resistance (h-IR) that is associated with priming of defence-related genes (Kathiria et al., 2010; Luna et al., 2012; Rasmann et al., 2012; Slaughter et al., 2012). Together, these studies generated a foundation for further research into the epigenetic basis of plant immunity and its evolutionary implications thereof. Since then, DNA (de)methylation has been repeatedly highlighted as an epigenetic mark that (i) changes in response to biotic stress, (ii) regulates the expression of defence genes, and (iii) is required for the establishment and/or maintenance of immune memory (Wilkinson et al., 2019; Hannan Parker et al., 2022). Therefore, the next

Sections of this introductory Chapter will focus on the role of DNA (de)methylation in plant immune responses, from transcriptional regulation to the erasure of immune memory.

1.4.2 Transcriptional Regulation of Defence Genes by DNA (De)methylation

The primary function of DNA methylation to repress activity of TEs, invading viral DNA, and other potentially harmful genetic elements, has been adopted by the plant immune system. Consequently, immune activation by biotic stress is associated with widespread loss of DNA methylation (hypomethylation) at TE-rich regions, while mutations that reduce global DNA methylation enhance the plant's responsiveness to biotic stress (Dowen et al., 2012; López Sánchez et al., 2016; Geng et al., 2019; Atighi et al., 2020; Annacondia et al., 2021; Wilkinson et al., 2023; Lee et al., 2023). This link between plant immunity and TE hypomethylation can partially be explained by the discovery that immune activation leads to transcriptional repression of genes encoding RdDM components. For example, *AGO4* is repressed in Arabidopsis treated with the bacterial PAMP flagellin-22 (flg22), while *AGO4a* expression in the wheat diploid progenitor *Aegilops tauschii* is repressed in response to *Blumeria graminis f. sp. tritici* (*Bgt*) infection (Yu et al., 2013; Geng et al., 2019). But how does DNA hypomethylation at TEs stimulate defence gene expression?

1.4.2.1 Cis-Regulation

The most straightforward mechanism by which DNA hypomethylation at TEs regulates defence gene expression is through *cis*-regulation, where hypomethylation of the TE enhances the accessibility and affinity of the transcriptional machinery for adjacent genes. (Wilkinson et al., 2019) (Figure 1.3a). For example, DNA demethylation by ROS1 in the promoter of two defence genes, *RESISTANCE METHYLATED GENE 1* (*RMG1*) and *RECEPTOR LIKE PROTEIN 43* (*RLP43*), is required for the effective binding of flg22-responsive transcription factors (TFs), thereby enabling their transcriptional activation (Halter et al., 2021; Charvin et al., 2023). Similarly, expression of the NLR gene *PigmS* in rice, which controls resistance against the rice blast fungus *Magnaporthe oryzae*, is repressed by RdDM at *MITE*-family TEs in its promoter region (Deng et al., 2017). Whilst the vast majority of TFs in Arabidopsis have reduced binding capacity on methylated cytosines, a small subset preferentially bind methylated motifs (O'Malley et al., 2016). This, in addition to the activity of other protein complexes, means that methylated regions in the promoter of genes can also have a positive influence on transcription (Harris et al., 2018) (Figure 1.3a). For example, *ROS1* in Arabidopsis is positively regulated by DNA methylation in its promoter (Williams et al., 2015; Harris et al., 2018).

The methylation status of transposons or other repetitive regions can also influence the transcriptional start site, or transcriptional termination site of nearby genes, leading to the production of alternative transcript isoforms (Berthelier et al., 2023). This is a common form of *cis* regulation in NLR genes, such as *RPP7* and *RPP4* (Tsuchiya and Eulgem, 2013; Lei et al., 2014; Lai et al., 2020; Berthelier et al., 2023), as NLR genes tend to cluster with TEs (Quadrana et al., 2016). Interestingly, dysfunction of *IBM1*, which can be caused by a loss of intronic DNA methylation (see Section 1.3.5) (Figure 1.2), causes upregulation of SA-dependent defences and enhanced resistance to *Pst* (Rigal et al., 2012; Lv et al., 2022; Yinwen Zhang et al., 2024).

1.4.2.2 Trans-Regulation

Recent genome-wide DNA methylation and RNA profiling experiments suggest that the resistance-enhancing activity of hypomethylated TEs may not only be limited to *cis*-acting mechanisms (Wilkinson et al., 2019). Global transcriptome analysis of the Arabidopsis mutants *ros1-4* and *nrpe1-1,* which are oppositely affected in TE DNA methylation and basal resistance to virulent *Hyaloperonospora arabidopsidis* (*Hpa*), revealed that nearly half of the defence-related transcriptome during the early stages of *Hpa* colonisation were influenced by ROS1-dependent DNA demethylation and/or RdDM . However, only 15% of these genes were associated with nearby TEs, indicating that the majority of RdDM/ROS1-dependent defence genes were regulated indirectly by hypomethylated TEs (López Sánchez et al., 2016). Similarly, Halter et al. (2021) found that only 10% of all genes showing altered flg22-induced immune responses in the *ros1-3* mutant were hypermethylated, demonstrating that most ROS1-repressed defence genes are controlled indirectly by DNA methylation. In tomato, reductions in DNA methylation are linked to chemically induced primed defence against the fungal pathogen *Botrytis cinerea*, but most of these changes are not located near genes associated with the primed transcriptional response (Catoni et al., 2022). Furthermore, Furci et al. (2019) mapped epigenetic regions in a population of Arabidopsis epigenetic recombinant inbred lines (epiRILs) controlling heritable resistance to virulent *Hpa*. These epiRILs are derived from a cross, and subsequent backcrosses, between wild-type Arabidopsis and the severely hypomethylated *ddm1-2* mutant, resulting in near isogenic Arabidopsis lines with mosaic epigenomes due to the stable inheritance of hypomethylated DNA from the *ddm1-2* parent (Johannes et al., 2009). In this population, four hypomethylated epigenetic quantitative trait loci (epiQTL) were found to enhance *Hpa* resistance via genome-wide priming of defence genes (Furci et al., 2019). However, these epiQTLs were not located near defence genes; instead, they mapped to TE-rich pericentromeric regions, which lacked primed defence genes of resistant epiRILs. It was thus concluded that the hypo-methylated epiQTLs in resistant epiRILs prime defence genes via *trans*-acting mechanisms. In support, Cambiagno et al.

(2018) reported that transcriptional activation of TEs during infection by *Pst* leads to the increased accumulation of RdDM-dependent sRNAs which are complementary to both TEs and distal defence genes. After the initial immune response, these TEs were re-silenced, but the complementary defence genes remained active, suggesting that the sRNAs generated during TE silencing *trans*-activate distal defence genes (Cambiagno et al., 2018). Additional support for this model came from Liu et al. (2018), who showed that 21–22 nt sRNAs can *trans*-activate defence genes through interactions with nuclear-localised AGO1 and SWI/SNF chromatin remodelling complexes (CRCs). Wilkinson et al. (2023) subsequently provided evidence that long-term priming of jasmonic acid (JA)-induced defence against the generalist herbivore *Spodoptera littoralis* is controlled by AGO1-associated sRNAs from TEs, which become hypomethylated in response to primary JA stress and share sequence complementarity to the primed defence genes. Collectively, these results indicate a new model of plant immune memory, in which stress-induced sRNAs from epigenetically de-silenced TEs associate with nuclear AGO1 to *trans*-prime distal defence genes contributing to long-term IR (Figure 1.3c) (Hannan Parker et al., 2022).

The sequence complementarity of TEs also means that RdDM-derived small RNAs (sRNAs) from one TE has the potential to maintain RdDM at other, potentially distal TEs, even in cases of imperfect sequence homology (Martins and Law, 2023). For instance, using the same *ddm1*-derived epiRIL population as mentioned above (Johannes et al., 2009), Baduel et al. (2024) demonstrated that the vast majority of 23–24 nt sRNAs responsible for reverting *ddm1*-induced hypomethylated epialleles in chromosome arms originate from pericentromeric TEs. Furthermore, many of the TE-associated epialleles in chromosome arms were located in promoters of protein-coding genes having durect influence on their expression (Baduel et al., 2024). This recent study illustrates that defence gene expression is concertedly modulated by DNA methylation both via *cis*- and *trans*-regulatory mechanisms (Figure 1.3d).

Lastly, long-range chromatin interactions represent a mechanism by which changes in DNA methylation can *trans*-regulate distal defence genes (Wilkinson et al., 2019; Furci et al., 2019; Yueying Zhang et al., 2024). For instance, in Arabidopsis, mild cold stress treatment triggers the formation of extensive promoter-promoter interaction networks, both within and between chromosomes, which are highly enriched in transcription factor binding sites (Yueying Zhang et al., 2024). These genes display primed induction upon subsequent cold stress and show evidence of co-expression (Yueying Zhang et al., 2024). Therefore, the formation of promoter-promoter interactions may facilitate gene priming and co-regulation of environmentally responsive genes during stress. Under normal conditions, chromatin contacts in plants are enriched for repressive chromatin marks like H3K9me2, H3K27me3, and DNA methylation (S. Feng et al., 2014; Nützmann et al., 2020; Huang et al., 2021). However, loss of these marks

in epigenetic mutants leads to the establishment of new interactions between euchromatic gene regions (S. Feng et al., 2014; Huang et al., 2021). It is thus plausible that stress-induced repression of these silencing marks disrupts heterochromatin contacts and subsequently promotes interactions with alternative regions, such as the promoters of stress-responsive genes, resulting in co-transcriptional activation (Figure 1.3e

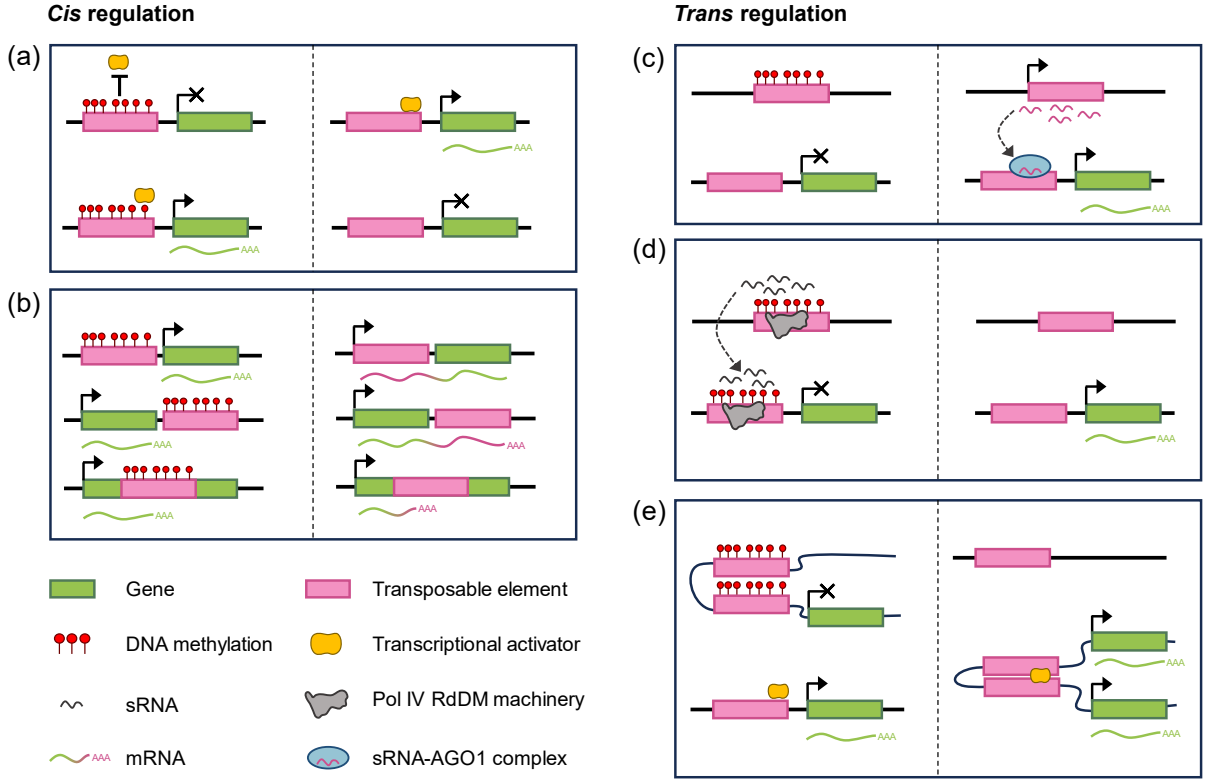


Figure 1.3 Models of *cis*- (a,b) and *trans*-regulatory (c,d,e) mechanisms of gene expression by DNA methylation of transposable elements (TEs).
(a) DNA methylation at TEs in the promoter region of genes can influence transcription factor binding to promote or repress gene expression (Harris et al., 2018; Halter et al., 2021). (b) DNA methylation status of TEs in gene proximal regions can influence transcript isoform usage (Berthelier et al., 2023). (c) 21 nucleotide (nt) small RNAs (sRNAs) from transcriptionally active TEs that have lost DNA methylation can associate with nuclear-localised AGO1 to promote the expression of distal genes with sequence homology in their promoter (Liu et al., 2018; Wilkinson et al., 2023). (d) 23–24 nt sRNAs associated with Pol IV-dependent RNA directed DNA methylation (RdDM) can silence distal TEs with sequence homology (Baduel et al., 2024), which may negatively regulate defence genes in *cis* (a,b). (e) Loss of DNA methylation can disrupt long-range chromatin interactions between heterochromatic TEs and instead promote interactions with euchromatic promoter regions that facilitate co-regulation of genes (S. Feng et al., 2014; Huang et al., 2021; Yueying Zhang et al., 2024).

1.5 Epigenetic Control of the Establishment, Maintenance, and Erasure of Immune Memory

As discussed above, plants can 'remember' past environmental signals to respond more effectively to future exposures, which in the case of stress, can result in IR phenotypes. Epigenetic mechanisms offer a plausible explanation for the long-term nature of IR against pathogens, because they can direct changes in basal expression and/or responsiveness of genes that remain stable across cell divisions and even generations (Wilkinson et al., 2019) (Figure 1.3). Indeed, previous research has shown that artificial manipulation of DNA methylation in the pericentromeres can induce priming-related resistance that remains stable over at least 8 generations (Furci et al., 2023). For stress-induced immune memory, the severity of the initial stress is proportional to the effectiveness and stability of the heritable induced resistance (h-IR) across generations (López Sánchez et al., 2021; Furci et al., 2023). Thus, stress intensity influences the rate at which immune memory is erased or ‘forgotten’ (Bozarth and Ross, 1964; López Sánchez et al., 2021). The reversibility of immune memory implies it can be maladaptive under certain conditions, requiring plants to reset the memory in the absence of recurrent stress to prevent enduring costs (Crisp et al., 2016). Indeed, h-IR following disease stress by biotrophic pathogens mounts an effective h-IR response to pathogens of the same lifestyle in the following generations, but causes heritable induced susceptibility (h-IS) against necrotrophic pathogens and abiotic stress (López Sánchez et al., 2021). Similarly, 2-week-old seedlings treated with JA exhibit IR against the generalist herbivore *Spodoptera littoralis* three weeks later, but display increased susceptibility to both a necrotrophic pathogen (*Plectosphaerella cucumerina*) and a hemi-biotrophic pathogen (*Pst*) (Wilkinson et al., 2023). Furthermore, even if stress in the parental and progeny environments are matched, primed plants can show reductions in growth and seed set (van Hulten et al., 2006; López Sánchez et al., 2021). Therefore, in addition to the protective benefits of 'remembering' stress, it is ecologically beneficial for plants to 'forget' stress memory to ensure plasticity in a dynamic environment. However, as articulated by Crisp et al. (2016) and Harris et al. (2023), few studies have focussed on the mechanisms driving this erasure of immune memory.

There are some studies that provide clues about the mechanisms of stress memory erasure. Using Arabidopsis lines carrying a transcriptionally silent transgene that transiently becomes activated following heat stress, Iwasaki and Paszkowski (2014) demonstrated that two key regulators of DNA methylation, DDM1 (Zemach et al., 2013) and MORPHEUS' MOLECULE 1 (MOM1) (Li et al., 2023), play a role in the epigenetic re-silencing of this transgene. It was shown that the transiently expressed activation of this heat-inducible

transgene persisted within and across a generation in *ddm1 mom1* double mutants (Iwasaki and Paszkowski, 2014), thus mimicking long-term constitutive priming following a stress stimulus (Figure 1.1). Interestingly, elicitation of h-IR in Arabidopsis by *Pst* stress is associated with a prolonged downregulation of *DDM1*, which persists in the F1 progeny of the *Pst*-stressed plants (Furci et al., 2023). Thus, suppression of DDM1 may facilitate long-term activation of defence genes associated with h-IR against *Pst*. Further clues about the negative regulation of immune memory derives from an experiment using the chemical priming agent β-aminobutyric acid (BABA; Zimmerli et al., 2000; Ton et al., 2005; Luna et al., 2014; Cohen et al., 2016; Tao et al., 2022). Sub-active doses of BABA, which are insufficient to trigger IR against *Pst* in wild-type Arabidopsis, result in robust IR in *ddm1* and *met1* mutants, suggesting that DDM1- and MET1-dependent DNA methylation antagonises BABA-induced immune memory (Lee et al., 2023). Taken together these studies highlight a critical role for DNA re-methylation by DDM1 in the erasure of immune memory.

Apart from DDM1, many other regulators of DNA methylation have been implicated in the establishment and/or maintenance of immune memory (see Harris et al., 2023 for a more comprehensive list). Most notably, plants lacking a functional ROS1 enzyme fail to express heritable induced resistance (h-IR) against biotrophic pathogens (López Sánchez et al., 2016), whereas plants affected in non-CG DNA methylation mimic the priming phenotype of h-IR by parental *Pst* stress or BABA (Luna et al., 2012, 2014). Furthermore, both *nrpe1-11* and *ros1-4* mutants, which are deficient in Pol V-dependent RdDM and ROS1-dependent DNA demethylation, respectively, cannot establish long-term within-generation JA-IR against the chewing herbivore *Spodoptera littoralis* (Wilkinson et al., 2023). Therefore, in addition to the suppression of epigenetic memory erasure, the fine-tuned homeostasis of DNA methylation seems critical for the epigenetic endurance of immune memory across cell divisions and generations. Furthermore, these findings highlight a central role of the DNA demethylase ROS1 in regulating immune memory, both within a generation and across generations, in response to pests and diseases.

1.6 Scope of This Thesis

Based on the central role of the DNA de-methylase ROS1 in the regulation of within- and transgenerational immune memory, my PhD research has investigated how this protein orchestrates the onset, maintenance, and erasure of immune memory in Arabidopsis (Figure 1.1). To address the complexity and confounding variability of natural memory responses triggered by biotic stress, I utilised a chemically inducible ROS1 gene construct, which enabled precise and transient inductions of genome-wide ROS1-dependent DNA demethylation in Arabidopsis.

In the first experimental Chapter (Chapter 3), I highlight and discuss some of the difficulties of investigating the roles of DNA (de)methylation in h-IR assays using natural plant-pathogen assays. Furthermore, while the use of genetic mutants has provided a wealth of knowledge in understanding the role of epigenetics in regulating plant immune responses, I will discuss the limitations associated with permanently modifying DNA methylation machinery to study an inherently dynamic epigenetic mark. By doing so, I further justify the use of a transiently inducible *ROS1* gene construct, which enables dose- and tissue-dependent induction of transient DNA demethylation to mimic the more variable DNA demethylation events occurring upon natural stresses. The Chapter provides evidence that 2 successive ectopic inductions of ROS1 induce immune memory against *Pst* and *Hpa*. However, progeny of ROS1-induced lines did not mount effective h-IR, indicating that the immune memory incurred by ROS1-dependent DNA demethylation is lost during the life of the parental plants and/or over meiosis.

The second experimental Chapter (Chapter 4) aims to characterise the transcriptional and epigenomic impacts of ectopically induced ROS1 during the onset of epigenetic immune memory against biotrophic pathogens. The results in Chapter 4 confirm that two successive inductions of ROS1 lead to global reductions in DNA demethylation, which are associated with an up-regulation of genes involved in SA-dependent immunity and DNA damage. Furthermore, the results of Chapter 4 uncover an unexpected disparity in sRNA profiles between chromosome arms and pericentromeric regions, which is explored further in the context of the establishment and erasure of immune memory in Chapter 5.

The third experimental Chapter (Chapter 5) focuses on the role of DNA methylation in centromeric and pericentromeric regions in the erasure of ROS1-induced immune memory. Taking advantage of direct long-read sequencing by Oxfrod Nanopore Technology (ONT), which provides better coverage of the more repetitive (peri)centromeric regions of the genome, it is shown that ROS1 induction not only induces widespread DNA hypomethylation in

chromosome arms, but simultaneously induces variable DNA hypermethylation in the (peri)centromeric regions in non-CG cytosine contexts. Based on my previous finding that ectopically induced ROS1 increases the accumulation of epigenomic sRNAs, this DNA hypermethylation in the (peri)centromeres points to a rapid RdDM response that counteracts widespread TE activation in these regions. Subsequent experiments with 5-azacytidine (5-Aza), which preferentially inhibits pericentromeric CHH methylation (Griffin et al., 2016), demonstrate that estradiol-induced ROS1 induction combined with 5-Aza enhances the strength and durability of ROS1-IR. These results are discussed in the context of a wider model of the epigenetic cross-communication between the chromosome arms and the (peri)centromeric regions, which may function to balance the establishment, maintenance, and erasure of ROS1-dependent immune memory.

The final experimental Chapter (Chapter 6) describes how variation in the spatiotemporal patterning of ROS1 induction determines the specificity of stress-specific memory against a diverse selection of biotic and abiotic stressors. Complementary ONT sequencing shows that these stress-specific IR responses mediated by ROS1 correlate with varying patterns of genome-wide DNA methylation. The results of Chapter 6 demonstrate the multifaceted involvement of ROS1 in plant stress memory and suggest that it is possible to exploit ROS1-mediated DNA demethylation for the epigenetic engineering of stress-specific resistance in plants.

Chapter 2. General Materials and Methods

Authors:

Hannan Parker, A[1]; Ton, J[1].

Affiliations:

[1] Plants, Photosynthesis and Soil, School of Biosciences, The University of Sheffield, Western Bank, Sheffield, S10 2TN, UK

Author contributions:

All material in this Chapter was created by AHP (the candidate). JT reviewed and provided comments on the Chapter.

Contribution of co-authored publications:

Section 2.4 ("generation of *XVE:ROS1-YFP* plants") has been adapted from the following published article with permission from Springer Nature.

Wilkinson, S. W.; Hannan Parker, A.; Muench, A.; Wilson, R. S.; Hooshmand, K.; Henderson, M. A.; Moffat, E. K.; Rocha, P. S. C. F.; Hipperson, H.; Stassen, J. H. M.; López Sánchez, A.; Fomsgaard, I. S.; Krokene, P.; Mageroy, M. H. and Ton, J. (2023). Long-Lasting Memory of Jasmonic Acid-Dependent Immunity Requires DNA Demethylation and ARGONAUTE1. Nature Plants, 9(1): 81–95.

2.1 Abstract

This Chapter outlines the materials and methods commonly used across multiple experiments from different experimental Chapters . Namely, (i) standard plant growth conditions, (ii) preparation of estradiol (E2) stocks, (iii) generation of *XVE:ROS1-YFP* transgenic plants, and (iv) cultivation and inoculation of the phytopathogens *Pseudomonas syringae* pv. *tomato* DC3000 *luxCDABE* (*Pst-Lux*) and *Hyaloperonospora arabidopsidis* isolate Waco9 (*Hpa*). Each experimental Chapter also includes a specific method Section, with references to this Chapter where appropriate.

2.2 Plant Materials and Growth Conditions

Unless specified otherwise, all experiments were carried out with *Arabidopsis thaliana* (Arabidopsis) accession Col-0. Prior to sowing, seeds were stratified at 4°C in the dark for 2-4 days. Seeds were then sown in 70 ml pots containing a soil mix consisting of sand and a peat-based compost (either Levington M3 or Levington F2+S) in a 2:1 ratio saturated with water. Plant were grown either under short day conditions (8.5:15.5 h day:night at 21°C at 45–70% RH; Valoya NS1 LED bulbs providing ~150 µmol photons m^{-2} s^{-1}), or long day conditions (16:8 h day:night at 21-23°C; 45–70% RH; Sylvania GroLux T8 36W bulbs providing ~ 150 µmol photons m^{-2} s^{-1}), and kept in covered trays to maintain 100% relative humidity (RH) until seeds had germinated and cotyledons had emerged. Pots were watered from the bottom 2-3 times a week with distilled water (dH_2O) and trays were frequently rotated within the chamber to prevent positional effects.

2.3 Preparation, Storage, and Treatments With Estradiol

Estradiol (17β-estradiol; Sigma-Aldrich; E8875) (E2) was dissolved in dimethyl sulfoxide (DMSO; Sigma-Aldrich D4540) to a final stock concentration of 50 mM. Small stock aliquots (25-100 ul) were stored at -20°C to prevent recurrent thawing and freezing. At the time of treatment, stock aliquots were diluted in ultra-pure distilled and deionised water (ddH_2O) to the desired working solution. Unless stated otherwise, induction treatments were carried out by spraying leaves until run-off with solutions containing the indicated E2 concentrations and 0.015% (v/v) silwet L-77 (silwet; LEHLE SEEDS, VIS-30). Mock solutions (DMSO treatments) consisted of ddH_2O containing 0.015% silwet and an equivalent volume of DMSO.

2.4 Generation of *XVE:ROS1-YFP* Plants

Genomic DNA (gDNA) was extracted from seedlings as described previously (Brunel, 1992), and used for PCR to amplify genomic ROS1 DNA (*gROS1*), using Phusion DNA polymerase (New England Biolabs, M0530S) and gene specific primers (ROS1-F1 and ROS1-R1) (primer list in Table 2.1). The amplified sequence included a CACC sequence at the 5'-end for site specific recombination and excluded the stop codon at the 3'-end of ROS1 to enable the C-terminal fusion with Venus YFP (YFP; Nagai et al. 2002). The amplified *gROS1* sequence was cloned into a *pENTR/D-TOPO* plasmid and transformed into TOP10 *E. coli* cells (*pE-gROS1;* Thermo Fisher Scientific, K240020). Using a MultiSite Gateway recombination reaction according to manufacturer's recommendations (MultiSite Gateway Three-Fragment Vector Construction Kit; Thermo Fisher Scientific, 12537-023), the entry plasmids *pE-gROS1*, *p1R4-p35S:XVE* (kindly provided by Ari Pekka Mähönen of the

University of Helsinki, Finland) (Siligato et al., 2016), and *p2R3a-VenusYFP-3AT* (containing the VenusYFP gene; Addgene, plasmid #71269) (Siligato et al., 2016) were integrated into the empty binary destination vector *pCAMkan0R4R3* (Addgene, plasmid #71275) (Siligato et al., 2016) to form the 19,117 bp *pR-XVE-gROS1-YFP* binary plasmid (*R* stands for recombinant) (Figure 3.4a). The plasmid was extracted from positive colonies, using the GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific, K0503), and were verified by restriction enzyme analysis, Sangar sequencing (provided by Eurofins Genomics) (Table 2.1), and whole plasmid sequencing (provided by PlasmidSaurus). Notably, plasmid construction was originally attempted with *ROS1* cDNA, but successful clones grew at a significantly slower rate. Colonies growing at a normal rate contained mutations rendering non-functional ROS1 and/or their inserts were cloned in the wrong orientation, suggesting that DNA glycosylase/lyase activity by functional ROS1 has a negative impact on TOP10 *E.* coli cells.

The *pR-XVE-gROS1-YFP* plasmid was transformed into *Agrobacterium tumefaciens* strain GV3101 by electroporation, using the 'Agr' preset on the MicroPulser Electroporator (BioRad) (0.1 cm cuvette, 2.2 kV, 1 pulse). Transformed GV3101 cells were used for floral-dip transformation of Arabidopsis (Col-0) (Clough and Bent, 1998), after which T1 transformants were selected on half-strength Murashige and Skoog agar (1%) plates, containing 50 µg/ml kanamycin (Harrison et al., 2006). Lines displaying Mendelian 3:1 segregation in the T2 generation of kanamycin resistance and YFP fluorescence upon E2 treatment were selected to yield single-insertion homozygous T3 lines, which were verified for homozygosity by kanamycin resistance and YFP fluorescence. The construct was transformed into a Col-0 background, rather than a *ros1* mutant background, as plants lacking *ROS1* are severely impaired in basal and induced defence against (hermi)biotrophic pathogens. This would greatly complicate investigations into immune memory using the construct as control plants (*ros1* mutants) would be unresponsive to memory-inducing treatments (López Sánchez et al., 2016; Halter et al., 2021). However, a caveat of this approach is that Col-0 plants have a high level of background *ROS1* expression (Zhang et al., 2022), which could complicate the detection and interpretation of certain ROS1 targets that are already very low in DNA methylation.

Table 2.1 PCR primers used for the creation and verification of the *pR-XVE-gROS1-YFP* plasmid

	Name	**Nucleotide sequence**
Amplification of genomic *ROS1* (*gROS1*) sequence.	ROS1-F1	CACCGAAATGGAGAAACAGAGGAGAGAAG
	ROS1-R1	GGCGAGGTTAGCTTGTTGTCCC
Sangar sequencing of *p2R3a-VenusYFP-3AT*	pYFP-3657R	GAAGTCGTGCTGCTTCATGTG
	pYFP-3947F	ACCACTACCAGCAGAACACC
	pYFP-4216R	GTGTGTGGGCAATGAAACTG
Sangar sequencing of *p1R4-p35S:XVE*	pXVE-393F	CCACTATCCTTCGCAAGACC
	pXVE-1346F	GCTACTGTTTGCTCCTAACTTGC
	pXVE-2603F	GGTGTCATCTATGTTACTAGATCG
	pXVE-2679R	TGCCATAATACTCGAACAAGC
Sangar sequencing of *pCAMkan0R4R3*	pCAM-10221R	CGTATGTTGTGTGGAATTGTGAG
	nos-T	CCTTCTATCGCCTTCTTGACG
	npt-F	AAATGCTCCACTGACGTTCC
	npt-R	CCAAATGTTTGAACGATCTGC
	T7	TAATACGACTCACTATAGGG
Sangar sequencing of *gROS1* in *pR-XVE-gROS1-YFP*	M13-F	TGAGTTTCGTCACCAGTA
	M13-R	CAGGAAACAGCTATGACC
	SQ-ROS1-1069F	CAAGGAAGGCAAGAGTCTGAG
	SQ-ROS1-1967R	GTTTCCTCTGAGTCCAAGTACG
	SQ-ROS1-2243F	GATGGTGTCTGATGCTCCTC
	SQ-ROS1-2475R	AGGTTCACTACTATCAACAGAGG
	SQ-ROS1-2629F	GATTGGAAGGCAATACGAGC
	SQ-ROS1-3530F	AAACTGCTTCTCTTCCAATGC

2.5 Pathogen Cultivations and Inoculations

Pseudomonas syringae pv. *tomato* DC3000 *luxCDABE* (*Pst-Lux*; Fan, Crooks, and Lamb 2008) was originally provided by Dr Jun Fan (John Innes Centre, UK) and was stored in 15% (v/v) glycerol stocks at -80°C. The afternoon before inoculation, a glycerol stock was thawed on ice before adding 500 µL of the stock to 100 ml of King's B (KB) solution, supplemented with rifampicin (50 µg/ml) and kanamycin (50 µg/ml) to select for the *luxCDABE* cassette. Cultures were grown for approximately 16 hours at 28°C and 200 rpm, during which plant trays were covered with transparent lids to obtain a RH of 100%. *Pst-Lux* cultures were split into 2 x 50 ml falcon tubes, and centrifuged (4000 x *g* for 3 mins), after which the supernatant was discarded. Each pellet was washed and resuspended in 20 ml of 10 mM $MgSO_4$, before further centrifugation at 4000 x *g* for 3 mins. A single pellet was resuspended in 25 mL of 10 mM $MgSO_4$ and transferred into the second tube to resuspend the second pellet. The optical density of the merged suspension was measured at 600 nm (OD_{600}) and adjusted to a final density of 0.2 OD_{600} before adding 0.015% v/v silwet (Furci et al., 2021). The resulting inoculum was sprayed onto the leaf surface of plants until run-off. Inoculated plants were covered with transparent lids (wrapped in parafilm to ensure all trays reached 100% RH).

Non-destructive quantification of *Pst-Lux* colonisation was performed using a G:BOX Chemi XRQ gel doc system and GeneSys software (Furci et al., 2021). Plants (9-12 pots per image) were placed in the dark room of the gel doc for 8 minutes to quench all chlorophyll autofluorescence before measuring bacterial bioluminescence. Images were acquired with a 4 min exposure time without external light source, using the binning set at "4 x 4 (0.25MP)". To estimate rosette area, a matching bright field (BF) image was acquired immediately after, using the internal LED upper white light source with 25 ms exposure time. A 2 pence coin was placed in the gel doc for size reference. Images were saved as 16-bit TIFF files, and plants were returned to the climate chamber and maintained at 100% RH to ensure ongoing infection and allow for recurrent analyses at later time-points. For image analysis, the brightfield image stack was imported (file > import > image sequence) into Fiji/ImageJ (Rueden et al., 2017) and converted to 8-bit format (image > type > 8-bit). The scale was set in mm per pixel using the diameter of 2 pence coin (25.9 mm) as reference. Automatic local thresholding was applied to the stack (image > adjust > auto local threshold), using the Phansalkar method with a radius of 30 and the option of "white objects on black background" selected. Plants that were connected after thresholding were separated using a black pencil tool, with a line width of 3, and disconnected leaves were reconnected using a white pencil tool with a line width of 3. Manual removal of non-plant selections (e.g., pot edge and soil debris) was done using the

black pencil tool (Figure 2.1a). Regions of interest (ROIs), corresponding to the outline of the plants (and thus, the area of the plant rosette), were generated (analyse > analyse particles > add to manager) and saved locally. The bioluminescent image stack was then imported, and the ROIs were overlayed to highlight plant regions. Five additional ROIs of varying size were drawn to select non-plant background regions (Figure 2.1b) and the mean-grey value of these background ROIs were measured for every image in the stack (more > multi measured > untick one row per slice). The mean of all background measurements plus 3x the standard deviation was used to subtract background noise from the bioluminescent images (process > math > subtract). Subsequently, the remaining pixels were used to calculate mean-grey values for each plant-derived ROI. Thus, the final measurement gives a metric of *Pst-Lux* bioluminescence intensity above background levels, which is normalised to plant size, as determined by the BF-derived ROI (Furci et al., 2021).

Hyaloperonospora arabidopsidis isolate Waco9 (*Hpa*) was maintained on hypersusceptible *NahG*-expressing Arabidopsis plants of the ecotype Wassilewskija (Ws) (Molina et al., 1998). To collect conidiospores, infected leaves from host plants were washed with deionised water, filtered through miracloth, and quantified using a hemocytometer (Improved Neubauer, Hawksley, UK; Depth 0.1 mm, 1/400 mm^2). Spore density was adjusted to $1x10^5$ conidiospores/mL and sprayed onto the surface of leaves until droplets covered all leaves (Asai et al., 2014; Schwarzenbacher et al., 2020). Trays were covered transparent lids and wrapped in parafilm to ensure 100% RH. For pathogen propagation, the spore suspension was sprayed onto fresh Ws *NahG* plants, and left for 7 days before collecting conidiospores. For pathogen bioassays, the spore suspension was sprayed onto the plants and left for 5-6 days before harvesting shoot material in 100% ethanol for *Hpa* colonisation scoring. Plant material was stained with 1 volume lactophenol-trypan blue solution (100 mg trypan blue, 50 mg phenol, 50 mL lactic acid, 50 mL glycerol, dissolved in 100 mL deionised water) and 2 volumes of 100% ethanol (300 ml) in falcon tubes (Koch and Slusarenko, 1990; Stassen et al., 2018). Falcon tubes were incubated for 60 seconds at 100°C. Immediately after, caps of falcon tubes were temporarily loosened to initiate a change in pressure and so aid the staining process. After 5 minutes, this was repeated but with a 40 second incubation at 100°C. Samples were left at room temperature for 4 hours before replacing the trypan blue solution with 60% w/v chloral hydrate. After ~16 h, this was replaced with fresh 60% w/v chloral hydrate. Individual leaves (no less than 50 per treatment) were microscopically analysed (Optika Italy Lab-20 binocular stereozoom microscope), based on previously reported hallmarks of *Hpa* colonisation and pathogenesis (Schwarzenbacher et al. 2020). Briefly, leaves were assigned to four *Hpa* colonisation classes: I: no hyphal colonisation; II: limited hyphal colonisation with 8 or less conidiophores; III: extensive hyphal colonisation with more than 8 conidiophores; and

class IV: extensive hyphal colonisation with conidiophores and oospore(s), as illustrated in Figure 2.2.

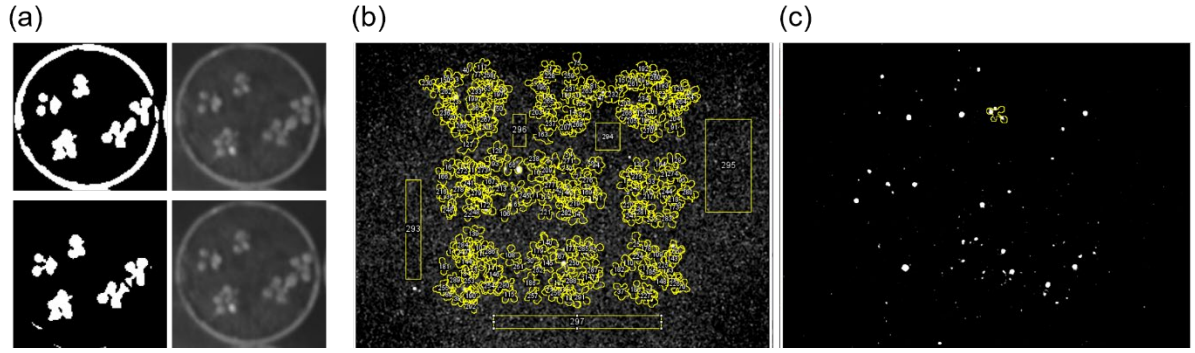


Figure 2.1 Quantification of bioluminescence by *Pst-Lux* in Arabidopsis.
(a) Examples of manual-touch ups following thresholding in imageJ. Disconnected leaves are rejoined, and connected plants are separated. Pot edges are removed. (b) Shown are all regions of interest (ROIs) corresponding to all plants in the image stack. The five yellow-contoured square ROIs (labelled 292 to 296) were used to measure and subtract background levels. (c) Selection of a single plant ROI (indicated by yellow contours), which contains white pixels corresponding to background-corrected *Pst-Lux* bioluminescence.

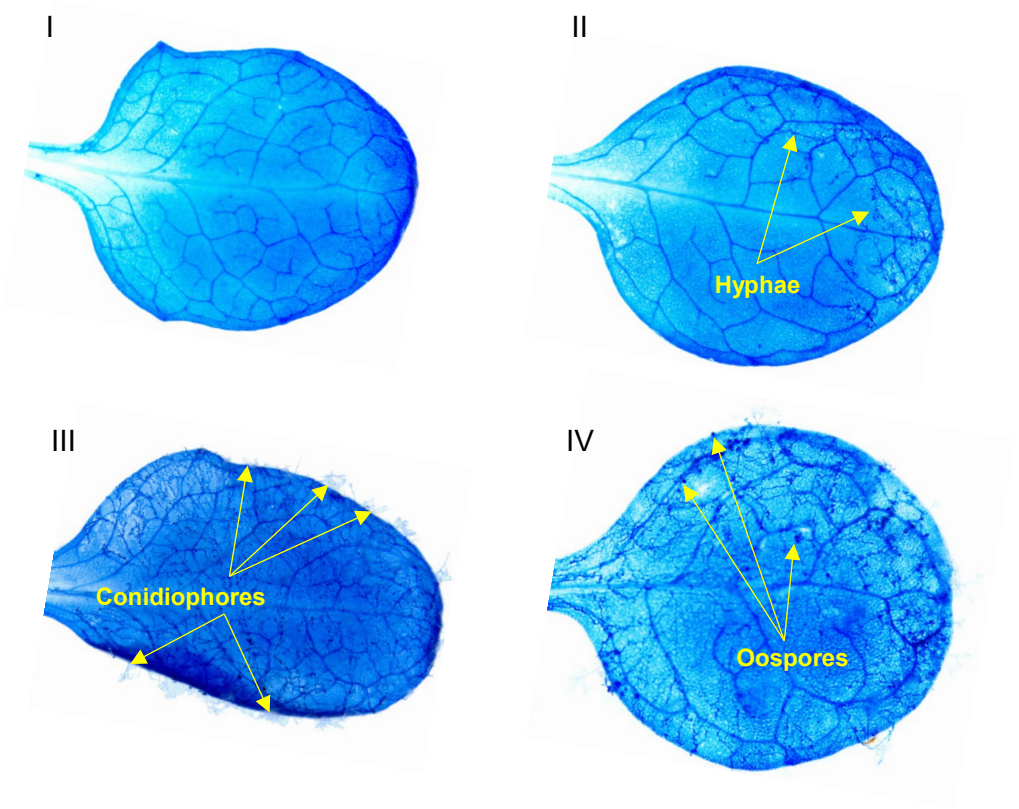


Figure 2.2 Representative trypan-blue stained leaves for each *Hpa* colonisation class.
The *Hpa* colonisation scoring is based on Schwarzenbacher et al (2020). (I) Class I: no hyphal colonisation. (II) Class II: limited hyphal colonisation, ≤ 8 conidiophores. (III) Class III extensive hyphal colonisation, > 8 conidiophores. (IV) Class IV: extensive hyphal colonisation with conidiophores and ≥ 1 oospore.

Chapter 3. A Novel Tool for Investigating ROS1-Dependent DNA Demethylation in the Regulation of Long-Term Immune Memory in Arabidopsis

Authors:

Hannan Parker, A[1]; Man, K. Y. [1]; Rocha, P. S. C. F. [1]; Wilkinson, S. W. [1]; Furci, L.[1]; Stassen, J. H. M. [1]; Rolfe S. A. [1]; Johannes, F.[2]; Smith, L. M.[1] and Ton, J[1].

Affiliations:

[1] Plants, Photosynthesis and Soil, School of Biosciences, The University of Sheffield, Western Bank, Sheffield, S10 2TN, UK.
[2] TUM School of Life Sciences, Plant Epigenomics, Technical University of Munich, Freising, 85354, Germany.

Author Contributions:

AHP (the candidate), SAR, LMS, JT, and FJ proposed the original ideas for the research. AHP conducted experiments and gathered data with assistance from PSCFR, KYM, SWW, LF, and JT. Data analysis was performed by AHP with assistance from SWW, JHMS, and SAR. AHP created all Figures and wrote all the text in the Chapter. JT reviewed and provided comments on the Chapter. Funding for the project came from grants/studentships awarded to AHP, LMS, FJ, and JT.

Contribution of co-authored publications:

Figure 3.4a, and Sections in the methods (section 3.3.3 and 3.3.4) have been adapted from the following published article. Reproduced with permission from Springer Nature.

Wilkinson, S. W.; Hannan Parker, A.; Muench, A.; Wilson, R. S.; Hooshmand, K.; Henderson, M. A.; Moffat, E. K.; Rocha, P. S. C. F.; Hipperson, H.; Stassen, J. H. M.; López Sánchez, A.; Fomsgaard, I. S.; Krokene, P.; Mageroy, M. H. and Ton, J. (2023). Long-Lasting Memory of Jasmonic Acid-Dependent Immunity Requires DNA Demethylation and ARGONAUTE1. Nature Plants, 9(1): 81–95.

Supplementary data 3.1 which lists papers evidencing h-IR is has been reproduced from the following published article with permission from Wiley. LF, AHP, and JT contributed to the creation of this Table in the original publication.

Furci, L.; Pascual-Pardo, D.; Tirot, L.; Zhang, P.; Hannan Parker, A. and Ton, J. (2023). Heritable Induced Resistance in Arabidopsis Thaliana: Tips and Tools to Improve Effect Size and Reproducibility. Plant Direct, 7(8): e523.

3.1 Abstract

Plants recovering from severe biotic stress typically develop long-term immune memory, which can persist across generations and manifest itself as long-term and even heritable induced resistance (h-IR). DNA methylation, an epigenetic mark that can be maintained across mitosis and meiosis, has been implicated in the regulation of h-IR. Nonetheless, the epigenetic loci underpinning h-IR remain elusive. In this Chapter, I first examine the changes in DNA methylation that are associated with h-IR in Arabidopsis against the biotrophic pathogen *Hyaloperonospora arabidopsidis* (*Hpa*), following parental disease with *Pseudomonas syringae* (*Pst*). Using whole-genome bisulfite sequencing, the DNA methylome between progeny from healthy and *Pst*-infected plants were compared in wild-type plants (Col-0) and the hypermethylated mutant *ros1-4*, which is impaired in h-IR. Although wild-type plants displayed a statistically significant h-IR response to *Hpa*, which was absent in *ros1-4*, the impacts of h-IR on global DNA methylation were highly variable, making it challenging to pinpoint h-IR-related epigenetic loci. Therefore, to study how ROS1-dependent DNA demethylation mediates epigenetic immune memory, Arabidopsis plants were transformed with an estradiol-inducible *XVE:ROS1-YFP* construct and tested for IR upon ROS1-YFP induction. Two successive estradiol treatments of independent *XVE:ROS1-YFP* lines resulted in significant IR against *Pst* and *Hpa*. However, the progeny of induced *XVE:ROS1-YFP* plants did not display h-IR against *Hpa*, indicating that the immune memory by ROS1-dependent DNA methylation is lost during the life of the parental plants and/or over meiosis. Thus, the estradiol-inducible *XVE:ROS1-YFP* construct serves as valuable tool for studying the establishment, maintenance, and often-overlooked erasure of immune memory.

3.2 Introduction

More than forty years ago, Roberts (1983) reported that progeny from tobacco mosaic virus-infected tobacco shows increased resistance to the same virus. Since then, heritable induced resistance (h-IR; De Kesel et al., 2021) phenotypes have been reported in at least 13 different plant species and 17 different types of biotic stress (Furci et al., 2023) (Supplementary Data 3.1). Molecular characterisation of h-IR responses has revealed that it typically involves the priming of defence genes (Kathiria et al., 2010; Luna et al., 2012; Rasmann et al., 2012; Slaughter et al., 2012). Furthermore, a recent study has shown that this immune memory can be stress-specific, with h-IR phenotypes only manifesting when the pathogens that established immune memory in the parental generation share similar lifestyles with those used to infect their progeny (López Sánchez et al., 2021). This holds similarity to IR phenotypes which have been reported within a generation, several weeks after a biotic stress stimulus has been applied (Bozarth and Ross, 1964; Luna et al., 2014; Wilkinson et al., 2022, 2023; Catoni et al., 2022). Hence, the establishment of immune memory upon exposure to biotic stress can be maintained into newly formed cell lines and tissues and, in some cases, even be transmitted to following generations.

A growing body of evidence highlights the significant role of epigenetic mechanisms in the regulation immune memory and associated IR phenotypes (Hannan Parker et al., 2022). Among these marks, DNA methylation has emerged as a key candidate, as it can be stably inherited through both mitotic and meiotic cell divisions (Hofmeister et al., 2017, 2020; Long et al., 2021). In support, numerous studies have reported that disrupting core regulators of DNA methylation in Arabidopsis are impacted in their ability to mount an effective IR response, both within a generation (Luna et al., 2014; Wilkinson et al., 2023), and across a generation (Luna et al., 2012; López Sánchez et al., 2016). Further evidence comes from studies using epigenetic recombinant inbred lines (epiRILs), which are derived from crosses of plants with identical genetic backgrounds, except for a mutation in one parent that causes severe losses in levels DNA methylation (Reinders et al., 2009; Johannes et al., 2009). The resulting populations exhibit a mosaic distribution of reduced DNA methylation (hypomethylation) across their chromosomes and show differential responses to phytohormones and varied susceptibility to plant pathogens (Latzel et al., 2012; Furci et al., 2019). Thus, variation in DNA methylation alone can determine plant defence phenotypes. Finally, several studies have shown that exposure to biotic stress is associated with genome-wide losses in DNA methylation, which can remain stable for several weeks, demonstrating that biotic stress has a direct influence on DNA methylation (Wilkinson et al., 2019; Hannan Parker et al., 2022; Catoni et al., 2022). However, few studies have directly investigated the impact of parental

biotic stress on DNA methylation in the progeny of these plants (Stassen et al., 2018), a gap that is addressed in this Chapter.

To establish immune memory, Arabidopsis plants were repeatedly infected with the bacterial pathogen *Pseudomonas syringae* pv. *tomato* DC3000 (*Pst*), a method previously shown to trigger h-IR against other (hemi-)biotrophic pathogens in the progeny (Luna et al., 2012; López Sánchez et al., 2016, 2021; Stassen et al., 2018; Furci et al., 2023). To test for h-IR, progeny were subsequently infected with the oomycete pathogen *Hyaloperonospora arabidopsidis* strain Waco9 (*Hpa*). While *Hpa* and *Pst* are taxonomically distinct pathogens with unique infection strategies (Slusarenko and Schlaich, 2003; Xin et al., 2016; Hu et al., 2022), both suppress defence responses associated with the phytohormone salicylic acid (SA), highlighting a critical role of SA in counteracting their infections (Brooks et al., 2005; Geng et al., 2012; Asai et al., 2014). As a result, exogenous application of SA effectively supresses infection of both *Hpa* and *Pst* (van Wees et al., 2000; Ton et al., 2002; Asai et al., 2014; Furci et al., 2021). SA-induced immunity is expected to occur through action of the SA-binding protein NONEXPRESSOR OF PATHOGENESIS-RELATED GENE 1 (NPR1), which associates with genic promoters to reprogram the expression of defence genes and other transcriptional regulators (Fan and Dong, 2002; Jin et al., 2018; Kumar et al., 2022). Highlighting the key role of SA and NPR1 in IR phenotypes, the Arabidopsis *npr1-1* mutant is unable to mount systemic acquired resistance (SAR) or h-IR responses to biotrophic pathogens (Cao et al., 1994, 1997; Luna et al., 2012).

Given the key role of SA-dependent defences and DNA methylation in regulating h-IR in Arabidopsis (Luna et al., 2012; López Sánchez et al., 2016), this Chapter first independently tests the importance of NPR1 and ROS1 in establishing immune memory and/or mounting h-IR responses against biotrophic pathogens by utilising the mutant lines *npr1-1* and *ros1-4*, respectively. Whole-genome DNA methylation sequencing was then performed on the progeny of Col-0 and *ros1-4* plants, which were capable and incapable of mounting h-IR responses, respectively. Upon failing to identify distinct regions of altered DNA methylation that could explain these phenotypes, I discuss the challenges of studying the molecular-epigenetic basis of variable h-IR responses in pooled progeny populations. Finally, I outline the development and validation of a novel molecular-genetic tool designed to study ROS1-dependent IR in Arabidopsis.

3.3 Materials and Methods

3.3.1 Plant Materials, Growth Conditions, and Chemical Treatments

For h-IR experiments (Figures 3.1, 3.2, 3.3), plants were grown in short-day conditions during parental treatments after which they were transferred to long-day conditions to trigger flowering. Growth conditions are detailed in Section 2.2. Arabidopsis genotypes Col-0, *ros1-4* (SALK_135293) (Tang et al., 2016), and *npr1-1* (Cao et al., 1994) were used. For details on the generation and selection of *XVE:ROS1-YFP* plants, see Section 2.4. Lines 5 and 7 (L5 and L7) were selected for further study as they displayed 3:1 segregation of kanamycin resistance and YFP fluorescence following 17β-estradiol (estradiol or E2; Sigma-Aldrich; E8875) treatment in the T2 generation, indicating the presence of a single *XVE:ROS1-YFP* construct.

For experiments with hydroponically grown plants (Figure 3.4c, 3.4d), seedlings were grown in long-day conditions (section 2.1) under sterile conditions. Seeds were vapour-phase sterilised (Lindsey et al., 2017) before sowing into 12-well plates. Each well contained ~5 seeds and 2 mL of ½ strength Murashige and Skoog (MS) basal salt mixture (Murashige and Skoog, 1962) without vitamins (MS; Sigma-Aldrich M5524) supplemented with 0.5% sucrose (pH = 5.7). Seeds in plates were stratified at 4°C in the dark for 2-4 days, after which the plates were moved to long-day conditions to trigger growth. After 7 days of growth, the MS media in the wells were replaced with fresh medium. The following day, 10 mM filter-sterilised (0.22 µm pore size) E2, dissolved in DMSO, was added to the wells at the desired concentration. An equivalent volume/volume percentage (v/v %) of DMSO was added in control wells. Further details about the preparation and storage of the E2 stocks are detailed in Section 2.3.

Plant cultivation for experiments with soil-grown *XVE:ROS1-YFP* plants (Figure 3.4e, 3.4f, 3.5) occurred under short-day conditions in 70 mL-pots containing a 2:1 soil:sand mixture, as detailed in Section 2.1. For the *XVE:ROS1-YFP* h-IR assay (Figure 3.5), plants were transferred to long-day conditions 28 DAS. Induction of the *XVE:ROS1:YFP* construct was performed by spraying 2-week old plants with 25 uM E2 or 0.05% DMSO supplemented with 0.015% (v/v) silwet L-77 (LEHLE SEEDS, VIS-30) (silwet) onto aerial tissue as described in Section 2.3. To standardise the dose, each tray containing 18 pots was sprayed with a total volume of 20 mL DMSO solution or E2 solution.

3.3.2 Pathogen Cultivation and Inoculation

Elicitation of h-IR in parental plants by *Pst* (p*Pst*) (Figure 3.1) was performed by 3 spray-inoculations at 21, 28, and 35 days after sowing (DAS) using GFP tagged *Pseudomonas*

syringae pv. tomato DC3000 *(Pst-GFP)* (Yu et al., 2013). This strain was stored in the form of 1 mL 20% glycerol stocks at -80°C, each derived from a single fluorescent colony. One day prior to inoculation, a glycerol stock was thawed on ice and then added to a sterile 100 mL Lysogeny Broth (LB) culture. The culture was left overnight (O/N) in an orbital shaker at 28°C and 200 rpm. *Pst-GFP* cells were harvested by spinning the O/N culture at 4000 x g for 8 mins at room temperature in two 50 mL falcon tubes. Cells were resuspended in an arbitrary volume of 10mM $MgSO_4$, after which the optical density was measured at 595 nm (OD_{595}) and adjusted to OD_{595} = 0.1 (~ 1 x 10^8 cells/mL). The final inoculum was supplemented with 0.01% (v/v) silwet and sprayed onto the shoots until saturated. Parental mock treatments (pMock) were performed using a solution of 10 mM $MgSO_4$ supplemented with 0.01% (v/v) silwet only. After each spray-treatment of *Pst-GFP* or mock, plants were maintained at 100% relative humidity (RH) for 2 days in sealed trays. At the end of the treatment course, plants were left to flower and seeds collected from 4 individual plants per treatment per genotype.

To quantify h-IR, 2-week-old generation 1 (G1) seedlings from *Pst*- and mock-inoculated parents (G0) were challenged with *Hyaloperonospora arabidopsidis* strain Waco9 (*Hpa*) at a spore density of 1 x 10^5 spores/mL in dH_2O and maintained at 100% RH in sealed trays. At 5 dpi, leaves were collected, stained with trypan blue, and assessed microscopically for *Hpa* colonisation. For details on propagation, maintenance, inoculation, staining, and scoring see Section 2.5.

To quantify pathogen colonisation in the *XVE:ROS1-YFP* bioassays (Figure 3.5), plants were either inoculated with *Pseudomonas syringae* pv. *tomato* DC3000 *luxCDABE* (*Pst-Lux*) (Fan, Crooks, and Lamb 2008) at an OD_{600} = 0.2, or inoculated with *Hpa* at a spore density of 1 x 10^5 spores/mL as described in Section 2.5. *Pst-Lux* colonisation was quantified non-destructively over multiple days by measuring *Pst-Lux* bioluminescence (Furci et al., 2021), as described in Section 2.5. Pathogen inoculations with *Hpa* and *Pst-Lux* in the *XVE:ROS1-YFP* lines were performed 48 hours after E2 treatment, and quantified, at earliest, 7 or 4 days after E2 treatment, respectively, which is much later than the ROS1 induction peak (Figure 3.4d) and time point at which ROS1-YFP protein is cleared from the nucleus (Figure 3.4f).

3.3.3 Sequencing and Analysis of DNA Methylation

To determine impacts of parental disease by *Pst* on the DNA methylome of Col-0 and *ros1-4* seedlings, seeds from 4 individual parents per genotype-treatment combination were sown and grown for 2-3 weeks in in long-day conditions (section 2.1). Between 20 and 35 seedlings of similar age were harvested, snap frozen in liquid nitrogen, and used to extract Genomic DNA with the GenElute Plant Genomic DNA Miniprep Kit (Sigma-Aldrich G2N350).

DNA quality and quantity were assessed using a Nanodrop 8000 spectrophotometer and a Qubit 3.0 Fluorometer. BGI Genomics performed the bisulfite conversion, library preparation, and 150 bp paired-end sequencing on a HiSeq X Ten System (Illumina). Read statistics of the whole-genome bisulfite sequencing (WGBS) are provided in Supplementary Data 3.2.

Reads were quality-assessed, using FastQC v0.12.1 and MuiltiQC v1.21 (Ewels et al., 2016). Adapters and poor-quality sequences were trimmed with Trimmomatic v0.39 (Bolger et al., 2014) (options: SLIDINGWINOW:4:25, HEADCROP:6, MINLEN:36). Quality-assessed trimmed reads then were aligned to the TAIR10 Arabidopsis genome using Bismark v0.24.1 (function: bismark; options: default parameters) (Krueger and Andrews, 2011), which utilises Bowtie2 v2.5.1 (Langmead and Salzberg, 2012). Using these default parameters, any reads which align ambiguously to the genome were excluded from downstream analyses of DNA methylation. While this carries the risk of reducing coverage at highly repetitive sequences (eg., TEs), it ensures that all methylation calls are of high confidence. PCR duplicate reads were removed using Bismark (function: deduplicate_bismark; options: default parameters). Binary scores for methylated cytosines (1) and unmethylated cytosines (0) in all contexts were generated and summarised for each position using Bismark (function: bismark_methylation_extractor; options: --paired-end, --no_overlap, --multicore 2, --comprehensive, --ignore_r2 2, --bedGraph, --CX). Bisulfite treatment non-conversion rates were calculated, using the unmethylated plastid genome as a control, and all samples passed the accepted threshold of <2%, with an average non-conversion rate of 0.54 % (Stuart et al., 2018) (Supplementary Data 3.2).

Principal component analysis (PCA) and hierarchical clustering analysis (HCA) were performed on all cytosine positions (allC) to investigate global shifts in DNA methylation. Positions with fewer than 5 mapped reads in any sample were discarded from the analysis. Levels of cytosine methylation were calculated for all remaining positions (Schultz et al., 2012). Sites which had a standard deviation less than or equal to the median of the standard deviations across all cytosine positions were removed to focus the computational analysis on the most variable positions. The PCA was conducted using base R v4.2.1 (function: prcomp; options: scale = FALSE, center = TRUE). A Euclidean distance matrix was generated in base R (function: dist; options: method = "euclidean") and HCA performed using base R (function: hclust; options: method = “ward.D2”). Plots were generated using the R package ggplot2 v3.5.0 and dendextend v1.17.1 (Galili, 2015).

To call differentially methylated regions (DMRs) of variable sizes, the R package DSS v2.46.0 was used (H. Feng et al., 2014; Park and Wu, 2016). As DSS accounts for read coverage, no filtering was performed on the data. Differentially methylated cytosines were first

called using DSS (function: DMLtest; options: equal.disp = FALSE, smoothing = FALSE) which were then used to call DMRs with DSS (function: callDMR; options: minlen = 25, dis.merge = 50, minCG = 3, p.threshold = 0.05, delta = 0.1, pct.sig = 0.5). This pipeline was performed on cytosines in all (allC), CG, CHG, or CHH contexts to generate context specific DMRs. Annotation of DMRs was based on Araport11 (Cheng et al., 2017) and the TAIR10 transposable element annotation file (downloaded from arabidopsis.org; accessed December 2023). The analysis was performed using the R packages genomation v1.34.0 and GenomicRanges v1.54.1 (Lawrence et al., 2013; Akalin et al., 2015). In the case where a DMR overlapped multiple features, a single annotation was assigned in the following order of preference: gene body > promoter > intergenic.

ROS1-dependent h-IR-associated DMRs were defined as DMRs that were hypermethylated in the *ros1-4* (pMock) vs Col-0 (pMock) comparison (*ros1-4* pMock hyper-DMRs), and hypomethylated in the Col-0 (p*Pst*) vs Col-0 (pMock) comparison (Col-0 p*Pst* hypo-DMRs) (Figure 3b). For a chosen set of DMRs, all hyper- or hypomethylated DMRs from all contexts were pooled. If DMRs overlapped between contexts, they were merged to form a single larger DMR using the R package GenomicRanges (function: reduce; options: default parameters). All hypomethylated DMRs linked to parental *Pst* stress in Col-0 from previously published independent experiments (Stassen et al., 2018) were pooled and merged to form a large DMR set called ‘Col-0 p*Pst* hypo-DMR (Stassen et al 2018)’ (Figure 3.3b). Overlapping ranges for these various comparisons were identified using GenomicRanges (function: findOverlaps; options: select = "all", maxgap = 0).

3.3.4 Epifluorescence Microscopy and Imaging Analysis of ROS1-YFP Accumulation

Epi-fluorescence microscopy to quantify accumulation of ROS1-YFP in living plants was performed using a Leica M165 FC fluorescent stereo microscope (Objective: 1x/0.06; ET-GFP filter set: 470/40 nm excitation, 495 nm dichroic and 525/50 nm emission) with a CoolLED pE-300 illumination system, controlled by Leica LAS X software v3.7.4.23463.

To quantify ROS1-YFP fluorescence intensity in hydroponically grown seedlings (Figures 3.4c, Figure 3.4d), bright field (BF) and YFP images were taken for each well separately. In addition, 10 wells containing only media were imaged using the same YFP and BF settings. Regions of interest were defined within a well using the oval tool in Fiji/imageJ. The background YFP signal across all images was removed (process > math > subtract) by calculating the average mean-grey value (Analyze > Analyze Particles) plus three times the standard deviation of the mean-grey value in the ten empty wells. The YFP intensity was then normalised by plant area within the well. To account for autofluorescence, ROS1-YFP signal was calculated as a relative fold change compared to DMSO-treated plants.

To validate nuclear localisation of ROS1-YFP signal, leaves were detached from soil-grown plants 24 h after spraying with 25 μM E2 and vacuum infiltrated with DAPI solution (2 μg/ml) for 2-3 minutes. Leaves were mounted on slides and imaged using a Leica DM6B epifluorescence microscope (light source: CoolLED pE-2; Objective: HC PL Fluotar 40x/0.80; ET-DAPI filter set: 350/50 nm excitation, 400 nm dichroic, and 460/50 nm emission; ET-GFP filter set: 470/40 nm excitation, 495 nm dichroic, and 525/50 nm emission) controlled by Leica LAS X v3.7.3.23245 software. Using Fiji/ImageJ (Rueden et al., 2017), the brightness and contrast were increased for the BF images (20% and 10%, respectively), DAPI images (60% and 70%, respectively), and YFP images (40% and 70%, respectively). DAPI and YFP images were then pseudo-coloured cyan and yellow, respectively, using Fiji/ImageJ.

3.3.5 Reverse Transcriptase-Quantitative Polymerase Chain Reactions (RT-qPCR)

ROS1 gene expression in soil grown Col-0, L5, and L7 seedlings spray-treated with 0.015% DMSO or 25 μM E2 (as described above) was quantified using RT-qPCR. For each biological replicate (n=4), 24 hours after chemical induction, aerial tissue of ~8 seedlings were harvested together, flash frozen in liquid nitrogen, and stored at -80°C. Subsequently, tissue was ground to a fine powder using a pestle and mortar with liquid nitrogen. 100 mg of ground tissue was used for RNA extractions using the RNeasy plant mini kit (Qiagen, #74904). RNA concentration was measured using a Nanodrop 8000 spectrophotometer, and 800 ng of RNA was used for cDNA synthesis using the 'Maxima First Strand cDNA Synthesis Kit for RT-qPCR, with dsDNase' kit (Thermo Scientific, K1672). Upon completion of cDNA synthesis, samples were diluted 1:5 using nuclease free water. Subsequently, qPCR was performed using SsoAdvanced SYRB Green Supermix (BioRad, 1725270), *ROS1*-specific primers (Table 3.1) (Duan et al., 2015), and a Rotor-Gene Q (Qiagen) real-time PCR cycler. Cycle conditions are as follows: 1 cycle of 30 seconds at 95°C, 40 cycles of 10 seconds at 95°C and 30 seconds at 60°C. Relative expression values were calculated as described in Wilkinson et al (2023), using *MON1 (AT2G28390)* and *UBC21 (AT5G25760)* as reference genes (Czechowski et al., 2005), and using a DMSO-treated Col-0 sample as a calibrator sample.

Table 3.1 Primers used for RT-qPCR of *ROS1*.

	Name	Nucleotide sequence
RT-qPCR of *ROS1*	ROS1-F1	AAAACTACCCCTCATCGCTG
	ROS1-R2	GTTAGTACGTGCATATTCCAAGC
RT-qPCR of reference gene *UBC21*	UBC21-F1	CTGCGACTCAGGGAATCTTCTAA
	UBC21-R1	TTGTGCCATTGAATTGAACCC
RT-qPCR of reference gene *MON1* from RNA	MON1-F1	AACTCTATGCAGCATTTGATCCACT
	MON1-R1	TGATTGCATATCTTTATCGCCATC

3.4 Results and Discussion

3.4.1 ROS1 and NPR1 are Key for Basal Resistance and Heritable Induced Resistance Against *Hpa*

To investigate the roles of NPR1 and ROS1 in basal resistance and long-term heritable induced resistance in Arabidopsis, a standard heritable induced resistance (h-IR) assay was performed (Furci et al., 2023) using the wild-type ecotype Col-0 and the mutants *ros1-4* (Tang et al., 2016) and *npr1-1* (Cao et al., 1994). Parental stress was induced by spray-inoculating parental plants three times with GFP-tagged *Pseudomonas syringae* pv. *tomato* DC3000 (*Pst-GFP*) (Yu et al., 2013) or a mock solution. Progeny (generation 1; G1) from four parental plants (G0) per treatment-genotype combination were challenged with *Hyaloperonospora arabidopsidis* strain Waco9 (*Hpa*) (Figure 3.1a). *Hpa* colonisation was more severe in *ros1-4* and *npr1-1* mutant plants compared to Col-0 (Figure 3.1b) ($p_{adj} < 0.05$). Disease by *Pst-GFP* in the previous generation had no influence on *Hpa* resistance in either mutant line. In contrast, wild-type Col-0 plants derived from *Pst-GFP*-infected parents were more resistant to *Hpa* compared to those that experienced no stress in the previous generation (Figure 3.1b). This indicates that both ROS1 and NPR1 are required for the establishment, maintenance and/or activation of h-IR against *Hpa* following parental *Pst* infection, supporting previous findings from this h-IR model system (Luna et al., 2012; López Sánchez et al., 2016; Furci et al., 2023).

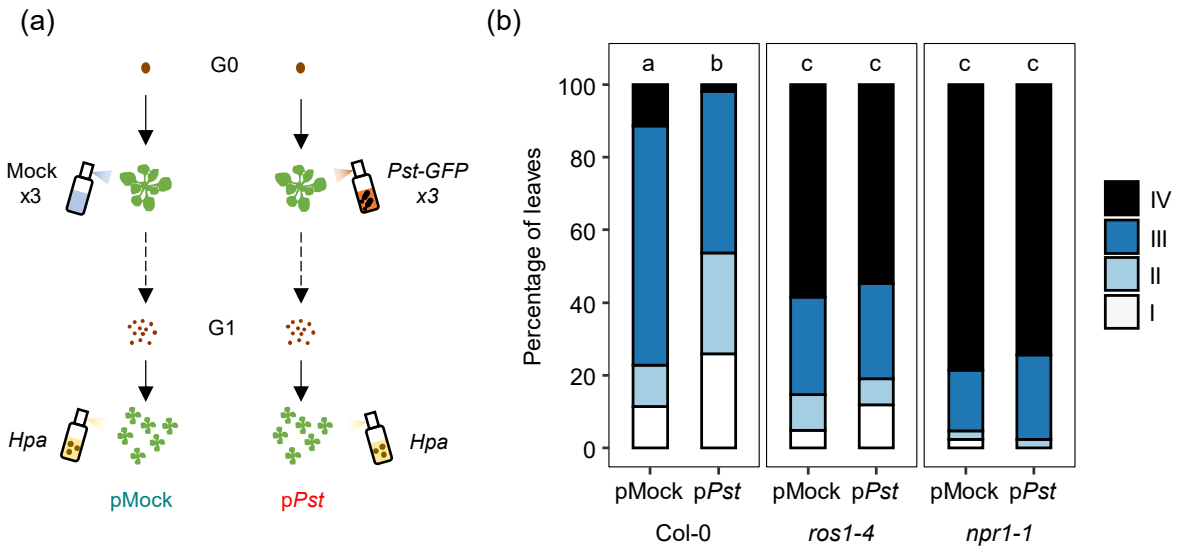


Figure 3.1 Heritable induced resistance (h-IR) phenotypes in Col-0, *ros1-4*, and *npr1-1.*
(a) Schematic outlining the heritable induced resistance (h-IR) assay. Four parental plants (G0) from a single genotype were subject to 3 spray-inoculations of GFP-tagged *Pseudomonas syringae* pv. *tomato* DC3000 (*Pst-GFP*) (Yu et al., 2013) (OD_{600} = 0.1) or a 10 mM $MgSO_4$ mock treatment, supplemented with 0.01% silwet-L77. After flowering, the seeds from these plants (G1) were collected, sown, and grown for 3 weeks before spray inoculation with *Hyaloperonospora arabidopsidis* (*Hpa*) at a spore density of $1x10^5$ condiospores ml^{-1}. Parental *Pst-GFP* treatment = p*Pst*; parental mock treatment = pMock. (b) Levels of h-IR to *Hpa* in Col-0, *ros1-4*, and *npr1-1,* following parental *Pst-GFP* treatment (p*Pst*), or parental mock treatment (pMock). Five days post infection (dpi), seedlings were pooled per parental treatment, and stained with trypan blue. Individual leaves (n = 35-54) were microscopically assigned to four different *Hpa* colonisation classes: I: no hyphal colonisation; II: limited hyphal colonisation, ≤ 8 conidiophores; III: extensive hyphal colonisation, > 8 conidiophores; class IV: extensive hyphal colonisation with conidiophores and oospore(s), as described in Section 2.5 (Schwarzenbacher et al., 2020). Shown are the proportion of leaves assigned to each colonisation class. Treatments with the same letter are not significantly different (Fisher's exact test; *q*-value > 0.05; FDR correction).

3.4.2 Exposure to *Pst* in the Previous Generation Limited but Detectable Impact on DNA Methylation

The DNA demethylase ROS1 of Arabidopsis reduces DNA methylation and is particularly active at transposon-gene boundaries (Tang et al., 2016). Consequently, this chromatin re-modeller enzyme has been implicated in the transcriptional regulation of environmentally responsive genes (Halter et al., 2021). Since the mutations in *ROS1* reduce basal resistance to biotrophic pathogens (Yu et al., 2013; López Sánchez et al., 2016) and affect to h-IR against *Hpa* (López Sánchez et al., 2016; Figure 3.1b), the ROS1-dependent contribution to the DNA methylome of h-IR was investigated by whole-genome bisulfite sequencing (WGBS) of seedlings from the same pool of G1 seeds that had been used for the *Hpa* bioassay (Figure 3.1a). For each sample, 20-35 seedlings were pooled for DNA

extractions and sequencing (n=4). Principal component analysis (PCA; Figure 3.2a) and hierarchical cluster analysis (HCA; Figure 3.2b) of DNA methylation at a single cytosine resolution revealed that parental treatment had limited impact on the methylome of seedlings, and most of the variation in the dataset is driven by the *ros1-4* mutation. Nonetheless, the HCA analysis grouped seedlings of p*Pst*-treated Col-0 parents separately from seedlings of pMock-treated Col-0 parents (Figure 3.2b), which was not apparent for *ros1-4*. This indicates that Col-0 seedlings have subtle infection-dependent changes in DNA methylation that do not occur in ROS1-deficient plants, which aligns with the findings of a previous DNA methylome analysis of h-IR (Stassen et al., 2018).

To identify methylation patterns associated with h-IR, differentially methylated regions (DMRs) across the genome were searched using DSS (Park and Wu, 2016). The analysis was performed for all cytosine contexts (allC) as well as distinct cytosine contexts (CG, CHG, and CHH). The comparison between *ros1-4* seedlings (pMock) and Col-0 seedlings (pMock) yielded thousands of DMRs with a significant bias towards hyper-methylated DMRs (two-sided exact binomial test, $p < 0.001$), corroborating previous results and validating the WGBS analysis (Figure 3.2c) (Tang et al. 2016). In contrast, for both Col-0 and *ros1-4*, only a marginal number of DMRs (228 and 248, respectively) could be detected between progeny from *Pst-GFP*-treated plants and mock-treated plants (Figure 3.2c). Hence, exposure to *Pst* in the previous generation has only a marginal impact on DNA methylation as detected in our dataset.

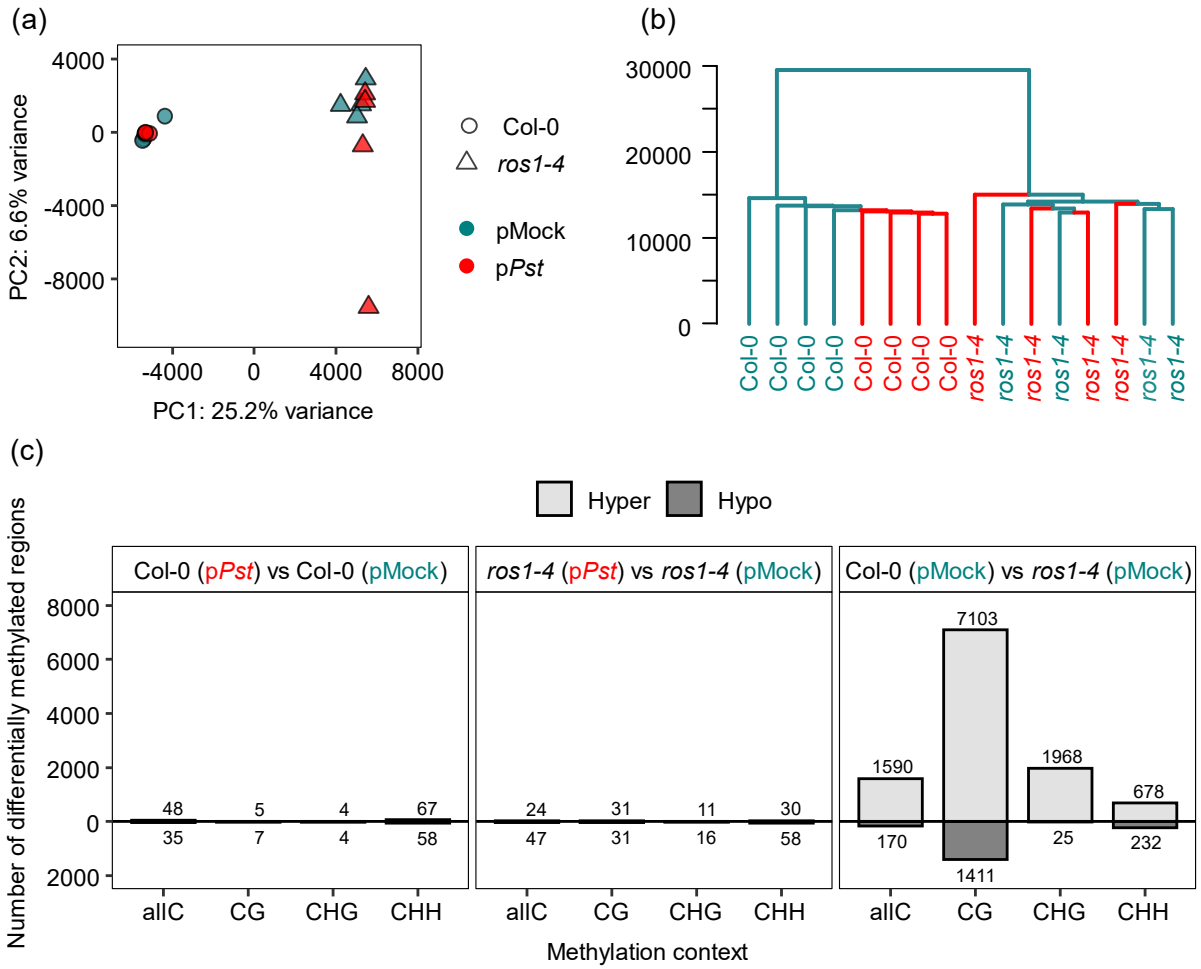


Figure 3.2 Global DNA methylation patterns in progeny from Col-0 and *ros1-4* progeny of mock- and *Pseudomonas syringae* pv. *tomato* DC300 (*Pst*)-inoculated plants.
(a) PCA plot displaying variation in global cytosine methylation (all cytosine contexts) for Col-0 (circle) and *ros1-4* (triangle) seedlings, derived from parental plants treated with mock (pMock; blue) or *Pst* (p*Pst*; red) as illustrated in Figure 3.1a. A total of 20-35 seedlings from a single parent were pooled to form one biologically replicated sample (n=4). (b) Hierarchical clustering (Euclidean distance, Ward) of global cytosine methylation (all cytosine contexts) for genotypes Col-0 and *ros1-4*, and parental treatments (pMock: blue and p*Pst*: red). (c) Frequency of differentially methylated regions (DMRs) for the following comparisons: Col-0 p*Pst* vs Col-0 pMock (left panel), *npr1-4* p*Pst* vs. *ros1-4* pMock (middle panel), and Col-0 pMock vs. pMock *ros1-4.* Hyper- and hypomethylated DMRs are shown above and below the x-axis in light and dark grey shades, respectively.

3.4.3 The Variability of ROS1-Dependent h-IR and Associated DMRs Pose an Obstacle to Study the Underpinning Mechanisms

To better understand the genomic context in which h-IR-related DMRs occur, the association of DMRs with gene bodies, gene promoters, intergenic regions, and transposable elements (TEs) was determined (Supplementary Data 3.3) (Figure 3.3a). DMRs between progeny from *Pst*-and Mock-treated Col-0 occurred predominantly in allC and CHH contexts and were largely associated with intergenic TEs (Figure 3.3a). Interestingly, in the *ros1-4* background, the same comparison revealed fewer DMRs in the allC and CHH contexts but an increase in DMRs within the CG and CHG contexts (Figure 3.3a).This difference is potentially relevant as loss of DNA methylation in TE-rich pericentromeric regions is linked to heritable quantitative resistance against *Hpa* (Furci et al., 2019). However, more *pPst*-associated DMRs in Col-0 were hypermethylated than hypomethylated (Figure 3.3a), and the DMRs were relatively small, averaging ~40 bp in width (Supplementary Data 3.3).

Many of the hypermethylated DMRs in *ros1-4* occurred within gene bodies or gene promoters (Figure 3.3a). As DNA methylation can prevent efficient binding of immune-responsive transcription factors (O'Malley et al., 2016; Charvin et al., 2023), the enhanced susceptibility of *ros1-4* to *Hpa* (Figure 3.1a) may be caused by DMRs repressing defence genes in *cis*. For instance, a hypermethylated DMR has been reported in the promoter of *RLP43* in the *ros1-3* mutant, which affects basal immunity against *Pst* (Halter et al., 2021). This same DMR has been detected in our dataset for the *ros1-4* mutant (Supplementary Data 3.3). *Cis*-regulatory DMRs like these could play a role in immunity against *Hpa*, making them promising candidates for involvement in ROS1-dependent h-IR. Therefore, to search for candidate hypomethylated DMRs contributing to h-IR, hypermethylated DMRs in *ros1-4* (ros1-4 pMock hyper-DMRs) were overlapped with hypomethylated DMRs from the progeny of *Pst*-treated Col-0 plants (Col-0 p*Pst* hypo-DMRs). Additionally, hypomethylated DMRs from the progeny of *Pst*-treated *ros1-4* plants, which could not establish h-IR, were overlapped as a negative control (*ros1-4* p*Pst* hypo-DMRs) (Figure 3.3b). Lastly, hypomethylated DMRs from the progeny of *Pst*-treated Col-0 plants which were associated with h-IR to *Hpa* in a previous study (Stassen et al., 2018) (Col-0 pPst hypo-DMRs Stassen et al. 2018) were included in the analysis. In all 4 of these DMR sets, DMRs were merged across all cytosine contexts for simplicity, and are summarised in Supplementary Data 3.4. This analysis revealed only 4 hypermethylated DMRs in *ros1-4* that mapped to the same location as hypomethylated DMRs in the progeny of *Pst*-treated Col-0 (Figure 3.3b). Of these, 3 mapped to TEs, and 1 mapped to the second intron of the gene locus AT4G10603 (Figure 3.3c), which is predicted to encode a defensin-like protein (Silverstein et al., 2005). Interestingly, this same h-IR-associated DMR was identified in Stassen et al. (2018). However, AT4G10603 is exclusively expressed in flower

buds (Klepikova et al., 2016) (Supplementary Figure 3.1) and has been identified as a pollen coat protein (L. Wang et al., 2023). Therefore, while this DMR correlates well with h-IR, it is unlikely to have a role in defence against *Hpa* in leaf tissues. Beyond this single DMR, no other h-IR-associated DMRs were consistent with those identified by Stassen et al. (2018) (Figure 3.3b).

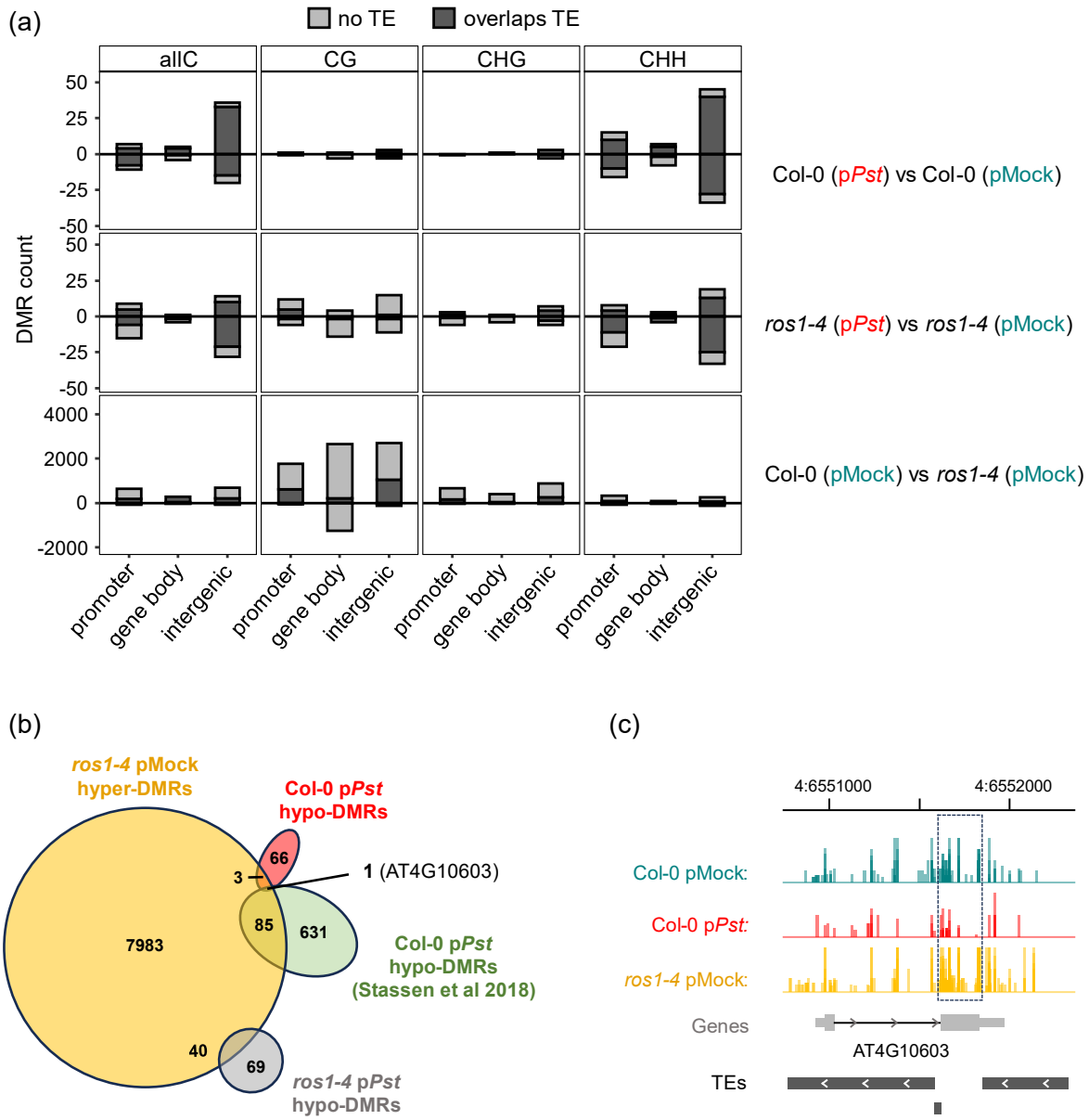


Figure 3.3 Annotated counts of differentially methylated regions (DMRs) between progenies of *Pst* (*pPst*)- or mock- (pMock)-inoculated Col-0 and *ros1-4* plants.
(a) Number of DMRs annotated as overlapping with gene promoters (< 1 kb upstream of gene transcriptional start site - TSS), gene body (TSS to transcriptional termination site - TTS), or intergenic (> 1 kb upstream of gene TSS) for each comparison. DMRs that also overlap with a transposable element (TE) within each region are indicated by dark gray shading. p*Pst* and pMock: progeny of *Pst*- or mock-inoculated parents, respectively (see Figure 3.1). (b) Venn diagram overlapping context-merged hypomethylated DMRs (hypo-DMRs) or hypermethylated DMRs (hypo-DMRs) for the following comparisons: hyper-DMRs from *ros1-4* pMock vs Col-0 pMock (yellow), hypo-DMRs from Col-0 p*Pst* vs Col-0 pMock (red), hypo-DMRs from previously published Col-0 p*Pst* vs Col-0 pMock (Stassen et al., 2018; any of 3 described experiments) (green), and hypo-DMRs from *ros1-4* p*Pst* vs *ros1-4* pMock (grey). (c) Integrative Genomics Viewer snapshot of AT4G10603 showing methylation percentage (limits: 0-100%) for Col-0 pMock (track 1; blue), Col-0 p*Pst* (track 2; red), and *ros1-4* pMock (track 3; yellow). Bottom track shows annotation of genes (grey) and TEs (dark grey). Methylation tracks contain 4 overlayed tracks from 4 biological replicates.

3.4.4 Technical Challenges Associated With the Identification of Stress-Induced Heritable Changes in DNA Methylation

A plausible explanation for the failure of the WGBS analysis to identify consistent candidate DMRs underlying h-IR lies in the experimental design. Firstly, stress-induced changes in DNA methylation likely vary among individuals in the parental generation. For example, when exposed to jasmonic acid, Arabidopsis plants generally exhibit a reduction in DNA methylation at TEs in the ATREP2 family, but the specific TEs affected differ among plants (Wilkinson et al. 2023). Secondly, during reproduction, there is significant reprogramming of the epigenetic landscape, with only some epigenetic marks being faithfully transmitted to the next generation (Calarco et al., 2012; Long et al., 2021; Tirot and Jullien, 2022). As a result, plants derived from a single parent and grown in identical conditions will diverge and fluctuate in their DNA methylation patterns over successive generations (Becker et al., 2011; Johannes and Schmitz, 2019; Zhang et al., 2021; Briffa et al., 2023). Interestingly, a recent study has shown that biotic stress elevates the frequency of spontaneous epimutations, resulting in increased divergence in DNA methylation patterns over time in stressed plants (Zhang et al., 2024).Therefore, comparing the methylome of 20-35 pooled seedlings from 4 parental plants introduces multiple layers of variability that are extremely challenging to disentangle. One way to overcome this would involve harvesting leaves of individual plants for DNA methylation sequencing and subsequently phenotyping the same plants for h-IR responses. However, harvesting material could trigger wound-associated defence responses, potentially influencing the resistance of the plants to pathogens. It would also be challenging to harvest enough material from a single plant to extract enough DNA for next-generation sequencing. Furthermore, the increased sensitivity of this approach would incur substantial financial costs, as many plants would need to be sequenced for DNA methylation to characterise h-IR-associated DMRs.

Alternatively, the use of mutant lines that are compromised in DNA methylation machinery, such as the *ros1-4* mutant, can generate more consistent and pronounced changes in DNA methylation (as evidenced in Figures 3.2 and 3.3), making them valuable tools for understanding the functional roles of the associated proteins. To further dissect the role of *ros1* in establishing or maintaining h-IR, crosses between wild-type plants and *ros1* mutants could be created. The segregating F_1 population would then be subjected to a standard h-IR assay (Figure 3.1a) (Furci et al., 2023). If h-IR phenotypes co-segregate with the presence/absence of ROS1 in the F_2 population, this would indicate a critical role of ROS1 in maintaining h-IR-associated epialleles over reproduction. However, if h-IR phenotypes did not co-segregate with the *ros1* mutation, and all progeny displayed h-IR phenotypes, it would indicate that ROS1 is important for the establishment of immune memory in the parental

generation, but not necessarily the maintenance of these h-IR epialleles over reproduction. However, permanent changes caused by genetic mutation do not reflect the reversable nature of stress-induced epigenetic changes and often become progressively more severe over generations, sometimes rendering plants inviable (Jeddeloh et al., 1998; Mathieu et al., 2007; Saze et al., 2008; Williams and Gehring, 2017; Zhang et al., 2021). Furthermore, dysfunction of one protein can lead to hyperactivity of another gene/protein through feedback mechanisms. For example, mutants in RNA-directed DNA methylation (RdDM) machinery, which exhibit lower levels of DNA methylation, have an approximately 10-fold decrease in *ROS1* transcript abundance, resulting in compensatory accumulation of DNA methylation (Williams et al., 2015; Tang et al., 2016). Similarly, plants lacking METHYLTRANSFERASE 1 (MET1) have drastic reductions in CG DNA methylation but exhibit ectopic accumulation of non-CG methylation and H3K9me2 marks due to reduced expression of *IBM1* (Rigal et al., 2012; Yinwen Zhang et al., 2024) (see Figure 1.2 in Chapter 1).

Therefore, the use of an inducible system that allows for controlled introduction of ROS1-dependent DNA demethylation is proposed and characterised below. This system is expected to mimic a transient epigenetic shock more closely, where DNA methylation is lost in response to natural stress (Dowen et al., 2012; López Sánchez et al., 2016; Geng et al., 2019; Atighi et al., 2020; Annacondia et al., 2021; Wilkinson et al., 2023; Lee et al., 2023), compared to the use of a genetic mutation lines such as *ros1-4*, which harbour permanent and non-reversible changes. Furthermore, since ROS1 has distinct targets (Tang et al., 2016), the induced changes in DNA methylation are expected to be less variable than those caused by pathogens. Altogether, this tool could provide a more sensitive and robust system to characterise regions of DNA methylation that contribute to basal immunity and/or long-term induced resistance.

3.4.5 The XVE Inducible System Enables Controlled Activation of ROS1 in Arabidopsis

To address the challenges associated with the inherent variability of h-IR, a system allowing for controlled induction of DNA demethylation was developed. The results from the h-IR experiment (Figure 3.1b), along with findings from previous studies (López Sánchez et al., 2016; Halter et al., 2021), highlight the importance of the DNA demethylase ROS1 in regulating innate and induced immunity. Therefore, transgenic Arabidopsis lines were generated carrying an inducible ROS1 construct using the XVE transactivating system (Zuo et al., 2000; Siligato et al., 2016) (Figure 3.4a). XVE is a chimeric protein formed of a DNA-binding domain for the bacterial repressor LexA (X), the acidic transactivating domain of VP16 (V), and the carboxyl region of the human estrogen receptor (E) (Zuo et al., 2000). When the

human hormone 17β-estradiol (E2) binds to the E domain of XVE, the protein undergoes a conformational change which facilitates its binding to LexA operator sequences, which are fused to a minimal 35S promoter (Benfey et al., 1990; Zuo et al., 2000). In this construct, the E2-induced binding of XVE to the LexA operator sequences induces transcription of the *ROS1* gene fused with *YFP* (*ROS1-YFP*). This plasmid was named *pR-XVE-gROS1-YFP* and was transformed into Col-0 Arabidopsis plants to form transgenic *XVE:ROS1-YFP* plants (for details on the construction and transformation of this plasmid see Section 2.4) (Figures 3.4a, 3.4b).

To validate the functionality of the *XVE:ROS1-YFP* construct, the accumulation of the ROS1-YFP protein was quantified by epi-fluorescence microscopy imaging in hydroponically grown Col-0 seedlings (serving as the negative control) and two independent *XVE:ROS1-YFP* lines, designated as line 5 (L5) and line 7 (L7), 24 hours after supplementing the medium with increasing concentrations of E2 (Figure 3.4c). Pairwise t-tests with a false discovery rate (FDR) correction on the *p*-values (resulting in *q*-values) revealed that the relative YFP signal in L5 and L7 was significantly higher than the background level in Col-0 at E2 concentrations ≥1 μM ($q < 0.05$). Moreover, the relative YFP signal showed a dose-dependent increase in both *XVE:ROS1-YFP* lines until saturation at approximately 25 μM E2 (Figure 3.4c).

A time-course analysis was then performed at two E2 doses, 5 μM and 25 μM (Figure 3.4d), using hydroponically grown Col-0, L5, and L7 seedlings (Figure 3.4d). There was no significant effect of the interaction between genotype, E2 dose, and time on relative YFP signal ($F = 0.1$; $df = 4, 36$; $p > 0.05$). However, the interaction between time and genotype, irrespective of dose, had a highly significant effect on relative YFP signal ($F = 6.4$; $df = 4, 36$; $p < 0.001$). For both 5 μM and 25 μM treatments, relative YFP signal in L5, but not L7, was significantly higher than Col-0 as early as 6 hours post-treatment ($q < 0.05$) (Figure 3.4d). At 24 hours post treatment, both *XVE:ROS1-YFP* lines exhibited a significantly increased YFP signal compared to Col-0 ($q < 0.01$), which disappeared by 48 hours post-E2 treatment ($q > 0.05$) (Figure 3.4d). Therefore, E2 induces transient accumulation of ROS1-YFP, indicating that the recombinant protein is no longer expressed and/or is degraded 24 hours after E2 treatment, regardless of the dose.

The induction of the construct in soil-grown plants was tested by spraying 2-week-old seedlings with 25 μM E2 or 0.05% DMSO, supplemented with 0.015% Silwet L-77 (silwet). As a surfactant, silwet ensures a more homogeneous distribution of the chemical solution on a the hydrophobic surface of leaves (Chengara et al., 2007). Relative accumulation of ROS1 transcripts increased approximately 200-fold in L5 and L7 24 hours after E2 treatment, compared to native ROS1 expression in Col-0 (Figure 3.4e). Furthermore, no changes in

ROS1 expression or YFP signal were detected in Col-0 plants in response to E2 (Figures 3.4e, 3.4f). DAPI staining showed that the enhanced ROS1-YFP fluorescence in L5 and L7 treated with 25 μM E2 predominantly co-localised with nuclear DAPI fluorescence, confirming nuclear localisation of the recombinant protein (Figure 3.4f). While DMSO-treated L5 and L7 had a mild but significant increase in ROS1 transcript abundance (1.7x and 2.3x, respectively) compared to DMSO-treated Col-0 (t-test, $p < 0.05$), there was no detectable increase in ROS1-YFP signal (Figures 3.4e, 3.4f). Therefore, even with mild 'leakiness' of the construct, the *XVE:ROS1-YFP* system is tightly controlled by E2 induction (Zuo et al., 2000), allowing for high levels of transient ROS1-YFP accumulation in the nucleus of plants.

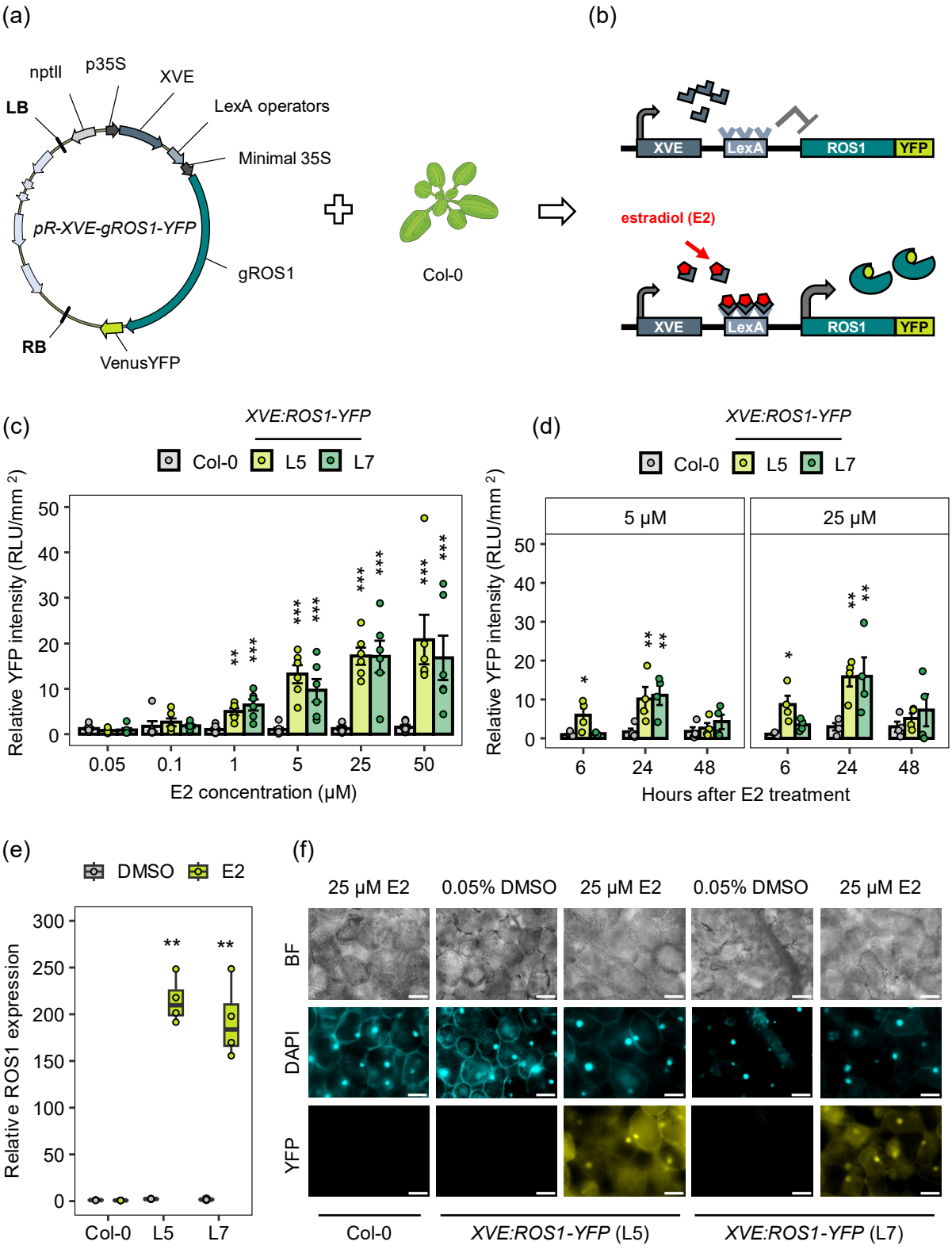


Figure 3.4 Characterisation of *XVE:ROS1-YFP* lines.
(a) Plasmid map of estradiol (E2)-inducible *XVE:ROS1-YFP* construct used to transform Col-0. (b) Schematic demonstrating the XVE E2-inducible system. See main text for details. (c) ROS1-YFP protein accumulation upon treatment with increasing E2 concentrations of hydroponically grown Col-0, *XVE:ROS1-YFP* (L5) and *XVE:ROS1-YFP* (L7) seedlings at 24 hours after treatment of 7-day-old seedlings. Data points represent relative YFP fluorescence intensities from biologically replicated seedlings (n = 6), normalised against the mean relative YFP fluorescence intensity in DMSO-treated

control seedlings. Asterisks indicate a statistically significant difference in YFP fluorescence intensity compared to Col-0 plants (t-test; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). (d) ROS1-YFP protein accumulation at 6, 24, and 48 hours after treatment with 5 µM or 25 µM E2 to 7-day-old hydroponically grown seedlings, measured as relative light units (RLU) per plant area (mm^2). Data points represent YFP fluorescence intensities from biological replicates (n = 4) relative to the mean YFP fluorescence intensity value of DMSO-treated controls for a given genotype. For each time-point, pairwise t-tests against Col-0 were performed and *p*-values adjusted for using the FDR correction (*q*; * $q < 0.05$, ** $q < 0.01$, *** $q < 0.001$). (e) Quantification of *ROS1* transcripts by RT-qPCR at 24 hours after spraying 2-week-old soil-grown Col-0, L5, and L7 seedlings with 25 µM E2 or 0.05% DMSO, supplemented with 0.015% silwet. Data points represent expression values of individual biological replicates (n=4) relative to the mean relative expression value of DMSO-treated Col-0 seedlings. Asterisks indicate a statistically significant difference between treatments within a genotype (t-test; ** $p < 0.01$, *** $p < 0.001$). (f) ROS1-YFP protein accumulation was imaged at 24 hours after the same treatment as (e) by epi-fluorescence microscopy. Brightness and contrast of all raw images were increased for the bright field (BF) images (20% and 10%, respectively), DAPI images (60% and 70%, respectively), and YFP images (40% and 70%). DAPI and YFP images were pseudo-coloured cyan and yellow, respectively. Scale bars = 50 µm.

3.4.6 Transient Induction of ROS1 Induces Resistance Against Biotrophic pathogens.

The *ros1-4* mutant is more susceptible to the biotrophic oomycete *Hpa* and the hemi-biotrophic bacterium *Pst* (Figure 3.1b) (Halter et al., 2021). To investigate how transient overexpression of *ROS1* influences resistance against *Hpa* and *Pst*, Col-0, L5, and L7 seedlings were treated twice with E2 prior to infection (Figure 3.5a). The first induction was performed 14 days after sowing (14 DAS), using a concentration of 10 µM E2. A second induction treatment was performed 4 days later (18 DAS) using a concentration of 25 µM E2 (Figure 3.5a). The equivalent volume/volume percentage (v/v %) of DMSO was used as a control treatment and all solutions were supplemented with 0.015% silwet. Two days after the final E2 treatment, plants were spray-inoculated with either *Hpa* (1×10^5 spores ml^{-1} in dH_2O) or the bioluminescent *Pst luxCDABE* (*Pst-Lux*) strain (Fan et al., 2008) ($OD_{600} = 0.2$ in 10 mM $MgSO_4$ with 0.015% silwet) (Figure 3.5a). This dual E2 treatment was chosen based on the hypothesis that it would more accurately mimic the escalating stress encountered during infection by a virulent pathogen which proliferates within plant tissues. In agreement, data presented in Chapter 6 of this thesis highlight that multiple treatments of E2 are required to establish resistance against *Pst* (Figure 6.2).

Colonisation by *Hpa* was microscopically analysed in trypan blue-stained leaves at 5 days post inoculation (dpi), and individual leaves (n = 72-80) were assigned to *Hpa* colonisation classes, ranging from I (no colonisation) to IV (severe colonisation and sporulation) (see Section 2.5) (Figure 3.5b). Pairwise statistical comparisons of the distribution of leaves across the four colonisation classes were performed using Fisher's exact test with FDR-corrected *p*-values (*q*-values). Wild-type Col-0 plants treated with DMSO or E2 showed no

difference in susceptibility to *Hpa* (q > 0.05). However, E2-treated lines harbouring the *XVE:ROS1-YFP* construct (L5 and L7) were more resistant compared to those treated with DMSO (L5: q < 0.001; L7: q < 0.001) (Figure 3.5b). *Hpa* colonisation in L5 and L7 showed no statistically significant difference within either the DMSO-treated or E2-treated groups (q > 0.05). Furthermore, DMSO-treated *XVE:ROS1-YFP* lines were as susceptible as DMSO-treated Col-0 and E2-treated Col-0 seedlings (q > 0.05) (Figure 3.5b). It can thus be concluded that E2 has no impact on *Hpa* resistance in Col-0 and that the transient accumulation of ROS1-YFP upon E2 treatment in *XVE:ROS1-YFP*-harbouring lines induces resistance against *Hpa*.

Colonisation by *Pst-Lux* was non-destructively monitored from 1 to 4 dpi using bioluminescence as a quantitative measure of disease (Fan et al., 2008; Furci et al., 2021) (Figures 3.5c, 3.5d). A linear mixed-effects model, which accounted for the random effects of dpi and pot, was fitted to the square-root transformed data from 2-4 dpi. The omission of data from 1 dpi was necessary to comply with the assumption of normality, as most plants exhibited little to no *Pst-Lux* bioluminescence at this time point. Using likelihood ratio tests (LRTs) between the full model and reduced models, significant effects of genotype (Col-0, L5, L7) (X^2 = 46.0, df = 4, p < 0.001), treatment (DMSO, E2) (X^2 = 31.3, df = 3, p < 0.001), and the interaction between genotype and treatment (X^2 = 15.7, df = 2, p < 0.001) were detected. Post-hoc pairwise comparisons revealed that both lines harbouring the *XVE:ROS1-YFP* construct significantly supressed the progression of *Pst-Lux* when treated with E2 (multiple comparisons using estimated marginal means and a Tukey p value adjustment < 0.05) (Figure 3.5c). All DMSO-treated and E2-treated Col-0 seedlings displayed similar progression of *Pst-Lux* colonisation (p > 0.05) (Figures 3.5c, 3.5d). Therefore, E2 does not influence the progression of *Pst-Lux* colonisation in Col-0. By contrast, transient induction of *ROS1-YFP* by E2 in L5 and L7 effectively induces resistance against *Pst-Lux*.

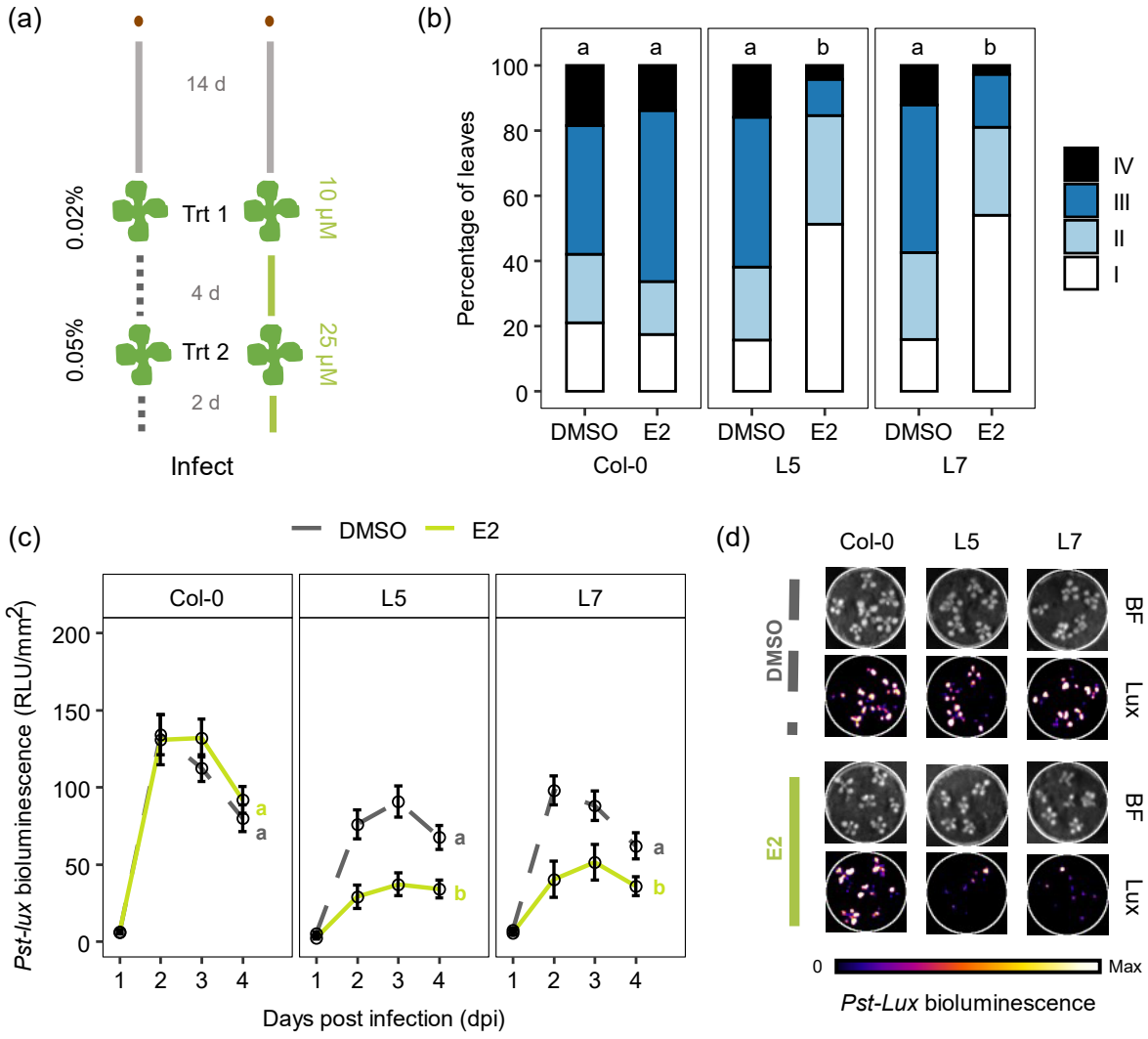


Figure 3.5 Impacts of estradiol treatment on resistance to *Hpa* and *Pseudomonas syringae* pv. *tomato* DC3000 *luxCDABE* (*Pst-Lux*) in wild-type Col-0 and two *XVE:ROS1-YFP* lines (L5 and L7).
(a) Experimental setup to analyse the impacts of E2-induced ROS1 on resistance to (hemi)biotrophic pathogens *Hpa* and *Pst-Lux*. Col-0, L5, and L7 plants were treated twice with E2 or DMSO at 14 days after sowing (DAS) (10 uM E2 or 0.02% DMSO) and 18 DAS (25 uM E2 or 0.05% DMSO). 2 days after treatment 2, seedlings were inoculated with *Hpa* or *Pst-Lux*. (b) At 5 dpi with *Hpa*, individual leaves (n = 72-80) were microscopically assigned to 4 different *Hpa* colonisation classes (details in Figure 3.4 legend and Section 2.5). The relative frequency of leaves assigned to each *Hpa* colonisation class is shown. Treatments with the same letter are not significantly different (Fisher's exact test; FDR-correction; $q > 0.05$). (c) Colonisation by *Pst-Lux* was quantified 1-4 dpi using bioluminescence intensity (section 2.5), reported as plant area-normalised relative light units (RLU/mm^2). Points with error bars represent mean *Pst-Lux* bioluminescence values (n = 31-41) ± standard error of the mean. Treatments with the same letter are not significantly different in their progression of disease between 2-4 dpi (linear mixed effects model; estimated marginal means post-hoc; Tukey $p_{adj} < 0.05$). (d) Visualisation of *Pst-Lux* bioluminescence in inoculated seedlings of Col-0, L5, and L7. Shown are representative pots for each genotype-treatment combination. Bright field (BF) images are used to define plant regions. Long-exposure images taken in the dark (4 mins) are used to quantify *Pst-Lux* bioluminescence (section 2.5). Inoculation and quantification of *Pst-Lux* and *Hpa* are detailed in Section 2.5.

3.4.7 Transient Induction of ROS1 in the Parental Generation Does not Enable Heritable Induced Resistance (h-IR) to *Hpa* in the Progeny

The results presented above confirm that ROS1 is required for the establishment, maintenance, and/or activation of h-IR against *Hpa* following parental *Pst* infection (Figure 3.1) (López Sánchez et al., 2016). Furthermore, ectopic induction of ROS1 conferes induced resistance against (hemi-)biotrophic pathogens (Figure 3.5), providing further evidence that ROS1 is a positive regulator of SA-dependent defences (Yu et al., 2013; López Sánchez et al., 2016; Halter et al., 2021). Therefore, to determine whether ROS1 activity alone can drive the establishment of long-term immune memory preceding h-IR, Col-0, L5, and L7 seedlings were treated twice with DMSO or E2, as previously described (Figure 3.5), and their 3-week-old progeny were tested for resistance against *Hpa* (Figure 3.6a). Parental E2 treatment (pE2) was not associated with heightened immunity compared to parental DMSO (pDMSO) treatment in any of the tested genotypes (Fisher's exact test, $q > 0.05$). Therefore, while ectopic induction of ROS1 can provide short-term immune memory (48 hours) against *Pst* and *Hpa*, this memory is not stable across a generation, and as a result, no h-IR is observed in the progeny. Having observed that this memory is initially established, but subsequently lost, this *XVE:ROS1-YFP* system acts as an excellent tool to study the establishment, maintenance, and erasure of immune memory. Thus, in the following Chapters of this thesis, I first characterise molecular changes associated with the establishment of this immune memory (Chapter 4), and subsequently investigate the within-generation durability (i.e., the erasure) of this memory (Chapter 5).

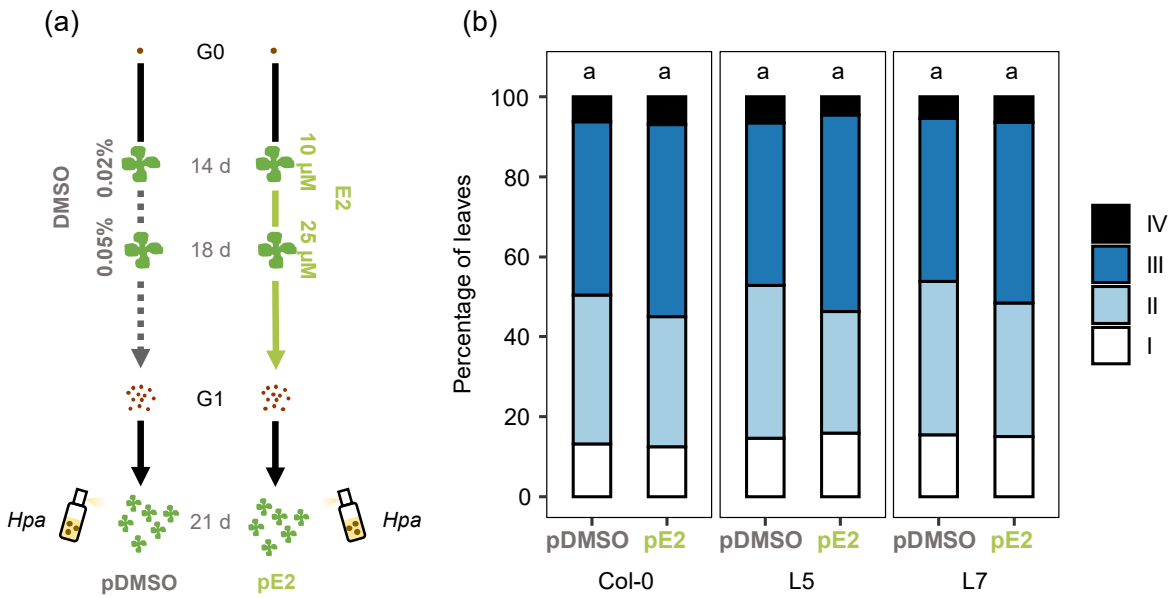
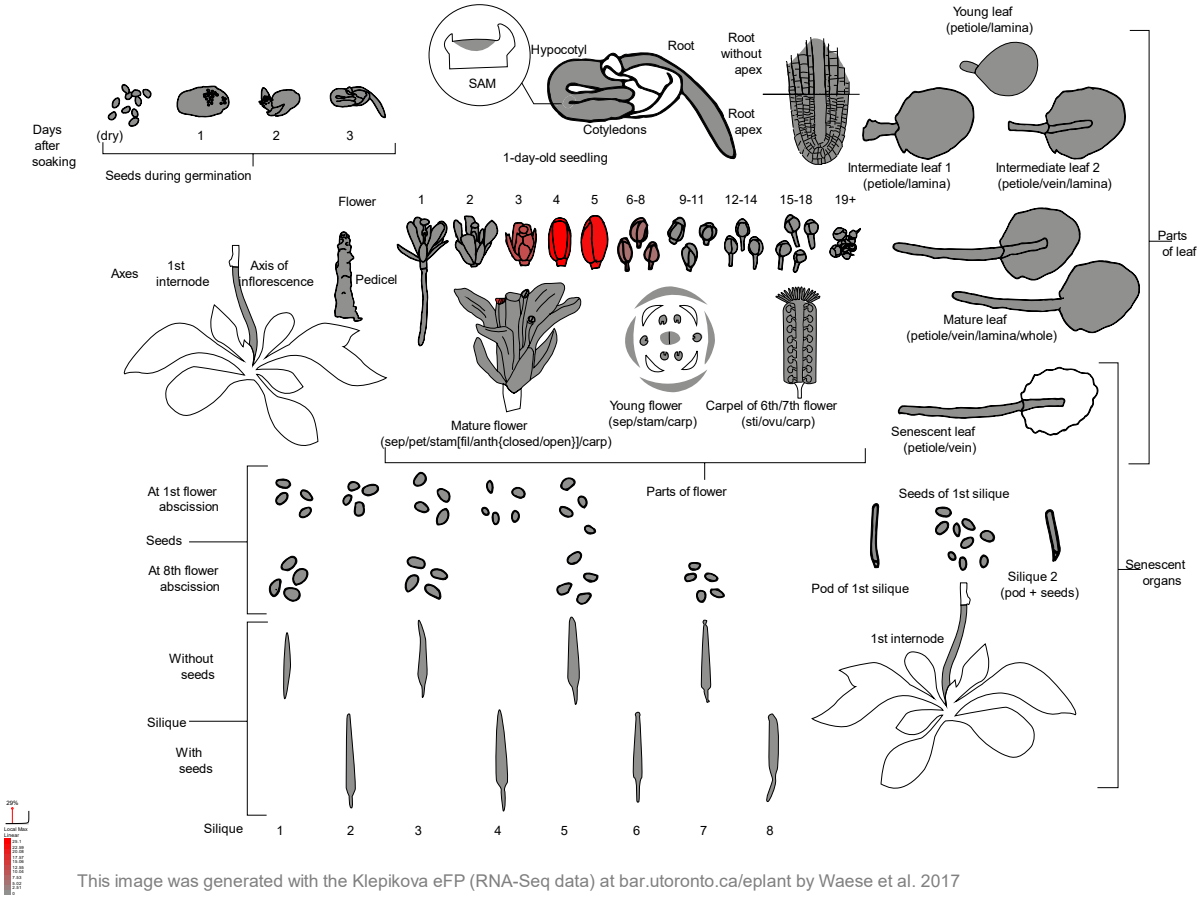


Figure 3.6 Resistance of progeny to *Hyaloperonospora arabidopsidis* from DMSO- or E2-treated Col-0, L5, and L7 seedlings.
(a) Col-0, L5, and L7 plants (generation 0; G0) were treated twice with E2 or DMSO at 14 and 18 DAS, as described in Figure 3.5a. Seeds (G1) were collected from three individual plants per genotype for each treatment and infected with *Hpa* 21 DAS. (b) At 5 dpi with *Hpa*, seedlings were stained with trypan blue. Individual leaves from 8 plants per parental line (3 parental plants per genotype per treatment) were microscopically assigned to four different *Hpa* colonisation classes as detailed in the legend of Figure 3.1b and in Section 2.5. Shown are the relative frequencies of leaves assigned to each *Hpa* colonisation class, pooled by parental treatment and genotype (n = 123–132). Treatments with the same letter are not significantly different (Fisher's exact test; FDR-correction; $q > 0.05$).

3.5 Supplementary Figures



This image was generated with the Klepikova eFP (RNA-Seq data) at bar.utoronto.ca/eplant by Waese et al. 2017

Supplementary Figure 3.1 Snapshot of the Klepikova Arabidopsis Atlas for the Arabidopsis gene locus AT4G10603.
Screenshot from from ePlant (Waese et al., 2017) (http://bar.utoronto.ca/eplant; accessed 12/10/2024). Figure shows RNA sequencing derived expression values of the gene in different tissues, with grey indicating no expression and red indicating high expression. This Figure demonstrates that the gene is exclusively expressed in flower buds and the onset of flower opening.

Chapter 4. Epigenomic and Transcriptomic Profiling of Arabidopsis at the Onset of ROS1-Driven Immune Memory

Authors:

Hannan Parker, A[1]; Rolfe S. A. [1]; Smith, L. M.[1] and Ton, J[1].

Affiliations:

[1] Plants, Photosynthesis and Soil, School of Biosciences, The University of Sheffield, Western Bank, Sheffield, S10 2TN, UK

Author Contributions:

AHP (the candidate), SAR, LMS, and JT proposed the original ideas for the research. AHP conducted experiments, gathered data, and performed all data analysis. AHP created all Figures and wrote all the text in the Chapter. JT reviewed and provided comments on the Chapter. Funding for the project came from studentships/grants awarded to AHP and JT.

4.1 Abstract

Plants can retain memory from past stress events, which enables a more efficient defence response to recurrent stress exposures. Genome-wide reductions in DNA methylation occur in response to biotic stress and can persist long after recovery from the stress. Mutations in genes controlling DNA methylation have been shown to disrupt this ability, rendering plants 'forgetful' and preventing them to maintain induced resistance (IR). In the previous Chapter, I demonstrated that transient induction of the DNA demethylase ROS1 in estradiol-treated *XVE:ROS1-YFP* plants results in IR against (hemi-)biotrophic pathogens. In this Chapter, I have combined whole-genome bisulfite sequencing, whole-transcriptome sequencing, and small RNA (sRNA) sequencing to characterise the epigenetic and transcriptomic changes during the onset of ROS1-dependent immune memory. This integrated analysis firstly revealed that transient induction of ROS1 reduces DNA methylation and 23-24 nucleotide (nt) sRNAs at the boundaries between transposable elements (TEs) and protein-coding genes in the chromosome arms. By contrast, intergenic DNA demethylation in the (peri)centromeric regions was less pronounced, which correlated with a dramatic increase in sRNA accumulation. Functional analysis of all protein-coding genes induced by ROS1 revealed an enrichment of gene ontology terms related to salicylic acid (SA)-dependent defence. Moreover, integrating all sequencing data to identify protein-coding genes that are antagonistically *cis*-regulated by ROS1-induced DNA-demethylation and RNA-directed DNA methylation (RdDM) identified key regulators of salicylic acid (SA)-dependent defences and DNA damage responses. This included a locus in the promoter of *NPR1* – a master regulator of SA-dependent defences and systemic immune memory. Collectively, this Chapter provides new insights into the epigenetic and transcriptomic changes occurring during the establishment of epigenetic immune memory. The results not only reveal novel functions of ROS1 and RdDM in epigenomic homeostasis but also identify potential targets for the epigenetic engineering of stress memory in plants.

4.2 Introduction

After recovering from an infection or infestation by pathogens or pests, plants possess the ability to remember the stress, which enables them to respond more effectively to a similar stress later in life (Wilkinson et al., 2019; Hannan Parker et al., 2022). A common response of plants to biotic stress is genome-wide DNA de-methylation. This epigenetic change influences the expression of genes and transposable elements (TEs) and can be maintained over cell divisions (Wilkinson et al., 2019, 2023; Hannan Parker et al., 2022; Catoni et al., 2022). Several studies have shown that genetic mutations disrupting DNA (de)methylation in Arabidopsis de-regulate the ability of plants to remember stress and/or mount an induced resistance (IR) response (Luna et al., 2014; López Sánchez et al., 2016; Wilkinson et al., 2023). Notably, mutants impaired in the DNA demethylase ROS1 fail to express heritable induced resistance (h-IR) against biotrophic pathogens (López Sánchez et al., 2016), whereas plants affected in non-CG DNA methylation mimic the priming of salicylic acid-dependent defences that drives h-IR upon parental disease stress and long-term IR by chemical priming agent BABA (Luna et al., 2012, 2014). These results have led to the hypothesis that transient disease stress induces an 'epigenetic shock' via ROS1-dependent DNA demethylation, which results in the establishment and long-term maintenance of immune memory (López Sánchez et al., 2016; Wilkinson et al., 2019, 2023; Hannan Parker et al., 2022).

Despite the evidence from genetic mutations affecting DNA methylation homeostasis, it remains challenging to separate cause and effect of DNA demethylation during the establishment of immune memory. For instance, mutations affecting DNA (de)methylation have a permanent impact on the epigenetic machinery in the plant, which contrasts the transient changes occurring during stress-induced immune memory. Consequently, it remains unclear whether transient changes in the activity of DNA (de)methylation machinery are merely a byproduct of the immune response, or whether they serve as a critical signal in the establishment of immune memory. To address this question, the study presented in this Chapter has taken advantage of the transgenic *XVE:ROS1-YFP* lines described in Chapter 3 (Figure 3.4). These plants develop IR against the (hemi-)biotrophic pathogens *Hyaloperonospora arabidopsidis* (*Hpa*) and *Pseudomonas syringae* pv. *tomato* DC3000 (*Pst)* upon two successive applications of estradiol (E2; Figure 3.5), which supports the notion that the epigenetic shock by transient ROS1-induced DNA demethylation is causal to the establishment of epigenetic immune memory against biotrophic pathogens.

To characterise the epigenetic and transcriptomic changes associated with this ROS1-dependent immune memory, this Chapter describes the DNA methylome and transcriptome of two independent *XVE:ROS1-YFP* lines and wild-type Col-0 plants as controls after two

resistance-inducing treatments with E2. These results not only confirm that the recombinant ROS1-YFP protein functions as an active DNA demethylase enzyme but also reveal gene expression patterns consistent with enhanced SA-dependent defences that explain the IR responses observed in Chapter 3 (Figure 3.5). Furthermore, the patterns of global DNA demethylation and small RNA distribution induced by ROS1 highlight a notable difference in the response between chromosome arms and pericentromeric regions, suggesting a difference in the regulation of DNA (de)methylation in these genomic regions.

4.3 Materials and Methods

4.3.1 Plant Materials, Growth Conditions, Chemical Treatments, and Material Harvest

Three Arabidopsis genotypes were used in this Chapter: wild-type Col-0 and two *XVE:ROS1-YFP* lines, line 5 (L5) and line 7 (L7). For details on the generation of *XVE:ROS1-YFP* plants, see Section 2.4 and for characterisation of L5 and L7 see Figures 3.4, and 3.5. Plants were cultivated in short-day conditions using 70 mL pots containing a 2:1 soil:sand mixture following a 4 day stratification period, as detailed in Section 2.1. Ten seeds were sown per pot, and 6 pots per genotype were grown within a single tray. Plants were watered from the bottom 2-3 times per week.

Aerial parts of seedlings were sprayed twice with E2 or DMSO. The first spray was performed at 14 days after sowing (DAS) with 10 µM E2 or 0.02% (v/v) DMSO; the second spray was performed at 18 DAS with 25 µM E2 or 0.05% DMSO (v/v). Control treatments were based on equivalent DMSO sprays without E2. All solutions were prepared in dH_2O and were supplemented with 0.015% (v/v) silwet L-77 (LEHLE SEEDS, VIS-30) (silwet). Each tray of 18 pots was sprayed with a total volume of 20 mL E2 solution or DMSO solution. Further details about the preparation and storage of E2 stocks are detailed in Section 2.3.

Green leaf area (GLA) (Supplementary Figure 4.2) was calculated from the previous *Pst-Lux* assay in Chapter 3 (Figure 3.5), using the plant area regions of interest (ROIs) from bright field images as a measure of GLA. The effect of genotype and treatment, and the interaction between the two, on GLA was tested using a two-way ANOVA. GLA was log transformed to meet the assumptions of normality and homogeneity of variance. Tukey post-hoc tests were performed using the R package agricolae v1.3-7 (function: HSD.test).

For each biological replicate (n=3), two days after the final E2 or DMSO treatment (20 DAS), aerial tissues from approximately 60 seedlings per genotype per treatment were collected from a single tray, snap frozen in liquid nitrogen, and stored at -80°C. The material was then ground to fine powder in a pre-cooled pestle and mortar and split between several ~100 mg aliquots in pre-cooled 2 mL Eppendorf tubes. These aliquots were placed back in liquid nitrogen and stored at -80°C. Individual aliquots, derived from the same pool of ground up tissue, were used for DNA and RNA extractions.

4.3.2 Whole Genome Bisulfite Sequencing (WGBS) Analysis

4.3.2.1 DNA Extraction and WGBS

Genomic DNA was extracted from biologically replicated (n = 3) E2- or DMSO-treated Col-0, L5, and L7 seedlings (18 samples in total) using the DNeasy Plant Mini Kit (Qiagen 69104). DNA quality and quantity were assessed using a Nanodrop 8000 spectrophotometer and a Qubit 3.0 Fluorometer. In cases where total DNA yield was low, a second extraction was performed and the eluate from the first extraction was used to elute the second extraction, thereby increasing both concentration and yield. The bisulfite conversion, library preparation, and 100 bp paired-end sequencing was performed by BGI Genomics using their DNA nanoball (DNB) sequencing (DNBSEQ) platform with standard protocols.

Raw sequencing data were filtered by BGI Genomics using SOAPnuke (Y. Chen et al., 2018) with the following parameters: '-n 0.001', '-l 20', '-q 0.4, '--adaMR 0.25', '--minReadLen 100'. This filtering process was conducted to remove adapter sequences and low-quality reads. Subsequently, the absence of adapters and the overall quality of the provided filtered sequencing data were independently verified in-house using FastQC v0.12.1 and MultiQC v1.21 (Ewels et al., 2016). Reads were then aligned to the TAIR10 Arabidopsis genome, using Bismark v0.24.1 (function: bismark; options: default parameters) (Krueger and Andrews, 2011), which utilises Bowtie2 v2.5.1 (Langmead and Salzberg, 2012). Using these default parameters, any reads which align ambiguously to the genome were excluded from downstream analyses of DNA methylation. While this carries the risk of reducing coverage at highly repetitive sequences (eg., TEs), it ensures that all methylation calls are of high confidence. PCR duplicate reads were removed (function: deduplicate_bismark; options: default parameters) and Bismark was used to generate and summarise binary scores for methylated cytosines (1) and unmethylated cytosines (0) in all contexts for each position (function: bismark_methylation_extractor; options: --comprehensive --bedGraph --CX --cytosine_report --ignore_r2 4). The unmethylated plastid genome was used to calculate non-conversion rates, and all samples passed the accepted threshold of <2%, with an average non-conversion rate of 0.42 % (Stuart et al., 2018) (Supplementary Data 4.1).

4.3.2.2 Principal Component Analysis (PCA)

PCA was performed using positional cytosine methylation levels (%) in any context (all cytosines; allC). Cytosines with <5 mapped reads were dropped from the analysis. Furthermore, to focus the computational analysis on the most variable regions, cytosines with standard deviations that were less than the median of the standard deviation across all cytosines were removed from the analysis. PCA was conducted in base R v4.2.1 (function:

prcomp; options: scale = FALSE, center = TRUE) and plots were generated using the R package ggplot2 v3.5.0.

4.3.2.3 Chromosome Metaplots

Plots are based on average levels of cytosine DNA methylation in 100 kb bins of the TAIR10 genome. Bins were created using bedtools v2.31.0 (function: makewindows; options: -w 100000 -s 100000) (Quinlan and Hall, 2010). Only cytosines with ≥ 5 mapped reads were included in the analysis, after which bins with at least 60% of the expected number of sites for a given context based on averages of the nuclear TAIR10 genome, rounded to the nearest thousand, were retained. This resulted in the following selection: bins must contain ≥ 3000 sites for CG and CHG contexts, ≥ 16000 sites for CHH context, and ≥ 22000 sites for allC, all with sufficient coverage (≥ 5 reads). To define chromosome arms and (peri)centromeric regions, the 100 kb windows with TE densities > protein coding gene densities were identified. For each nuclear chromosome, the start of the pericentromere was defined as the midpoint of the first bin in a string of at least 5 neighbouring bins for which TE density > gene density. The end of the pericentromere was defined as the midpoint of the last bin in a string of at least 5 neighbouring bins where TE density > gene density (Chr1: 12450000-17650000, Chr2: 1250000-7050000, Chr3: 10450000-16250000, Chr4: 1650000-6350000, Chr5: 9750000-14650000). The pericentromeric midpoint for each chromosome was defined as the midpoint between the start and end of the associated pericentromeric coordinates. Metaplots of chromosomes plot the coordinates of the 100 kb window relative to the pericentromeric midpoint along the x-axis and average cytosine methylation of windows on the y axis. The pericentromeric region in the metaplots was defined as the maximum absolute relative distance for any pericentromeric window (as defined above) for both the start and end points.

4.3.2.4 Gene, Transposable Element (TE), and miRNA Metaplots

Plots are based on 20 equally sized bins across protein-coding (transcriptional start site to transcriptional termination site; TSS to TTS), TE, or miRNA regions, as defined by the Araport11 (Cheng et al., 2017) and TAIR10 TE annotation (https://www.arabidopsis.org; accessed 10/02/2024) files. Similarly, the 1 kb region upstream and downstream of the feature were divided into 20 equally sized bins, equating to 50 bp per bin. Average allC methylation was calculated for each bin using bedtools (function: map; options -o mean,count) (Quinlan and Hall, 2010). Only cytosines with ≥ 5 mapped reads were included in the analysis.

4.3.2.5 Analysis of Differentially Methylated Regions (DMRs)

The R package DSS v2.46.0 (H. Feng et al., 2014; Park and Wu, 2016) was used to calculate DMRs. As DSS accounts for read coverage, no prior data filtering was performed.

Differentially methylated cytosines were first called using DSS (function: DMLtest; options: equal.disp = FALSE, smoothing = FALSE) which were then used to call DMRs with DSS (function: callDMR; options: minlen = 25, dis.merge = 50, minCG = 3, p.threshold = 0.05, delta = 0.1, pct.sig = 0.5). This pipeline was performed on cytosines in allC, CG, CHG, and CHH sequence contexts to generate context specific DMRs. Annotation of DMRs were based on Araport11 (Cheng et al., 2017) and TAIR10 transposable element annotation (https://www.arabidopsis.org; accessed 10/02/2024) files using the R packages genomation v1.34.0 (Akalin et al., 2015) and GenomicRanges v1.54.1 (Lawrence et al., 2013). In cases where a DMR overlapped with multiple features, a single annotation was assigned in the following order of preference: exon > intron > promoter > downstream > non-coding RNA > intergenic. Promoters were defined as 1 kb regions upstream of the transcriptional start site (TSS) of protein-coding genes. Downstream gene sequences were defined as 1 kb regions from the transcriptional termination site (TTS) of protein-coding genes. Non-coding RNA included any defined transcribed locus in Araport11 that is not a protein-coding gene. Intergenic regions were defined as regions that do not overlap any of these above defined features. DMRs were annotated as overlapping/not overlapping with a TE using the TAIR10 TE annotation file.

4.3.2.6 DMR Enrichment Plots

ROS1-hypomethylated DMRs (hypo-DMRs) were merged across contexts in E2-treated L5 and L7 plants to form set of non-duplicate hypo-DMRs. These merged DMRs were then cross-referenced between both lines and overlapping DMRs were extended such that the start coordinate was the minimum value between L5 and L7, and the end coordinate was the maximum value between L5 and L7. These so-called ‘consistent DMRs’ were then annotated using the Araport11 and TAIR10 TE annotations, as described above. Consistent DMRs were generated using the functions ‘reduce’ and ‘findOverlaps’ with default parameters from the R package GenomicRanges v1.54.1.

Based on the consistent DMRs, four random control sets of DMRs of equal size and number were generated by randomly selecting a chromosome start coordinate in the TAIR10 genome. To account for the varying size of chromosomes, proportional probability weights were assigned to each randomly selected chromosome. Annotation was performed exactly as done for the set of consistent DMRs. Counts of real DMRs mapping to a given annotation type was tested for enrichment using a two-sided Fisher's exact test with false discovery rate (FDR) correction, taking the counts from a single random DMR set as the expected count value. Four rounds of tests were performed (using a different random DMR set each time) and averages were taken for the test statistics (fold change, *q*-value). Standard errors of the mean were

calculated for the fold change and *q*-value and plotted as error bars. Random DMRs were generated using a custom R script which uses functions from base R v4.3.2 for selecting chromosomes (function: sample; options: prob = chr_size_proportion) and coordinates (function: runif; options: n = 1, min = 1, max = max_chr_size). Fisher's exact test was performed using the R package rstatix v0.7.2 (function: pairwise_fisher_test; options: alternative = "two.sided", p.adjust.method = "fdr") and plots were generated using the R package ggplot2 v3.5.0.

4.3.3 Ribosome-Depleted Whole Transcriptome RNA Sequencing Analysis

4.3.3.1 Total RNA Extraction and Whole Transcriptome Sequencing

Total RNA was extracted from biologically replicated (n = 3) E2- or DMSO-treated Col-0, L5, and L7 seedlings (18 samples in total) using ~100 mg aliquots of ground material and 1 mL of TRIzol Reagent (Thermo Fisher Scientific 15596026), as per manufacturer guidelines. Following chloroform-phenol separation, precipitation of RNA from the aqueous phase was performed using 0.8 M sodium citrate and isopropanol at a ratio of 2:1:1, respectively. Following centrifugation for 15 minutes at 12,000 × *g* at 4°C, the pellet was washed twice with 70% ethanol and resuspended in 50 µL nuclease-free water. To aid resuspension of the pellet, tubes were vortexed for 10-15 seconds, incubated in a heat block for 10 minutes at 60°C, and then kept on ice. The quality of the extracted RNA was first assessed using a Nanodrop 8000 spectrophotometer. In cases where samples had low 260/280 ratios (< 1.7), RNA was reprecipitated overnight at -20°C by adding 0.1x volume of 3 M sodium acetate (pH 5.2-5.5 prepared in nuclease-free H_2O) and 2.5x volume pre-cooled 100% ethanol. Centrifugation, wash steps, and resuspension were then repeated as before. Final RNA extracts were checked for quality and quantity using a Nanodrop 8000 spectrophotometer and an Agilent 2200 TapeStation (Analysis type: Eukaryotic Total RNA QC). All samples had an RNA integrity number (RIN) of at least 6.7. A ribosomal RNA-depleted stranded library was constructed using BGI Genomics' standard protocol, and 150 bp paired-end reads were sequenced on their DNBSEQ platform. Raw sequencing data were filtered by BGI Genomics using SOAPnuke (Chen et al. 2018) with the following parameters: '-n 0.001', '-l 20', '-q 0.4, '--adaMR 0.25', '--polyX 50', '--minReadLen 150'. This filtering process was conducted to remove adapter sequences and low-quality reads. Subsequently, the absence of adapters and the overall quality of the provided filtered sequencing data were independently verified in-house using FastQC v0.12.1 and MultiQC v1.21 (Ewels et al., 2016). Read statistics are provided in Supplementary Data 4.2.

4.3.3.2 Read Alignment and Counting

An index of the TAIR10 genome was generated using STAR v2.7.9a (function: STAR; options: --runMode genomeGenerate --genomeSAindexNbases 12) (Dobin et al., 2013) using a custom annotation file that merged Araport11 annotations (Cheng et al., 2017) and a gene-like annotation file of transposable elements (TEs) (Panda and Slotkin, 2020). TE loci from Araport11 were excluded from the merged file if they were also annotated by Panda and Slotkin (2020) to prevent multiple annotations of the same locus. Reads were aligned and numbers of reads mapping to each annotated gene were counted using STAR v2.7.9a (function: STAR; options: --quantMode GeneCounts --alignIntronMin 60 --alignIntronMax 6000 --outMultimapperOrder Random). Since this is a reverse-stranded library, read counts from the fourth column of the output file were used, which represent the counts where the reverse strand aligned with the annotated locus.

4.3.3.3 Library Normalisation

Count tables were loaded into R v4.3.2 and filtered to retain only features where ≥6 samples had ≥10 read counts at the given locus. Counts were initially normalised using the default median-of-ratios method implemented by the R package DESeq2 v1.42.1 (function: DESeq) (Anders and Huber, 2010; Love et al., 2014). This normalisation method assumes that most genes are not differentially expressed, leading to a normalisation factor, or size factor, that typically centres around 1. However, as discussed by Love et al. (2014), there exists a potential "pathological case" where size factors are confounded with experimental conditions. In such cases, features that are truly unchanged in expression between conditions could be incorrectly identified as differentially expressed due to artifacts introduced during library normalisation. To investigate this, size factors were extracted in R using DEseq2 v1.42.1 (function: sizeFactors) and plotted with the R package ggplot2 v3.5.0. This revealed noticeably larger size factors of E2-treated *XVE:ROS1-YFP* lines compared to all other sample groups (Figure 4.1a). To address issues where size factor is confounded with treatment, DESeq2 implemented the ability for users to generate size factors from a set of features that should not be changing with respect to the treatment conditions (Love et al., 2014). This is also discussed in the DESeq2 user guide on Bioconductor (Love et al. 2023, guide version 1.42.0). These features can be either external spike-in controls or internal reference genes, which, in some cases, facilitate more representative library normalisation (Bullard et al. 2010; Robinson and Oshlack 2010; Jiang et al. 2011; Lovén et al. 2012; Risso et al. 2014; Berghoff et al. 2017; Laosuntisuk et al., 2024). Therefore, size factors were regenerated using 15 of the most stably expressed genes in Arabidopsis (Czechowski et al., 2005; Hong et al., 2010) using the R package DESeq2 (function: estimateSizeFactors; options: controlGenes = `stableGenes_Czec_and_Hong`). These 15 genes were selected based on the following

criteria: (i) the genes were among the 22 most stably expressed genes identified from a dataset covering 323 experimental conditions (Czechowski et al. 2005); and (ii) the genes were consistently experimentally verified as stably expressed in 3 independent biological replicates under at least 2 different experimental conditions out of the 6 tested by Hong et al. (2010). The experimental conditions tested by Hong et al. (2010) included seedling age, photoperiod, ambient temperature, diurnal expression, distinct tissue types, and defined developmental stages. The 15 genes used for normalisation are listed in Supplementary Data 4.3. This normalisation technique reduced the size factors of E2-treated *XVE:ROS1-YFP* lines, and visually, all samples better cantered around 1 (Figure 4.1a). Plotting the mean relative expression of these 15 stably expressed genes, based on the change in normalised counts relative to DMSO-treated Col-0 seedlings, revealed statistically significant differences between treatment groups under the default DESeq2 normalisation method (H = 30.6, df = 5, p < 0.001) (Figure 4.1b). Specifically, E2-treated *XVE:ROS1-YFP* lines exhibited a significant reduction in the relative expression of these genes compared to all other treatment groups (Conover-Iman test, p < 0.05) (Figure 4.1b). This suggests that the size factors derived from the default normalisation technique may be inflated (Figure 4.1a), leading to overcorrection of gene expression data and resulting in artificially reduced normalised count values. As comparison, no significant differences in relative expression were detected for these genes following the reference-guided normalisation technique (H = 1.1, df = 5, p > 0.05). Therefore, the control gene normalisation method was used instead. Kruskal-Wallis tests and Conover-Iman tests were performed using base R v4.3.2 (function: kruskal.test) and the R package DescTools v0.99.54 (function: ConoverTest). All plots were generated using the R package ggplot2 v3.5.0.

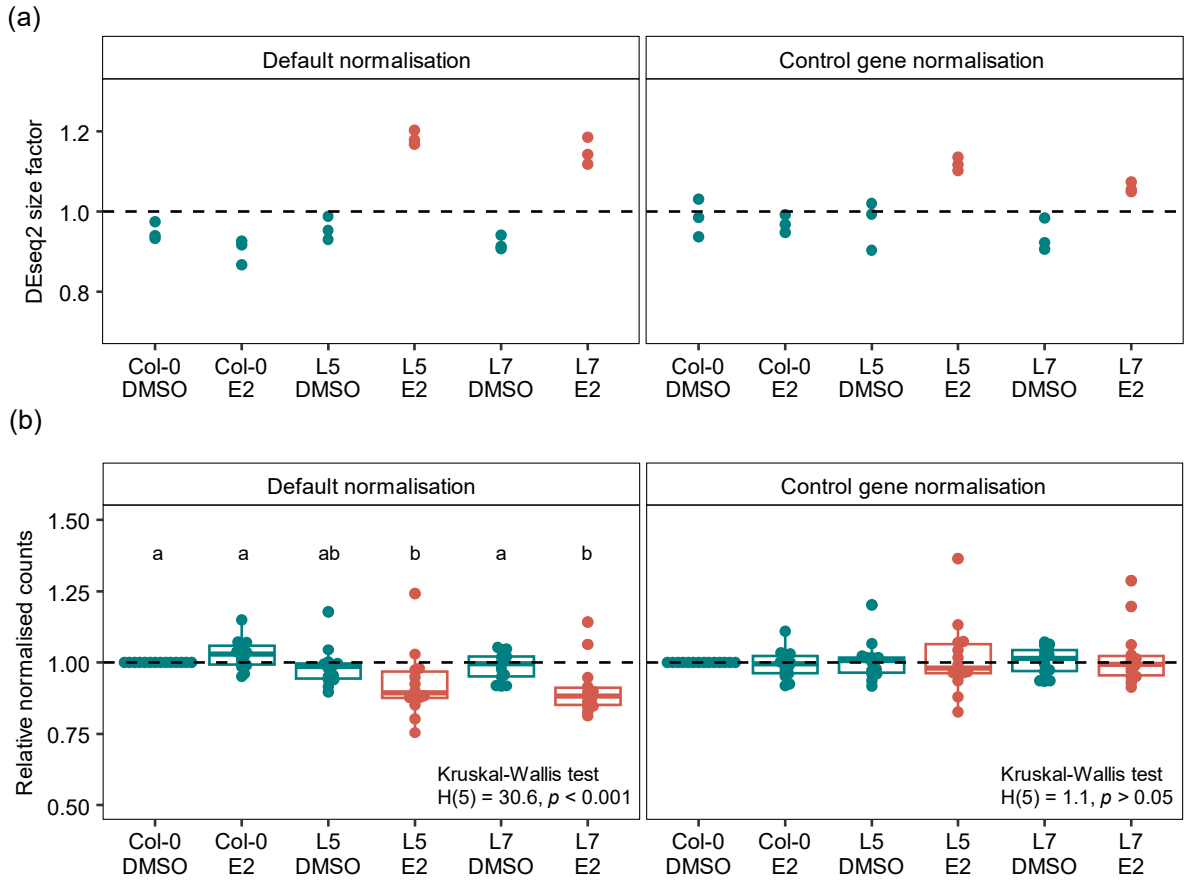


Figure 4.1 Comparison of transcriptome library normalisation techniques in the R package DESeq2. (a) Library normalisation weightings ("size factors") assigned to each sample and treatment under the default median-of-ratios method ("default normalisation") or after defining a set of 15 stably expressed genes in Arabidopsis ("control gene normalisation") in DESeq2 (see text for details). E2-treated *XVE:ROS1-YFP* lines (L5 and L7) are highlighted in red. (b) Relative normalised read counts of 15 stably expressed genes in Arabidopsis for each treatment group. Each point represents the mean normalised read count of a single gene, relative to the mean of DMSO-treated Col-0 samples. Conover-Iman post-hoc tests performed only upon detecting a significant effect of treatment group on relative expression (Kruskal Wallis test, $p < 0.05$). Treatment groups with the same letter do not have significantly different relative expression values (Conover-Iman, $q < 0.05$).

4.3.3.4 Principal Component Analysis (PCA)

Normalised read counts were transformed with a regularised logarithm transformation using the R package DESeq2 v1.42.1 (function: rlog; options: blind=FALSE) (Love et al., 2014). PCA of the transformed data was conducted using base R v4.3.2 (function: prcomp; options: default parameters) and plotted with the R package ggplot2 v3.5.0.

4.3.3.5 Identification and Clustering of Differentially Expressed Features

Differentially expressed features (*i.e.*, coding, and noncoding loci) associated with E2 treatment were identified by comparing DMSO-treated and E2-treated seedlings in Col-0, L5,

and L7 using the R package DESeq2 v1.42.1 (model: ~ Genotype + Treatment + Genotype:Treatment) (all DEGs are listed in Supplementary Data 4.4). For simplicity, differentially expressed features were broadly defined as differentially expressed genes (DEGs). A locus was considered differentially expressed if it exhibited a $\log_2$ fold-change greater than 0 and an adjusted *p*-value (p_{adj}) less than 0.05 in the E2 vs DMSO comparison within a given genotype. The rlog-transformed expression values of all identified DEGs across all samples were visualised in a gene-wise and sample-wise clustered heatmap using the R package pheatmap v1.0.12 (function: pheatmap; options: cluster_rows=TRUE, cluster_cols=TRUE, clustering_distance_cols = "correlation", clustering_distance_rows = "correlation", clustering_method = "ward.D2"). Two distinct gene clusters emerged: one containing DEGs upregulated in E2-treated *XVE:ROS1-YFP* lines (the 'up' cluster), and another containing DEGs downregulated in these lines (the 'down' cluster). The DEGs associated with these clusters were isolated using the “cutree” function from base R v4.3.2.

4.3.3.6 Statistical Enrichment of Gene Ontology (GO) Terms

Enrichment of biological process GO terms was performed using R packages clusterProfiler v4.10.1 (Wu et al. 2021; Yu et al. 2012) and org.At.tair.db v3.18.0. GO term enrichment of protein-coding genes associated with the ‘up’ and ‘down’ cluster were identified using clusterProfiler 4.10.1 (function: enrichGO; options: pvalueCutoff = 0.05, qvalueCutoff = 0.05, OrgDb = "org.At.tair.db", keyType = "TAIR", ont="BP", minGSSize = 10, maxGSSize = 500, universe = all_proteinGenes_≥10counts_≥6samples). Biological process GO terms were classed as enriched if they had p_{adj}-values < 0.05 (Benjamini and Hochberg correction). GO semantic similarity analysis was performed using the R package GOSemSim v2.28.1 (Yu 2020) (function: godata; options: OrgDb = "org.At.tair.db", ont = "BP"). The object resulting from this analysis was used in conjunction with clusterProfiler v4.10.1 to create a simplified GO term list (function: simplify; options: cutoff=0.7, measure="Jiang", semData=SemSimAnalysis_object, by="p.adjust", select_fun=min). Bar charts of $-\log_{10}$-transformed p_{adj} values were plotted for selected GO terms using ggplot2 v3.5.0. Significantly enriched GO terms, both pre- and post-simplification, are provided in Supplementary Data 4.5.

4.3.3.7 Identification of Differential Transcript Usage

Protein-coding genes with differential transcript usage in E2-treated plants vs DMSO-treated plants within a given genotype were investigated using Salmon v1.10.0 (Patro et al., 2017) and the web-based 3D RNA-Seq analysis application (Guo et al., 2021) (https://3drnaseq.hutton.ac.uk/app_direct/3DRNAseq/). Firstly, a protein-coding transcript reference was built using Salmon (function: index; options: default parameters) and the TAIR10 cDNA reference transcriptome (EnsemblPlants release version 58). Estimates of

reads mapping to each transcript were aligned and quantified using Salmon v1.10.0 (function: quant; options: --validateMappings --seqBias --gcBias). Read counts were then uploaded onto the 3D RNA-Seq platform using default settings and the following optional parameters: only transcripts with ≥3 counts per million (CPM) in ≥3 samples were considered; the 'TMM' method was used to normalise read counts (Bullard et al., 2010); and the 'RUVr' method was selected to remove batch effects due to time of harvest (Risso et al., 2014) (see Figure 4.1). A protein-coding gene was considered differentially alternatively spliced (DAS) if it had an associated transcript with a $\log_2$ fold-change > 0 and a $p_{adj} < 0.05$ in the E2 vs DMSO comparison within a given genotype.

4.3.3.8 Analysis of *35S:TET1* Arabidopsis Lines

Raw RNA sequencing data from two *35S:TET1* lines and three corresponding Col-0 wild-type controls were obtained from Ji et al., (2018) under GEO accession GSE93024. Adapters and poor quality sequences were removed from reads using trimmomatic v0.39 (options: SLIDINGWINDOW:4:25, MINLEN:36, ILLUMINACLIP:adapters/TruSeq3-SE.fa:2:30:10) (Bolger et al., 2014). The absence of adapters and the overall quality of the trimmed reads were assessed using FastQC v0.12.1 and MultiQC v1.21 (Ewels et al., 2016). Read alignment, read counting, library normalisation, calling of DEGs (*35S:TET1* vs Col-0), and GO term analysis were performed exactly as described above. Sequencing read statistics, DEGs, and GO terms are presented in Supplementary Data 4.2, 4.4, and 4.5, respectively.

4.3.3.9 Transcription Factor Enrichment Analysis of Upregulated DEGs

Protein-coding DEGs associated with upregulation in E2-treated *XVE:ROS1-YFP* lines ('up' cluster) were scanned for motif enrichments using PlantRegMap (https://plantregmap.gao-lab.org/tf_enrichment.php; accessed 10/10/2024) with the following options: 'species: Arabidopsis thaliana', 'methods: all' and 'Threshold p-value ≤ 0.05'. This scans a list of genes for TFs that have previously identified in the literature from genome-wide scan data such as ChIP-sequencing data or inferred by combining TF binding motifs with regulatory elements data (Tian et al., 2020).

4.3.4 Small RNA Sequencing Analysis

4.3.4.1 Small RNA Library Preparation, Sequencing, and Read Filtering

Small RNA sequencing was performed for 3 biologically replicated samples of each DMSO and E2 treated Col-0, L5, and L7 using RNA from the same extract that was prepared for the whole transcriptome sequencing. Details on material harvesting and total RNA extraction are outlined above. A standard unique molecular identifier (UMI) small RNA (sRNA) library was prepared by BGI Genomics following a gel size selection of reads 18-30 nt in size.

Sequencing of 50 bp single-end reads was performed on the BGI Genomics DNBSEQ platform. Raw sequencing data were filtered by BGI Genomics using SOAPnuke (Y. Chen et al., 2018) with the following parameters: '-n 0.001', '-l 13', '-q 0.1', '--highA 1', '--minReadLen 15', '-- maxReadLen 44' '--ada_trim'. This filtering process was conducted to remove adapter sequences, low-quality reads, reads with 100% adenosine nucleotides, and any reads outside of the range of 15-44 nt. The absence of adapters and the overall quality of the provided filtered sequencing data were independently verified in-house using FastQC v0.12.1 and MultiQC v1.21 (Ewels et al., 2016). To match the original gel size selection, reads were further filtered for size using TrimGalore v0.6.10 (function: trim_galore; options: --quality 0 --length 18 --max_length 30 -a X), retaining only those between 18-30 nt. To remove rRNA reads from the data, reads were mapped to all known *Arabidopsis thaliana* rRNA sequences deposited on RNAcentral (accessed: 23/01/2024) (RNAcentral Consortium, 2021) using Bowtie v1.3.1 (function: bowtie; options: -q --un $1.unmapped_reads.fq --best -v 0). Reads that perfectly matched known rRNA sequences were removed and all unmapped reads were retained for further downstream analyses. All read statistics are provided in Supplementary Data 4.6.

4.3.4.2 Read Alignment, sRNA Cluster Identification, and Read Quantification

ShortStack v4.0.3 software was used to align reads to the TAIR10 genome, identify and characterise *de novo* sRNA clusters, and quantify the number of reads in each sample mapping to these clusters (Axtell, 2013; Shahid and Axtell, 2014; Johnson et al., 2016). ShortStack uses Bowtie v1.3.1 (Langmead et al., 2009) for genome indexing and read alignment. Known Arabidopsis miRNAs sequences were downloaded from miRBase (accessed 26/01/2024) (Kozomara et al., 2019) and ShortStack v4.0.3 was run with the following parameters: '--genomefile TAIR10.fa', '--known_miRNAs miRBase_ ath_known_miRNAs.fasta', '--mmap u', '--dicermin 20', '--dicermax 24', '--pad 200', '--mincov 2'. Using these parameters, ShortStack first identified all predicted and/or known miRNA clusters. It then defined additional clusters based on the dominant read size, only considering those where ≥80% of the reads were of a single size, ranging from 20 to 24 nucleotides (nt). Clusters that did not meet these criteria were classified as "not defined" (ND) clusters.

4.3.4.3 Count Normalisation, Principal component analysis (PCA), and Identification of Differentially Expressed sRNA Clusters

Count tables produced by ShortStack were loaded into R v4.3.2 and clusters which had ≥10 read counts in ≥6 samples were kept. Counts were normalised using the default median-of-ratios method implemented by the R package DESeq2 v1.42.1 (function: DESeq) (Anders and Huber 2010; Love, Huber, and Anders 2014). A control-gene normalisation is not possible as, unlike protein-coding genes, which have well-defined stably expressed genes

(Czechowski et al., 2005; Hong et al., 2010), equivalent stably expressed sRNA clusters have not been described in Arabidopsis. Normalised read counts were then transformed with a regularised logarithm (rlog) transformation using the R package DESeq2 v1.42.1 (function: rlog; options: blind=FALSE) (Love et al., 2014). PCA of the transformed data was conducted using base R v4.3.2 (function: prcomp; options: default parameters) and plotted with the R package ggplot2 v3.5.0.

Differentially expressed sRNA clusters (DECs) associated with E2 treatment were identified by comparing DMSO-treated and E2-treated seedlings in Col-0, L5, and L7 using the R package DESeq2 v1.42.1 (model: ~ Genotype + Treatment + Genotype:Treatment) (Supplementary Data 4.9). A cluster was considered differentially expressed if it exhibited a $\log_2$ fold-change greater than 0 and a $p_{adj} \leq 0.05$ in the E2 vs DMSO comparison within a given genotype. The rlog-transformed count values of all identified DECs across all samples were visualised in a cluster-wise and sample-wise clustered heatmap using the R package pheatmap v1.0.12 (function: pheatmap; options: cluster_rows=TRUE, cluster_cols=TRUE, clustering_distance_cols = "correlation", clustering_distance_rows = "correlation", clustering_method = "ward.D2"). Two distinct DEC clusters emerged: one containing DECs upregulated in E2-treated *XVE:ROS1-YFP* lines (the 'up' cluster), and another containing DECs downregulated in *XVE:ROS1-YFP* lines (the 'down' cluster). The DECs associated with these clusters were isolated using the “cutree” function from base R v4.3.2.

4.3.4.4 Identification of miRNA Targets

The predicted targets of miRNAs that were identified in the ‘up’ DEC cluster were identified using targetFinder v1.7 (Fahlgren et al., 2007; Srivastava et al., 2014) and the TAIR10 cDNA reference transcriptome (EnsemblPlants release version 58) with a prediction score cutoff of 3 (function: targetfinder.pl; options: -c 3, -d Arabidopsis_thaliana.TAIR10.cdna.all.fa). Biological process GO enrichments associated with the targeted protein-coding genes were identified as described above. All targets and GO terms are listed in Supplementary Data 4.7. Network plots to visualise the targets of each upregulated miRNA were created using the R packages igraph v2.0.3 and ggraph v2.2.1.

4.3.4.5 sRNA Read Length Distributions

For each sample, the total number of 18-30 nt, non-rRNA sRNA reads which were successfully aligned to the nuclear TAIR10 genome by ShortStack were calculated from “.bam” files using Samtools v1.3.1 (Danecek et al., 2021) (function: view). The percentage of reads for each size, relative to all mapped reads within a sample, were plotted using the R package using the R package ggplot2 v3.5.0.

4.3.4.6 DEC Annotation and Chromosome Plots

As above, annotation of DECs was based on Araport11 and the TAIR10 TE annotation file using the R packages genomation v1.34.0 and GenomicRanges v1.54.1. In cases where a DEC overlapped with multiple features, a single annotation was assigned in the following order of preference: miRNA > gene body (exon or intron) > promoter > tRNA > snRNA > snoRNA > lncRNA > other ncRNA > downstream > intergenic. Promoters were defined as 1 kb regions upstream of the transcriptional start site of protein-coding genes and downstream regions were defined as 1 kb regions downstream of the transcriptional termination site. Other ncRNA included any defined transcribed locus in Araport11 that was not a miRNA, snRNA, snoRNA, or lncRNA (e.g, rRNA). Intergenic regions were defined as regions that do not overlap any of the above defined features. For each annotation type, DECs were also annotated as overlapping/not overlapping with a TEs using the TAIR10 TE annotation file. Plots were generated using the R package ggplot2 v3.5.0.

To visualise the position of each DEC along the 5 nuclear chromosomes of Arabidopsis, the R packages karyoploteR v1.28.0 (Gel and Serra, 2017) and GenomicRanges v1.54.1 (Lawrence et al., 2013) were used.

4.3.5 Identification of Protein-coding Genes C*is*-regulated by RdDM and ROS1.

4.3.5.1 Genome-wide Scan of *Cis*-regulated Targets

Cis-regulated targets in this Chapter refer to genes which are negatively regulated by DNA methylation in their promoter region, as has been described for known ROS1 targets (Tang et al., 2016; Halter et al., 2021) (Chapter 1, Figure 1.4a). Furthermore, in chromosome arms, DNA methylation is antagonistically regulated by Pol IV-dependent RNA directed DNA methylation (RdDM), which is associated with 23-24 nt sRNAs, and ROS1-dependent DNA demethylation (Tang et al., 2016; Erdmann and Picard, 2020; Harris et al., 2024). As Pol IV-dependent RdDM requires DNA methylation to be present for its targeting (Erdmann and Picard, 2020; Sigman et al., 2021), loss of DNA methylation is expected to reduce Pol IV-dependent RdDM. To test this, hypomethylated DMRs merged across all contexts were overlapped with downregulated 23-24 nt DECs using the R package GenomicRanges v1.54.1 (Lawrence et al., 2013) (function: findOverlaps; options: select = “all”). In L5 and L7, 65.6% and 69.5% of the downregulated DECs overlapped with a hypomethylated DMR in the same line, respectively. Thus, downregulation of 23-24 nt sRNAs is associated with a loss of DNA methylation. As a result, *cis*-regulated genes were defined as protein-coding genes that were

significantly upregulated in E2-treated plants and were associated with a hypomethylated DMR in the promoter region and/or a downregulated 23-24 nt DEC in the promoter region.

4.3.5.2 Verification of *NPR1 cis* Regulation Using qPCR

Verification of the *NPR1 cis*-regulated locus (Supplementary Figure 4.5) was performed using quantitative qPCR techniques in an independent experiment using *XVE:ROS1-YFP* (L5). Seedlings were grown, treated with DMSO or E2, aerial tissue harvested, and DNA extracted exactly as described above. RNA was also extracted from the same pool of material but using the RNeasy plant mini kit (Qiagen, #74904). RNA and DNA concentrations were quantified using a Nanodrop 8000 spectrophotometer. To measure DNA methylation, 500 ng of gDNA was subject to McrBC (New England Biolabs, M0272) digestion for 16 hours at 37 °C, as described by Rausch and Laubinger (2016). Mock digestions were performed using all reagents minus McrBC enzyme. After 16 hours, reactions were terminated by treating samples at 65°C for 20 minutes. Subsequently, 2 µL was used from each sample (digested and mock-digested) for quantitative PCR (qPCR) using SsoAdvanced SYRB Green Supermix (BioRad, 1725270), target-specific primers (Table 4.1), and a Rotor-Gene Q (Qiagen) real-time PCR cycler. Cycle conditions are as follows: 1 cycle of 3 minutes at 98°C, 40 cycles of 10 seconds at 95°C and 30 seconds at 60°C. To determine relative methylation levels, the following equation was used (Rausch and Laubinger, 2016):

$$Relative\ DNA\ methylation\ (\%) = \left(1 - E^{(Ct_{mock} - Ct_{McrBC})}\right) \cdot 100$$

Where '*E'* is the mean estimated amplification efficiency across the assay, and 'Ct_{mock}' and 'Ct_{McrBC}' are the 'take-off' values of mock- and McrBC-digested samples, respectively, as calculated by the Rotor-Gene Q 2.3.5 software.

To measure gene expression from RNA samples, genomic DNA removal and cDNA synthesis were performed using 'Maxima First Strand cDNA Synthesis Kit for RT-qPCR, with dsDNase' (Thermo Scientific, K1672) using 800 ng of RNA. Upon completion of cDNA synthesis, samples were diluted 1:5 using nuclease free water. Subsequently, qPCR was performed using SsoAdvanced SYRB Green Supermix (BioRad, 1725270), gene-specific primers (Table 4.1), and a Rotor-Gene Q (Qiagen) real-time PCR cycler. Cycle conditions are as follows: 1 cycle of 30 seconds at 95°C, 40 cycles of 10 seconds at 95°C and 30 seconds at 60°C. Relative expression values were calculated as described in Wilkinson et al (2023), using *MON1 (AT2G28390)* and *UBC21 (AT5G25760)* as reference genes (Czechowski et al., 2005), and using a DMSO-treated L5 sample as a calibrator sample.

Table 4. 1 Primers used for McrBC-qPCR and RT-qPCR for *NPR1*.

	Name	**Nucleotide sequence**
McrBC-qPCR of *NPR1* promoter region from genomic DNA.	pNPR1-F1	ATTGGTGTTTTCCGGACTCG
	pNPR1-R1	ACGTTGTTGTATTCATAGGCATC
RT-qPCR of *NPR1*	NPR1-F1	GGTGGATTTCATGTTGGAGGT
	NPR1-R2	CGTCCAATAAGTGCCTCTGATAG
RT-qPCR of reference gene *UBC21*	UBC21-F1	CTGCGACTCAGGGAATCTTCTAA
	UBC21-R1	TTGTGCCATTGAATTGAACCC
RT-qPCR of reference gene *MON1*	MON1-F1	AACTCTATGCAGCATTTGATCCACT
	MON1-R1	TGATTGCATATCTTTATCGCCATC

4.4 Results

4.4.1 Ectopic Induction of ROS1 Induces Genome-Wide Shifts in DNA Methylation and Transcription of Coding- and Non-Coding RNAs

The previous Chapter has shown that ectopic induction of ROS1-YFP by two successive E2 treatments in *XVE:ROS1-YFP* lines leads to significant reductions in *Pst* and *Hpa* colonisation (Chapter 3, Figure 3.5). To investigate the molecular mechanisms underpinning the establishment of ROS1-dependent IR, seedlings from Col-0, L5, and L7 genotypes were subjected to an identical dual treatment course of E2 or DMSO before harvesting material for DNA and RNA extractions (Figure 4.2a). Material was harvested 48 hours after the final E2/DMSO treatment (20 DAS), which coincided with the same time point of inoculation with *Pst* and *Hpa* in Chapter 3 (Figure 3.5). For each treatment group, DNA and RNA extractions were performed from the same pool of tissues (n=3). Changes in DNA methylation were analysed by whole genome bisulfite sequencing (WGBS). Transcriptional changes, both coding and non-coding, were examined by ribosome-depleted RNA sequencing (whole transcriptome sequencing; WTS), while changes in small RNA (sRNA) abundance were determined by sRNA sequencing (sRNA-seq).

PCA of DNA methylation revealed a clear separation along the first principle component (PC1) between E2-treated *XVE:ROS1-YFP* seedlings (L5 and L7) and all other treatment groups (Figure 4.2b). No separation was observed between DMSO-treated and E2-treated Col-0 seedlings, indicating that E2 treatment of wild type plants does not have a global impact on DNA methylation. Some separation by genotype was apparent along PC2, but this only explained about 6% of the overall variation in the data (Figure 4.2b) and can be explained by spontaneous drifts in DNA methylation as reported previously (Becker et al., 2011; Stassen et al., 2018). It can thus be concluded E2-treatment of two independent *XVE:ROS1-YFP* lines (L5 and L7) causes a consistent global change in DNA methylation, which is not related to direct chemical effects by E2.

PCA of normalised transcriptomic (Figure 4.2c) and sRNA (Figure 4.3d) read counts also revealed a clear separation along PC1 between E2-treated *XVE:ROS1-YFP* seedlings and all other samples. Unlike the DNA methylome PCA, no clear genotypic separation was observed between DMSO-treated L5, L7 and Col-0 for either transcriptomic or sRNA read counts. Thus, the minor drift observed for ROS1-independent methylome changes between lines does not correspond to transcriptional differences. In contrast, the transcriptome exhibited more substantial separation along PC2 (~20% of the variation) compared to the DNA methylome and sRNAome, with this variation correlating to the time of sample harvesting over

a 2.5-hour period, suggesting the effect is driven by diurnal variations in gene expression. Together these global analyses indicate that ectopic induction of ROS1-YFP in L5 and L7 has a major global impact on DNA methylation, the transcriptome, and the sRNAome, whilst having no effect on Col-0.

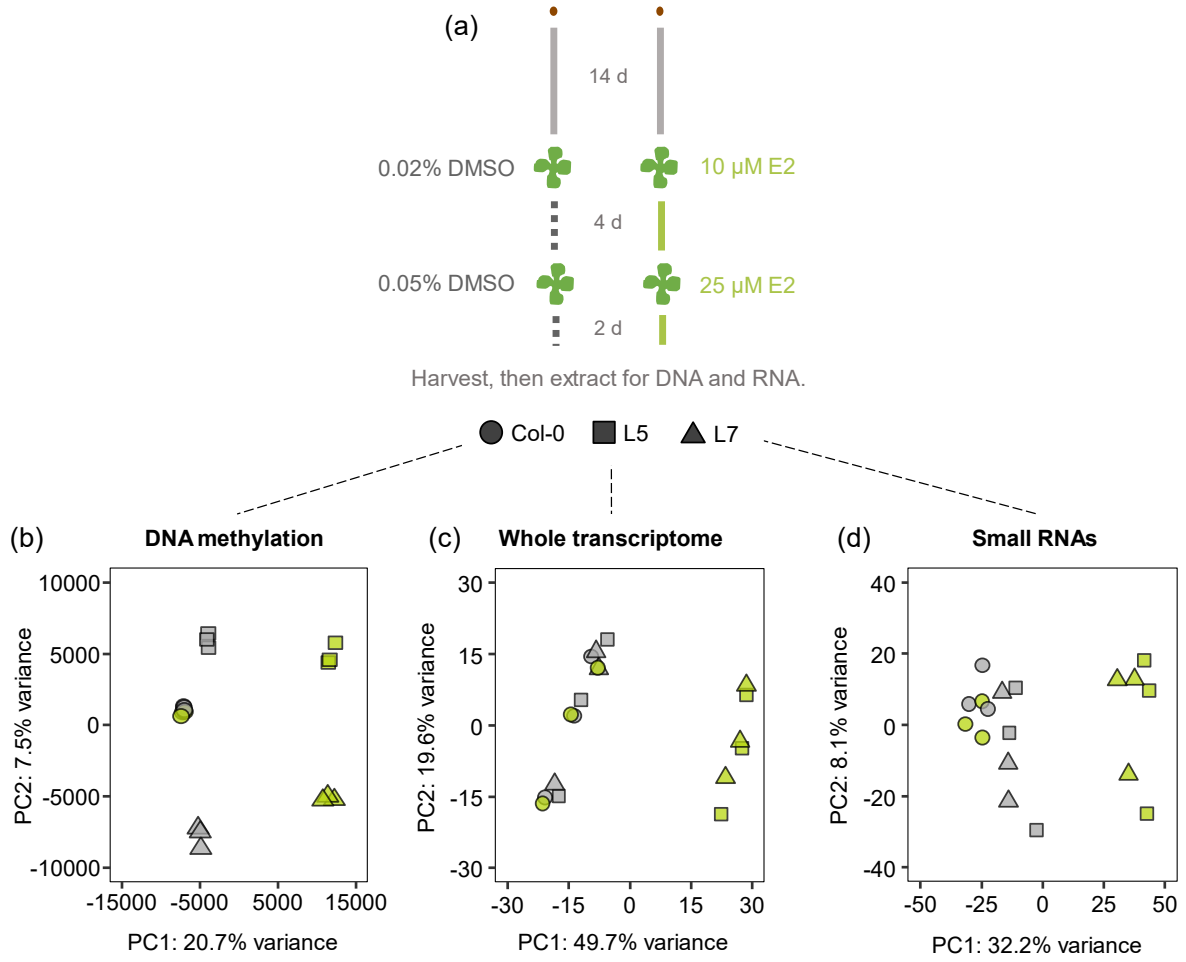


Figure 4.2 Profiling global patterns of DNA methylation, the whole transcriptome, and the expression of small RNAs in Col-0, and two *XVE:ROS1-YFP* lines (L5 and L7) following two successive treatments with estradiol (E2) or DMSO.
(a) Experimental setup to analyse the molecular changes following E2-induced ROS1, exactly as was done in Chapter 3 (Figure 3.5). Col-0, L5, and L7 seedlings were treated twice with E2 (green) or DMSO (grey) at 14 days after sowing (DAS) (10 uM E2 or 0.02% DMSO) and 18 DAS (25 uM E2 or 0.05% DMSO). All aerial tissues were collected for DNA and RNA extractions 48 hours after the second chemical treatment (20 DAS) for three biologically replicated samples per genotype-treatment combination. (b) PCA plot displaying variation in global cytosine DNA methylation for E2- and DMSO-treated (green and grey, respectively) Col-0 (circle), L5 (square), and L7 (triangle) seedlings. (c, d) PCA plots displaying global variation in (a) the normalised expression of coding and non-coding transcripts (whole transcriptome), and (b) global variation in normalised expression of small RNA clusters (small RNAs) in Col-0 (circle), L5 (square), and L7 (triangle), treated with either DMSO (grey) or E2 (green).

4.4.2 Induction of ROS1 Leads to Global Reductions in DNA Methylation, Particularly at TE-Gene Boundaries in the Chromosome Arms

DNA methylation in E2-treated seedlings was assessed at the chromosomal level by averaging DNA methylation over 100 kb windows and expressing these values relative to DMSO-treated seedlings. The relative DNA methylation was then plotted against their distance from the centre of the pericentromeric regions (see methods for details) (Figure 4.3a). This analysis revealed that for all cytosine contexts (allC), and distinct cytosine contexts (CG, CHG, CHH), there was a clear reduction in DNA methylation along the chromosome arms in E2-treated L5 and L7, which was not apparent in E2-treated Col-0 seedlings (Figure 4.3a). Interestingly, loss of DNA methylation in E2-treated L5 and L7 was smaller in the pericentromeric regions, with some windows being hypermethylated towards the midpoint of the pericentromere, particularly in the CHG context (Figure 4.3a). Thus, ectopic induction of ROS1 in *XVE:ROS1-YFP* lines causes DNA demethylation in the chromosome arms.

Next, DNA methylation at all cytosine contexts (allC) was assessed at protein coding genes, transposable elements (TEs), and the flanking regions (± 1 kb) of these features. While E2-treated Col-0 seedlings showed no changes in DNA methylation in any of these regions, the E2-treated L5 and L7 seedlings showed clear reductions in DNA methylation in both TE and genic regions (Figure 4.3b). Interestingly, the reduction in DNA methylation was more striking upstream and downstream of protein-coding genes (L5: -23.9% and -20.8% respectively; L7: -25.5% and -22.8% respectively) compared to the gene body itself (L5: -14.4%; L7: -17.1%). This is in contrast with TEs, where the relative reduction in DNA methylation was greatest within the TE body (L5: -12.4%; L7: -14.2%), rather than the upstream and downstream regions of the TE (L5: -8.56% and -8.53% respectively; L7: -10.1% for both) (Figure 4.3b). Hence, E2-induced ROS1-YFP in L5 and L7 leads to a preferential reduction in DNA methylation in TE bodies and flanking regions of protein-coding genes.

To further analyse the impacts of ROS1-YFP induction on DNA methylation, the R package DSS v2.46.0 (Park and Wu, 2016) was used to identify differentially methylated regions (DMRs). The analysis was performed for all cytosine contexts (allC) and the three distinct cytosine contexts (CG, CHG, and CHH). Few DMRs were identified in E2-treated Col-0 seedlings, with only 22 hyper-methylated DMRs and 25 hypo-methylated identified across the four tested cytosine contexts (Figure 4.3c). By contrast, the E2-treated L5 and L7 seedlings had thousands of DMRs across the four cytosine contexts, with 11,067 hypo-methylated and 18 hyper-methylated DMRs in L5, and 12,158 hypo-methylated and 30 hyper-methylated DMRs in L7. In E2-treated *XVE:ROS1-YFP* lines, the greatest number of DMRs occurred in CG contexts, most of which were not associated with TEs (Figure 4.3c). In contrast, in non-

CG contexts, approximately half of ROS1-induced hypomethylated DMRs overlapped TEs (Figure 4.3c) (all DMRs listed in Supplementary Data 4.8).

To better understand where changes in DNA methylation tend to occur in the genome, all hypomethylated DMRs (hypo-DMRs) were merged within a line, such that any overlapping hypo-DMRs from different contexts were combined to form a larger single DMR. This generated a merged and simplified set of 7,829 and 8,951 hypo-DMRs in L5 and L7, respectively, of which 4,971 were shared between L5 and L7. The latter set of consistent ROS1-associated hypo-DMRs (Figure 4.3d; Supplementary Data 4.8) was annotated using Araport11 gene annotations (Cheng et al., 2017) and TAIR10 TE annotations. Annotation enrichment was tested by comparing these hypo-DMRs against randomly selected and annotated sets of DMRs of equal size and number. The consistent hypo-DMRs were significantly enriched in gene-TE boundary regions (Figure 4.3e). The most significant enrichment was in gene promoters containing one or more TEs, occurring on average 5.1x more than in the random DMR sets (Figure 4.3e) (Fisher's exact test, $q < 0.001$). Gene promoters lacking TEs were also significantly enriched, but only on average 1.6x more than in the random DMR sets (Fisher's exact test, $q < 0.001$). All other TE-gene boundary annotations were also significantly enriched, namely downstream regions, exons, and introns that overlapped or shared a DMR with a TE ($\log_2$ fold change > 0; Fisher's exact test $q < 0.05$) (Figure 4.3e). In contrast, except for promoters, all other genic regions that were not associated with TEs had no (under)enrichment ($q > 0.05$) or were significantly underrepresented for ROS1-dependent DNA demethylation ($\log_2$ fold change < 0; Fisher's exact test $q < 0.05$). This pattern complements previous characterisations of ROS1-targeted loci in *ros1* mutants, which propose ROS1 functions to prevent the spread of DNA methylation from TE regions to gene bodies (Tang et al., 2016). Interestingly, however, TEs occurring in intergenic regions were also underrepresented in the ROS1-associated hypo-DMR set, occurring on average 1.7x less than in the random DMR sets (Fisher Exact Test; $q < 0.001$) (Figure 4.3e). This may reflect the lack of DNA hypomethylation observed in pericentromeric regions (Figure 4.3a) which are TE-rich and gene-poor. Thus, two successive inductions of ROS1-YFP result in highly specific patterns of DNA demethylation, exhibiting distinct differences depending on the genomic region and loci.

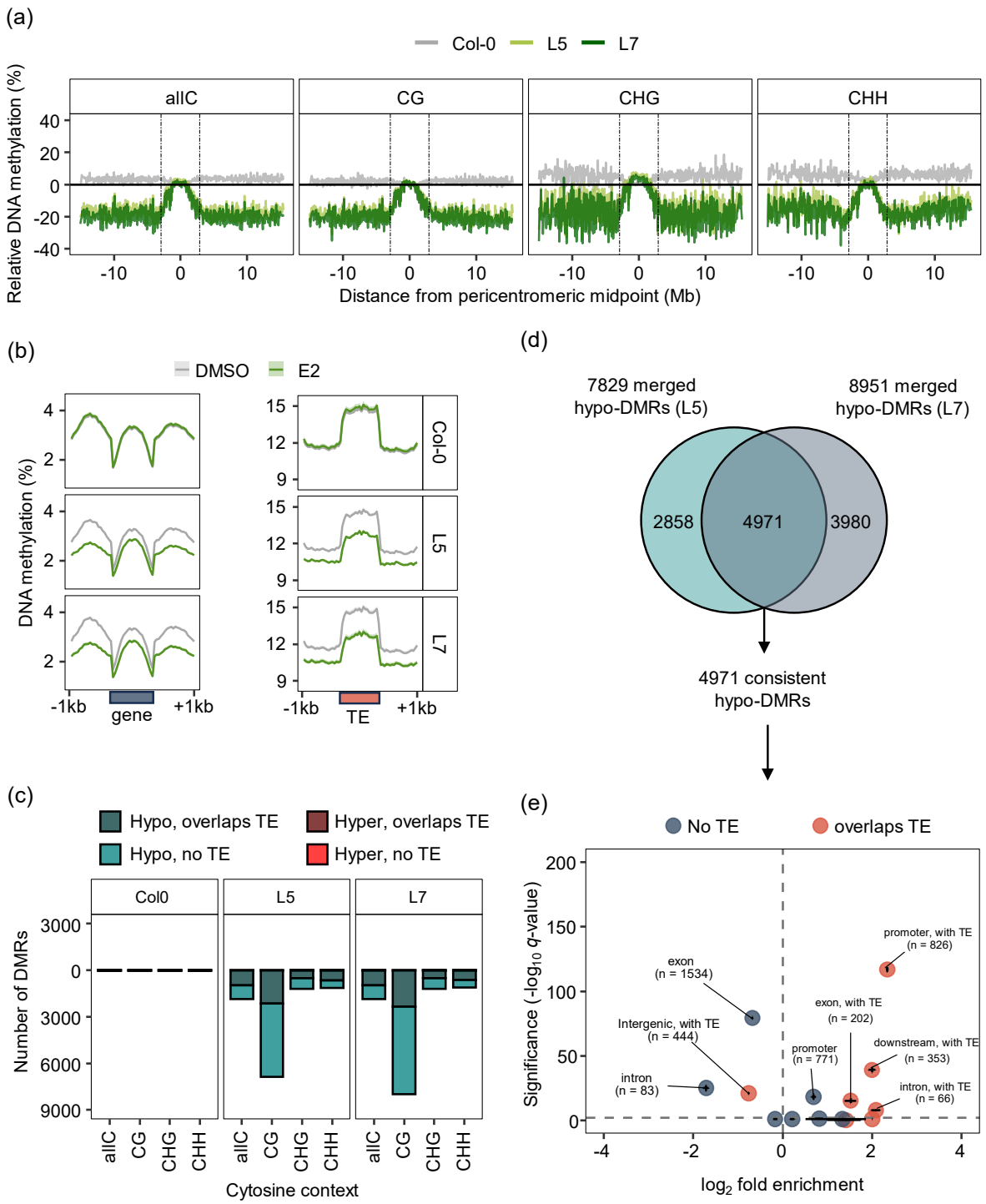


Figure 4.3 Characterisation of global DNA methylation changes following two successive E2 treatments in Col-0, and two *XVE:ROS1-YFP* lines (L5 and L7).
(a) Chromosome-level metaplots of DNA methylation in E2-treated plants are shown as a percentage relative to DMSO-treated plants in Col-0 (grey), L5 (light green), and L7 (dark green). For each chromosome, DNA methylation levels for all cytosine contexts (allC), CG, CHG, and CHH were binned into 100 kb windows and averaged based on their relative distance to the midpoint of the pericentromere. Dashed vertical lines indicate the minimum and maximum pericentromeric start and end positions relative to the pericentomeric midpoint. (b) Metaplots showing average allC DNA methylation in E2-treated (green) and DMSO-treated (grey) Col-0, L5, and L7 seedlings within protein-coding genes (gene; left panel), and transposable elements (TEs; right panel). For each gene/TE, DNA methylation was

binned into 20 equal sized bins for the gene/TE body (for genes: transcription start site to transcription termination site; TSS to TTS), the upstream region (-1 kb) and the downstream region (+1 kb). (c) Counts of differentially methylated regions (DMRs) in E2-treated vs DMSO-treated Col-0, L5, and L7 for allC, CG, CHG, and CHH contexts. Hypomethylated DMR (Hypo) counts (blue) are shown below the y=0 line, hypermethylated DMR (Hyper) counts (red) are shown above the y=0 line. Counts of DMRs overlapping TEs are shown in dark blue for hypomethylated DMRs, and dark red for hypermethylated DMRs. (d) Hypomethylated DMRs (hypo-DMRs) from all cytosine contexts (c) were merged and simplified for L5 and L7. A total of 4,971 of these merged hypo-DMRs were shared between L5 and L7, which were used as a single set of consistent hypo-DMRs for annotation and enrichment analysis (e). (e) Consistently hypomethylated DMRs by ROS1 (consistent hypo-DMRs) were annotated with Araport11 (Cheng et al., 2017). For each hypo-DMR overlapping with an assigned annotation feature (one of exon, intron, promoter, downstream, non-coding RNA, intergenic), DMRs were also analysed for overlaps with TEs (coloured orange if true), as determined by the TAIR10 TE annotation file (https://www.arabidopsis.org/). Four sets of random DMRs of equal size and number were generated and annotated for comparison. Two-sided Fisher Exact Test's on the number of hypo-DMRs matching a given annotation feature were performed between selections of consistent hypo-DMRs and one of the four random DMR sets. The average FDR-adjusted *p* value (*q*-value) for these tests were plotted against the average fold change (real count / random count) on a $-\log_{10}$ and $\log_2$ scale, respectively. Error bars indicate the standard error of the mean, based on comparisons with four random DMR sets. The horizontal dashed line shows $q = 0.05$. Significantly under- or over-represented annotation features are labelled with the number of real hypo-DMRs mapping to the given annotation feature.

4.4.3 Induction of ROS1 Leads to Transcriptional Activation of Genes Involved in DNA Repair and Plant Immune Responses

Ribosome-depleted RNA sequencing was employed to investigate impacts of ROS1-dependent DNA demethylation on the whole transcriptome, including coding and non-coding loci. Differentially expressed loci in response to E2 treatment, referred to as differentially expressed genes (DEGs), were identified in Col-0, L5, and L7 using DEseq2 (Love et al., 2014) ($p_{adj} < 0.05$). E2-treated L5 and L7 displayed 2,530 and 2,777 upregulated DEGs, respectively, whereas E2-treated Col-0 had only 1 upregulated DEG compared to the DMSO control. Furthermore, there were 1,108 downregulated DEGs in L5, 853 downregulated DEGs in L7, and 0 downregulated DEGs in Col-0 (Supplementary Data 4.4). *AT3G07105* was the only DEG identified in Col-0, which encodes a noncoding natural antisense transcript to the protein coding gene *MILDEW RESISTANCE LOCUS O 3* (*MLO3*). This long non-coding RNA (lncRNA) gene was also upregulated in E2-treated L7 plants but not in E2-treated L5 plants (Supplementary Figure 4.1). Importantly, however, the differential expression of his lncRNA gene did not affect *MLO3* expression in Col-0. Thus, E2 has a negligible impact on the transcriptome in wild-type Col-0 plants, but induces drastic changes in the transcription of coding and non-coding RNA in *XVE:ROS1-YFP* lines (L5 and L7).

Hierarchical clustering of all DEGs separated samples into two distinct clusters, E2-treated *XVE:ROS1-YFP* lines, and all other treatment groups (E2-treated Col-0, DMSO-

treated Col-0, and DMSO-treated *XVE:ROS1* lines). There were no further subclusters associated with genotype or chemical treatment (Figure 4.4a). Clustering of genes split the DEGs into two clusters: upregulated in E2-treated *XVE:ROS1-YFP* lines ('up' cluster) and downregulated E2 *XVE:ROS1-YFP* lines ('down' cluster) (Figure 4.4a). The upregulated cluster comprised of 3,460 loci, of which 3,297 were protein-coding genes (mRNA), 88 were non-coding genes (ncRNA), and 75 were TEs. The downregulated cluster comprised of 1,264 loci, of which 1,219, 29, and 16 were mRNA, lncRNA, and TE loci, respectively (Figure 4.4b). The proportion of each feature type in both clusters did not significantly differ from the expected distributions based on their relative proportions in the Araport11-annotated genome (Chi-squared test; $p > 0.05$).

Protein-coding genes within the upregulated cluster were enriched with gene ontology (GO) terms related to DNA repair and immune responses, including "DNA damage response" and "systemic acquired resistance" (Figure 4.4c). Downregulated protein-coding genes were enriched with GO terms linked to photosynthetic processes, including "photosynthesis" and "plastid organisation" (Figure 4.4c) (all GO terms listed in Supplementary Data 4.5). This is associated with reduced green leaf areas in E2-treated L5 and L7 at the time of harvest (Supplementary Figure 4.2). Thus, activation of ROS1 in E2-treated *XVE:ROS1-YFP* generally leads to widespread transcriptional activation of defence responses with a concurrent suppression of photosynthetic pathways.

Transcriptional changes of loci in L5 and L7 were positively correlated ($r = 0.94$, n = 4724, $p < 0.001$). Of all DEGs in L5 or L7, 53.8% were shared in both lines in the same direction, called overlap DEGs (Figure 4.4d). An additional 45.5% were significantly differentially expressed in one line but not the other, although both had fold-changes in the same direction, called unique concordant DEGs (Figure 4.3d). No DEGs were significantly expressed in opposite directions in E2-treated L5 and E2-treated L7 (opposite DEGs) (Figure 4.3d). However, a small proportion (0.66%) of DEGs were significantly differentially expressed in one line and, while not significantly differentially expressed in the other line, had a fold-change in the opposite direction, which were referred to as unique discordant DEGs (Figure 4.3d). Hence, activation of ROS1-YFP in the two independent *XVE:ROS1-YFP* lines by E2 treatment causes a robust and highly consistent change in the transcription of coding and non-coding RNA.

In addition to identifying DEGs, protein-coding genes were tested for differences in the abundance of alternatively spliced or alternatively polyadenylated transcripts in response to E2 treatment, using the 3D RNA-seq analysis pipeline (Guo et al., 2021). For this analysis, differential transcript usage is referred to here as differentially alternatively spliced (DAS)

genes. In E2-treated Col-0, only one DAS gene was identified, whereas E2 treatment of L5 and L7 yielded 15 and 12 DAS genes, respectively (Figure 4.4e). As these numbers were relatively low, no significant GO terms were associated with DAS genes in L5 and L7 (p_{adj} > 0.05). However, it is noteworthy that the histone demethylase *INCREASED IN BONSAI METHYLATION 1* (*IBM1*) was identified as a significant DAS gene in L5, which is a H3K9 deacetylase implicated in the regulation of plant defences (Rigal et al., 2012; Chan and Zimmerli, 2019; Lv et al., 2022) (Chapter 1, Figure 1.2). E2 treatment promoted accumulation of the short and non-functional *IBM1-S* transcript variant and reduced accumulation of the long functional *IBM1-L* variant transcript (Figure 4.4f) (Rigal et al., 2012). While *IBM1* was not statistically classified as a DAS gene in L7, the pattern of transcript usage in this line was similar to that in L5 (Figure 4.4f). Interestingly, the elevated expression of *IBM1-S* is associated with an intronic DMR at the alternative polyadenylation site following *ROS1-YFP* induction in both E2-treated L5 and L7 (Supplementary Figure 4.3a).

In Arabidopsis, ROS1 and its homologues DEMETER (DME), DME-LIKE2 (DML2), and DML3, are plant-specific bifunctional enzymes that excise 5-methylcytosine (5mC) bases and cleave the DNA sugar-phosphate backbone via β- or δ-elimination (Zhang et al., 2022). Base-excision repair (BER) pathways then fill the nicked DNA with an unmodified cytosine, thereby removing DNA methylation (Gong et al., 2002; Agius et al., 2006; Ponferrada-Marín et al., 2010; Lee et al., 2014; Zhang et al., 2022; Du et al., 2023). In contrast, in animals, DNA demethylation is mediated by TEN-ELEVEN TRANSLOCATION (TET) family proteins (TETs) that convert 5mC to 5-hydroxymethyl-, 5-formyl- and 5-carboxyl-cytosine (5hmC, 5fC, 5caC, respectively). These oxidised products (5hmC, 5fC, and 5caC) are poor DNA METHYLTRASFERASE 1 (DNMT1) substrates, resulting in passive loss of DNA methylation during replication (Valinluck and Sowers, 2007; Ji et al., 2014; X. Zhang et al., 2023). Alternatively, 5fC and 5caC are recognised and excised by thymine-DNA glycosylate (TDG), returning the site to unmethylated status through BER (Zhang et al. 2023). In plants, there are no known orthologs of TET or TDG enzymes (Iyer et al., 2009; Grin et al., 2023) (OrthoDB v12.0; https://www.orthodb.org/). Therefore, plants and animals have evolved independent mechanisms to regulate cytosine DNA methylation.

Despite this, In Arabidopsis, Ji et al. (2018) reported that constitutive expression of TET1 by a 35S promoter induces genome-wide reductions in DNA methylation. As plants are not expected to have enzymes that can recognise oxidised 5mC products to carry out active DNA demethylation, this loss is likely driven by a passive process, whereby DNA methyltransferases fail to recognise these marks, as occurs with DNMT1 in animals (Valinluck and Sowers, 2007; Ji et al., 2014; X. Zhang et al., 2023). Despite this different mechanism, TET1-dependent DNA demethylation activated the expression of thousands of genes (Ji et al.,

2018) (Supplementary Data 4.4), similar to ROS1-dependent DNA demethylation in this study (Figures 4.4a, 4.4g). By comparing both datasets, 532 genes were commonly upregulated in *35S:TET1* lines (*35S:TET1-1* and *35S:TET1-2*) and *XVE:ROS1-YFP* lines (L5 or L7) (Figure 4.4g). Notably, the top 5 most significantly enriched GO terms for these commonly upregulated genes by TET and ROS1 were associated with salicylic acid (SA)-dependent defence responses. These terms included “response to salicylic acid”, “systemic acquired resistance”, and “response to molecule of bacterial origin” (Figure 4.4g). Genes uniquely upregulated in the *XVE:ROS1-YFP* lines were enriched for DNA-damage-associated GO terms such as “DNA damage response”, “response to ionizing radiation” and “chromatin organisation”. In contrast, genes uniquely upregulated in *35S:TET1* lines were linked to jasmonic acid (JA)-dependent pathways and photosynthesis. The significantly enriched GO term for these genes included “photosynthesis”, “jasmonic acid mediated signalling pathway”, and “defence response to insect” (all GO terms are presented in Supplementary Data 4.5) (Figure 4.4g). Therefore, loss of DNA methylation, independent of the mechanism driving it, is associated with the activation of SA-dependent defence pathways in Arabidopsis.

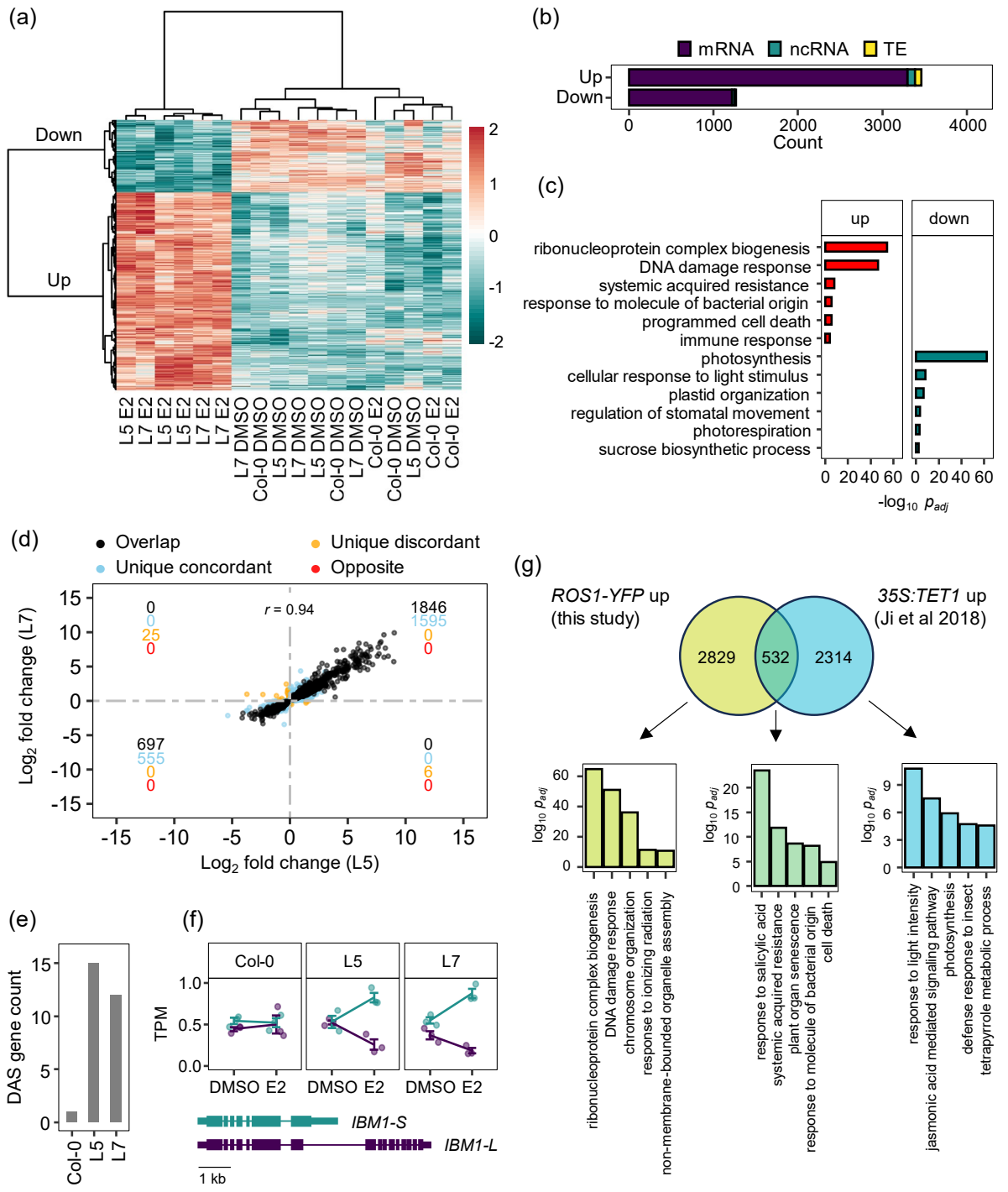


Figure 4.4 Transcriptional responses upon E2 treatment in Col-0, L5, and L7 (Figure 4.2a).
(a) Heatmap presenting expression levels of all 4,724 differentially expressed loci/genes (DEGs) between E2- and DMSO-treated plants within one genotype using DESeq2 (Love et al., 2014). Heatmap-projected values represent per gene *Z*-scores of transformed normalised read counts from 3 biological replicates for each genotype-treatment combination. Scores were clustered by DEG and sample (Pearson correlation, Ward). Two distinct gene clusters were identified, labelled “Down” and “Up”, reflecting their expression patterns in E2-treated L5 and L7 relative to all other genotype-treatment combinations. (b) Feature type counts of all DEGs identified in the “Up” and “Down” cluster from the heatmap in Figure 4.4a. Each DEG was annotated either as a protein-coding gene (mRNA; purple) or a transposable element (TE; yellow). All other features were assigned as a noncoding locus (ncRNA; blue). (c) Selected significant gene ontology (GO) terms for biological processes associated with mRNA

loci in the “Up” (red) and “Down” (blue) cluster from the heatmap in Figure 4.4a (p_{adj} < 0.05). All GO terms are listed in Supplementary Data 4.5. (d) Correlation between $\log_2$ fold-change (FC) values of DEGs identified in L5 or L7. Pearson correlation coefficient (r) = 0.94 (p < 0.001). Each point represents a DEG identified in E2-treated L5 or L7 classified as ‘overlap’ (both p_{adj} < 0.05; both $\log_2$FC > 0 or both $\log_2$FC < 0), ‘unique concordant’ (blue) (only one p_{adj} < 0.05; both $\log_2$FC > 0 or both $\log_2$FC < 0), ‘unique discordant’ (orange) (only one p_{adj} < 0.05; one $\log_2$FC < 0 and one $\log_2$FC > 0), or ‘opposite’ (red) (both p_{adj} < 0.05; one $\log_2$FC < 0 and one $\log_2$FC > 0). Coloured text (from top to bottom: ‘overlap’, ‘unique concordant, ‘unique discordant’, ‘opposite’) shows the sum of DEGs corresponding to these classes in each quadrant of the graph (centred around 0, 0). (e) Counts of protein-coding genes exhibiting differential transcript usage following E2-treatment in L5, L7 and Col-0, referred to as differentially alternatively spliced (DAS) genes. DAS genes were identified using the 3D RNA-seq pipeline (Guo et al., 2021). (f) Transcript per million (TPM) counts of short non-functional *IBM1* transcripts (*IBM1-S* / AT3G07610.2) (blue) and long functional *IBM1* transcripts (*IBM1-L* / AT3G07610.1) (purple) in Col-0, L5, and L7 upon two succesive treatments with DMSO or E2 (Figure 4.1a). Transcript models are shown to scale (black bar = 1 kb). (g) A Venn diagram showing the overlap of significantly upregulated DEGs (p_{adj} < 0.05) consistently identified in E2-treated *XVE:ROS1-YFP* lines (L5 and L7) and significantly upregulated DEGs (p_{adj} < 0.05) in *35S:TET1* lines (Ji et al., 2018). Top 5 GO enrichment terms (in order of statistical significance) for biological process in the overlapping DEG set (green), and the unique non-overlapping DEG sets (yellow and blue for *XVE:ROS1-YFP* and *35S:TET1*, respectively).

4.4.4 Induction of ROS1 Alters Genome-Wide Distributions of 23-24 nt sRNAs

Plant small RNAs (sRNAs) in the size range of 21-24 nucleotides (nt) are ubiquitous regulators of protein-coding genes and TE genes via post-transcriptional gene silencing (PTGS; Borges and Martienssen 2015; Zhan and Meyers 2023) and the RNA-directed DNA methylation (RdDM) pathways (Matzke and Mosher, 2014; Cuerda-Gil and Slotkin, 2016; Erdmann and Picard, 2020; Sigman et al., 2021). In response to stress, some sRNAs can even promote the expression of defence-related genes via the secondary nuclear function of the sRNA-binding protein ARGONAUTE 1 (AGO1) (Liu et al. 2018; Wilkinson et al. 2023). To investigate the link between ROS1-induced DNA demethylation and sRNAs, the same material that was subjected to WBGS and RNA-seq analysis (Figure 4.2a) was sequenced for sRNAs (18-30 nt).

Differentially expressed sRNA clusters (DECs) were identified using ShortStack and DESeq2 (Axtell, 2013; Shahid and Axtell, 2014; Love et al., 2014). E2-treated L5 and L7 seedlings had thousands of DECs compared to DMSO-treated seedlings, with 57.2% overlapping between the two lines (Figure 4.5a) (Supplementary Data 4.9). Of the DECs uniquely identified as significant in L5 or L7, 93.7% were concordantly expressed between the lines (at least one line with p_{adj} <0.05; and both $\log_2$ FC > 0 or both $\log_2$ FC < 0). E2-treated Col-0 had 17 DECs compared to DMSO-treated Col-0 seedlings (Figure 4.5a). Of these, only 1 DEC was also identified in E2-treated *XVE:ROS1-YFP* lines (L5 or L7).

Hierarchical clustering of all DECs grouped samples into two distinct clusters: E2-treated *XVE:ROS1-YFP* lines, and all other treatment groups (E2-treated Col-0, DMSO-treated Col-0, and DMSO-treated *XVE:ROS1* lines). No further subclusters were observed based on genotype or chemical treatment (Figure 4.5b). Two distinct clusters of DEC expression were identified: those upregulated in E2-treated *XVE:ROS1-YFP* lines ('up' cluster) and those downregulated E2 *XVE:ROS1-YFP* lines ('down' cluster) (Figure 4.5b). The number of DECs in the 'up' and 'down' cluster were almost balanced, with 2079 and 2449 DECs in each, respectively (Figure 4.5c). Most DECs were 24 nt in size (Figure 4.5c), comprising 96.8% of the 'down' cluster and 82.2% of the 'up' cluster, demonstrating that ROS1 has a large impact on the abundance of 24 nt sRNAs, which are associated with polymerase IV (Pol IV)-dependent RdDM (Cuerda-Gil and Slotkin, 2016). Interestingly, the 'up' cluster showed greater diversity in DEC classes compared to the downregulated cluster (Fisher Exact Test, $p < 0.001$) (Figure 4.5c).

Out of the entire DEC set associated with *XVE:ROS1-YFP* lines, only 31 predicted microRNAs (miRNAs) were identified, representing just 0.4% of identified DECs. Notably, all these miRNAs were upregulated in one or both of the E2-treated *XVE:ROS1-YFP* lines. MiRNAs are well characterised master regulators of gene expression with diverse roles in a variety of biological processes (Zhan and Meyers, 2023). For example, miRNAs can promote the posttranscriptional degradation of gene transcripts through associations with AGO1 (Zhan and Meyers, 2023). Interestingly, a small proportion (1.3%) of upregulated DECs were categorised as "not defined" (ND), indicating no dominant sRNA read size within the cluster and/or a majority of reads outside of the expected 20-24 nt range for siRNAs (Zhan and Meyers, 2023). These ND DECs, which often mapped to gene bodies (63.8% of all ND DECs), comprised a range of sRNA read sizes, which indicates they derive from degraded transcripts. To test whether this may be driven by miRNAs, the predicted targets of the upregulated miRNAs in E2-treated *XVE:ROS1-YFP* lines were identified using targetFinder (Fahlgren et al., 2007; Srivastava et al., 2014). However, none of the corresponding targets matched the ND DECs. Instead, 16 of these miRNAs had predicted protein-coding gene targets, which were enriched for GO terms related to development, including "meristem initiation", "regulation of development, heterochronic", and "regulation of auxin metabolic process" (Supplementary Figure 4.4a) (Supplementary Data 4.7). This enrichment could partially explain the observed growth reduction and down-regulation of development associated transcripts (Figure 4.3c) seen in E2-treated *XVE:ROS1-YFP* lines (Supplementary Figure 4.4). However, analysis of the gene bodies and flanking regions (± 1 kb TSS / TTS) of their associated *MIR* genes revealed low levels of DNA methylation, with no changes in E2-treated *XVE:ROS1-YFP* lines (two-sample t-tests, $p > 0.05$; Supplementary Figure 4.4c). This suggests that ROS1 does not

regulate *MIR* gene expression through local changes in DNA methylation. Instead, ROS1 may influence *MIR* expression through unknown *trans*-acting mechanisms. For example, ROS1-mediated DNA demethylation might enhance the expression of certain transcription factors, which in turn could promote *MIR* transcription (Zhan and Meyers, 2023). Nevertheless, these findings suggest that miRNAs only play a minor role in the transcriptional changes associated with ROS1-dependent DNA demethylation.

On a global level, E2-treated *XVE:ROS1-YFP* lines showed changes in the distribution of sRNA read-length distributions (Figure 4.5d). Conducting pairwise t-tests on the percent mapped reads for each read size between E2-treated and DMSO-treated plants within a genotype revealed no significant differences in Col-0 (t-test, $p > 0.05$). In L7, there were significant reductions in the proportion of 23 and 24 nt sRNAs mapping to the genome in E2 treated seedlings (t-test, $p < 0.05$). Likewise, in L5, a significant reduction in the proportion of mapped 23 nt sRNAs was observed (t-test, $p < 0.05$), but the reduction for 24 nt sRNAs was not significant (t-test, $p > 0.05$) (Figure 4.5d). Pol IV-dependent RdDM, also known as canonical RdDM, results in the production of predominantly 23-24 nt sRNAs, and requires methylation to be present to be active (Erdmann and Picard, 2020; Sigman et al., 2021). In contrast, Pol II-dependent RdDM, which relies on the active transcription of loci such as TEs, is additionally guided by 21-22 nt sRNAs (Matzke and Mosher, 2014; Cuerda-Gil and Slotkin, 2016; Sigman et al., 2021). Therefore, the relative proportion of 21-22 nt to 23-24 nt sRNAs in the genome provides a rough estimate of the respective contributions of Pol II-dependent and Pol IV-dependent RdDM, respectively. The average ratio of 21-22 nt to 23-24 nt sRNAs in DMSO-treated seedlings was 0.48 in Col-0, 0.65 in L5, and 0.64 in L7. In E2-treated seedlings, this ratio remained nearly identical in Col-0 at 0.49. However, in L5 and L7, the ratios increased to 0.96 and 0.82, representing relative increases of 48% and 28%, respectively. Despite an overall bigger difference between DMSO and E2 treated seedlings in L5, the shift was only detected as significantly different in L7 (t-test, $p < 0.05$). Nevertheless, the increased proportion of 21-22 nt sRNAs in E2-treated *XVE:ROS1-YFP* lines suggests an increase in Pol II-dependent RdDM and/or a decrease in Pol IV-dependent RdDM. This could, in part, be driven by the upregulation of TEs following DNA demethylation in E2-treated L5 and L7 seedlings (Figure 4.3; Figure 4.4b) (Cuerda-Gil and Slotkin, 2016).

To characterise the genomic location of all DECs, those in E2-treated L5 and L7 were annotated with Araport11 (Cheng et al., 2017) and examined for overlaps with TEs, as per the TAIR10 TE annotation file. Few DECs mapped to Araport11-defined non-coding sRNA loci, including miRNAs, tRNAs, snRNAs, and snoRNAs (Figure 4.5e). For both L5 and L7, >75% of all downregulated DECs occurred in protein-coding genic regions (gene body ± 1 kb), with >55% also overlapping with a TE (Figure 4.5e). Among all annotation classes, the

promoters of protein-coding genes containing TEs had the highest number of downregulated DECs for both lines, which is consistent with the previously observed enrichment of hypo-DMRs at TE-gene boundaries in *XVE:ROS1-YFP* lines (Figure 4.3e). Indeed, in L5 and L7, 65.6% and 69.5% of the downregulated DECs overlapped with a hypomethylated DMR in the same line, indicating a tight association between DNA hypomethylation and downregulation of sRNAs. In contrast, only 2.6% and 3.5% of the upregulated DECs in L5 and L7 overlapped with a DMR (hypermethylated or hypomethylated), respectively. Furthermore, the majority of upregulated DECs occurred in TE-containing intergenic regions which were significantly underrepresented for hypomethylated DMRs in E2-treated *XVE:ROS1-YFP* lines (Figure 4.3e).

Plotting the genomic location of all DECs across the 5 nuclear chromosomes revealed a clear disparity in their distribution: downregulated DECs were predominantly located along chromosome arms, whereas upregulated DECs were mostly found in (peri)centromeric regions. This pattern aligns with the observation that E2-induced DNA demethylation in *XVE:ROS1-YFP* lines is most pronounced along chromosome arms (Figure 4.3a). This indicates that ROS1-mediated DNA demethylation supresses Pol IV-dependent RdDM in chromosome arms, as >97% of all DECs are associated with 23-24 nt sRNAs (Figure 4.5c) (Erdmann and Picard, 2020). In contrast, (peri)centromeric regions, which exhibit a lack of DNA demethylation and some evidence of DNA hypermethylation in E2-treated *XVE:ROS1-YFP* lines (Figure 4.3a), demonstrate enhanced activity of Pol IV-dependent RdDM as most upregulated DECs are also associated with 24 nt sRNAs. Thus, ROS1-induced DNA demethylation leads to a genomic redistribution of sRNAs associated with Pol IV-dependent RdDM, which may specifically function to counteract DNA demethylation in TE-rich (peri)centromeric regions.

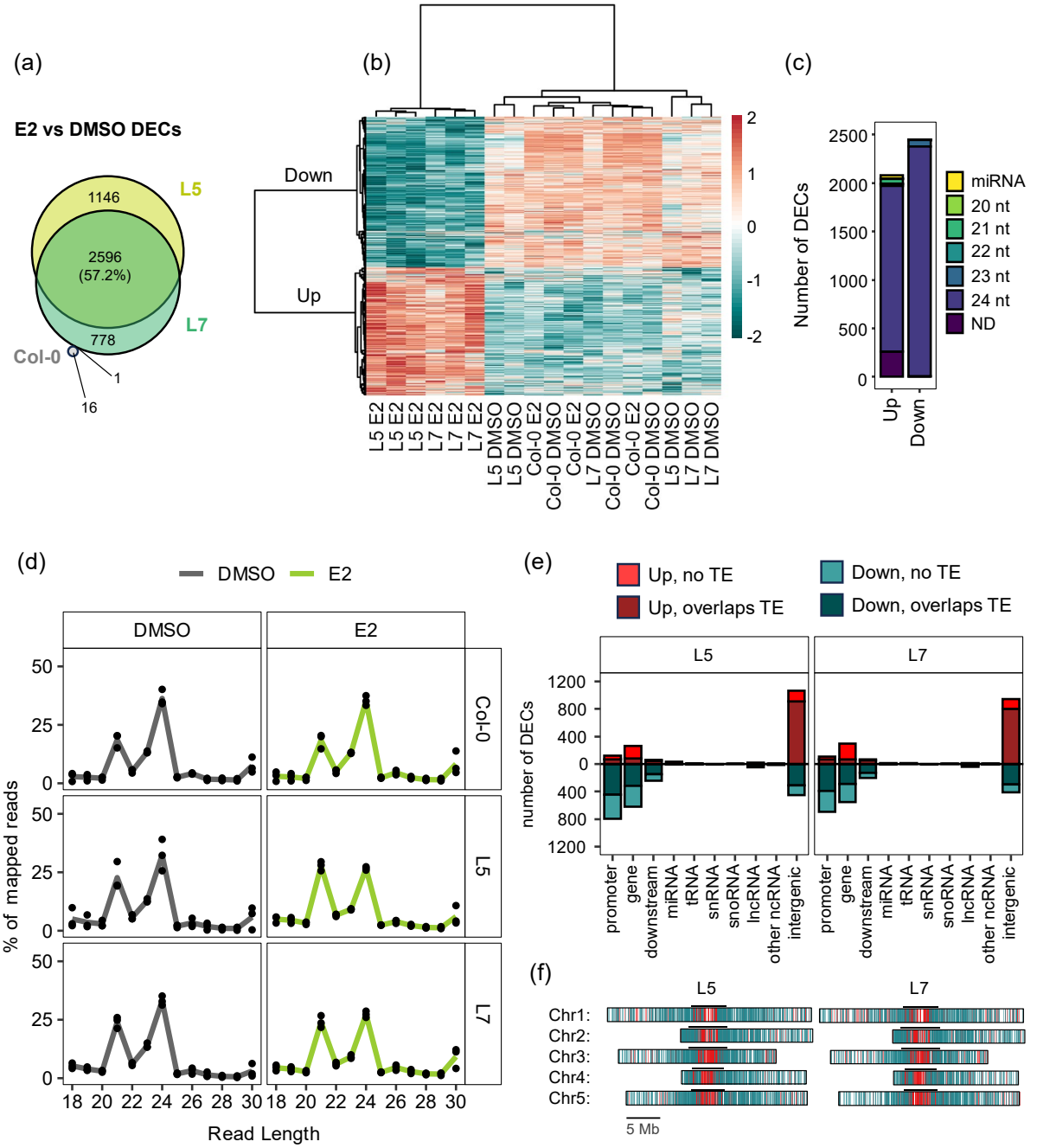


Figure 4.5 Characterisation of small RNA (sRNA) changes following E2 treatment in Col-0, and two *XVE:ROS1-YFP* lines (L5 and L7).

(a) Differentially expressed sRNA clusters (DECs) between E2- and DMSO-treated seedlings L5 (yellow), L7 (green), and Col-0 (grey) genotypes, identified using ShortStack and DESeq2 (Axtell, 2013; Shahid and Axtell, 2014; Love et al., 2014). (b) Heatmap projection of sRNA abundance of all 4,520 DECs between E2- and DMSO-treated L5, L7, and Col-0 plants, using DESeq2 (Love et al., 2014). Heatmap-projected values represent per cluster Z-scores of transformed and normalised read counts from 3 biological replicates for each genotype-treatment combination. Scores were clustered by DEC and sample (Pearson correlation, Ward). Two distinct clusters were identified, labelled “Down” and “Up”, reflecting their expression pattern in E2-treated L5 and L7 relative to all other genotype-treatment combinations. (c) Cluster type counts, as defined by ShortStack, of all DECs identified in the “Up” and “Down” cluster from the heatmap (b). Clusters where >80% of reads were of a single size, ranging from 20 to 24 nucleotides (nt), were assigned to that specific size category. Any clusters associated with predicted or known miRNAs were assigned as a miRNA cluster. Any other clusters with no dominant

read size, or a dominant read size outside of the 20-24 nt range, were classified as not defined (ND). Further characterisation of all DECs is shown in Figure 4.4e. (d) The distribution of mapped sRNA read lengths for each genotype-treatment combination, aligned to the TAIR10 Arabidopsis genome using ShortStack (Axtell 2013; Shahid and Axtell 2014). Lines represent the average percentage of mapped reads for each size in the 18-30 nt range from three biological replicates, which are depicted as points. E2-treated lines are coloured green, while DMSO-treated lines are coloured grey. (e) Annotation counts of DECs from E2-treated L5 and L7, as compared to DMSO-treated L5 and L7, respectively. DECs were annotated using Araport11 (Cheng et al., 2017), and the TAIR10 TE annotation file (https://www.arabidopsis.org/). See methods for details. Upregulated DECs are represented in red, while downregulated DECs are shown in blue. DECs overlapping TEs are depicted in dark red (upregulated) and dark blue (downregulated). Counts of upregulated and downregulated DECs are shown above and below the y = 0 line, respectively. (f) Distribution of DECs across the five nuclear chromosomes of Arabidopsis for L5 (left) and L7 (right). Chromosomes are depicted as rectangular horizontal boxes, with a scale bar representing 5 mega bases (Mb). Red and blue vertical rectangles superimposed on the chromosomes represent upregulated and downregulated DECs, respectively. Chromosomes are centred around their centromeres, with black lines above the chromosomes marking the pericentromeric regions.

4.4.5 Identification of Novel *Cis*-Regulated Genes by RdDM and ROS1 in the Chromosome Arms

The association between DNA hypomethylation and downregulated DECs in the gene-rich chromosome arms, triggered by ROS1 induction in E2-treated *XVE:ROS1-YFP* lines, holds significant potential to impact gene expression. Indeed, changes in DNA methylation of promoters have been implicated to alter the expression and/or transcriptional responsiveness of nearby genes (Gong et al., 2002; Williams et al., 2015; Gallego-Bartolomé et al., 2018; Kim et al., 2019; Halter et al., 2021). For instance, DNA methylation in the promoter region of some defence-associated genes of Arabidopsis, such as *RLP43*, has been associated with reduced transcriptional activity of the gene (Halter et al., 2021). This mechanism, whereby changes in DNA methylation influence the transcriptional activity of proximal genes, are broadly defined as *cis*-regulatory mechanisms (Chapter 1, Figure 1.4) (Wilkinson et al., 2019; Hannan Parker et al., 2022).

Notably, *ROS1-YFP* induction in L5 and L7 leads to thousands of hypomethylated DMRs in the promoters of protein-coding genes (Figure 4.6e), which coincided with reduced abundance of RdDM-associated 23-24 nt sRNAs (Figure 4.6b), and a general up-regulation of protein-coding genes (Figure 4.6c). Since RdDM-dependent DNA methylation and ROS1-dependent DNA demethylation can negatively and positively affect gene transcription, respectively (Halter et al., 2021), a genome wide scan for novel *cis*-regulated gene targets (Figure 4.6d) was performed. Protein coding genes which contained a hypomethylated DMR or a downregulated 23-24 nt sRNA cluster and concomitantly showed enhanced expression ($log_2FC > 0$, $p_{adj} < 0.05$) in E2-treated *XVE:ROS1-YFP* lines, as compared to DMSO treated

XVE:ROS1-YFP lines (Figure 4.1a) were defined as putative *cis*-regulated targets (Figure 4.6d). For L5 and L7, 229 and 251 putative *cis*-regulated genes were identified, respectively, of which 156 loci were shared between both lines (Figure 4.6e) (Supplementary Data 4.10). Notably, this analysis pipeline identified the *RLP43* gene, which had previously been reported to be *cis*-regulated by ROS1 (Halter et al., 2021), hence confirming the validity of the approach (Supplementary Figure 4.3d).

In addition to transcriptional silencing by RdDM-mediated DNA methylation in the promoter of genes, the presence of 21 nt sRNAs have the potential to stimulate gene expression via association with nuclear-localised AGO1 in the promoter (Figure 1.3c) (Liu et al., 2018; Hannan Parker et al., 2022; Wilkinson et al., 2023). Therefore, up-regulated protein-coding genes ($log_2FC > 0$, $p_{adj} < 0.05$) with an up-regulated 21 nt DEC ($log_2FC > 0$, $p_{adj} < 0.05$) in their promoter were identified. In total, only 2 and 1 gene(s) matched these criteria in L5 (*AT2G27402*, *AT1G66730*) and L7 (*AT1G66730*), respectively. The single common gene, *AT1G66730*, is a DNA ligase involved in seed germination (Waterworth et al., 2010). No DMR was associated with this upregulated 21 nt DEC. Therefore, based on these results, further analyses in this Chapter were focussed on a typical *cis* regulatory model, whereby DNA methylation and sRNAs act to transcriptionally silence genes (Figure 4.6d).

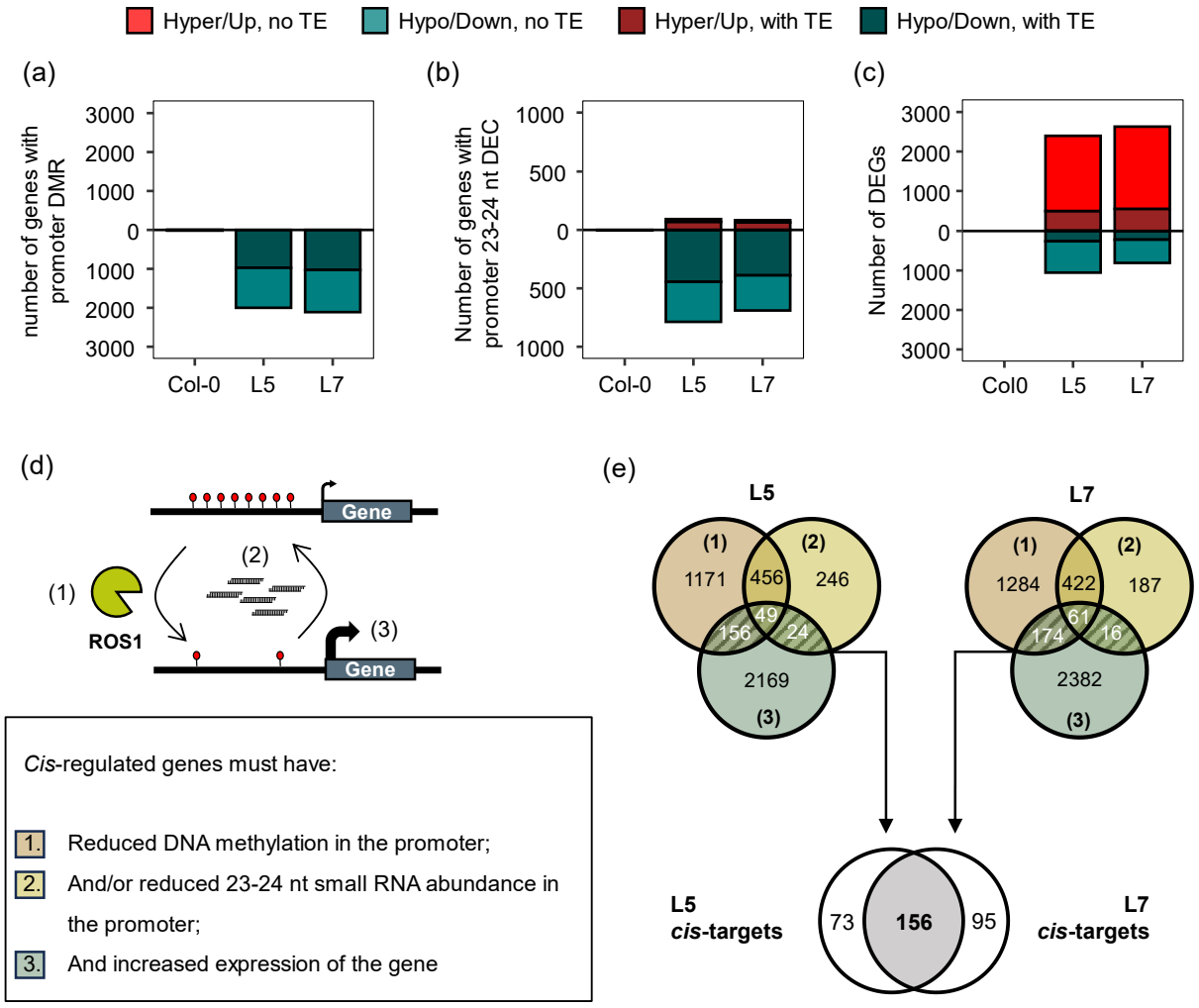


Figure 4.6 Genome wide identification of genes *cis* regulated by RNA directed DNA methylation (RdDM) and ROS1 in the promoter region of protein coding genes using data derived from E2-treated vs. DMSO-treated Col-0, L5, and L7.

(a) Count of differentially methylated regions (DMRs) within promoter regions (1 kb upstream of transcriptional start site). Hyper- and hypomethylated DMR counts are shown above and below the y = 0 line, respectively. Dark colours indicate counts overlapping with transposable elements (TE). (b) Counts of differentially expressed clusters (DECs) within promoter regions, considering only those classed as 23 or 24 nt in size by ShortStack (Axtell, 2013; Shahid and Axtell, 2014). Upregulated and downregulated counts are shown above and below the y = 0 line, respectively. Dark colours indicate counts overlapping with TEs. (c) Counts of protein-coding DEGs. Upregulated and downregulated counts are shown are shown above and below the y = 0 line, respectively. Dark colours indicate genes containing one or more TEs within their promoter. (d) Schematic demonstrating *cis*-regulatory model under consideration. (1) ROS1 removes DNA methylation from the promoter region of a gene, (2) and/or antagonises RdDM in the region driven by 23-24 nt sRNAs, (3) which leads to increased transcriptional activity of the gene. Box defines what is considered a *cis*-regulated gene under this model. (e) Venn diagram showing numbers of *cis*-regulated genes by ROS1 in L5 and L7, as defined in (d).

4.4.6 ROS1 Targets the Promoters of *NPR1* and *TRXh5*, Which is Associated With Transcriptional Activation of NPR1-Mediated Defence Responses

No GO-term enrichment (p_{adj} > 0.05) was found for the 156 consistent *cis*-regulated loci, suggesting no preference for genes related to any biological processes. Therefore, the 156 candidates were individually checked for functions related to plant immunity based on the observation of enhanced resistance to (hemi)biotrophic pathogens (Chapter 3, Figure 3.5) and GO-term enrichments related to immunity in E2-treated *XVE:ROS1-YFP* lines (Figure 4.4). Of note, *NON-EXPRESSOR OF PATHOGENESIS-RELATED GENES 1* (*NPR1*), which is a master regulator of SA-induced defence expression (Zavaliev and Dong, 2024), and *THEOREDOXIN HOMOLOG 5* (*TRXh5*), which promotes the monomerization and subsequent nuclear translocation of NPR1 (Figure 4.7) (Tada et al., 2008; Kneeshaw et al., 2014), were both identified as *cis*-targets in L5 and L7 (Supplementary Figures 4.3b, 4.3c). Using L5 in an independent experiment, the reduction of DNA methylation and increased expression of *NPR1* by ROS1-YFP was validated using McrBC-mediated chop-qPCR and RT-qPCR, respectively (Supplementary Figures 4.5b, 4.5c). Furthermore, using data from Chapter 3, this region is hypermethylated in the ROS1 mutant *ros1-4* mutants and is hypomethylated in the Pol IV-dependent RdDM mutant *nrpe1-11*, further demonstrating that DNA methylation in the promoter of *NPR1* is regulated antagonistically by ROS1 and RdDM (Supplementary Figure 4.5d).

Interestingly, many well characterised downstream targets of NPR1 were upregulated in E2-treated *XVE:ROS1-YFP* lines including WRKY DNA-BINDING PROTEIN (WRKY) transcription factors (TFs) (Wang et al., 2006), *PATHOGENESIS RELATED* (*PR*) genes (Cao et al., 1994; Zhou et al., 2000; Li et al., 2018), and *SAR DEFICIENT 1* (*SARD1*) (Ding et al., 2018; Nomoto et al., 2021; Yun et al., 2024), which regulates the biosynthesis of SA and the systemic acquired resistance (SAR) signalling molecule N-hydroxy-pipecolic acid (NHP) (Sun et al., 2015; Hartmann et al., 2018) (Figure 4.6). NPR1 physically interacts with TGACG MOTIF-BINDING FACTOR (TGA) TFs to promote the expression of genes (Kumar et al., 2022; Zavaliev and Dong, 2024). While no TGA TFs were significantly upregulated in E2-treated L5 or L7 (Figure 4.7) (Supplementary Data 4.4), the promoters of protein coding genes upregulated in E2-treated *XVE:ROS1-YFP* lines (Figure 4.4a) were significantly enriched for the binding motifs of 7 TGA TFs (p_{adj} < 0.05), as identified by the PlantRegMap TF enrichment tool (Tian et al., 2020) (Supplementary Data 4.11). These include the TGA TFs 2 through to 7, which have all been shown to interact with NPR1 to transcriptionally coactivate SA-associated defence genes (Zhou et al., 2000; Fan and Dong, 2002; Jin et al., 2018; Budimir et al., 2021; Kumar et al., 2022; Han et al., 2022; Yildiz et al., 2023). Furthermore, TGAs 2-7 are highly sensitive to DNA methylation, exhibiting an average of 9.8 times less binding affinity to

methylated DNA compared to unmethylated DNA (O'Malley et al., 2016). Therefore, ROS1 DNA demethylation may facilitate SA-dependent defences threefold: (i) promoting the expression of *NPR1* in *cis* (Supplementary Figures 4.3c, 4.5), (ii) promoting the expression of *TRXh5* in *cis* (Supplementary Figure 4.3b), which facilitates the nuclear translocation of NPR1 and subsequent interaction with TGA TFs (Tada et al., 2008; Kneeshaw et al., 2014), and (iii) removing DNA methylation from other regions of the genome to enhance DNA binding of TGA TFs that interact with NPR1 (O'Malley et al., 2016) (Figure 4.7).

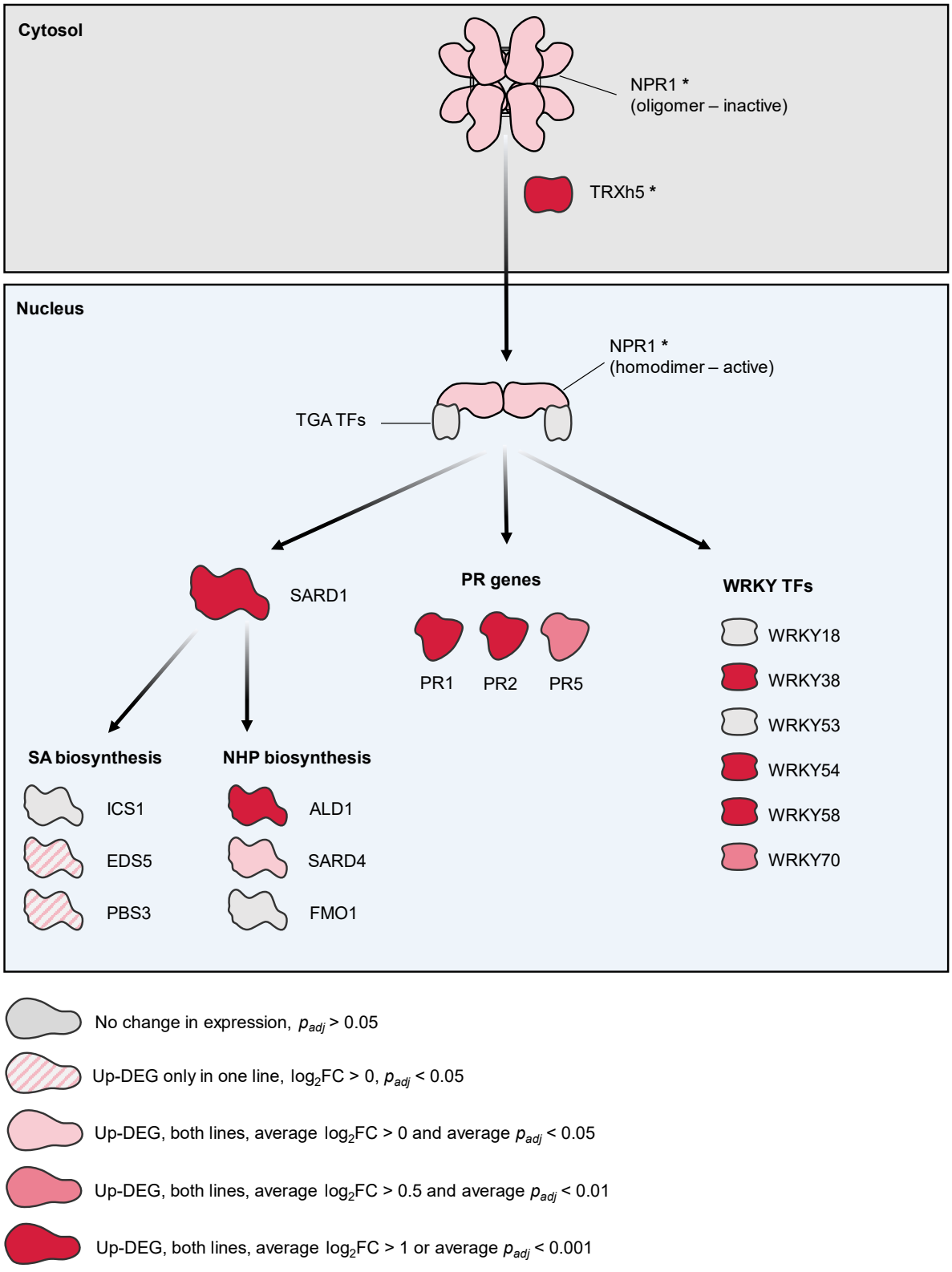


Figure 4.7 Simplified schematic of NPR1 signalling pathway, partially adapted from Zavaliev and Dong (2024).
Figure is showing colour/pattern-coded differential expression between E2-treated and DMSO-treated L5 and L7 seedlings. *NPR1* and *TRXh5* (marked with *) were identified as *cis*-regulated genes (Figure 4.6; Supplementary Figure 4.5). For details about genes and pathways, see main text. Raw $\log_2$ fold change and p_{adj} values for each component are listed in Supplementary Data 4.12.

4.4.7 ROS1 Induction Activates the *ATR*-Dependent DNA Damage Response Pathway

Upregulated genes in E2-treated *XVE:ROS1-YFP* lines were highly enriched for the GO term “DNA damage response” (Figure 4.4). Interestingly, two components of the DNA damage response were amongst the 156 *cis*-regulated targets (Supplementary Data 4.10), namely *ATAXIA TELANGIECTASIA AND RAD3-RELATED* (*ATR*) and the transcription factor *E2FC* (Figure 4.8). ATR is typically activated in response to single-stranded breaks (SSBs) and/or replicative stress and subsequently phosphorylates the transcription factor SUPPRESSOR OF GAMMA RADIATION 1 (SOG1), which was also upregulated in E2-treated *XVE:ROS1-YFP* lines (Figure 4.9) (Waterworth et al., 2011; Sjogren et al., 2015; Herbst et al., 2024). SOG1 directly regulates the transcription of a plethora of other genes with roles in DNA repair, defence, and transcriptional regulation (Bourbousse et al., 2018; Herbst et al., 2024) (Figure 4.8). One direct target of SOG1, with critical roles in DNA damage-induced growth repression is *WEE1*, which was also upregulated in E2-treated L5 and L7 (De Schutter et al., 2007; Bourbousse et al., 2018; Herbst et al., 2024). DNA damage caused by double stranded DNA breaks (DSBs) are recognised and repaired via the ATAXIA TELANGIECTASIA MUTATED (ATM) dependent pathway (Garcia et al., 2003; Waterworth et al., 2011; Herbst et al., 2024). However, no genes in this pathway were consistently upregulated in both E2-treated *XVE:ROS1-YFP* lines (Figure 4.8), which is consistent with the ROS1 mode of action inducing SSBs (Gong et al., 2002). E2FC forms part of a larger protein complex known as the DREAM complex, which negatively regulates cell cycle progression in response to DNA damage, thereby reducing further SSB DNA damage associated with DNA replication (Lang et al., 2021). Furthermore, the DREAM complex has important roles in regulating genome-wide levels of DNA methylation by negatively regulating the expression of *ROS1* and other genes associated in the regulation of DNA methylation including *MET1*, *CMT3*, *DDM1*, and *KYP* (Ning et al., 2020). Thus, the observed impact of ROS1 induction on the ATR-dependent DNA damage pathway likely represents a protective mechanism, whereby *cis*-regulated induction of the *ATR* and *E2FC* genes rapidly activates *ATR* and *E2FC* to increase capacity for SSB DNA repair and thus, reduce ROS1-dependent DNA damage (Figure 4.5).

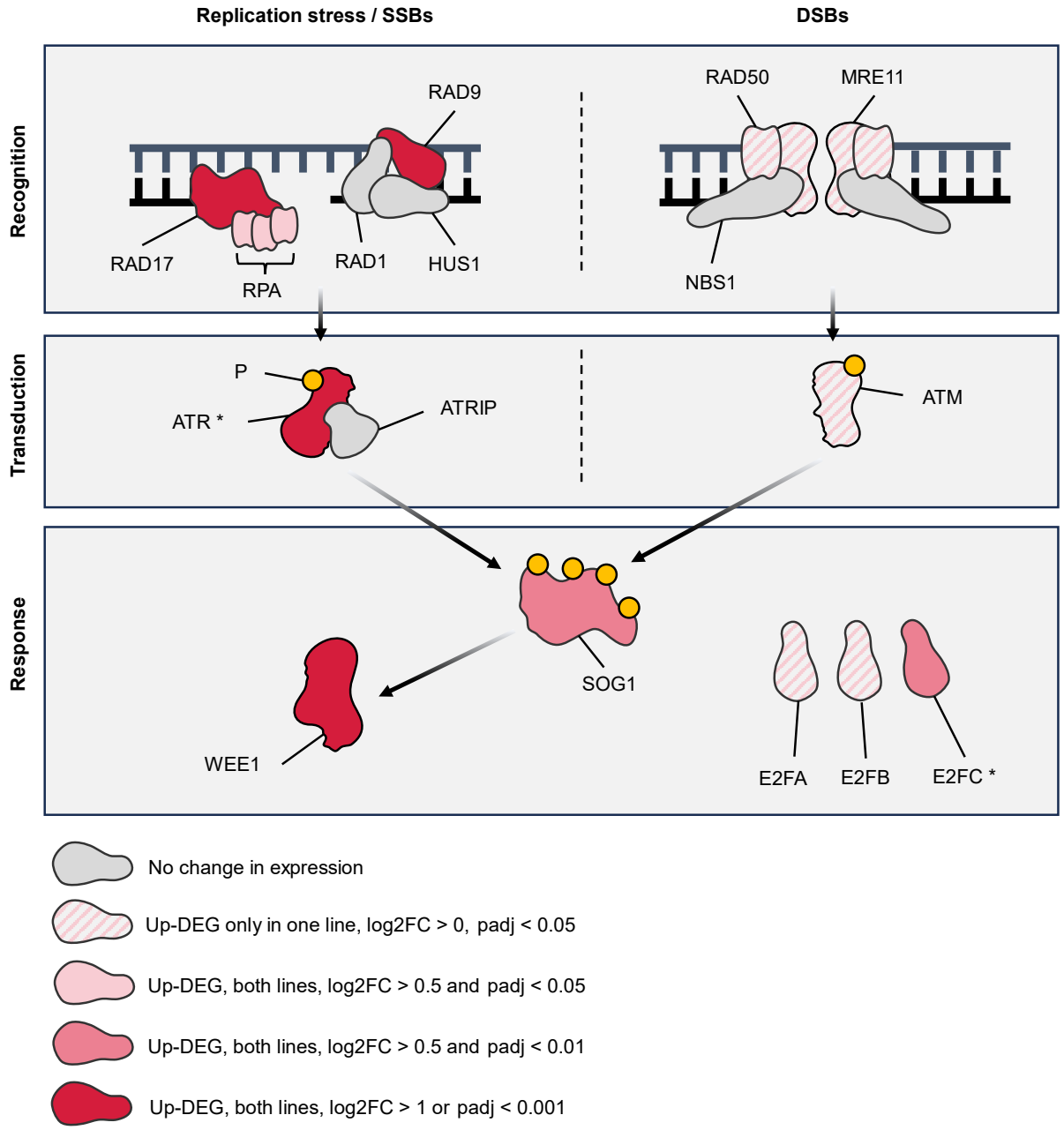


Figure 4.8 Simplified schematic of the DNA damage response, adapted from Herbst et al. (2024). Figure is showing colour/pattern-coded differential expression between E2-treated and DMSO-treated L5 and L7 seedlings (Figure 4.2a). *ATR* and *E2FC* (marked with *) were identified as *cis*-regulated genes by ROS1 (Figure 4.6). Single stranded breaks (SSBs) (left) and double stranded breaks (DSBs) (right) are detected and transduced by different proteins, as demonstrated by dashed line. Raw $\log_2$ fold change and p_{adj} values for each component are listed in Supplementary Data 4.12.

4.5 Discussion

4.5.1 Characterising ROS1 Targets Using an Inducible Construct

Over the last two decades, numerous studies have demonstrated the importance of ROS1 in regulating transcriptional responses to environmental stress in Arabidopsis (Gong et al., 2002; Yu et al., 2013; López Sánchez et al., 2016; Liu et al., 2021; Halter et al., 2021; Yang et al., 2022; Wilkinson et al., 2023). However, all reports investigating the link between ROS1-dependent DNA demethylation and plant stress responses have relied on the use of T-DNA insertion lines that knockout or knockdown *ROS1* expression and thus generated DNA hypermethylated plants. Whilst this has provided the community with very valuable insights into the predicted targeting of *ROS1* (Tang et al., 2016; Halter et al., 2021), it fails to resolve some targets of ROS1. For instance, loci with maximal levels of DNA methylation in wild-type Col-0 plants, that are still targeted by ROS1 for DNA demethylation under certain conditions, would not be identified as hypermethylated regions in *ros1* mutants, and thus, would not be identified as ROS1 targets. Furthermore, the permanent nature of the mutations poorly reflects the responsive and reversible nature of epigenetic marks and may lead to the activation of other compensatory epigenetic mechanisms, as discussed in Chapter 3. In this Chapter, the use of the *XVE:ROS1-YFP* system has enabled detailed characterisation of genome-wide changes in cytosine DNA methylation, small RNA abundance, and transcriptomic changes in response to transient increases in ROS1 expression. Furthermore, as the ROS1 induction course was effective in activating immunity against (hemi-)biotrophic pathogens (Chapter 3; Figure 3.5), the data are directly relevant to the onset of ROS1-dependent immune memory in Arabidopsis.

4.5.2 ROS1-*Cis*-Regulated Genes

As reported before, removal of DNA methylation by ROS1 is enriched at TEs located near protein-coding genes (Figure 4.3e) (Tang et al., 2016). However, only 156 of the 1,846 genes consistently upregulated in E2-treated *XVE:ROS1-YFP* lines exhibited hypomethylated DMRs or reduced sRNA abundance within their promoters (Figure 4.6). This suggests that the vast majority of these genes are not directly *cis*-regulated by ROS1. However, of the 156 *cis*-regulated genes, many had roles in plant immune responses and DNA damage responses (Supplementary Data 4.10). Interestingly, two core components of the SA signalling pathway, namely *NPR1* and *TRXh5* (Kneeshaw et al., 2014; Zavaliev and Dong, 2024; Spoel and Dong, 2024), were both identified as *cis*-regulated targets. NPR1 activates a cascade of transcriptional responses related to defence as it promotes the expression of many transcription factors, which have multiple gene targets themselves (Yun et al. 2024; Wang,

Amornsiripanitch, and Dong 2006). Overexpression of Arabidopsis' *NPR1* has been shown to promote resistance to pathogens in many plant species (Zavaliev and Dong, 2024), even with relative increases in protein levels as low as 1.4x (Cao, Li, and Dong 1998), demonstrating its potent effect on plant immune responses. Therefore, the observed ~1.3x increase in NPR1 expression following ROS1 induction in E2-treated *XVE:ROS1-YFP* lines (Supplementary Data 4.12; Supplementary Figure 4.5; Figure 4.7) could certainly still influence plant immune responses to (hemi)biotrophic pathogens. Furthermore, in order to translocate into the nucleus and coactivate the expression of genes with TGA TFs, NPR1 must first be reduced from its oligomer form, a process that involves TRXh5-mediated denitrosylation (Tada et al., 2008; Kneeshaw et al., 2014). Thus, TRXh5 also has crucial roles in promoting plant immunity against pathogens in Arabidopsis and is upregulated in response to salicylic acid, pathogen infection, and other environmental stresses (Laloi et al., 2004; Tada et al., 2008; Kneeshaw et al., 2014; Arnaiz et al., 2023). Interestingly, *TRXh5* had an approximately 2 fold increase in expression (Supplementary Data 4.12; Figure 4.7) in E2-treated *XVE:ROS1-YFP* lines and was also identified as a *cis*-regulated gene. Finally, while TGA transcription factors (TFs), which are essential for NPR1-mediated activation of defence genes were not upregulated in E2-treated *XVE:ROS1-YFP* lines, upregulated protein-coding genes in E2-treated *XVE:ROS1-YFP* lines were enriched for 7 out of 10 TGA TF binding sites (Jakoby et al., 2002) (Supplementary Data 4.11). These include all TGAs known to constitutively interact with NPR1 (TGA2, TGA3, TGA5, TGA6, and TGA7) (Després et al., 2000; Zhou et al., 2000; Fan and Dong, 2002), as well as TGA4, which only interacts with NPR1 in SA-induced leaves (Després et al., 2003; Kesarwani et al., 2007; Budimir et al., 2021). Since the association between NPR1 and TGA TFs is essential for NPR1-mediated activation of SA defences and systemic acquired resistance (SAR) (Zhang et al., 2003; Yildiz et al., 2023), this enrichment further highlights a central role of NPR1 in the transcriptomic response of E2-treated *XVE:ROS1-YFP* lines. Therefore, while most upregulated genes in E2-treated *XVE:ROS1-YFP* were not identified as *cis*-regulated loci, the significant association with biological processes such as "response to salicylic acid", "systemic acquired resistance", and "response to molecule of bacterial origin" (Figure 4.3c), could be a result of a transcriptional cascade following *cis*-regulation of a few key activators such as *NPR1* and *TRXh5* (Figure 4.7).

4.5.3 Relationship Between ROS1 and the DNA Damage Response

As well as plant immune responses, many genes related to the DNA damage response were upregulated following ROS1 induction in *XVE:ROS1-YFP* lines (Figure 4.4c). ROS1 inherently causes single stranded breaks as it removes DNA methylation via a base-excision repair pathway (Gong et al., 2002; Agius et al., 2006; Ponferrada-Marín et al., 2010; Lee et

al., 2014; Zhang et al., 2022; Du et al., 2023). Many of the genes encoding proteins involved in repairing these single stranded breaks, were upregulated in E2-treated *XVE:ROS1-YFP* lines (Figure 4.8). Strikingly, *ATR*, which encodes a kinase that activates the master transcription factor SOG1 in response to DNA damage (Herbst et al., 2024), was identified as a *cis*-regulated gene (Figure 4.8; Supplementary Data 4.10). This hints to a protective mechanism whereby increased ROS1 activity boosts the expression of *ATR* by removing DNA methylation from its promoter, subsequently enhancing the DNA damage capacity required for repairing single-stranded breaks caused by ROS1. Future studies should establish whether increased expression of *ATR* would enhance tolerance to genotoxic stress, by running bioassays with chemicals that induce single-stranded breaks (e.g., Zeocin) (Povirk et al., 1977; Vladejić et al., 2024) in *ATR* overexpressing lines.

4.5.4 ROS1-Driven DNA Damage and its Potential Role in Plant Immunity

Activation of the DNA damage response following active DNA demethylation by ROS1 may also have a secondary role in promoting plant immune responses. Pathogenic infections can induce DNA damage in Arabidopsis (Lucht et al., 2002; Song and Bent, 2014), and the DNA damage response pathway is known to transduce and amplify SA-dependent defence pathways (Durrant et al., 2007; Wang et al., 2010; Yan et al., 2013; Vega-Muñoz et al., 2023; Jeong et al., 2023). To distinguish transcriptomic changes caused by DNA hypomethylation from those triggered by the DNA damage response, transcriptomic changes caused by ROS1-dependent DNA hypomethylation were compared with those induced by TET1-dependent DNA hypomethylation (Ji et al., 2018) (Figure 4.4g). None of the genes consistently upregulated in *XVE:ROS1-YFP* lines, associated with ATR-mediated repair of single stranded breaks - such as *RAD17*, *RAD9*, *RPA*, *ATR*, *SOG1*, and *WEE1* (Figure 4.8) - were upregulated in *35S:TET1* lines (Supplementary Data 4.4). In contrast, every gene consistently upregulated in *XVE:ROS1-YFP* lines in the NPR1-dependent signalling pathway (marked in Figure 4.7) was also upregulated in *35S:TET1* lines, with the exception of *ALD1* (Supplementary Data 4.4). This aligns with the observation that the most enriched GO terms for upregulated genes in both *35S:TET1* lines and E2-treated *XVE:ROS1-YFP* lines were associated with SA-dependent immunity (Figure 4.4g). Meanwhile, upregulated genes unique to E2-treated *XVE:ROS1-YFP* lines were enriched with DNA damage response pathways. Therefore, whilst DNA damage responses may contribute to the enhanced defence phenotype of E2-treated *XVE:ROS1-YFP* plants, DNA (de)methylation alone is a crucial regulator of defence responses against pathogens. This is in line with previous findings that hypomethylated Arabidopsis mutants tend to be more resistant to (hemi)biotrophic pathogens, whereas

hypermethylated mutants, such as *ros1*, tend to be more susceptible (Luna et al., 2012; Yu et al., 2013; López Sánchez et al., 2016; Furci et al., 2019; Halter et al., 2021).

4.5.5 ROS1-Mediated DNA Demethylation and its Impact on *IBM1* Splicing and Epigenetic Regulation

Few genes were found to be alternatively spliced following ROS1-dependent DNA demethylation. Of these few genes, the H3K9 demethylase *IBM1* was identified as a locus that had differential transcript usage in E2-treated *XVE:ROS1-YFP* lines. E2-treated L5 and L7 had reduced methylation in the 7th intron of *IBM1* and increased production of the short prematurely polyadenylated transcripts (*IBM1-S*) (Rigal et al., 2012; Lei et al., 2014). This methylated region in *IBM1* has been proposed to act as a sensing locus for epigenetic marks, as occurs in the promoter region of *ROS1* (discussed in Chapter 1; Figure 1.2) (Yinwen Zhang et al., 2024). Under basal conditions, the 7th intron of *IBM1* contains DNA methylation in all contexts. DNA methylation in CHG contexts promotes binding of KRYPTONITE (KYP), which deposits H3K9me2 thus creating a heterochromatic environment within the intron (Chapter 1; Figure 1.2). This promotes the maintenance of DNA methylation in the region as the CHG DNA methyltransferase CHROMOMETHYLASE 3 (CMT3) is recruited by H3K9me2 in plants (Fang et al., 2022). The resultant heterochromatic environment is recognised by the INCREASE IN BONSAI METHYLATION2/ASI1-IMMUNOPRECIPITATED PROTEIN1/ENHANCED DOWNY MILDEW2 (IBM2/AIPP1/EDM2) protein complex, which promotes distal polyadenylation and production of functional *IBM1-L* transcripts. As a result, the functional IBM1 enzyme antagonises the silencing H3K9me2 mark and CHG formation in gene-rich euchromatic regions (Rigal et al. 2012; Saze et al. 2013; Wang et al. 2013; Saze et al. 2008; Miura et al. 2009), thus acting in a negative feedback loop as a heterochromatin-sensing locus to ensure epigenome homeostasis (Yinwen Zhang et al., 2024) (Chapter 1; Figure 1.2). The activity of this feedback loop is exemplified by the fact that hypomethylated mutants, such as *cmt3, kyp,* and *met1* mutants, have reduced production of *IBM1-L* transcripts and ectopic accumulation of CHG DNA methylation, particularly in chromosome arms (Rigal et al., 2012). Furthermore, intronic methylation in *IBM1* is prevalent in orthologous genes in flowering plants, despite the underlying DNA sequence exhibiting dynamic evolution, and there appears to be an evolutionary link between natural reductions/losses in *CMT3* expression and functional *IBM1* expression (Yinwen Zhang et al., 2024). The data presented in this Chapter demonstrate that this 7th intron is a target of active DNA demethylation by ROS1 (Figure 4.4f; Supplementary Figure 4.3a), indicating that this sensing locus also detects ROS1-dependent DNA demethylation to ensure epigenetic homeostasis in Arabidopsis. However, no ectopic accumulation of CHG DNA methylation was observed along the chromosome arms or in gene

bodies of E2-treated *XVE:ROS1-YFP* lines (Figure 4.2). Instead, ectopic accumulation of CHG DNA methylation was observed in pericentromeric regions (Figure 4.3a), which is not a canonical target of *IBM1*-mediated demethylation of H3K9me2 (Rigal et al., 2012). The lack of CHG hypermethylation in chromosome arms indicates that ROS1-mediated DNA methylation counteracts the accumulation of repressive marks in *IBM1-S*-expessing plants. Thus, ROS1 and IBM1 may form a larger epigenetic homeostatic pathway that enables plants to self-regulate DNA methylation and H3K9me2 across the genome. Chromatin immunoprecipitation (ChIP) sequencing studies need to be performed to confirm ROS1 localisation at the 7th intron of *IBM1* and subsequently, the genomic distribution of IBM1 under different environmental stress conditions.

4.5.6 ROS1 Activity and its Differential Impact on Chromosome Arms Versus (Peri)centromeric Regions.

An interesting observation from the data presented above is that ROS1-dependent loss of DNA methylation in E2-treated *XVE:ROS1-YFP* lines was more pronounced in the chromosome arms compared to the pericentromeric regions (Figure 4.3a). Notably, hypomethylated DMRs in E2-treated L5 and L7 were significantly underrepresented in intergenic TE regions (Figure 4.3e), which are predominantly located within pericentromeric regions. A simple explanation for this could be that ROS1-YFP is not targeted and/or cannot access pericentromeric regions for active DNA demethylation. However, the specific regions of the genome accessible to ROS1 remain unclear due to the lack of direct mechanistic evidence, such as high-throughput sequencing data of ROS1 DNA binding sites (e.g., ChIP-seq). Current understanding is instead based on previous evidence of epigenetic marks associated with ROS1-dependent DNA demethylation, mostly using mutant lines. For instance, Tang et al (2016) found that the hypermethylated DMRs in *ros1* mutant lines were enriched with the histone marks H3K18ac and H3K27me3 and depleted of H3K27me1 and H3K9me2. H3K27me1 and H3K9me2 are known to be enriched in pericentromeric regions (Sequeira-Mendes et al., 2014) and might therefore repel ROS1. However, *in vitro* binding of ROS1 to H3 was unaffected by the methylation status of H3K (Parrilla-Doblas et al., 2022). Thus, it is unlikely that these marks explain the contrasting effects of ROS1-YFP between chromosome arms and pericentromeric regions.

Furthermore, under basal conditions, some ROS1-targeted loci are regulated by a protein complex containing INCREASE IN DNA METHYLATION 1 (IDM1) and the methyl-CpG-binding protein MBD7 (Nie et al. 2019; Zhang, Gong, and Zhu 2022). IDM1 and MBD7 both have methyl binding domains that preferentially bind to highly methylated, CG-dense regions of the genome (Zhang, Gong, and Zhu 2022; Lang et al. 2015). IDM1 catalyses the

acetylation of H3 residues K14, K18, and K23, which promotes H2A.Z deposition and subsequently recruits ROS1 for DNA demethylation (Nie et al. 2019; Lang et al. 2015; Qian et al. 2012; Li et al. 2015; Qian et al. 2014). This explains why H3K18ac marks are enriched at *ros1* hypermethylated DMRs (Tang et al., 2016). Therefore, pericentromeric regions may lack permissive marks required for ROS1 activity, rather than contain antagonistic marks that repel ROS1. Supporting this, it was discovered that the histone deacetylase HDA6 maintains heterochromatin in pericentromeric regions by preventing IDM1-dependent accumulation of H3K18ac, thus inhibiting active DNA demethylation by ROS1 and related DNA demethylases (Wang et al. 2021). However, IDM1 regulates only about one-third of ROS1 targets (Nie et al. 2019; Zhang, Gong, and Zhu 2022) and ROS1 has been shown to effectively remove DNA methylation at sites lacking H3K18ac and H2A.Z entirely (Liu et al. 2021). In these cases, ROS1-dependent DNA demethylation is achieved through association via a different protein complex that includes ROS1-ASSOCIATED WD40 DOMAIN-CONTAINING PROTEIN (RWD40), the DNA methylation binding protein ROS1-ASSOCIATED METHYL-DNA BINDING PROTEIN 1 (RMB1), and ROS1-ASSOCIATED HOMEODOMAIN PROTEIN 1 (RHD1) (Liu et al. 2021). While the genome-wide binding of the ROS1-RWD40-RMB1-RHD1 complex has not been fully characterized, it has been shown to target distinct loci including the promoter of *ROS1* itself (Liu et al. 2021). Interestingly, RMB1 can bind to DNA methylation in all cytosine contexts, differing from the CG-dense preference of MBD9 in the IDM1-containing complex (Liu et al. 2021). Thus, ROS1-dependent DNA demethylation could occur in pericentromeric regions through IDM1-independent mechanisms, potentially mediated by alternative protein interactions, such as the RWD40-containing complex. Therefore, ROS1 access to the pericentromeric regions cannot be ruled out. Instead, the reduced level of DNA demethylation in the pericentromeric regions of E2-treated *XVE:ROS1-YFP* lines (Figure 4.2a) may reflect a rapid remethylation response in these regions that does not occur in chromosome arms.

4.5.7 ROS1 Activity and its Influence on Pol IV-Dependent RdDM

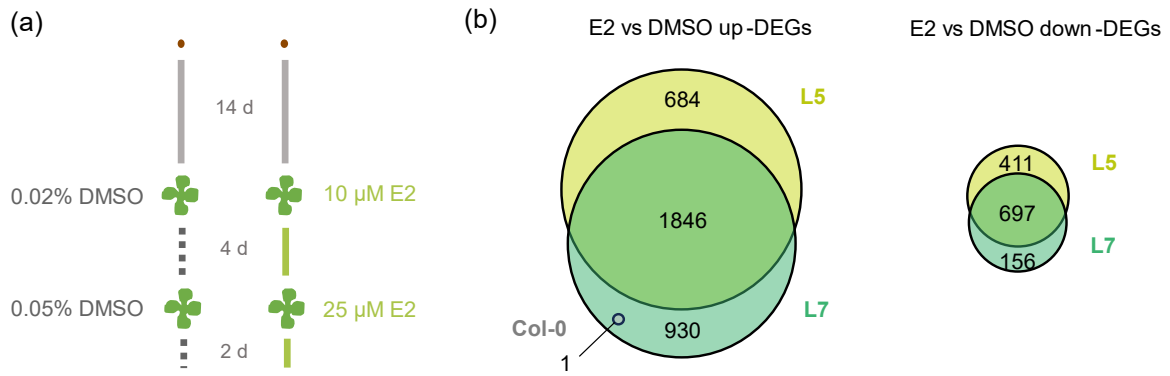
Interestingly, downregulated sRNA clusters primarily occurred along chromosome arms, whilst upregulated sRNA clusters were mainly found within pericentromeric regions (Figure 4.4e). Furthermore, over 90% of all DECs were 23-24 nt in size (Figure 4.5c), which is characteristic of Pol IV-dependent RdDM (Matzke and Mosher, 2014; Cuerda-Gil and Slotkin, 2016; Erdmann and Picard, 2020). Therefore, ROS1 appears to have a significant influence on the distribution and activity of Pol IV-dependent RdDM. The Pol IV complex is recruited by the H3K9me reader SAWADEE HOMEODOMAIN HOMOLOG 1 (SHH1) at approximately 50% of Pol IV-associated 24 nt sRNA sites through its association with CLASSY1 (CLSY1) and CLSY2 (Law et al. 2013; Matzke and Mosher 2014; Cuerda-Gil and Slotkin 2016; Zhou,

Palanca, and Law 2018). Since H3K9me marks are mechanistically linked to DNA methylation (Soppe et al. 2002; Law et al. 2013; Johnson, Cao, and Jacobsen 2002; Fang et al. 2022), the removal of DNA methylation by ROS1 at these sites would likely lead to a passive loss of H3K9me. Furthermore, Pol V, which produces scaffold transcripts to associate with Pol IV-derived 23-24 nt sRNAs, is continually recruited to sites with DNA methylation by association with the proteins SUVH2 and SUVH9 (Liu et al., 2014; Johnson et al., 2014). Therefore, ROS1-dependent DNA demethylation is expected to result in fewer viable sites for Pol IV recruitment, transcription, and subsequent 23-24 nt sRNA production. Indeed, on average, 68% of the downregulated 23-24 nt DECs overlapped with a hypomethylated DMR in E2-treated *XVE:ROS1-YFP* lines. Notably, there were no changes in the expression of the largest subunits of Pol IV (*NRPD1*), Pol V (*NRPE1*), or genes that guide Pol IV-dependent RdDM (*SHH1*, *CLSY1*, and *CLSY2*) (Supplementary Data 4.4). Therefore, the capacity for Pol IV-dependent RdDM is the same in DMSO- and E2-treated *XVE:ROS1-YFP* lines, but the number of viable sites in chromosome arms are significantly reduced in E2-treated *XVE:ROS1-YFP* lines as ROS1 preferentially targets gene-TE boundaries (Tang et al., 2016) (Figure 4.3e). As a result, the proteins associated with Pol IV-dependent RdDM would likely be forced to shift to alternative, non-canonical sites, such as those in the pericentromeric regions which have an abundance of methylated DNA. This, in turn, would drive excessive RdDM in pericentromeric regions, resulting in the hyperaccumulation of DNA in these regions, as observed in Figure 4.3a. Future analyses should more thoroughly characterise the subcontext in which cytosines in the pericentromeric regions become hypermethylated after ROS1-YFP induction. If there are no obvious biases, it would provide an early indication that this indeed is occurring via the RdDM pathway rather than by CMT2/3 methyltransferases (Gouil and Baulcombe, 2016). Interestingly, the gene encoding CLSY3, which facilitates pericentromeric Pol IV recruitment independently of SHH1 and H3K9me (Zhou, Palanca, and Law 2018), was also strongly upregulated in E2-treated *XVE:ROS1-YFP* lines, with an average 7.3-fold increase in expression relative to DMSO-treated seedlings (Supplementary Data 4.4). Therefore, in addition to the passive redistribution of Pol IV-dependent RdDM driven by changes in DNA methylation, the loss of DNA methylation by ROS1 may promote the expression of genes, like *CLSY3*, that actively recruit Pol IV-dependent RdDM to TE-rich pericentromeric regions.

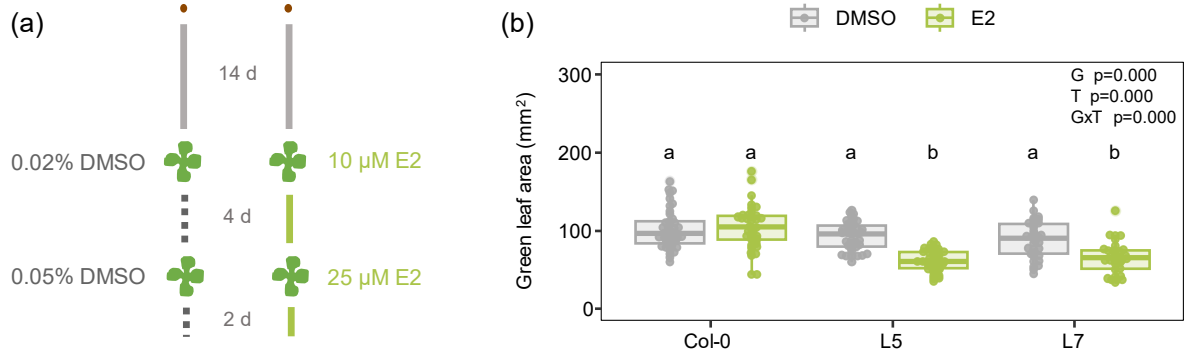
Therefore, the lack of DNA demethylation in pericentromeric regions in E2-treated *XVE:ROS1-YFP* lines may not result from reduced ROS1-YFP accessibility, but rather from the rapid re-methylation of these TE-rich regions by RdDM following ROS1-induced DNA demethylation. It is plausible that this mechanism serves to safeguard genome stability under severe epigenetic shock, by preventing TE mobilisation and ectopic recombination events in

(peri)centromeric regions (Hannan Parker, Wilkinson, and Ton 2022; Underwood et al. 2018; Wilkinson et al. 2019). The potential biological implications of this mechanism, particularly in relation to long-term immune memory to pathogens, are explored in Chapter 5 of this thesis.

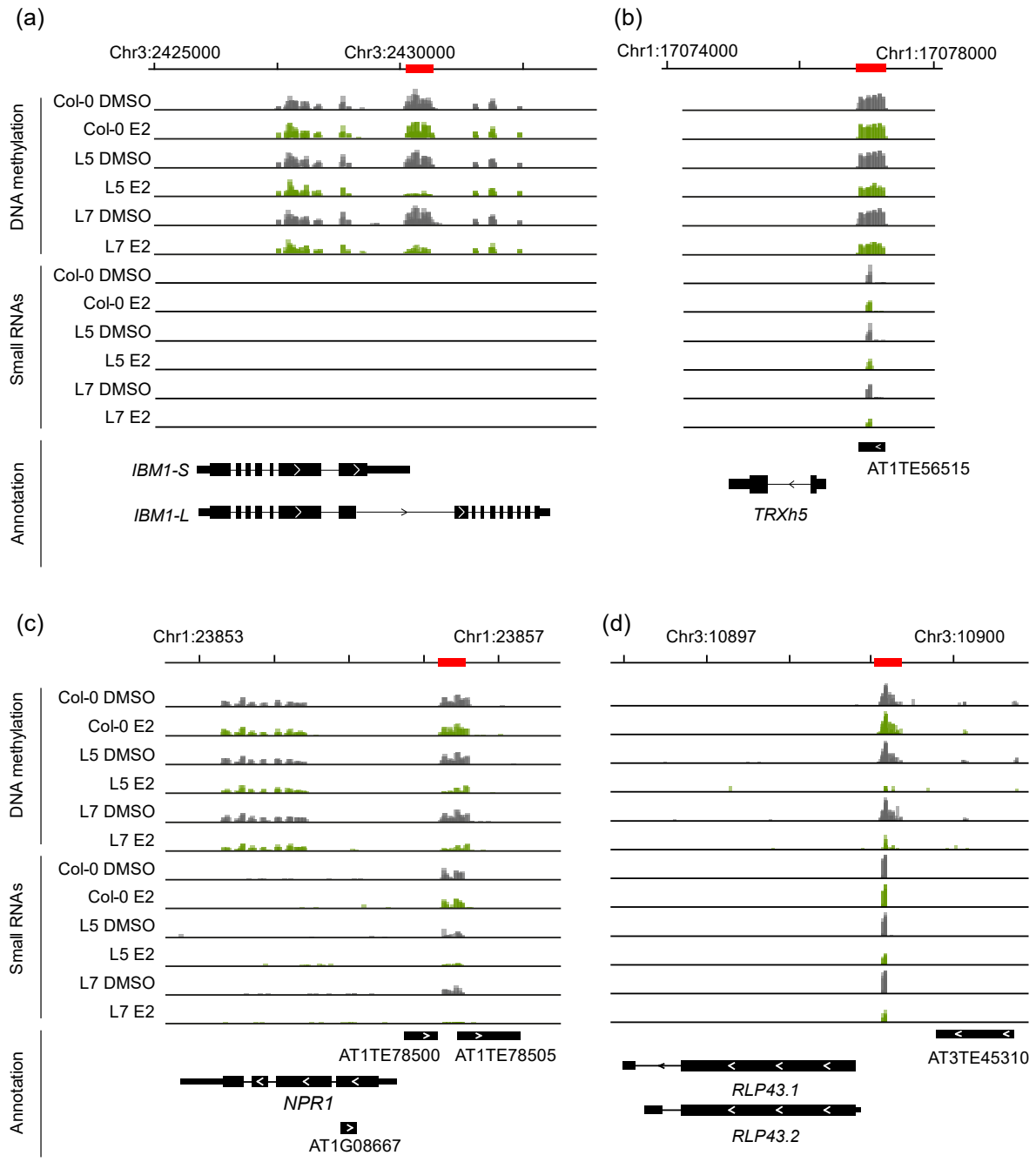
4.6 Supplementary Figures



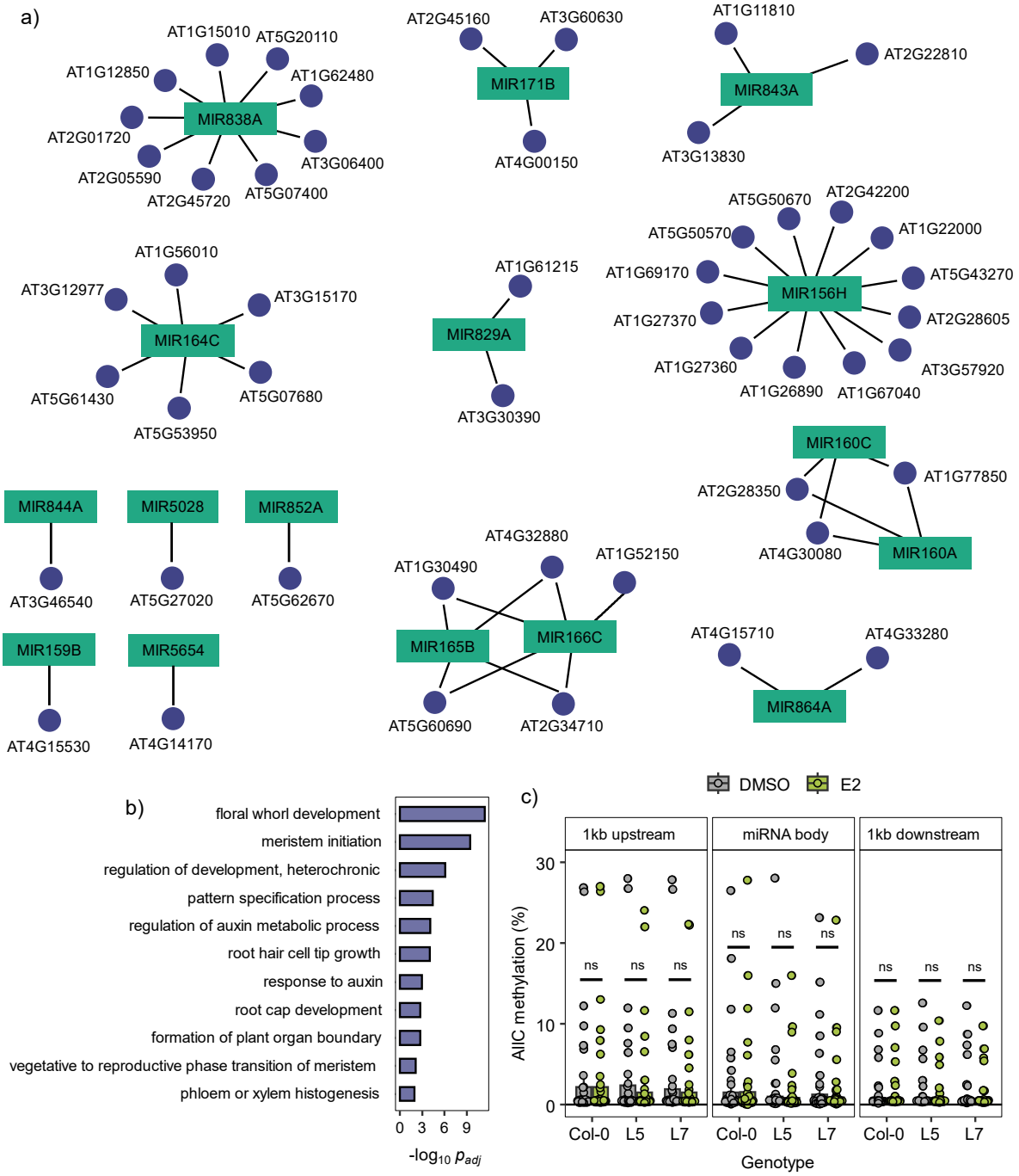
Supplementary Figure 4.1 Overlaps of differentially expressed loci/genes (DEGs) from whole transcriptome sequencing.
(a) Treatment course used to treat Col-0, L5, and L7 seedlings, as in Figure 4.2. (b) Venn diagram overlapping significantly upregulated DEGs (fold change > 0; p_{adj} < 0.05) (left) and significantly downregulated DEGs (fold change < 0; p_{adj} < 0.05) (right) in E2-treated Col-0, L5, and L7, as compared to DMSO controls.



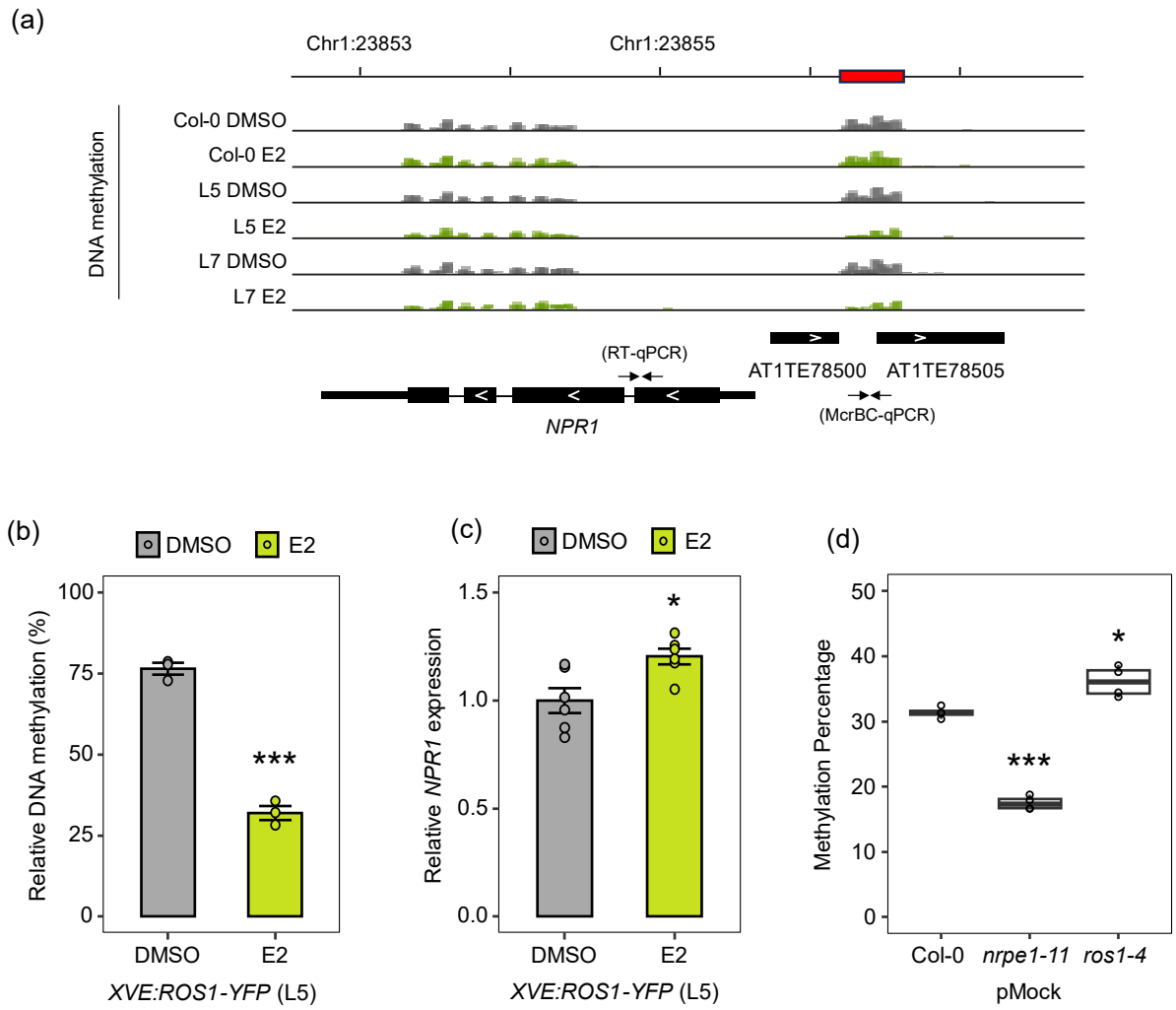
Supplementary Figure 4.2 Green leaf area (GLA) of Col-0 and *XVE:ROS1-YFP* lines (L5 and L7) after treating twice E2 or DMSO.
(a) Treatment course used to treat Col-0, L5, and L7 seedlings, as in Figure 4.2. (b) GLA of individual plants (32-42) calculated from the *Pst* assay in Chapter 3 (Figure 3.5). The effect of genotype and treatment, and the interaction between the two, on log transformed GLA was tested using a two-way ANOVA. Following identification of significant interaction effect, Tukey post-hoc analysis was performed. Treatments with the same letter are not significantly different (p_{adj} > 0.05).



Supplementary Figure 4.3 Genome browser snapshots of *cis*-regulated loci.
Genome browser snapshots showing DNA methylation (scale = 0-100%) and small RNAs (scale variable, read counts) in DMSO- and E2-treated Col-0, L5, and L7. Annotations are from Araport11 and TAIR10 TE annotation files. Each track is an overlay of all three biological replicates. Figures created from Integrative Genomics Viewer (IGV) (Thorvaldsdóttir et al., 2013). Red bars highlight DMRs and/or DECs. (a) *IBM1* intronic DMR identified E2-treated *XVE:ROS1-YFP* lines (L5, L7). Small RNA count limits = 0-10. (b) *TRXh1* promoter DMR and DEC overlapping a TE in E2-treated *XVE:ROS1-YFP* lines (L5, L7). Small RNA count limits = 0-300. (c) *NPR1* promoter DMR and DEC overlapping a TE in E2-treated *XVE:ROS1-YFP* lines (L5, L7). Small RNA count limits = 0-3. (d) *RLP43* promoter DMR and DEC in E2-treated *XVE:ROS1-YFP* lines (L5, L7). Small RNA count limits = 0-125.



Supplementary Figure 4.4 Characterisation of the protein-coding gene targets of upregulated miRNAs in E2-treated *XVE:ROS1-YFP* lines (L5 and L7).
(a) Network plot showing each upregulated miRNA identified in L5 or L7 (green box) and all their predicted protein-coding gene targets (blue dots). Targets were predicted using targetFinder v1.7 (Fahlgren et al., 2007; Srivastava et al., 2014). (b) GO enrichment for biological process of all predicted protein-coding targets identified in (a). (c) Average DNA methylation (allC) 1 kb upstream, 1 kb downstream, and across the body (transcriptional start site to transcriptional termination site) of all upregulated miRNAs in E2-treated *XVE:ROS1-YFP* lines. ns = not significantly different in E2-treated plants compared to DMSO-treated plants within a genotype (t-test, $p > 0.05$).



Supplementary Figure 4.5 Verification of antagonistic *cis*-regulation in the promoter of *NPR1* by ROS1 and RdDM.

(a) Genome browser snapshots showing WGBS DNA methylation (scale = 0-100%) in DMSO- and E2-treated Col-0, L5, and L7 at the *NPR1* locus, and the approximate location of designed primers for McrBC-qPCR and RT-qPCR (Table 4.1). Annotations are from Araport11 and TAIR10 TE annotation files. Each track is an overlay of all three biological replicates. Figures created from Integrative Genomics Viewer (IGV) (Thorvaldsdóttir et al., 2013). Red bar highlights DMR detected in E2-treated *XVE:ROS1-YFP* lines in the WGBS data. (b and c) qPCR results from an independent repeat experiment using DMSO- and E2-treated *XVE:ROS1-YFP* L5 seedlings (as illustrated in Figure 4.2a). (b) Relative DNA methylation in the promoter of *NPR1* as measured by McrBC-qPCR. (b) Relative *NPR1* expression, as measured by RT-qPCR. (d) DNA methylation (allC) of the *NPR1* promoter region DMR (as indicated by red bar in (a)) in Col-0, *nrpe1-11* and *ros1-4* seedlings that had a stress-free parental treatment (pMock). Data is from the WGBS data presented in Chapter 2 using cytosines with ≥5 mapped reads. Mean cytosine DNA methylation of the region was calculated using bedtools v2.31.0 (Quinlan and Hall, 2010).

Chapter 5. Investigating the Roles of (Peri)centromeric DNA (De)methylation in the Establishment and Erasure of ROS1-Driven Immune Memory.

Authors:

Hannan Parker, A[1]; Peijun Zhang[1]; Hipperson H.; Rolfe S. A. [1]; Smith, L. M.[1] and Ton, J[1].

Affiliations:

[1] Plants, Photosynthesis and Soil, School of Biosciences, The University of Sheffield, Western Bank, Sheffield, S10 2TN, UK

Author contributions:

AHP (the candidate), SAR, LMS, and JT proposed the original ideas for the research. AHP conducted experiments and gathered data with assistance from PZ and HH. Data analysis was performed by AHP. AHP created all Figures and wrote all the text in the Chapter. JT reviewed and provided comments on the Chapter. Funding for the project came from studentships/grants awarded to AHP and JT.

Contribution of co-authored publications:

Parts of the methods related to the sequencing and analysis of the Oxford Nanopore Technology sequencing data has been adapted from the following published article with permission from Springer Nature:

Wilkinson, S. W.; Hannan Parker, A.; Muench, A.; Wilson, R. S.; Hooshmand, K.; Henderson, M. A.; Moffat, E. K.; Rocha, P. S. C. F.; Hipperson, H.; Stassen, J. H. M.; López Sánchez, A.; Fomsgaard, I. S.; Krokene, P.; Mageroy, M. H. and Ton, J. (2023). Long-Lasting Memory of Jasmonic Acid-Dependent Immunity Requires DNA Demethylation and ARGONAUTE1. Nature Plants, 9(1): 81–95.

5.1 Abstract

Changes in DNA methylation play a critical role in the establishment and maintenance of plant immune memory. However, remarkably little is understood about the molecular mechanisms that underpin the loss, or erasure, of this immune memory in the absence of stress. Previous characterisation of epigenetically modified inbred lines of Arabidopsis revealed that heritable reductions in DNA methylation in pericentromeric regions can provide plants with a strong and highly stable immune memory phenotype. However, the genetic and epigenetic characterisation of (peri)centromeric regions by short-read sequencing applications has proven technically challenging due to the highly repetitive nature of transposable elements (TEs) and satellite repeats in these regions. In this Chapter, I have leveraged long-read Oxford Nanopore Sequencing (ONT-seq) to investigate the epigenetic changes within (peri)centromeric regions at the onset of ROS1-driven immune memory. Strikingly, while estradiol (E2)-indued ROS1 *in XVE:ROS1-YFP* plants caused widespread DNA demethylation in the chromosome arms, DNA methylation in the (peri)centromeric regions was increased, potentially reflecting a rapid DNA re-methylation response to counter ROS1-induced DNA demethylation and erase epigenetic immune memory. However, blocking CMT3-dependent DNA (re)methylation in CHG contexts did not alter the strength or durability of ROS1-induced resistance (ROS1-IR) against the hemi-biotrophic pathogen *Pseudomonas syringae* pv. *tomato* DC3000 (*Pst*). By contrast, the chemical inhibitor of DNA methyltransferases 5-azacytidine, which particularly perturbs CHH DNA methylation in pericentromeric regions, not only increased the strength of ROS1-IR against *Pst* but also prolonged the immune memory by at least one week. Combined with my previous finding that transient induction of ROS1 increases the accumulation of pericentromeric small RNAs (Chapter 4), these results suggest that RNA-directed re-methylation of the (peri)centromeric regions erases epigenetic immune memory. The results are discussed in a context of a new model, which presents plant immune memory as a function of epigenetic interactions between the (peri)centromeric regions and chromosome arms.

5.2 Introduction

Induced resistance (IR) occurs when plants respond more rapidly and/or robustly to a stress after being exposed to an immune-activating stimulus earlier in their life (De Kesel et al., 2021). This enhanced defence mechanism enables genetically susceptible plants to mount a more efficient response against recurrent attacks. In order to achieve IR, plants must 'remember' the primary stress stimulus – illustrating how IR serves as a form of immune memory in plants (Cooper and Ton, 2022). Previous research has shown that IR can be effective against a wider range of biotic stresses, both over relatively short periods and longer timescales, including inter- and trans-generational IR responses (Mauch-Mani et al., 2017; López Sánchez et al., 2021; Yassin et al., 2021; De Kesel et al., 2021; Furci et al., 2023). Furthermore, studies of Arabidopsis and tomato have highlighted the role of epigenetic modifications such as DNA (de)methylation in both the establishment and maintenance of immune memory (Conrath, 2011; Luna et al., 2014; López Sánchez et al., 2016; Catoni et al., 2022; Wilkinson et al., 2023). In support of this notion, the previous two Chapters of this thesis have demonstrated that enhanced activity of the DNA demethylase ROS1 in young Arabidopsis seedlings induces immune memory, resulting in IR against (hemi-)biotrophic pathogens.

Despite decades of research into IR, our understanding of how the underpinning immune memory is reversed remains remarkably limited. In 2014, Luna et al. showed that IR by the priming agent β-aminobutyric acid (BABA) against the (hemi-)biotrophic pathogens *Hyaloperonospora arabidopsidis* (*Hpa*) and *Pseudomonas syringae* pv *tomato* DC3000 is maintained for up to four weeks in Arabidopsis. However, disruption of the of the histone methyltransferase KYP, which forms a reinforcing loop with the CHG methyltransferase CMT3 (Fang et al., 2022), resulted in a more rapid reversal of immune memory, with IR responses observed up to 7 days after treatment but not 3 weeks after treatment (Luna et al., 2014). Similarly, plants lacking functional NPR1 could only maintain the BABA-IR for 2 weeks and had lost the IR by 3 weeks after BABA treatment (Luna et al., 2014). These results indicate that KYP and CMT3 ensure long-term maintenance of transcriptional priming of NPR1-dependent defence genes against (hemi-)biotrophic pathogens. On the other hand, plants defective in Pol IV-dependent RNA-directed DNA methylation (RdDM) or the DNA demethylase ROS1 can mount effective IR against the generalist herbivore *Spodoptera littoralis* one day after JA treatment, but are unable to sustain this IR three weeks after seedling treatment with JA (Wilkinson et al., 2023). These same mutants were also found to be affected in heritable IR (h-IR) against *Hpa* following parental disease by *Pst* (López Sánchez et al., 2016). Therefore, regulation of DNA methylation by Pol IV-dependent RdDM and ROS1 is

involved in the long-term and heritable maintenance of immune memory against chewing herbivores and biotrophic pathogens, respectively. Finally, *ddm1* plants, which are severely compromised in DNA methyltransferase activity, show DNA hypomethylation in all sequence contexts, and are highly resistant to the biotrophic pathogen *Hpa* (Zemach et al., 2013; López Sánchez et al., 2016; Furci et al., 2019). Remarkably, even after backcrossing these *ddm1* plants with wild-type plants and selfing wild-type progeny for 8 generations, many epigenetic recombinant inbred lines (epiRIL) maintained high levels of immunity against *Hpa* (Reinders et al., 2009; Johannes et al., 2009; Furci et al., 2019). Interestingly, 60% of the variation in *Hpa* resistance in this epiRIL population could be attributed to four epigenetic quantitative trait loci (epiQTLs). However, none of these epiQTLs contained genes displaying increased transcription or primed transcription upon *Hpa* inoculation; instead, they mapped near the centromeres of four of the five Arabidopsis nuclear chromosomes (chromosomes 1, 2, 4, 5; Furci et al., 2019). Therefore, loss of DNA methylation in TE-rich pericentromeric regions promotes long-term maintenance and heritability of epigenetic immune memory against biotrophic pathogens, which is based on a priming of defence genes via *trans*-acting mechanisms (Wilkinson et al., 2019; Furci et al., 2019).

In the previous Chapter, the DNA hypomethylation response upon ectopic induction of ROS1 was notably less pronounced in (peri)centromeric regions and correlated with an increase in sRNAs associated with Pol IV-dependent RdDM. However, in the past, research into the function, evolution, and epigenetic features associated with (peri)centromeric regions has been a challenging task, due to the complexity of aligning short sequencing reads (~ 50-150 nt) to these regions, which contain long and highly repetitive sequences (Naish et al., 2021). Recent advancements in long-read sequencing technologies, like Oxford Nanopore Technologies (ONT) and PacBio high-fidelity (HiFi), have greatly improved our understanding of these regions in both humans and plants (Naish and Henderson, 2024). In 2021, the first Arabidopsis genome (Col-0 ecotype) with complete assembly of the centromeric regions, and the epigenetic features associated with them, was published in Science, and subsequently named the Col-CEN genome assembly (Naish et al., 2021). In this Chapter, I leverage the benefits of long-read ONT sequencing (ONT-seq) and the Col-CEN genome assembly (Naish et al., 2021) to characterise DNA methylation in the (peri)centromeric regions during the establishment of ROS1-driven immune memory in *XVE:ROS1-YFP* lines. In support of the (short-read) whole genome bisulfite sequencing (WGBS) analysis in Chapter 4, ROS1 induction caused robust reductions in DNA methylation along the chromosome arms. Strikingly, however, ROS1 activity was also associated with significant, albeit more variable, DNA hypermethylation in centromeric and pericentromeric regions in non-CG contexts. Based on the findings of Furci et al (2019), which established a causal contribution of severe DNA

demethylation to epigenetically heritable disease resistance, and my earlier finding that that ROS1-YFP induction dramatically increases the accumulation of pericentromeric small RNAs (Chapter 4), I have investigated the hypothesis that this variable DNA hypermethylation in the (peri)centromeres reflects a counter response to ROS1-induced DNA demethylation, resulting in the erasure of ROS1-driven immune memory against *Pst*. While the mutation of CMT3, which maintains DNA methylation in CHG sequence contexts, did not extend the durability of ROS1-driven immune memory, concurrent inhibition of DNA methyltransferase activity by 5-Azacytidine, a chemical that preferentially inhibits pericentromeric CHH methylation, was effective in prolonging immune memory by at least 1 week. Integrating these results with those from Chapter 4, I propose a role of Pol IV-dependent RdDM in the erasure of ROS1-driven immune memory against *Pst*.

5.3 Materials and Methods

5.3.1 Plant Material

Four *Arabidopsis thaliana* (Arabidopsis) genotypes were used in this Chapter: wild-type accession Col-0, *XVE:ROS1-YFP* line 5 (L5), the T-DNA insertion mutant *cmt3-11* (SALK_148381) (Alonso et al., 2003), and a cross between *XVE:ROS1-YFP L5* and *cmt3-11*. The *cmt3-11* mutant and the *XVE:ROS1-YFP* L5 transformant are in the genetic background of Col-0. Details on the generation of *XVE:ROS1-YFP* L5 plants are detailed in Section 2.4 of this thesis. Phenotypic and genomic characterisation of *XVE:ROS1-YFP* L5 can be found in Chapter 3 and Chapter 4 of this thesis. The *cmt3-11* null allele was first described by Chan et al. (2006) and was verified to be homozygous for the T-DNA insertion using PCR and insertion-specific primers (Table 5.1), as described in López Sánchez et al. (2016).

5.3.2 Growth Conditions, Chemical Treatments, and Material Harvest

Plants were cultivated in short-day conditions using 70 mL pots containing a 2:1 soil:sand mixture following a 4 day stratification period, as detailed in Section 2.2. Ten seeds were sown per pot, and 6 pots per genotype were grown within a single tray. Plants were watered from the bottom 2-3 times per week. For the Oxford Nanopore Technology sequencing (ONT-seq) experiment, aerial parts of seedlings were sprayed twice with estradiol (E2) or dimethylsulfoxide (DMSO). The first spray was performed at 14 days after sowing (DAS) with 10 uM E2 or 0.02% (v/v) DMSO; the second spray was performed at 18 DAS with 25 uM E2 or 0.05% DMSO (v/v). Control treatments were based on equivalent DMSO sprays without E2. All solutions were prepared in dH_2O and were supplemented with 0.015% (v/v) silwet L-77 (LEHLE SEEDS, VIS-30) (silwet). Each tray of 18 pots was sprayed with a total volume of 20 mL E2 solution or DMSO solution. Further details about the preparation and storage of E2 stocks are detailed in Section 2.3. For each biological replicate (n=3), 48 hours after the second spray treatment and following 2 hours in the dark, aerial tissues from approximately 100 seedlings (~500 mg fresh weight) were harvested, snap frozen in liquid N_2, and stored at -80˚C.

5.3.3 Oxford Nanopore Technology (ONT) Library Preparing and Sequencing

High molecular weight (HMW) DNA was extracted from frozen tissue using the NucleoBond® HMW DNA kit (Macherey Nagel, 740160.20) according to the manufacturer's instructions. DNA quality, quantity and size were assessed using a Nanodrop 8000 spectrophotometer, a Qubit 3.0 Fluorometer and an Agilent Femto Pulse. Library preparation was conducted using the Rapid Barcoding Sequencing kit (ONT, SQK-RBK004) with 400 ng

of extracted HMW genomic DNA. The prepared library was loaded into the FLO-MIN106 flow cell (FC) of a MinION sequencer (ONT). After 2 days, ~37% of FC pores were recovered using the Flow Cell Wash Kit (ONT, EXP-WSH004), and a fresh library preparation was loaded. The FC was then run to exhaustion (~2 days). After sequencing, raw .fast5 files were base-called and filtered using Guppy v6.0.1 (function: guppy_basecaller; options; --config dna_r9.4.1_450bps_hac.cfg). Summary statistics were obtained using NanoPlot v1.39.0 (Supplementary Data 5.1).

5.3.4 Cytosine Methylation Calling From ONT-seq

Methylated cytosines were identified using DeepSignal-plant v0.1.4, as recommended in the GitHub user guide (Ni et al., 2021) (https://github.com/PengNi/deepsignal-plant). Base-called sequences were annotated onto raw .fast5 files (function: tombo preprocess annotate_raw_with_fastqs; options: --basecall-group Basecall_1D_000) and resquiggled (raw signal to reference sequence alignment) (function: tombo resquiggle; options: --signal-length-range 0 10000000, --sequence-length-range 0 1000000, --basecall-group Basecall_1D_000) using Tombo v1.5.1 (https://github.com/nanoporetech/tombo) and the Arabidopsis reference genome Col-CEN v1.2 (Naish et al., 2021) (https://github.com/schatzlab/Col-CEN/tree/main/v1.2). Methylation predictions for all cytosines in the genome were called, using DeepSignal-plant with the model 'model.dp2.CNN.arabnrice2-1_120m_R9.4plus_tem.bn13_sn16.both_bilstm.epoch6.ckpt' (function: deepsignal_plant call_mods; options: --corrected_group RawGenomeCorrected_000, --motifs C) (Ni et al., 2021). The frequency of methylation at each CG, CHG, and CHH site was then called by DeepSignal-plant (function: deepsignal_plant call_freq; options: default parameters). The unmethylated plastid genome was used for estimating false-positive methylation calls, with all samples showing an average plastud methylation percentage < 1% and an average methylation percentage across all samples of 0.74% (Supplementary Data 5.1), which, for reference, is below the accepted non-conversion rate of WGBS data (< 2%) (Stuart et al., 2018).

5.3.5 Genome Alignment Between TAIR10 and Col-CEN Assemblies

Alignment between nuclear genomes of Col-CEN v1.2 (Naish et al. 2021) and TAIR10 were performed using Minimap2 v2.28 (Li, 2018) (function: minimap2; options: -x asm5 --secondary = no). Alignments were only kept if they had a mapping score ≥ 60. Plots were generated using the R package syntenyPlotteR v1.0.0 (function: draw.linear) (Quigley et al., 2023) and ggplot2 v3.5.0. The analysis and generation of this plot was done to resemble the Figure presented in Naish et al. (2021).

5.3.6 Comparison of ONT-seq and WGBS Data Mapping to TAIR10 and Col-CEN Genome Assemblies

Processed whole genome bisulfite sequencing (WGBS) data derived from *XVE:ROS1-YFP* L5 seedlings, treated identically with E2 and DMSO (see Chapter 4), were aligned to the Col-CEN v1.2 Arabidopsis genome (Naish et al., 2021), using Bismark v0.24.1 (Krueger and Andrews, 2011) (function: bismark; options: default parameters), which utilises Bowtie2 v2.5.1 (Langmead and Salzberg, 2012). PCR duplicate reads were subsequently removed using Bismark (function: deduplicate_bismark; options: default parameters). Bismark was also used to generate and summarise binary scores for methylated cytosines (1) and unmethylated cytosines (0) in all contexts for every cytosine in the Col-CEN assembly (function: 'bismark_methylation_extractor; options: --comprehensive --bedGraph --CX --cytosine_report --ignore_r2 4). Mapping WGBS data to the TAIR10 assembly were based on the methylation files described in Chapter 4. Long-read ONT-seq data were processed as described above, using the TAIR10 reference genome during the Tombo resquiggle step instead of the Col-CEN reference genome. As DeepSignal-plant only reports DNA methylation information at cytosines which have at least 1 mapped read bedtools v2.31.0 (Quinlan and Hall, 2010) was used to intersect the Bismark cytosine reference files with DeepSignal-plant methylation files (function: bedtools intersect; options: -a ref_CX_file.bed -b DeepSignalMeth.bed -wao), thereby ensuring that any unmapped cytosines from the ONT-seq data have a read count of 0. Windows of 100 kb in size were created for Col-CEN and TAIR10 using bedtools (function: bedtools makewindows; options: -g ChromLengths.bed -w 100000 -s 100000) and the average number of reads mapping to cytosines within 100 kb windows were calculated for all WGBS and ONT-seq samples and mapped to both TAIR10 and Col-CEN, using bedtools (function: bedtools map; options: -o mean). Coverage plots were generated using the R package ggplot2 v3.5.0.

5.3.7 Analysis of Single Cytosine DNA Methylation

5.3.7.1 Principal Component Analysis (PCA)

PCA was performed using positional cytosine methylation levels (%) in any context (allC). Cytosines with <5 mapped reads were dropped from the analysis. Furthermore, to focus the analysis on the most variable regions, cytosines with standard deviations less than the median of the standard deviation across all cytosines were removed from the analysis. PCA was conducted in base R v4.2.1 (function: prcomp; options: scale = FALSE, center = TRUE) and plots were generated using R package ggplot2 v3.5.0.

5.3.7.2 Chromosome-Level Metaplots

Plots are based on average levels of cytosine DNA methylation in 100 kb bins of the Col-CEN genome in allC, CG, CHG, and CHH contexts. Only cytosines with 5 or more reads were included in the analysis, after which bins with at least 60% of the expected number of sites for a given context based on averages of the nuclear Col-CEN genome, rounded to the nearest thousand, were retained. This resulted in the following selection: bins must contain ≥ 3000 sites for CG and CHG contexts, ≥ 16000 sites for CHH context, and ≥ 22000 sites for allC, all with sufficient coverage (≥ 5 reads). From this selection, windows were only kept if they survived filtering steps in at least two of the three samples within the DMSO or E2 treatment. To define pericentromeric regions, the 100 kb windows with TE densities > protein coding gene densities were identified. For each nuclear chromosome, the start of the pericentromere was defined as the midpoint of the first bin in a string of at least 5 neighbouring bins for which TE density > gene density. The end of the pericentromere was defined as the midpoint of the last bin in a string of at least 5 neighbouring bins where TE density > gene density. The pericentromeric ranges for Col-CEN were defined as: Chr1: 12050000-19850000, Chr2: 1350000-9850000, Chr3: 10350000-18950000, Chr4: 1750000-9850000, Chr5: 9650000-18150000. The pericentromeric midpoints, rounded to the nearest hundred thousand, were calculated using these coordinate ranges. The centromeric ranges were defined in Naish et al (2021) as follows: Chr1: 14840000-17560000, Chr2: 3823000-6046000, Chr3: 13597000-15734000, Chr4: 4204000-6978000, Chr5: 11784000-1456000. The position of 100 kb cytosine DNA methylation windows were expressed relative to the midpoint of the associated chromosome pericentromere, and average 100 kb cytosine DNA methylation levels were calculated within a treatment based on their relative distance to the pericentromeric midpoint. The pericentromeric and centromeric regions in the metaplots were defined as the maximum absolute relative distance for any (peri)centromeric window from the midpoint of the pericentromere for both the start and end coordinates.

5.3.7.3 Average Cytosine Methylation Plots

Average cytosine DNA methylation (%) was calculated for each sample genome-wide in allC, CG, CHG, and CHH sequence contexts, for all cytosines with ≥ 5 mapped reads. Cytosine DNA methylation was also calculated within chromosome arms, pericentromeric regions, and centromeric regions, using bedtools v2.31.0 (Quinlan and Hall, 2010) (function: bedtools intersect). Pericentromeric regions were defined as above, but with the centromeric regions removed, as defined by Naish et al (2021). Chromosome arms included all other regions of the nuclear genome.

5.3.8 Differentially Methylated Region (DMR) Analyses of Cytosines

5.3.8.1 Identification of DMRs From ONT-seq

To identify consistent and stochastic DMRs between E2-treated and DMSO-treated *XVE:ROS1-YFP* plants, methylation frequency files from DeepSignal-plant were analysed, using R package DSS v2.46.0 (H. Feng et al., 2014; Park and Wu, 2016). To identify consistent DMRs between replicates a statistical comparison was made between 3 replicates from the DMSO treatment against 3 replicates from the E2 treatment ('3 vs 3 comparison'). To this end, differentially methylated cytosines in allC, CG, CHG, and CHH contexts were first identified in this 3 vs 3 comparison using DSS (function: DMLtest; options: equal.disp = FALSE, smoothing = FALSE), after which DMRs were called with the function 'callDMR', using parameters 'delta =0.1', 'p.threshold = 0.05', 'minlen = 25', 'minCG = 3', 'dis.merge = 50', and 'pct.sig = 0.5'. To identify stochastic DMRs that vary between replicates, a statistical comparison was made between individual replicates from the DMSO treatment versus the E2 treatment ('1 vs 1 comparison'). To do so, each E2-treated sample was tested against each DMSO-treated sample, using DSS as above, but enabling the assumption of equal dispersion between samples (function: DMLtest; options: equal.disp = TRUE, smoothing = FALSE), which allows for testing without biological replicates (H. Feng et al., 2014; Park and Wu, 2016; Wilkinson et al., 2023). DMRs were then called, using the same parameters as in the 3 vs 3 comparison. To eliminate multiple counts of the same DMR in the 1 vs 1 comparison, DMRs of the same context and direction (hypermethylated or hypomethylated), were merged and simplified across all comparisons using the R package GenomicRanges v1.54.1 (function: reduce) (Lawrence et al., 2013). Furthermore, using GenomicRanges, all DMRs were checked for overlaps with transposable elements, using the Col-CEN v1.2 TE annotation file ("t2t-col.20210610.fasta.mod.EDTA.TEanno.gff3") downloaded from the Col-CEN v1.2 GitHub page (https://github.com/schatzlab/Col-CEN/tree/main/v1.2; accessed 21/09/2023).

5.3.8.2 DMR Chromosome Plots and Distribution Analysis

Hypo-DMRs and hyper-DMRs from the 3 vs 3 comparison and the 1 vs 1 comparison were merged and simplified across all contexts separately using the R package GenomicRanges v1.54.1 (function: reduce) (Lawrence et al., 2013). To visualise the position of each DMR along the 5 nuclear chromosomes of the Col-CEN v1.2 assembly (Naish et al., 2021), the R package karyoploteR v1.28.0 (Gel and Serra, 2017) was used. The number of DMRs falling within pericentromeric, centromeric, and chromosome arm regions (as defined above) for a given direction (hypo/hyper) and comparison type (3 vs 3 / 1 vs 1) were identified using GenomicRanges (function: findOverlaps; options: type = "within"). The proportion of base pairs in the nuclear genome of Col-CEN v1.2 falling within centromeres, pericentromeres,

and chromosome arms, as defined above, was 9.6%, 21.9%, and 68.5%, respectively. Thus, the expected number of DMRs for a given direction (hypo/hyper), comparison type (3 vs 3 / 1 vs 1), and genomic region (centromeres / pericentromeres / chromosome arms) was calculated by multiplying the total number of DMRs with the proportion of each region in the nuclear Col-CEN v1.2 genome. Chi-squared tests were used to compare the observed distribution of DMRs against the expected distribution of DMRs. To correct for low count values, a Fisher's exact test was used for the analysis of the 3 vs 3 hyper-DMR set.

5.3.8.3 DMR Comparison From WGBS and ONT-seq

To compare the methylation calls from WGBS and ONT-seq paired with DeepSignal-plant and Bismark, respectively, hypo-DMRs in the 3 vs 3 comparison were generated for both datasets and mapped onto TAIR10 and Col-CEN. Hypo-DMRs were merged across all contexts within a data type and reference genome combination using the R package GenomicRanges v1.54.1 (function: 'reduce') (Lawrence et al., 2013). Overlapping DMRs between WGBS and ONT-seq data were identified both in TAIR10 and Col-CEN using GenomicRanges (function: findOverlaps; options: select = "all").

5.3.9 Identification of the Genomic Insertion Site of the *XVE:ROS1-YFP* Cassette and Selection of Homozygous *cmt3-11 XVE:ROS1-YFP* Plants

To identify the insertion site of *XVE:ROS1-YFP* L5 seedlings, Oxford Nanopore Technology sequencing (ONT-seq) reads from two *XVE:ROS1-YFP* L5 samples were checked for presence of the Venus YFP sequence (Nagai et al., 2002), using local BLAST (Camacho et al., 2009). Reads with high confidence matches (*e*-value < 1^{-200}) were used for *de novo* assembly, using Flye v2.9.2 (Kolmogorov et al., 2019) (function flye; options: default parameters). This analysis revealed a single insertion site of *XVE:ROS1-YFP* in L5 in the second intron of gene AT5G17060, which had undergone an inverted tandem duplication, centred exactly at the right border (RB) T-DNA repeat (Figure 5.1). To enable identification of the *XVE:ROS1-YFP* L5 insertion by PCR in F_2 plants from a cross between L5 and *cmt3-11*, primers flanking the insertion site (XVE LP + XVE RP) were designed to detect the wild-type allele. A primer within the right border (XVE RB) of the insertion site was used with the RP to detect the *XVE:ROS1-YFP*-inserted L5 allele (Table 5.1). The wild-type allele of *CMT3* was detected using primers flanking the T-DNA insertion site (CMT3 LP + CMT3 RP) and a primer specific to SALK T-DNA insertion lines (LBb1.3) in combination with the flanking CMT3 primer (CMT3 RP) was used to detect the T-DNA-inserted *cmt3-11* allele (Alonso et al., 2003). Plant 28 in the F_2 population was identified as homozygous for both the *cmt3-11* T-DNA insertion and the *XVE:ROS1-YFP L5* insertion (Figure 5.1).

Table 5. 1 PCR primers used for genotyping in this Chapter.

	Name	Nucleotide sequence
Genotyping the T-DNA insertion in *cmt3-11* (SALK_148381)	CMT3 LP	CCCTCAACAATTAACTGACGC
	CMT3 RP	ATAAGAGAAGGAGCTGCTGCC
	LBb1.3	ATTTTGCCGATTTCGGAAC
Genotyping the *XVE:ROS1-YFP* cassette insertion in L5	XVE_L5 LP	AAGTCTTTTGCAAGTTCTGAGGT
	XVE_L5 RP	TGACAGGTCGTTATGCTGGG
	XVE_L5 RB	ATCATGGTGGAAAATGGCCG

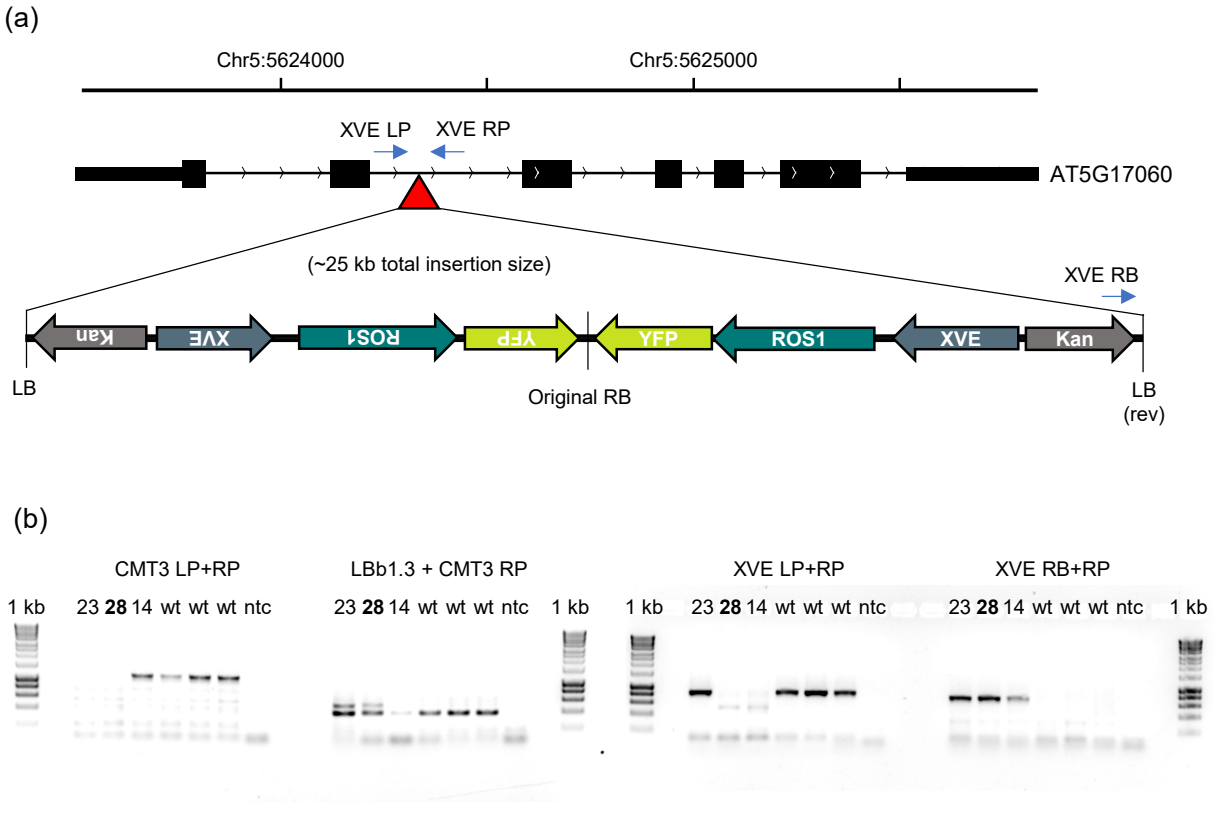


Figure 5.1 Insertion site of *XVE:ROS1-YFP L5* and genotyping of *cmt3-11 / XVE:ROS1-YFP* L5 F_2 crosses.
(a) Identification of the genomic insertion site of the *XVE:ROS1-YFP* casette in L5 (red triangle), using long-read Oxford Nanopore Technology sequencing (ONT-seq). The insertion site is in the second intron of AT5G17060 and the construct has undergone an inverted duplication at the insertion site, cantered around the right border (RB) of the construct. The AT5G17060 gene model is drawn to scale; components of the *XVE:ROS1-YFP* construct (*Kan, XVE*, *ROS1, YFP*) are not drawn to scale. Primers are shown as blue arrows, with XVE LP and XVE RP complementary to the regions flanking the *XVE:ROS1-YFP* insertion site in L5, and XVE RB complementary to the right border (RB) of the *XVE:ROS1-YFP L5* insertion, which corresponds to an inverted copy of the left border (LB) sequence (LB rev). (b) Agarose gel electrophoresis analysis showing PCR products from 3 F_2 plants of a *cmt3-11* and *XVE:ROS1-YFP* L5 cross. Three wild type Col-0 controls ("wt") were included per PCR reaction, as well as a non-template control (ntc). Plant 28 was identified as homozygous for both the *cmt3-11* mutation and *XVE:ROS1-YFP* L5 insertion.

5.3.10 Testing the Durability of ROS1-Induced Resistance in Col-0 and *cmt3-11* Backgrounds

Col-0 and *cmt3-11* plants with or without the *XVE:ROS1-YFP* L5 insertion were treated twice with E2 or DMSO, as per all previous sequencing experiments, before challenge inoculation with *Pseudomonas syringae* pv. *tomato* DC3000 *luxCDABE* (*Pst-Lux*) (Fan et al., 2008; Furci et al., 2021) at 1 day, 1 week, or 2 weeks after the final E2/DMSO treatment (Figure 5.5a). For each infection time point ~18 plants per genotype per treatment were infected with *Pst-Lux* with an OD_{600} = 0.2 (section 2.5). *Pst-Lux* bioluminescence was captured

and analysed at 2- and 3-days post inoculation (dpi), as described in Section 2.5. Data within an infection stage was selected based on the dpi which had the highest mean infection across all groups. However, in cases where variance in *Pst-Lux* bioluminescence was exceptionally high for one of the two dpi's (twice as high as the other dpi), the dpi with the lower variance was selected even if the mean was lower. The experiment was independently repeated 3 times, and *Pst-Lux* bioluminescence values were expressed relative to the mean of DMSO-treated Col-0 plants for each experiment before pooling data. Two-way ANOVA followed by Tukey post-hoc tests were used to determine statistically significant differences in relative *Pst-Lux* bioluminescence between treatment groups at each infection stage, using base R v4.3.2. Plots were generated using the R package ggplot2 v3.5.0.

5.3.11 Testing the Durability of ROS1-Induced Resistance in 5-Azacytidine-treated Plants

Col-0 and *XVE:ROS1-YFP* L5 plants, treated twice with DMSO or E2 (as described above), were also treated with either 100 µM 5-Azacytidine (5-Aza; Sigma #A2385) or H_2O at 7, 9, 11, 13 and 16 days after sowing (DAS). This corresponds to 5 treatments of 5-Aza or H_2O every 2-3 days starting 1 week before the first DMSO/E2 treatment and ending 1 day before the final DMSO/E2 treatment. To minimise health risks of spraying 5-Aza solution, a single droplet containing either 5-Aza or a mock solution (H_2O), supplemented with 0.015% (v/v) surfactant (silwet; LEHLE SEEDS, VIS-30), was carefully applied to the apical meristem of every seedling. Droplet volumes were 6 µL at 7 and 9 DAS, 8 µL at 11 DAS, and 10 µL at 13 DAS. *Pst-Lux* inoculations and analysis of bioluminescence were performed and analysed as described for above. The (epi)genotype group was defined as the combination of genotype and 5-Aza treatment. Two-way ANOVAs were performed to test the impacts of (epi)genotype group, spray treatment, and their interaction on relative *Pst-Lux* bioluminescence. Tukey post-hoc tests were used to determine significant differences in relative *Pst-Lux* bioluminescence between (epi)genotype-treatment combinations at each infection stage using base R v4.3.2. Plots were generated using the R package ggplot2 v3.5.0.

5.4 Results

5.4.1 Enhanced (Peri)centromeric Resolution Using Long-Read Sequencing With the Col-CEN Genome Assembly

Syntenic alignment plots of nuclear chromosome sequences from TAIR10 (Ensembl Plants release v58) and Col-CEN (v1.2) (Naish et al., 2021) revealed no major differences along the chromosome arms, consistent with Naish et al.'s (2021) findings (Figure 5.2a). However, the Col-CEN assembly provided greater sequence information within (peri)centromeric regions compared to TAIR10, offering better resolution of these areas. This increased detail accounts for the longer lengths observed in all five chromosomes in the Col-CEN assembly (Figure 5.2a).

XVE:ROS1-YFP L5 seedlings (L5 seedlings) were treated twice with either DMSO or E2 and harvested, as described in Chapter 4. Biologically replicated samples (n=3) were then subjected to high molecular weight DNA extractions for long-read Oxford Nanopore Technology (ONT) sequencing (ONT-seq). When mapped onto the TAIR10 genome, the average read coverage of cytosines was roughly consistent across all five nuclear chromosomes for both whole genome bisulfite sequencing (WGBS) and ONT-seq data (Figure 5.2b). However, mapping short-read WGBS data to the Col-CEN genome revealed sharp reductions in cytosine coverage within (peri)centromeric regions (Figure 5.2b). In contrast, ONT-seq data showed relatively even cytosine coverage across all five nuclear chromosomes in the Col-CEN assembly (Figure 5.2b). Thus, short-read sequencing data are not well-suited for the Col-CEN genome assembly, especially when focusing on (peri)centromeric regions. For unknown reasons, spurious spikes in coverage were observed (Figure 5.2b), particularly in pericentromeric regions, which could derive from sequencing or mapping artifacts.

Next, hypomethylated regions (hypo-DMRs) in E2-treated L5 seedlings relative to DMSO-treated L5 seedlings were identified for both WGBS and ONT-seq data and mapped onto both the TAIR10 and Col-CEN assemblies. Cytosine methylation in all contexts (allC, CG, CHG, CHH) was called using Bismark (Krueger and Andrews, 2011) and DeepSignal-plant (Ni et al., 2021) for WGBS data and ONT-seq, respectively. Hypomethylated DMRs for both sequencing methods were called by DSS (H. Feng et al., 2014; Park and Wu, 2016) using identical parameters. Subsequently, hypo-DMRs from both methods were combined and simplified across all contexts. In the TAIR10 assembly, 49% of hypo-DMRs were consistent between WGBS and ONT-seq, while in the Col-CEN assembly, 50% of hypo-DMRs were consistent between both sequencing methods (Figure 5.2c). Since the WGBS and ONT-seq data are derived from completely independent experiments, this overlap underscores the

robustness of the *XVE:ROS1-YFP* system and indicates a high degree of agreement in cytosine methylation calls across sequencing methods and analysis software. For comparison, 42% of hypo-DMRs were consistent between L5 and L7 seedlings from the same experiment, using the same sequencing and analysis pipelines (Chapter 4; Figure 4.3d). Thus, to explore the effects of ROS1 induction on (peri)centromeric cytosine methylation, long-read ONT-seq data mapped onto the Col-CEN genome were analysed for cytosine methylation using DeepSignal-plant (Ni et al., 2021).

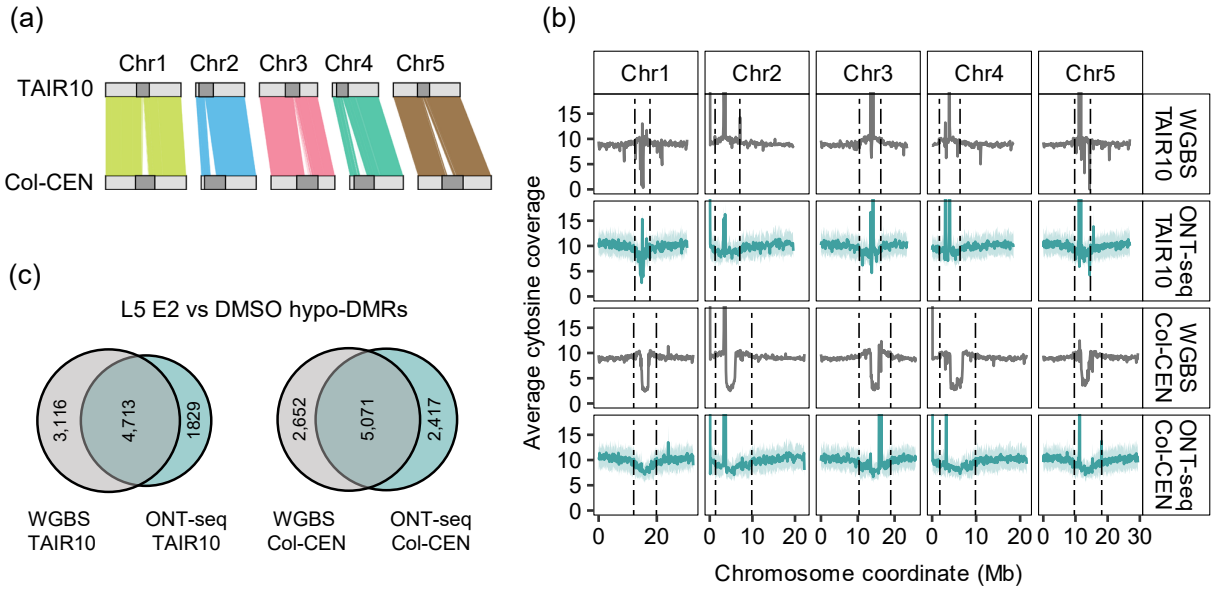


Figure 5.2 Comparison of Col-CEN and TAIR10 genome assemblies and their performance with short-read and long-read sequencing data.
(a) Syntenic alignment between TAIR10 and Col-CEN (v1.2) genomes, recreated in a similar style to Naish et al. (2021). Vertical coloured lines indicate high-confidence Minimap2 alignments (mapping score ≥ 60) based on sequence complementarity. White gaps represent regions with no alignment between genomes. Light grey boxes depict chromosomes proportional to their lengths, and dark grey boxes indicate (peri)centromeric regions (see Methods for the definition of pericentromeric regions). (b) Average cytosine coverage along the five nuclear chromosomes following alignment of short-read WGBS data (Chapter 4) and long-read ONT-seq data, using Bismark (Krueger and Andrews, 2011) and DeepSignal-plant (Ni et al. 2021), respectively. Solid lines show average cytosine coverage across all 6 samples (3 DMSO-treated and 3 E2-treated L5 samples) within 100 kb windows. Shaded areas represent the standard error of the mean for each window. Dashed vertical lines mark the (peri)centromeric regions. (c) Overlap of hypomethylated regions (hypo-DMRs) in E2-treated L5 seedlings between short-read WGBS and long-read ONT-seq data, mapped to the Col-CEN (left) or TAIR10 (right) assemblies. Cytosine methylation in all contexts (allC, CG, CHG, CHH) was called by Bismark (Krueger and Andrews 2011) and DeepSignal-plant (Ni et al. 2021) for WGBS data and ONT-seq, respectively. DMRs were identified by DSS and simplified by GenomicRanges (see methods for more details on hypo-DMR set creation).

5.4.2 ROS1 Induction Causes Concurrent Hypomethylation in Chromosome Arms and Hypermethylation in (Peri)centromeres

Principal component analysis (PCA) of single cytosine methylation (%) revealed a distinct separation between DMSO-treated and E2-treated seedlings along the first principal component (Figure 5.3a). To further explore these differences, cytosine methylation in E2-treated seedlings across all (allC) and distinct (CG, CHG, and CHH) sequence contexts was assessed at the chromosomal level. Average cytosine methylation levels were calculated over 100 kb windows relative to DMSO-treated seedlings (%) and plotted relative to the midpoint of pericentromeres. These windows were then averaged across all five nuclear chromosomes based on their relative positions to generate chromosome-level metaplots (Figure 5.3b). In E2-treated seedlings, DNA methylation was reduced relative to DMSO-treated seedlings along the chromosome arms and across all sequence contexts (Figure 5.3b). However, this reduction was notably diminished within the (peri)centromeric regions. Remarkably, DNA methylation appeared to increase within the centromeric regions of E2-treated seedlings, particularly in CHG and CHH contexts (Figure 5.3b). This pattern was consistent when analysed for each chromosome individually (Supplementary Figure 5.1).

When cytosine methylation levels were averaged at a base-pair resolution across the entire nuclear genome, no significant differences were observed between DMSO-treated and E2-treated seedlings in any contexts (t-test, $p > 0.05$) (Figure 5.3c). However, statistically significant reductions in cytosine methylation were detected along the chromosome arms in all contexts except CHG (t-test, $p < 0.05$) (Figure 5.3c). Notably, no contexts were significantly hypomethylated within the pericentromeric and centromeric regions of E2-treated seedlings (Figure 5.3c). Instead, E2-treated seedings displayed cytosine hypermethylation in CHG context within the pericentromeric regions and in CG and CHG contexts within the centromeric regions (t-test, $p < 0.05$; Figure 5.2c).

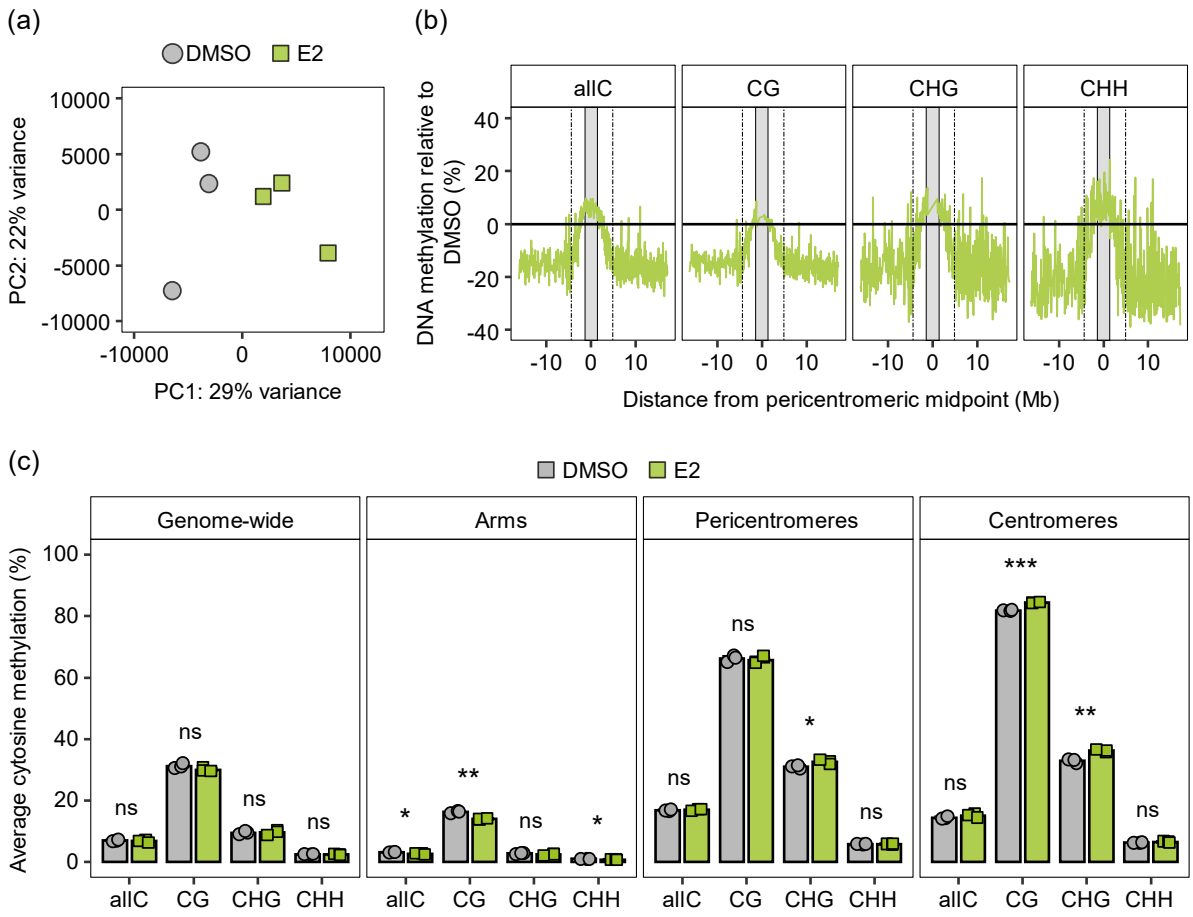


Figure 5.3 Global cytosine methylation analysis of DMSO- and E2-treated *XVE:ROS1-YFP* L5 seedlings (L5 seedlings).
(a) Principal component analysis (PCA) plot showing variation in cytosine DNA methylation across all sequence contexts (allC) between DMSO-treated (grey circles) and E2-treated (green squares) L5 seedlings. (b) Chromosome-level metaplots of cytosine DNA methylation in E2-treated L5 seedlings, expressed as a percentage relative to DMSO-treated plants. DNA methylation levels across all cytosines (allC), and cytosines in CG, CHG, and CHH sequence contexts were binned into 100 kb windows and averaged based on their distance from the midpoint of the pericentromere. Dashed vertical lines indicate the pericentromeric borders and shaded grey boxes represent centromeric regions relative to the pericentromere midpoint. (c) Relative cytosine DNA methylation across all contexts (allC, CG, CHG, CHH) for the entire genome (genome-wide), chromosome arms excluding (peri)centromeric regions (arms), pericentromeric regions excluding centromeric regions (pericentromeres), and centromeric regions only (centromeres), as defined by Naish et al. (2021). Shown are averaged values from 3 biological replicates.

5.4.3 ROS1 Induction Causes Consistent Hypomethylation Along Chromosome Arms and Variable Hypermethylation in (Peri)centromeric Regions.

To identify smaller genomic regions showing consistent changes in DNA methylation upon E2 treatment of L5 seedlings, differentially methylated regions (DMRs) were identified using DSS (H. Feng et al., 2014; Park and Wu, 2016) by comparing all three biological replicates of E2-treated seedlings against all 3 biological replicates of DMSO-treated seedlings (3 vs 3 DMR comparison) (Figure 5.4a). As was observed in the WGBS analysis of Chapter 4 (Figure 4.3c), E2-treated L5 seedlings displayed thousands of hypomethylated DMRs (10,659), with the majority (6,396) occurring in CG contexts (Figure 5.4b). When merging the hypomethylated DMRs across all sequence contexts, a simplified set of 7,486 hypo-DMRs was identified, of which none were located within centromeric regions, 1,992 (27%) were located within pericentromeric regions, and 5,494 (73%) were in the chromosome arms (Figures 5.4c, 5.4d). This distribution significantly deviates from the expected distribution across chromosome arms, pericentromeres, and centromeres, which constitute 68.5 %, 21.9%, and 9.6% of the genome, respectively ($X^2 = 764.5$, $p < 0.001$) (Figure 5.4d). By contrast, only 34 simplified hyper-DMRs were identified in the 3 vs 3 comparison. Of these, 22 (69 %) were in pericentromeric regions, 1 (3%) in centromeric regions, and 11 (32%) in the chromosome arms (Figure 5.4d). Despite the low number of hypermethylated DMRs, this distribution is significantly different to the null expectation of a proportional distribution, with a higher-than-expected occurrence in pericentromeric regions (Fisher's Exact Test, $p < 0.001$) (Figure 5.4d). Thus, the distribution of consistent hypo- and hyper-DMRs along chromosomes supports the global cytosine methylation analysis (Figures 5.3b, 5.3c) and shows that ROS1-induced cytosine hypomethylation predominantly occurs in the chromosome arms and is statistically underrepresented in the (peri)centromeric regions (Figures 5.4c, 5.4d). However, the relatively low number of hypermethylated DMRs (Figure 5.4b, 5.4c) does not reflect the substantial and statistically significant increases in DNA methylation observed in the global cytosine methylation analyses (Figures 5.3b, 5.3c).

A possible cause of the discrepancy in (peri)centromeric DNA methylation between the global analyses (Figures 5.3b, 5.3c) and the 3 vs 3 DMR comparison (Figures 5.4a, 5.4b, 5.4c, 5.4d) is that observed (peri)centromeric hyper-methylation is more variable between replicates, resulting in an underestimation by this 3 vs 3 approach. For a DMR to be identified by this method, the specific region must exhibit similar changes in DNA methylation across all 3 biological replicates of E2-treated L5 seedlings. Since ROS1 is expected to target distinct regions in the genome (Tang et al., 2016), this analysis pipeline is highly effective for detecting regions of reduced DNA methylation. However, it is possible that the (peri)centromeric hypermethylation observed by the global analyses arises from a feedback response that is a

more stochastic and does not consistently occur at relatively small scales at which DMRs are called (tens to thousands of base pairs). Therefore, to identify more variable DMRs, each individual DMSO replicate was compared against each individual E2 replicate in L5 (Figure 5.4e). DMRs were then simplified for each context by merging overlapping DMRs between two or more comparisons into a single DMR. Numbers and patterns of hypo-DMRs in the 1 vs 1 comparison (Figure 5.4f) were similar to those observed in the 3 vs 3 comparison (Figure 5.4b), particularly for CG and allC contexts. However, approximately twice as many hypo-DMRs in the 1 vs 1 comparison were detected in CHG and CHH contexts, particularly at TEs (90% of all hyper-DMRs), suggesting that ROS1-induced hypomethylation is more variable in these regions (Figure 5.4f). Strikingly, 1,327 and 784 hypermethylated DMRs were detected in CHG and CHH contexts, respectively, with fewer detected in CG (264) and allC contexts (386), which is consistent with the global cytosine methylation analysis (Figures 5.3b, 5.3c). Merging all nuclear hypermethylated DMRs across all contexts resulted in 2,310 simplified hyper-DMRs, of which 192 are located in centromeres (8%), 1,615 in pericentromeres (70%), and 503 along chromosome arms (22%) (Figures 5.4g, 5.4h). This distribution is statistically different to the expected distribution of proportional DMRs across these regions (Figure 5.4h) (X^2 = 1140, p < 0.001). Thus, DNA hypermethylation in E2-treated *XVE:ROS1-YFP* L5 seedlings occurs within the pericentromeric and centromeric regions but there is more variation in the specific locations of these hyper-DMRs between biological replicates.

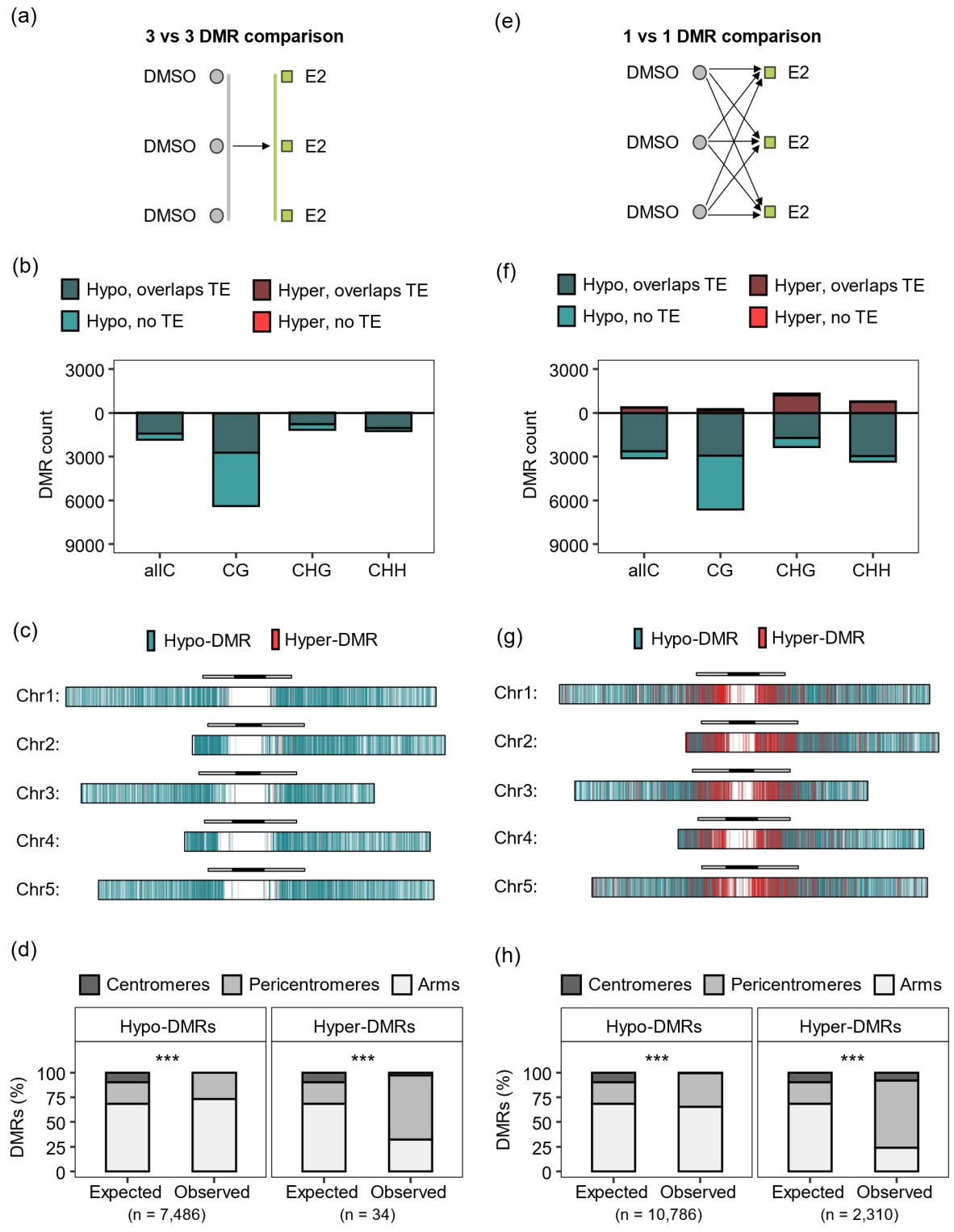


Figure 5.4 Analysis of differentially methylated regions (DMRs) in E2-treated *XVE:ROS1-YFP* L5 seedlings (L5 seedlings).
(a, e) Schematics of the DMR calling approach for identification of: (a) consistent DMRs detected by statistical comparison between all three DMSO-treated replicates (grey circles) and all three E2-treated replicates (green squares) of L5 ('3 vs 3 DMR comparison'); (b) variable DMRs detected by comparing individual DMSO-treated replicates against individual E2-treated replicates of L5 replicates ('1 vs 1 DMR comparison'). See this the Methods Section of this Chapter for details. (b, f) Counts of differentially methylated regions (DMRs) in E2-treated vs DMSO-treated L5 seedlings for allC, CG, CHG, and CHH

contexts in the (b) 3 vs 3 DMR comparison and (f) 1 vs 1 DMR comparison. Hypomethylated DMRs (hypo-DMRs; blue) and hypermethylated DMRs (hyper-DMRS) are shown below and above the y=0 line, respectively. DMRs overlapping with TEs are shown in dark blue (hypo-DMRs) and dark red (hyper-DMRs). (c, g) Distribution of merged hyper-DMRs and hypo-DMRs across the five nuclear chromosomes of the Col-CEN assembly by (c) the 3 vs 3 DMR comparison and (g) the 1 vs 1 DMR comparison. Large white rectangles represent chromosomes, and small rectangles above the chromosomes indicate (peri)centromeric regions, with centromeres shaded in dark grey and pericentromeres in light grey. Vertical lines within the chromosomes represent hypo-DMRs (blue) and hyper-DMRs (red). (d, h) Proportion of merged hypo-DMRs and hyper-DMRs within the centromeric regions (dark grey), pericentromeric regions (grey), and chromosome arms (light grey) detected by (d) the 3 vs 3 DMR comparison and (h) the 1 vs 1 DMR comparison. Expected distributions of DMRs reflect the proportionate genomic distribution across chromosome arms, pericentromeres, and centromeres. The total number of merged hypo-DMRs and hyper-DMRs for each comparison are shown below the x-axis labels. Asterisks indicate statistically significant differences between expected and observed distributions (Fisher's Exact test for the 3 vs 3 comparison of hyper-DMRs; Chi-squared test for all other comparisons; *** $p < 0.001$).

5.4.4 Reduction of CHG DNA Methylation Alone Does not Prolong ROS1-Driven Immune Memory Against *Pst-Lux*

As shown in Chapter 3, E2-treated *XVE:ROS1-YFP* seedlings effectively mounted an IR response against (hemi)biotrophic pathogens 48 hours after the final E2 treatment. However, the progeny of these E2-treated plants did not exhibit heritable induced resistance (h-IR) against *Hyaloperonospora arabidopsidis* (*Hpa*) isolate Waco9, suggesting that the immune memory underpinning this IR response was erased at some point between the final E2 treatment in the previous generation and the point of infection in the progeny. Interestingly, Furci et al. (2019) reported high level of heritable *Hpa* resistance in epigenetic recombinant inbred lines (epiRILs), which were generated by backcrossing F1 plants from a cross between wild-type Col-0 and the severely hypomethylated *ddm1-2* mutant into a wild-type background (Reinders et al., 2009; Johannes et al., 2009). Consequently, these epiRILs are genetically nearly identical but display a mosaic of stably inherited DNA hypomethylation patterns, predominantly within (peri)centromeric regions. Indeed, the epigenetic resistance reported by Furci et al. (2019) was quantitatively linked to four (peri)centromeric regions, demonstrating that highly stable epigenetic disease resistance in this population is controlled by hypomethylated (peri)centromeric DNA. In this regard, the stochastic hypermethylation observed in the (peri)centromeric regions of E2-treated *XVE:ROS1-YFP* seedlings (Figures 5.3c, 5.4f, 5.4g, 5.4h) could reflect a feedback response to ROS1 activity, resulting in progressive erasure of the immune memory. Accordingly, prevention of (peri)centromeric hypermethylation may prolong the immune memory against (hemi-)biotrophic pathogens established by increases in ROS1 activity.

Given the significant increases in DNA methylation at cytosines in CHG contexts within both centromeric and pericentromeric regions (Figure 5.3c), and the observation that most hypermethylated DMRs in the pericentromeres occurred in CHG contexts (Figure 5.4f), *XVE:ROS1-YFP* L5 was crossed with the *cmt3-11* mutant (SALK_148381; background Col-0) (Chan et al., 2006). CMT3 is a DNA methyltransferase responsible for maintaining CHG methylation, particularly in (peri)centromeric regions, meaning that *cmt3* mutants exhibit significant reductions in CHG DNA methylation in these regions (Lindroth et al., 2001; Bartee et al., 2001; Du et al., 2012; Ning et al., 2020) (Supplementary Figure 5.2a).

Thus, to determine whether CMT3 functions in the erasure of ROS1-induced immune memory, Col-0 and *cmt3-11* plants with and without the *XVE:ROS1-YFP* L5 insertion, were treated twice with E2 or DMSO (Figure 5.5a), and challenged with *Pseudomonas syringae* pv. *tomato* DC3000 *luxCDABE* (*Pst-Lux*) (Fan et al., 2008; Furci et al., 2021) at 1 day, 1 week, or 2 weeks after the final E2/DMSO treatment (Figure 5.5a). A two-way ANOVA indicated a statistically significant interaction between genotype and treatment on relative *Pst-Lux* bioluminescence 1 day ($F = 11.12$; $df = 3,424$; $p < 0.001$) and 1 week ($F = 11.99$; $df = 3,372$; $p < 0.001$) after the final E2 or DMSO treatment (Figure 5.5b). Post-hoc analysis revealed that relative *Pst-Lux* bioluminescence in E2-treated or DMSO-treated Col-0 and *cmt3-11* plants lacking the *XVE:ROS1-YFP* L5 insertion were not significantly different from each other, both 1 day and 1 week after the final E2/DMSO treatment (Tukey, $p_{adj} > 0.05$) (Figure 5.5b). In contrast, plants carrying the *XVE:ROS1-YFP* L5 insertion had reduced relative *Pst-Lux* bioluminescence when treated with E2 compared with DMSO-treated plants, both 1 day and 1 week after the final E2 or DMSO treatment. However, the reduction in *Pst-Lux* bioluminescence was not significantly different between E2-treated Col-0 and *cmt3-11* plants containing the *XVE:ROS1-YFP* L5 insertion at either time point (Tukey, $p_{adj} > 0.05$). Furthermore, 2 weeks after the final E2/DMSO treatments, there was no statistically significant effect of genotype ($F = 2.20$; $df = 3,374$; $p > 0.05$) or interaction between genotype and treatment ($F = 1.09$; $df = 3,374$; $p > 0.05$) on *Pst-Lux* bioluminescence. However, a statistically significant effect of treatment alone was observed ($F = 1.09$; $df = 3,374$; $p = 0.03$) (Figure 5.5b), which might be driven by a subtle reduction in *Pst-Lux* bioluminescence across E2-treated plants containing the *XVE:ROS1-YFP* L5 insertion. Nevertheless, the *cmt3-11* mutation did not enhance IR phenotypes against *Pst-Lux* following the establishment of ROS1-driven immune memory at any time point. Furthermore, this mutation did not prolong the durability of ROS1-driven immune memory against *Pst-Lux* beyond 1-2 weeks (Figure 5.5b). Thus, CMT3-dependent CHG methylation does not account for the erasure of ROS1-driven immune memory against *Pst-Lux* despite the *cmt3* mutation having a strong influence on (peri)centromeric DNA methylation levels (Supplementary Figure 5.2a). This suggests that other mechanisms might drive

(peri)centromeric hypermethylation in E2-treated *XVE:ROS1-YFP* lines. Therefore, to address whether DNA (re)methylation has any role in erasing ROS1-driven immune memory, a broad approach was employed in the following section using the general DNA methyltransferase inhibitor 5-Azacytidine (5-Aza) (Pecinka and Liu, 2014; Griffin et al., 2016).

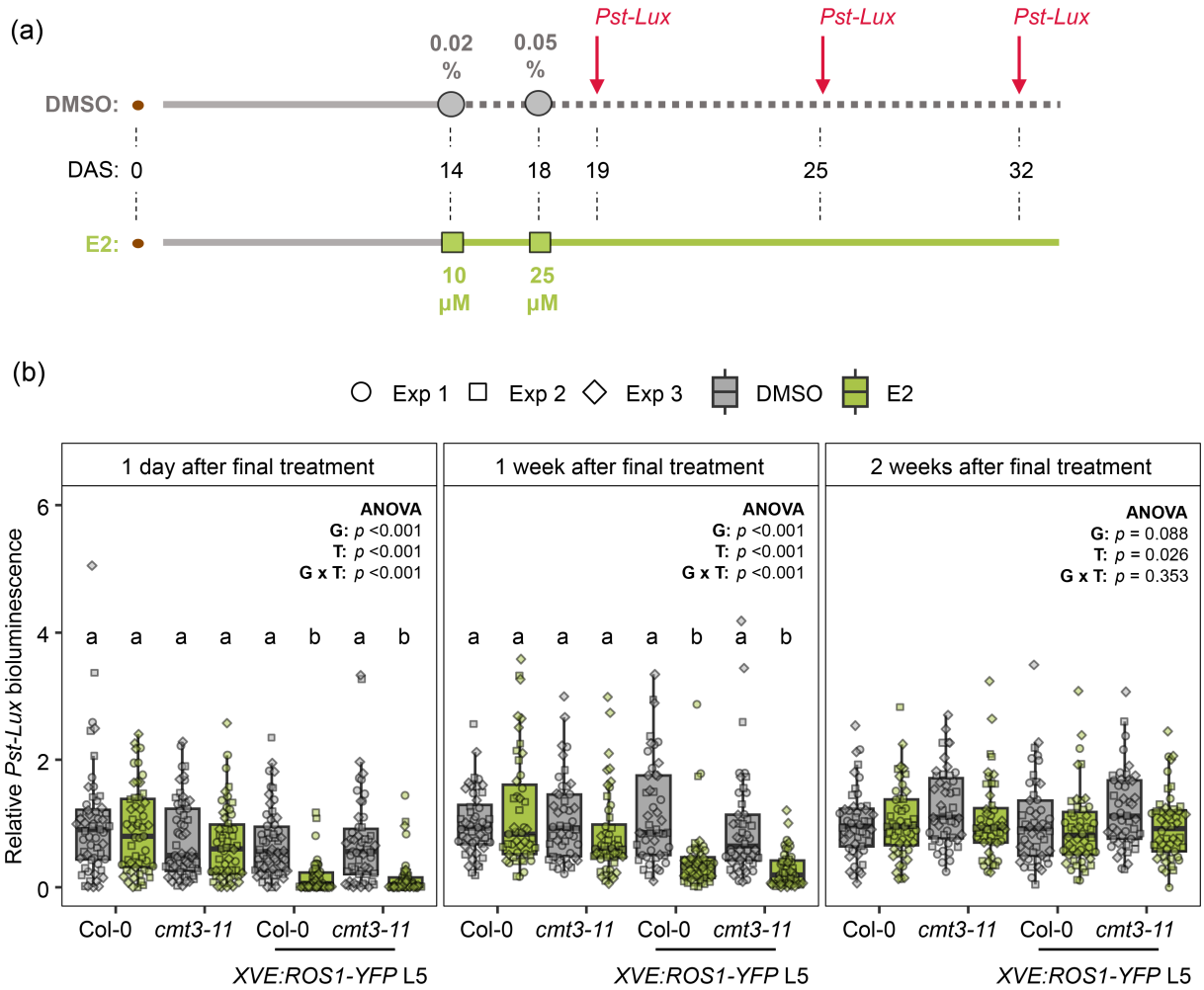


Figure 5.5 Duration of ROS1-driven immune memory, assessed by the strength of induced resistance (IR) against *Pseudomonas syringae* pv. *tomato* DC3000 *luxCDABE* (*Pst-Lux*) in the Col-0 and *cmt3-11* genetic backgrounds.
(a) Schematic of the experimental design. Col-0 and *cmt3-11* genotypes containing or not containing the *XVE:ROS1-YFP* L5 insertion were spray-treated twice with DMSO (0.02% and 0.05%) or E2 (10 µM and 25 µM), at 14 and 18 days after sowing (DAS). All plants were then inoculated with *Pst-Lux* at 1 day, 1 week, or 2 weeks after the final E2/DMSO treatment. (b) Quantification of *Pst-Lux* bioluminescence at 2-3 days post-inoculation (dpi) at 1 day, 1 week, and 2 weeks after the final DMSO (grey) or E2 (green) treatment. Data from three independent experiments (Exp1, Exp2, Exp3) were normalized to the mean *Pst-Lux* bioluminescence values of DMSO-treated Col-0 plants at each time point and subsequently pooled for analysis. Two-way ANOVA was used to assess the significance of genotype (G), treatment (T), and their interaction (G x T) on relative *Pst-Lux* bioluminescence. Post-hoc analysis was conducted for timepoints at which there was a significant interaction ($p < 0.05$). Treatment-genotype combinations sharing the same letter are not significantly different from each other (Tukey post-hoc, $p_{adj} < 0.05$). For details about *Pst-Lux* inoculation and bioluminescence quantification, see Section 2.5.

5.4.5 Chemical Inhibition of DNA Methyltransferases Using 5-Azacytidine Prolongs ROS1-Driven Immune Memory Against *Pst-Lux*

While blocking CMT3-dependent CHG methylation did not enhance the strength or durability of ROS1-driven immune memory (Figure 5.5), it is plausible that a combination of DNA methyltransferases is responsible for the erasure of this memory. Indeed, apart from CHG hypermethylation, statistically significant increases in CG DNA methylation were detected in the centromeres (Figure 5.4c). Furthermore, the chromosome-level metaplots revealed progressively reduced hypomethylation at all sequence contexts towards the centromeres, with a trend to be hypermethylated nearest the midpoint of pericentromeres (Figure 5.3b). This DNA hypermethylation response was particularly clear from the 1 vs 1 comparisons, which capture more variable changes in DNA methylation between E2- and DMSO-treated L5 plants, revealing 386, 264, 1,327, and 784 hyper-DMRs, occurring in allC, CG, CHG, and CHH contexts, respectively (Figure 5.4f), which occurred almost exclusively in (peri)centromeric regions (Figure 5.4g, 5.4h). This indeed suggests that multiple DNA methyltransferases play a role in the (re)methylation of cytosines in (peri)centromeric regions after the establishment of immune memory through enhanced ROS1 activity.

To test whether a (re)methylation response in all sequence contexts contributes to the erasure of ROS1-driven immune memory, ROS1 was induced by E2 with or without the DNA methyltransferase inhibitor 5-Azacytidine (5-Aza) (Pecinka and Liu, 2014; Griffin et al., 2016) (Figure 5.6a). In Arabidopsis, treatment with 5-Aza results in dose-dependent reductions in DNA methylation throughout the genome, with a pronounced effect on (peri)centromeric regions, particularly in CHH contexts, even at lower concentrations (Griffin et al., 2016) (Supplementary Figure 5.2b). Five treatments with 100 µM 5-Aza or H_2O were performed over a 2-week period, starting 1 week before the first E2/DMSO treatment, and ending 1 day before the final E2/DMSO treatment (Figure 5.6a). Plants were then challenged with *Pst-Lux* either 1 day, 1 week, or 2 weeks after the final E2/DMSO treatment. Aside from a mild reduction in growth, plants treated with 5-Aza had no obvious physiological defects (Supplementary Figure 5.3).

There was a significant effect of the interaction between (epi)genotype group (combination of genotype and 5-Aza/H_2O treatment) and treatment (E2 or DMSO) on relative *Pst-Lux* bioluminescence at all measured timepoints: 1 day ($F = 6.97$; $df = 3,404$; $p < 0.001$), 1 week ($F = 13.00$; $df = 3,346$; $p < 0.001$), and 2 weeks ($F = 8.59$; $df = 3,346$; $p < 0.001$) after the final E2/DMSO treatment (Figure 5.6b). Post-hoc analysis revealed that E2-treated *XVE:ROS1-YFP* L5 plants, both with and without additional 5-Aza treatment, were statistically different to their DMSO-treated counterparts at 1 day and 1 week after the final E2 treatment

(Figure 5.6b; Tukey, $p_{adj} < 0.05$). Although pre-treatments of Col-0 and *XVE:ROS1-YFP* plants with 5-Aza had little to no impact on *Pst-Lux* colonisation (Figure 5.6b), combining E2 treatment of *XVE:ROS1-YFP* plants with 5-Aza resulted in statistically lower levels of *Pst-Lux* colonisation than E2 treatment alone, which was apparent at 1 day and 1 week after the final E2 treatment (Figure 5.6b) (Tukey, $p_{adj} < 0.05$). Thus, there is not a direct additive effect of 5-Aza pre-treatment and ROS1 induction on *Pst-Lux* resistance, but rather a synergistic effect of both treatments combined. Notably, ROS1-IR was observed 2 weeks after the final DMSO/E2 treatment in *XVE:ROS1-YFP* L5 plants treated with 5-Aza and E2, but not in any other treatment groups (Figure 5.6b). This demonstrates that chemical inhibition of methyltransferases not only prolongs ROS1-driven immune memory against *Pst-Lux* but also enhances IR responses linked to this memory at time points where immune memory persists in the absence of DNA methyltransferase inhibition.

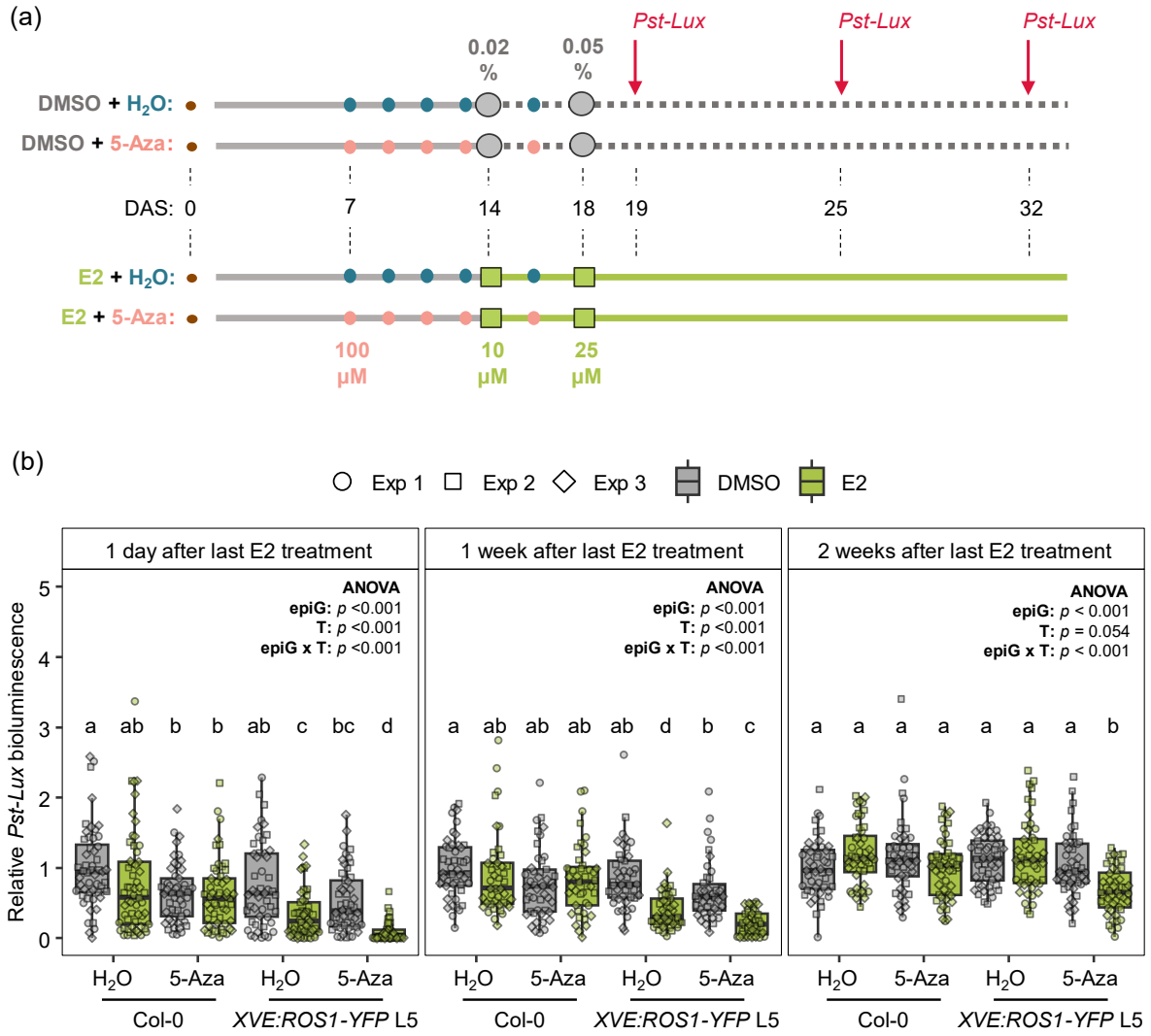


Figure 5.6 Duration of ROS1-driven immune memory, assessed by the strength of induced resistance (IR) against *Pseudomonas syringae* pv. *tomato* DC3000 *luxCDABE* (*Pst-Lux*) in the presence and absence of the chemical DNA methyltransferase inhibitor 5-Azacytidine (5-Aza).
(a) Schematic of the experimental design. Droplets of 100 µM 5-Aza (pink) or H_2O (blue; mock) were applied to the leaves of Col-0 seedlings, both with and without the *XVE:ROS1-YFP L5* insertion, every 2-3 days between 7 and 17 days after sowing (DAS), resulting in a total of 5 successive 5-Aza treatments. Seedlings were also spray-treated twice with DMSO (0.02% and 0.05%) or E2 (10 µM and 25 µM), at 14 and 18 DAS. Plants were then challenged with *Pst-Lux* at 1 day, 1 week, or 2 weeks after the final E2/DMSO treatment. (b) Quantification of *Pst-Lux* bioluminescence at 2-3 days post-challenge and at 1 day, 1 week, and 2 weeks after the final DMSO (grey) or E2 (green) treatment. At each time point, data from three independent experiments (Exp1, Exp2, Exp3) were normalised to the average *Pst-Lux* bioluminescence values of H_2O- and DMSO-treated Col-0 plants and pooled. Two-way ANOVA was used to assess the statistical significance of (epi)genotype group (epiG), treatment (T), and their interaction (epiG x T) on relative *Pst-Lux* bioluminescence. The (epi)genotype group was defined as the combination of the genotype and 5-Aza/H_2O pre-treatment. For each timepoint, treatments sharing the same letter are not significantly different from each other (Tukey, $p_{adj} < 0.05$). For details about *Pst-Lux* inoculation and bioluminescence quantification, see Section 2.5.

5.5 Discussion

5.5.1 ROS1-Dependent Regulation of Immune Memory Through (Peri)centromeric DNA Methylation

Stably inherited epigenetic disease resistance in Arabidopsis has been linked to hypomethylated (peri)centromeric regions of Arabidopsis (Furci et al., 2019). In this Chapter, long-read ONT-seq data and the Col-CEN genome assembly (Naish et al., 2021) were used to investigate the link between immune memory driven by increased activity of ROS1 and DNA methylation in (peri)centromeric regions of Arabidopsis. Surprisingly, while ectopic induction of the DNA demethylase ROS1 in E2-treated *XVE:ROS1-YFP* L5 seedlings caused hypomethylation of cytosines along chromosome arms, a contrasting hypermethylation response was observed within the (peri)centromeric regions (Figures 5.3, 5.4). Furthermore, whilst all biological replicates exhibited similar average increases in cytosine methylation across the (peri)centromeres (Figures 5.3b, 5.3c), the specific features affected within the (peri)centromeres, as quantified by DMR analyses, varied between replicates (Figure 5.4). Thus, the hypermethylation response within the (peri)centromeres is more stochastic than the hypomethylation response in the chromosome arms. It is plausible that this variable DNA hypermethylation reflects a counter response to ROS1 activity, which acts to prevent excessive DNA hypomethylation in TE-rich (peri)centromeres. As *ddm1*-induced (peri)centromeric hypomethylation has been shown induce heritable disease resistance (López Sánchez et al., 2016; Furci et al., 2019), DNA (re)methylation in the (peri)centromeres may drive the erasure of immune memory established by ROS1 activity. Experimental support for this hypothesis comes from the finding that co-application of E2 with the DNA methyltransferase inhibitor 5-Aza, which preferentially inhibits maintenance of DNA methylation in (peri)centromeric regions (Supplementary Figure 5.2) (Griffin et al., 2016), resulted in enhanced ROS1-IR and prolonged ROS1-driven immune memory against *Pst-Lux* (Figure 5.6). Hence, DNA methylation has a role both in the establishment and erasure of immune memory, which is potentially coordinated by the (peri)centromeric regions. In the next Section, I expand on several *trans*-regulatory mechanisms proposed in the literature, as introduced in Chapter 1 (Figure 1.3), which may explain how DNA methylation at (peri)centromeric regions influences the transcriptional activity of genes involved in immune memory as observed in this Chapter. Building on these models, and the observation in Chapter 4 that small RNAs accumulate in the (peri)centromeric regions after ROS1 induction (Figure 4.5), I propose a model in which epigenetic regulation of pericentromeric regions following

increased ROS1 activity plays a key role in balancing the maintenance and erasure of plant immune memory.

5.5.2 Epigenetic *Trans*-Regulation Of Chromosome Arm Genes by (Peri)centromeric Regions

5.5.2.1 Roles of Small RNAs

The functional role of pericentromeric and centromeric regions in regulating gene transcription and stress responses is an emerging but challenging area of research. As discussed above, hypomethylation of (peri)centromeric regions in Arabidopsis causes highly stable resistance to the biotrophic pathogen *Hpa* (Furci et al., 2019). Interestingly, the epigenetic quantitative trait loci (epiQTLs) identified in this paper did not include any hypomethylated regions proximal to known defence genes, suggesting a broader *trans* regulatory role of (peri)centromeres in regulating plant immune responses. Whilst many models of *trans* regulation have been proposed in the regulation of plant immune responses (Chapter 1; Figure 1.3) (Wilkinson et al., 2019; Furci et al., 2019; Hannan Parker et al., 2022), few have been conclusively characterised. Nevertheless, significant advancements have been made in understanding the role of sRNAs and long-range chromatin interactions in the *trans* regulation of genes. Firstly, Liu et al. (2018) characterised a positive role for small RNAs (sRNAs) in regulating gene expression by associating with nuclear-localised ARGUONATE 1 (AGO1). Their study demonstrated that sRNA-AGO1 complexes can bind to genic regions and promote their expression; a process that is enhanced under stress-stimulating conditions (Liu et al., 2018). Therefore, sRNAs derived from one region in the genome have the potential to influence expression of genes in another part of the genome given that they have similar nucleotide sequences (Figure 1.3). For instance, Cambiagno et al. (2018) highlighted that (peri)centromeric sRNAs had sequence complementarity to well characterised defence genes, and thus, could have a role in regulating their expression following *Pst* infection. Furthermore, the establishment of long-term immune memory against *Spodoptera littoralis* is associated with the hypomethylation of euchromatic *ATREP2* TEs which produce sRNAs that can associate with nuclear AGO1 (Wilkinson et al., 2023). These sRNAs shared significant sequence complementarity with genes which were associated with the IR response (Wilkinson et al., 2023). These models highlight a positive role of sRNAs in controlling stress-responsive genes.

Despite this, sRNAs are more commonly associated with their roles in (post)transcriptional gene silencing than with transcriptional activation. For instance, 21 nt

sRNAs loaded onto cytoplasmic AGO1 can direct transcript degradation or suppress translation (Zhan and Meyers, 2023). Additionally, 24 nt sRNAs guide RdDM machinery, which is known to suppress the transcriptional activity of genes, particularly when associated with gene promoters (Halter et al., 2021; Zhan and Meyers, 2023). Recent work investigating the remethylation of hypomethylated TEs in epiRIL populations demonstrated that the majority of remethylation in promoter regions is driven by *trans*-acting sRNAs derived from (peri)centromeric regions (Baduel et al., 2024). Furthermore, this remethylation of promoter regions was associated with transcriptional (re)silencing of the associated gene (Baduel et al., 2024). Therefore, sRNAs derived from (peri)centromeric regions have the potential to both positively and negatively regulate the transcription of chromosome arm genes in *trans*.

5.5.2.2 Roles of Higher Order Chromatin Structure

Regions of the genome with similar epigenetic states, such as (peri)centromeric regions, tend to form strong chromatin interactions within- and between-chromosomes (S. Feng et al., 2014; Nützmann et al., 2020; Yadav et al., 2022; Yueying Zhang et al., 2024). However, mutations in genes involved in the deposition of DNA methylation, such as *DDM1*, *MET1*, and *MORC6* (Kankel et al., 2003; Zemach et al., 2013; Xue et al., 2021), can cause spurious interactions between (peri)centromeric regions and gene-rich chromosome arms (S. Feng et al., 2014). This demonstrates that DNA methylation plays a key role in shaping chromatin structure and thus, has a major influence on the long-range chromatin interactions between (peri)centromeric regions and chromosome arms. Interestingly, Nützmann et al (2020) found that the thalianol gene cluster in Arabidopsis, which is transcriptionally active in roots but inactive in shoots, exhibits distinct chromatin interactions between tissues. The gene cluster topologically associates with chromosome arms in roots, where it is expressed, but forms interactions with pericentromeric regions in shoots, where it is repressed. Therefore, (peri)centromeric regions may contribute to gene repression by encouraging the formation of a heterochromatic state through long-range chromatin interactions, which results in transcriptional silencing.

These genomic interactions could also shift in response to environmental stress, potentially altering gene expression to adapt to changing conditions. For instance, heat stress in Arabidopsis causes a reduction in chromatin interactions within- and between-pericentromeric regions and other heterochromatic regions of the genome and a complementary increase in the interactions between pericentromeric regions and chromosome arms (Sun et al., 2020). Similarly, in response to cold stress, genes that exhibited primed transcriptional regulation following a mild cold pretreatment were found to form extensive promoter-promoter interaction networks. The authors suggested that these closely

interacting promoters facilitate rapid co-transcriptional activation by shared transcription factors (Yueying Zhang et al., 2024). However, while these examples all demonstrate how (peri)centromeric regions can change their interaction with chromosome arms, its consequence on transcriptional activity of genes within the interacting region, if any, remains unclear. Combining chromatin conformation capture technologies, such as Hi-C and CAP-C (Ouyang et al., 2021; Yueying Zhang et al., 2024), with the *XVE:ROS1-YFP* system could provide deeper insights into the regulatory roles of long-range chromatin interactions between (peri)centromeric regions and chromosome arms in the context of immune memory.

5.5.3 A Potential Role of RdDM in Reestablishing (Peri)Centromeric DNA Methylation Following ROS1-Dependent Demethylation

In Chapter 4, a discrepancy in the distribution of 24 nt sRNAs between chromosome arms and pericentromeric regions was observed in E2-treated *XVE:ROS1-YFP* lines (Figure 4.5f). This pattern mirrors the DNA methylation patterns observed in these lines (both in Chapter 4 and in this Chapter), whereby hypomethylation occurs along the chromosome arms and contrasting DNA hypermethylation occurs within the (peri)centromeric regions. While the sRNA data in Chapter 4 was mapped to the TAIR10 genome and the ONT-seq data in this chapter was mapped to the Col-CEN genome, these two genomes only differ in their resolution of centromeric regions(Figure 5.2a). Taken together, these results suggest that active DNA demethylation by ROS1 triggers a counteracting response of sRNA-dependent hypermethylation at some genomic loci. This response appears to be either exclusive to (peri)centromeric regions or tends to occur more rapidly there compared to chromosomal arms. In support of this rapid (re)methylation response, Cambiagno et al. (2018) found that TEs from the (peri)centromeric-associated families *ATHILA6A* and *TSI* were transcriptionally downregulated in response to *Pst* infection, accompanied by hyperaccumulation of 24 nt sRNAs. In contrast, TEs from the families *ATLANTYS2A* and *TA11*, which are typically embedded in chromosome arms were actively transcribed (Cambiagno et al., 2018). Further analysis revealed that *TSI* (peri)centromeric TEs experienced an initial boost in expression at early infection time points (3-7 hpi) before being rapidly supressed at 24 hpi (Cambiagno et al., 2018). Conversely, euchromatic-associated TEs activated at 24 hpi did not remain induced 5 dpi in another study (Dowen et al., 2012). Thus, *Pst*-induced activation of TEs is quickly reversed in (peri)centromeric regions compared to TEs in chromosome arms.

In Arabidopsis, RdDM activity, driven by the *de novo* DNA methyltransferases DRM1 and DRM2, is typically confined to the euchromatic chromosome arms, whilst DNA methyltransferase activity by CMT3, CMT2, and MET1, maintain DNA methylation in (peri)centromeric regions (Zemach et al., 2013; Stroud et al., 2014, 2014; Harris et al., 2024).

However, in the event of extensive hypomethylation of (peri)centromeric TEs, such as is caused by the *ddm1* mutation, RdDM activity is drastically increased in these regions, with the strongest effect observed at actively transcribed TEs (Zemach et al., 2013; Panda and Slotkin, 2020; Shimada et al., 2024) (Supplementary Figure 5.4a). Notably, a greater proportion of (peri)centromeric TEs become transcriptionally active in hypomethylated mutants compared to TEs in the chromosome arms, making these regions particularly sensitive to losses of DNA methylation (Panda and Slotkin, 2020) (Supplementary Figure 5.4b). This heightened sensitivity enables and necessitates a much faster recovery of DNA methylation via RdDM pathways following episodes of hypomethylation. Thus, the presence and insertion of active TEs in (peri)centromeric regions is crucial for maintaining a stable epigenetic state and genomic stability (Shimada et al., 2024).

The initiation of RdDM at hypo-methylated (peri)centromeric regions typically begins with the transcription and activation of TEs by Pol II. These TE transcripts can then be processed into small RNAs (sRNAs) of 21, 22, or 24 nt in length, which facilitate the initiation or 'first-round' of RdDM in a sRNA-size-independent manner (Panda et al., 2016; Cuerda-Gil and Slotkin, 2016; Sigman et al., 2021). This is exemplified in *ddm1* mutants, whereby strongly activated TEs have a drastic increase in Pol II-dependent RdDM pathways, including the DCL3-dependent and RDR6-dependent pathways (Panda et al., 2016; Cuerda-Gil and Slotkin, 2016; Panda and Slotkin, 2020) (Supplementary Figure 5.4a). However, for RdDM to be initiated, a complementary region of the genome must be transcribed by Pol II (Sigman et al., 2021). As (peri)centromeric regions of Arabidopsis contain a larger proportion of TEs that are transcriptionally active (Panda and Slotkin, 2020) (Supplementary Figure 5.4b), and because these regions have a high density of closely related TEs (Quesneville, 2020), RdDM initiation is more likely to occur at (peri)centromeric regions compared to chromosome arms. Thus, this could explain why (peri)centromeric regions are more rapidly remethylated in E2-treated *XVE:ROS1-YFP* plants compared to chromosome arms.

Following RdDM initiation, methylated cytosines are able to be recognised and bound by SUVH2 and SUVH9, which recruit Pol V and continually promote the canonical methylation-dependent RdDM pathway (Erdmann and Picard, 2020; Sigman et al., 2021). Likewise, heterochromatic marks mechanistically linked to DNA methylation enable further recruitment of Pol IV, enabling for sRNA biogenesis in the absence of Pol II transcription (Matzke and Mosher, 2014; Erdmann and Picard, 2020). Further amplification of RdDM can also occur as DNA methylation spreads into adjacent regions (Henderson and Jacobsen, 2008; Ahmed et al., 2011; Liu et al., 2023) and sRNAs bind to other genomic regions with sequence complementarity (Sigman et al., 2021). Once established, DNA methylation can then continue

to be maintained by RdDM-independent mechanisms, such as DNA methyltransferase activity (Erdmann and Picard, 2020; Sigman et al., 2021).

Evidence of TE-dependent RdDM activation in E2-treated *XVE:ROS1-YFP* lines exists in the transcriptome and sRNA-ome data presented in Chapter 4. In total, 75 TEs were significantly upregulated in E2-treated *XVE:ROS1-YFP* L5 or L7 (Figure 4.4b; Supplementary Data 4.4), of which 48 (64%) were within (peri)centromeric regions. Of these TEs, 24 (50%) overlapped a significantly upregulated sRNA cluster (Supplementary Data 4.9). To conclusively test the role of RdDM in coordinating the remethylation of pericentromeric regions following active DNA demethylation by ROS1, the *XVE:ROS1-YFP* construct should be introduced into an RdDM-deficient mutant such as *nrpe1-11* or *rdr1;2;6* (Stroud et al., 2013; Shimada et al., 2024). If the response is not dependent on RdDM, a similar hypermethylation response should still occur following E2 treatment, even in the absence of core RdDM machinery.

5.5.4 Potential Mechanisms Driving the Redistribution Of RdDM Into (Peri)centromeric Regions

If RdDM is involved in the rapid remethylation of (peri)centromeres of E2-treated *XVE:ROS1-YFP* lines, the mechanisms driving its redistribution from the chromosome arms to (peri)centromeric regions during the initial epigenetic shock by ROS1 still remains a puzzle. In Chapter 4, the possible role of CLSY3, which facilitates (peri)centromeric Pol IV recruitment, was highlighted – noting that E2-treated *XVE:ROS1-YFP* lines have on average 7.3x more expression of *CLSY3* relative to DMSO-treated lines (Supplementary Data 4.4). More recent work by (Harris et al., 2024) demonstrated that the linker histone H1 has important roles in restricting RdDM to euchromatic regions under basal conditions, and thus may be involved in redistributing RdDM in E2-treated *XVE:ROS1-YFP* lines. H1 is normally enriched in (peri)centromeric regions and has a negative association with Pol V, thereby limiting RdDM activity in these regions (Zemach et al., 2013; Harris et al., 2024). As a result, mutation of *h1* leads to increased occupancy of the Pol V subunit NRPE1 across the genome, particularly in (peri)centromeric regions. Consequently, *h1* mutants have drastic increases in 24 nt sRNA production at (peri)centromeric TEs, and show concurrent increases in non-CG DNA methylation at these regions (Papareddy et al., 2020; Choi et al., 2021; Harris et al., 2024). Furthermore, TEs that become hypermethylated in *h1* mutants show a strong overlap with sRNA clusters associated with CLSY3 and CLSY4, suggesting a potential mechanistic link between H1 and CLSY3/4 (Choi et al., 2021). As the 24 nt sRNA distributions and DNA methylome patterns of *h1* mutants closely mimic those observed in E2-treated *XVE:ROS1-YFP* lines, H1 may also have an important role in the establishment and eventual erasure of

ROS1-dependent immune memory. Interestingly, *h1* mutants have previously been described to have enhanced resistance against *Pst* but are unable to form immune memory following treatment with the bacterial peptide flg22 in Arabidopsis (Sheikh et al., 2023). This further indicates a key role of H1 in regulating plant immune memory. Despite this, none of the three genes encoding H1 variants were found to be differentially expressed in E2-treated lines (Supplementary Data 4.4), and no previous studies have established a link between ROS1 and H1. Interestingly, H1 has been found to interact *in vitro* with the closely related DNA demethylase DEMETER (DME) (Rea et al., 2012) and may therefore interact with ROS1 too. To address this, ROS1-YFP protein interactors need to be identified under this E2 treatment regime using co-immunoprecipitation (Co-IP) or proximity labelling techniques, such as TurboID (Mair et al., 2019). Furthermore, introducing the *XVE:ROS1-YFP* construct into the mutants *h1* and *clsy3* would help elucidate their role in coordinating the DNA hypermethylation and 24 nt sRNA hyperaccumulation response in (peri)centromeric regions.

5.5.5 A Proposed Multifunctional Role of (Peri)centromeric sRNAs in Establishing, Maintaining, and Resetting Plant Immune Memory

Integrating the discussion points above, a surge in sRNA production from (peri)centromeric regions following active DNA demethylation of TEs by ROS1 may serve three key functions in regulating plant immune memory: (**i**) sRNAs derived from activated (peri)centromeric TEs ensure the rapid reinitiation and reestablishment of DNA methylation and transcriptional silencing of TEs in these regions to maintain genomic stability; (**ii**) some TE-derived 21 nt sRNAs share sequence complementarity with immune-responsive genes and thus positively regulate gene transcription in *trans* via association with nuclear AGO1; (**iii**) (re)methylation of DNA at (peri)centromeric TEs drives strong recruitment of canonical RdDM machinery, leading to a large increase in Pol IV-derived 24 nt sRNAs, which can guide gene-repressive RdDM to complementary regions of the distal chromosome arms, such as the promoters of *cis*-regulated immune-responsive genes with embedded TEs (Halter et al., 2021; Baduel et al., 2024) (Figure 5.7). Under this model, inhibition of DNA methyltransferases that silence TEs in the (peri)centromeric regions by 5-Aza (Zemach et al., 2013; Griffin et al., 2016) (Figure 5.7a) would enhance the transcriptional activation associated with increased ROS1 activity (Figure 5.7b). Thus, the production of 21 nt sRNAs derived from degraded TEs would also increase (Panda and Slotkin, 2020) (Supplementary Figure 5.4), which could enhance the *trans* activation of chromosome arm genes via association with nuclear localised AGO1 (Figure 5.7c) (Liu et al., 2018; Wilkinson et al., 2023). This could explain the synergistic enhancement of ROS1-IR caused by E2-induced ROS1 activity and 5-Aza treatment (Figure 5.6b). Furthermore, as 5-Aza would also block the activity of RdDM-associated DNA

methyltransferases DRM1 and DRM2 (Zhong et al., 2014; Griffin et al., 2016) this would slow the RdDM-associated (re)methylation of (peri)centromeric regions (Figure 5.7d) and consequently, the RdDM-mediated (re)methylation of gene promoter regions with sequence complementarity to the (peri)centromeres (Baduel et al., 2024) (Figure 5.7e). Consequently, the erasure of immune memory would be antagonised under 5-Aza treatments, as demonstrated in Figure 5.6b.

Therefore, these results, and the model outlined below, would implicate (peri)centromeric-derived sRNAs in both the establishment and erasure of immune memory in plants (Figure 5.7). Future work integrating the use of epiRILs and priming treatments that establish immune memory in Arabidopsis would provide valuable insights into the role of (peri)centromeric DNA (hypo)methylation in the establishment, maintenance, and/or erasure of plant immune memory (Figure 5.7). As epiRILs vary in their extent of (peri)centromeric DNA methylation, the strength and durability of IR established in each line could be linked to the DNA methylation status of the (per)centromeric regions. This would also help determine whether the increased basal resistance observed in some epiRILs (Furci et al., 2019) is determined by the same epigenomic features that regulate the establishment, maintenance, and/or erasure of plant immune memory. Furthermore, introducing the *XVE:ROS1-YFP* construct into a Pol V-deficient mutant would help distinguish the effects of de novo DNA methyltransferases involved in RdDM pathways (DRM1/2) from those of maintenance DNA methyltransferases, such as CMT2, CMT3, and MET1. Furthermore, introducing this construct into Pol IV-deficient mutants could help differentiate the presumed memory-stimulating effects of 21 nt sRNAs, derived from Pol II-dependent RdDM (Figure 5.7c), from the presumed memory-erasing effects of 24 nt sRNAs associated with Pol IV-dependent RdDM (Figure 5.7d).

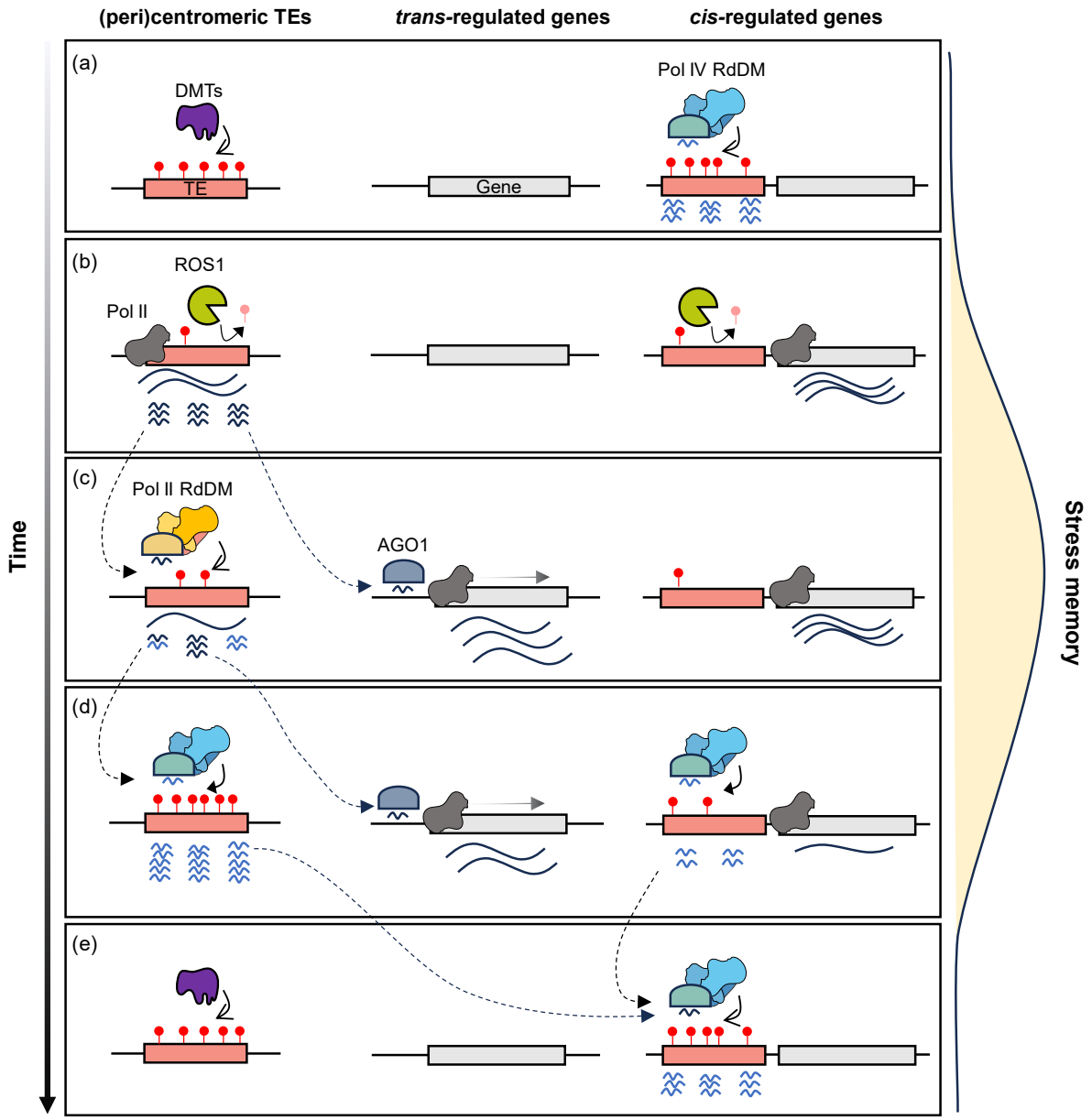
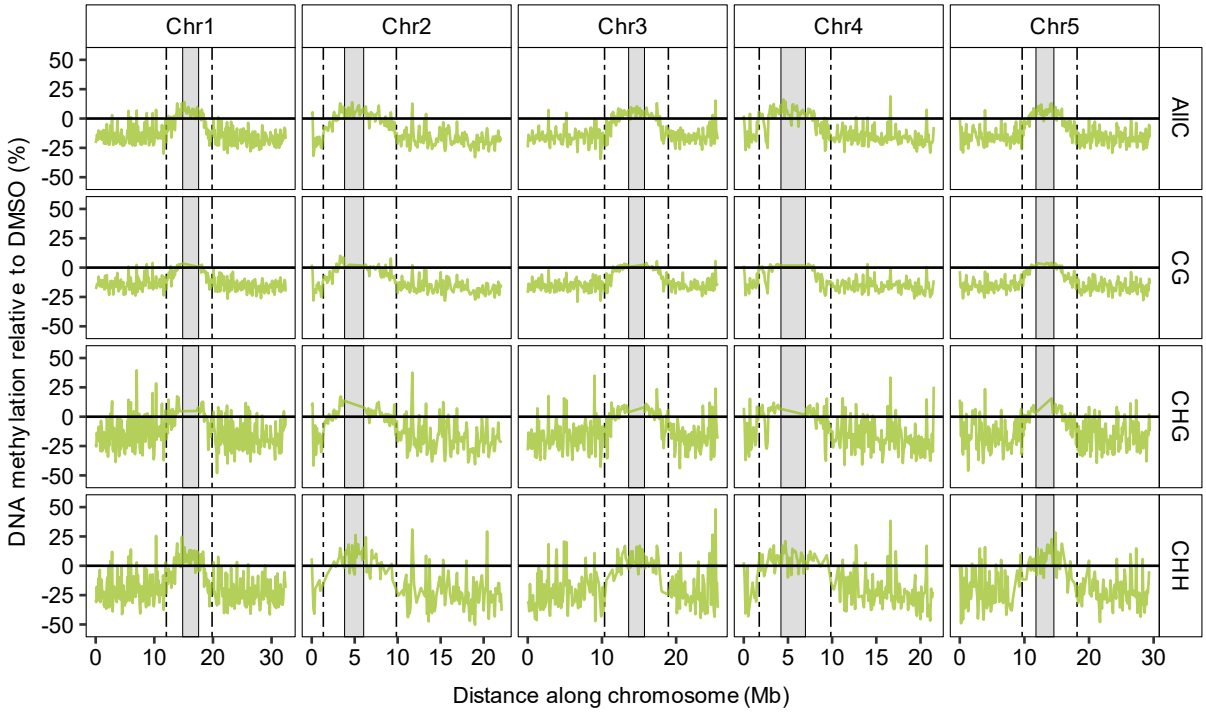


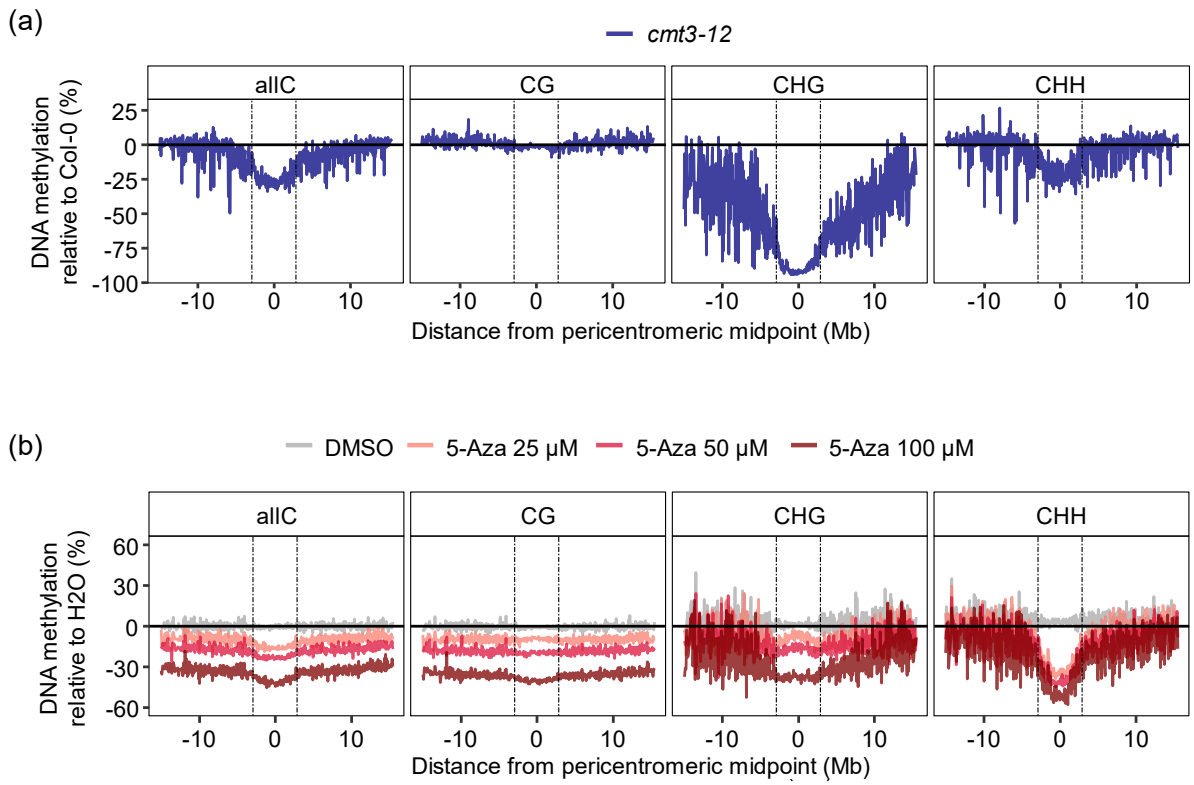
Figure 5.7 Proposed regulatory roles of (peri)centromeric TE-derived sRNAs in plant immune memory. (a) Naïve state: DNA methylation at transposable elements (TEs) in (peri)centromeric regions (left) and chromosomal arms (right) is primarily maintained by DDM1-associated DNA methyltransferases (DMTs) and the RNA polymerase IV-dependent RNA-directed DNA methylation (Pol IV-dependent RdDM) pathway, respectively. (b) Onset of immune memory: Active DNA demethylation by ROS1 at transposable elements (TEs) facilitates the transcription of pericentromeric TEs by Pol II (Panda and Slotkin, 2020) (Chapter 4; Figure 4.4b) (Supplementary Figure 5.3b) and promotes the expression of cis-regulated genes (Chapter 4; Figure 4.6) (Halter et al., 2021). TE transcripts are processed into 21, 22, and 24 nt sRNAs (Sigman et al., 2021) (black squiggles). (c) Onset/maintenance of immune memory: TE transcript-derived sRNAs initiate the remethylation of other transcribed TEs with sequence complementarity via Pol II-dependent RdDM (yellow complex) pathways (Cuerda-Gil and Slotkin, 2016; Sigman et al., 2021). Some of the 21-22 nt sRNAs can also associate with nuclear AGO1 to promote gene expression in *trans* (Liu et al., 2018). (d) Erasure of immune memory: Propagation of DNA methylation in (peri)centromeres drives strong recruitment of Pol V and canonical Pol IV-dependent RdDM pathways (blue complex) (Sigman et al. 2021). As a result, TEs are transcriptionally silenced, and the biogenesis of Pol IV-derived 24 nt sRNAs (blue squiggles) increases. Since nuclear AGO1 predominantly associates with 21-22 nt sRNAs, the *trans*-regulation of genes decreases. Additionally,

the increased abundance of Pol IV-derived 24 nt sRNAs promote the reactivation/redistribution of Pol IV-dependent RdDM at distal regions with sequence complementarity, such as the promoters of *cis*-regulated genes that contain TEs (Baduel et al., 2024). (e) Return to naïve state: As RdDM transitions back to DMT-dependent maintenance of DNA methylation in (peri)centromeric regions, sRNA biogenesis from these regions declines, leading to the termination of *trans*-regulation of stress-responsive genes. Concurrently, RdDM fully restores DNA methylation at chromosomal arm regions, such as gene promoters with embedded TEs, thereby resetting any the transcriptional activity of *cis*-regulated genes.

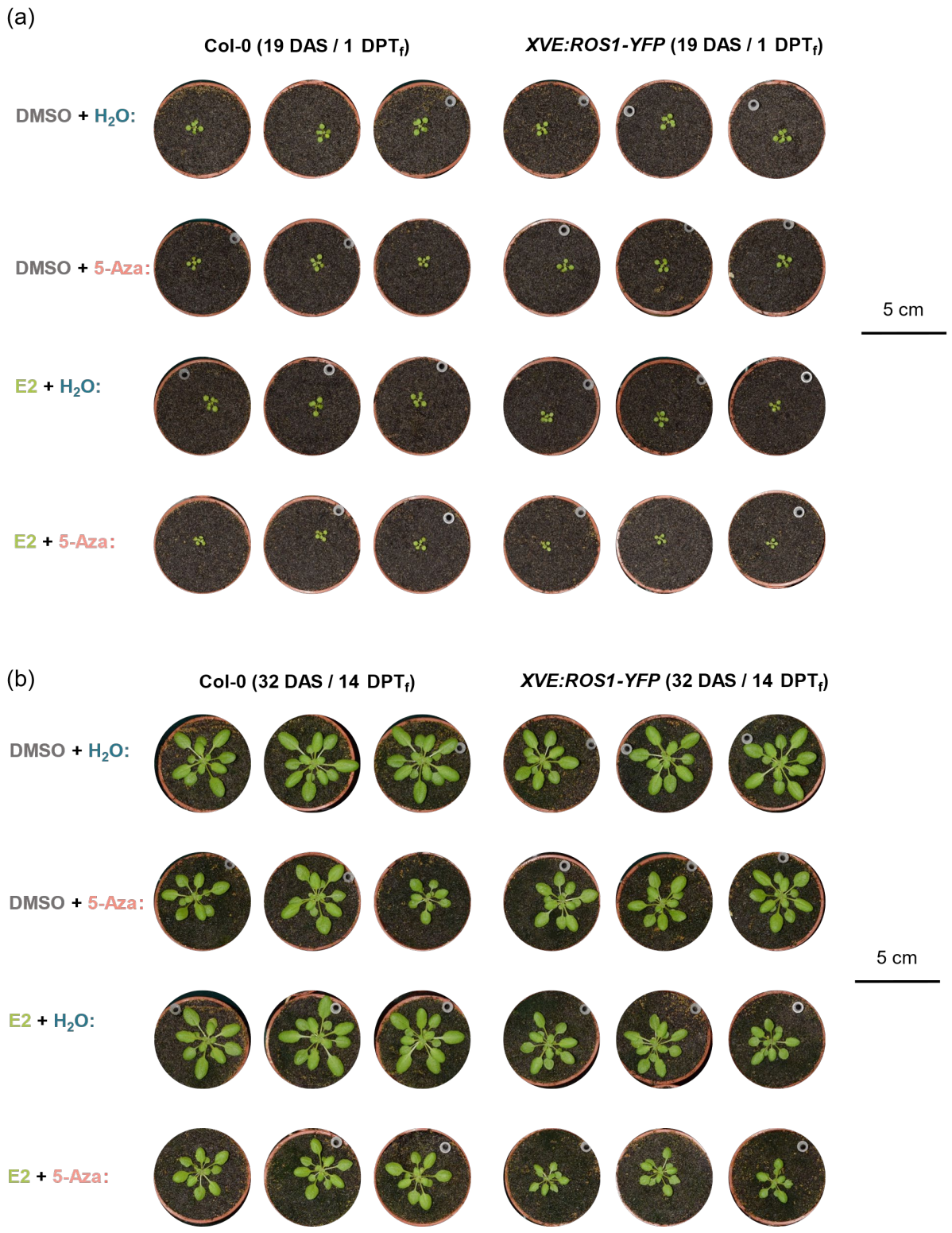
5.6 Supplementary Figures



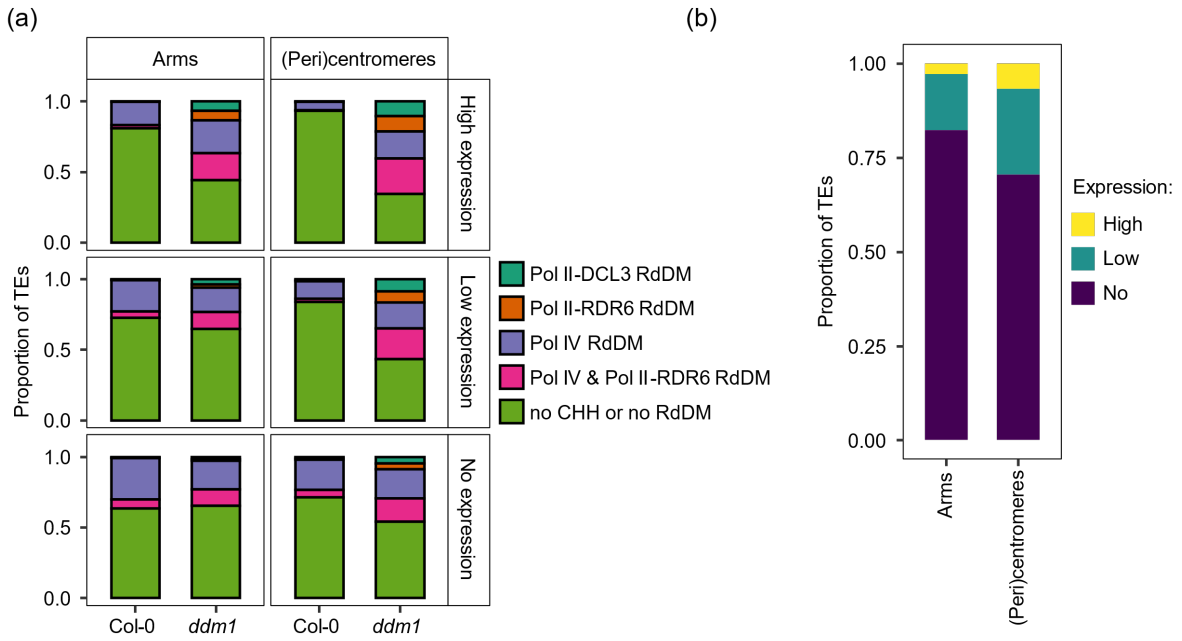
Supplementary Figure 5.1 Average cytosine DNA methylation across the five nuclear chromosomes of Col-CEN (v1.2) in E2-treated *XVE:ROS1-YFP* L5 seedlings.
Average cytosine methylation in allC, CG, CHG, and CHH contexts is shown as a percentage relative to DMSO-treated seedlings. Methylation levels were calculated in 100 kb windows across the five nuclear chromosomes. Dashed vertical lines indicate the pericentromeric regions, and shaded grey boxes highlight the centromeric regions, as defined by Naish et al (2021).



Supplementary Figure 5.2 Chromosome-level metaplots of WGBS-derived cytosine DNA methylation in *cmt3-12* plants and 5-Azacytidine treated plantsChromosome-level metaplots of WGBS-derived cytosine DNA methylation in (a) *cmt3-12* plants (Ning et al., 2020) and (b) Col-0 plants treated with various doses of 5-Azacytidine (5-Aza) (Griffin et al., 2016) across allC, CG, CHG, and CHH contexts. Methylation data were binned into 100 kb windows across the TAIR10 genome and averaged based on their relative distance from the pericentromere midpoint. Average cytosine DNA methylation levels are expressed as a percentage relative to (a) Col-0 plants or (b) Col-0 plants treated with H_2O. Dashed vertical lines indicate the minimum and maximum pericentromeric start and end positions relative to the centromere.



Supplementary Figure 5.3 Representative images (n=3) of Col-0 and *XVE:ROS1-YFP* L5 plants under all treatment combinations from the 5-Azacytidine experiment (see Figure 5.6 for details).
(a) Top grid shows plants 1 day post final E2 treatment (1 DPT_f), which equates to 19 days after sowing (19 DAS). (b) Bottom grid shown plants 14 DPT_f / 32 DAS. Plants in the top and bottom grid are from the same pot. Scale bar = 5 cm.



Supplementary Figure 5.4 TE expression in chromosome arms and pericentromeric regions. Data analysed and plotted from Panda and Slotkin (2020) showing (a) the forms of RdDM (as defined by Panda et al. 2016) targeting TEs in the Arabidopsis wild-type ecotype Col-0 and the hypomethylated mutant *ddm1*, which exhibits TE activation. TEs are categorised based on expression levels (no, low, high) in hypomethylated mutants and by their location in either chromosome arms (Arms) or centromeric/pericentromeric regions ((peri)centromeres), as defined by Panda and Slotkin (2020). (b) The proportion of TEs exhibiting high, low, and no expression in chromosome arms and (peri)centromeric regions in hypomethylated mutants, as defined by Panda and Slotkin (2020).

Chapter 6. Establishing Stress-Specific Memory In Arabidopsis By Modulating the Spatiotemporal Activity of ROS1

Authors:

Hannan Parker, A[1]; Hipperson H.; Rolfe S. A. [1]; Smith, L. M.[1] and Ton, J[1].

Affiliations:

[1] Plants, Photosynthesis and Soil, School of Biosciences, The University of Sheffield, Western Bank, Sheffield, S10 2TN, UK

Author Contributions:

AHP (the candidate), SAR, LMS, and JT proposed the original ideas for the research. AHP conducted experiments and gathered data with assistance from HH. Data analysis was performed by AHP. AHP created all Figures and wrote all the text in the Chapter. JT reviewed and provided comments on the Chapter. Funding for the project came from studentships/grants awarded to AHP and JT.

6.1 Abstract

Induced resistance (IR) occurs when plants respond more rapidly and robustly to stress after exposure to an immune-activating stimulus earlier in life. While IR can provide enhanced defence against taxonomically unrelated attackers, it is often most effective when the subsequent challenge involves a similar infection strategy or lifestyle to that of the initial stimulus. Interestingly, however, the DNA demethylase ROS1 has been implicated in the regulation of IR to different stress types, including biotrophic pathogens, chewing herbivores and even abiotic stresses. Here, I have tested the hypothesis that the spatiotemporal pattern of ROS1-dependent DNA demethylation, which may differ between stress types, determines the specificity of the resulting IR response. Using an estradiol-inducible gene construct for the DNA demethylase ROS1 in Arabidopsis (*XVE:ROS1-YFP*), plants were exposed to 3 different estradiol induction regimens, varying both spatially (tissue type) and temporally (timing/frequency). Subsequently, plants were challenged with a biotrophic bacterial pathogen (*Pseudomonas syringae* pv. *tomato* DC3000), a necrotrophic fungal pathogen (*Plectospherella cucumerina*), or an abiotic stress treatment (hypoxia caused by flooding). The data presented in this Chapter confirm that the spatiotemporal patterning of ROS1 induction determines the effectiveness of IR later in life. Global analysis of DNA methylation of all treatment regimens at the time-point of challenge revealed a significant difference in both the extent and distribution of ROS1-driven DNA demethylation. These findings not only explain how a single DNA demethylase can control IR responses against a range of biotic and abiotic stresses, but it also provides proof-of-concept that controlled modification of DNA methylation by ROS1 can be exploited to engineer epigenetic resistance against specific stress types.

6.2 Introduction

Due to their sessile lifestyle, plants cannot escape from the varying and often hostile environmental conditions they encounter during their life cycles, forcing them to rely on internal defence mechanisms to survive. However, continuous activation of all available defence mechanisms imposes significant costs on growth, yield, development, and ultimately, reproductive success. Therefore, plants have evolved to the ability to acquire resistance through priming, which enables them to sensitise specific defence responses based on the environmental stress they experienced earlier in life (van Hulten et al., 2006; Crisp et al., 2016; Wolinska and Berens, 2019; López Sánchez et al., 2021; Cooper and Ton, 2022). Priming represents of form of immune memory in plants and underpins the well-characterised phenomenon of induced resistance (IR), whereby genetically susceptible plants acquire resistance against (a)biotic stresses upon exposure to a stress-associated stimulus earlier in life (Mauch-Mani et al., 2017; Wilkinson et al., 2019; López Sánchez et al., 2021; Yassin et al., 2021; De Kesel et al., 2021; Cooper and Ton, 2022).

As IR typically involves the augmentation of basal defence mechanisms that are mechanistically related to pattern triggered immunity (PTI) (Walters et al., 2005, 2013; Ahmed et al., 2011; Mauch-Mani et al., 2017; Yassin et al., 2021), it offers protection against a wider spectrum of biotic stresses than effector trigger immunity (ETI), which relies on the detection of effector molecules that are limited to a taxonomically narrow range of attackers (Remick et al., 2023). Nonetheless, recent work on the specificity of long-term stress memory has demonstrated that certain IR responses coincide with induced susceptibility (IS) responses to other types of stress. For instance, early exposure to drought in rice is associated with IS against the pathogenic fungus *Magnaporthe oryzae* (Bidzinski et al., 2016). In Arabidopsis, early exposure (2 week old plants) to the phytohormone jasmonic acid (JA), which is involved in defence against necrotrophic pathogens and herbivores (Pieterse et al., 2012; Erb and Reymond, 2019), leads to IR against the generalist herbivore *Spodoptera littoralis* 3 weeks later, but simultaneously causes IS against the necrotrophic fungus *Plectosphaerella cucumerina* (*Pc*) and the hemi-biotrophic bacterial pathogen *Pseudomonas syringae* pv. *tomato* DC3000 (*Pst*) (Wilkinson et al., 2023). Furthermore, this trade-off between IR and IS can last over multiple generations. For instance, the progeny of Arabidopsis plants that have been infected with *Pst* are more susceptible to necrotrophic fungi, and vice versa (Luna et al., 2012; López Sánchez et al., 2021). Likewise, the progeny of salt-treated plants are more susceptible to both biotrophic and necrotrophic pathogens (López Sánchez et al., 2021). It thus seems that long-term stress-memory enables IR against specific stresses that

are controlled by the same innate defence response as the eliciting stimulus, which is associated with IS to stresses that are controlled by other pathways.

Interestingly, while stress memory can be specific as outlined above, epigenetic mechanisms, such as DNA methylation, have been implicated in the establishment and/or regulation of stress memory against a diverse range of (a)biotic stress types (Wilkinson et al., 2019; Harris et al., 2023). In particular, the DNA demethylase ROS1 of Arabidopsis appears to play a ubiquitous role in plant defence responses to a range of environmental stresses. For instance, ROS1 is required for long-term and heritable IR against both the chewing herbivore *Spodoptera littoralis* (Wilkinson et al., 2023) and the biotrophic pathogen *Hyaloperonospora arabidopsidis (Hpa)* (López Sánchez et al., 2016), following seedling exposure to JA and parental disease by *Pst,* respectively. However, it remains unclear how this single DNA demethylase can serve as a common mediator of stress-specific memory against different (a)biotic stress types. As different stresses are expected to differ in their duration and their severity on different tissues, it is possible that the spatiotemporal activity of stress-induced DNA demethylation determines the specificity of ROS1-driven stress memory.

In this final study of my PhD, I have explored whether the developmental stage (spatial) and/or the duration/frequency (temporal) of induced ROS1 activity influences the specificity of the resulting memory response. To this end, transgenic *XVE:ROS1-YFP* plants, which carry an estradiol (E2)-inducible transgene for ROS1, were treated with E2 at varying points in their development over a 3-week time window, after which they were challenged with *Pst*, *Pc*, and flooding-induced hypoxia when 4 weeks old. The primary hormonal defence pathways involved in the immediate defence reaction to these stresses are salicylic acid (SA), jasmonic acid (JA), and ethylene, respectively, thereby reflecting a broad-range of defence mechanisms (Cameron and Zaton, 2004; Pieterse et al., 2012; Erb and Reymond, 2019; Hartman et al., 2019, 2021). The results in this Chapter show that the effectiveness of IR to different stress types is indeed dependent on the spatiotemporal patterning of ROS1 activity experienced earlier in life, which is associated with distinct patterns of global DNA (de)methylation across the genome. As ROS1 can target distinct genomic loci (Tang et al., 2016), these results suggest that the cumulative increase in activity and the developmental context in which ROS1 is induced determines its genomic targeting and the subsequent effectiveness of IR.

6.3 Materials and methods

6.3.1 Plant Material and Growth Conditions

For all experiments in this Chapter, *XVE:ROS1-YFP* line 5 (L5) plants were used. This line is a single insertion transformant (Figure 5.1) carrying an estradiol (E2)-inducible construct of the DNA demethylase ROS1 in the genetic background of the Arabidopsis ecotype Col-0. Details about the generation of *XVE:ROS1-YFP* L5 plants are presented in Section 2.4 of this thesis. Plants were cultivated in short-day conditions using 70 mL pots containing a 2:1 soil:sand mixture following a 4 day stratification period, as detailed in Section 2.2. Plants were watered from the bottom 2-3 times per week.

6.3.2 Chemical Treatments

Four different chemical treatment regimens were applied to *XVE:ROS1-YFP* L5 seedlings, as illustrated in Figure 6.2a. During the 4-day stratification period at 4°C, approximately 100 seeds were placed in four separate Eppendorf tubes containing 2 mL of distilled H_2O (dH_2O). For the 'Seed' treatment course, the water was supplemented with 25 µM estradiol (E2), while all other tubes were supplemented with 0.05% dimethylsulfoxide (DMSO). A similar seed treatment in rice containing an XVE-controlled GUS transgene was shown to effectively induce expression in germinating embryos (Chen et al., 2017). After stratification, ~5 seeds were sown per pot, with a total of 18 pots per treatment, each representing a biological replicate. All pots were thinned to a single central plant per pot 10–12 days after sowing (DAS). Between 14 and 21 DAS, four spray applications of either 0.05% DMSO or 25 µM estradiol (E2) were applied to all seedlings, prepared using dH_2O and supplemented with 0.015% surfactant (v/v) (silwet L-77; LEHLE SEEDS, VIS-30) (silwet). For the 'DMSO' and 'Seed' treatment courses, all four sprays consisted of 0.05% DMSO. For the 'Multi' treatment course, all four sprays consisted of 25 µM E2. For the 'Single' treatment course, the first three spray treatments were performed using 0.05% DMSO, but the final spray treatment at 21 DAS was with 25 µM E2. At 28 DAS, all plants were challenged with a stress treatment, as detailed in the next paragraph, or harvested for long-read Oxford Nanopore Technology sequencing (ONT-seq). Pots were randomised for treatment before (0-13 DAS) and after (15 DAS onwards) the spray-treatment period. Preparation of E2 chemical stocks is detailed in Section 2.3.

6.3.3 Stress treatments and Analysis

The hemi-biotrophic bacterial pathogen *Pseudomonas syringae* pv. *tomato* DC3000 *luxCDABE* (*Pst-Lux*) (Fan et al., 2008; Furci et al., 2021) was prepared as an aqueous

inoculum in 10 mM $MgSO_4$ at an OD_{600} = 0.2, supplemented with 0.015% silwet. The inoculum was sprayed onto the surface of leaves until saturation and 2 days post infection (dpi), bacterial colonisation was analysed by measuring *Pst*-lux bioluminescence intensity on plants (Furci et al., 2021). Further information about the preparation, inoculation, and analysis of *Pst-Lux* infections are detailed in Section 2.5. Examples of *Pst-Lux* bioluminescence images are shown in Figure 6.1a.

The fungal pathogen *Plectosphaerella cucumerina* strain BMM (*Pc*) (Ton and Mauch-Mani, 2004), which was continuously cultured on potato dextrose agar (PDA) (Oxoid, CM0139) in the dark at 15–25°C, was propagated onto a fresh PDA plate 4 weeks prior to infection. Spores were resuspended into sterilised dH_2O by gently scraping the surface of the agar. Mycelial debris was removed by filtering the spore suspension through Miracloth (Merck, 475855) and spore density was quantified using a hemocytometer (Improved Neubauer, Hawksley, UK; Depth 0.1 mm, 1/400 mm^2). The final *Pc* spore density was diluted to $5x10^6$ spores/mL and for each plant, 4 leaves of a comparable physiological age were inoculated using 6 µl droplets of the inoculum to ensure a nectrotrophic infection strategy of *Pc* (Ton and Mauch-Mani, 2004; Pétriacq et al., 2016). Plants were kept at 100% relative humidity (RH) for 6 days before measuring lesion diameters with digital vernier callipers. The average lesion diameter across all 4 infected leaves was reported for each replicate (n = 18 per treatment). Examples of infected leaves are shown in Figure 6.1b.

To induce hypoxia, plants were submerged in the dark in tanks with a water level of 10 cm above leaf level, starting at the end of the light period, as described by Loreti et al (2020). After 5 days of submergence, plants were carefully removed and left to recover for 3 days under normal short-day growth conditions (section 2.2). Afterwards, leaves of individual plants were categorized into four groups: (i) healthy leaves (“healthy”); (ii) leaves with water-soaked, necrotic lesions but still turgid or full (“necrotic”); (iii) leaves that are completely decayed or shrivelled, indicating a total loss of moisture and structure (“dead”). Representative images of each category are shown in Figure 6.1c. Categorical scoring of susceptibility to hypoxia was then calculated per plant (n = 18 per treatment) using the following formula:

$$\text{Susceptibility score} = (1 \cdot f_{healthy}) + (2 \cdot f_{necrotic}) + (3 \cdot f_{dead})$$

Where *f* = relative frequency of each leaf category within a plant, multiplied by an arbitrary weight value ranging from the most resistant leaf category (healthy) to the most susceptible category (dead) (Figure 6.1c).

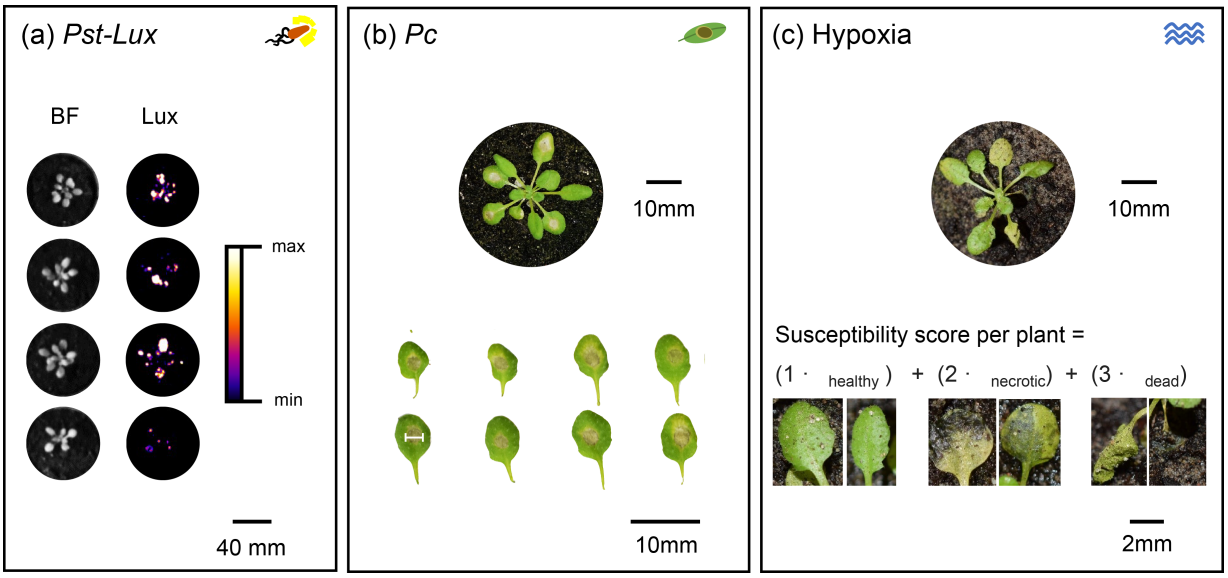


Figure 6.1 Methods for assessing susceptibility to selected stress types.
(a) infection with the hemibiotrophic bacterial pathogen *Pseudomonas syringae* pv. *tomato* DC3000 *luxCDABE* (*Pst-Lux*), (b) infection with the necrotrophic fungal pathogen *Plectosphaerella cucumerina* strain BMM (*Pc*), and (c) abiotic submergence-induced hypoxia. (a) *Pst-Lux* abundance was measured by capturing bright field (BF) images followed by a 4-minute exposure of bioluminescence (Lux) in complete darkness. Images were analysed using ImageJ as described in Section 2.5. Shown are Lux images false coloured with ImageJ's "Fire" LUT, as indicated by the colour scale. (b) *Pc* lesion diameters were measured (white bar, bottom left leaf). Average lesion diameter was calculated from 4 leaves per plant. (c) After 5 days of submergence in the dark, plants were left to recover for 3 days before scoring leaf phenotypes (healthy, necrotic, dead). Representative images for each category are shown. A susceptibility score was calculated per plant as shown by the equation.

6.3.4 Statistical Analysis of Green Leaf Area to Estimate Plant Growth

Whole-plant-coloured images of plants were taken using a Nikon D5300 digital camera. Green leaf area (GLA) was determined using the "color threshold" tool in Fiji/ImageJ (Rueden et al., 2017). GLA measurements from all three experiments (*Pst-Lux*, *Pc*, and hypoxia) were pooled. The effect of experiment, treatment, and the interaction between the two was tested with a two-way ANOVA having met the assumptions of normality and homogeneity of variance. Tukey post-hoc tests were performed in R v4.3.2 using the R package agricolae v1.3-7 (function: HSD.test). Plots for all experiments were generated using the R package ggplot2 v3.5.0 and ggbeeswarm v0.7.2.

6.3.5 Statistical Analysis of Stress Responses

For each stress treatment type (*Pst-Lux*, *Pc*, and Hypoxia), the effect of treatment on stress response was assessed using a Kruskal-Wallis test. If significant ($p < 0.05$), the Conover-Iman post-hoc test (Conover and Iman, 1979) was applied to compare each treatment group against the DMSO control group using the R package DescTools v0.99.54 (function: ConoverTest).

6.3.6 Quantification of ROS1-YFP Fluorescence by Epifluorescence Microscopy

Fluorescence microscopy for quantifying *in planta* accumulation of ROS1-YFP in whole-plants was performed using a Leica M165 FC fluorescent stereo microscope (Objective: 1x/0.06; ET-GFP filter set: 470/40 nm excitation, 495 nm dichroic and 525/50 nm emission) with a CoolLED pE-300 illumination system, using the Leica LAS X software v3.7.4.23463. Images were analysed using Fiji/ImageJ (Rueden et al., 2017). Plant regions were differentiated from background regions using the magic wand tool. ROS1-YFP intensity was reported as the relative mean-grey value, which is the sum of all grey values (ie., relative brightness values) in the selection divided by the total area (mm^2) of the selection. To account for background noise, ROS1-YFP signal intensity at each measured day was expressed relative to the average ROS1-YFP signal intensity of DMSO-treated plants. Statistical significance was determined from t-test FDR-adjusted *p*-values (*q*-values) using using the R package rstatix v0.7.2 (function: t_test; options: paired = FALSE, p.adjust.method = "fdr").

6.3.7 Oxford Nanopore Technology (ONT) Library Preparing and Sequencing

For ONT-seq, 3 biological replicates per treatment (12 samples total) were generated by harvesting aerial tissue from 12 individual plants at 28 DAS (1 week after the final E2/DMSO treatment), and snap freezing material in liquid N_2. Prior to harvesting, plants were kept in the dark for 2 hours. For each sample, material was ground to a fine powder using a pestle and mortar with liquid N_2 and split into aliquots of ~500 mg fresh weight which were stored at -80˚C.

High molecular weight (HMW) DNA was extracted from ground tissue aliquots using the NucleoBond® HMW DNA kit (Macherey Nagel, 740160.20) according to the manufacturer's instructions. DNA quality, quantity and size were assessed using a Nanodrop 8000 spectrophotometer, a Qubit 3.0 Fluorometer and an Agilent Femto Pulse. Library preparation was conducted using the Native Barcoding Kit 96 V14 (SQK-NBD114.96) and 400 ng of extracted HMW genomic DNA. The library was then loaded into a FLO-PRO114M flow-cell (FC) which was docked to a PromethION 2 Solo sequencing device. During the run, the flow cell was washed twice using the Flow Cell Wash Kit (ONT, EXP-WSH004), and fresh library loaded. The data output from this initial run was low with ~14 gigabases (Gb) called in 121 hours, with many pores becoming inactive relatively quickly. As discussed in the Native Barcoding Kit 96 V14 (SQK-NBD114.96) manual, very long reads can sometimes block/inhibit pore activity. Therefore, for the second sequencing attempt, 400 ng of HMW DNA in a total volume of 100 uL nuclease free H_2O was sheared using a Diagenode Megaruptor 3 device at a shear speed of 39 to generate fragment lengths of approximately 10 kb. Library construction

and sequencing was repeated as above, which yielded twice as much data (~30 Gb passed in 153 hours). Data output for barcode 87 was particularly low, so a fresh simplex library for the barcode was prepared using the Ligation Sequencing Kit V14 (SQK-LSK114). Sequencing was performed as above until ~1.5 Gb of passed sequencing reads was achieved. In total 42 Gb were successfully sequenced across all samples (Supplementary Data 6.1).

6.3.8 Processing and Base Calling in ONT-seq Data

As the sequencing data derives from PromethION flow cells with R10.4.1 chemistry and a 5 kHz sample rate, current versions of DeepSignal-Plant (Ni et al., 2021) are incompatible for calling cytosine methylation, as was done in Chapter 5 of this thesis. Therefore, the raw sequencing data in .pod5 format were pooled by barcode, trimmed, base called, and aligned to the Col-CEN v1.2 reference genome (Naish et al., 2021) using Dorado v0.7.0+71cc744 and the DNA base calling model 'dna_r10.4.1_e8.2_400bps_hac@v5.0.0' (https://github.com/nanoporetech/dorado; accessed: 23/05/2024). Modified base detection was enabled for the epigenetic marks 5-Methylcytosine (5mC), 5-Hydroxymethylcytosine (5hmC), and DNA N^6-methyladenosine (6mA) (function: dorado basecaller; options: --trim all, --reference Col-CEN_v1.2.fasta, --modified-bases 5mC_5hmC 6mA). Summary statistics were obtained from the output .bam files using NanoPlot v1.39.0 (Supplementary Data 6.1).

Summary counts of modified and unmodified bases in bedMethyl format were generated using Modkit v0.3.0 (https://github.com/nanoporetech/modkit; accessed: 20/05/2024). For cytosine methylation (5mC), calls were summarised in CG, CHG, and CHH contexts (function: modkit pileup; options: --filter-threshold 0.75 --motif CG 0 --motif CHG 0 --motif CHH 0 --reference Col-CEN_v1.2.fasta). Outputs from these distinct contexts were concatenated to form an allC summary file. For 5hmC and 6mA, calls were summarised for all cytosines and adenosines, by setting the modkit options to '--motif C 0' and '--motif A 0', respectively. The '--filter-threshold' parameter was kept at 0.75 for both.

Nuclear averages of 5mC, 5hmC, and 6mA levels (%) were calculated using only bases with read coverage ≥ 5 (Supplementary Data 6.1). Similarly, bases with ≥ 5 reads in the plastid genome, which is expected to lack 5mC, were used to roughly estimate false-positive 5mC calls. All samples had an average plastid 5mC level < 0.4% (Supplementary Data 6.1), which, for reference, is below the accepted non-conversion rate of WGBS data (< 2%) (Stuart et al., 2018).

6.3.9 Principal Component Analysis (PCA) of Modified Base Calls

PCA was performed using positional modified base levels (%) for 5mC, 5hmC, and 6mA. Bases with <5 mapped reads were dropped from the analysis. Furthermore, to focus the computational analysis on the most variable regions for a given modification, bases with standard deviations less than the median of the standard deviation across all other bases were removed from the analysis. PCA was also performed for 5mC in the sequence contexts CG, CHG, and CHH specifically. PCA was conducted in base R v4.2.1 (function: prcomp; options: scale = FALSE, center = TRUE) and plots were generated using the R packages ggplot2 v3.5.0 and viridis v0.6.5.

6.3.10 Chromosome-Level Metaplots of Cytosine DNA Methylation

Plots are based on average levels of 5mC in 100 kb bins of the Col-CEN genome assembly of Arabidopsis (Naish et al., 2021) in allC, CG, CHG, and CHH contexts. Only cytosines with ≥5 reads were included in the analysis. Furthermore, bins that did not occur in at least 1 replicate of each of the four treatment groups (DMSO, Seed, Single, Multi) were removed from the analysis. For each nuclear chromosome, the start of the peri-centromere was defined as the midpoint of the first bin in a string of at least 5 neighbouring bins for which TE density > protein-coding gene density. The end of the pericentromere was defined as the midpoint of the last bin in a string of at least 5 neighbouring bins where TE density > gene density. This resulted in the following defined pericentromeric ranges: Chr1: 12050000-19850000, Chr2: 1350000-9850000, Chr3: 10350000-18950000, Chr4: 1750000-9850000, Chr5: 9650000-18150000. Annotations of TEs and genes were downloaded from the Col-CEN v1.2 GitHub page (https://github.com/schatzlab/Col-CEN/tree/main/v1.2; accessed 21/09/2023). The position of 100 kb windows were expressed relative to the midpoint of the chromosome's pericentromere, and average 5mC levels within the bin were averaged by treatment and relative distance to the midpoint of the pericentromeres. The pericentromeric and centromeric regions in the metaplots were defined as the maximum absolute relative distance for any (peri)centromeric window from the midpoint of the pericentromere for both the start and end points. Col-CEN centromeric ranges were used as defined by Naish et al (2021) (Chr1: 14840000-17560000, Chr2: 3823000-6046000, Chr3: 13597000-15734000, Chr4: 4204000-6978000, Chr5: 11784000-1456000). Plots were generated using the R packages ggplot2 v3.5.0 and viridis v0.6.5.

6.3.11 Average Cytosine DNA Methylation Plots

Average 5mC (%) in nuclear chromosomes was calculated for each sample in all cytosine (allC) contexts, as well as CG, CHG, and CHH sequence contexts, using only cytosines with ≥ 5 mapped reads. Cytosine DNA methylation was also calculated within chromosome arms, pericentromeric regions, and centromeric regions specifically using bedtools v2.31.0 (Quinlan and Hall, 2010) (function: bedtools intersect). Centromeric ranges were used as defined by Naish et al (2021), pericentromeric regions were defined as above, but with the centromeric regions removed. Chromosome arms included all other regions of the nuclear genome.

6.3.12 Identification of Differentially Methylated Regions (DMRs) for Cytosine DNA Methylation

To identify 5mC DMRs between 'DMSO'-treated replicates and all other treatments ('Seed', 'Single', 'Multi') (n = 3 per treatment), 5mC summary files from Modkit were used to generate files for analysis with the R package DSS v2.46.0 (H. Feng et al., 2014; Park and Wu, 2016). Differentially methylated cytosines (DMLs) were first called using the DSS function 'DMLtest' with parameters: 'equal.disp = FALSE, and 'smoothing = FALSE'. DMRs were then called with the function 'callDMR' with parameters: 'delta = 0.1', 'p.threshold = 0.05', 'minlen = 25', 'minCG = 3', 'dis.merge = 50', and 'pct.sig = 0.5'. Using the R package GenomicRanges v1.54.1 (Lawrence et al., 2013), DMRs were checked for overlaps with transposable elements, using the Col-CEN v1.2 (Naish et al., 2021) TE annotation file 't2t-col.20210610.fasta.mod.EDTA.TEanno.gff3' (accessed 21/09/2023).

Hyper- and hypomethylated DMRs were analysed separately to identify overlaps between 'Seed' vs. 'DMSO' ('Seed DMRs'), 'Single' vs. 'DMSO' ('Single DMRs'), and 'Multi' vs. 'DMSO' ('Multi DMRs') comparisons. Firstly, all DMRs for each direction (hypo- or hypermethylation) and treatment ('Seed', 'Single', or 'Multi') were merged across all cytosine sequence contexts (allC, CG, CHG, and CHH DMRs), using the 'reduce' function from the R package GenomicRanges v1.54.1 (Lawrence et al., 2013).

Next, Multi DMRs were overlapped with Single DMRs using GenomicRanges (function: findOverlaps; options: select = "all"), and overlapping regions were merged by taking the minimum start and maximum end positions of each pair of overlapping DMRs. This set of overlapping DMRs ('Single-Multi DMRs') was further simplified by merging any adjacently overlapping regions using GenomicRanges (function: 'reduce').

Single-Multi DMRs were next overlapped with Seed DMRs to create a set of DMRs common across all treatments ('Seed-Single-Multi DMRs'). To identify regions shared between 'Single' and 'Multi' but absent in 'Seed', the 'Seed-Single-Multi' DMRs were subtracted from the 'Single-Multi' DMRs, forming the 'Single-Multi-noSeed' DMR set. Similarly, systematic comparisons and exclusions were performed to identify 'Seed unique', 'Single unique', 'Multi unique', 'Seed-Multi-noSingle', and 'Seed-Single-noMulti' DMR sets. The number of DMRs in each category was used to generate a Venn diagram using the R package eulerr v7.0.2.

Random DMRs of equal size and number to the 'Single-Multi-noSeed' hypomethylated DMR set were generated by randomly selecting a chromosome start coordinate in the Col-CEN v1.2 genome (Naish et al., 2021), using base R v4.3.2 (function: runif; options: n = 1, min = 1, max = max_chr_size). To account for the varying size of chromosomes, proportional probability weights were assigned for chromosomal selection (function: sample; options: prob = chr_size_proportion).

6.3.13 Annotation of DMRs

Merged hypomethylated DMRs were annotated, using the Col-CEN v1.2 TE ('t2t-col.20210610.fasta.mod.EDTA.TEanno.gff3') and gene ('Col-CEN_v1.2_genes.araport11.gff3.gz') annotation files (Naish et al., 2021). Annotation was performed, using the R packages genomation v1.34.0 (Akalin et al., 2015) and GenomicRanges v1.54.1 (Lawrence et al., 2013). In cases where a DMR overlapped with multiple features, a single annotation was assigned in the following order of preference: gene body > promoter > downstream > non-coding RNA > intergenic. Promoters were defined as 1 kb regions upstream of the transcriptional start site (TSS) of protein-coding genes. Downstream gene sequences were defined as 1-kb regions from the transcriptional termination site (TTS) of protein-coding genes. Gene bodies include everything from the TSS to TTS of protein-coding genes. Non-coding RNA (ncRNA) included any defined transcribed locus in the Col-CEN Araport11 annotation file that does not match a protein-coding gene. Intergenic regions were defined as regions that do not overlap any of these above defined features. DMRs were annotated as overlapping/not overlapping with a TEs using the Col-CEN v1.2 TE annotation file.

To test whether the genomic distribution of DMRs differed between each hypo-DMR set (Figure 6.4), pairwise Fisher's exact tests were performed in base R v4.2.1 for all annotated DMR sets using a Monte Carlo simulation to calculate *p*-values (function: fisher.test; options: simulate.p.value = TRUE, B = 10000) which were then corrected for using the FDR method (*q*-values) (function = p.adjust; options: method = "fdr"). Compact letter display (CLD)

letters were generated using the R package multcompView v0.1-10 (function: multcompLetters; options: threshold = 0.05).

6.3.14 Gene ontology (GO) Enrichment of IR-associated Hypomethylated DMRs

Hypomethylated DMRs associated with each stress type mapping to promoter or downstream regions of protein-coding genes were tested for gene ontology enrichment (GO) for both biological process (BP) and molecular function (MF), using the R package clusterProfiler v4.10.1 (Yu et al., 2012; Wu et al., 2021) and org.At.tair.db v3.18.0 (function: enrichGO; options: universe = 'ColCEN_prot_genes', pvalueCutoff = 0.05, qvalueCutoff = 0.05, OrgDb = 'org.At.tair.db', keyType = "TAIR", pAdjustMethod = "Bonferroni", ont = "BP" / "MF", minGSSize = 10, maxGSSize = 500). No update to GO terms were provided in the Col-CEN v1.2 genome release and so, TAIR10 GO annotations were used.

6.4 Results

6.4.1 The Spatiotemporal Patterning of ROS1 Induction Determines the Level of Resistance Induced Against Specific Biotic and Abiotic Stressors

XVE:ROS1-YFP L5 plants were treated with 25 µM E2 or a mock DMSO treatment (0.05%) at different stages of growth and development between stratification to 21 days after sowing (DAS) (Figure 6.2a). 'Seed' treated plants were stratified at 4°C in the dark in a water solution containing 25 µM E2. The aerial tissue of 'Single'-treated plants was sprayed with 25 µM E2 at 21 DAS, whereas the 'Multi'-treated plants were sprayed four times with 25 µM E2 between 14 and 21 DAS, at 2–3-day intervals (Figure 6.2a). The control treatment ('DMSO') consisted of identical treatment timepoints but using the equivalent volume of DMSO present in the E2 treatment solutions (Figure 6.2a). Plants from all treatment courses were then exposed to one of three environmental stressors at 28 DAS: (i) the hemi-biotrophic bacterial pathogen *Pseudomonas syringae* pv. *tomato* DC3000 *luxCDABE* (*Pst-Lux*) (Fan et al., 2008; Furci et al., 2021), (ii) the necrotrophic fungal pathogen *Plectosphaerella cucumerina* strain BMM (*Pc*) (Ton and Mauch-Mani, 2004; Pétriacq et al., 2016), and (iii) abiotic submergence-induced hypoxia (Loreti et al., 2020).

As is shown in Figures 6.2b and c, all E2 treatment regimens led to a reduction in *XVE:ROS1-YFP* L5 plant sizes relative to 'DMSO'-treated plants. Pooling data from all three experiments at 27 DAS revealed a significant effect of E2-treatment course (ANOVA: $F=19.11$, $df=3,212$, $p<0.001$) and experiment (ANOVA: $F=8.92$, $df=2,212$, $p<0.001$) but not the interaction between these factors on green leaf area (GLA) (ANOVA: $F=1.65$, $df=6,212$, $p>0.05$). Thus, all ectopic ROS1 induction regimens had a consistent effect on plant size across all three stress experiments. Post-hoc analysis on the effect of E2 treatment revealed that 'Multi'-treated plants had the most severe reduction in green leaf area (GLA) just before stress induction (27 DAS) (-23 %; Tukey, $p_{adj} < 0.05$), with 'Seed' (-13 %) and 'Single' (-10 %) treatments having intermediate but significant reductions in GLA compared to 'DMSO'-treated plants (Tukey, $p_{adj} < 0.05$), which were not statistically different from each other (Tukey, $p_{adj} > 0.05$) (Figures 6.2b, 6.2c).

'Multi'-treated plants had induced resistance (IR) to all three tested stress treatments (Figure 6.2 d-f). Interestingly, however, 'Single'-treated plants had IR against *Pc* and hypoxia, but not against *Pst-Lux*, whereas 'Seed'-treated plants only had IR against hypoxia. To ensure the growth reductions and stress-specific IR responses to E2 treatment are not caused by a chemical effect from E2 rather than ROS1, a repeat experiment was conducted using both Col-0 and *XVE:ROS1-YFP* genotypes, after which both plant growth (GLA) and tolerance to

flooding was quantified (Supplementary Figure 6.1). As is shown in Supplementary Figure 6.1, no significant differences in GLA or hypoxia susceptibility was observed between 'Seed'- and 'DMSO'-treated Col-0 plants (Supplementary Figure 6.5), whereas the E2-induced growth reduction and IR to flooding in *XVE:ROS1-YFP* were similar to those observed in the first experiments (Figure 6.2) (Supplementary Figure 6.1). Thus, the spatiotemporal patterning of transient ROS1 activity determines the specificity of IR responses, which can be evident up to 4 weeks after induction.

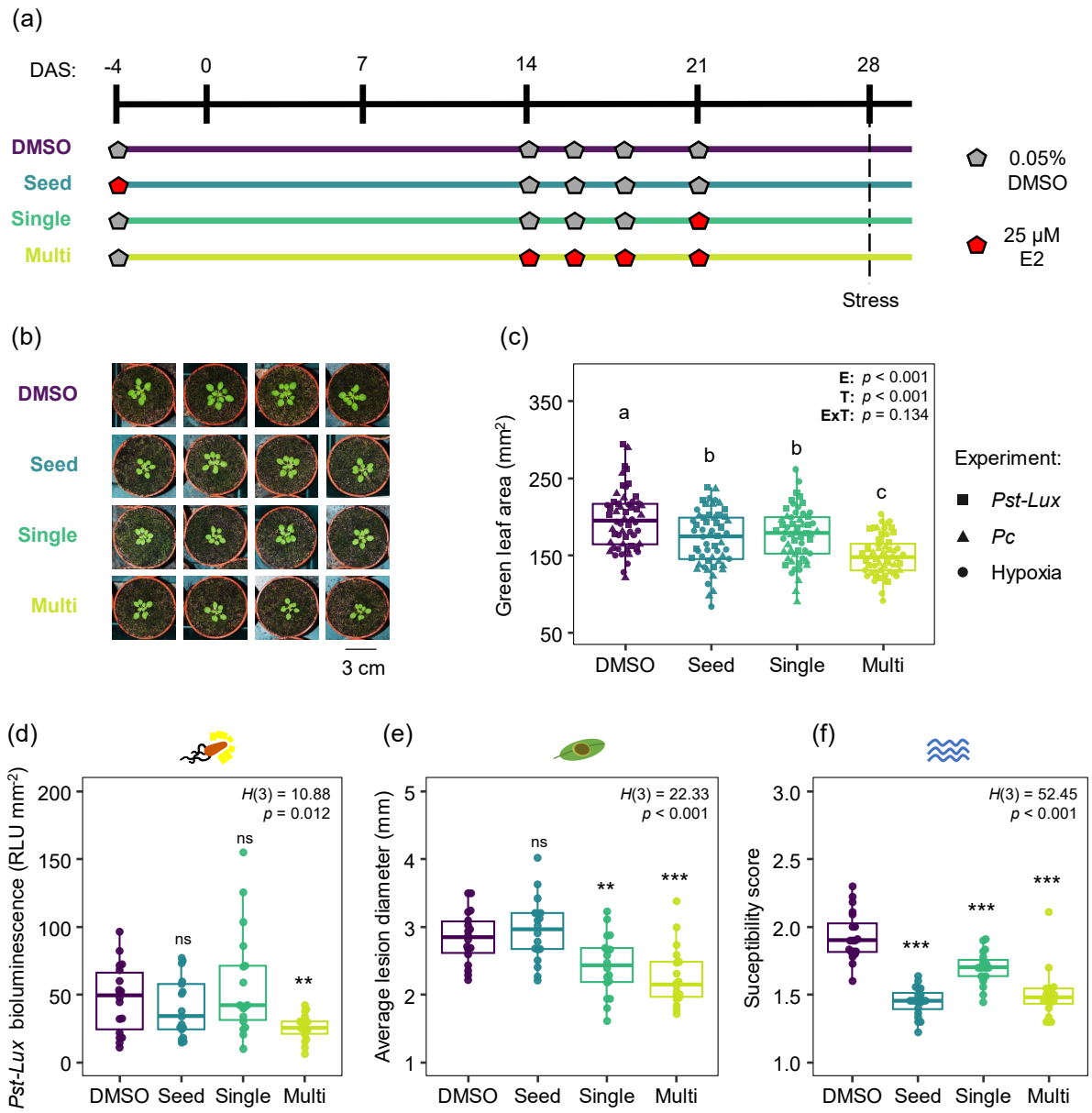


Figure 6.2 Effects of various estradiol (E2) treatments on plant growth and stress resistance in *XVE:ROS1-YFP* L5 plants.

(a) Schematic of four treatment regimens: 'DMSO' (control), 'Seed', 'Single', and 'Multi'. Treatments were applied at indicated days after sowing (DAS), with stress treatments at 28 DAS. Grey and red hexagons correspond to when DMSO (0.05 %) and E2 (25 µM) treatments were applied for each treatment course, respectively. See methods for details on how chemical and stress treatments were applied. (b) Representative images of plants from each treatment group at 27 DAS. (c) Green leaf area (GLA) analysis at 27 DAS (n = 3 experiments). ANOVA revelead significant effects of experiment (E; $F = 8.92$, $df = 2,212$, $p < 0.001$) and treatment (T; $F = 19.11$, $df = 3,212$, $p < 0.001$), but not their interaction (E x T; $F = 1.65$, $df = 6,212$, $p > 0.05$) on GLA. Groups not sharing letters are significantly different following post-hoc analysis on the effect of treatment (Tukey post-hoc, $p_{adj} < 0.05$). (d, e, f) Quantification of stress resposes. (d) *Pseudomonas syringae* pv. *tomato* DC3000 *luxCDABE* (*Pst-Lux*) bioluminescence (Fan et al., 2008; Furci et al., 2021) 2 days post infection. Points represent *Pst-Lux* bioluminescence intensity of individual plants. (e) Average necrotic lesion size 6 days post-infection with *Plectosphaerella cucumerina* strain BMM (*Pc*); each point represents the average lesion diameter measured from 4 leaves of an individual plant. (f) Plant susceptibility to hypoxia following 5 days of submergence and 3 days of recovery. Each point represents the susceptibility score of an individual

plant, calculated a weighted formula based on the relative proportions of healthy, necrotic, and dead leaves. See methods for details. Significant effects of treatment on quantified stress responses were detected using a Kruskal-Wallis test ($p < 0.05$). Asterisks represent statistical difference compared to DMSO-treated plants following a Conover-Iman post-hoc test (Conover and Iman, 1979) (ns $p > 0.05$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

6.4.2 The Timing and Frequency of ROS1 Induction Differentially Impacts Genome-Wide Cytosine DNA Methylation Levels

Since each treatment regimen resulted in IR to a specific set of stress types (Figure 6.2), it is likely that each treatment uniquely affects genome-wide DNA methylation at the time of infection (28 DAS). Thus, *XVE:ROS1-YFP* L5 plants under each treatment regimen ('DMSO', 'Seed', Single, Multi) were harvested and used for long-read Oxford Nanopore Technology sequencing (ONT-seq) of DNA.

Principal component analysis (PCA) of DNA methylation at all cytosines (allC) and in specific sequence contexts (CG, CHG, CHH) showed clear separation of 'Single' and 'Multi' treatments from 'DMSO' and 'Seed' treatments along the first principal component (PC1) (Figure 6.3a). 'Seed'-treated seedlings clustered closely with 'DMSO'-treated samples in both allC and CG contexts, suggesting similar global cytosine DNA methylation patterns. However, in CHH and CHG patterns, 'Seed'-treated samples separated along PC2, indicating greater variability in cytosine DNA methylation within these contexts compared to 'DMSO'-treated samples.

PCA of 5-Hydroxymethylcytosine (5hmC), and DNA N^6-methyladenosine (6mA) at all cytosines and adenosines, respectively, revealed no separation linked to treatment (Supplementary Figure 6.2). Thus, ROS1 has little to no impact on genome-wide levels of 5hmC or 6mA. Furthermore, across all samples, the nuclear abundance of 5hmC was 0.22% which is less than the average abundance of cytosine DNA methylation (5mC) in the plastid genome (0.28%), which is expected to lack 5mC entirely. This supports previous findings that the epigenetic mark 5hmC is largely absent from the Arabidopsis genome (Erdmann et al., 2014).

Chromosome-level metaplots revealed clear reductions in DNA methylation in 'Single'- and 'Multi'-treated plants along the chromosome arms in all cytosine contexts (allC, CG, CHG, CHH) (Figure 6.3e). However, the extent of hypomethylation in chromosome arms was greater in 'Single'-treated plants compared to 'Multi'-treated plants (Figure 6.3e) (Supplementary Figure 6.3). This difference aligns with the observation that 'Multi'-treated plants showed lower

levels of ROS1-YFP fluorescence at 22 DAS (i.e., 24 hours after the final spray treatment), indicating either increased pre- or post-transcriptional silencing of the construct in 'Multi'-treated plants, or increased posttranslational degradation of the ROS1-YFP protein (Supplementary Figure 6.4). However, irrespective of the exact mechanisms involved, the reduced ROS1-YFP abundance at 22 DAS offers a plausible explanation for the reduced global cytosine DNA hypomethylation in 'Multi'-treated plants relative to 'Single'-treated plants (Figure 6.3e) (Supplementary Figure 6.3). In addition, it is possible that the DNA-remethylation response in 'Multi'-treated plants is more pronounced, considering that there was more time to activate this recovery response between the first induction treatment and the time-point of sampling (7 days).

In allC, CG, and CHG sequence contexts, 'Seed'-treated plants had no clear differences in their abundance of DNA methylation relative to DMSO-treated control plants. However, CHH cytosine DNA methylation in the chromosome arms was reduced by a similar extent in 'Seed'-treated plants (-14.1 %) and 'Multi'-treated plants (-13.6 %) compared to 'DMSO'-treated plants (Figure 6.3e). Despite this, on a single cytosine level, no significant differences between Seed- and DMSO-treated plants could be detected (Supplementary Figure 6.3). ROS1-dependent DNA hypomethylation in the (peri)centromeric regions in Single- and Multi-treated seedlings were visibly reduced compared to chromosome arms (Figure 6.3e). On a single-cytosine level, 'Single'-treated plants were significantly hypomethylated in allC (-9.6 %), CG (-6.8 %), and CHH (-15.6 %) contexts compared to 'DMSO'-treated plants (t-test, $q < 0.05$) (Supplementary Figure 6.3). In contrast, 'Multi'-treated seedlings were not significantly different to 'DMSO'-treated plants but had subtle increases in cytosine methylation in CHG (+3.0 %) and CHH (+1.4 %) contexts. Within centromeric regions, cytosines in both 'Single'- and 'Multi'-treated plants were significantly hypermethylated in CG and CHG contexts compared to DMSO-treated plants (Supplementary Figure 6.3). In CG contexts, hypermethylation levels were similar between 'Single'- and 'Multi'-treated seedlings, with increases of 3.1% and 2.7%, respectively. However, in CHG contexts, hypermethylation was noticeably higher in 'Multi'-treated seedlings, with increases of 6.6% in 'Single'-treated plants and 8.9% in 'Multi'-treated plants. Although not statistically significant (t-test, $p_{adj} > 0.05$), a similar trend was detected in CHH contexts, with 'Single'- and 'Multi'-treated plants having relative increases of 0.9% and 5.1% in centromeric regions, respectively (Supplementary Figure 6.3) (Figure 6.3e). In contrast, no significant differences in (peri)centromeric DNA methylation were detected in 'Seed'-treated plants at any sequence context. Thus, 'Seed', 'Single', and 'Multi' treatments each have distinct effects on genome-wide cytosine methylation levels at 28 DAS, offering a plausible explanation for the stress-specific IR responses between these treatments.

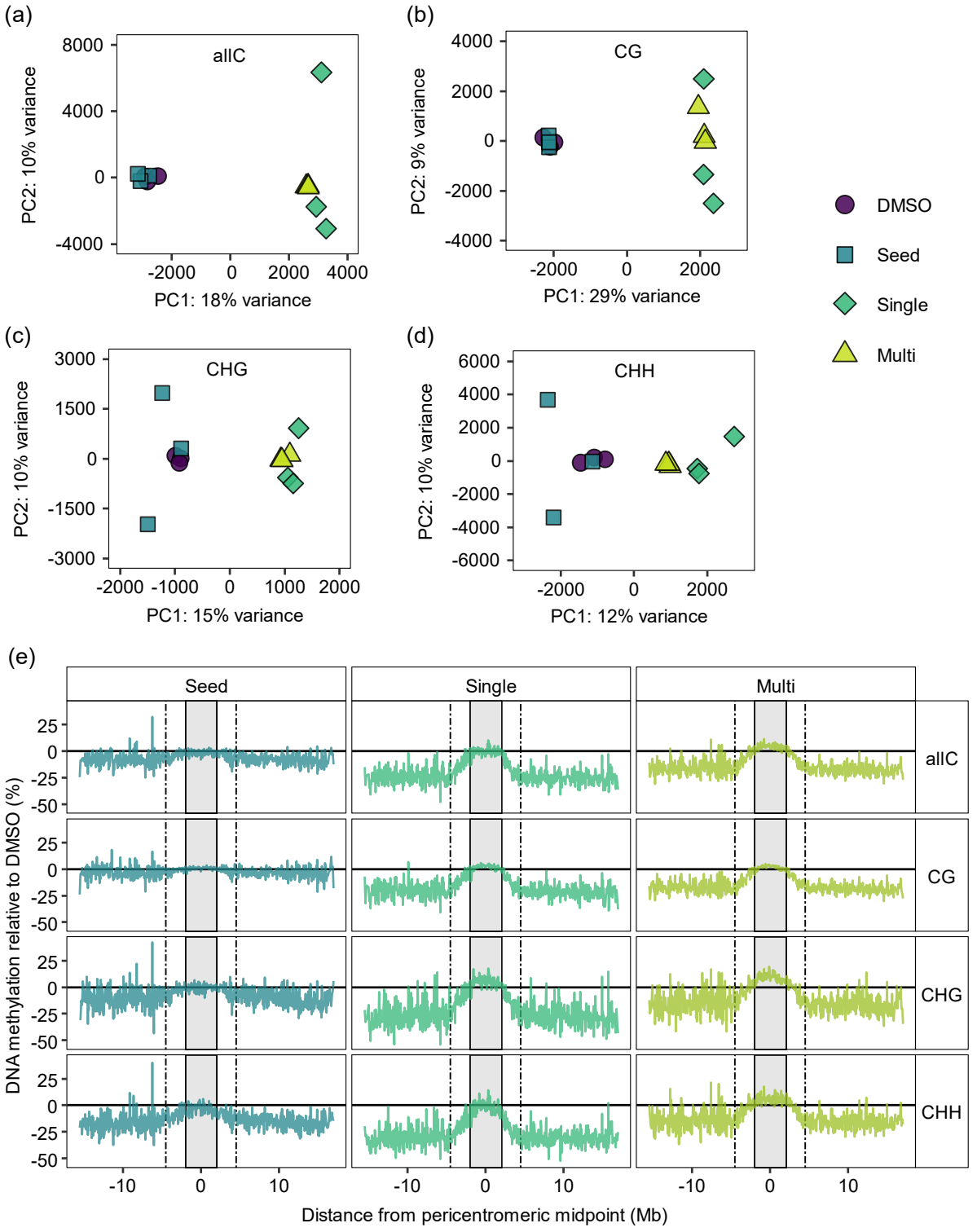


Figure 6.3 Global cytosine DNA methylation analysis of 'DMSO'-, 'Seed'-, 'Single'-, and 'Multi'-treated *XVE:ROS1-YFP* L5 plants (Figure 6.2a).
(a-d) Principal component analysis (PCA) plot showing variation in cytosine DNA methylation across all cytosine positions (allC) and context-specific cytosine positions (CG, CHG, and CHH). (e) Chromosome-level metaplots of average cytosine DNA methylation in 'Seed'-, 'Single'-, and 'Multi'-treated plants, expressed as a percentage relative to the average of 'DMSO'-treated plants. DNA methylation levels across all cytosines (allC), and cytosines in CG, CHG, and CHH sequence contexts were calculated for 100 kb windows of the Col-CEN genome assembly (Naish et al., 2021) and averaged based on their distance from the midpoint of the pericentromere in megabases (Mb). Dashed

vertical lines indicate the minimum and maximum pericentromeric start and end positions relative to the pericentromere midpoint. Shaded grey boxes represent the minimum and maximum centromeric regions relative to the pericentromere midpoint.

6.4.3 The Specific Genomic Regions Targeted by ROS1 Vary Depending on the Timing and Frequency of Induction

Genomic regions with significant changes in cytosine DNA methylation were identified by comparing 'Seed'-, 'Single'-, and 'Multi'-treated samples against 'DMSO'-treated samples using the R package DSS (H. Feng et al., 2014; Park and Wu, 2016). Differentially methylated regions (DMRs) for cytosine DNA methylation were identified in allC, CG, CHG, and CHH contexts (Supplementary Data 6.2) (Figure 6.4a). Few DMRs were identified in 'Seed'-treated samples, with 44 and 71 hyper- and hypo-methylated DMRs across all contexts. In contrast, 'Single'- and 'Multi'-treated samples had thousands of hypomethylated DMRs across all contexts, with the most occurring in CG contexts. Few hypermethylated DMRs were identified, but the majority occurred in CHG contexts in both 'Single' (24; 63 %) and 'Multi' (26; 60%), which reflects the CHG hypermethylation detected in the cytosine-resolution analysis (Figure 6.3e; Supplementary Figure 6.3). More hypomethylated DMRs were detected in 'Multi'-treated plants compared to 'Single'-treated plants, despite more extensive hypomethylation across in the genome in 'Single'-treated plants (Figure 6.3e; Supplementary Figure 6.3). It is possible that this discrepancy is due to a statistical artifact resulting from the fact that 'Multi'-treated samples, on average, had higher coverage per cytosine (10x) compared to 'Single'-treated samples (7x) (Supplementary Data 6.1). As a result, the 'Multi'-treated samples had greater statistical power per cytosine and region (H. Feng et al., 2014; Ziller et al., 2015), enabling more sensitive detection of DMRs at the defined statistical threshold ($p_{adj} < 0.05$).

To identify DMRs that are associated with resistance to *Pst-Lux*, *Pc*, and hypoxia (Figures 6.2d, 6.2e, 6.2f), hypo- and hyper-methylated DMRs were merged separately across contexts for a given comparison ('Seed', 'Single', 'Multi') against 'DMSO'. Overlaps between comparisons for merged hyper- and hypo-methylated DMRs (hyper-DMRs and hypo-DMRs) were then identified between comparison groups. Few merged hyper-DMRs were identified across all treatments, and only 2 overlapping hyper-DMRs, between 'Single' and 'Multi', were identified (Supplementary Figure 6.5). Thus, the analysis was focussed on the merged hypo-DMR set (Figure 6.4b). Hypo-DMRs associated with IR against *Pst-Lux* (*Pst-Lux* IR DMRs) were those associated with 'Multi' treatment only (n = 3,602) (Figures 6.2a, 6.4b), *Pc* IR DMRs were those overlapping between 'Single' and 'Multi' but not 'Seed' (n = 4,167) (Figures 6.2a, 6.4b), and hypoxia IR DMRs were those overlapping in 'Seed', 'Single', and 'Multi' (n = 12) (Figures 6.2a, 6.4b).

DMRs associated with IR against *Pst-Lux*, *Pc*, and Hypoxia all had an annotation profile different to that of randomly distributed DMRs across the genome (Fisher's Exact Test, $q < 0.05$) with overrepresentation of promoter and downstream regions in all three DMR groups (Figure 6.4c). Furthermore, *Pst-Lux* IR DMRs and *Pc* IR DMRs had significantly different genomic distributions from each other (Fisher's Exact Test, $q < 0.05$), with *Pst-Lux* IR DMRs mapping more to intergenic regions and less to promoter regions (Figure 6.4c). However, Hypoxia IR DMRs were not significantly different in their genomic distribution compared with *Pst-Lux* IR DMRs and *Pc* IR DMRs (Fisher's Exact Test, $q > 0.05$). Thus, the genomic distribution of hypomethylated DMRs varies based on the spatiotemporal patterning of ROS1 induction, which could be a determinate of the IR achieved.

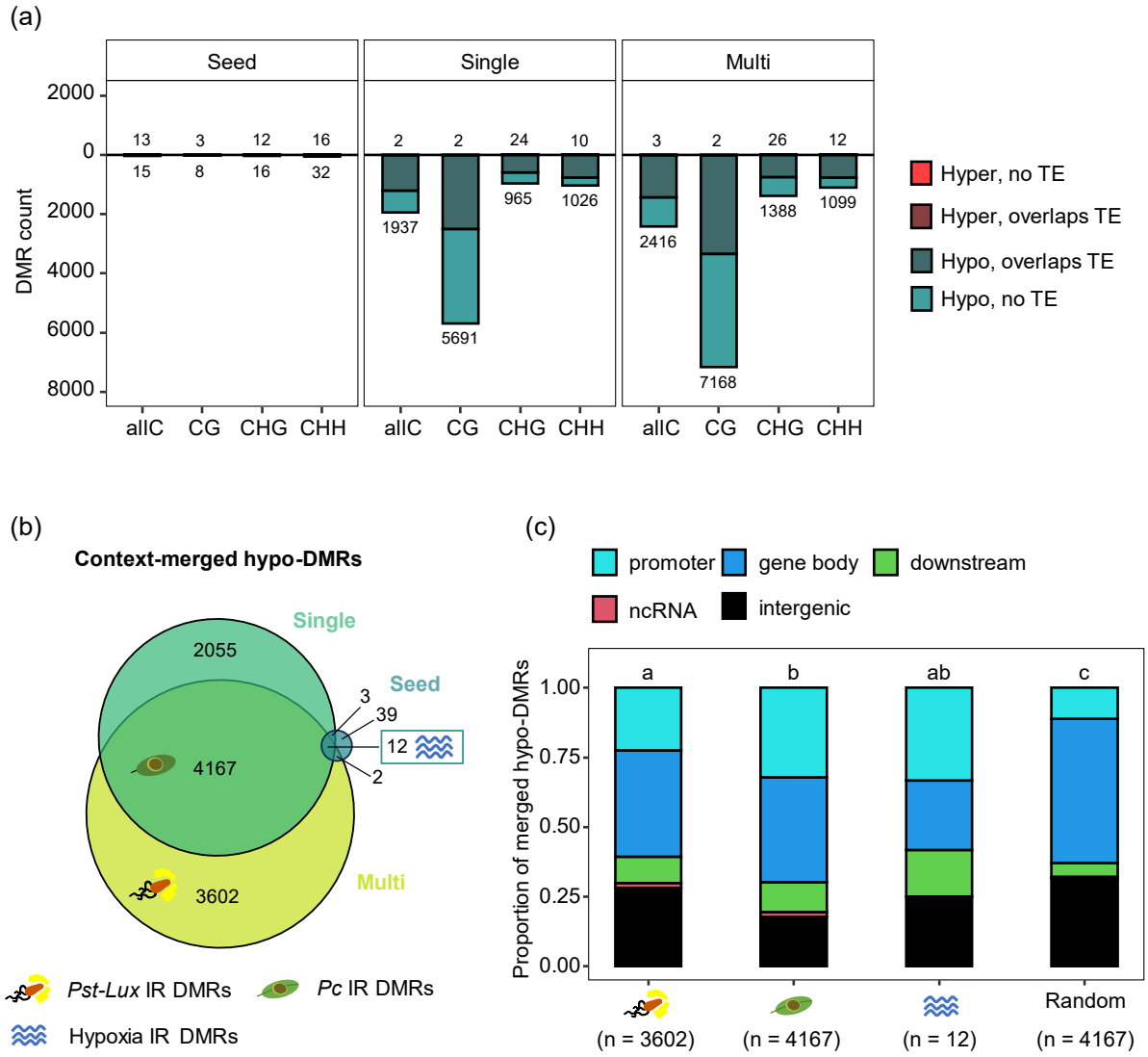


Figure 6.4 Analysis of differentially methylated regions (DMRs) of cytosine DNA methylation for Seed-, Single-, and Multi-treated *XVE:ROS1-YFP* plants (Figure 6.2), as compared to DMSO-treated plants. (a) Counts of DMRs for each E2 treatment regimen for allC, CG, CHG, and CHH sequence contexts. Hypomethylated DMRs (hypo-DMRs; blue) are shown below the y=0 line, and hypermethylated DMRs

(hyper-DMRs; red) are shown above. DMRs overlapping TEs are represented in dark blue (hypo-DMRs) and dark red (hyper-DMRs). Total count of all hyper- and hypo-DMRs for each context shown. (b) Venn diagram showing overlaps of hypo-DMRs merged across all cytosine sequence contexts for the three E2 treatment regiments (Seed, Single, Multi; filled in green, blue, and yellow respectively). Hypo-DMR sets are named and assigned a symbol based on their induced resistance (IR) phenotypes to *Pst-Lux* (bacteria symbol), *Pc* (leaf symbol), and Hypoxia (water symbol) (Figure 6.2). (c) Proportion of DMRs mapping to specific annotated regions of the Col-CEN genome (Naish et al., 2021) (see methods for definition of each annotation category). Symbols along the x-axis reflect the set of DMRs also annotated with a symbol in the Venn diagram (b). A random set of DMRs of equal size and number to the *Pc* IR DMRs was generated (see methods for details). DMR sets which share the same letter are not significantly different ($q > 0.05$) (Fisher's exact test).

6.4.4 DMRs Associated With Protein-Coding Genes are Associated With Different Molecular Functions Depending on the Spatiotemporal Patterning of ROS1 Induction

As downstream and promoter regions were particularly enriched in all sets of IR-associated DMRs, gene ontology (GO) enrichment for biological process and molecular function was performed on protein-coding genes within each IR-associated DMR category that contained flanking (promoter or downstream) hypo-DMRs. No biological process was enriched for any of the three DMR sets ($p_{adj} > 0.05$). However, *Pst-Lux* IR genic DMRs were enriched for the molecular function terms "cis-regulatory region sequence-specific DNA binding" and "ligase activity, forming carbon-nitrogen bonds", whereas *Pc* IR genic DMRs were enriched for the molecular function term "protein kinase regulator activity" (Supplementary Data 6.3). The 27 targeted genes associated with the term "cis-regulatory region sequence-specific DNA binding" in *Pst-Lux* IR DMRs all encoded transcription factors (TFs) from a variety of families with the largest representation occurring from the MYB TFs (n = 8), which have known functions in biotic stress responses (Dubos et al., 2010; Biswas et al., 2023) (Supplementary Data 6.3). Thus, 'Multi'-treatment, but not 'Single'-treatment, is enriched in the targeting of a suite of TF genes that may contribute to the specific establishment of IR against *Pst-Lux*.

Pc IR DMRs were enriched in flanking regions of genes with "protein kinase regulator activity" (n = 18), most of which were CYCLIN genes (n = 10), which have roles in regulating cell cycle progression and development (Wang et al., 2004; Zheng, 2022) (Supplementary Data 6.3). Their role, if any, in regulating resistance to *Pc* is unclear.

No significant GO enrichment ($p_{adj} > 0.05$) for molecular function was identified for Hypoxia IR genic DMRs. Of the 12 DMRs identified, half occurred in downstream or promoter regions of protein-coding genes (Supplementary Data 6.2; Figure 6.4c). DMRs in the promoter of 4 protein-coding genes were detected, of which 3 have unknown functions (AT3G30160,

AT1G29980, AT2G34730), and one encodes a thylakoid-membrane-bound protein, HIGH CHLOROPHYLL FLUORESCENCE 164 (HCF164), with no known role in abiotic stress tolerance. DMRs were identified in the downstream regions of *PWWP DOMAIN PROTEIN 3* (*PDP3*) and *XERICO* (*XER*). PDP3 and XER have functions in regulating flowering time and ABA metabolism, respectively. Interestingly, *XER* is induced by several abiotic stress treatments and mutation of this gene generates Arabidopsis plants with reduced levels of ABA, altered stomatal development, and stomata that are less responsive to abiotic stress (Ko et al., 2006; Vonapartis et al., 2022; Mohamed et al., 2023). Whilst ethylene is generally regarded the dominant plant hormone involved in hypoxic stress responses in Arabidopsis (Hartman et al., 2019, 2021), ABA-dependent signalling also plays a role (Hsu et al., 2011; Bui et al., 2020). Therefore, future research should explore the transcriptional regulation of *XER* under hypoxic conditions, and test mutants for hypoxic susceptibility. In addition to the *XER*-associated DMR, a Hypoxia IR hypo-DMR was present 1,118 bp upstream of the gene encoding BCL2-ASSOCIATED ATHANOGENE 1 (BAG1), a molecular chaperone that has roles in regulating growth under abiotic stress (Lee et al., 2016; Jiang et al., 2023). This too could contribute to the hypoxic tolerance difference observed in ‘Seed’ -treated plants.

6.5 Discussion

6.5.1 Stress-Specific IR Depends on the Developmental Stage and Frequency of ROS1 Induction

The establishment of stress memory enables plants to respond faster and/or stronger to a subsequent stress later in life. This induced resistance (IR) has been observed across a broad range of stress types, triggered by exposure to stress-associated stimuli such as mild stress or synthetic chemical treatments, and is thus, often regarded as a broad-spectrum defence mechanism in plants. However, recent research into the specificity of long-lasting stress-induced memory in Arabidopsis has revealed that induced resistance (IR) is only observed when the original stimulus closely resembles the subsequent stress that later challenges the same plants or their progeny (Luna et al., 2012; López Sánchez et al., 2021; Wilkinson et al., 2023). Instead, exposure to different stress types can lead to an induced susceptibility (IS) response (Luna et al., 2012; López Sánchez et al., 2021). Interestingly, the active DNA demethylase ROS1 has been identified as a crucial regulator of these long-lasting stress-specific memories induced by jasmonic acid (JA) and parental *Pst* infection (López Sánchez et al., 2016; Wilkinson et al., 2023). This indicates that ROS1, which is thought to target distinct genomic loci (Tang et al., 2016), can facilitate the establishment, maintenance and/or transcriptional activation of stress-specific memory. In support, the results in this Chapter demonstrate that the spatiotemporal context in which ROS1 is induced can signal for the establishment of stress specific memory.

A striking example of this comes from the IR phenotypes observed against the hemi-biotrophic pathogen *Pst-Lux* (Figure 6.2d). Whilst ‘Single’ and ‘Multi’ treatments both involved the induction of ROS1 7 days before infection with *Pst-Lux* (Figure 6.2a), ‘Single’ completely lacked IR whereas ‘Multi’ had a robust IR response. This indicates that multiple episodes of increased ROS1 activity are required for the establishment of stress memory against *Pst-Lux*. Interestingly in Chapter 3 of this thesis, I demonstrate that two inductions of ROS1 can result in the short-term memory against *Pst-Lux* (Chapter 3; Figure 3.5c). Why is ROS1-mediated IR against *Pst-Lux* only observed following two or more inductions? One possibility is that the secondary induction of ROS1 leads to distinct post-translational modifications (PTMs) or protein interactions that enable ROS1 to target different genomic regions. Interestingly, ROS1 has a predicted ubiquitination site (Maor et al., 2007; Kim et al., 2013), which may favour the degradation of ROS1, as well as a SUMOylation site which is known to promote the stability of this protein (Kong et al., 2020). Furthermore, ROS1 has been shown to interact with at least two distinct protein complexes that direct the targeting of ROS1 to distinct regions of the genome (Zhang et al., 2022). Thus, speculatively, initial increases in ROS1 activity may serve

to enhance the transcription of genes linked to specific protein complexes that interact with ROS1 and/or regulate the expression of genes responsible for modulating the PTMs of ROS1. This would then subsequently direct ROS1 to target different loci that may be associated with the defence against specific stress types, like *Pst*. However, to date, no studies have directly investigated the protein interactors of ROS1 under differential environmental stress conditions, nor the PTM modifications associated with ROS1 under such conditions. In fact, no studies have described the genome-wide binding sites of ROS1 under either basal or stressed conditions using high-throughput sequencing techniques such as chromatin immunoprecipitation followed by sequencing (ChIP-seq). Consequently, the regulation of ROS1 protein remains a significant knowledge gap in the field. The *XVE:ROS1-YFP* construct would be an excellent tool to better understand the dynamic regulation of ROS1 protein in the context of the establishment of stress-specific memory by modulating the spatiotemporal patterning of induction and subsequently isolating ROS1-YFP protein for ChIP-seq and/or proteomic characterisation using mass spectrometry-based techniques.

6.5.2 Seed-treatment has Significant Physiological Impacts Despite few Detectable DMRs

The limited number of DMRs associated with ‘Seed’ treatment is puzzling considering the strong phenotype associated with hypoxic stress (Figure 6.2f) and the significant reduction in plant size at 27 DAS (Figure 6.2c), that is not observable in Col-0 plants (Supplementary Figure 6.1). Principal component analysis (PCA) of DNA methylation at the single-cytosine level showed that ‘Seed’-treated samples showing higher variability in both CHG and CHH methylation than DMSO-treated samples. Accordingly, it is possible that ROS1-IR against flooding is determined by a partially stochastic and gradual DNA remethylation response involving AGO1-associated sRNAs, similarly as reported for long-lasting JA-IR in Arabidopsis (Wilkinson et al., 2019). In support, previous work identified both AGO1 and AGO4, key components of RNA directed DNA methylation (RdDM) pathways, as essential regulators of the transcriptional response to hypoxia stress in Arabidopsis (Loreti et al., 2020).

E2-induced expression of *ROS1* in germinating *XVE:ROS1-YFP* seeds can be expected to cause a reduction in DNA methylation in embryonic tissues. Indeed, a similar E2 seed treatment of rice expressing an XVE-controlled GUS transgene effectively induced the GUS expression in embryos (Chen et al., 2017). Interestingly, however, GUS staining was most prominent in the emerging radical, the first part of the embryo to emerge during germination. As the radical develops into the primary root, it is possible that the ‘Seed’ treatment resulted in more severe DNA hypomethylation in root tissues as compared to shoot tissues in germinating *XVE:ROS1-YFP* L5 seeds. As ONT-seq was conducted on DNA

extracted from aerial tissues, the potential long-lasting changes to DNA methylation in root tissues remain unknown. Considering that submergence creates a hypoxic environment that is most extreme in belowground tissues (Lee et al., 2011; Vashisht et al., 2011), DNA methylation levels in root tissues may play a key role in regulating transcriptional responses to hypoxic stress.

6.5.3 Interactions Between Developmental-Associated Epigenetic Changes and ROS1 Activity

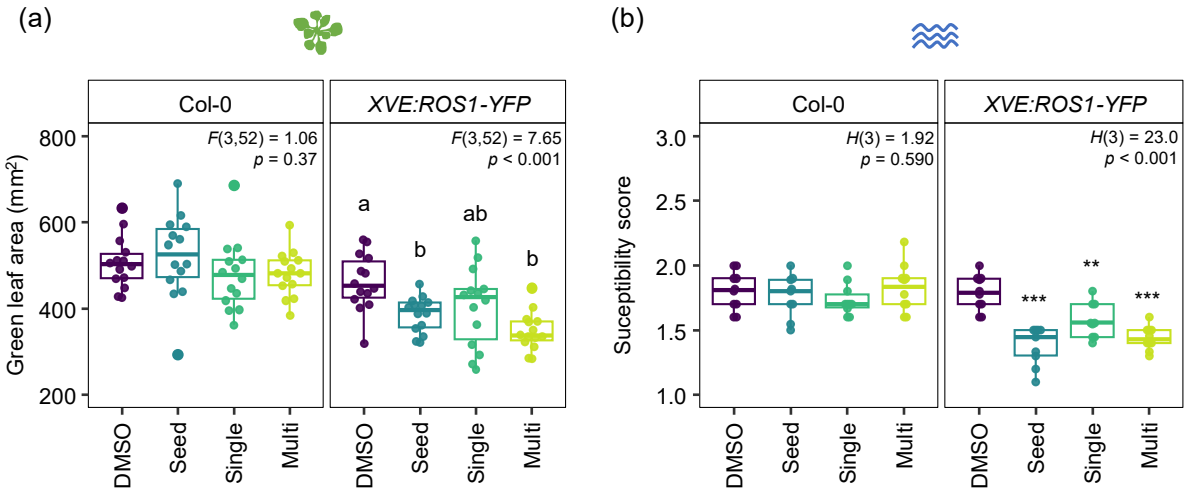
The spatiotemporal patterning of ROS1 induction may generate differential DNA hypomethylation patterns, as observed in this Chapter, which can be caused by a variety of different mechanisms. Firstly, DNA methylation and chromatin accessibility dynamically change during the development of Arabidopsis. For instance, CHH DNA methylation rapidly decreases in the pericentromeric regions upon seed imbibition (Kawakatsu et al., 2017), and this continues to change throughout the early stages of seedling development (Bouyer et al., 2017). Comparing average cytosine DNA methylation in DMSO-treated *XVE:ROS1-YFP* L5 plants of 4 weeks in this Chapter and DMSO-treated *XVE:ROS1-YFP* (L5) plants of ~3 weeks (20 days) in Chapter 5, average cytosine methylation increased by 10% across the genome (Supplementary Data 6.4). The biggest change occurred in CHH sequence contexts in pericentromeric regions, with a 39% increase in DNA methylation. Whilst these results are from independent experiments with slightly different DMSO treatment courses, it supports the idea that DNA methylation levels are dynamically changing during the development of Arabidopsis. Therefore, ROS1-targeted regions may also differ at distinct developmental stages. Furthermore, depending on where these targeted regions are in the genome, the ROS1-induced changes of DNA methylation may also be reset at different rates. Thus, the impact of ROS1-dependent DNA demethylation may vary depending on the developmental stage of the plant, during which ROS1 was active, which could, in turn, influence the type of stress memory established.

The basal level of resistance to environmental stress in Arabidopsis is known to vary across plant development and could be influenced by its epigenetic state at each stage of development. For instance, 2-week-old Arabidopsis seedlings are significantly more tolerant to hypoxic conditions compared to 3-week-old plants (Bui et al., 2020). This is associated with transcriptional silencing of stress-responsive genes which fall in regions of the genome which move towards a heterochromatic state with adulthood (Bui et al., 2020). In contrast, resistance to *Pst* and the fungal pathogen *Sclerotinia sclerotiorum* increases as Arabidopsis plants mature, a phenomena known as age-related resistance (Kus et al., 2002; Wilson et al., 2017; Xu et al., 2018). Although the epigenetic mechanisms underlying age-related resistance are

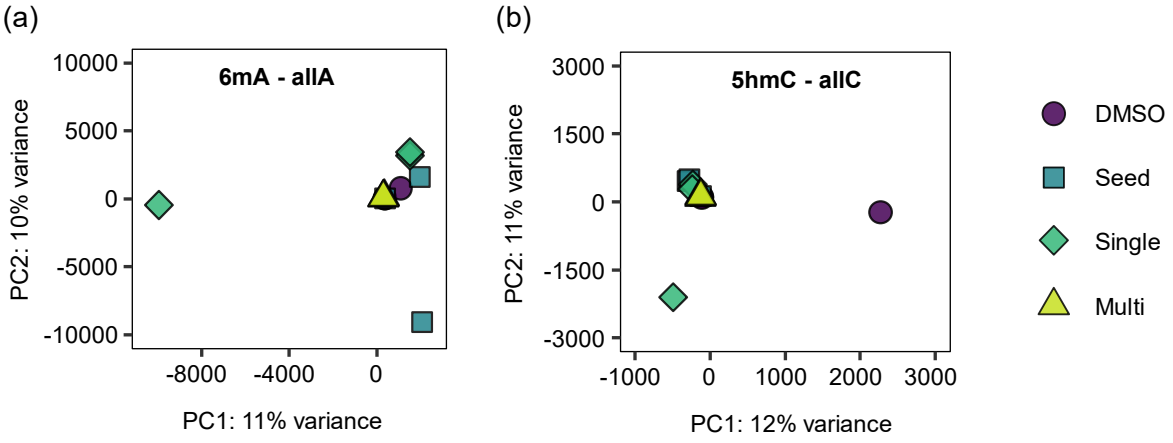
not well understood, DNA methylation and other epigenetic marks likely play a role in regulating the transcription of genes associated with this resistance. The spatiotemporal patterning of ROS1 activity could interact with these developmental epigenetic changes to alter age-related resistance to biotic and abiotic stress types. For example, ROS1-dependent reductions in DNA methylation may counteract the hypermethylation associated with the developmental transition from 3 weeks to 4 weeks (Supplementary Data 6.4), and thus prevent age-related susceptibility to hypoxia (Bui et al., 2020). Interestingly, heritable IR against *Pst-Lux* is more effectively established when plants in the previous generation are stressed at a young age (2 weeks) compared to an old age (5 weeks) which is likely a function of age-related resistance (Furci et al., 2023). This supports the findings presented in this Chapter, where Multi-treated plants, in which ROS1-YFP was first induced at 2 weeks, could establish within-generation IR against *Pst-Lux*, while those treated at 3 weeks (Single-treated) failed to establish IR despite generating greater reductions in DNA methylation at 4 weeks (Figure 6.3).

In conclusion, the results in this Chapter demonstrate that Arabidopsis can establish ROS1-dependent IR in a stress-specific manner, which is determined by the spatiotemporal patterning of ROS1 activity. Speculatively, in a natural system, certain environmental stresses may influence ROS1 transcription, localisation, and/or interactions with other proteins in differing ways to promote the establishment of stress-specific immune memory. To better understand this in natural systems, experimental conditions which effectively establish long-term immune memory in a stress-specific manner (e.g., López Sánchez et al., 2021) need to be identified and tested for their effects on the transcription and localisation of ROS1. Long-term, understanding the epigenetic changes required to establish stress-specific memory could facilitate the development of crop protection strategies that harness and enhance internal plant defence mechanisms, enabling more robust IR responses.

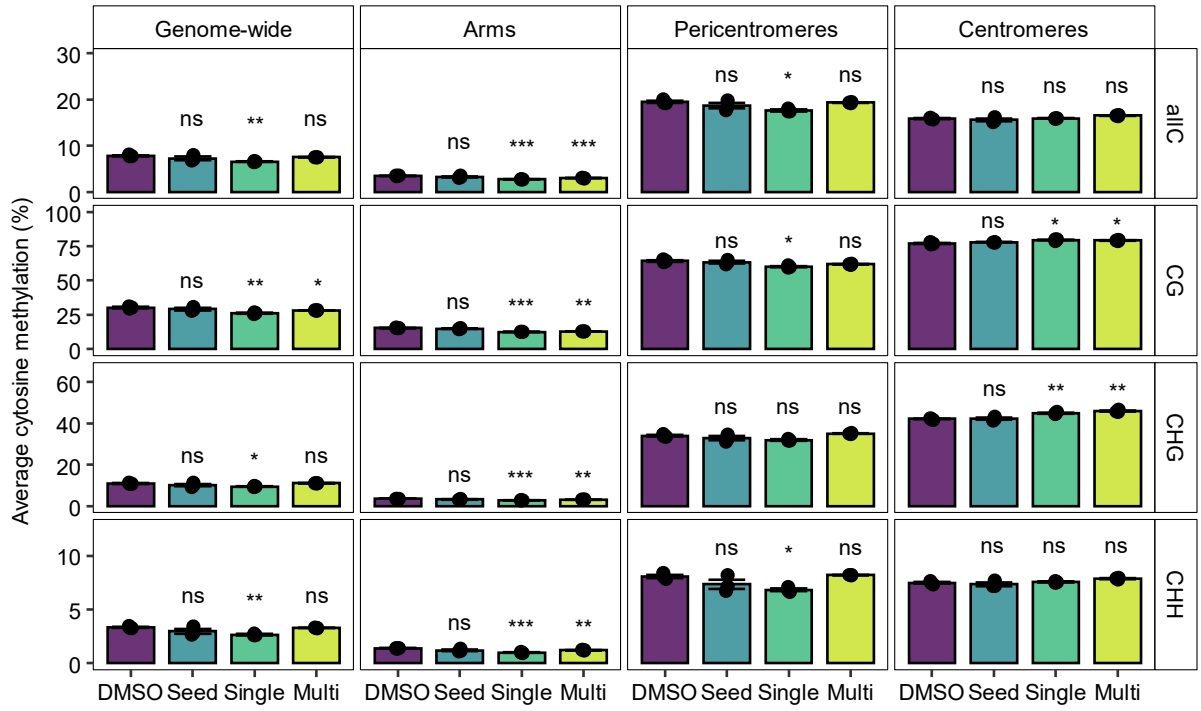
6.6 Supplementary Figures



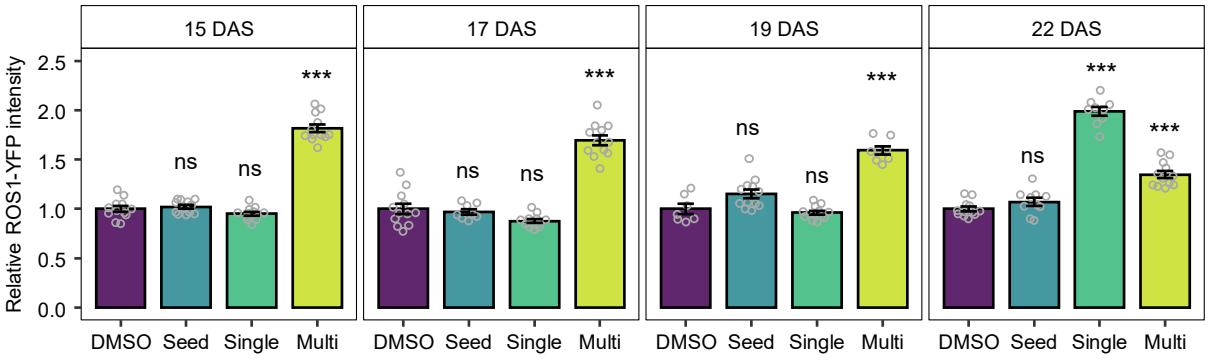
Supplementary Figure 6.1: Repeat experiment measuring GLA and hypoxia tolerance in E2-treated Col-0 plants.
(a) Repeat experiment (Figure 6.2) using both wild-type Col-0 and *XVE:ROS1-YFP* L5 plants. One-way ANOVA performed for each genotype to test the effect of treatment ('DMSO', 'Seed', 'Single', 'Multi') on green leaf area (GLA), ANOVA statistics are shown in the top right of each plot. Post-hoc analysis was performed following detection of a significant effect of treatment ($p < 0.05$), and groups which same the same letter are not statistically different from each other (Tukey, $p_{adj} < 0.05$). Plants were grown under different growth conditions (10 hour light period, 200 µmol s^{-1} m^{-2} light intensity, 21:19°C day:night temperatures), hence the differences in GLA compared to Figure 6.2c.



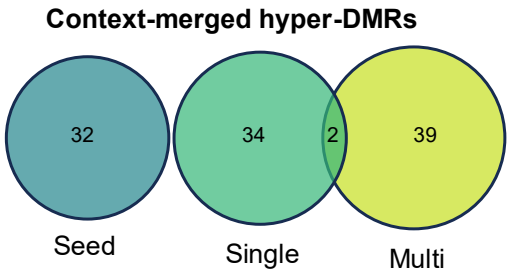
Supplementary Figure 6.2 PCA plots of different epigenetic marks under different E2 treatment regimens.
Principal component analysis (PCA) plots showing variation in (a) N^6-adenine DNA methylation (6mA) and (b) 5-hydroxymethylcytosine DNA methylation (5hmC) in 'DMSO'-, 'Seed'-, 'Single'-, and 'Multi'-treated plants (Figure 6.2b) at all adenosines and cytosines across the genome with ≥5 reads.



Supplementary Figure 6.3 Average cytosine DNA methylation levels under each E2 treatment regimen. Mean cytosine DNA methylation levels across all contexts (allC, CG, CHG, CHH) are shown for the entire genome (genome-wide), chromosome arms excluding (peri)centromeric regions (arms), pericentromeric regions excluding centromeric regions (pericentromeres), and centromeric regions only (centromeres). See methods for details on the definition of (peri)centromeric regions. Average cytosine methylation calculated using cytosines with ≥5 reads.



Supplementary Figure 6.4 ROS1-YFP signal intensity for 'DMSO'-, 'Seed'-, 'Single'-, and 'Multi'-treated plants at 15, 17, 19, and 22 days after sowing (DAS).
Measurements taken 24 hours after each spray treatment with 25 µM E2 or 0.05% DMSO at 14, 16, 19, and 21 DAS (as shown in Figure 6.2). Each point represents an independent plant per treatment per day (n = 7-12). Bar plot shows average relative ROS1-YFP intensity with error bars representing the standard error of the mean. ROS1-YFP signal intensity at each day expressed relative to the average ROS1-YFP intensity of DMSO-treated plants (background noise). Asterisks above treatments represent statistically different ROS1-YFP intensity as compared to DMSO-treated plants (t-test with FDR correction); * $q < 0.05$, ** $q < 0.01$, *** $q < 0.001$, ns $q > 0.05$.



Supplementary Figure 6.5 Venn diagram showing overlaps of hyper-DMRs merged across all cytosine sequence contexts for the three E2 treatment regimens.
'Seed', 'Single', 'Multi' filled in green, blue, and yellow respectively.

Chapter 7. General Discussion

Partially adapted from:

Hannan Parker, A.; Wilkinson, S. W. and Ton, J. (2022). Epigenetics: A Catalyst of Plant Immunity Against Pathogens. New Phytologist, 233(1): 66–83.

Authors:

Hannan Parker, A[1]; Wilkinson, S. W. [1]; Ton, J[1].

Affiliations:

[1] Plants, Photosynthesis and Soil, School of Biosciences, The University of Sheffield, Western Bank, Sheffield, S10 2TN, UK

Author Contributions:

AHP (the candidate) and JT proposed the original idea and scope of the review. All authors contributed to the writing of the article and/or production of the Figures in the original publication. All material in this Chapter, which is not in the publication, was created by AHP. JT reviewed and provided comments on the Chapter. Funding for the project came from studentships/grants awarded to AHP, SWW, and JT.

7.1 Summary and Highlights From Experimental Chapters

The aim of this PhD was to gain knowledge about the function of DNA (de)methylation in the establishment, maintenance, and eventual erasure of immune memory in plants. In Arabidopsis, it had been reported previously that mutation of the DNA demethylase ROS1 (*ros1* mutants) prevents plants from mounting effective basal defence and induced resistance (IR) against (hemi)biotrophic pathogens (Yu et al., 2013; López Sánchez et al., 2016; Halter et al., 2021). These reports not only revealed that ROS1-mediated DNA demethylation is important for the activation of plant immune responses, but also suggested involvement of ROS1 in the establishment and/or maintenance of immune memory that underpins IR. Indeed, the results presented in Chapter 3 show that progeny from *ros1-4* plants that had been severely infected by *Pseudomonas syringae* pv. *tomato* DC3000 (*Pst*) failed to exhibit h-IR against *Hyaloperonospora arabidopsidis* isolate Waco9 (*Hpa*), which confirms the findings reported by López Sánchez et al (2016).

However, while research using *ros1* mutants has significantly advanced our understanding of plant immune memory, relying on these mutants to demonstrate the absence of phenotype(s) has its limitations. Crucially, the permanent nature of genetic mutant lines makes it is impossible to resolve the phenotype in time – a fundamental aspect of plant immune memory. On the other hand, as previously reported (Chapter 1; Stassen et al., 2018; Catoni et al., 2022; Wilkinson et al., 2023) investigating the lasting changes in DNA methylation associated with stress-induced immune memory in wild-type plants is challenging due to significant variation among individuals, which is difficult to resolve when methylome sequencing is performed on replicates derived from many individual plants.

Therefore, this PhD project utilised an estradiol-inducible transgene to enable the controlled and transient activation of the Arabidopsis DNA demethylase ROS1. Chapter 3 demonstrates that transient inductions of ROS1 is a sufficient signal to establish short-term immune memory against *Pst* and *Hpa*. Furthermore, the observation that this memory did not persist over a generation highlighted an unexpected and novel benefit of using this tool to enhance our understanding of the erasure of immune memory. Chapter 4 explores the onset of immune memory through a combination of DNA methylation, whole transcriptome, and small RNA (sRNA) sequencing. Highlights from this work include a clear demonstration of highly consistent patterns of DNA demethylation following ROS1 induction, particularly at transposable element (TE)-gene boundaries in the chromosome arms. Furthermore, this was accompanied by a suppression of sRNAs associated with Pol IV-dependent RNA-directed DNA methylation (RdDM) and the activation of pathogen defence genes co-regulated by NPR1 and TGA transcription factors. Strikingly, the promoters of *NPR1* and *TRXh5*, which promotes

the nuclear translocation of NPR1 (Tada et al., 2008; Kneeshaw et al., 2014), were direct targets of ROS1. This was linked to increased transcription of both genes, suggesting that key components of the SA-dependent immune response are subject to antagonistic regulation by ROS1 and RdDM.

Chapter 5 further explores an observation from Chapter 4: ROS1 activation induces DNA demethylation in the chromosome arms, while its effect is significantly diminished in regions near the centromere (pericentromeric regions). Work by Furci et al. (2019) demonstrated that hypomethylation in pericentromeric regions is quantitatively associated with heritable primed defence against *Hpa*. Thus, long-read Oxford Nanopore Technology (ONT) sequencing and a centromere-complete genome assembly (Naish et al., 2021) were used to better characterise DNA methylation in these regions in the context of ROS1-driven immune memory. As observed in Chapter 4, ROS1 activity was associated with the loss of DNA methylation in chromosome arms, but there was a contrasting lack of demethylation – and even hypermethylation – of cytosines in (peri)centromeric regions. Interestingly, chemically inhibiting the activity of all DNA methyltransferases by 5-azacytidine, which causes strong reductions in DNA methylation in (peri)centromeric regions (Griffin et al., 2016), increased the durability of ROS1-driven immune memory by at least 1 week. This supports previous findings that DNA methylation can also function in the erasure of stress memory (Iwasaki and Paszkowski, 2014; Furci et al., 2023; Lee et al., 2023).

Finally, Chapter 6 demonstrates that ROS1 activity can establish stress-specific memory depending on the spatiotemporal patterning of its induction. While multiple events of ROS1 activity provided protection against all tested stress types – namely, the necrotrotphic pathogen *Plectosphaerella cucumerina* (*Pc*), hypoxic stress driven by submergence of plants (hypoxia), and the hemi-biotrophic pathogen *Pst* – a single induction at 3 weeks only established immune memory against *Pc*. Furthermore, induction of ROS1 in germinating seeds effectively induced resistance to hypoxic stress 4 weeks later but had no effect on resistance to *Pc* or *Pst*. Interestingly, despite the expectation of highly specific ROS1 targeting, as shown in Chapter 4 and by Tang et al. (2016), the long-term changes in DNA methylation varied significantly depending on the spatiotemporal patterning of induced ROS1 activity. These findings provide early but encouraging results that epigenetic changes can be exploited to engineer stress-specific resistance in plants, and thus provide novel avenues for crop protection strategies.

7.2 Control of DNA Methylation at Transposable Elements: From Regulation to Evolution of Plant Immunity

The previous Chapters of this thesis have built on an ever-growing body of evidence that implicates DNA demethylation at transposable elements (TEs) in the transcriptional activation and/or priming of defence genes to activate and enhance plant immune responses (López Sánchez et al., 2016; Wilkinson et al., 2019; Hannan Parker et al., 2022). Furthermore, as DNA methylation dynamically changes in response to environmental cues, such as the onset of disease (Wilkinson et al., 2019; Hannan Parker et al., 2022), the regulatory links between TE DNA (de)methylation and the expression of genes (as highlighted in Chapter 1; Figure 1.3) enables plants to fine-tune transcriptional responses to environmental stress.

However, the results of this thesis demonstrate that not all TEs are regulated equally. While DNA demethylation by ROS1 was enriched at TEs near protein-coding genes in the chromosome arms, there was contrasting rapid (re)methylation of TEs in (peri)centromeric regions (Chapter 5). A major difference between TEs in chromosome arms and those in pericentromeric regions is that, on average, the latter are evolutionarily young (Quesneville, 2020) and have greater potential to become transcriptionally active in the absence of DNA methylation (Panda and Slotkin, 2020; Shimada et al., 2024). Thus, the pericentromeric regions harbour a reservoir of potent large-effect (epi)mutagens that can accelerate the diversification and evolution of plant genomes (Quadrana et al., 2019; Baduel and Quadrana, 2021; Baduel et al., 2024). If left unchecked, runaway TE mobilisation can have deleterious consequences for both the organism's survival and the fate of the TEs themselves (Baduel and Quadrana, 2021). Therefore, tight epigenetic regulation of these elements, as appears to occur in response to ectopic increases in ROS1 activity, would be beneficial for both the host organism and the TEs themselves. Despite this, some TEs in the pericentromeric regions still showed transcriptional activation following increased ROS1 activity (Chapter 4, Chapter 5), which has interesting implications related to the evolution of new environmentally responsive genes.

Under extreme conditions, relaxation of epigenetic silencing mechanisms like DNA methylation, as occurs in response to severe biotic stress (Wilkinson et al., 2019; Hannan Parker et al., 2022; Furci et al., 2023), can facilitate adaptive evolution in plant genomes through TE mobilisation (Quadrana et al., 2019; Baduel and Quadrana, 2021; Baduel et al., 2024). For instance, a recent study demonstrated that natural Arabidopsis populations that are geographically associated with heavy herbicide use, such as railways and field margins, contain a TE insertion within the intron of the floral-repressor gene *FLOWERING LOCUS C* (*FLC*) (Raingeval et al., 2024). In response to stress, such as heat shock or a sub-lethal

herbicide doses, DNA methylation is lost at this intronic TE, leading to the production of non-functional transcripts in an IBM2-dependent manner (Raingeval et al., 2024). Thus, this evolutionarily young *FLC* TE insertion created a new locus of epigenetic regulation that enables plants to rapidly adapt to environmental stress, such as the application of herbicides, and shows signs of strong positive selection in these populations (Raingeval et al., 2024). Interestingly, the histone demethylase *IBM1* and the resistance gene *RPP7*, both of which regulate plant immune responses, also contain TE insertions within their introns that are regulated by the IBM2 complex (Chapter 1; Figure 1.2) (Lei et al., 2014; Lai et al., 2019; Lv et al., 2022). Notably, while the genetic sequence in the 7th intron of *IBM1* is diverse across flowering plants, there is strong conservation of enriched DNA methylation in this region, driven by the presence of TE insertions or simple repeat elements (Yinwen Zhang et al., 2024). Thus, the presence of these elements may have arisen independently in different species, resulting in a favourable epiallele that enables dynamic regulation of heterochromatin in response to stress (Chapter 1; Figure 1.2). In addition to intronic insertions, there are numerous other examples of transcriptional regulation by DNA methylation at TEs inserted in the promoters of defence genes (Deng et al., 2017; Barco et al., 2019; Halter et al., 2021; Baduel et al., 2024). Notably, in this PhD project, I identified that ROS1 removes DNA methylation at TEs in promoter of two master regulators of immunity against pathogens, *NPR1* and *TRXh5* (Chapter 4) (Tada et al., 2008; Kneeshaw et al., 2014; Zavaliev and Dong, 2024). Some TE families in plants show preferential insertion at genomic loci marked by H2A.Z, a chromatin modification associated with the promoters of environmentally responsive genes (Coleman-Derr and Zilberman, 2012; Quadrana et al., 2019; Hannan Parker et al., 2022). Finally, TEs and the epigenetic regulation they carry have been proposed to drive the diversification, evolution, and higher-order arrangements of resistance (*R*) gene clusters and biosynthetic gene clusters in plants (Wilkinson et al., 2019; Cawood and Ton, 2024).

Therefore, taken together, the multifaceted impacts of epigenetic regulation at TEs on the plant immune system can be conceptualised as a circular process operating over expanding time scales (Figure 7.1). Firstly, in Stage 1, TEs undergo stress-induced epigenetic changes that regulate the expression of genes in response to environmental cues. In Stage 2, this relaxed epigenetic silencing can promote the transcription and mobilisation of TEs, resulting in large-scale (epi)mutations within the genome. Finally, in Stage 3, some TEs that insert into genic regions may enhance the diversification of genes and/or create novel regulatory mechanisms that facilitate environmentally responsive transcription (i.e., Stage 1) (Figure 7.1).

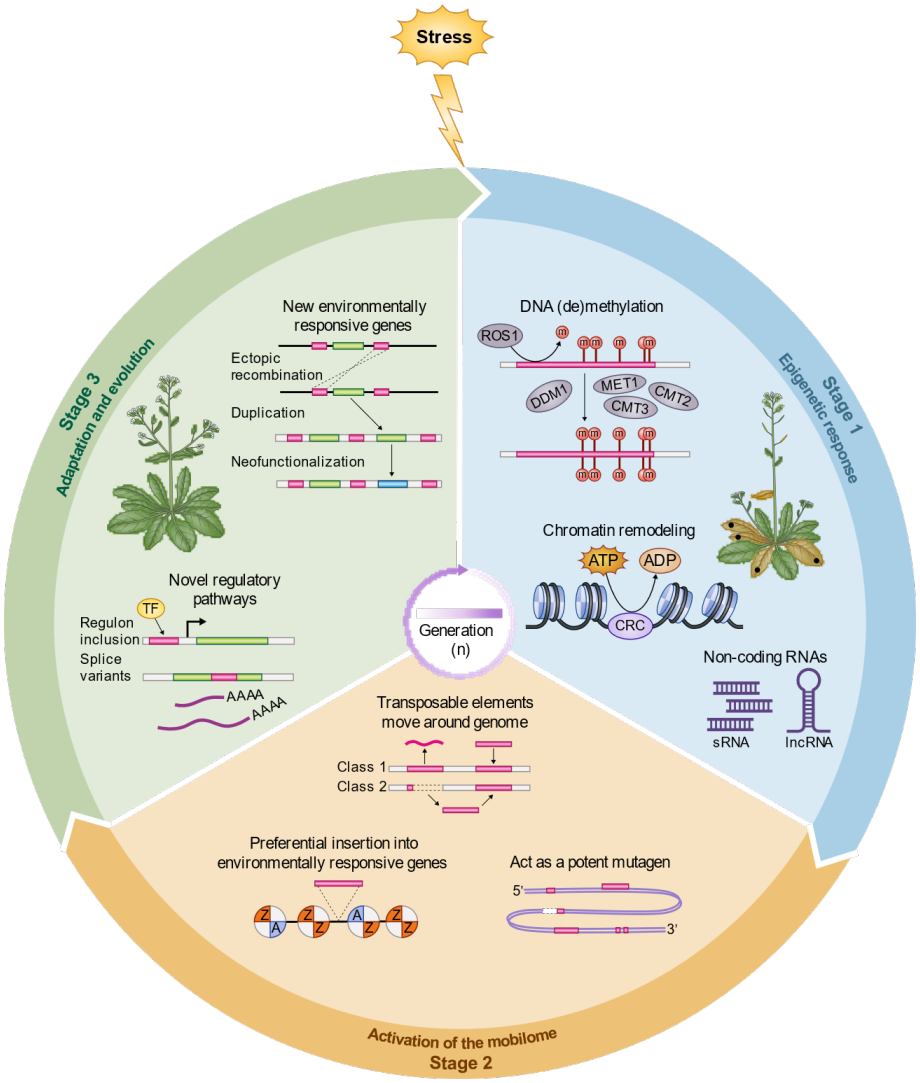


Figure 7.1 The effects of stress-induced epigenetic changes on short- and long-term adaptation of the plant immune system.
Stage 1: upon recognition of biotic stress, the plant epigenome undergoes changes that enable long-term up-regulation and/or priming of defence genes. This epigenetic stress memory can be transmitted to following generations and involves changes in the silencing of transposable elements (TEs) by DNA methylation, histone modifications, and noncoding RNAs. Stage 2: enduring stress increases the mutagenic activities of functional class I ('copy and paste') and class II ('cut and paste') TEs, collectively referred to as the mobilome. This process results in small and large mutations at the sites of excision (class 2) and insertion (class 1 and 2). Since TE integration is guided by the histone variant H2A.Z (Quadrana et al., 2019), the mobilome preferentially targets environmentally responsive genes (ERGs), which are enriched with H2A.Z (Coleman-Derr and Zilberman, 2012). Stage 3: mobilome-induced mutations increase the genetic diversity of ERGs, thereby accelerating the evolution of novel defence regulatory genes. Since TEs are tightly regulated by epigenetic mechanisms, the newly evolved defence genes and associated pathways remain under stress-dependent epigenetic control, thereby diversifying both the genetic and epigenetic regulatory potential to resist biotic stress.
Figure from Hannan Parker et al. (2022), *New Phytologist*, Volume 233, © 2021 The Authors, © 2021 New Phytologist Foundation. Reproduced with permission from Wiley.

7.3 Future Perspectives and Applicability of Epigenetic Immune Memory in Plant Protection

While future directions of research have been discussed at the end of each Chapter in this thesis, a larger question remains: how can the insights gained from studying epigenetically controlled immune memory be effectively translated into practical applications for enhancing crop resilience against pests and diseases? Furthermore, how can these technologies help tackle the challenges posed by climate change and emerging pathogens in agricultural systems?

The use of chemicals or synthetic microbiome communities that establish immune memory and facilitate induced resistance (IR) in crops holds significant promise (Yassin et al., 2021; Martins et al., 2023). Studying the epigenetic mechanisms underpinning the establishment, maintenance, and eventual erasure of this immune memory has potential to improve the efficacy and durability or these technologies. For instance, the research in this project led to the finding in Chapter 5, whereby chemically inhibiting DNA methyltransferases effectively prolonged the durability of immune memory established by ROS1 DNA demethylation. This also compliments a recent finding that reduced levels of DNA methylation in Arabidopsis can significantly enhance the efficacy of the IR-eliciting agent β-aminobutyric acid (BABA) (Cohen et al., 2016; Yassin et al., 2021; Lee et al., 2023). Thus, in the long term, the combined use of chemicals that (i) establish immune memory and (ii) supress DNA (re)methylation could reduce the dosage and frequency of chemical applications needed to effectively protect plants from pests and diseases.

Intentional disruption of epigenetic modifications may also accelerate the development of novel and adaptive (epi)genetic traits that enhance plant resistance to pests and diseases. For instance, TEgenesis (https://epibreed.com/) is an innovative breeding technology that involves chemically inhibiting Pol II prior to exposing plants to stress. This approach prevents the (re)silencing of transposable elements (TEs) that lose DNA methylation in response to stress, enabling their mobilisation and inducing large-scale (epi)mutations that can be selected for in breeding programmes (Thieme et al., 2017). Moreover, plants possess a remarkable ability to regenerate, allowing them to grow, divide, and differentiate into complete organisms from a mass of dedifferentiated cells known as callus. This can be utilised to induce epigenetic change in relatively few cells at the callus stage, which can then be inherited by other cells during plant regeneration, resulting in long-lasting phenotypic changes. For example, regeneration of sugarcane (*Saccharum spp*.) from calli treated with high doses of the DNA methyltransferase inhibitor 5-azacytidine has been shown to generate extensive (epi)genetic and phenotypic variation suitable for selection (Munsamy et al., 2013). Notably, plants

regenerated from 5-azacytidine-treated calli exhibited increased resistance to smut disease (*Sporisorium scitamineum*) six months later (Munsamy et al., 2013). Likewise, in Norway spruce (*Picea abies*), heat stress during zygotic or somatic embryogenesis (i.e., regeneration) results in plants with delayed bud set in summer, a phenomenon that can persist for up to 20 years and is associated with alterations in DNA methylation (Fossdal et al., 2024). In Arabidopsis, the epigenetic imprints of the organ (root or shoot) tissue used for regeneration are maintained into new tissues, and this feature persists even after sexual reproduction (Wibowo et al., 2018). Similarly, changes in DNA methylation associated with callus formation and regeneration have been shown to be heritable in other species including maize, rice, and triticale (Lee et al., 2024). Therefore, regenerating plants from tissue exposed to treatments that induce widespread changes in DNA methylation, such as severe stress, ectopic induction of ROS1, or chemical inhibition of DNA methyltransferases, may be an effective strategy to introduce very stable epigenetic changes in crops that provide long-lasting and heritable protection against pests and diseases.

Lastly, while TEs comprise only ~15% of Arabidopsis genome, they are far more abundant in most crop species, comprising more than 70% of the maize, wheat, and sunflower genomes (Pedro et al., 2021). Thus, genome-wide suppression of DNA methylation carries a more severe risk of causing deleterious (epi)mutations in these species. Consequently, epigenetic variation may also need to be introduced in a much more controlled manner to mitigate the potential for these harmful (epi)mutations. For instance, recent work in cassava demonstrated that introducing DNA methylation into the promoter of a gene associated with increased susceptibility to bacterial blight (*Xanthomonas phaseoli* pv. *manihotis*) effectively prevented its expression after infection, thereby enhancing disease resistance through precise epigenome editing. Significant advancements have been made in the development of sequence-targeted epimutagenic tools that harness CRISPR-based technologies that facilitate the introduction or removal of epigenetic marks (Gallego-Bartolomé et al., 2018; Papikian et al., 2019; Nuñez et al., 2021; Gardiner et al., 2022; M. Wang et al., 2023). However, while conceptually promising, these technologies have not been used outside of model species like Arabidopsis. Furthermore, they face two significant challenges (i) the complexity and constitutive expression of these transgenes makes them prone to post-transcriptional gene silencing by the host plant (M. Wang et al., 2023; Casas-Mollano et al., 2023), and (ii) legislative restrictions in many parts of the world surrounding genetically modified organisms (GMOs) limit the use of such technologies in commercial crop breeding.

Overall, building on the fundamental research presented in this thesis, alongside a stronger translational focus on epigenetic regulation in the establishment, maintenance, and

erasure of immune memory in crop species, will deepen our understanding of plant immunity and drive innovative advances in epigenetic-based plant protection technologies.

References

Agius, Fernanda; Kapoor, Avnish and Zhu, Jian-Kang (2006). Role of the Arabidopsis DNA Glycosylase/Lyase ROS1 in Active DNA Demethylation. *Proceedings of the National Academy of Sciences*, 103(31): 11796–11801.

Ahmed, Ikhlak; Sarazin, Alexis; Bowler, Chris; Colot, Vincent and Quesneville, Hadi (2011). Genome-Wide Evidence for Local DNA Methylation Spreading from Small RNA-Targeted Sequences in Arabidopsis. *Nucleic Acids Research*, 39(16): 6919–6931.

Akalin, Altuna; Franke, Vedran; Vlahoviček, Kristian; Mason, Christopher E. and Schübeler, Dirk (2015). Genomation: A Toolkit to Summarize, Annotate and Visualize Genomic Intervals. *Bioinformatics*, 31(7): 1127–1129.

Alonso, José M.; Stepanova, Anna N.; Leisse, Thomas J.; Kim, Christopher J.; Chen, Huaming; Shinn, Paul; Stevenson, Denise K.; Zimmerman, Justin; Barajas, Pascual; Cheuk, Rosa; Gadrinab, Carmelita; Heller, Collen; Jeske, Albert; Koesema, Eric; Meyers, Cristina C.; Parker, Holly; Prednis, Lance; Ansari, Yasser; Choy, Nathan; Deen, Hashim; Geralt, Michael; Hazari, Nisha; Hom, Emily; Karnes, Meagan; Mulholland, Celene; Ndubaku, Ral; Schmidt, Ian; Guzman, Plinio; Aguilar-Henonin, Laura; Schmid, Markus; Weigel, Detlef; Carter, David E.; Marchand, Trudy; Risseeuw, Eddy; Brogden, Debra; Zeko, Albana; Crosby, William L.; Berry, Charles C. and Ecker, Joseph R. (2003). Genome-Wide Insertional Mutagenesis of Arabidopsis Thaliana. *Science*, 301(5633): 653–657.

Alvarez, María E.; Nota, Florencia and Cambiagno, Damián A. (2010). Epigenetic Control of Plant Immunity. *Molecular Plant Pathology*, 11(4): 563–576.

Anders, Simon and Huber, Wolfgang (2010). Differential Expression Analysis for Sequence Count Data. *Genome Biology*, 11(10): R106.

Annacondia, Maria Luz; Markovic, Dimitrije; Reig-Valiente, Juan Luis; Scaltsoyiannes, Vassilis; Pieterse, Corné M. J.; Ninkovic, Velemir; Slotkin, R. Keith and Martinez, German (2021). Aphid Feeding Induces the Relaxation of Epigenetic Control and the Associated Regulation of the Defense Response in Arabidopsis. *The New Phytologist*, 230(3): 1185–1200.

Armstrong, Lyle (2014). *Epigenetics*. Garland Science.

Arnaiz, Ana; Romero-Puertas, Maria C.; Santamaria, M. Estrella; Rosa-Diaz, Irene; Arbona, Vicent; Muñoz, Alfonso; Grbic, Vojislava; González-Melendi, Pablo; Mar Castellano, M.; Sandalio, Luisa Maria; Martinez, Manuel and Diaz, Isabel (2023). The Arabidopsis Thioredoxin TRXh5regulates the S-Nitrosylation Pattern of the TIRK Receptor Being Both Proteins Essential in the Modulation of Defences to Tetranychus Urticae. *Redox Biology*, 67: 102902.

Arteaga-Vazquez, Mario; Sidorenko, Lyudmila; Rabanal, Fernando A.; Shrivistava, Roli; Nobuta, Kan; Green, Pamela J.; Meyers, Blake C. and Chandler, Vicki L. (2010). RNA-Mediated Trans-Communication Can Establish Paramutation at the B1 Locus in Maize. *Proceedings of the National Academy of Sciences of the United States of America*, 107(29): 12986–12991.

Asai, Shuta; Rallapalli, Ghanasyam; Piquerez, Sophie J. M.; Caillaud, Marie-Cécile; Furzer, Oliver J.; Ishaque, Naveed; Wirthmueller, Lennart; Fabro, Georgina; Shirasu, Ken and Jones, Jonathan D. G. (2014). Expression Profiling during Arabidopsis/Downy Mildew

Interaction Reveals a Highly-Expressed Effector That Attenuates Responses to Salicylic Acid. *PLOS Pathogens*, 10(10): e1004443.

Atighi, Mohammad Reza; Verstraeten, Bruno; Meyer, Tim De and Kyndt, Tina (2020). Genome-Wide DNA Hypomethylation Shapes Nematode Pattern-Triggered Immunity in Plants. *New Phytologist*, 227(2): 545–558.

Axtell, Michael J. (2013). ShortStack: Comprehensive Annotation and Quantification of Small RNA Genes. *RNA*, 19(6): 740–751.

Baduel, Pierre; Oliveira, Louna De; Caillieux, Erwann; Bohl-Viallefond, Grégoire; Messaoudi, Mounia El; Xu, Ciana; Barois, Matteo; Singh, Vipin; Sarazin, Alexis; Boccara, Martine; Gilbault, Elodie; France, Antoine de; Quadrana, Leandro; Loudet, Olivier and Colot, Vincent (2024). Transposable Elements Are Prevalent Vectors of Transgenerational Epigenetic Inheritance in Nature. : 2024.09.20.614076. available at https://www.biorxiv.org/content/10.1101/2024.09.20.614076v1 [14 October 2024].

Baduel, Pierre and Quadrana, Leandro (2021). Jumpstarting Evolution: How Transposition Can Facilitate Adaptation to Rapid Environmental Changes. *Current Opinion in Plant Biology*, 61: 102043.

Bannister, Andrew J. and Kouzarides, Tony (2011). Regulation of Chromatin by Histone Modifications. *Cell Research*, 21(3): 381–395.

Barco, Brenden; Kim, Yoseph and Clay, Nicole K. (2019). Expansion of a Core Regulon by Transposable Elements Promotes Arabidopsis Chemical Diversity and Pathogen Defense. *Nature Communications*, 10(1): 3444.

Bartee, Lisa; Malagnac, Fabienne and Bender, Judith (2001). Arabidopsis Cmt3 Chromomethylase Mutations Block Non-CG Methylation and Silencing of an Endogenous Gene. *Genes & Development*, 15(14): 1753–1758.

Bäurle, Isabel (2018). Can't Remember to Forget You: Chromatin-Based Priming of Somatic Stress Responses. *Seminars in Cell & Developmental Biology*, 83: 133–139.

Bebber, Daniel Patrick (2015). Range-Expanding Pests and Pathogens in a Warming World. *Annual Review of Phytopathology*, 53(Volume 53, 2015): 335–356.

Becker, Claude; Hagmann, Jörg; Müller, Jonas; Koenig, Daniel; Stegle, Oliver; Borgwardt, Karsten and Weigel, Detlef (2011). Spontaneous Epigenetic Variation in the Arabidopsis Thaliana Methylome. *Nature*, 480(7376): 245–249.

Benfey, P. N.; Ren, L. and Chua, N. H. (1990). Tissue-Specific Expression from CaMV 35S Enhancer Subdomains in Early Stages of Plant Development. *The EMBO journal*, 9(6): 1677–1684.

Berghoff, Bork A.; Karlsson, Torgny; Källman, Thomas; Wagner, E. Gerhart H. and Grabherr, Manfred G. (2017). RNA-Sequence Data Normalization through in Silico Prediction of Reference Genes: The Bacterial Response to DNA Damage as Case Study. *BioData Mining*, 10(1): 30.

Berthelier, Jérémy; Furci, Leonardo; Asai, Shuta; Sadykova, Munissa; Shimazaki, Tomoe; Shirasu, Ken and Saze, Hidetoshi (2023). Long-Read Direct RNA Sequencing Reveals Epigenetic Regulation of Chimeric Gene-Transposon Transcripts in Arabidopsis Thaliana. *Nature Communications*, 14(1): 3248.

Bhadouriya, Sneha Lata; Mehrotra, Sandhya; Basantani, Mahesh K.; Loake, Gary J. and Mehrotra, Rajesh (2021). Role of Chromatin Architecture in Plant Stress Responses: An Update. *Frontiers in Plant Science*, 11. available at https://www.frontiersin.org/articles/10.3389/fpls.2020.603380/full [6 March 2021].

Bidzinski, Przemyslaw; Ballini, Elsa; Ducasse, Aurélie; Michel, Corinne; Zuluaga, Paola; Genga, Annamaria; Chiozzotto, Remo and Morel, Jean-Benoit (2016). Transcriptional Basis of Drought-Induced Susceptibility to the Rice Blast Fungus Magnaporthe Oryzae. *Frontiers in Plant Science*, 7. available at https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2016.01558/full [23 October 2024].

Bigeard, Jean; Colcombet, Jean and Hirt, Heribert (2015). Signaling Mechanisms in Pattern-Triggered Immunity (PTI). *Molecular Plant*, 8(4): 521–539.

Biswas, Dew; Gain, Hena and Mandal, Arunava (2023). MYB Transcription Factor: A New Weapon for Biotic Stress Tolerance in Plants. *Plant Stress*, 10: 100252.

Bolger, Anthony M.; Lohse, Marc and Usadel, Bjoern (2014). Trimmomatic: A Flexible Trimmer for Illumina Sequence Data. *Bioinformatics*, 30(15): 2114–2120.

Borg, Michael; Jiang, Danhua and Berger, Frédéric (2021). Histone Variants Take Center Stage in Shaping the Epigenome. *Current Opinion in Plant Biology*, 61: 101991.

Borges, Filipe and Martienssen, Robert A. (2015). The Expanding World of Small RNAs in Plants. *Nature Reviews Molecular Cell Biology*, 16(12): 727–741.

Bourbousse, Clara; Vegesna, Neeraja and Law, Julie A. (2018). SOG1 Activator and MYB3R Repressors Regulate a Complex DNA Damage Network in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, 115(52): E12453–E12462.

Boutrot, Freddy and Zipfel, Cyril (2017). Function, Discovery, and Exploitation of Plant Pattern Recognition Receptors for Broad-Spectrum Disease Resistance. *Annual Review of Phytopathology*, 55: 257–286.

Bouyer, Daniel; Kramdi, Amira; Kassam, Mohamed; Heese, Maren; Schnittger, Arp; Roudier, François and Colot, Vincent (2017). DNA Methylation Dynamics during Early Plant Life. *Genome Biology*, 18: 179.

Bozarth, R. F. and Ross, A. F. (1964). Systemic Resistance Induced by Localized Virus Infections: Extent of Changes in Uninfected Plant Parts. *Virology*, 24(3): 446–455.

Briffa, Amy; Hollwey, Elizabeth; Shahzad, Zaigham; Moore, Jonathan D.; Lyons, David B.; Howard, Martin and Zilberman, Daniel (2023). Millennia-Long Epigenetic Fluctuations Generate Intragenic DNA Methylation Variance in *Arabidopsis* Populations. *Cell Systems*, 14(11): 953-967.e17.

Brooks, David M.; Bender, Carol L. and Kunkel, Barbara N. (2005). The Pseudomonas Syringae Phytotoxin Coronatine Promotes Virulence by Overcoming Salicylic Acid-Dependent Defences in Arabidopsis Thaliana. *Molecular Plant Pathology*, 6(6): 629–639.

Bruce, Toby J. A.; Matthes, Michaela C.; Napier, Johnathan A. and Pickett, John A. (2007). Stressful “Memories” of Plants: Evidence and Possible Mechanisms. *Plant Science*, 173(6): 603–608.

Brunel, D (1992). An Alternative, Rapid Method of Plant DNA Extraction for PCR Analyses. *Nucleic Acids Research*, 20(17): 4676.

Budimir, Jelena; Treffon, Katrin; Nair, Aswin; Thurow, Corinnna and Gatz, Christiane (2021). Redox-Active Cysteines in TGACG-BINDING FACTOR 1 (TGA1) Do Not Play a Role in Salicylic Acid or Pathogen-Induced Expression of TGA1-Regulated Target Genes in Arabidopsis Thaliana. *New Phytologist*, 230(6): 2420–2432.

Bui, Liem T.; Shukla, Vinay; Giorgi, Federico M.; Trivellini, Alice; Perata, Pierdomenico; Licausi, Francesco and Giuntoli, Beatrice (2020). Differential Submergence Tolerance between Juvenile and Adult Arabidopsis Plants Involves the ANAC017 Transcription Factor. *The Plant Journal*, 104(4): 979–994.

Bullard, James H.; Purdom, Elizabeth; Hansen, Kasper D. and Dudoit, Sandrine (2010). Evaluation of Statistical Methods for Normalization and Differential Expression in mRNA-Seq Experiments. *BMC Bioinformatics*, 11(1): 94.

van den Burg, Harrold A. and Takken, Frank L. W. (2009). Does Chromatin Remodeling Mark Systemic Acquired Resistance? *Trends in Plant Science*, 14(5): 286–294.

Burgess, Diane; Chow, Hiu Tung; Grover, Jeffrey W; Freeling, Michael and Mosher, Rebecca A (2022). Ovule siRNAs Methylate Protein-Coding Genes in Trans. *The Plant Cell*, 34(10): 3647–3664.

Calarco, Joseph P.; Borges, Filipe; Donoghue, Mark T. A.; Van Ex, Frédéric; Jullien, Pauline E.; Lopes, Telma; Gardner, Rui; Berger, Frédéric; Feijó, José A.; Becker, Jörg D. and Martienssen, Robert A. (2012). Reprogramming of DNA Methylation in Pollen Guides Epigenetic Inheritance via Small RNA. *Cell*, 151(1): 194–205.

Camacho, Christiam; Coulouris, George; Avagyan, Vahram; Ma, Ning; Papadopoulos, Jason; Bealer, Kevin and Madden, Thomas L. (2009). BLAST+: Architecture and Applications. *BMC Bioinformatics*, 10(1): 421.

Cambiagno, Damián A.; Nota, Florencia; Zavallo, Diego; Rius, Sebastián; Casati, Paula; Asurmendi, Sebastián and Alvarez, María E. (2018). Immune Receptor Genes and Pericentromeric Transposons as Targets of Common Epigenetic Regulatory Elements. *The Plant Journal*, 96(6): 1178–1190.

Cameron, Robin K. and Zaton, Kasia (2004). Intercellular Salicylic Acid Accumulation Is Important for Age-Related Resistance in *Arabidopsis* to *Pseudomonas Syringae*. *Physiological and Molecular Plant Pathology*, 65(4): 197–209.

Cao, H.; Bowling, S. A.; Gordon, A. S. and Dong, X. (1994). Characterization of an Arabidopsis Mutant That Is Nonresponsive to Inducers of Systemic Acquired Resistance. *The Plant Cell*, 6(11): 1583–1592.

Cao, Hui; Glazebrook, Jane; Clarke, Joseph D.; Volko, Sigrid and Dong, Xinnian (1997). The Arabidopsis NPR1 Gene That Controls Systemic Acquired Resistance Encodes a Novel Protein Containing Ankyrin Repeats. *Cell*, 88(1): 57–63.

Cao, Hui; Li, Xin and Dong, Xinnian (1998). Generation of Broad-Spectrum Disease Resistance by Overexpression of an Essential Regulatory Gene in Systemic Acquired Resistance. *Proceedings of the National Academy of Sciences of the United States of America*, 95(11): 6531–6536.

Cao, Xiaofeng and Jacobsen, Steven E. (2002). Locus-Specific Control of Asymmetric and CpNpG Methylation by the DRM and CMT3 Methyltransferase Genes. *Proceedings of the National Academy of Sciences*, 99(suppl_4): 16491–16498.

Casas-Mollano, J Armando; Zinselmeier, Matthew H; Sychla, Adam and Smanski, Michael J (2023). Efficient Gene Activation in Plants by the MoonTag Programmable Transcriptional Activator. *Nucleic Acids Research*, 51(13): 7083–7093.

Catoni, Marco; Alvarez-Venegas, Raul; Worrall, Dawn; Holroyd, Geoff; Barraza, Aarón; Luna, Estrella; Ton, Jurriaan and Roberts, Michael R. (2022). Long-Lasting Defence Priming by β-Aminobutyric Acid in Tomato Is Marked by Genome-Wide Changes in DNA Methylation. *Frontiers in Plant Science*, 13: 836326.

Cawood, George Lister and Ton, Jurriaan (2024). Decoding Resilience: Ecology, Regulation, and Evolution of Biosynthetic Gene Clusters. *Trends in Plant Science*. available at https://www.sciencedirect.com/science/article/pii/S1360138524002413 [21 October 2024].

Chaloner, Thomas M.; Gurr, Sarah J. and Bebber, Daniel P. (2021). Plant Pathogen Infection Risk Tracks Global Crop Yields under Climate Change. *Nature Climate Change*, 11(8): 710–715.

Chan, Ching and Zimmerli, Laurent (2019). The Histone Demethylase IBM1 Positively Regulates Arabidopsis Immunity by Control of Defense Gene Expression. *Frontiers in Plant Science*, 10. available at https://www.frontiersin.org/articles/10.3389/fpls.2019.01587 [29 January 2023].

Chan, Simon W.-L.; Henderson, Ian R.; Zhang, Xiaoyu; Shah, Govind; Chien, Jason S.-C. and Jacobsen, Steven E. (2006). RNAi, DRD1, and Histone Methylation Actively Target Developmentally Important Non-CG DNA Methylation in Arabidopsis. *PLOS Genetics*, 2(6): e83.

Chandler, Vicki L. (2007). Paramutation: From Maize to Mice. *Cell*, 128(4): 641–645.

Charvin, Magali; Halter, Thierry; Blanc-Mathieu, Romain; Barraud, Pierre; Aumont-Nicaise, Magali; Parcy, François and Navarro, Lionel (2023). Single-Cytosine Methylation at W-Boxes Repels Binding of WRKY Transcription Factors through Steric Hindrance. *Plant Physiology*, 192(1): 77–84.

Chen, Taiping and Li, En (2004). Structure and Function of Eukaryotic DNA Methyltransferases, 55–89, in: *Current Topics in Developmental Biology*. Stem Cells in Development and Disease. Academic Press. available at https://www.sciencedirect.com/science/article/pii/S0070215304600032 [11 October 2024].

Chen, Yun-Chu; Holmes, Eric C.; Rajniak, Jakub; Kim, Jung-Gun; Tang, Sandy; Fischer, Curt R.; Mudgett, Mary Beth and Sattely, Elizabeth S. (2018). N-Hydroxy-Pipecolic Acid Is a Mobile Metabolite That Induces Systemic Disease Resistance in Arabidopsis. *Proceedings of the National Academy of Sciences*, 115(21): E4920–E4929.

Chen, Yuxin; Chen, Yongsheng; Shi, Chunmei; Huang, Zhibo; Zhang, Yong; Li, Shengkang; Li, Yan; Ye, Jia; Yu, Chang; Li, Zhuo; Zhang, Xiuqing; Wang, Jian; Yang, Huanming; Fang, Lin and Chen, Qiang (2018). SOAPnuke: A MapReduce Acceleration-Supported Software for Integrated Quality Control and Preprocessing of High-Throughput Sequencing Data. *GigaScience*, 7(1): gix120.

Chen, Zaijie; Cheng, Qianqian; Hu, Chanquan; Guo, Xinrui; Chen, Ziqiang; Lin, Yan; Hu, Taijiao; Bellizzi, Maria; Lu, Guodong; Wang, Guo-Liang; Wang, Zonghua; Chen, Songbiao and Wang, Feng (2017). A Chemical-Induced, Seed-Soaking Activation Procedure for Regulated Gene Expression in Rice. *Frontiers in Plant Science*, 8. available at https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2017.01447/full [7 October 2024].

Cheng, Chia-Yi; Krishnakumar, Vivek; Chan, Agnes P.; Thibaud-Nissen, Françoise; Schobel, Seth and Town, Christopher D. (2017). Araport11: A Complete Reannotation of the Arabidopsis Thaliana Reference Genome. *The Plant Journal*, 89(4): 789–804.

Chengara, Anoop; Nikolov, Alex D. and Wasan, Darsh T. (2007). Spreading of a Water Drop Triggered by the Surface Tension Gradient Created by the Localized Addition of a Surfactant. *Industrial & Engineering Chemistry Research*, 46(10): 2987–2995.

Choi, Jaemyung; Lyons, David B and Zilberman, Daniel (2021). Histone H1 Prevents Non-CG Methylation-Mediated Small RNA Biogenesis in Arabidopsis Heterochromatin (R. Amasino, D. Weigel, C. S. Pikaard, and M. D. Nodine, Eds.). *eLife*, 10: e72676.

Choi, Yeonhee; Gehring, Mary; Johnson, Lianna; Hannon, Mike; Harada, John J.; Goldberg, Robert B.; Jacobsen, Steven E. and Fischer, Robert L. (2002). DEMETER, a DNA Glycosylase Domain Protein, Is Required for Endosperm Gene Imprinting and Seed Viability in Arabidopsis. *Cell*, 110(1): 33–42.

Chow, Hiu Tung and Mosher, Rebecca A (2023). Small RNA-Mediated DNA Methylation during Plant Reproduction. *The Plant Cell*, 35(6): 1787–1800.

Clapier, Cedric R.; Iwasa, Janet; Cairns, Bradley R. and Peterson, Craig L. (2017). Mechanisms of Action and Regulation of ATP-Dependent Chromatin-Remodelling Complexes. *Nature Reviews Molecular Cell Biology*, 18(7): 407–422.

Clough, Steven J. and Bent, Andrew F. (1998). Floral Dip: A Simplified Method for -Mediated Transformation Of. *The Plant Journal*, 16(6): 735–743.

Cohen, Yigal; Vaknin, Moshe and Mauch-Mani, Brigitte (2016). BABA-Induced Resistance: Milestones along a 55-Year Journey. *Phytoparasitica*, 44(4): 513–538.

Coker, Tim L. R.; Rozsypálek, Jiří; Edwards, Anne; Harwood, Tony P.; Butfoy, Louise and Buggs, Richard J. A. (2019). Estimating Mortality Rates of European Ash (Fraxinus Excelsior) under the Ash Dieback (Hymenoscyphus Fraxineus) Epidemic. *PLANTS, PEOPLE, PLANET*, 1(1): 48–58.

Coleman-Derr, Devin and Zilberman, Daniel (2012). Deposition of Histone Variant H2A.Z within Gene Bodies Regulates Responsive Genes. *PLOS Genetics*, 8(10): e1002988.

Combes, Matt; Webber, Joan and Boddy, Lynne (2024). Current Understanding and Future Prospects for Ash Dieback Disease with a Focus on Britain. *Forestry: An International Journal of Forest Research*, 97(5): 678–691.

Conover, W. J. and Iman, R. L. (1979). *Multiple-Comparisons Procedures. Informal Report*. Los Alamos National Lab. (LANL), Los Alamos, NM (United States). available at https://www.osti.gov/biblio/6057803 [26 September 2024].

Conrath, Uwe (2011). Molecular Aspects of Defence Priming. *Trends in Plant Science*, 16(10): 524–531.

Cooper, Agatha and Ton, Jurriaan (2022). Immune Priming in Plants: From the Onset to Transgenerational Maintenance. *Essays in Biochemistry*, 66(5): 635–646.

Corwin, Jason A. and Kliebenstein, Daniel J. (2017). Quantitative Resistance: More Than Just Perception of a Pathogen. *The Plant Cell*, 29(4): 655–665.

Crevillén, Pedro; Yang, Hongchun; Cui, Xia; Greeff, Christiaan; Trick, Martin; Qiu, Qi; Cao, Xiaofeng and Dean, Caroline (2014). Epigenetic Reprogramming That Prevents Transgenerational Inheritance of the Vernalized State. *Nature*, 515(7528): 587–590.

Crisp, Peter A.; Ganguly, Diep; Eichten, Steven R.; Borevitz, Justin O. and Pogson, Barry J. (2016). Reconsidering Plant Memory: Intersections between Stress Recovery, RNA Turnover, and Epigenetics. *Science Advances*, 2(2): e1501340.

Cuerda-Gil, Diego and Slotkin, R. Keith (2016). Non-Canonical RNA-Directed DNA Methylation. *Nature Plants*, 2(11): 1–8.

Czechowski, Tomasz; Stitt, Mark; Altmann, Thomas; Udvardi, Michael K. and Scheible, Wolf-Rüdiger (2005). Genome-Wide Identification and Testing of Superior Reference Genes for Transcript Normalization in Arabidopsis. *Plant Physiology*, 139(1): 5–17.

Danecek, Petr; Bonfield, James K; Liddle, Jennifer; Marshall, John; Ohan, Valeriu; Pollard, Martin O; Whitwham, Andrew; Keane, Thomas; McCarthy, Shane A; Davies, Robert M and Li, Heng (2021). Twelve Years of SAMtools and BCFtools. *GigaScience*, 10(2): giab008.

De Kesel, Jonas; Conrath, Uwe; Flors, Víctor; Luna, Estrella; Mageroy, Melissa H.; Mauch-Mani, Brigitte; Pastor, Victoria; Pozo, María J.; Pieterse, Corné M. J.; Ton, Jurriaan and Kyndt, Tina (2021). The Induced Resistance Lexicon: Do's and Don'ts. *Trends in Plant Science*. available at https://www.sciencedirect.com/science/article/pii/S1360138521000029 [6 March 2021].

De Lucia, Filomena; Crevillen, Pedro; Jones, Alexandra M. E.; Greb, Thomas and Dean, Caroline (2008). A PHD-Polycomb Repressive Complex 2 Triggers the Epigenetic Silencing of FLC during Vernalization. *Proceedings of the National Academy of Sciences*, 105(44): 16831–16836.

De Schutter, Kristof; Joubès, Jérôme; Cools, Toon; Verkest, Aurine; Corellou, Florence; Babiychuk, Elena; Van Der Schueren, Els; Beeckman, Tom; Kushnir, Sergeï; Inzé, Dirk and De Veylder, Lieven (2007). Arabidopsis WEE1 Kinase Controls Cell Cycle Arrest in Response to Activation of the DNA Integrity Checkpoint. *The Plant Cell*, 19(1): 211–225.

Delgado-Baquerizo, Manuel; Guerra, Carlos A.; Cano-Díaz, Concha; Egidi, Eleonora; Wang, Jun-Tao; Eisenhauer, Nico; Singh, Brajesh K. and Maestre, Fernando T. (2020). The Proportion of Soil-Borne Pathogens Increases with Warming at the Global Scale. *Nature Climate Change*, 10(6): 550–554.

Deng, Yiwen; Zhai, Keran; Xie, Zhen; Yang, Dongyong; Zhu, Xudong; Liu, Junzhong; Wang, Xin; Qin, Peng; Yang, Yuanzhu; Zhang, Guomin; Li, Qun; Zhang, Jianfu; Wu, Shuangqing; Milazzo, Joëlle; Mao, Bizeng; Wang, Ertao; Xie, Huaan; Tharreau, Didier and He, Zuhua (2017). Epigenetic Regulation of Antagonistic Receptors Confers Rice Blast Resistance with Yield Balance. *Science*, 355(6328): 962–965.

Deremetz, Aurélie; Le Roux, Clémentine; Idir, Yassir; Brousse, Cécile; Agorio, Astrid; Gy, Isabelle; Parker, Jane E. and Bouché, Nicolas (2019). Antagonistic Actions of FPA and IBM2 Regulate Transcript Processing from Genes Containing Heterochromatin. *Plant Physiology*, 180(1): 392–403.

Després, Charles; Chubak, Catherine; Rochon, Amanda; Clark, Rena; Bethune, Terry; Desveaux, Darrell and Fobert, Pierre R. (2003). The Arabidopsis NPR1 Disease Resistance Protein Is a Novel Cofactor That Confers Redox Regulation of DNA Binding Activity to the Basic Domain/Leucine Zipper Transcription Factor TGA1. *The Plant Cell*, 15(9): 2181.

Després, Charles; DeLong, Catherine; Glaze, Sarah; Liu, Enwu and Fobert, Pierre R. (2000). The Arabidopsis NPR1/NIM1 Protein Enhances the DNA Binding Activity of a Subgroup of the TGA Family of bZIP Transcription Factors. *The Plant Cell*, 12(2): 279–290.

Ding, Yuli; Sun, Tongjun; Ao, Kevin; Peng, Yujun; Zhang, Yaxi; Li, Xin and Zhang, Yuelin (2018). Opposite Roles of Salicylic Acid Receptors NPR1 and NPR3/NPR4 in Transcriptional Regulation of Plant Immunity. *Cell*, 173(6): 1454-1467.e15.

Dobin, Alexander; Davis, Carrie A.; Schlesinger, Felix; Drenkow, Jorg; Zaleski, Chris; Jha, Sonali; Batut, Philippe; Chaisson, Mark and Gingeras, Thomas R. (2013). STAR: Ultrafast Universal RNA-Seq Aligner. *Bioinformatics*, 29(1): 15–21.

Dowen, Robert H.; Pelizzola, Mattia; Schmitz, Robert J.; Lister, Ryan; Dowen, Jill M.; Nery, Joseph R.; Dixon, Jack E. and Ecker, Joseph R. (2012). Widespread Dynamic DNA Methylation in Response to Biotic Stress. *Proceedings of the National Academy of Sciences*, 109(32): E2183–E2191.

Du, Jiamu; Johnson, Lianna M.; Jacobsen, Steven E. and Patel, Dinshaw J. (2015). DNA Methylation Pathways and Their Crosstalk with Histone Methylation. *Nature Reviews Molecular Cell Biology*, 16(9): 519–532.

Du, Jiamu; Zhong, Xuehua; Bernatavichute, Yana V.; Stroud, Hume; Feng, Suhua; Caro, Elena; Vashisht, Ajay A.; Terragni, Jolyon; Chin, Hang Gyeong; Tu, Andy; Hetzel, Jonathan; Wohlschlegel, James A.; Pradhan, Sriharsa; Patel, Dinshaw J. and Jacobsen, Steven E. (2012). Dual Binding of Chromomethylase Domains to H3K9me2-Containing Nucleosomes Directs DNA Methylation in Plants. *Cell*, 151(1): 167–180.

Du, Xuan; Yang, Zhenlin; Xie, Guohui; Wang, Changshi; Zhang, Laixing; Yan, Kaige; Yang, Maojun; Li, Sisi; Zhu, Jian-Kang and Du, Jiamu (2023). Molecular Basis of the Plant ROS1-Mediated Active DNA Demethylation. *Nature Plants*, 9(2): 271–279.

Duan, Cheng-Guo; Wang, Xingang; Tang, Kai; Zhang, Huiming; Mangrauthia, Satendra K.; Lei, Mingguang; Hsu, Chuan-Chih; Hou, Yueh-Ju; Wang, Chunguo; Li, Yan; Tao, W. Andy and Zhu, Jian-Kang (2015). MET18 Connects the Cytosolic Iron-Sulfur Cluster Assembly Pathway to Active DNA Demethylation in Arabidopsis. *PLOS Genetics*, 11(10): e1005559.

Duan, Cheng-Guo; Wang, Xingang; Zhang, Lingrui; Xiong, Xiansong; Zhang, Zhengjing; Tang, Kai; Pan, Li; Hsu, Chuan-Chih; Xu, Huawei; Tao, W. Andy; Zhang, Heng and Zhu, Jian-Kang (2017). A Protein Complex Regulates RNA Processing of Intronic Heterochromatin-Containing Genes in Arabidopsis. *Proceedings of the National Academy of Sciences*, 114(35): E7377–E7384.

Dubos, Christian; Stracke, Ralf; Grotewold, Erich; Weisshaar, Bernd; Martin, Cathie and Lepiniec, Loïc (2010). MYB Transcription Factors in *Arabidopsis*. *Trends in Plant Science*, 15(10): 573–581.

Duc, Céline; Sherstnev, Alexander; Cole, Christian; Barton, Geoffrey J. and Simpson, Gordon G. (2013). Transcription Termination and Chimeric RNA Formation Controlled by Arabidopsis Thaliana FPA. *PLOS Genetics*, 9(10): e1003867.

Durrant, Wendy E.; Wang, Shui and Dong, Xinnian (2007). Arabidopsis SNI1 and RAD51D Regulate Both Gene Transcription and DNA Recombination during the Defense Response. *Proceedings of the National Academy of Sciences*, 104(10): 4223–4227.

Erb, Matthias and Reymond, Philippe (2019). Molecular Interactions Between Plants and Insect Herbivores. *Annual Review of Plant Biology*, 70(Volume 70, 2019): 527–557.

Erdmann, Robert M. and Picard, Colette L. (2020). RNA-Directed DNA Methylation. *PLOS Genetics*, 16(10): e1009034.

Erdmann, Robert M.; Souza, Amanda L.; Clish, Clary B. and Gehring, Mary (2014). 5-Hydroxymethylcytosine Is Not Present in Appreciable Quantities in Arabidopsis DNA. *G3: Genes|Genomes|Genetics*, 5(1): 1–8.

Ewels, Philip; Magnusson, Måns; Lundin, Sverker and Käller, Max (2016). MultiQC: Summarize Analysis Results for Multiple Tools and Samples in a Single Report. *Bioinformatics*, 32(19): 3047–3048.

Fahlgren, Noah; Howell, Miya D.; Kasschau, Kristin D.; Chapman, Elisabeth J.; Sullivan, Christopher M.; Cumbie, Jason S.; Givan, Scott A.; Law, Theresa F.; Grant, Sarah R.; Dangl, Jeffery L. and Carrington, James C. (2007). High-Throughput Sequencing of Arabidopsis microRNAs: Evidence for Frequent Birth and Death of MIRNA Genes. *PLoS ONE*, 2(2): e219.

Fan, Jun; Crooks, Casey and Lamb, Chris (2008). High-Throughput Quantitative Luminescence Assay of the Growth in Planta of Pseudomonas Syringae Chromosomally Tagged with Photorhabdus Luminescens luxCDABE. *The Plant Journal*, 53(2): 393–399.

Fan, Weihua and Dong, Xinnian (2002). In Vivo Interaction between NPR1 and Transcription Factor TGA2 Leads to Salicylic Acid–Mediated Gene Activation in Arabidopsis. *The Plant Cell*, 14(6): 1377–1389.

Fang, Jian; Jiang, Jianjun; Leichter, Sarah M.; Liu, Jie; Biswal, Mahamaya; Khudaverdyan, Nelli; Zhong, Xuehua and Song, Jikui (2022). Mechanistic Basis for Maintenance of CHG DNA Methylation in Plants. *Nature Communications*, 13(1): 3877.

Fang, Jian; Leichter, Sarah M.; Jiang, Jianjun; Biswal, Mahamaya; Lu, Jiuwei; Zhang, Zhi-Min; Ren, Wendan; Zhai, Jixian; Cui, Qiang; Zhong, Xuehua and Song, Jikui (2021). Substrate Deformation Regulates DRM2-Mediated DNA Methylation in Plants. *Science Advances*, 7(23): eabd9224.

FAO (2022). *FAO's Plant Production and Protection Division*. FAO ; available at https://openknowledge.fao.org/handle/20.500.14283/cc2447en [12 October 2024].

FAO (2024). Pesticides Use and Trade, 1990–2022. available at https://openknowledge.fao.org/handle/20.500.14283/cd1486en [12 October 2024].

Feng, Hao; Conneely, Karen N. and Wu, Hao (2014). A Bayesian Hierarchical Model to Detect Differentially Methylated Loci from Single Nucleotide Resolution Sequencing Data. *Nucleic Acids Research*, 42(8): e69.

Feng, Suhua; Cokus, Shawn J.; Schubert, Veit; Zhai, Jixian; Pellegrini, Matteo and Jacobsen, Steven E. (2014). Genome-Wide Hi-C Analyses in Wild Type and Mutants Reveal High-Resolution Chromatin Interactions in Arabidopsis. *Molecular cell*, 55(5): 694–707.

Fossdal, Carl Gunnar; Krokene, Paal; Olsen, Jorunn Elisabeth; Strimbeck, Richard; Viejo, Marcos; Yakovlev, Igor and Mageroy, Melissa H (2024). Epigenetic Stress Memory in Gymnosperms. *Plant Physiology*, 195(2): 1117–1133.

Franco-Zorrilla, José Manuel; Valli, Adrián; Todesco, Marco; Mateos, Isabel; Puga, María Isabel; Rubio-Somoza, Ignacio; Leyva, Antonio; Weigel, Detlef; García, Juan Antonio and Paz-Ares, Javier (2007). Target Mimicry Provides a New Mechanism for Regulation of microRNA Activity. *Nature Genetics*, 39(8): 1033–1037.

Fu, Zheng Qing and Dong, Xinnian (2013). Systemic Acquired Resistance: Turning Local Infection into Global Defense. *Annual Review of Plant Biology*, 64(1): 839–863.

Furci, L.; Pascual-Pardo, D.; Tirot, L.; Zhang, P.; Hannan Parker, A. and Ton, J. (2023). Heritable Induced Resistance in Arabidopsis Thaliana: Tips and Tools to Improve Effect Size and Reproducibility. *Plant Direct*, 7(8): e523.

Furci, Leonardo; Jain, Ritushree; Stassen, Joost; Berkowitz, Oliver; Whelan, James; Roquis, David; Baillet, Victoire; Colot, Vincent; Johannes, Frank and Ton, Jurriaan (2019). Identification and Characterisation of Hypomethylated DNA Loci Controlling Quantitative Resistance in Arabidopsis (D. J. Kliebenstein and C. S. Hardtke, Eds.). *eLife*, 8: e40655.

Furci, Leonardo; Pascual-Pardo, David and Ton, Jurriaan (2021). A Rapid and Non-Destructive Method for Spatial–Temporal Quantification of Colonization by Pseudomonas Syringae Pv. Tomato DC3000 in Arabidopsis and Tomato. *Plant Methods*, 17(1): 126.

Galili, Tal (2015). Dendextend: An R Package for Visualizing, Adjusting and Comparing Trees of Hierarchical Clustering. *Bioinformatics*, 31(22): 3718–3720.

Gallego-Bartolomé, Javier; Gardiner, Jason; Liu, Wanlu; Papikian, Ashot; Ghoshal, Basudev; Kuo, Hsuan Yu; Zhao, Jenny Miao-Chi; Segal, David J. and Jacobsen, Steven E. (2018). Targeted DNA Demethylation of the Arabidopsis Genome Using the Human TET1 Catalytic Domain. *Proceedings of the National Academy of Sciences*, 115(9): E2125–E2134.

Gallusci, Philippe; Agius, Dolores R.; Moschou, Panagiotis N.; Dobránszki, Judit; Kaiserli, Eirini and Martinelli, Federico (2023). Deep inside the Epigenetic Memories of Stressed Plants. *Trends in Plant Science*, 28(2): 142–153.

Garcia, Valérie; Bruchet, Hugues; Camescasse, Delphine; Granier, Fabienne; Bouchez, David and Tissier, Alain (2003). AtATM Is Essential for Meiosis and the Somatic Response to DNA Damage in Plants[W]. *The Plant Cell*, 15(1): 119–132.

Gardiner, Jason; Ghoshal, Basudev; Wang, Ming and Jacobsen, Steven E (2022). CRISPR–Cas-Mediated Transcriptional Control and Epi-Mutagenesis. *Plant Physiology*, 188(4): 1811–1824.

Gel, Bernat and Serra, Eduard (2017). karyoploteR: An R/Bioconductor Package to Plot Customizable Genomes Displaying Arbitrary Data. *Bioinformatics*, 33(19): 3088–3090.

Geng, Shuaifeng; Kong, Xingchen; Song, Gaoyuan; Jia, Meiling; Guan, Jiantao; Wang, Fang; Qin, Zhengrui; Wu, Liang; Lan, Xiujin; Li, Aili and Mao, Long (2019). DNA Methylation Dynamics during the Interaction of Wheat Progenitor Aegilops Tauschii with the Obligate Biotrophic Fungus Blumeria Graminis f. Sp. Tritici. *New Phytologist*, 221(2): 1023–1035.

Geng, Xueqing; Cheng, Jiye; Gangadharan, Anju and Mackey, David (2012). The Coronatine Toxin of Pseudomonas Syringae Is a Multifunctional Suppressor of Arabidopsis Defense. *The Plant Cell*, 24(11): 4763–4774.

Gerken, Jacob; Vincent, Gear Thomas; Zapata, Demi; Barron, Ileana G. and Zapata, Isain (2024). Comprehensive Assessment of Pesticide Use Patterns and Increased Cancer Risk. *Frontiers in Cancer Control and Society*, 2. available at https://www.frontiersin.org/journals/cancer-control-and-society/articles/10.3389/fcacs.2024.1368086/full [12 October 2024].

Glazebrook, Jane (2005). Contrasting Mechanisms of Defense against Biotrophic and Necrotrophic Pathogens. *Annual Review of Phytopathology*, 43: 205–227.

Goll, Mary Grace and Bestor, Timothy H. (2005). EUKARYOTIC CYTOSINE METHYLTRANSFERASES. *Annual Review of Biochemistry*, 74(Volume 74, 2005): 481–514.

Gong, Zhizhong; Morales-Ruiz, Teresa; Ariza, Rafael R.; Roldán-Arjona, Teresa; David, Lisa and Zhu, Jian-Kang (2002). ROS1, a Repressor of Transcriptional Gene Silencing in Arabidopsis, Encodes a DNA Glycosylase/Lyase. *Cell*, 111(6): 803–814.

Gouil, Quentin and Baulcombe, David C. (2016). DNA Methylation Signatures of the Plant Chromomethyltransferases. *PLOS Genetics*, 12(12): e1006526.

Gouil, Quentin; Novák, Ondřej and Baulcombe, David C. (2016). SLTAB2 Is the Paramutated SULFUREA Locus in Tomato. *Journal of Experimental Botany*, 67(9): 2655–2664.

Griffin, Patrick T.; Niederhuth, Chad E. and Schmitz, Robert J. (2016). A Comparative Analysis of 5-Azacytidine- and Zebularine-Induced DNA Demethylation. *G3: Genes|Genomes|Genetics*, 6(9): 2773–2780.

Grin, Inga R.; Petrova, Daria V.; Endutkin, Anton V.; Ma, Chunquan; Yu, Bing; Li, Haiying and Zharkov, Dmitry O. (2023). Base Excision DNA Repair in Plants: Arabidopsis and Beyond. *International Journal of Molecular Sciences*, 24(19). available at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10573277/ [19 June 2024].

Guo, Wenbin; Tzioutziou, Nikoleta A.; Stephen, Gordon; Milne, Iain; Calixto, Cristiane PG; Waugh, Robbie; Brown, John W. S. and Zhang, Runxuan (2021). 3D RNA-Seq: A Powerful and Flexible Tool for Rapid and Accurate Differential Expression and Alternative Splicing Analysis of RNA-Seq Data for Biologists. *RNA Biology*, 18(11): 1574.

Halter, Thierry; Wang, Jingyu; Amesefe, Delase; Lastrucci, Emmanuelle; Charvin, Magali; Singla Rastogi, Meenu and Navarro, Lionel (2021). The Arabidopsis Active Demethylase ROS1 Cis-Regulates Defence Genes by Erasing DNA Methylation at

Promoter-Regulatory Regions (D. Zilberman, D. Weigel, D. Zilberman, J.-K. Zhu, and J. Ton, Eds.). *eLife*, 10: e62994.

Han, Qing; Tan, Wenrong; Zhao, Yuqing; Yang, Feng; Yao, Xiuhong; Lin, Honghui and Zhang, Dawei (2022). Salicylic Acid-activated BIN2 Phosphorylation of TGA3 Promotes Arabidopsis PR Gene Expression and Disease Resistance. *The EMBO Journal*, 41(19): e110682.

Han, Soon-Ki; Wu, Miin-Feng; Cui, Sujuan and Wagner, Doris (2015). Roles and Activities of Chromatin Remodeling ATPases in Plants. *The Plant Journal*, 83(1): 62–77.

Hannan Parker, Adam; Wilkinson, Samuel W. and Ton, Jurriaan (2022). Epigenetics: A Catalyst of Plant Immunity against Pathogens. *New Phytologist*, 233(1): 66–83.

Harris, C. Jake; Amtmann, Anna and Ton, Jurriaan (2023). Epigenetic Processes in Plant Stress Priming: Open Questions and New Approaches. *Current Opinion in Plant Biology*, 75: 102432.

Harris, C. Jake; Scheibe, Marion; Wongpalee, Somsakul Pop; Liu, Wanlu; Cornett, Evan M.; Vaughan, Robert M.; Li, Xueqin; Chen, Wei; Xue, Yan; Zhong, Zhenhui; Yen, Linda; Barshop, William D.; Rayatpisheh, Shima; Gallego-Bartolome, Javier; Groth, Martin; Wang, Zonghua; Wohlschlegel, James A.; Du, Jiamu; Rothbart, Scott B.; Butter, Falk and Jacobsen, Steven E. (2018). A DNA Methylation Reader Complex That Enhances Gene Transcription. *Science*, 362(6419): 1182–1186.

Harris, C Jake; Zhong, Zhenhui; Ichino, Lucia; Feng, Suhua and Jacobsen, Steven E (2024). H1 Restricts Euchromatin-Associated Methylation Pathways from Heterochromatic Encroachment (R. Amasino and D. Weigel, Eds.). *eLife*, 12: RP89353.

Harrison, Samuel J.; Mott, Ellie K.; Parsley, Kate; Aspinall, Sue; Gray, John C. and Cottage, Amanda (2006). A Rapid and Robust Method of Identifying Transformed Arabidopsis Thaliana Seedlings Following Floral Dip Transformation. *Plant Methods*, 2(1): 19.

Hartman, Sjon; Liu, Zeguang; van Veen, Hans; Vicente, Jorge; Reinen, Emilie; Martopawiro, Shanice; Zhang, Hongtao; van Dongen, Nienke; Bosman, Femke; Bassel, George W.; Visser, Eric J. W.; Bailey-Serres, Julia; Theodoulou, Frederica L.; Hebelstrup, Kim H.; Gibbs, Daniel J.; Holdsworth, Michael J.; Sasidharan, Rashmi and Voesenek, Laurentius A. C. J. (2019). Ethylene-Mediated Nitric Oxide Depletion Pre-Adapts Plants to Hypoxia Stress. *Nature Communications*, 10(1): 4020.

Hartman, Sjon; Sasidharan, Rashmi and Voesenek, Laurentius A. C. J. (2021). The Role of Ethylene in Metabolic Acclimations to Low Oxygen. *The New Phytologist*, 229(1): 64–70.

Hartmann, Michael; Zeier, Tatyana; Bernsdorff, Friederike; Reichel-Deland, Vanessa; Kim, Denis; Hohmann, Michele; Scholten, Nicola; Schuck, Stefan; Bräutigam, Andrea; Hölzel, Torsten; Ganter, Christian and Zeier, Jürgen (2018). Flavin Monooxygenase-Generated N-Hydroxypipecolic Acid Is a Critical Element of Plant Systemic Immunity. *Cell*, 173(2): 456-469.e16.

Heil, Martin and Ton, Jurriaan (2008). Long-Distance Signalling in Plant Defence. *Trends in Plant Science*, 13(6): 264–272.

Henderson, Ian R. and Jacobsen, Steven E. (2008). Tandem Repeats Upstream of the Arabidopsis Endogene SDC Recruit Non-CG DNA Methylation and Initiate siRNA Spreading. *Genes & Development*, 22(12): 1597–1606.

Henikoff, S.; Mckittrick, E. and Ahmad, K. (2004). Epigenetics, Histone H3 Variants, and the Inheritance of Chromatin States. *Cold Spring Harbor Symposia on Quantitative Biology*, 69: 235–244.

Herbst, Josephine; Li, Qian-Qian and De Veylder, Lieven (2024). Mechanistic Insights into DNA Damage Recognition and Checkpoint Control in Plants. *Nature Plants*, 10(4): 539–550.

Hickey, Lee T.; N. Hafeez, Amber; Robinson, Hannah; Jackson, Scott A.; Leal-Bertioli, Soraya C. M.; Tester, Mark; Gao, Caixia; Godwin, Ian D.; Hayes, Ben J. and Wulff, Brande B. H. (2019). Breeding Crops to Feed 10 Billion. *Nature Biotechnology*, 37(7): 744–754.

Hofmeister, Brigitte T.; Denkena, Johanna; Colomé-Tatché, Maria; Shahryary, Yadollah; Hazarika, Rashmi; Grimwood, Jane; Mamidi, Sujan; Jenkins, Jerry; Grabowski, Paul P.; Sreedasyam, Avinash; Shu, Shengqiang; Barry, Kerrie; Lail, Kathleen; Adam, Catherine; Lipzen, Anna; Sorek, Rotem; Kudrna, Dave; Talag, Jayson; Wing, Rod; Hall, David W.; Jacobsen, Daniel; Tuskan, Gerald A.; Schmutz, Jeremy; Johannes, Frank and Schmitz, Robert J. (2020). A Genome Assembly and the Somatic Genetic and Epigenetic Mutation Rate in a Wild Long-Lived Perennial Populus Trichocarpa. *Genome Biology*, 21: 259.

Hofmeister, Brigitte T.; Lee, Kevin; Rohr, Nicholas A.; Hall, David W. and Schmitz, Robert J. (2017). Stable Inheritance of DNA Methylation Allows Creation of Epigenotype Maps and the Study of Epiallele Inheritance Patterns in the Absence of Genetic Variation. *Genome Biology*, 18(1): 155.

Hollick, Jay B. (2017). Paramutation and Related Phenomena in Diverse Species. *Nature Reviews Genetics*, 18(1): 5–23.

Hong, Sung Myun; Bahn, Sung Chul; Lyu, Aram; Jung, Hye Seung and Ahn, Ji Hoon (2010). Identification and Testing of Superior Reference Genes for a Starting Pool of Transcript Normalization in Arabidopsis. *Plant and Cell Physiology*, 51(10): 1694–1706.

Hsu, Fu-Chiun; Chou, Mei-Yi; Peng, Hsiao-Ping; Chou, Shu-Jen and Shih, Ming-Che (2011). Insights into Hypoxic Systemic Responses Based on Analyses of Transcriptional Regulation in Arabidopsis. *PLOS ONE*, 6(12): e28888.

Hu, Yezhou; Ding, Yanxia; Cai, Boying; Qin, Xiaohui; Wu, Jingni; Yuan, Minhang; Wan, Shiwei; Zhao, Yang and Xin, Xiu-Fang (2022). Bacterial Effectors Manipulate Plant Abscisic Acid Signaling for Creation of an Aqueous Apoplast. *Cell Host & Microbe*, 30(4): 518-529.e6.

Huang, Ying; Sicar, Sanchari; Ramirez-Prado, Juan S.; Manza-Mianza, Deborah; Antunez-Sanchez, Javier; Brik-Chaouche, Rim; Rodriguez-Granados, Natalia Y.; An, Jing; Bergounioux, Catherine; Mahfouz, Magdy M.; Hirt, Heribert; Crespi, Martin; Concia, Lorenzo; Barneche, Fredy; Amiard, Simon; Probst, Aline V.; Gutierrez-Marcos, Jose; Ariel, Federico; Raynaud, Cécile; Latrasse, David and Benhamed, Moussa (2021). Polycomb-Dependent Differential Chromatin Compartmentalization Determines Gene Coregulation in Arabidopsis. *Genome Research*, 31(7): 1230–1244.

van Hulten, Marieke; Pelser, Maaike; van Loon, L. C.; Pieterse, Corné M. J. and Ton, Jurriaan (2006). Costs and Benefits of Priming for Defense in Arabidopsis. *Proceedings of the National Academy of Sciences*, 103(14): 5602–5607.

Iwasaki, Mayumi and Paszkowski, Jerzy (2014). Identification of Genes Preventing Transgenerational Transmission of Stress-Induced Epigenetic States. *Proceedings of the National Academy of Sciences of the United States of America*, 111(23): 8547–8552.

Iyer, Lakshminarayan M.; Tahiliani, Mamta; Rao, Anjana and Aravind, L. (2009). Prediction of Novel Families of Enzymes Involved in Oxidative and Other Complex Modifications of Bases in Nucleic Acids. *Cell cycle (Georgetown, Tex.)*, 8(11): 1698.

Jacob, Yannick; Feng, Suhua; LeBlanc, Chantal A.; Bernatavichute, Yana V.; Stroud, Hume; Cokus, Shawn; Johnson, Lianna M.; Pellegrini, Matteo; Jacobsen, Steven E. and Michaels, Scott D. (2009). ATXR5 and ATXR6 Are H3K27 Monomethyltransferases Required for Chromatin Structure and Gene Silencing. *Nature Structural & Molecular Biology*, 16(7): 763–768.

Jakoby, Marc; Weisshaar, Bernd; Dröge-Laser, Wolfgang; Vicente-Carbajosa, Jesus; Tiedemann, Jens; Kroj, Thomas; Parcy, François; and bZIP Research Group (2002). bZIP Transcription Factors in Arabidopsis. *Trends in Plant Science*, 7(3): 106–111.

Jamge, Bhagyshree; Lorković, Zdravko J; Axelsson, Elin; Osakabe, Akihisa; Shukla, Vikas; Yelagandula, Ramesh; Akimcheva, Svetlana; Kuehn, Annika Luisa and Berger, Frédéric (2023). Histone Variants Shape Chromatin States in Arabidopsis (P. Casati and J. Kleine-Vehn, Eds.). *eLife*, 12: RP87714.

Jaskiewicz, Michal; Conrath, Uwe and Peterhänsel, Christoph (2011). Chromatin Modification Acts as a Memory for Systemic Acquired Resistance in the Plant Stress Response. *EMBO reports*, 12(1): 50–55.

Jean Finnegan, E. and Dennis, Elizabeth S. (1993). Isolation and Identification by Sequence Homology of a Putative Cytosine Methyltransferase from Arabidopsis Thaliana. *Nucleic Acids Research*, 21(10): 2383–2388.

Jeddeloh, Jeffrey A.; Bender, Judith and Richards, Eric J. (1998). The DNA Methylation Locus DDM1 Is Required for Maintenance of Gene Silencing in Arabidopsis. *Genes & Development*, 12(11): 1714–1725.

Jeong, Hwi-Won; Ryu, Tae Ho; Lee, Hyo-Jeong; Kim, Kook-Hyung and Jeong, Rae-Dong (2023). DNA Damage Triggers the Activation of Immune Response to Viral Pathogens via Salicylic Acid in Plants. *The Plant Pathology Journal*, 39(5): 449–465.

Ji, Debin; Lin, Krystal; Song, Jikui and Wang, Yinsheng (2014). Effects of Tet-Induced Oxidation Products of 5-Methylcytosine on Dnmt1- and DNMT3a-Mediated Cytosine Methylation. *Molecular bioSystems*, 10(7): 1749.

Ji, Lexiang; Jordan, William T.; Shi, Xiuling; Hu, Lulu; He, Chuan and Schmitz, Robert J. (2018). TET-Mediated Epimutagenesis of the Arabidopsis Thaliana Methylome. *Nature Communications*, 9. available at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5832761/ [19 June 2024].

Jiang, Hailong; Liu, Xiaoya; Xiao, Peixiang; Wang, Yan; Xie, Qihui; Wu, Xiaoxia and Ding, Haidong (2023). Functional Insights of Plant Bcl-2–Associated Ahanogene (BAG)

Proteins: Multi-Taskers in Diverse Cellular Signal Transduction Pathways. *Frontiers in Plant Science*, 14: 1136873.

Jiang, Lichun; Schlesinger, Felix; Davis, Carrie A.; Zhang, Yu; Li, Renhua; Salit, Marc; Gingeras, Thomas R. and Oliver, Brian (2011). Synthetic Spike-in Standards for RNA-Seq Experiments. *Genome Research*, 21(9): 1543–1551.

Jin, Hongshi; Choi, Sun-Mee; Kang, Min-Jeong; Yun, Se-Hun; Kwon, Dong-Jin; Noh, Yoo-Sun and Noh, Bosl (2018). Salicylic Acid-Induced Transcriptional Reprogramming by the HAC–NPR1–TGA Histone Acetyltransferase Complex in Arabidopsis. *Nucleic Acids Research*, 46(22): 11712–11725.

Johannes, Frank; Porcher, Emmanuelle; Teixeira, Felipe K.; Saliba-Colombani, Vera; Simon, Matthieu; Agier, Nicolas; Bulski, Agnès; Albuisson, Juliette; Heredia, Fabiana; Audigier, Pascal; Bouchez, David; Dillmann, Christine; Guerche, Philippe; Hospital, Frédéric and Colot, Vincent (2009). Assessing the Impact of Transgenerational Epigenetic Variation on Complex Traits. *PLOS Genetics*, 5(6): e1000530.

Johannes, Frank and Schmitz, Robert J. (2019). Spontaneous Epimutations in Plants. *New Phytologist*, 221(3): 1253–1259.

Johnson, Lianna; Cao, Xiaofeng and Jacobsen, Steven (2002). Interplay between Two Epigenetic Marks. DNA Methylation and Histone H3 Lysine 9 Methylation. *Current biology: CB*, 12(16): 1360–1367.

Johnson, Lianna M.; Du, Jiamu; Hale, Christopher J.; Bischof, Sylvain; Feng, Suhua; Chodavarapu, Ramakrishna K.; Zhong, Xuehua; Marson, Giuseppe; Pellegrini, Matteo; Segal, David J.; Patel, Dinshaw J. and Jacobsen, Steven E. (2014). SRA- and SET-Domain-Containing Proteins Link RNA Polymerase V Occupancy to DNA Methylation. *Nature*, 507(7490): 124–128.

Johnson, Nathan R.; Yeoh, Jonathan M.; Coruh, Ceyda and Axtell, Michael J. (2016). Improved Placement of Multi-Mapping Small RNAs. *G3: Genes|Genomes|Genetics*, 6(7): 2103–2111.

Jones, Jonathan D. G. and Dangl, Jeffery L. (2006). The Plant Immune System. *Nature*, 444(7117): 323–329.

Jones, Jonathan D. G.; Staskawicz, Brian J. and Dangl, Jeffery L. (2024). The Plant Immune System: From Discovery to Deployment. *Cell*, 187(9): 2095–2116.

Jones, Jonathan D. G.; Vance, Russell E. and Dangl, Jeffery L. (2016). Intracellular Innate Immune Surveillance Devices in Plants and Animals. *Science*, 354(6316): aaf6395.

Jullien, Pauline E.; Susaki, Daichi; Yelagandula, Ramesh; Higashiyama, Tetsuya and Berger, Frédéric (2012). DNA Methylation Dynamics during Sexual Reproduction in Arabidopsis Thaliana. *Current Biology*, 22(19): 1825–1830.

Kankel, Mark W.; Ramsey, Douglas E.; Stokes, Trevor L.; Flowers, Susan K.; Haag, Jeremy R.; Jeddeloh, Jeffrey A.; Riddle, Nicole C.; Verbsky, Michelle L. and Richards, Eric J. (2003). Arabidopsis MET1 Cytosine Methyltransferase Mutants. *Genetics*, 163(3): 1109–1122.

Kapranov, Philipp; Cheng, Jill; Dike, Sujit; Nix, David A.; Duttagupta, Radharani; Willingham, Aarron T.; Stadler, Peter F.; Hertel, Jana; Hackermüller, Jörg; Hofacker, Ivo L.; Bell, Ian;

Cheung, Evelyn; Drenkow, Jorg; Dumais, Erica; Patel, Sandeep; Helt, Gregg; Ganesh, Madhavan; Ghosh, Srinka; Piccolboni, Antonio; Sementchenko, Victor; Tammana, Hari and Gingeras, Thomas R. (2007). RNA Maps Reveal New RNA Classes and a Possible Function for Pervasive Transcription. *Science*, 316(5830): 1484–1488.

Kathiria, Palak; Sidler, Corinne; Golubov, Andrey; Kalischuk, Melanie; Kawchuk, Lawrence M. and Kovalchuk, Igor (2010). Tobacco Mosaic Virus Infection Results in an Increase in Recombination Frequency and Resistance to Viral, Bacterial, and Fungal Pathogens in the Progeny of Infected Tobacco Plants. *Plant Physiology*, 153(4): 1859–1870.

Kawakatsu, Taiji; Nery, Joseph R.; Castanon, Rosa and Ecker, Joseph R. (2017). Dynamic DNA Methylation Reconfiguration during Seed Development and Germination. *Genome Biology*, 18(1): 171.

Kesarwani, Meenu; Yoo, Jungmin and Dong, Xinnian (2007). Genetic Interactions of TGA Transcription Factors in the Regulation of Pathogenesis-Related Genes and Disease Resistance in Arabidopsis. *Plant Physiology*, 144(1): 336.

Kim, Do-Young; Scalf, Mark; Smith, Lloyd M. and Vierstra, Richard D. (2013). Advanced Proteomic Analyses Yield a Deep Catalog of Ubiquitylation Targets in Arabidopsis. *The Plant Cell*, 25(5): 1523–1540.

Kim, June-Sik; Lim, Joo Young; Shin, Hosub; Kim, Beom-Gi; Yoo, Sang-Dong; Kim, Woo Taek and Huh, Jin Hoe (2019). ROS1-Dependent DNA Demethylation Is Required for ABA-Inducible NIC3 Expression. *Plant Physiology*, 179(4): 1810–1821.

Kimathi, Emily; Tonnang, Henri E. Z.; Subramanian, Sevgan; Cressman, Keith; Abdel-Rahman, Elfatih M.; Tesfayohannes, Mehari; Niassy, Saliou; Torto, Baldwyn; Dubois, Thomas; Tanga, Chrysantus M.; Kassie, Menale; Ekesi, Sunday; Mwangi, David and Kelemu, Segenet (2020). Prediction of Breeding Regions for the Desert Locust Schistocerca Gregaria in East Africa. *Scientific Reports*, 10(1): 11937.

Klepikova, Anna V.; Kasianov, Artem S.; Gerasimov, Evgeny S.; Logacheva, Maria D. and Penin, Aleksey A. (2016). A High Resolution Map of the Arabidopsis Thaliana Developmental Transcriptome Based on RNA-Seq Profiling. *The Plant Journal*, 88(6): 1058–1070.

Kneeshaw, Sophie; Gelineau, Silvère; Tada, Yasuomi; Loake, Gary J. and Spoel, Steven H. (2014). Selective Protein Denitrosylation Activity of Thioredoxin-*h*5 Modulates Plant Immunity. *Molecular Cell*, 56(1): 153–162.

Ko, Jae-Heung; Yang, Seung H. and Han, Kyung-Hwan (2006). Upregulation of an Arabidopsis RING-H2 Gene, XERICO, Confers Drought Tolerance through Increased Abscisic Acid Biosynthesis. *The Plant Journal: For Cell and Molecular Biology*, 47(3): 343–355.

Koch, E and Slusarenko, A (1990). Arabidopsis Is Susceptible to Infection by a Downy Mildew Fungus. *The Plant Cell*, 2(5): 437–445.

Kolmogorov, Mikhail; Yuan, Jeffrey; Lin, Yu and Pevzner, Pavel A. (2019). Assembly of Long, Error-Prone Reads Using Repeat Graphs. *Nature Biotechnology*, 37(5): 540–546.

Kong, Xiangfeng; Hong, Yechun; Hsu, Yi-Feng; Huang, Huan; Liu, Xue; Song, Zhe and Zhu, Jian-Kang (2020). SIZ1-Mediated SUMOylation of ROS1 Enhances Its Stability and

Positively Regulates Active DNA Demethylation in Arabidopsis. *Molecular Plant*, 13(12): 1816–1824.

Kong, Yimeng; Cao, Lei; Deikus, Gintaras; Fan, Yu; Mead, Edward A.; Lai, Weiyi; Zhang, Yizhou; Yong, Raymund; Sebra, Robert; Wang, Hailin; Zhang, Xue-Song and Fang, Gang (2022). Critical Assessment of DNA Adenine Methylation in Eukaryotes Using Quantitative Deconvolution. *Science*, 375(6580): 515–522.

Kornienko, Aleksandra E; Nizhynska, Viktoria; Molla Morales, Almudena; Pisupati, Rahul and Nordborg, Magnus (2024). Population-Level Annotation of lncRNAs in Arabidopsis Reveals Extensive Expression Variation Associated with Transposable Element–like Silencing. *The Plant Cell*, 36(1): 85–111.

Kozomara, Ana; Birgaoanu, Maria and Griffiths-Jones, Sam (2019). miRBase: From microRNA Sequences to Function. *Nucleic Acids Research*, 47(D1): D155–D162.

Krueger, Felix and Andrews, Simon R. (2011). Bismark: A Flexible Aligner and Methylation Caller for Bisulfite-Seq Applications. *Bioinformatics*, 27(11): 1571–1572.

Kumar, Shivesh; Zavaliev, Raul; Wu, Qinglin; Zhou, Ye; Cheng, Jie; Dillard, Lucas; Powers, Jordan; Withers, John; Zhao, Jinshi; Guan, Ziqiang; Borgnia, Mario J.; Bartesaghi, Alberto; Dong, Xinnian and Zhou, Pei (2022). Structural Basis of NPR1 in Activating Plant Immunity. *Nature*, 605(7910): 561–566.

Kus, Julianne V.; Zaton, Kasia; Sarkar, Raani and Cameron, Robin K. (2002). Age-Related Resistance in Arabidopsis Is a Developmentally Regulated Defense Response to Pseudomonas Syringae. *The Plant Cell*, 14(2): 479–490.

Lai, Yan; Cuzick, Alayne; Lu, Xueqing M.; Wang, Jianqiang; Katiyar, Neerja; Tsuchiya, Tokuji; Roch, Karine Le; McDowell, John M.; Holub, Eric and Eulgem, Thomas (2019). The Arabidopsis RRM Domain Protein EDM3 Mediates Race-Specific Disease Resistance by Controlling H3K9me2-Dependent Alternative Polyadenylation of RPP7 Immune Receptor Transcripts. *The Plant Journal*, 97(4): 646–660.

Lai, Yan; Lu, Xueqing Maggie; Daron, Josquin; Pan, Songqin; Wang, Jianqiang; Wang, Wei; Tsuchiya, Tokuji; Holub, Eric; McDowell, John M.; Slotkin, R. Keith; Roch, Karine G. Le and Eulgem, Thomas (2020). The Arabidopsis PHD-Finger Protein EDM2 Has Multiple Roles in Balancing NLR Immune Receptor Gene Expression. *PLOS Genetics*, 16(9): e1008993.

Laloi, Christophe; Mestres-Ortega, Dominique; Marco, Yves; Meyer, Yves and Reichheld, Jean-Philippe (2004). The Arabidopsis Cytosolic Thioredoxin H5 Gene Induction by Oxidative Stress and Its W-Box-Mediated Response to Pathogen Elicitor. *Plant Physiology*, 134(3): 1006–1016.

Lang, Lucas; Pettkó-Szandtner, Aladár; Elbaşı, Hasibe Tunçay; Takatsuka, Hirotomo; Nomoto, Yuji; Zaki, Ahmad; Dorokhov, Stefan; Jaeger, Geert De; Eeckhout, Dominique; Ito, Masaki; Magyar, Zoltán; Bögre, László; Heese, Maren and Schnittger, Arp (2021). The DREAM Complex Represses Growth in Response to DNA Damage in Arabidopsis. *Life Science Alliance*, 4(12). available at https://www.life-science-alliance.org/content/4/12/e202101141 [26 February 2024].

Lang, Zhaobo; Lei, Mingguang; Wang, Xingang; Tang, Kai; Miki, Daisuke; Zhang, Huiming; Mangrauthia, Satendra K.; Liu, Wenshan; Nie, Wenfeng; Ma, Guojie; Yan, Jun; Duan, Cheng-Guo; Hsu, Chuan-Chih; Wang, Chunlei; Tao, W. Andy; Gong, Zhizhong and

Zhu, Jian-Kang (2015). The Methyl-CpG-Binding Protein MBD7 Facilitates Active DNA Demethylation to Limit DNA Hyper-Methylation and Transcriptional Gene Silencing. *Molecular Cell*, 57(6): 971–983.

Langmead, Ben and Salzberg, Steven L. (2012). Fast Gapped-Read Alignment with Bowtie 2. *Nature Methods*, 9(4): 357–359.

Langmead, Ben; Trapnell, Cole; Pop, Mihai and Salzberg, Steven L. (2009). Ultrafast and Memory-Efficient Alignment of Short DNA Sequences to the Human Genome. *Genome Biology*, 10(3): R25.

Laosuntisuk, Kanjana; Vennapusa, Amaranatha; Somayanda, Impa M.; Leman, Adam R.; Jagadish, SV Krishna and Doherty, Colleen J. A Normalization Method That Controls for Total RNA Abundance Affects the Identification of Differentially Expressed Genes, Revealing Bias toward Morning-Expressed Responses. *The Plant Journal*, n/a(n/a). available at https://onlinelibrary.wiley.com/doi/abs/10.1111/tpj.16654 [4 March 2024].

Latzel, Vít; Zhang, Yuanye; Karlsson Moritz, Kim; Fischer, Markus and Bossdorf, Oliver (2012). Epigenetic Variation in Plant Responses to Defence Hormones. *Annals of Botany*, 110(7): 1423–1428.

Law, Julie A.; Du, Jiamu; Hale, Christopher J.; Feng, Suhua; Krajewski, Krzysztof; Palanca, Ana Marie S.; Strahl, Brian D.; Patel, Dinshaw J. and Jacobsen, Steven E. (2013). Polymerase-IV Occupancy at RNA-Directed DNA Methylation Sites Requires SHH1. *Nature*, 498(7454): 385–389.

Lawrence, Michael; Huber, Wolfgang; Pagès, Hervé; Aboyoun, Patrick; Carlson, Marc; Gentleman, Robert; Morgan, Martin T. and Carey, Vincent J. (2013). Software for Computing and Annotating Genomic Ranges. *PLOS Computational Biology*, 9(8): e1003118.

Lee, Dong Wook; Kim, Su Jin; Oh, Young Jun; Choi, Bongsoo; Lee, Juhun and Hwang, Inhwan (2016). *Arabidopsis* BAG1 Functions as a Cofactor in Hsc70-Mediated Proteasomal Degradation of Unimported Plastid Proteins. *Molecular Plant*, 9(10): 1428–1431.

Lee, Jiyoon; Jang, Hosung; Shin, Hosub; Choi, Woo Lee; Mok, Young Geun and Huh, Jin Hoe (2014). AP Endonucleases Process 5-Methylcytosine Excision Intermediates during Active DNA Demethylation in Arabidopsis. *Nucleic Acids Research*, 42(18): 11408–11418.

Lee, Seung Cho; Mustroph, Angelika; Sasidharan, Rashmi; Vashisht, Divya; Pedersen, Ole; Oosumi, Teruko; Voesenek, Laurentius A.C.J. and Bailey-Serres, Julia (2011). Molecular Characterization of the Submergence Response of the Arabidopsis Thaliana Ecotype Columbia. *New Phytologist*, 190(2): 457–471.

Lee, Seunga; Park, Young Seo; Rhee, Ji Hoon; Chu, Hyojeong; Frost, Jennifer M. and Choi, Yeonhee (2024). Insights into Plant Regeneration: Cellular Pathways and DNA Methylation Dynamics. *Plant Cell Reports*, 43(5): 120.

Lee, Seungchul; Choi, Jaemyung; Park, Jihwan; Hong, Chang Pyo; Choi, Daeseok; Han, Soeun; Choi, Kyuha; Roh, Tae-Young; Hwang, Daehee and Hwang, Ildoo (2023). DDM1-Mediated Gene Body DNA Methylation Is Associated with Inducible Activation of Defense-Related Genes in Arabidopsis. *Genome Biology*, 24(1): 106.

Lei, Mingguang; La, Honggui; Lu, Kun; Wang, Pengcheng; Miki, Daisuke; Ren, Zhizhong; Duan, Cheng-Guo; Wang, Xingang; Tang, Kai; Zeng, Liang; Yang, Lan; Zhang, Heng; Nie, Wenfeng; Liu, Pan; Zhou, Jianping; Liu, Renyi; Zhong, Yingli; Liu, Dong and Zhu, Jian-Kang (2014). Arabidopsis EDM2 Promotes IBM1 Distal Polyadenylation and Regulates Genome DNA Methylation Patterns. *Proceedings of the National Academy of Sciences*, 111(1): 527–532.

Lei, Mingguang; Zhang, Huiming; Julian, Russell; Tang, Kai; Xie, Shaojun and Zhu, Jian-Kang (2015). Regulatory Link between DNA Methylation and Active Demethylation in Arabidopsis. *Proceedings of the National Academy of Sciences*, 112(11): 3553–3557.

Lescot, Thierry (2020). Banana Genetic Diversity. *FruiTrop*, FruiTrop n°269: 98–102.

Li, Heng (2018). Minimap2: Pairwise Alignment for Nucleotide Sequences. *Bioinformatics*, 34(18): 3094–3100.

Li, Min; Chen, Huan; Chen, Jian; Chang, Ming; Palmer, Ian A.; Gassmann, Walter; Liu, Fengquan and Fu, Zheng Qing (2018). TCP Transcription Factors Interact With NPR1 and Contribute Redundantly to Systemic Acquired Resistance. *Frontiers in Plant Science*, 9: 1153.

Li, Qi; Wang, Chenggang and Mou, Zhonglin (2020). Perception of Damaged Self in Plants. *Plant Physiology*, 182(4): 1545–1565.

Li, Qi; Wang, Xiaokang; Sun, Han; Zeng, Jun; Cao, Zhendong; Li, Yan and Qian, Weiqiang (2015). Regulation of Active DNA Demethylation by a Methyl-CpG-Binding Domain Protein in Arabidopsis Thaliana. *PLOS Genetics*, 11(5): e1005210.

Li, Zheng; Wang, Ming; Zhong, Zhenhui; Gallego-Bartolomé, Javier; Feng, Suhua; Jami-Alahmadi, Yasaman; Wang, Xinyi; Wohlschlegel, James; Bischof, Sylvain; Long, Jeff A. and Jacobsen, Steven E. (2023). The MOM1 Complex Recruits the RdDM Machinery via MORC6 to Establish de Novo DNA Methylation. *Nature Communications*, 14(1): 4135.

Liang, Zhe; Shen, Lisha; Cui, Xuean; Bao, Shengjie; Geng, Yuke; Yu, Guoliang; Liang, Fan; Xie, Shang; Lu, Tiegang; Gu, Xiaofeng and Yu, Hao (2018). DNA N6-Adenine Methylation in Arabidopsis Thaliana. *Developmental Cell*, 45(3): 406-416.e3.

Lindroth, Anders M.; Cao, Xiaofeng; Jackson, James P.; Zilberman, Daniel; McCallum, Claire M.; Henikoff, Steven and Jacobsen, Steven E. (2001). Requirement of CHROMOMETHYLASE3 for Maintenance of CpXpG Methylation. *Science*, 292(5524): 2077–2080.

Lindsey, Benson E.; Rivero, Luz; Calhoun, Chistopher S.; Grotewold, Erich and Brkljacic, Jelena (2017). Standardized Method for High-Throughput Sterilization of Arabidopsis Seeds. *Journal of Visualized Experiments : JoVE*, (128): 56587.

Liu, Chang; Xin, Ying; Xu, Le; Cai, Zhaokui; Xue, Yuanchao; Liu, Yong; Xie, Daoxin; Liu, Yule and Qi, Yijun (2018). Arabidopsis ARGONAUTE 1 Binds Chromatin to Promote Gene Transcription in Response to Hormones and Stresses. *Developmental Cell*, 44(3): 348-361.e7.

Liu, Pan; Nie, Wen-Feng; Xiong, Xiansong; Wang, Yuhua; Jiang, Yuwei; Huang, Pei; Lin, Xueqiang; Qin, Guochen; Huang, Huan; Niu, Qingfeng; Du, Jiamu; Lang, Zhaobo; Lozano-Duran, Rosa and Zhu, Jian-Kang (2021). A Novel Protein Complex That

Regulates Active DNA Demethylation in Arabidopsis. *Journal of Integrative Plant Biology*, 63(4): 772–786.

Liu, Xinyue; Zhang, Dongxiao and He, Xiaogang (2024). Unveiling the Role of Climate in Spatially Synchronized Locust Outbreak Risks. *Science Advances*, 10(7): eadj1164.

Liu, Zhang-Wei; Liu, Jie; Liu, Fengquan and Zhong, Xuehua (2023). Depositing Centromere Repeats Induces Heritable Intragenic Heterochromatin Establishment and Spreading in Arabidopsis. *Nucleic Acids Research*: gkad306.

Liu, Zhang-Wei; Shao, Chang-Rong; Zhang, Cui-Jun; Zhou, Jin-Xing; Zhang, Su-Wei; Li, Lin; Chen, She; Huang, Huan-Wei; Cai, Tao and He, Xin-Jian (2014). The SET Domain Proteins SUVH2 and SUVH9 Are Required for Pol V Occupancy at RNA-Directed DNA Methylation Loci. *PLoS genetics*, 10(1): e1003948.

Long, Jincheng; Walker, James; She, Wenjing; Aldridge, Billy; Gao, Hongbo; Deans, Samuel; Vickers, Martin and Feng, Xiaoqi (2021). Nurse Cell-–Derived Small RNAs Define Paternal Epigenetic Inheritance in Arabidopsis. *Science*, 373(6550): eabh0556.

López, Ana; Ramírez, Vicente; García-Andrade, Javier; Flors, Victor and Vera, Pablo (2011). The RNA Silencing Enzyme RNA Polymerase V Is Required for Plant Immunity. *PLOS Genetics*, 7(12): e1002434.

López Sánchez, Ana; Pascual-Pardo, David; Furci, Leonardo; Roberts, Michael R. and Ton, Jurriaan (2021). Costs and Benefits of Transgenerational Induced Resistance in Arabidopsis. *Frontiers in Plant Science*, 12. available at https://www.frontiersin.org/articles/10.3389/fpls.2021.644999/full [8 March 2021].

López Sánchez, Ana; Stassen, Joost H.M.; Furci, Leonardo; Smith, Lisa M. and Ton, Jurriaan (2016). The Role of DNA (de)Methylation in Immune Responsiveness of Arabidopsis. *The Plant Journal: For Cell and Molecular Biology*, 88(3): 361–374.

Loreti, Elena; Betti, Federico; Ladera-Carmona, Maria Jose; Fontana, Fabrizia; Novi, Giacomo; Valeri, Maria Cristina and Perata, Pierdomenico (2020). ARGONAUTE1 and ARGONAUTE4 Regulate Gene Expression and Hypoxia Tolerance. *Plant Physiology*, 182(1): 287–300.

Love, Michael I.; Huber, Wolfgang and Anders, Simon (2014). Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2. *Genome Biology*, 15(12): 550.

Lovén, Jakob; Orlando, David A.; Sigova, Alla A.; Lin, Charles Y.; Rahl, Peter B.; Burge, Christopher B.; Levens, David L.; Lee, Tong Ihn and Young, Richard A. (2012). Revisiting Global Gene Expression Analysis. *Cell*, 151(3): 476–482.

Lucht, Jan M.; Mauch-Mani, Brigitte; Steiner, Henry-York; Metraux, Jean-Pierre; Ryals, John and Hohn, Barbara (2002). Pathogen Stress Increases Somatic Recombination Frequency in Arabidopsis. *Nature Genetics*, 30(3): 311–314.

Luger, Karolin; Mäder, Armin W.; Richmond, Robin K.; Sargent, David F. and Richmond, Timothy J. (1997). Crystal Structure of the Nucleosome Core Particle at 2.8 Å Resolution. *Nature*, 389(6648): 251–260.

Luna, Estrella; Bruce, Toby J. A.; Roberts, Michael R.; Flors, Victor and Ton, Jurriaan (2012). Next-Generation Systemic Acquired Resistance. *Plant Physiology*, 158(2): 844–853.

Luna, Estrella; López, Ana; Kooiman, Jaap and Ton, Jurriaan (2014). Role of NPR1 and KYP in Long-Lasting Induced Resistance by β-Aminobutyric Acid. *Frontiers in Plant Science*, 5. available at https://www.frontiersin.org/articles/10.3389/fpls.2014.00184/full [6 March 2021].

Lv, Suhui; Yang, Yu; Yu, Gang; Peng, Li; Zheng, Shuai; Singh, Sunil Kumar; Vílchez, Juan Ignacio; Kaushal, Richa; Zi, Hailing; Yi, Dian; Wang, Yuhua; Luo, Shaofan; Wu, Xiaoxuan; Zuo, Ziwei; Huang, Weichang; Liu, Renyi; Du, Jiamu; Macho, Alberto P.; Tang, Kai and Zhang, Huiming (2022). Dysfunction of Histone Demethylase IBM1 in Arabidopsis Causes Autoimmunity and Reshapes the Root Microbiome. *The ISME Journal*: 1–12.

Macho, Alberto P. and Zipfel, Cyril (2014). Plant PRRs and the Activation of Innate Immune Signaling. *Molecular Cell*, 54(2): 263–272.

Mair, Andrea; Xu, Shou-Ling; Branon, Tess C; Ting, Alice Y and Bergmann, Dominique C (2019). Proximity Labeling of Protein Complexes and Cell-Type-Specific Organellar Proteomes in Arabidopsis Enabled by TurboID (F. L. Menke and C. S. Hardtke, Eds.). *eLife*, 8: e47864.

Maor, Rudy; Jones, Alex; Nühse, Thomas S.; Studholme, David J.; Peck, Scott C. and Shirasu, Ken (2007). Multidimensional Protein Identification Technology (MudPIT) Analysis of Ubiquitinated Proteins in Plants *. *Molecular & Cellular Proteomics*, 6(4): 601–610.

Martinho, Cláudia; Wang, Zhengming; Ghigi, Andrea; Buddle, Sarah; Barbour, Felix; Yarur, Antonia; Gouil, Quentin; Müller, Sebastian; Stam, Maike; Liu, Chang and Baulcombe, David C. (2022). CHROMOMETHYLTRANSFERASE3/KRYPTONITE Maintains the Sulfurea Paramutation in Solanum Lycopersicum. *Proceedings of the National Academy of Sciences*, 119(13): e2112240119.

Martins, Laura M. and Law, Julie A. (2023). Moving Targets: Mechanisms Regulating siRNA Production and DNA Methylation during Plant Development. *Current Opinion in Plant Biology*, 75: 102435.

Martins, Samuel J.; Pasche, Josephine; Silva, Hiago Antonio O.; Selten, Gijs; Savastano, Noah; Abreu, Lucas Magalhães; Bais, Harsh P.; Garrett, Karen A.; Kraisitudomsook, Nattapol; Pieterse, Corné M. J. and Cernava, Tomislav (2023). The Use of Synthetic Microbial Communities to Improve Plant Health. *Phytopathology®*, 113(8): 1369–1379.

Mathieu, Olivier; Reinders, Jon; Čaikovski, Marian; Smathajitt, Chotika and Paszkowski, Jerzy (2007). Transgenerational Stability of the *Arabidopsis* Epigenome Is Coordinated by CG Methylation. *Cell*, 130(5): 851–862.

Matzke, Marjori A. and Mosher, Rebecca A. (2014). RNA-Directed DNA Methylation: An Epigenetic Pathway of Increasing Complexity. *Nature Reviews Genetics*, 15(6): 394–408.

Mauch-Mani, Brigitte; Baccelli, Ivan; Luna, Estrella and Flors, Victor (2017). Defense Priming: An Adaptive Part of Induced Resistance. *Annual Review of Plant Biology*, 68(1): 485–512.

Mitchell, Ruth J.; Pakeman, Robin J.; Broome, Alice; Beaton, Joan K.; Bellamy, Paul E.; Brooker, Rob W.; Ellis, Chris J.; Hester, Alison J.; Hodgetts, Nick G.; Iason, Glenn R.; Littlewood, Nick A.; Pozsgai, Gabor; Ramsay, Scot; Riach, David; Stockan, Jenni A.; Taylor, Andy F. S. and Woodward, Steve (2016). How to Replicate the Functions and

Biodiversity of a Threatened Tree Species? The Case of Fraxinus Excelsior in Britain. *Ecosystems*, 19(4): 573–586.

Miura, Asuka; Nakamura, Miyuki; Inagaki, Soichi; Kobayashi, Akie; Saze, Hidetoshi and Kakutani, Tetsuji (2009). An Arabidopsis jmjC Domain Protein Protects Transcribed Genes from DNA Methylation at CHG Sites. *The EMBO Journal*, 28(8): 1078–1086.

Mohamed, Deka; Vonapartis, Eliana; Corcega, Dennedy Yrvin and Gazzarrini, Sonia (2023). ABA Guides Stomatal Proliferation and Patterning through the EPF-SPCH Signaling Pathway in Arabidopsis Thaliana. *Development*, 150(23): dev201258.

Molina, Antonio; Hunt, Michelle D. and Ryals, John A. (1998). Impaired Fungicide Activity in Plants Blocked in Disease Resistance Signal Transduction. *The Plant Cell*, 10(11): 1903–1914.

Mundt, Christopher C. (2014). Durable Resistance: A Key to Sustainable Management of Pathogens and Pests. *Infection, Genetics and Evolution*, 27: 446–455.

Munsamy, Ashlin; Rutherford, R. Stuart; Snyman, Sandy J. and Watt, M. Paula (2013). 5-Azacytidine as a Tool to Induce Somaclonal Variants with Useful Traits in Sugarcane (Saccharum Spp.). *Plant Biotechnology Reports*, 7(4): 489–502.

Murashige, Toshio and Skoog, Folke (1962). A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiologia Plantarum*, 15(3): 473–497.

Nagai, Takeharu; Ibata, Keiji; Park, Eun Sun; Kubota, Mie; Mikoshiba, Katsuhiko and Miyawaki, Atsushi (2002). A Variant of Yellow Fluorescent Protein with Fast and Efficient Maturation for Cell-Biological Applications. *Nature Biotechnology*, 20(1): 87–90.

Naish, Matthew; Alonge, Michael; Wlodzimierz, Piotr; Tock, Andrew J.; Abramson, Bradley W.; Schmücker, Anna; Mandáková, Terezie; Jamge, Bhagyshree; Lambing, Christophe; Kuo, Pallas; Yelina, Natasha; Hartwick, Nolan; Colt, Kelly; Smith, Lisa M.; Ton, Jurriaan; Kakutani, Tetsuji; Martienssen, Robert A.; Schneeberger, Korbinian; Lysak, Martin A.; Berger, Frédéric; Bousios, Alexandros; Michael, Todd P.; Schatz, Michael C. and Henderson, Ian R. (2021). The Genetic and Epigenetic Landscape of the Arabidopsis Centromeres. *Science*, 374(6569): eabi7489.

Naish, Matthew and Henderson, Ian R. (2024). The Structure, Function, and Evolution of Plant Centromeres. *Genome Research*, 34(2): 161–178.

Ngou, Bruno Pok Man; Ahn, Hee-Kyung; Ding, Pingtao and Jones, Jonathan D. G. (2021). Mutual Potentiation of Plant Immunity by Cell-Surface and Intracellular Receptors. *Nature*, 592(7852): 110–115.

Ngou, Bruno Pok Man; Ding, Pingtao and Jones, Jonathan D G (2022). Thirty Years of Resistance: Zig-Zag through the Plant Immune System. *The Plant Cell*, 34(5): 1447–1478.

Ni, Peng; Huang, Neng; Nie, Fan; Zhang, Jun; Zhang, Zhi; Wu, Bo; Bai, Lu; Liu, Wende; Xiao, Chuan-Le; Luo, Feng and Wang, Jianxin (2021). Genome-Wide Detection of Cytosine Methylations in Plant from Nanopore Data Using Deep Learning. *Nature Communications*, 12(1): 5976.

Nicolopoulou-Stamati, Polyxeni; Maipas, Sotirios; Kotampasi, Chrysanthi; Stamatis, Panagiotis and Hens, Luc (2016). Chemical Pesticides and Human Health: The Urgent

Need for a New Concept in Agriculture. *Frontiers in Public Health*, 4. available at https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2016.00148/full [12 October 2024].

Nie, Wen-Feng; Lei, Mingguang; Zhang, Mingxuan; Tang, Kai; Huang, Huan; Zhang, Cuijun; Miki, Daisuke; Liu, Pan; Yang, Yu; Wang, Xingang; Zhang, Heng; Lang, Zhaobo; Liu, Na; Xu, Xuechen; Yelagandula, Ramesh; Zhang, Huiming; Wang, Zhidan; Chai, Xiaoqiang; Andreucci, Andrea; Yu, Jing-Quan; Berger, Frederic; Lozano-Duran, Rosa and Zhu, Jian-Kang (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.

Ning, Yong-Qiang; Liu, Na; Lan, Ke-Ke; Su, Yin-Na; Li, Lin; Chen, She and He, Xin-Jian (2020). DREAM Complex Suppresses DNA Methylation Maintenance Genes and Precludes DNA Hypermethylation. *Nature Plants*, 6(8): 942–956.

Nomoto, Mika; Skelly, Michael J.; Itaya, Tomotaka; Mori, Tsuyoshi; Suzuki, Takamasa; Matsushita, Tomonao; Tokizawa, Mutsutomo; Kuwata, Keiko; Mori, Hitoshi; Yamamoto, Yoshiharu Y.; Higashiyama, Tetsuya; Tsukagoshi, Hironaka; Spoel, Steven H. and Tada, Yasuomi (2021). Suppression of MYC Transcription Activators by the Immune Cofactor NPR1 Fine-Tunes Plant Immune Responses. *Cell Reports*, 37(11): 110125.

Nuñez, James K.; Chen, Jin; Pommier, Greg C.; Cogan, J. Zachery; Replogle, Joseph M.; Adriaens, Carmen; Ramadoss, Gokul N.; Shi, Quanming; Hung, King L.; Samelson, Avi J.; Pogson, Angela N.; Kim, James Y. S.; Chung, Amanda; Leonetti, Manuel D.; Chang, Howard Y.; Kampmann, Martin; Bernstein, Bradley E.; Hovestadt, Volker; Gilbert, Luke A. and Weissman, Jonathan S. (2021). Genome-Wide Programmable Transcriptional Memory by CRISPR-Based Epigenome Editing. *Cell*, 184(9): 2503-2519.e17.

Nützmann, Hans-Wilhelm; Doerr, Daniel; Ramírez-Colmenero, América; Sotelo-Fonseca, Jesús Emiliano; Wegel, Eva; Di Stefano, Marco; Wingett, Steven W.; Fraser, Peter; Hurst, Laurence; Fernandez-Valverde, Selene L. and Osbourn, Anne (2020). Active and Repressed Biosynthetic Gene Clusters Have Spatially Distinct Chromosome States. *Proceedings of the National Academy of Sciences*, 117(24): 13800–13809.

Oberkofler, Vicky; Pratx, Loris and Bäurle, Isabel (2021). Epigenetic Regulation of Abiotic Stress Memory: Maintaining the Good Things While They Last. *Current Opinion in Plant Biology*, 61: 102007.

O'Malley, Ronan C.; Huang, Shao-shan Carol; Song, Liang; Lewsey, Mathew G.; Bartlett, Anna; Nery, Joseph R.; Galli, Mary; Gallavotti, Andrea and Ecker, Joseph R. (2016). Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. *Cell*, 165(5): 1280–1292.

Ortega-Galisteo, Ana Pilar; Morales-Ruiz, Teresa; Ariza, Rafael R. and Roldán-Arjona, Teresa (2008). Arabidopsis DEMETER-LIKE Proteins DML2 and DML3 Are Required for Appropriate Distribution of DNA Methylation Marks. *Plant Molecular Biology*, 67(6): 671–681.

Ouyang, Weizhi; Xiao, Qin; Li, Guoliang and Li, Xingwang (2021). Technologies for Capturing 3D Genome Architecture in Plants. *Trends in Plant Science*, 26(2): 196–197.

Panda, Kaushik; Ji, Lexiang; Neumann, Drexel A.; Daron, Josquin; Schmitz, Robert J. and Slotkin, R. Keith (2016). Full-Length Autonomous Transposable Elements Are

Preferentially Targeted by Expression-Dependent Forms of RNA-Directed DNA Methylation. *Genome Biology*, 17(1): 170.

Panda, Kaushik and Slotkin, R. Keith (2020). Long-Read cDNA Sequencing Enables a "Gene-Like" Transcript Annotation of Transposable Elements. *The Plant Cell*, 32(9): 2687–2698.

Papareddy, Ranjith K.; Páldi, Katalin; Paulraj, Subramanian; Kao, Ping; Lutzmayer, Stefan and Nodine, Michael D. (2020). Chromatin Regulates Expression of Small RNAs to Help Maintain Transposon Methylome Homeostasis in Arabidopsis. *Genome Biology*, 21(1): 251.

Papikian, Ashot; Liu, Wanlu; Gallego-Bartolomé, Javier and Jacobsen, Steven E. (2019). Site-Specific Manipulation of Arabidopsis Loci Using CRISPR-Cas9 SunTag Systems. *Nature Communications*, 10(1): 729.

Park, Jin-Sup; Frost, Jennifer M.; Park, Kyunghyuk; Ohr, Hyonhwa; Park, Guen Tae; Kim, Seohyun; Eom, Hyunjoo; Lee, Ilha; Brooks, Janie S.; Fischer, Robert L. and Choi, Yeonhee (2017). Control of DEMETER DNA Demethylase Gene Transcription in Male and Female Gamete Companion Cells in Arabidopsis Thaliana. *Proceedings of the National Academy of Sciences of the United States of America*, 114(8): 2078–2083.

Park, Yongseok and Wu, Hao (2016). Differential Methylation Analysis for BS-Seq Data under General Experimental Design. *Bioinformatics*, 32(10): 1446–1453.

Parrilla-Doblas, J. T.; Morales-Ruiz, T.; Ariza, R. R.; Martínez-Macías, M. I. and Roldán-Arjona, T. (2022). The C-Terminal Domain of *Arabidopsis* ROS1 DNA Demethylase Interacts with Histone H3 and Is Required for DNA Binding and Catalytic Activity. *DNA Repair*, 115: 103341.

Patro, Rob; Duggal, Geet; Love, Michael I.; Irizarry, Rafael A. and Kingsford, Carl (2017). Salmon Provides Fast and Bias-Aware Quantification of Transcript Expression. *Nature Methods*, 14(4): 417–419.

Pecinka, Ales and Liu, Chun-hsin (2014). Drugs for Plant Chromosome and Chromatin Research. *Cytogenetic and Genome Research*, 143(1–3): 51–9.

Pedro, Daniel Longhi Fernandes; Amorim, Tharcisio Soares; Varani, Alessandro; Guyot, Romain; Domingues, Douglas Silva and Paschoal, Alexandre Rossi (2021). An Atlas of Plant Transposable Elements. *F1000Research*, 10: 1194.

Pei, Liuling; Zhang, Lin; Li, Jianying; Shen, Chao; Qiu, Ping; Tu, Lili; Zhang, Xianlong and Wang, Maojun (2019). Tracing the Origin and Evolution History of Methylation-Related Genes in Plants. *BMC Plant Biology*, 19(1): 307.

Pétriacq, Pierre; Stassen, Joost H.M. and Ton, Jurriaan (2016). Spore Density Determines Infection Strategy by the Plant Pathogenic Fungus Plectosphaerella Cucumerina. *Plant Physiology*, 170(4): 2325–2339.

Pieterse, Corné M.J.; Van der Does, Dieuwertje; Zamioudis, Christos; Leon-Reyes, Antonio and Van Wees, Saskia C.M. (2012). Hormonal Modulation of Plant Immunity. *Annual Review of Cell and Developmental Biology*, 28(1): 489–521.

Pieterse, Corné M.J.; Zamioudis, Christos; Berendsen, Roeland L.; Weller, David M.; Van Wees, Saskia C.M. and Bakker, Peter A.H.M. (2014). Induced Systemic Resistance by Beneficial Microbes. *Annual Review of Phytopathology*, 52(1): 347–375.

Ploetz, Randy C. (2005). Panama Disease: An Old Nemesis Rears Its Ugly Head: Part 1. The Beginnings of the Banana Export Trades. *Plant Health Progress*, 6(1): 18.

Ploetz, Randy C. (2006). Panama Disease: An Old Nemesis Rears Its Ugly: Head Part 2. The Cavendish Era and Beyond. *Plant Health Progress*, 7(1): 36.

Poland, Jesse A.; Balint-Kurti, Peter J.; Wisser, Randall J.; Pratt, Richard C. and Nelson, Rebecca J. (2009). Shades of Gray: The World of Quantitative Disease Resistance. *Trends in Plant Science*, 14(1): 21–29.

Ponferrada-Marín, María Isabel; Martínez-Macías, María Isabel; Morales-Ruiz, Teresa; Roldán-Arjona, Teresa and Ariza, Rafael R. (2010). Methylation-Independent DNA Binding Modulates Specificity of Repressor of Silencing 1 (ROS1) and Facilitates Demethylation in Long Substrates. *The Journal of Biological Chemistry*, 285(30): 23032.

Povirk, L. F.; Wübter, W.; Köhnlein, W. and Hutchinson, F. (1977). DNA Double-Strand Breaks and Alkali-Labile Bonds Produced by Bleomycin. *Nucleic Acids Research*, 4(10): 3573–3580.

Prime-A-Plant Group; Conrath, Uwe; Beckers, Gerold J. M.; Flors, Victor; García-Agustín, Pilar; Jakab, Gábor; Mauch, Felix; Newman, Mari-Anne; Pieterse, Corné M. J.; Poinssot, Benoit; Pozo, María J.; Pugin, Alain; Schaffrath, Ulrich; Ton, Jurriaan; Wendehenne, David; Zimmerli, Laurent and Mauch-Mani, Brigitte (2006). Priming: Getting Ready for Battle. *Molecular plant-microbe interactions: MPMI*, 19(10): 1062–1071.

Qian, Weiqiang; Miki, Daisuke; Lei, Mingguang; Zhu, Xiaohong; Zhang, Huiming; Liu, Yunhua; Li, Yan; Lang, Zhaobo; Wang, Jing; Tang, Kai; Liu, Renyi and Zhu, Jian-Kang (2014). Regulation of Active DNA Demethylation by an α-Crystallin Domain Protein in Arabidopsis. *Molecular Cell*, 55(3): 361–371.

Qian, Weiqiang; Miki, Daisuke; Zhang, Heng; Liu, Yunhua; Zhang, Xi; Tang, Kai; Kan, Yunchao; La, Honggui; Li, Xiaojie; Li, Shaofang; Zhu, Xiaohong; Shi, Xiaobing; Zhang, Kangling; Pontes, Olga; Chen, Xuemei; Liu, Renyi; Gong, Zhizhong and Zhu, Jian-Kang (2012). A Histone Acetyltransferase Regulates Active DNA Demethylation in Arabidopsis. *Science*, 336(6087): 1445–1448.

Quadrana, Leandro; Bortolini Silveira, Amanda; Mayhew, George F; LeBlanc, Chantal; Martienssen, Robert A; Jeddeloh, Jeffrey A and Colot, Vincent (2016). The Arabidopsis Thaliana Mobilome and Its Impact at the Species Level (D. Zilberman, Ed.). *eLife*, 5: e15716.

Quadrana, Leandro; Etcheverry, Mathilde; Gilly, Arthur; Caillieux, Erwann; Madoui, Mohammed-Amin; Guy, Julie; Bortolini Silveira, Amanda; Engelen, Stefan; Baillet, Victoire; Wincker, Patrick; Aury, Jean-Marc and Colot, Vincent (2019). Transposition Favors the Generation of Large Effect Mutations That May Facilitate Rapid Adaption. *Nature Communications*, 10(1): 3421.

Quesneville, Hadi (2020). Twenty Years of Transposable Element Analysis in the Arabidopsis Thaliana Genome. *Mobile DNA*, 11(1): 28.

Quigley, Sarah; Damas, Joana; Larkin, Denis M and Farré, Marta (2023). syntenyPlotteR: A User-Friendly R Package to Visualize Genome Synteny, Ideal for Both Experienced and Novice Bioinformaticians. *Bioinformatics Advances*, 3(1): vbad161.

Quinlan, Aaron R. and Hall, Ira M. (2010). BEDTools: A Flexible Suite of Utilities for Comparing Genomic Features. *Bioinformatics*, 26(6): 841–842.

Raingeval, Mathieu; Leduque, Basile; Baduel, Pierre; Edera, Alejandro; Roux, Fabrice; Colot, Vincent and Quadrana, Leandro (2024). Retrotransposon-Driven Environmental Regulation of FLC Leads to Adaptive Response to Herbicide. *Nature Plants*: 1–10.

Rasmann, Sergio; Vos, Martin De; Casteel, Clare L.; Tian, Donglan; Halitschke, Rayko; Sun, Joel Y.; Agrawal, Anurag A.; Felton, Gary W. and Jander, Georg (2012). Herbivory in the Previous Generation Primes Plants for Enhanced Insect Resistance. *Plant Physiology*, 158(2): 854–863.

Rausch, Stephanie and Laubinger, Sascha (2016). Rapid Assessment of DNA Methylation Changes in Response to Salicylic Acid by Chop-qPCR. *Methods in Molecular Biology (Clifton, N.J.)*, 1398: 345–356.

Rea, Matthew; Zheng, Wenguang; Chen, Ming; Braud, Christopher; Bhangu, Drutdaman; Rognan, Tara N. and Xiao, Wenyan (2012). Histone H1 Affects Gene Imprinting and DNA Methylation in Arabidopsis. *The Plant Journal: For Cell and Molecular Biology*, 71(5): 776–786.

Reimer-Michalski, Eva-Maria and Conrath, Uwe (2016). Innate Immune Memory in Plants. *Seminars in Immunology*, 28(4): 319–327.

Reinders, Jon; Wulff, Brande B. H.; Mirouze, Marie; Marí-Ordóñez, Arturo; Dapp, Mélanie; Rozhon, Wilfried; Bucher, Etienne; Theiler, Grégory and Paszkowski, Jerzy (2009). Compromised Stability of DNA Methylation and Transposon Immobilization in Mosaic Arabidopsis Epigenomes. *Genes & Development*, 23(8): 939–950.

Remick, Brenna C.; Gaidt, Moritz M. and Vance, Russell E. (2023). Effector-Triggered Immunity. *Annual Review of Immunology*, 41(Volume 41, 2023): 453–481.

Rigal, Mélanie; Kevei, Zoltán; Pélissier, Thierry and Mathieu, Olivier (2012). DNA Methylation in an Intron of the IBM1 Histone Demethylase Gene Stabilizes Chromatin Modification Patterns. *The EMBO Journal*, 31(13): 2981–2993.

Risso, Davide; Ngai, John; Speed, Terence P. and Dudoit, Sandrine (2014). Normalization of RNA-Seq Data Using Factor Analysis of Control Genes or Samples. *Nature Biotechnology*, 32(9): 896–902.

Ristaino, Jean B.; Anderson, Pamela K.; Bebber, Daniel P.; Brauman, Kate A.; Cunniffe, Nik J.; Fedoroff, Nina V.; Finegold, Cambria; Garrett, Karen A.; Gilligan, Christopher A.; Jones, Christopher M.; Martin, Michael D.; MacDonald, Graham K.; Neenan, Patricia; Records, Angela; Schmale, David G.; Tateosian, Laura and Wei, Qingshan (2021). The Persistent Threat of Emerging Plant Disease Pandemics to Global Food Security. *Proceedings of the National Academy of Sciences*, 118(23): e2022239118.

RNAcentral Consortium (2021). RNAcentral 2021: Secondary Structure Integration, Improved Sequence Search and New Member Databases. *Nucleic Acids Research*, 49(D1): D212–D220.

Roberts, D. A. (1983). Acquired Resistance to Tobacco Mosaic Virus Transmitted to the Progeny of Hypersensitive Tobacco. *Virology*, 124(1): 161–163.

Roberts, Janet M.; Carvalhais, Lilia C.; O'Dwyer, Cecilia; Rincón-Flórez, Vivian A. and Drenth, Andre (2024). Diagnostics of Fusarium Wilt in Banana: Current Status and Challenges. *Plant Pathology*, 73(4): 760–776.

Robinson, Mark D. and Oshlack, Alicia (2010). A Scaling Normalization Method for Differential Expression Analysis of RNA-Seq Data. *Genome Biology*, 11(3): R25.

Rueden, Curtis T.; Schindelin, Johannes; Hiner, Mark C.; DeZonia, Barry E.; Walter, Alison E.; Arena, Ellen T. and Eliceiri, Kevin W. (2017). ImageJ2: ImageJ for the next Generation of Scientific Image Data. *BMC Bioinformatics*, 18(1): 529.

Ruiz-Velasco, Mariana and Zaugg, Judith B. (2017). Structure Meets Function: How Chromatin Organisation Conveys Functionality. *Current Opinion in Systems Biology*, 1: 129–136.

Rutowicz, Kinga; Lirski, Maciej; Mermaz, Benoît; Teano, Gianluca; Schubert, Jasmin; Mestiri, Imen; Kroteń, Magdalena A.; Fabrice, Tohnyui Ndinyanka; Fritz, Simon; Grob, Stefan; Ringli, Christoph; Cherkezyan, Lusik; Barneche, Fredy; Jerzmanowski, Andrzej and Baroux, Célia (2019). Linker Histones Are Fine-Scale Chromatin Architects Modulating Developmental Decisions in Arabidopsis. *Genome Biology*, 20(1): 157.

Salih, Abubakr A. M.; Baraibar, Marta; Mwangi, Kenneth Kemucie and Artan, Guleid (2020). Climate Change and Locust Outbreak in East Africa. *Nature Climate Change*, 10(7): 584–585.

Saze, Hidetoshi; Kitayama, Junko; Takashima, Kazuya; Miura, Saori; Harukawa, Yoshiko; Ito, Tasuku and Kakutani, Tetsuji (2013). Mechanism for Full-Length RNA Processing of Arabidopsis Genes Containing Intragenic Heterochromatin. *Nature Communications*, 4(1): 2301.

Saze, Hidetoshi; Shiraishi, Akiko; Miura, Asuka and Kakutani, Tetsuji (2008). Control of Genic DNA Methylation by a jmjC Domain-Containing Protein in Arabidopsis Thaliana. *Science*, 319(5862): 462–465.

Schoft, Vera K.; Chumak, Nina; Choi, Yeonhee; Hannon, Mike; Garcia-Aguilar, Marcelina; Machlicova, Adriana; Slusarz, Lucyna; Mosiolek, Magdalena; Park, Jin-Sup; Park, Guen Tae; Fischer, Robert L. and Tamaru, Hisashi (2011). Function of the DEMETER DNA Glycosylase in the Arabidopsis Thaliana Male Gametophyte. *Proceedings of the National Academy of Sciences*, 108(19): 8042–8047.

Schultz, Matthew D.; Schmitz, Robert J. and Ecker, Joseph R. (2012). 'Leveling' the Playing Field for Analyses of Single-Base Resolution DNA Methylomes. *Trends in Genetics*, 28(12): 583–585.

Schumann, Ulrike; Lee, Joanne; Kazan, Kemal; Ayliffe, Michael and Wang, Ming-Bo (2017). DNA-Demethylase Regulated Genes Show Methylation-Independent Spatiotemporal Expression Patterns. *Frontiers in Plant Science*, 8. available at https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2017.01449/full [12 October 2024].

Schwarzenbacher, Roland E.; Wardell, Grace; Stassen, Joost; Guest, Emily; Zhang, Peijun; Luna, Estrella and Ton, Jurriaan (2020). The IBI1 Receptor of β-Aminobutyric Acid

Interacts with VOZ Transcription Factors to Regulate Abscisic Acid Signaling and Callose-Associated Defense. *Molecular Plant*, 13(10): 1455–1469.

Sequeira-Mendes, Joana; Aragüez, Irene; Peiró, Ramón; Mendez-Giraldez, Raul; Zhang, Xiaoyu; Jacobsen, Steven E.; Bastolla, Ugo and Gutierrez, Crisanto (2014). The Functional Topography of the Arabidopsis Genome Is Organized in a Reduced Number of Linear Motifs of Chromatin States[C][W]. *The Plant Cell*, 26(6): 2351–2366.

Shahid, Saima and Axtell, Michael J. (2014). Identification and Annotation of Small RNA Genes Using ShortStack. *Methods (San Diego, Calif.)*, 67(1): 20–27.

Sharma, Anket; Kumar, Vinod; Shahzad, Babar; Tanveer, Mohsin; Sidhu, Gagan Preet Singh; Handa, Neha; Kohli, Sukhmeen Kaur; Yadav, Poonam; Bali, Aditi Shreeya; Parihar, Ripu Daman; Dar, Owias Iqbal; Singh, Kirpal; Jasrotia, Shivam; Bakshi, Palak; Ramakrishnan, M.; Kumar, Sandeep; Bhardwaj, Renu and Thukral, Ashwani Kumar (2019). Worldwide Pesticide Usage and Its Impacts on Ecosystem. *SN Applied Sciences*, 1(11): 1446.

Sheikh, Arsheed H; Nawaz, Kashif; Tabassum, Naheed; Almeida-Trapp, Marilia; Mariappan, Kiruthiga G; Alhoraibi, Hanna; Rayapuram, Naganand; Aranda, Manuel; Groth, Martin and Hirt, Heribert (2023). Linker Histone H1 Modulates Defense Priming and Immunity in Plants. *Nucleic Acids Research*: gkad106.

Shimada, Atsushi; Cahn, Jonathan; Ernst, Evan; Lynn, Jason; Grimanelli, Daniel; Henderson, Ian; Kakutani, Tetsuji and Martienssen, Robert A. (2024). Retrotransposon Addiction Promotes Centromere Function via Epigenetically Activated Small RNAs. *Nature Plants*: 1–13.

Sigman, Meredith J.; Panda, Kaushik; Kirchner, Rachel; McLain, Lauren L.; Payne, Hayden; Peasari, John Reddy; Husbands, Aman Y.; Slotkin, R. Keith and McCue, Andrea D. (2021). An siRNA-Guided ARGONAUTE Protein Directs RNA Polymerase V to Initiate DNA Methylation. *Nature Plants*, 7(11): 1461–1474.

Siligato, Riccardo; Wang, Xin; Yadav, Shri Ram; Lehesranta, Satu; Ma, Guojie; Ursache, Robertas; Sevilem, Iris; Zhang, Jing; Gorte, Maartje; Prasad, Kalika; Wrzaczek, Michael; Heidstra, Renze; Murphy, Angus; Scheres, Ben and Mähönen, Ari Pekka (2016). MultiSite Gateway-Compatible Cell Type-Specific Gene-Inducible System for Plants. *Plant Physiology*, 170(2): 627–641.

Silverstein, Kevin A.T.; Graham, Michelle A.; Paape, Timothy D. and VandenBosch, Kathryn A. (2005). Genome Organization of More Than 300 Defensin-Like Genes in Arabidopsis. *Plant Physiology*, 138(2): 600–610.

Singh, Brajesh K.; Delgado-Baquerizo, Manuel; Egidi, Eleonora; Guirado, Emilio; Leach, Jan E.; Liu, Hongwei and Trivedi, Pankaj (2023). Climate Change Impacts on Plant Pathogens, Food Security and Paths Forward. *Nature Reviews Microbiology*, 21(10): 640–656.

Sjogren, Caroline A.; Bolaris, Stephen C. and Larsen, Paul B. (2015). Aluminum-Dependent Terminal Differentiation of the Arabidopsis Root Tip Is Mediated through an ATR-, ALT2-, and SOG1-Regulated Transcriptional Response. *The Plant Cell*, 27(9): 2501–2515.

Slaughter, Ana; Daniel, Xavier; Flors, Victor; Luna, Estrella; Hohn, Barbara and Mauch-Mani, Brigitte (2012). Descendants of Primed Arabidopsis Plants Exhibit Resistance to Biotic Stress. *Plant Physiology*, 158(2): 835–843.

Slotkin, R. Keith; Vaughn, Matthew; Borges, Filipe; Tanurdzić, Milos; Becker, Jörg D.; Feijó, José A. and Martienssen, Robert A. (2009). Epigenetic Reprogramming and Small RNA Silencing of Transposable Elements in Pollen. *Cell*, 136(3): 461–472.

Slusarenko, Alan J. and Schlaich, Nikolaus L. (2003). Downy Mildew of Arabidopsis Thaliana Caused by Hyaloperonospora Parasitica (Formerly Peronospora Parasitica). *Molecular Plant Pathology*, 4(3): 159–170.

Smith, Lisa M. (2019). Epigenetic Regulation of mRNA Polyadenylation Site Selection. *Plant Physiology*, 180(1): 7–9.

Song, Jie; Angel, Andrew; Howard, Martin and Dean, Caroline (2012). Vernalization – a Cold-Induced Epigenetic Switch. *Journal of Cell Science*, 125(16): 3723–3731.

Song, Junqi and Bent, Andrew F. (2014). Microbial Pathogens Trigger Host DNA Double-Strand Breaks Whose Abundance Is Reduced by Plant Defense Responses. *PLOS Pathogens*, 10(4): e1004030.

Soppe, Wim J. J.; Jasencakova, Zuzana; Houben, Andreas; Kakutani, Tetsuji; Meister, Armin; Huang, Michael S.; Jacobsen, Steven E.; Schubert, Ingo and Fransz, Paul F. (2002). DNA Methylation Controls Histone H3 Lysine 9 Methylation and Heterochromatin Assembly in Arabidopsis. *The EMBO journal*, 21(23): 6549–6559.

Spoel, Steven H and Dong, Xinnian (2024). Salicylic Acid in Plant Immunity and Beyond. *The Plant Cell*, 36(5): 1451–1464.

Srivastava, Prashant K.; Moturu, Taraka Ramji; Pandey, Priyanka; Baldwin, Ian T. and Pandey, Shree P. (2014). A Comparison of Performance of Plant miRNA Target Prediction Tools and the Characterization of Features for Genome-Wide Target Prediction. *BMC Genomics*, 15(1): 348.

Stassen, Joost H. M.; López, Ana; Jain, Ritushree; Pascual-Pardo, David; Luna, Estrella; Smith, Lisa M. and Ton, Jurriaan (2018). The Relationship between Transgenerational Acquired Resistance and Global DNA Methylation in Arabidopsis. *Scientific Reports*, 8(1): 14761.

Statello, Luisa; Guo, Chun-Jie; Chen, Ling-Ling and Huarte, Maite (2021). Gene Regulation by Long Non-Coding RNAs and Its Biological Functions. *Nature Reviews Molecular Cell Biology*, 22(2): 96–118.

Stroud, Hume; Do, Truman; Du, Jiamu; Zhong, Xuehua; Feng, Suhua; Johnson, Lianna; Patel, Dinshaw J. and Jacobsen, Steven E. (2014). Non-CG Methylation Patterns Shape the Epigenetic Landscape in Arabidopsis. *Nature Structural & Molecular Biology*, 21(1): 64–72.

Stroud, Hume; Greenberg, Maxim V. C.; Feng, Suhua; Bernatavichute, Yana V. and Jacobsen, Steven E. (2013). Comprehensive Analysis of Silencing Mutants Reveals Complex Regulation of the Arabidopsis Methylome. *Cell*, 152(1): 352–364.

Stuart, Tim; Buckberry, Sam and Lister, Ryan (2018). Approaches for the Analysis and Interpretation of Whole Genome Bisulfite Sequencing Data, 299–310, in: Jeltsch, A.

and Rots, M. G. (Eds.), *Epigenome Editing: Methods and Protocols*. New York, NY: Springer. available at https://doi.org/10.1007/978-1-4939-7774-1_17 [10 May 2024].

Sun, Linhua; Jing, Yuqing; Liu, Xinyu; Li, Qi; Xue, Zhihui; Cheng, Zhukuan; Wang, Daowen; He, Hang and Qian, Weiqiang (2020). Heat Stress-Induced Transposon Activation Correlates with 3D Chromatin Organization Rearrangement in Arabidopsis. *Nature Communications*, 11(1): 1886.

Sun, Tongjun; Zhang, Yaxi; Li, Yan; Zhang, Qian; Ding, Yuli and Zhang, Yuelin (2015). ChIP-Seq Reveals Broad Roles of SARD1 and CBP60g in Regulating Plant Immunity. *Nature Communications*, 6: 10159.

Sundström, Jens F.; Berlin, Anna; Phuong, Nam Kieu; Karlsson, Milla and Andreasson, Erik (2024). New Genomic Techniques Can Contribute to Reduced Pesticide Usage in Europe. *Plants, People, Planet*, n/a(n/a). available at https://onlinelibrary.wiley.com/doi/abs/10.1002/ppp3.10559 [12 October 2024].

Tada, Yasuomi; Spoel, Steven H.; Pajerowska-Mukhtar, Karolina; Mou, Zhonglin; Song, Junqi; Wang, Chun; Zuo, Jianru and Dong, Xinnian (2008). Plant Immunity Requires Conformational Changes [Corrected] of NPR1 via S-Nitrosylation and Thioredoxins. *Science (New York, N.Y.)*, 321(5891): 952–956.

Tang, Fiona H. M.; Lenzen, Manfred; McBratney, Alexander and Maggi, Federico (2021). Risk of Pesticide Pollution at the Global Scale. *Nature Geoscience*, 14(4): 206–210.

Tang, Kai; Lang, Zhaobo; Zhang, Heng and Zhu, Jian-Kang (2016). The DNA Demethylase ROS1 Targets Genomic Regions with Distinct Chromatin Modifications. *Nature Plants*, 2(11): 1–10.

Tao, Chia-Nan; Buswell, Will; Zhang, Peijun; Walker, Heather; Johnson, Irene; Field, Katie; Schwarzenbacher, Roland and Ton, Jurriaan (2022). A Single Amino Acid Transporter Controls the Uptake of Priming-Inducing Beta-Amino Acids and the Associated Tradeoff between Induced Resistance and Plant Growth. *The Plant Cell*, 34(12): 4840–4856.

Thieme, Michael; Lanciano, Sophie; Balzergue, Sandrine; Daccord, Nicolas; Mirouze, Marie and Bucher, Etienne (2017). Inhibition of RNA Polymerase II Allows Controlled Mobilisation of Retrotransposons for Plant Breeding. *Genome Biology*, 18(1): 134.

Thorvaldsdóttir, Helga; Robinson, James T. and Mesirov, Jill P. (2013). Integrative Genomics Viewer (IGV): High-Performance Genomics Data Visualization and Exploration. *Briefings in Bioinformatics*, 14(2): 178–192.

Tian, Feng; Yang, De-Chang; Meng, Yu-Qi; Jin, Jinpu and Gao, Ge (2020). PlantRegMap: Charting Functional Regulatory Maps in Plants. *Nucleic Acids Research*, 48(D1): D1104–D1113.

Tian, Hainan; Wu, Zhongshou; Chen, Siyu; Ao, Kevin; Huang, Weijie; Yaghmaiean, Hoda; Sun, Tongjun; Xu, Fang; Zhang, Yanjun; Wang, Shucai; Li, Xin and Zhang, Yuelin (2021). Activation of TIR Signalling Boosts Pattern-Triggered Immunity. *Nature*, 598(7881): 500–503.

Tirot, Louis and Jullien, Pauline E. (2022). Epigenetic Dynamics during Sexual Reproduction: At the Nexus of Developmental Control and Genomic Integrity. *Current Opinion in Plant Biology*, 69: 102278.

To, Taiko K.; Saze, Hidetoshi and Kakutani, Tetsuji (2015). DNA Methylation within Transcribed Regions. *Plant Physiology*, 168(4): 1219–1225.

Ton, Jurriaan; Jakab, Gabor; Toquin, Valérie; Flors, Victor; Iavicoli, Annalisa; Maeder, Muriel N.; Métraux, Jean-Pierre and Mauch-Mani, Brigitte (2005). Dissecting the β-Aminobutyric Acid–Induced Priming Phenomenon in Arabidopsis. *The Plant Cell*, 17(3): 987–999.

Ton, Jurriaan and Mauch-Mani, Brigitte (2004). β-Amino-Butyric Acid-Induced Resistance against Necrotrophic Pathogens Is Based on ABA-Dependent Priming for Callose. *The Plant Journal*, 38(1): 119–130.

Ton, Jurriaan; Van Pelt, Johan A.; Van Loon, L. C. and Pieterse, Corné M. J. (2002). Differential Effectiveness of Salicylate-Dependent and Jasmonate/Ethylene-Dependent Induced Resistance in Arabidopsis. *Molecular Plant-Microbe Interactions®*, 15(1): 27–34.

Tsuchiya, Tokuji and Eulgem, Thomas (2013). An Alternative Polyadenylation Mechanism Coopted to the Arabidopsis RPP7 Gene through Intronic Retrotransposon Domestication. *Proceedings of the National Academy of Sciences*, 110(37): E3535–E3543.

Tsuchiya, Tokuji and Eulgem, Thomas (2014). The PHD-Finger Module of the Arabidopsis Thaliana Defense Regulator EDM2 Can Recognize Triply Modified Histone H3 Peptides. *Plant Signaling & Behavior*, 9(7): e29202.

Underwood, Charles J.; Choi, Kyuha; Lambing, Christophe; Zhao, Xiaohui; Serra, Heïdi; Borges, Filipe; Simorowski, Joe; Ernst, Evan; Jacob, Yannick; Henderson, Ian R. and Martienssen, Robert A. (2018). Epigenetic Activation of Meiotic Recombination near Arabidopsis Thaliana Centromeres via Loss of H3K9me2 and Non-CG DNA Methylation. *Genome Research*, 28(4): 519–531.

Valinluck, Victoria and Sowers, Lawrence C. (2007). Endogenous Cytosine Damage Products Alter the Site Selectivity of Human DNA Maintenance Methyltransferase DNMT1. *Cancer Research*, 67(3): 946–950.

Vashisht, D.; Hesselink, A.; Pierik, R.; Ammerlaan, J. M. H.; Bailey-Serres, J.; Visser, E. J. W.; Pedersen, O.; van Zanten, M.; Vreugdenhil, D.; Jamar, D. C. L.; Voesenek, L. a. C. J. and Sasidharan, R. (2011). Natural Variation of Submergence Tolerance among Arabidopsis Thaliana Accessions. *New Phytologist*, 190(2): 299–310.

Vega-Muñoz, Isaac; Herrera-Estrella, Alfredo; Martínez-de la Vega, Octavio and Heil, Martin (2023). ATM and ATR, Two Central Players of the DNA Damage Response, Are Involved in the Induction of Systemic Acquired Resistance by Extracellular DNA, but Not the Plant Wound Response. *Frontiers in Immunology*, 14. available at https://www.frontiersin.org/journals/immunology/articles/10.3389/fimmu.2023.1175786/full [16 July 2024].

Vladejić, Jovanka; Kovacik, Martin; Zwyrtková, Jana; Szurman-Zubrzycka, Miriam; Doležel, Jaroslav and Pecinka, Ales (2024). Zeocin-Induced DNA Damage Response in Barley and Its Dependence on ATR. *Scientific Reports*, 14(1): 3119.

Vonapartis, Eliana; Mohamed, Deka; Li, Jingru; Pan, Wenqiang; Wu, Jian and Gazzarrini, Sonia (2022). CBF4/DREB1D Represses to Attenuate ABA, Osmotic and Drought Stress Responses in Arabidopsis. *The Plant Journal*, 110(4): 961–977.

Waddington, C. H. (1942). Canalization of Development and the Inheritance of Acquired Characters. *Nature*, 150(3811): 563–565.

Waese, Jamie; Fan, Jim; Pasha, Asher; Yu, Hans; Fucile, Geoffrey; Shi, Ruian; Cumming, Matthew; Kelley, Lawrence A.; Sternberg, Michael J.; Krishnakumar, Vivek; Ferlanti, Erik; Miller, Jason; Town, Chris; Stuerzlinger, Wolfgang and Provart, Nicholas J. (2017). ePlant: Visualizing and Exploring Multiple Levels of Data for Hypothesis Generation in Plant Biology. *The Plant Cell*, 29(8): 1806–1821.

Walker, James; Zhang, Jingyi; Liu, Yalin; Vickers, Martin; Dolan, Liam; Nakajima, Keiji and Feng, Xiaoqi (2021). Extensive N4 Cytosine Methylation Is Essential for Marchantia Sperm Function. : 2021.02.12.428880. available at https://www.biorxiv.org/content/10.1101/2021.02.12.428880v1 [11 October 2024].

Walters, Dale R.; Ratsep, Jaan and Havis, Neil D. (2013). Controlling Crop Diseases Using Induced Resistance: Challenges for the Future. *Journal of Experimental Botany*, 64(5): 1263–1280.

Walters, Dale; Walsh, David; Newton, Adrian and Lyon, Gary (2005). Induced Resistance for Plant Disease Control: Maximizing the Efficacy of Resistance Elicitors. *Phytopathology*, 95(12): 1368–1373.

Wang, Dong; Amornsiripanitch, Nita and Dong, Xinnian (2006). A Genomic Approach to Identify Regulatory Nodes in the Transcriptional Network of Systemic Acquired Resistance in Plants. *PLOS Pathogens*, 2(11): e123.

Wang, Guanfang; Kong, Hongzhi; Sun, Yujin; Zhang, Xiaohong; Zhang, Wei; Altman, Naomi; dePamphilis, Claude W. and Ma, Hong (2004). Genome-Wide Analysis of the Cyclin Family in Arabidopsis and Comparative Phylogenetic Analysis of Plant Cyclin-Like Proteins. *Plant Physiology*, 135(2): 1084–1099.

Wang, Hsiao-Lin V. and Chekanova, Julia A. (2017). Long Noncoding RNAs in Plants. *Advances in experimental medicine and biology*, 1008: 133–154.

Wang, Ludi; Lau, Yui-Leung; Fan, Lian; Bosch, Maurice and Doughty, James (2023). Pollen Coat Proteomes of Arabidopsis Thaliana, Arabidopsis Lyrata, and Brassica Oleracea Reveal Remarkable Diversity of Small Cysteine-Rich Proteins at the Pollen-Stigma Interface. *Biomolecules*, 13(1): 157.

Wang, Ming; Zhong, Zhenhui; Gallego-Bartolomé, Javier; Li, Zheng; Feng, Suhua; Kuo, Hsuan Yu; Kan, Ryan L.; Lam, Hoiyan; Richey, John Curtis; Tang, Linli; Zhou, Jessica; Liu, Mukun; Jami-Alahmadi, Yasaman; Wohlschlegel, James and Jacobsen, Steven E. (2023). A Gene Silencing Screen Uncovers Diverse Tools for Targeted Gene Repression in Arabidopsis. *Nature Plants*, 9(3): 460–472.

Wang, Qianwen; Bao, Xiucong; Chen, Shengjie; Zhong, Huan; Liu, Yaqin; Zhang, Li; Xia, Yiji; Kragler, Friedrich; Luo, Ming; Li, Xiang David; Lam, Hon-Ming and Zhang, Shoudong (2021). AtHDA6 Functions as an H3K18ac Eraser to Maintain Pericentromeric CHG Methylation in Arabidopsis Thaliana. *Nucleic Acids Research*, 49(17): 9755–9767.

Wang, Shui; Durrant, Wendy E.; Song, Junqi; Spivey, Natalie W. and Dong, Xinnian (2010). Arabidopsis BRCA2 and RAD51 Proteins Are Specifically Involved in Defense Gene Transcription during Plant Immune Responses. *Proceedings of the National Academy of Sciences*, 107(52): 22716–22721.

Wang, Xingang; Duan, Cheng-Guo; Tang, Kai; Wang, Bangshing; Zhang, Huiming; Lei, Mingguang; Lu, Kun; Mangrauthia, Satendra K.; Wang, Pengcheng; Zhu, Guohui; Zhao, Yang and Zhu, Jian-Kang (2013). RNA-Binding Protein Regulates Plant DNA Methylation by Controlling mRNA Processing at the Intronic Heterochromatin-Containing Gene IBM1. *Proceedings of the National Academy of Sciences*, 110(38): 15467–15472.

Waterworth, Wanda M.; Drury, Georgina E.; Bray, Clifford M. and West, Christopher E. (2011). Repairing Breaks in the Plant Genome: The Importance of Keeping It Together. *New Phytologist*, 192(4): 805–822.

Waterworth, Wanda M.; Masnavi, Ghzaleh; Bhardwaj, Rajni M.; Jiang, Qing; Bray, Clifford M. and West, Christopher E. (2010). A Plant DNA Ligase Is an Important Determinant of Seed Longevity. *The Plant Journal: For Cell and Molecular Biology*, 63(5): 848–860.

van Wees, Saskia C. M.; de Swart, Erik A. M.; van Pelt, Johan A.; van Loon, Leendert C. and Pieterse, Corné M. J. (2000). Enhancement of Induced Disease Resistance by Simultaneous Activation of Salicylate- and Jasmonate-Dependent Defense Pathways in Arabidopsis Thaliana. *Proceedings of the National Academy of Sciences*, 97(15): 8711–8716.

Wibowo, Anjar; Becker, Claude; Durr, Julius; Price, Jonathan; Spaepen, Stijn; Hilton, Sally; Putra, Hadi; Papareddy, Ranjith; Saintain, Quentin; Harvey, Sarah; Bending, Gary D.; Schulze-Lefert, Paul; Weigel, Detlef and Gutierrez-Marcos, Jose (2018). Partial Maintenance of Organ-Specific Epigenetic Marks during Plant Asexual Reproduction Leads to Heritable Phenotypic Variation. *Proceedings of the National Academy of Sciences*, 115(39): E9145–E9152.

Wilkinson, S. W.; Hannan Parker, A.; Muench, A.; Wilson, R. S.; Hooshmand, K.; Henderson, M. A.; Moffat, E. K.; Rocha, P. S. C. F.; Hipperson, H.; Stassen, J. H. M.; López Sánchez, A.; Fomsgaard, I. S.; Krokene, P.; Mageroy, M. H. and Ton, J. (2023). Long-Lasting Memory of Jasmonic Acid-Dependent Immunity Requires DNA Demethylation and ARGONAUTE1. *Nature Plants*, 9(1): 81–95.

Wilkinson, Samuel W.; Dalen, Lars S.; Skrautvol, Thomas O.; Ton, Jurriaan; Krokene, Paal and Mageroy, Melissa H. (2022). Transcriptomic Changes during the Establishment of Long-Term Methyl Jasmonate-Induced Resistance in Norway Spruce. *Plant, Cell & Environment*, 45(6): 1891–1913.

Wilkinson, Samuel W.; Magerøy, Melissa H.; López Sánchez, Ana; Smith, Lisa M.; Furci, Leonardo; Cotton, T.E. Anne; Krokene, Paal and Ton, Jurriaan (2019). Surviving in a Hostile World: Plant Strategies to Resist Pests and Diseases. *Annual Review of Phytopathology*, 57(1): 505–529.

Williams, Ben P. and Gehring, Mary (2017). Stable Transgenerational Epigenetic Inheritance Requires a DNA Methylation-Sensing Circuit. *Nature Communications*, 8(1): 2124.

Williams, Ben P.; Pignatta, Daniela; Henikoff, Steven and Gehring, Mary (2015). Methylation-Sensitive Expression of a DNA Demethylase Gene Serves As an Epigenetic Rheostat. *PLOS Genetics*, 11(3): e1005142.

Wilson, Daniel C.; Kempthorne, Christine J.; Carella, Philip; Liscombe, David K. and Cameron, Robin K. (2017). Age-Related Resistance in Arabidopsis Thaliana Involves the MADS-Domain Transcription Factor SHORT VEGETATIVE PHASE and Direct Action of

Salicylic Acid on Pseudomonas Syringae. *Molecular Plant-Microbe Interactions®*, 30(11): 919–929.

Wolinska, Katarzyna Wiktoria and Berens, Matthias Leonhard (2019). Optimal Defense Theory 2.0: Tissue-Specific Stress Defense Prioritization as an Extra Layer of Complexity. *Communicative & Integrative Biology*, 12(1): 91–95.

Wu, Hua-Jun; Wang, Zhi-Min; Wang, Meng and Wang, Xiu-Jie (2013). Widespread Long Noncoding RNAs as Endogenous Target Mimics for MicroRNAs in Plants. *Plant Physiology*, 161(4): 1875–1884.

Wu, Tianzhi; Hu, Erqiang; Xu, Shuangbin; Chen, Meijun; Guo, Pingfan; Dai, Zehan; Feng, Tingze; Zhou, Lang; Tang, Wenli; Zhan, Li; Fu, Xiaocong; Liu, Shanshan; Bo, Xiaochen and Yu, Guangchuang (2021). clusterProfiler 4.0: A Universal Enrichment Tool for Interpreting Omics Data. *The Innovation*, 2(3): 100141.

Xiao, Jun; Lee, Un-Sa and Wagner, Doris (2016). Tug of War: Adding and Removing Histone Lysine Methylation in Arabidopsis. *Current Opinion in Plant Biology*, 34: 41–53.

Xin, Xiu-Fang; Nomura, Kinya; Aung, Kyaw; Velásquez, André C.; Yao, Jian; Boutrot, Freddy; Chang, Jeff H.; Zipfel, Cyril and He, Sheng Yang (2016). Bacteria Establish an Aqueous Living Space in Plants Crucial for Virulence. *Nature*, 539(7630): 524–529.

Xu, You-Ping; Lv, Lin-Hui; Xu, Ya-Jing; Yang, Juan; Cao, Jia-Yi and Cai, Xin-Zhong (2018). Leaf Stage-Associated Resistance Is Correlated with Phytohormones in a Pathosystem-Dependent Manner. *Journal of Integrative Plant Biology*, 60(8): 703–722.

Xue, Yan; Zhong, Zhenhui; Harris, C. Jake; Gallego-Bartolomé, Javier; Wang, Ming; Picard, Colette; Cao, Xueshi; Hua, Shan; Kwok, Ivy; Feng, Suhua; Jami-Alahmadi, Yasaman; Sha, Jihui; Gardiner, Jason; Wohlschlegel, James and Jacobsen, Steven E. (2021). Arabidopsis MORC Proteins Function in the Efficient Establishment of RNA Directed DNA Methylation. *Nature Communications*, 12(1): 4292.

Yaari, Rafael; Katz, Aviva; Domb, Katherine; Harris, Keith D.; Zemach, Assaf and Ohad, Nir (2019). RdDM-Independent de Novo and Heterochromatin DNA Methylation by Plant CMT and DNMT3 Orthologs. *Nature Communications*, 10(1): 1613.

Yadav, Vikash Kumar; Singh, Swadha; Yadav, Amrita; Agarwal, Neha; Singh, Babita; Jalmi, Siddhi Kashinath; Yadav, Vrijesh Kumar; Tiwari, Vipin Kumar; Kumar, Verandra; Singh, Raghvendra and Sawant, Samir Vishwanath (2022). Stress Conditions Modulate the Chromatin Interactions Network in Arabidopsis. *Frontiers in Genetics*, 12. available at https://www.frontiersin.org/journals/genetics/articles/10.3389/fgene.2021.799805/full [12 September 2024].

Yan, Shunping; Wang, Wei; Marqués, Jorge; Mohan, Rajinikanth; Saleh, Abdelaty; Durrant, Wendy E.; Song, Junqi and Dong, Xinnian (2013). Salicylic Acid Activates DNA Damage Responses to Potentiate Plant Immunity. *Molecular Cell*, 52(4): 602–610.

Yang, Liping; Lang, Chenjing; Wu, Yanju; Meng, Dawei; Yang, Tianbo; Li, Danqi; Jin, Taicheng and Zhou, Xiaofu (2022). ROS1-Mediated Decrease in DNA Methylation and Increase in Expression of Defense Genes and Stress Response Genes in Arabidopsis Thaliana Due to Abiotic Stresses. *BMC Plant Biology*, 22: 104.

Yassin, Mustafa; Ton, Jurriaan; Rolfe, Stephen A; Valentine, Tracy A; Cromey, Matthew; Holden, Nicola and Newton, Adrian C (2021). The Rise, Fall and Resurrection of Chemical-Induced Resistance Agents. *Pest Management Science*, 77(9): 3900–3909.

Yildiz, Ipek; Gross, Marlene; Moser, Denise; Petzsch, Patrick; Köhrer, Karl and Zeier, Jürgen (2023). N-Hydroxypipecolic Acid Induces Systemic Acquired Resistance and Transcriptional Reprogramming via TGA Transcription Factors. *Plant, Cell & Environment*, 46(6): 1900–1920.

Yildiz, Ipek; Mantz, Melissa; Hartmann, Michael; Zeier, Tatyana; Kessel, Jana; Thurow, Corinna; Gatz, Christiane; Petzsch, Patrick; Köhrer, Karl and Zeier, Jürgen (2021). The Mobile SAR Signal N-Hydroxypipecolic Acid Induces NPR1-Dependent Transcriptional Reprogramming and Immune Priming. *Plant Physiology*, 186(3): 1679–1705.

Yoshida, Kentaro; Schuenemann, Verena J; Cano, Liliana M; Pais, Marina; Mishra, Bagdevi; Sharma, Rahul; Lanz, Chirsta; Martin, Frank N; Kamoun, Sophien; Krause, Johannes; Thines, Marco; Weigel, Detlef and Burbano, Hernán A (2013). The Rise and Fall of the Phytophthora Infestans Lineage That Triggered the Irish Potato Famine (D. Baulcombe, Ed.). *eLife*, 2: e00731.

Yu, Agnès; Lepère, Gersende; Jay, Florence; Wang, Jingyu; Bapaume, Laure; Wang, Yu; Abraham, Anne-Laure; Penterman, Jon; Fischer, Robert L.; Voinnet, Olivier and Navarro, Lionel (2013). Dynamics and Biological Relevance of DNA Demethylation in Arabidopsis Antibacterial Defense. *Proceedings of the National Academy of Sciences*, 110(6): 2389–2394.

Yu, Guangchuang (2020). Gene Ontology Semantic Similarity Analysis Using GOSemSim, 207–215, in: Kidder, B. L. (Ed.), *Stem Cell Transcriptional Networks: Methods and Protocols*. New York, NY: Springer US. available at https://doi.org/10.1007/978-1-0716-0301-7_11 [8 August 2024].

Yu, Guangchuang; Wang, Li-Gen; Han, Yanyan and He, Qing-Yu (2012). clusterProfiler: An R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A Journal of Integrative Biology*, 16(5): 284–287.

Yuan, Minhang; Jiang, Zeyu; Bi, Guozhi; Nomura, Kinya; Liu, Menghui; Wang, Yiping; Cai, Boying; Zhou, Jian-Min; He, Sheng Yang and Xin, Xiu-Fang (2021). Pattern-Recognition Receptors Are Required for NLR-Mediated Plant Immunity. *Nature*, 592(7852): 105–109.

Yuan, Yuxiang; Liu, Yujie; Han, Lu; Li, Yan and Qi, Yijun (2024). An RdDM-Independent Function of Pol V Transcripts in Gene Regulation and Plant Defence. *Nature Plants*, 10(10): 1562–1575.

Yun, Se-Hun; Khan, Irfan Ullah; Noh, Bosl and Noh, Yoo-Sun (2024). Genomic Overview of INA-Induced NPR1 Targeting and Transcriptional Cascades in Arabidopsis. *Nucleic Acids Research*: gkae019.

Zavaliev, Raul and Dong, Xinnian (2024). NPR1, a Key Immune Regulator for Plant Survival under Biotic and Abiotic Stresses. *Molecular Cell*, 84(1): 131–141.

Zemach, Assaf; Kim, M. Yvonne; Hsieh, Ping-Hung; Coleman-Derr, Devin; Eshed-Williams, Leor; Thao, Ka; Harmer, Stacey L. and Zilberman, Daniel (2013). The Arabidopsis Nucleosome Remodeler DDM1 Allows DNA Methyltransferases to Access H1-Containing Heterochromatin. *Cell*, 153(1): 193–205.

Zeng, Wenjie; Huang, Huan; Lin, Xueqiang; Zhu, Chen; Kosami, Ken-ichi; Huang, Chaofeng; Zhang, Huiming; Duan, Cheng-Guo; Zhu, Jian-Kang and Miki, Daisuke (2021). Roles of DEMETER in Regulating DNA Methylation in Vegetative Tissues and Pathogen Resistance. *Journal of Integrative Plant Biology*, 63(4): 691–706.

Zhan, Junpeng and Meyers, Blake C. (2023). Plant Small RNAs: Their Biogenesis, Regulatory Roles, and Functions. *Annual Review of Plant Biology*, 74(Volume 74, 2023): 21–51.

Zhang, Heng; Gong, Zhizhong and Zhu, Jian-Kang (2022). Active DNA Demethylation in Plants: 20 Years of Discovery and Beyond. *Journal of Integrative Plant Biology*, 64(12): 2217–2239.

Zhang, Qian; Liang, Zhe; Cui, Xuean; Ji, Changmian; Li, Yun; Zhang, Pingxian; Liu, Jingrong; Riaz, Adeel; Yao, Pu; Liu, Min; Wang, Yunpeng; Lu, Tiegang; Yu, Hao; Yang, Donglei; Zheng, Hongkun and Gu, Xiaofeng (2018). N6-Methyladenine DNA Methylation in Japonica and Indica Rice Genomes and Its Association with Gene Expression, Plant Development, and Stress Responses. *Molecular Plant*, 11(12): 1492–1508.

Zhang, Xinchao; Zhang, Yue; Wang, Chaofu and Wang, Xu (2023). TET (Ten-Eleven Translocation) Family Proteins: Structure, Biological Functions and Applications. *Signal Transduction and Targeted Therapy*, 8(1): 1–20.

Zhang, Yinwen; Jang, Hosung; Luo, Ziliang; Dong, Yinxin; Xu, Yangyang; Kantamneni, Yamini and Schmitz, Robert J. (2024). Dynamic Evolution of the Heterochromatin Sensing Histone Demethylase IBM1. *PLOS Genetics*, 20(7): e1011358.

Zhang, Yinwen; Jang, Hosung; Xiao, Rui; Kakoulidou, Ioanna; Piecyk, Robert S.; Johannes, Frank and Schmitz, Robert J. (2021). Heterochromatin Is a Quantitative Trait Associated with Spontaneous Epiallele Formation. *Nature Communications*, 12(1): 6958.

Zhang, Yue; Zhang, Qian; Yang, Xingyu; Gu, Xiaofeng; Chen, Jinming and Shi, Tao (2023). 6mA DNA Methylation on Genes in Plants Is Associated with Gene Complexity, Expression and Duplication. *Plants*, 12(10): 1949.

Zhang, Yuelin; Tessaro, Mark J.; Lassner, Michael and Li, Xin (2003). Knockout Analysis of Arabidopsis Transcription Factors TGA2, TGA5, and TGA6 Reveals Their Redundant and Essential Roles in Systemic Acquired Resistance. *The Plant Cell*, 15(11): 2647–2653.

Zhang, Yueying; Dong, Qianli; Wang, Zhen; Liu, Qinzhe; Yu, Haopeng; Sun, Wenqing; Cheema, Jitender; You, Qiancheng; Ding, Ling; Cao, Xiaofeng; He, Chuan; Ding, Yiliang and Zhang, Huakun (2024). A Fine-Scale Arabidopsis Chromatin Landscape Reveals Chromatin Conformation-Associated Transcriptional Dynamics. *Nature Communications*, 15(1): 3253.

Zhang, Zhilin; Hannan Parker, Adam; Symeonidi, Aikaterini; Wilkinson, Samuel W.; Stassen, Joost H. M.; Smith, Lisa M.; Ton, Jurriaan and Johannes, Frank (2024). Multi-Generational Biotic Stress Increases the Rate of Spontaneous Epimutations in a ROS1-Dependent Manner. : 2024.09.19.613880. available at https://www.biorxiv.org/content/10.1101/2024.09.19.613880v1 [16 October 2024].

Zheng, Zhi-Liang (2022). Cyclin-Dependent Kinases and CTD Phosphatases in Cell Cycle Transcriptional Control: Conservation across Eukaryotic Kingdoms and Uniqueness to Plants. *Cells*, 11(2): 279.

Zhong, Xuehua; Du, Jiamu; Hale, Christopher J.; Gallego-Bartolome, Javier; Feng, Suhua; Vashisht, Ajay A.; Chory, Joanne; Wohlschlegel, James A.; Patel, Dinshaw J. and Jacobsen, Steven E. (2014). Molecular Mechanism of Action of Plant DRM De Novo DNA Methyltransferases. *Cell*, 157(5): 1050.

Zhou, Jun-Ma; Trifa, Youssef; Silva, Herman; Pontier, Dominique; Lam, Eric; Shah, Jyoti and Klessig, Daniel F. (2000). NPR1 Differentially Interacts with Members of the TGA/OBF Family of Transcription Factors That Bind an Element of the PR-1 Gene Required for Induction by Salicylic Acid. *Molecular Plant-Microbe Interactions*®, 13(2): 191–202.

Zhou, Ming; Palanca, Ana Marie S. and Law, Julie A. (2018). Locus-Specific Control of the de Novo DNA Methylation Pathway in Arabidopsis by the CLASSY Family. *Nature Genetics*, 50(6): 865–873.

Ziller, Michael J.; Hansen, Kasper D.; Meissner, Alexander and Aryee, Martin J. (2015). Coverage Recommendations for Methylation Analysis by Whole Genome Bisulfite Sequencing. *Nature methods*, 12(3): 230–232.

Zimmerli, Laurent; Jakab, Gabor; Métraux, Jean-Pierre and Mauch-Mani, Brigitte (2000). Potentiation of Pathogen-Specific Defense Mechanisms in Arabidopsis by β-Aminobutyric Acid. *Proceedings of the National Academy of Sciences*, 97(23): 12920–12925.

Zubko, Elena; Gentry, Matt; Kunova, Andrea and Meyer, Peter (2012). De Novo DNA Methylation Activity of METHYLTRANSFERASE 1 (MET1) Partially Restores Body Methylation in Arabidopsis Thaliana. *The Plant Journal*, 71(6): 1029–1037.

Zuo, Jianru; Niu, Qi-Wen and Chua, Nam-Hai (2000). An Estrogen Receptor-Based Transactivator XVE Mediates Highly Inducible Gene Expression in Transgenic Plants. *The Plant Journal*, 24(2): 265–273.