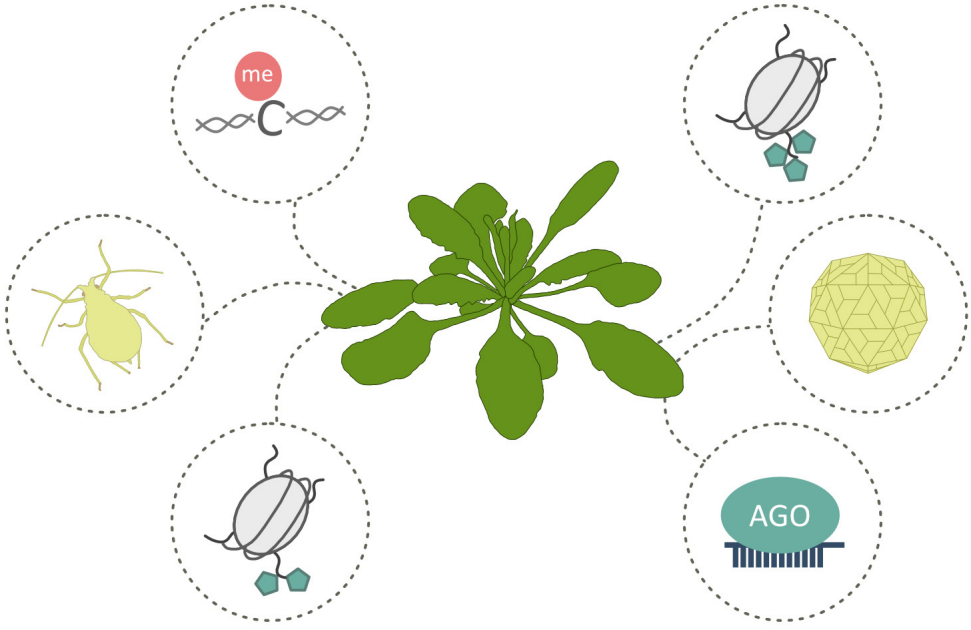
Acta Universitatis Agriculturae Sueciae

Doctoral Thesis No. 2021:58
Faculty of Natural Resources and Agricultural Sciences

Insights into biotic stress response in *Arabidopsis thaliana*: the roles of epigenetics and RNA silencing

María Luz Annacondia López



Insights into biotic stress response in *Arabidopsis thaliana*: the roles of epigenetics and RNA silencing

María Luz Annacondia López
Faculty of Natural Resources and Agricultural Sciences
Department of Plant Biology
Uppsala

SLU
SWEDISH UNIVERSITY OF AGRICULTURAL SCIENCES

DOCTORAL THESIS
Uppsala 2021

Acta Universitatis agriculturae Sueciae
2021:58

Cover: Graphical summary of the topics covered in this doctoral thesis: the roles of epigenetics and RNA silencing during aphid infestation and CMV infection (María Luz Annacondia).

ISSN 1652-6880
ISBN (print version) 978-91-7760-794-6
ISBN (electronic version) 978-91-7760-795-3

Uppsala
Print: SLU Service/Repro, Uppsala 2021

Insights into biotic stress response in *Arabidopsis thaliana*: the roles of epigenetics and RNA silencing

Abstract

Plants live in fluctuating environments that can lead to stressful conditions that threat their survival. To deal with these disturbances, plants have developed a fascinating genome plasticity that enables them to reprogram their gene expression and adapt to the new conditions. Among the different genome plasticity mechanisms, two are key regulators of the stress response: epigenetic regulation and RNA silencing. In this work we explored the roles of epigenetics and RNA silencing during two important biotic stresses, aphid (*Myzus persicae*) infestation and Cucumber mosaic virus (CMV) infection. We focused on the impact of changes on small RNAs (sRNAs), DNA methylation and repressive histone marks on gene expression and, hence, their role in the stress response. In both stresses, the sRNA populations are connected to the canonical role of the RNA-directed DNA methylation mechanism. Nevertheless, in the case of CMV, its interaction with the RNA silencing is more complex, as it involves the production of viral siRNAs (vsiRNAs) that interfere with the host ARGONAUTE (AGO) proteins and impact host gene expression. Additionally, we showed that DNA methylation plays an essential role in the regulation of the transcriptional response to both aphids and viruses, although it follows different trends. While aphid infestation leads to an important loss of DNA methylation, CMV causes an overall gain. Moreover, histone modifications also have essential functions during both stresses, highlighted by the enhance resistant of *kyp* to aphids and the reorganization of H3K9me2 and H3K27me3 observed during CMV infection, which are also connected to the changes on DNA methylation and the regulation of the transcriptional defense response. Altogether, this doctoral thesis provides key contributions to improve our knowledge about epigenetics, RNA silencing and their involvement in the regulation of the biotic stress response in *Arabidopsis thaliana*.

Keywords: Epigenetics, RNA silencing, gene expression, TEs, DNA methylation, histone modifications, sRNAs, AGOs, defense response

Author's address: María Luz Annacondia López, SLU, Department of Plant Biology, P.O. Box 7080, SE-750 07, Uppsala, Sweden

Insikter i biotisk stressrespons i Arabidopsis thaliana: funktionen av epigenetik och RNA silencing

Abstract

Växter lever i fluktuerande miljöer som kan leda till stressfulla förhållande som hotar deras överlevnad. För att hantera dessa störningar så har växter utvecklat en fascinerande genom-formbarhet som gör det möjligt att omprogrammera deras genuttryck och anpassa sig till dem nya förhållandena. Bland de olika mekanismerna som formar genomet, två är nyckelregulatorer av stressresponsen: epigenetisk reglering och RNA silencing. I det här arbetet utforskade vi rollen av epigenetik och RNA silencing under två viktiga biotiska stressorer, bladlus (*Myzus persicae)* infektion och Cucumber mosaic virus (CMV) infektion. Vi fokuserade på påverkan av förändringar på small RNAs (sRNAs), DNA metylering och undertryckande histonmärken på genuttryck och, därmed, deras roll i stressresponsen. Vid båda stressorer, sRNA:s populationer är kopplade till den kanoniska rollen för den RNA-riktade DNA-metylerings mekanismen. Ändå, i fallet med CMV, dess interaktion med RNA silencing är mer komplex, eftersom det involverar produktion av viral siRNA:er (vsiRNA) som interfererar med värdens ARGONAUTE (AGO) protein och påverkar värdens genuttryk. Dessutom, vi visar att DNA metylering spelar en viktig roll i regleringen transkriptionellt svar på både bladlöss och virus, även om det följer olika trender. Medan bladlöss infektionen leder till en viktig förlust av DNA metylering, CMV orsakar en total vinst. Dessutom, histon-modifikation har också en viktig roll för båda stressorer, markerad av den förbättrade resistansen till ***kyp*** av bladlöss och reorganisationen av H3K9me2 och H3K27me3 observerat under CMV infektion, vilket också är kopplat till förändringar i DNA metylering och reglering av det transkriptionella försvarets svar. Sammanlagt, den här doktorsavhandlingen ger nyckelbidrag till vår kunskap om epigenetik, RNA silencing och deras involvering i reglering av biotiska stress svar i *Arabidopsis thaliana*.

Nyckelord: Epigenetik, RNA silencing, genuttryck, Tes, DNA metylering, histonmärken, sRNAs, AGOs, försvarsrespons

Författarens adress: María Luz Annacondia López, SLU, Department of Plant Biology, P.O. Box 7080, SE-750 07, Uppsala, Sweden

Estudio de la respuesta a estrés biótico en *Arabidopsis thaliana*: el papel de la regulación epigenética y el silenciamiento génico por ARN

Resumen

Las plantas viven en ambientes cambiantes que pueden causar situaciones estresantes. Para lidiar con estas perturbaciones, han desarrollado una gran plasticidad génica que les permite reprogramar su transcripción rápidamente y adaptarse a las nuevas condiciones. Entre los distintos mecanismos de plasticidad génica, destacan la epigenética y el silenciamiento por ARN. En esta tesis, exploramos el papel de la epigenética y el silenciamiento por ARN durante dos estreses bióticos, la infestación de áfidos (*Mizus persicae*) y la infección por el virus mosaico del pepino (CMV). Estudiamos el impacto de los cambios en RNA pequeños (sRNAs), metilación del ADN y las marcas de histonas represivas en la expresión génica y, por tanto, su papel en la respuesta al estrés. En ambos estreses, la población de sRNAs está conectada a la función canónica de la metilación del ADN dirigida por ARN (RdDM). Sin embargo, en el caso de CMV, su interacción con el silenciamiento por ARN es más compleja, ya que implica la producción de sRNAs derivados del virus (vsRNAs) que interfieren con las proteínas ARGONAUTA (AGO) del huésped y afectan a su expresión génica. Además, demostramos que la metilación del ADN es esencial para la regulación de la respuesta transcriptómica contra áfidos y virus, aunque sigue tendencias distintas. Mientras que la infestación por áfidos causa una importante disminución en la metilación del ADN, CMV cause una ganancia global. Asimismo, las marcas de histonas también tienen un papel fundamental, destacado por la resistencia de los mutantes *kyp* a los áfidos y la reorganización de H3K9me2 y H3K27me3 durante la infección de CMV, la cual está relacionada con cambios en la metilación del ADN y la regulación de genes de defensa. En resumen, tesis doctoral aporta contribuciones clave para mejorar nuestros conocimientos sobre la epigenética, el silenciamiento por ARN y su función en la regulación de la respuesta a estrés biótico en *Arabidopsis thaliana*.

Palabras clave: Epigenética, silenciamiento génico por ARN, TEs, metilación del ADN, modificaciones de histonas, ARN pequeños, AGOs, respuesta de defensa

Dirección del autor: María Luz Annacondia López, SLU, Department of Plant Biology, Uppsala, Sweden

Dedication

Para mi familia.

Contents

List of publications

This thesis is based on the work contained in the following papers, referred to by Roman numerals in the text:

I. **Maria Luz Annacondia**, Dimitrije Markovic, Juan Luis Reig-Valiente, Vassilis Scaltsoyiannes, Corné M.J. Pieterse, Velemir Ninkovic, R. Keith Slotkin and German Martinez. 2021. Aphid feeding induces the relaxation of epigenetic control and the associated regulation of the defense response in Arabidopsis. *New Phytologist*, 230: 1185-1200.

II. **Maria Luz Annacondia** and German Martinez. 2021. Reprogramming of RNA silencing triggered by Cucumber mosaic virus infection in *Arabidopsis*. Manuscript in revision.

III. **Maria Luz Annacondia**, Juan Luis Reig-Valiente and German Martinez. 2021. Epigenomic landscape during Cucumber mosaic virus infection in *Arabidopsis thaliana*. Manuscript in preparation.

Paper I is reproduced with the permission of the publishers, following Wiley Journal's copyright guidelines. Paper II is reproduced as a pre-peer-reviewed submitted version. Paper III is reproduced as an author's version.

Additional publications produced during the course of this doctoral thesis but not included are listed below:

I. Cecilia Oliver, **Maria Luz Annacondia**, Zhenxing Wang, R. Keith Slotkin, Claudia Köhler, German Martinez. 2021. MicroRNA function transitions from regulating developmental genes to transposable elements during the maturation of pollen. Manuscript submitted.

II. Zahra Zangishei*, **Maria Luz Annacondia***, Alena Didriksen, Heidrun Gundlach, Rainer Schwacke, Björn Usadel, Julien Bruckmüller, Hooman Salari, Kirsten Krause and German Martinez. 2021. Parasitic plant sRNome analyses unveils parasite-specific signatures of miRNA retention, loss and gain. *Equal contribution. Manuscript submitted.

III. **Maria Luz Annacondia** and German Martinez. 2019. Chapter 12: Plant models of transgenerational epigenetic inheritance. Transgenerational Epigenetics (Second Edition), 236 - 282. Book chapter.

IV. **Maria Luz Annacondia**, Melissa H. Magerøy and German Martinez. 2018. Stress response regulation by epigenetic mechanisms: changing of the guards. *Physiologia Plantarum*, 162: 239-250. Review.

The contribution of María Luz Annacondia López to the papers included in this thesis was as follows:

I. Performed experiments, reviewed the manuscript.

II. Designed and performed experiments, generated high-throughput libraries, analyzed data, discussed results, and wrote the manuscript.

III. Designed and performed experiments, generated high-throughput libraries, analyzed data, discussed results, and wrote the manuscript.

List of tables

List of figures

Abbreviations

A	Arginine
ABA	Abscisic acid
ADK	*Adenosin kinase*
AFB	*AUXIN SIGNALING F-BOX* genes
AGO	ARGONAUTE
ARF	*AUXIN RESPONSE FACTOR*
ATX	ARABIDOPSIS TRITHORAX
ATXR	ARABIDOPSIS TRITHORAX RELATED proteins
BBSV	Beet black scorch virus
BCTV	Beet curly top virus
BER	Base-excision repair pathway
BIK1	*BOTRYTIS-INDUCED KINASE 1*
BRM	BRAHMA
BYV	Beet yellows virus
CaLCuV	Cabbage leaf curly virus
CaMV	Cauliflower mosaic virus
ChIP	Chromatin Immunoprecipitation
CHLI	*Chlorophyll biosynthetic* gene
CLF	CURLY LEAF
Clp-1	Ca^{2+}-dependent cysteine protease
CMV	Cucumber Mosaic Virus
CLSY1-4	CLASSY 1-4
crTMV	Crucifer-infecting tobacco mosaic virus
CymRSV	Cymbidum ringspot virus
DCL	DICER-LIKE
DDM1	DECREASED IN DNA METHYLATION 1

DEG	Differentially expressed genes
DMTases	DNA methyltransferases
dpi	Days post-infection
DRM2	DOMAINS REARRANGED METHYLTRANSFERASE 2
DRMs	Differentially methylated regions
dsRNA	Double-stranded RNA
eIF4A	*Translational initiation factor 4A*
EIN5	*ETHYLENE-INSENSITIVE 5*
ELF6	EARLY FLOWERING 6
EMF2	EMBRYONIC FLOWERING 2
ERF1	*ETHYLENE RESPONSIVE FACTOR 1*
ET	Ethylene
FACT	Facilitates Chromatin Transactions
FIS2	FERTILIZATION INDEPENDENT SEED 2
FLC	*FLOWERING LOCUS C*
FLD	*FLOWERING LOCUS D*
gbM	Gene body methylation
GO	Gene Ontology
GUS	β-glucuronidase
H3K9me	Methylation of lysine (K) 9 of histone H3
H3K9me2	Dimethylation of lysine (K) 9 of histone H3
H3K4	Methylation of lysine (K) 4 of histone H3
hc-siRNAs	Heterochromatic siRNAs
HC-Pro	Helper-component protease
HDA6	HISTONE DEACETYLASE 6
HDA19	HISTONE DEACETYLASE 19
HDP1-2	HARBINGER TRANSPOSON-DERIVED PROTEIN 1-2
HiC-15	C-15 hydroxylase
HKMTs	Histone lysine methyltransferases
HEN1	HUA ENHANCER 1
HESO1	HEN1 SUPPRESSOR 1
HSFB4	*HEAT SHOCK TRANSCRIPTION FACTOR B4*
HST	HASTY
HSVd	Hop stunt viroid
HYL1	HYPONASTIC LEAVES 1
IDM1-4	INCREASED DNA METHYLATION 1-4
IP	Immunoprecipitation

JA	Jasmonic acid
JmjC	Jumonji C
JMJ	JmjC domain-containing proteins
K	Lysine
KYP	Kryptonite
LAZ5	*LAZARUS 5*
LDL1	LSD1-LIKE 1
LSD1	Lysine-specific demethylase 1
lsiRNAs	Long small interfering RNAs
MEA	MEDEA
MET1	METHYLTRANSFERASE 1
miRNA	MicroRNA
MPK	*Mitogen activated protein kinases*
MSAP	Methylation-sensitive amplified fragment length polymorphism
MSI	MULTICOPY SUPPRESSOR OF IRA
NAT-sirnas	Natural antisense transcript siRNAs
NDR1	NON RACE-SPECIFIC DISEASE RESISTANCE 1
NERD	Needed for RDR2-independent DNA methylation
NBS-LRR	Nucleotide binding site and leucine-rich repeat domains
NRP1	*NON-EXPRESSOR OF PR GENE 1*
nt	Nucleotide
ORF	Open reading frame
PAD3	*PHYTOALEXIN DEFICIENT 3*
PAI	*PHOSPHORIBOSYLANTHRANILATE ISOMERASE*
PAMP	Pathogen-associated molecular pattern
PARE	Parallel Analysis of RNA Ends
PDF1.2	*PLANT DEFENSIN 1.2*
PepGMV	Pepper golden mosaic virus
phasiRNAs	Phased siRNAs
PKL	PICKLE
PolII	RNA polymerase II
PolIV	RNA polymerase IV
PolV	RNA polymerase V
PPDK	Pyruvate orthophosphate dikinase
PPR	*PENTATRICOPEPTIDE REPEAT*
PR1	*PATHOGENESIS RELATED 1*

PRXIIF	*Peroxiredoxin*
PTGS	Post-transcriptional gene silencing
PR	*PATHOGENESIS RELATED* genes
PRC2	Polycomb Repressive Complex 2
PSbMV	Pea seed-borne mosaic virus
PTGS	Post-transcriptional gene silencing
PVX	Potato virus X
PVY	Potato virus Y
R proteins	Resistance proteins
RAP2.4	*RELATED TO AP2.4*
rasiRNAs	Repeat-associated small interfering RNAs
RD	*RESPONSIVE TO DEHYDRATION*
RdDM	RNA-dependent DNA methylation
RAP	RNA-binding domain protein
REF6	RELATIVE OF EARLY FLOWERING 6
RISC	RNA-induced silencing complex
RLK	Receptor-like protein kinase
RLP43	*Receptor like protein 43*
RMV	Ribgrass mosaic virus
ROS	Reactive oxygen species
ROS1	*REPRESSOR OF SILENCING 1*
RP1	Pyruvate orthophosphate dikinase regulatory protein
RPS2	RESISTANT TO P. SYRINGAE 2
rRNA	Ribosomal RNA
RSV	Rice stripe virus
RT-qPCR	Reverse transcription quantitative polymerase chain reaction
RYSV	Rice yellow stunt virus
SA	Salicylic acid
SE	SERRATE
SEOR1	*SIEVE ELEMENT OCCLUSION-RELATED 1*
sgRNA	Subgenomic RNA
SGS3	SUPPRESSOR OF SILENCING 3
SHMV	Sunn-hemp mosaic virus
SPMMV	Sweet potato mild mottle virus
SSH1	SAWADEE-HOMEODOMAIN HOMOLOGUE 1
sidRNAs	Dicer-independent siRNAs
siRNA	Small interfering RNA

ssRNA	Single-stranded RNA
SUVH	SU(VAR)3-9 HOMOLOG
SWN	SWINGER
tasiRNAs	Trans-acting siRNAs
TE	Transposable element
TEV	Tobacco etch virus
TF	Transcription factor
TGS	Transcriptional gene silencing
TIR	Toll/interleukin 1
TIR1	*TRANSPORT INHIBITOR RESPONSE 1*
TGMV	Tomato golden mosaic virus
TMV	Tobacco Mosaic Virus
TRSV	Tobacco ringspot virus
TrxG	Trithorax group
TSS	Transcriptional start site
TTS	Transcriptional termination site
VRN2	VERNALIZATION 2
VIM1	VARIENT IN METHYLATION 1
vsiRNA	Virus-derived small interfering RNA
VSR	Viral Suppressor of RNA Silencing
WAK	*Cell wall-associated kinase*
WGBS	Whole genome bisulfite sequencing
Y-RNA	CMV Y-satellite RNA

1 Introduction

Plants are in a close relationship with their environment. Indeed, their ability to adapt to their surroundings affects their proper development and, ultimately, their survival. This is reflected on their different reproductive strategies, their variety of morphologies, their ability to develop specializations, or the different lifestyles that plants have evolved.

Additionally, plants live in a fluctuating environment that can present sudden adversities and lead to stressful scenarios **(section 1.1)**. To cope with these disturbances, plants have developed a fascinating genome transcriptional plasticity which allows them to quickly reprogram their transcription and adapt to new conditions. Genome transcriptional plasticity consists of four main different mechanism: alternative splicing, transcription factors (TFs), epigenetic regulation and RNA silencing (Baulcombe and Dean 2014; Khan et al. 2018; Laloum et al. 2018; Li and Wang 2019), being the latter two the focus of the PhD thesis **(sections 1.2 and 1.3, respectively)**.

1.1 Stress response in plants

Several authors have addressed the complex issue of defining plant stress. For instance, Larcher (1980) defined plant stress as "*changes in physiology that occur when species are exposed to extraordinary unfavorable conditions that need not represent a threat to life but will induce an alarm response*", while Lichtenthaler (1996) defined it as "*any unfavorable condition or substance that affects or blocks a plant's metabolism, growth or development*" (Gaspar et al. 2002; Kranner et al. 2010). In general, all definitions present a common concept, the deviation of the optimal status of the plant developmental program caused by an external factor. According to their nature, these external factors can be abiotic, which are constituted by nonliving conditions, or biotic, which are caused by living organisms. Abiotic

factors include temperature, light, water and nutrient availability, radiation, etc.; while biotic factors include herbivores, fungi, bacteria, viruses, and viroids (Gaspar et al. 2002).

In plants, the early signaling of the stress takes place by the alteration of the intracellular concentrations of Ca^{2+}, which is sensed by calcium-binding proteins. These proteins trigger a series of events in response to the stress, such as the activation of kinase cascades or the action of TFs that ultimately affect the transcription of stress-responsive genes. Furthermore, reactive oxygen species (ROS) act as a second stress messenger by, for instance, activating mitogen-activated protein kinase (MAPK) cascades (Verma et al. 2016; Aldon et al. 2018).

Although Ca^{2+} and ROS constitute the initial signaling of the stress, plant stress-response requires the action of several other components and pathways, where hormones play a key role. Among all the different plant hormones, salicylic acid (SA), jasmonic acid (JA), ethylene (ET) and abscisic acid (ABA) present the most important roles in stress response. Generally, SA, JA and ET are associated with the response to biotic stress, while ABA is mainly associated with abiotic stress (Verma et al. 2016).

Similar to Ca^{2+} and ROS, hormones modulate gene expression (Verma et al. 2016). For instance, increased levels of SA lead to the transcription of *PATHOGENESIS RELATED* (*PR*) genes, a group of genes that play a major role in biotic stress response. The transcription of these genes occurs via the activation of *NON-EXPRESSOR OF PR GENE 1* (*NPR1*), a master regulator of the defense signaling pathways that modulates the intersect between JA and SA/ET responses (Backer et al. 2019). In presence of a pathogen, SA turns *NPR1* into its active form, which moves to the nucleus where it interacts with members of the subclass TGA/OBF of basic domain/Leu zipper (bZIP) TF to facilitate the expression of *PR* genes (Spoel et al. 2003; van Loon et al. 2006). This interaction between a hormone and the expression of defense genes illustrates the complexity of the regulation of the response to stress in plants.

Ultimately, the objective of the stress response is to induce the transcription of beneficial genes but also to decrease the expression of genes that are not fundamental during the adverse conditions. This assumes that the resources, such as energy or nutrient availability, that a plant possesses are limited and, therefore, need to be utilized on the process that is the most beneficial for the plant's fitness. This phenomenon is known as the growth-defense tradeoff and consists of the allocation

of the limited resources into either growth or defense (Huot et al. 2014; Zust and Agrawal 2017), which is achieved by a dynamic genome regulation **(section 1.4)**.

The current thesis has explored two of these genome regulation mechanisms, epigenetic regulation and RNA silencing, in the context of two biotic stresses of high interest, both agricultural and scientifical, aphid infestation and Cucumber mosaic virus (CMV) infection.

1.1.1 Aphids (*Myzus persicae*)

Myzus persicae Sulzer (1776) (order Hemiptera, superfamily Aphididae), the green peach aphid or the peach-potato aphid, is a crop pest that is found globally on a wide range of hosts. It can infest over 400 species, including crops of economic importance, making it the most important aphid crop pest worldwide (Bass et al. 2014). Due to its economic interest, its ecology has been studied for decades making it one of the best analyzed herbivores (Kennedy et al. 1950; Mittler and Dadd 1966; Van Emden et al. 1969).

Although dependent on environmental conditions and geographic location (Blackman 1974; Kephalogianni et al. 2002), the life cycle of aphids in natural habitats starts in the spring with the parthenogenetic (asexual reproduction without the need of sperm) reproduction of female aphids in secondary hosts, like herbaceous plants, which were born from eggs on *Prunus* species (Bass et al. 2014; Özgökçe et al. 2018). During the fall, sexual females and males are developed and they produce eggs that experience diapause (a suspension on development) to survive the winter and become the new generation during the following spring (Braendle et al. 2006).

M. persicae is a polyphagous insect whose main source of food is the phloem sap of the host plant, which is rich in sucrose and few essential amino acids and, therefore, does not constitute a balanced source. Showing their ability of adaptation, aphids are able of stablishing symbiotic relations with bacteria capable of synthetizing the essential amino acids (Cao et al. 2018). Additionally, aphids have developed a number of adaptive traits to other aspects of their host plants, such as waxy surfaces, the presence of trichomes, nutrient composition or leaf thickness (Özgökçe et al. 2018).

The damage that *M. persicae* generates on the host plant is caused by the production of honeydew (a substance high in sugars), by their feeding strategy or the transmission of viruses. Interestingly, *M. persicae* is a vector for over 100 different plant viruses (Bass et al. 2014), including CMV, an insect-virus interaction that has

been widely explored for decades and remains of current interest due to the economic impact of both organisms (Mossop and Francki 1977; Chen and Francki 1990; Ng and Perry 1999; Krenz et al. 2015; Rhee et al. 2020; Tungadi et al. 2020).

Since *M. persicae* is both economically and ecologically important, its effects on the host at the molecular level have been studied in detail. In *Arabidopsis*, these studies were focused on aspects strictly related to the infestation, such as the production of defense metabolites or the expression of defense genes (De Vos et al. 2005; Kim and Jander 2007; Kusnierczyk et al. 2007; Kettles et al. 2013). Overall, these studies show that the defense against aphids considerably relies on the proper transcriptional reprogramming that leads to the production of dissuasive or detrimental compounds, such as glucosinolate or camalexin, as well as the induction of defense genes, such as *PR1* (De Vos et al. 2005; Kim and Jander 2007; Kusnierczyk et al. 2007; Kettles et al. 2013). Nevertheless, these studies were carried out using low resolution techniques like microarrays or reverse transcription quantitative polymerase chain reaction (RT-qPCR) (Kim and Jander 2007; Kusnierczyk et al. 2007; Kettles et al. 2013). Different analysis using RNA sequencing have revealed that aphid infestation induces a complex transcriptional response characterized by the elicitation of defense-related pathways, transcriptional factors and oxidative stress response as well as a complex transcriptional response from the aphid genome that includes aphid-transmitted mobile mRNAs that act as virulence factors in its host (Bansal et al. 2014; Chen et al. 2020; Duhlian et al. 2020; Zhang et al. 2020). Nonetheless, to better understand the interaction between hosts and *M. persicae*, there is a need of high-throughput sequencing studies of the global gene expression changes, as well as to find the potential causes of them, for instance, by exploring changes in epigenetic marks or the action of TFs.

1.1.2 Cucumber mosaic virus (CMV)

CMV was first described in 1926 as a disease discovered in cucurbits by two different authors, Doolittle in Michigan (Doolittle 1916), and Jagger in New York (Jagger 1916). Nevertheless, it was not purified until the 1960s (Scott 1963), and its genome organization and nucleotide sequence were not described until the 1990s (García-Arenal and Palukaitis 2008).

CMV has a worldwide distribution and infects more than 1200 species of monocots and dicots, such as crops, trees, and ornamental plants (Mochizuki and Ohki 2012), showing its tremendous economic and ecological importance. Moreover, it is transmitted by more than 80 different species of aphids, including *Aphis gossypii*

and *Myzus persicae* (García-Arenal and Palukaitis 2008), but also by parasitic plants of the genus *Cuscuta* (Mochizuki and Ohki 2012). Plants infected by CMV present symptoms that are not exclusive to CMV, such as chlorosis, stunting, mosaics or necrosis (García-Arenal and Palukaitis 2008).

CMV is the type member of the genus Cucumovirus, of the Bromoviridae family. Its genome is formed by three single-stranded RNAs (ssRNAs), RNA 1, RNA 2, and RNA 3, and two subgenomic RNAs, named RNA 4 and RNA 4A, which encode a total of five different proteins (Fig. 1). Proteins 1a, 2a and 3a are transcribed from RNA 1, RNA 2 and RNA 3, respectively, while proteins 2b and 3b come from RNA 4A and RNA 4, respectively (Mochizuki and Ohki 2012). The function of each of these proteins has been well characterized. Briefly, proteins 1a and 2a are replicases, containing a methyltransferase and helicase domains and an RNA-dependent RNA polymerase domain, respectively (Palukaitis and García-Arenal 2003). Proteins 3a and 3b participate in the movement of the virus, protein 3a is fundamental for intercellular movement, while protein 3b is necessary for both intercellular and long-distance movement (Canto et al. 1997), as well as aphid transmission (Perry et al. 1994). Finally, the 2b protein constitutes the viral suppressor of RNA silencing (VSR) (Mochizuki and Ohki 2012).

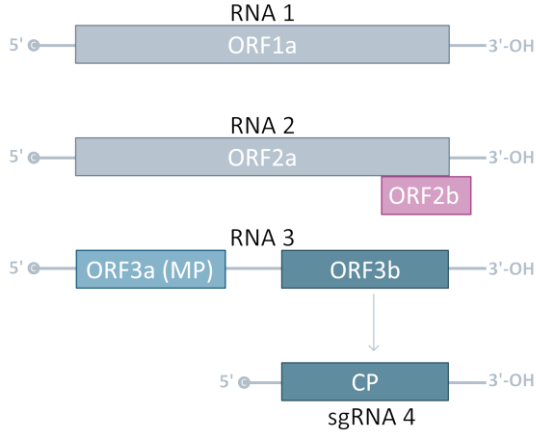


Figure 1. The genome of CMV is formed by three different RNAs (RNA 1, RNA 2 and RNA 3) encoding five different ORFs. MP: movement protein. CP: coat protein. sgRNA: subgenomic RNA. Modified from ViralZone, Expasy.

Apart from its economic and ecological importance, CMV has been widely studied as a model to understand the mechanism of action of VSRs, as its 2b protein is a potent VSR. This protein has been reported to interfere with different components

of the plant RNA silencing pathway, such as the ARGONAUTE (AGO) proteins, double-stranded RNAs (dsRNAs) and small RNAs (sRNAs) **(explained in section 1.3.2)**. Moreover, CMV has become a model system in plant virology and has been used to understand a wide range of topics, from plant-virus interactions (Jacquemond 2012) to RNA virus evolution (Roossinck 2001).

1.2 Epigenetics

Several definitions of the term "epigenetics" have been provided by different authors. In the early 40s, the geneticist and developmental biologist C. H. Waddington coined the term 'epigenetics' to refer to the genetic mechanisms which provide phenotypic differences (Waddington 1942; Waddington 1956). Since then, several definitions have been proposed, especially during the late 80s, when the term became widely used. R. Holliday proposed two different definitions (Holliday 1994). The first one claimed that epigenetics was "*the study of the changes in gene expression, which occur in organisms with differentiated cells, and the mitotic inheritance of given patterns of gene expression*"; the second one proposed that epigenetics was "*nuclear inheritance, which is not based on differences in DNA sequence*" (Holliday 1987). Later, Wu and Morris unified these two definitions and proposed that epigenetics is "*the study of changes in gene function that are mitotically and/or meiotically heritable and that do not entail change in DNA sequence*" (Wu and Morris 2001). In this PhD thesis, we consider epigenetics as the study of changes in gene expression that are not caused by alterations in the DNA sequence but caused by DNA methylation and histone modifications and are not necessarily inherited to the next generation.

These changes consist of reversible and covalent modifications which occur both on the DNA **(section 1.2.1)** and the histones **(section 1.2.2)**, without altering its sequence, and can potentially be inherited (Iwasaki and Paszkowski 2014). Epigenetic regulation plays a key role in the maintenance of genome stability by silencing transposable element (TE) expression, but also in other important biological processes, such as transcriptional gene regulation and genomic imprinting (Law and Jacobsen 2010; Iwasaki and Paszkowski 2014). Furthermore, in plants, epigenetic regulation is associated with the response to environmental cues, as it has been widely described in the case of vernalization (Baulcombe and Dean 2014), and the response to stress (Annacondia et al. 2018; Alonso et al. 2019; Chang et al. 2020).

1.2.1 DNA methylation

DNA methylation is the biochemical process by which a methyl group is taken from a S-adenosyl-L-methionine and added into the 5' position of a cytosine by the action of a DNA methyltransferase (DMTases) (Zhang et al. 2018). This was the first epigenetic mark to be described, and, curiously, it happened as the consequence of a failed experiment (Doerfler 2008). In 1978, researchers from Professor Walter Doerfler's group tried to digest the genome of the adenovirus type 12 (Ad12) by using the restriction enzyme HpaII and, instead of obtaining the expected small DNA fragments, the cleavage only occurred partially. Later, it was discovered that the methylation of the viral DNA prevented the action of HpaII, which is now known to be a methylation-sensitive restriction enzyme (Sutter et al. 1978; Sutter and Doerfler 1980; Doerfler 2008).

In plants, DNA methylation takes place in all possible cytosine contexts, namely CG, CHG and CHH (where H can be A, T or C), and overlaps with heterochromatin, the part of the genome enriched in repetitive sequences and TEs (Zhang et al. 2018). Indeed, the two first eukaryotic methylome analysis with single-base resolution, which interestingly were carried out in *Arabidopsis thaliana*, revealed that the distribution of DNA methylation is highly linked to the presence of TE sequences in the genome, although it can also be present in genes (Cokus et al. 2008; Lister et al. 2008).

DNA methylation is established *de novo*, by the RNA-dependent DNA methylation pathway (RdDM), and maintained through DNA replication, by the action of DMTases and the RdDM pathway (Zhang et al. 2018).

1.2.1.1 De novo methylation

In plants, *de novo* DNA methylation is established by the RdDM pathway, which mainly acts on TEs and repetitive sequences to guide methylation in all cytosine contexts (Matzke and Mosher 2014). Originally described in transgenic tobacco plants expressing Potato spindle tuber viroid (PSTVd) cDNAs and transgenic pea plants expressing the Pea seed-borne mosaic virus (PSbMV) replicase gene (Wassenegger et al. 1994; Jones et al. 1998), the RdDM pathway is based on the production of small interfering RNAs (siRNAs) by the RNA interference (RNAi) machinery and the action of two plant-specific RNA polymerases, PolIV and PolV (Matzke and Mosher 2014).

In *Arabidopsis thaliana*, the RdDM pathway starts with the recruitment of PolIV to its target sequences, mainly repeated sequences and TEs, to produce ssRNAs (Zhai

et al. 2015). In some loci, this recruitment is guided by SAWADEE-HOMEODOMAIN HOMOLOGUE 1 (SHH1), which binds to the methylated lysine 9 of histone H3 (H3K9) and the unmethylated lysine 4 of histone H3 (H3K4) and interacts with PolIV (Law et al. 2013). Moreover, the chromatin remodeler CLASSY 1 (CLSY1) also participates in this process (Copenhaver et al. 2011), probably facilitating the movement of PolIV throughout its genomic target (Matzke and Mosher 2014). Afterwards, the RNA-dependent RNA polymerase 2 (RDR2) copies the ssRNAs produced by PolIV into dsRNAs, which are then cleaved by DCL3 into 24-nucleotide (nt) siRNAs. These 24-nt are stabilized by the addition of a -OH group at their 3' end by the action of HUA ENHANCER 1 (HEN1) (Ji and Chen 2012) and loaded into AGO4 (Qi et al. 2006).

The AGO4-siRNAs complex interacts with the nascent PolV transcripts to recruit DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2) (Zhong et al. 2014). This enzyme catalyses *de novo* methylation at the homologous genomic sites at all methylation contexts (Quadrana and Colot 2016). Furthermore, the recruitment of DRM2 is facilitated by the presence of RNA-DIRECTED DNA METHYLATION 1 (RMD1), which interacts with both AGO4 and DRM2 and binds to methylated ssDNA (Zhong et al. 2014) (Fig. 2, upper panel). Generally, PolV is found at repetitive sequences and transposons that present considerable levels of DNA methylation and that give rise to 24-nt siRNAs (Wierzbicki et al. 2012; Zhong et al. 2012). The recruitment of PolV to these methylated sites takes place by means of SUVH2 and SUVH9 (Johnson et al. 2014; Liu et al. 2014b), two members of the histone methyltransferase family SU(VAR)3-9. These two proteins do not catalyze histone methylation (Johnson et al. 2008), but they are capable of binding methylated DNA to recruit PolV (Johnson et al. 2014; Liu et al. 2014b). Moreover, the DDR complex, which is formed by DEFECTIVE IN RNA-DIRECTED DNA METHYLATION 1 (DRD1), a CLSY1-related putative chromatin remodeler; DEFECTIVE IN MERISTEM SILENCING 3 (DMS3), a structural maintenance of chromosomes solo hinge protein; and RNA-DIRECTED DNA METHYLATION 1 (RDM1), a small plant-specific protein; interacts with SUVH2 and SUVH9 to promote the association of PolV with its target sequence (Law et al. 2010; Zhong et al. 2012). Nonetheless, PolV is also found, although at a much lower percentage, at target sites that do no present these characteristics and that correspond mainly to

genes, of which some present repetitive sequences within their coding sequence (Wierzbicki et al. 2012).

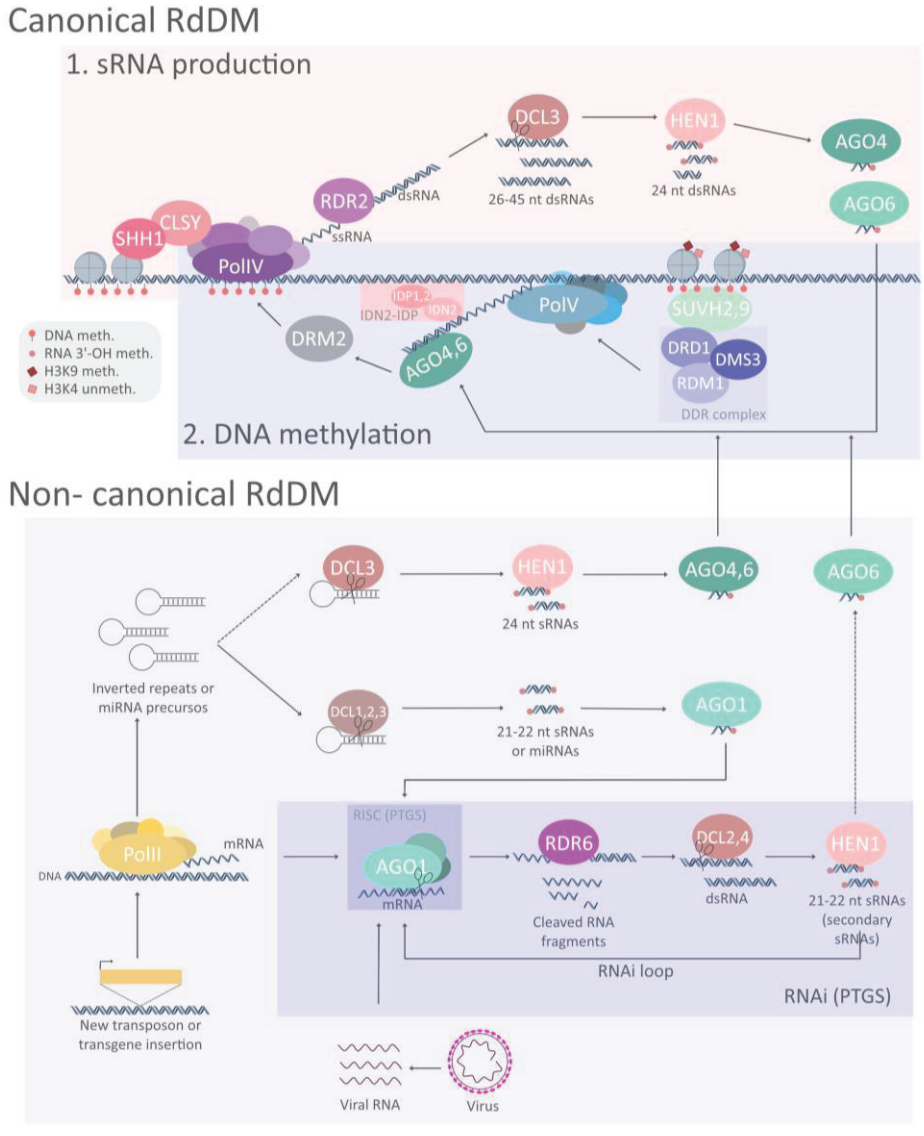


Figure 2. RdDM pathway in *Arabidopsis thaliana*. The upper panel represents the canonical pathway, the lower panel represents the non-canonical variation. Adapted from Erdmann and Picard, 2020.

Alternative variations of this canonical RdDM pathway have been described (Matzke and Mosher 2014). One of these alternative pathways is known as RDR6-dependent RdDM, which acts mainly on long and structurally autonomous TEs that are transcriptionally active (Nuthikattu et al. 2013; Cuerda-Gil and Slotkin 2016). In this case, RDR6 produces dsRNAs from PolII transcripts which are cleaved by DCL4 and DCL2 into 21- and 22-nt siRNAs, respectively. This 21-22-nt siRNAs are usually loaded into AGO1 to trigger post-transcriptional gene silencing (PTGS) of transposon mRNAs. In addition, these siRNAs can be loaded into AGO6 to interact with PolV nascent transcripts and recruit DRM2 (McCue et al. 2015; Cuerda-Gil and Slotkin 2016). AGO6-mediated methylation activates the canonical RdDM pathway (Matzke and Mosher 2014; Cuerda-Gil and Slotkin 2016) (Fig. 2, lower panel). Therefore, the aim of this non-canonical pathway is to recognize active transposons to trigger their silencing by the canonical RdDM (Nuthikattu et al. 2013; Cuerda-Gil and Slotkin 2016). Furthermore, there is another alternative to the canonical RdDM pathway, known as the RDR6-DCL3 pathway, that takes place when the number of reactivated TEs overcomes a certain threshold, saturating the activity of DCL2 and DCL4. In that situation, DCL3 processes RDR6-produced dsRNAs to produce 24-nt siRNAs that mediate *de novo* methylation and silence active TEs (Mari-Ordonez et al. 2013; Cuerda-Gil and Slotkin 2016).

Other non-canonical RdDM pathways are the PolII-DCL3 pathway, the PolV-Needed for RDR2-independent DNA methylation (NERD) pathway and the DCL-independent RdDM pathway. In the PolII-DCL3 pathway, transcripts generated by PolII are cleaved by DCL3, independently of RDR activity, into 24-nt sRNAs that can enter into the canonical RdDM. This alternative pathway probably occurs at transcripts that present imperfectly paired intramolecular dsRNAs or hairpins (Panda et al. 2016). In the PolIV-NERD pathway, NERD mediates the interaction of AGO2 with PolIV and PolV, which leads to the activity of the RdDM and the deposition of repressive histone marks on the target loci (Pontier et al. 2012). Finally, the dicer-independent pathway has been described in the triple mutant *dcl2 dcl3 dcl4* and the quadruple mutant *dcl1 dcl2 dcl3 dcl4*. In this case, non-cleaved dsRNAs, produced by PolIV-RDR2 or PolII-RDR6, are loaded into AGO4 and later processed at their 3' end by exosome-core complex nucleases to produce 21 to 24-nt siRNAs that participate in the silencing of their targets via canonical RdDM (Ye et al. 2016).

As the main targets of this pathway are both repeated sequences and TEs, RdDM plays an essential role on genome stability. This is important for TEs in purely heterochromatic regions, where RdDM activity is high on TE edges and for TEs found in euchromatic regions near genes, where Pol II transcriptional activity could

mediate the transcription of these TEs (Zemach et al. 2013; Sigman and Slotkin 2016). Therefore, the proper function of the RdDM prevents the activation of TEs, by maintaining their silencing, while allowing the transcription of the genes and, consequently, providing stability to the genome (Erdmann and Picard 2020).

Furthermore, the silencing of euchromatic TEs can also influence the transcription of nearby genes (Erdmann and Picard 2020) and, therefore, participate in the regulation of important biological processes, such as pathogen defense, stress responses, reproduction, development or genomic imprinting, among others (Matzke and Mosher 2014). One example that perfectly illustrates the importance of RdDM in development is the regulation of the gene *FLOWERING WAGENING* (*FWA)*. *FWA* encodes a homeodomain-containing transcription factor that participates in the regulation of flowering time (Soppe et al. 2000). The promoter of this gene contains two SINE-related direct repeats whose epigenetic regulation determines its transcriptional status (Soppe et al. 2000; Kinoshita et al. 2007). During normal development, *FWA* is exclusively expressed on the endosperm in an imprinted manner. In the rest of the plant tissues, the silencing of the promoter tandem repeats by the RdDM pathway leads to the transcriptional repression of the gene, which results in an appropriate flowering time. Nonetheless, if the methylation on the tandem repeats is removed, as it occurs in the mutants for *METHYLTRANSFERASE 1* (*MET1*) and *DECREASED DNA METHYLATION 1* (*DDM1*), *FWA* is transcribed ectopically, causing a late-flowering phenotype (Soppe et al. 2000; Kinoshita et al. 2007).

DNA methylation can also occur at gene bodies (termed gene body methylation, gbM). Usually, genes enriched in CG methylation have a housekeeping function, present an enrichment of CG methylation within their transcribed regions, and are accompanied by a depletion of the mark at both their transcriptional start site (TSS) and their transcriptional termination sites (TTS) (Bewick and Schmitz 2017). Recently, two studies have shed light into this issue and provided further knowledge on the role of gbM (Choi et al. 2020; Shahzad et al. 2021). These studies show that gbM participates, along with the deposition of H1, in the repression of antisense transcription (Choi et al. 2020), and that gbM is involved in the regulation of gene expression and that it is associated with the epigenetic inheritance of phenotypic traits related to the environmental responses (Shahzad et al. 2021).

1.2.1.2 Maintenance methylation

In plants, DNA methylation is maintained at all the three contexts throughout DNA replication by the action of several maintenance DMTases and the RdDM pathway.

In *Arabidopsis thaliana*, MET1 preserves the symmetrical CG methylation at both repeated sequences and genes (Finnegan and Dennis 1993; Finnegan et al. 1996; Kankel et al. 2003). MET1, with the participation of the accessory protein VARIENT IN METHYLATION 1 (VIM), recognizes the hemi-methylated CGs created during the replication process and conserves the methylation state (Kim et al. 2014).

On the other hand, the maintenance of the other symmetrical context, CHG, relies on the function of two plant specific DMTases CHROMOMETHYLASE 2 (CMT2) and CMT3, named after their chromodomain at their catalytic domain (Bartee et al. 2001; Stroud et al. 2013; Stroud et al. 2014). In this case, the recognition of the target takes place by a self-reinforcing loop between CHG and the dimethylation of H3K9 (H3K9me2). Through their chromodomain, CMT2 and CMT3 bind to H3K9me2, while the methyltransferases SUVH4/KRYPTONITE (KYP), SUVH5 and SUVH6, responsible for the H3K9 methylation, bind to the methylated CHG through their SRA domain (Du et al. 2015).

Methylation in the CHH context is maintained by CMT2 and DRM2. While CMT2 targets H1-containing heterochromatin, DRM2 maintains CHH methylation at RdDM-targeted regions (Zhang et al. 2018). The maintenance of DNA methylation at RdDM-targeted loci requires other important proteins, such as the chromatin-remodeling proteins DDM1 (Hirochika et al. 2000; Zemach et al. 2013), CLSY1-4 (Yang et al. 2018) or PICKLE (PKL) (Yang et al. 2017). DDM1 is an ATP-dependent chromatin remodeler which eases the access of the DMTases, especially CMT2, to the H1-containing heterochromatin (Zemach et al. 2013). Furthermore, *Arabidopsis ddm1* mutants experience genomic DNA hypomethylation and transposon reactivation, showing the essential role of DDM1 in maintenance of genome stability (Hirochika et al. 2000). Interestingly, the four CLSY (1-4) proteins act redundantly in the canonical RdDM pathway, but they are also important for DNA demethylation. Each individual CLSY protein participates in the demethylation of loci whose methylation depends in at least one of the other CLSY proteins. Therefore, mutation in one of these proteins leads to the hypomethylation of certain regions and the disruption of the balance between demethylation and methylation by the RdDM (Yang et al. 2018). Finally, PKL is a member of the Mi-2/CHD3 subfamily of ATP-dependent chromatin remodelers that potentially participates in the achievement of the correct chromatin environment to promote changes in DNA methylation. Specifically, this might happen through its nucleosome remodeling activity, which can both promote and repress DNA methylation. Nevertheless, the exact mechanism by which this occurs remains

unknown (Yang et al. 2017). Nonetheless, chromatin-remodeling proteins are not only associated with RdDM-targeted loci. For instance, BRAHMA (BRM) is a SNF2 chromatin-remodeling ATPase that controls the activity of the Polycomb Repressive Complex 2 (PRC2), avoiding aberrant deposition of H3K27me3 (Tang et al. 2008; Li et al. 2015a). Hence, the activity of BRM is crucial for the proper regulation of important developmental processes, such as seed maturation and flowering (Tang et al. 2008; Li et al. 2015a).

1.2.1.3 Demethylation

DNA methylation is removed by both passive and active mechanisms. The passive demethylation happens when this mark is not maintained or placed *de novo* and, therefore, it is gradually lost during replication. On the other hand, active demethylation occurs when it is removed by the action of DNA glycosylases. The genome of *Arabidopsis* encodes four different DNA glycosylases, ROS1, DEMETER (DME), DEMETER-LIKE 2 (DML2) and DML3 (Quadrana and Colot 2016). While DME acts exclusively during gametogenesis, ROS1, DML2 and DML3 are active in vegetative tissues and present redundant functions, albeit some locus specificity. Moreover, these three DNA glycosylases demethylate both transposons and genes located in or close to heterochromatin (Law and Jacobsen 2010).

These enzymes remove the methylated cytosine by means of their DNA glycosylase activity, which results in a single-nucleotide gap where an unmethylated cytosine is placed by the base-excision repair (BER) pathway (Liu and Lang 2020). To carry out their function, DNA demethylases need to be recruited to the target loci by the action of regulatory factors. For instance, the MBD7-IDM complex, constituted by methyl-CpG-binding domain protein 7 (MBD7), INCREASED DNA METHYLATION 1 (IDM1), IDM2, IDM3, HARBINGER TRANSPOSON-DERIVED PROTEIN (HDP1), and HDP2, recruits ROS1 to its target loci. This is achieved by generating the correct chromatin environment, which involves the acetylation of H3K18 and H3K23 by IDM1, which in turn is recruited by IDM2, IDM3, HDP1, HDP2, and the binding of MBD7 to methylated CG (Duan et al. 2017; Liu and Lang 2020). More specifically, in the central cell, the Facilitates Chromatin Transactions (FACT) protein complex assists on the recruitment of DME to the loci where it should act (Frost et al. 2018; Liu and Lang 2020).

In summary, plants have developed complex mechanisms of DNA methylation and demethylation that ensure genome stability by silencing TEs and accurately regulate the expression of genes.

1.2.2 Histone modifications

Histones are part of the fundamental unit of the chromatin, the nucleosomes, which consist of an octamer of histones (formed by two molecules of each of the four core histones, H2A, H2B, H3 and H4) wrapped around ~146 base pairs of DNA. The amino-terminal tails of these core histones undergo different post-transcriptional modifications: methylation, acetylation, phosphorylation, ubiquitination, sumoylation, glycosylation and ADP-ribosylation (Liu et al. 2010). These modifications are added, removed, and recognized by histone writers, erasers, and readers, respectively. Contrary to DNA methylation, which is mostly a repressive mark, histone modifications lead to both gene activation and repression (Zhao et al. 2019).

Histone acetylation and histone methylation have been widely studied, as they have been proved to be two fundamental marks for the regulation of gene expression. In general, histone acetylation acts as an active mark, as it increases DNA accessibility, while histone methylation constitutes both active and repressive marks (Ueda and Seki 2020).

1.2.2.1 Histone methylation

In plants, histone methylation is a fundamental regulatory mechanism of important developmental processes, as well as the maintenance of genome stability by the formation of heterochromatin. Moreover, histone methylation is one of the more complex histone modifications, as it can occur at two different residues, lysine (K) and arginine (R), and the number of added methyl groups can range from one to three. In *Arabidopsis*, histone lysine methyltransferases (HKMTs) are the writers of this mark and they act on different lysine residues of histone H3: Lys4 (K4), Lys9 (K9), Lys27 (K27) and Lys36 (K36) (Liu et al. 2010). Specifically, SET domain proteins are the putative writers of lysine methylation and, in plants, they are classified in four categories (Baumbusch et al. 2001): the SU(VAR)3-9 group, which includes SU(VAR)3-9 homologs (SUVH) and SU(VAR)3-9 related proteins (SUVR); the Enhancer of zeste [E(Z)] homologs; the trithorax (TRX) group, constituted by TRX-homologs and TRX-related proteins; and the absent, small, or homeotic discs 1 (ASH1) and ASH1-related proteins (ASHR).

Furthermore, this modification leads to both active and repressed transcriptional states, based on the methylated lysine and the level of methylation. Generally, the methylation of H3K4 and H3K36 are associated with active chromatin, while the methylation of H3K9 and H3K27 are involved in transcriptional silencing (Liu et al. 2010).

In *Arabidopsis*, H3K9 methylation mostly occurs as monomethylation (H3K9me1) or dimethylation (H3K9me2), while trimethylation (H3K9me3) happens at a very low level (Johnson et al. 2004). H3K9me1 and H3K9me2 are mainly found in chromocenters (Jasencakova et al. 2003), and H3K9me2 is also present in transposons and repeated sequences (Johnson et al. 2002; Jackson et al. 2004), which highlights the repressive nature of this mark.

The first plant histone H3K9 methyltransferase described was KYP, also named SUVH4, which was discovered in two independent genetic screens (Jackson et al. 2002; Malagnac et al. 2002). In one of these experiments, KYP was found to act as a suppressor of *stabilized clark kent* (*clk-st*), an epiallele of the gene *SUPERMAN* (*SUP*), which is involved in floral development (Jackson et al. 2002). In the other experiment, KYP was found to be a repressor of the gene *PHOSPHORIBOSYLANTHRANILATE ISOMERASE* (*PAI*), an intermediate enzyme in the tryptophan biosynthetic pathway (Malagnac et al. 2002). Interestingly, in both cases there was a reactivation of loci which were previously silenced by high levels of DNA methylation, which pointed to a potential function of H3K9 in DNA methylation-mediated silencing (Liu et al. 2010) **(discussed in section 1.2.3)**.

Another fundamental repressive histone mark is the methylation of H3K27. Like H3K9, H3K27 can also be mono-, di- or trimethylated (H3K27me1, H3K27me2 and H3K27me3, respectively) (Liu et al. 2010). While H3K27me1 is found at constitutive silenced heterochromatin (Fuchs et al. 2006), H3K27me3 and H3K27me2, the latter to a lesser extent, are located at euchromatin (Lindroth et al. 2004; Zhang et al. 2007), being H3K27me3 a crucial mark for development. Particularly, H3K27me3 is found preferentially at the 5' end of transcribed gene regions as well as upstream of their promoters (Zhang et al. 2007).

The trimethylation of H3K27 is established by the PRC2, which was first described in *Drosophila* (Cao et al. 2002; Czermin et al. 2002). The genome of *Arabidopsis* encodes all PRC2 components, including: three E(Z) homologs, CURLY LEAF (CLF), MEDEA (MEA) and SWINGER (SWN); three Su(z)12 homologs, FERTILIZATION INDEPENDENT SEED 2 (FIS2), EMBRYONIC FLOWER 2 (EMF2) and VERNALIZATION 2 (VRN2); five p55 homologs, MULTICOPY SUPPRESSOR OF IRA 1-5 (MSI1-5); and one homolog of Esc, FERTILIZATION-INDEPENDENT ENDOSPERM (FIE) (Kohler and Villar 2008).

In general, H3K27me3 is crucial for the proper development and growth of eukaryotic organisms. In *Arabidopsis*, *FLOWERING LOCUS C* (*FLC*) is a perfect example of a developmental gene that is regulated by H3K27me3 (among other

mechanisms). *FLC* encodes a MADS box transcriptional repressor that acts on genes related to the transition from vegetative to reproductive development and is epigenetically silenced by H3K27me3 as a result of exposure to cold. The regulation of this locus occurs in two different steps. First, long, antisense, non-coding RNAs (lncRNAs) are produced as a response to cold, preventing *FLC* transcription and, therefore, leading to its downregulation. Afterwards, PRC2 is recruited to establish H3K27me3 and completely silence its activity. Hence, this deposition of H3K27me3 reinforces the previously established downregulation, rather than initiating the silencing. When the temperatures start to increase, H3K27me3 is spread throughout *FLC*, to ensure a long-term epigenetically silenced state which can only be reset during reproductive development (Costa and Dean 2019).

On the contrary, H3K4 and H3K36 methylation are associated with an active chromatin state. In general, H3K4 is established by the action of Trithorax group (TrxG) proteins. The *Arabidopsis* genome contains five homologs, the ARABIDOPSIS TRITHORAX (ATX) proteins (Baumbusch et al. 2001), being ATX1 responsible for H3K4me3 and ATX2 responsible for H3K4me2 (Saleh et al. 2008), and seven ATX-related (ATXR) proteins, which present the overall structure of an ATX protein but they lack at least one of the C-terminal cysteines of the C-SAC motif (Baumbusch et al. 2001). This mark is found in genic regions, H3K4me1 and H3K4me2 are present in both active and inactive genes, while H3K4me3 is only seen on active genes. Specifically, H3K4me1 is deposited in gene body regions, closer to the 3' end, while H3K4me2 and H3K4me3 are found in promoters and the 5' end of coding regions (Zhang et al. 2009).

Arabidopsis has at least four ASH1 homologs and three ASH1-related proteins, which establish H3K36 methylation (Baumbusch et al. 2001). This mark occurs mostly as di- and trimethylation (H3K36me2 and H3K36me3, respectively) (Cheng et al. 2020). H3K36me2 is mainly found at the 3' end of transcriptionally active genes, whereas H3K36me3 is predominantly present at the 5' end (Roudier et al. 2011).

Histone methylation can be removed by the action of two different types of erasers, lysine-specific demethylase 1 (LSD1) and Jumonji C (JmjC) domain-containing proteins (Liu et al. 2010). In the genome of *Arabidopsis* there are four LSD1 homologs, FLOWERING LOCUS D (FLD), LSD1-LIKE 1 (LDL1), LDL2 and LDL3; and 21 JmjC domain-containing proteins (JMJ). In turn, the JMJ are divided into five subfamilies, based on their sequence similarities, the KDM5/JARID1

group; the KDM4/JHDM3 group; the KDM3/JHDM2 group; the JMJD6 group; and the JmjC domain-only group (Liu et al. 2010).

The removal of methylation from H3K9 is potentially carried out by the KDM3/JHDM2 group. INCREASE IN BONSAI METHYLATION 1 (IBM1)/JMJ35 is essential for the demethylation of H3K9me1 and H3K9me2, therefore, it prevents the spread of H3K9me2 (and CHG methylation) from TEs and repetitive sequences to close genes (Saze et al. 2008; Miura et al. 2009; Inagaki et al. 2010). Another important JMJ protein, known as JMJ27, participates in the demethylation of H3K9me1 and H3K9me2, playing a role in flowering time and defense (Dutta et al. 2017).

The KDM4/JHDM3 group is responsible for the removal of H3K27 methylation, being EARLY FLOWERING 6 (ELF6) (Crevillen et al. 2014) and RELATIVE OF EARLY FLOWERING 6 (REF6) (Lu et al. 2011) some of the better characterized H3K27 erasers. These two proteins act redundantly and are involved in the regulation of several developmental processes, for instance, flowering (Cheng et al. 2020). In this developmental process, REF6 acts as an FLC repressor, while ELF6 represses the photoperiod pathway (Noh et al. 2004). Moreover, ELF6 also participates in the epigenetic reprogramming of *FLC* to avoid transgenerational inheritance of the vernalized state (Crevillen et al. 2014).

The removal of the active histone marks H3K4 and H3K36 takes place by the action of LSD1 homologs and by JMJ30, respectively (Lee et al. 2018; Cheng et al. 2020). Particularly, three homologs of LSD1 have been found to be active in *Arabidopsis*, FLD, LDL1 and LDL2 (Jiang et al. 2007; Liu et al. 2007).

In summary, in plants, histone methylation is a complex epigenetic mark that is fundamental to maintain genome stability, for instance by repressing TEs, but also to regulate gene expression and, as a consequence, important developmental processes, such as flowering.

1.2.2.2 Histone acetylation

Histone acetylation is associated with permissive chromatin states. It regulates numerous important cellular processes, such as chromatin folding, nucleosome assembly, gene transcription, or the avoidance of heterochromatin spreading (Shahbazian and Grunstein 2007). The genome of *Arabidopsis* encodes 12 histone acetyltransferases (HATs) and 16 histones deacetylases (HDACs). Specifically, the group of HATs is constituted by five members of the GNAT/MYST superfamily (named in *Arabidopsis* as HAG), five members the CBP family (named in

Arabidopsis as HAC), and two members of the $TAF_{II}250$ family (named in *Arabidopsis* as *HAF*); while the group of HDACs is constituted by ten members of the RPD3/HDA1 superfamily (named in *Arabidopsis* as HAD), four members of the HD2 family (named in *Arabidopsis* as HDT), and two members of the SIR2 family (named in *Arabidopsis* as SRT) (Pandey et al. 2002).

In plants, histone acetylation and deacetylation play key roles in different developmental processes, such as seed, leaf and root development or flowering. Furthermore, it also participates in the responses to the environment, being crucial for the response to both biotic and abiotic stresses (Liu et al. 2016). As previously mentioned, the epigenetic regulation of *FLC* has been widely studied. Apart from H3K27me3, histone acetylation also participates in the regulation of this gene. Briefly, when flowering needs to be repressed and, therefore, *FCL* repressive activity is needed, the chromatin of this loci incorporates active histone marks, such as H3K36 and H3K4 methylation and histone acetylation (Berry and Dean 2015; Whittaker and Dean 2017). Furthermore, histone acetylation has also been associated with stress response, for instance, in the case of drought stress. Several important drought stress inducible genes, from the *RESPONSIVE TO DEHYDRATION* (*RD*) group and *RELATED TO AP2.4* (*RAP2*.4), experience an increase on histone acetylation, which leads to their transcriptional activation in response to the stress (Kim et al. 2008). Moreover, mutants for one of the main HDAs in *Arabidopsis*, *HISTONE DEACETYLASE 6* (*HDA6*), present enhanced tolerance towards this abiotic stress (Kim et al. 2017). HDA6 also plays an important role on the biotic stress response, for instance, during *Pseudomonas syringae* pv. tomato (*Pst*) infection in *Arabidopsis* (Wang et al. 2017). In this case, HDA6 regulates the expression of a considerable number of defense genes, including *PATHOGENESIS RELATED 1* (*PR1*) and *PR2*, as well as *WRKY* TFs. This was seen on a newly described *hda6* mutant allele (named *shi5*), where these defense genes were upregulated, providing the mutant with enhanced resistant towards *Pst* infection (Wang et al. 2017).

Overall, these examples illustrate the importance of histone acetylation as a mechanism to modulate plant development and stress response by regulating the expression of essential genes.

1.2.3 Interplay between DNA methylation and histone modifications

The different epigenetic mechanisms described above do not act independently, on the contrary, they are interconnected to constitute a complex regulatory network that

ensures the proper chromatin status. These interplays involve the different contexts of DNA methylation, different histone modifications, and histone variants.

One of the best studied interplays is the one formed by DNA methylation and H3K9 methylation. The methylation context CHG, maintained by CMT3, is recognized by the H3K9 methyltransferases KYP and SUVH6 by their N-terminal YDG/SRA domain. Once they recognize their target, KYP and SUVH6 methylate the adjacent histones via their SET domains. This interaction constitutes a feedback loop by which CHG and H3K9 methylation are maintained (Du et al. 2015). Moreover, SUVH2 and SUVH9, which also present a YDG/SRA domain, were proved to be necessary for the activity of the RdDM pathway. SUVH2 preferentially binds to CG, whereas SUVH9 binds to CHH (Johnson et al. 2008). In general, if H3K9 methylation is removed, non-CG methylation is also lost and, vice versa, when non-CG methylation is eliminated, H3K9 methylation levels decrease, highlighting the close relation between the two marks (Du et al. 2015). Interestingly, a recent study showed how H3K9me2 is also established by the action of sRNAs in *Arabidopsis* embryos (Parent et al. 2021). In this case, the methylation occurs independently of the action of SUVH4, SUVH5 and SUVH6, the H3K9 methyltransferases which have been described to be catalytically active, and partially independent of the action of DRM1 and DRM2. On the contrary, this mechanism relies on SUVH9, previously thought to be catalytically inactive, and potentially SUVH2, whose action is guided by sRNAs (Parent et al. 2021).

Recently, an interaction between DNA methylation and the recruitment of the H3K27 demethylase REF6 has been described (Qiu et al. 2019). In this study, it was shown that REF6 recognizes the CTCTGYTY motif (where Y can be either T or C), when these cytosines are unmethylated. Moreover, in the quadruple mutant *drm1 drm2 cmt2 cmt3* (*ddcc*), in which non-CG methylation is almost depleted, REF6 binds ectopically to short TEs located in euchromatic regions and, in some cases, this binding is accompanied by the transcriptional activation of the TEs or their nearby genes. This suggests that the preference of REF6 for non-methylated cytosines might contribute to a mechanism by which the activation of heterochromatin is avoided (Qiu et al. 2019).

Furthermore, DNA methylation is also associated with the role of the histone deacetylase HDA6 (Liu et al. 2012a). In the *hda6* mutant, its target loci do not only present changes in histone modification, but they also lose DNA methylation. Particularly, this DNA methylation loss happens in target loci that are not neighboring regions methylated by MET1. Furthermore, the upregulated loci found

in *hda6* greatly overlap with the upregulated loci found in *met1*, which suggests a functional connection between these two enzymes (To et al. 2011). Later on, HDA6 was shown to physically interact with MET1 (Liu et al. 2012b). Overall, this interaction regulates both the DNA methylation and histone acetylation and methylation of TEs to ensure their silencing (To et al. 2011; Liu et al. 2012b).

Additionally, active chromatin marks have also been associated with DNA methylation, specifically a relation between H3K4 demethylation and the activity of the RdDM pathway has been described (Greenberg et al. 2013). The mutation of the H3K4 demethylase *JMJ14* led to a decrease on DRM2-related DNA methylation, without affecting the MET1 and CMT3 pathways. Moreover, this decrease on DNA methylation levels was correlated with an increase on H3K4me2 and H3K4me3. Nevertheless, the observed DNA methylation loss was not as pronounced as in the *drm2* mutant, pointing to a redundant role of JMJ14 and other histone demethylases (Deleris et al. 2010). Furthermore, it has been shown that LDL1 and LDL2 also have an impact on DRM2-related DNA methylation, showing that histone demethylase potentially enhances the activity of the RdDM pathway in genes to balance the close activation by H3K4 methylation (Greenberg et al. 2013).

Histone variants have also been associated with DNA methylation, particularly the histone variant H3.3 and gbM (Wollmann et al. 2017) and the histone variant H2A.Z and DNA methylation (Zilberman et al. 2008). H3.3 is predominantly found close the transcription end sites (TES) of actively expressed genes, mainly related to development and responses to the environment. Interestingly, its profile is remarkably similar to that of gbM, which suggests an interaction between them (Wollmann et al. 2017). This interaction was confirmed on a *h3.3* knock-down mutant, in which the levels of gbM significantly decreased. Inversely, the presence of H2A.Z is negatively correlated with DNA methylation. Particularly, H2A.Z is not present in methylated TEs and methylated bodies of actively transcribed genes (Zilberman et al. 2008). Overall, these two studies show that DNA methylation participates on the deposition of different histone variants to ensure the proper chromatin structure, which ultimately leads to either transcription or gene silencing.

Altogether, these interplays show that the different histone marks do not act independently of DNA methylation and highlight the importance of considering them globally, as, ultimately, the status of the chromatin is defined by the balance of them.

1.3 RNA silencing

RNA silencing (known as RNA interference in animals) is a mechanism of genic regulation that constitutes one of the main protection tools against foreign genetic material, such as transgenes; deleterious sequences, such as TEs; or pathogens, such as viruses (Baulcombe 2002). In plants, it is fundamental for the maintenance of genome stability, regulation of developmental processes and the response to stress (Hohn and Vazquez 2011).

Even though RNA silencing emerged as a research field in the decade of 1990s (Eamens et al. 2008), one of the first articles associated with this phenomenon was published in 1928 (Wingard 1928). In that study, Wingard showed that the lower leaves of a plant infected by Tobacco ringspot virus (TRSV) showed symptoms, whereas the upper leaves remained healthy (Wingard 1928). At that time this phenomenon could not be explained, however, we know now that it was due to the action of the RNA silencing machinery (Baulcombe 2004). In plants, one of the first examples directly linked to the activity of RNA silencing was described in *Petunia hybrida* (Napoli et al. 1990; van der Krol et al. 1990). When the *chalcone synthase A* (*chsA*) gene, a key enzyme in anthocyanin biosynthesis, was introduced as a transgene in petunia plants, both the endogene and the transgene were silenced due to the overexpression of the transgene leading to the production of sRNAs. This led to a loss of anthocyanin pigments and, therefore, the flowers of these plants were white instead of purple (Napoli et al. 1990; van der Krol et al. 1990; Metzlaff et al. 1997).

RNA silencing encompasses different RNA-based mechanisms that rely on sequence-specific inhibition of gene expression, at the level of transcription, mRNA-stability, or translation, and share three common characteristics: (1) the production of dsRNAs; (2) the production of 20 to 26-nt sRNAs from the dsRNAs; and (3) the inhibition of the target RNA or DNA by the action of sRNAs (Brodersen and Voinnet 2006). These sRNAs can be classified as siRNAs or microRNAs (miRNAs), depending on their origin. siRNAs originate from a dsRNA produced by an RDR, but also from folded inverted-repeat sequences, hybridized unrelated RNA molecules, hybridized sense and antisense transcripts, while miRNAs originate from endogenous genes, known as *MIRNA* (*MIR*) genes (Borges and Martienssen 2015).

Moreover, two essential components are always present in all of the different RNA silencing mechanisms, the RNase III-type enzymes named Dicer, characterized by having dsRNA binding, RNA helicase, RNase III and PAZ (Piwi/Argonaute/Zwille) domains (termed Dicer-like in plants, DCL), and the AGO proteins, which have a

sRNA binding, PAZ and PIWI domains which provides the endonucleolytic activity (Brodersen and Voinnet 2006). Interestingly, the loading of both siRNAs and miRNA into AGO proteins is preferentially determined by their 5' terminal nucleotide, AGO1 prefers uridine, AGO2 and AGO4 adenosine and AGO5 cytosine (Mi et al. 2008).

The different pathways that mediate RNA silencing can be classified as transcriptional gene silencing (TGS) or post-transcriptional gene silencing (PTGS). TGS maintains the silencing of heterochromatic regions throughout the action of the RdDM pathway, and its variants, as previously explained (Hohn and Vazquez 2011; Rosa et al. 2018). On the other hand, PTGS occurs via miRNAs and siRNAs. Generally, PTGS via miRNAs results in the cleavage of the target mRNA, or its translation repression. In plants, PTGS is essential for both the silencing of foreign genetic material and for the regulation of important developmental processes (Borges and Martienssen 2015).

Plant miRNAs originate from specific stem regions of single-stranded hairpin precursors, encoded from *MIR* genes, and characterized by their unique secondary structure. Their biogenesis starts when PolII transcribes the *MIR* genes into pri-miRNAs, which are then folded into hairpin-like structures that contain a terminal-loop, an upper stem, the miRNA/miRNA* region, a lower stem and two arms (Wang et al. 2019). This structure is recognized and processed by DCL1, with the help of the double-stranded RNA-binding protein HYPONASTIC LEAVES 1 (HYL1) and the zinc-finger protein SERRATE (SE) to generate the miRNA/miRNA* duplex in a two-step process (Kurihara et al. 2006; Yang et al. 2006; Dong et al. 2008). First, DCL1 cleaves the pri-miRNA containing the stem loop, leaving it with a two nucleotide 3' overhang and releasing the 5' phosphate. Afterwards, the miRNA/miRNA* duplex is released (Grabowska et al. 2020).

Once produced, HEN1, a small RNA methyltransferase, stabilizes the duplex by methylating it at the 2'-OH position (Yu et al. 2005). In the absence of the activity of HEN1, HEN1 SUPPRESSOR 1 (HESO1) uridylates the 3' end of the duplex, leading to its degradation (Grabowska et al. 2020). Then, the selected strand from the miRNA/miRNA* duplex, named the guide strand, is loaded into AGO1 to form the RNA-induced silencing complex (RISC). Even though the loading was thought to happen in the cytoplasm, where the miRNA/miRNA* duplex is exported by HASTY (HST) (Park et al. 2005), recent studies points to a nuclear loading and a latter export (Bologna et al. 2018; Grabowska et al. 2020). Finally, the miRNA guides the RISC to the target RNA and triggers either its slicing or its translational

repression (Hohn and Vazquez 2011). In plants, the most common silencing mechanism is the cleavage of the target mRNA, which generates fragments with exposed hydroxyl and phosphate groups on the 5' and 3' ends and, therefore, causes its degradation (Grabowska et al. 2020). Nevertheless, in some cases the RISC can interfere with the activity of ribosomes, interrupting the translation of the target mRNA (Lanet et al. 2009; Grabowska et al. 2020) (Fig. 3A).

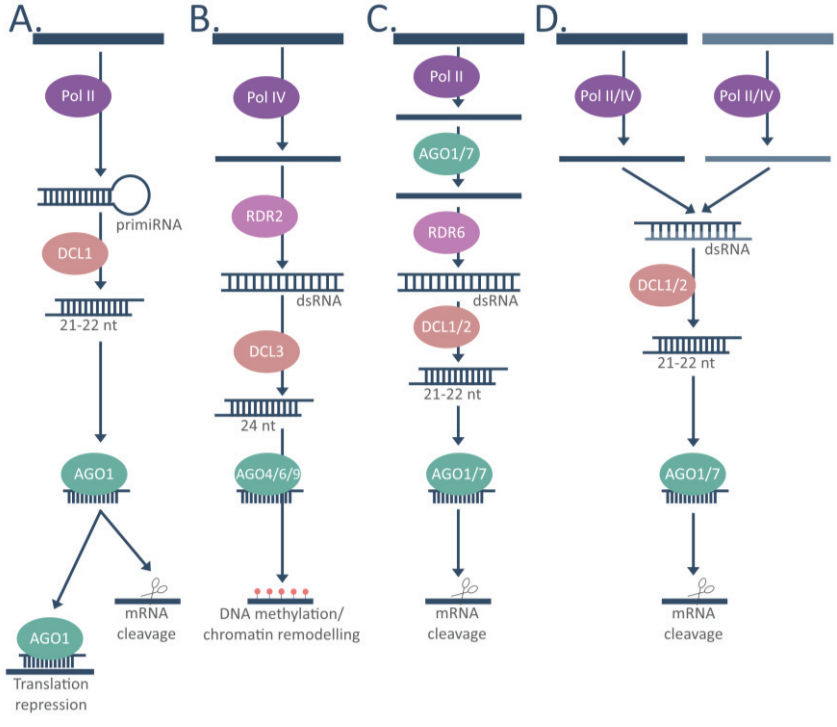


Figure 3. sRNAs biogenesis pathways. A. miRNA. B. hc-siRNAs. C. Secondary siRNAs. D. NAT-siRNAs. Adapted from Lionel Morgado. 2020.

siRNAs can be classified into three main categories, heterochromatic siRNAs (hc-siRNAs), secondary siRNAs and natural antisense transcript siRNAs (NAT-siRNAs) (Axtell 2013). hc-siRNAs are those siRNAs that are 23-24-nt long and are produced from intergenic or repetitive regions. Secondary siRNAs are those siRNAs that rely on an upstream sRNA trigger and RDR activity to be synthetized. In turn, they can be phased siRNAs (phasiRNAs), which present a uniformly defined terminus that leads to the production of siRNAs in phase series, or trans-acting siRNAs (tasiRNAs), which are those secondary siRNAs that can act on one or more targets that are not their origin locus. Finally, NAT-siRNAs originate from the

hybridization of independently and complementary transcribed RNAs and they can be cis-NAT-siRNAs or trans-NAT-siRNAs, depending on their origin. Cis-nat-siRNAs are originated from overlapping genes in opposite polarities, while trans-NAT-siRNAs are originated from RNAs that have complementarity and are not overlapping (Axtell 2013).

The biogenesis of hc-siRNAs is mainly carried out by the PolIV-RDR2-DCL3 pathway, although some of them are produced by PolV to generate RNA scaffolds that recruit DRM2 and lead to *de novo* methylation via the RdDM pathway (Morgado 2020) (Fig. 3B). Interestingly, the production of this type of siRNAs has also been associated with the methylation status during the defense response against pathogens (Katiyar-Agarwal and Jin 2010).

The production of phasiRNAs is triggered when a miRNA, usually of 22-nt length, cleaves either its target mRNA or a lncRNA (Liu et al. 2020). Then, RDR6, with the assistance of SUPPRESSOR OF SILENCING 3 (SGS3), generates a dsRNA that is cleaved by DCL2/4 in a reiterative way, producing the 21-nt phasiRNAs. These phasiRNAs are then loaded into AGO1 or AGO7 to promote the cleavage of the target RNA (Liu et al. 2020; Morgado 2020) (Fig. 3C). The loci that generate phasiRNAs, known as *PHAS* loci, are mainly protein-coding genes with a wide variety of roles, such as nucleotide-binding leucine-rich repeat (*NLR*), *PENTATRICOPEPTIDE REPEAT* (*PPR*), *AUXIN RESPONSE FACTOR* (*ARF*) genes, or *NAC* and *MYB* TFs (Liu et al. 2020). Moreover, tasiRNAs are produced by the same RNA silencing components, but they originate from *TAS* loci, and they act in *trans*, meaning they can target genes that are different to those they are produced from (Liu et al. 2020). Secondary siRNAs play key roles in different biological processes, including vegetative development, reproduction, seed germination or disease resistance (Liu et al. 2020).

NAT transcripts are those that share complementarity with other RNA transcripts to which they are not related to, and they can either be protein-coding loci or have a different nature. The cleavage of these transcripts can be carried out by DCL1, giving rise to 21-nt siRNAs which lead to PTGS, or by DCL3, producing 24-nt siRNAs which promote DNA methylation (Fig. 3D). Interestingly, it has been shown that nat-siRNAs accumulate in different types of stresses, both biotic and abiotic (Morgado 2020).

Altogether, the action of all these different types of sRNAs protects the genome from detrimental genetic material, as well as it regulates the expression of endogenous genes to ensure proper development and growth.

1.3.1 Antiviral role

Viruses are intracellular pathogens that depend entirely on the host machinery to complete their life cycle. To fulfill their replication, they need to produce replicative forms, which are recognized as a foreign genetic material by the plant RNA silencing machinery, triggering its antiviral activity (Li and Wang 2019). Therefore, viruses are potent inducers of the activity of the RNA silencing machinery.

The viral RNAs generated during their replication can produce dsRNAs with different origins that are recognized by the RNA silencing machinery: (1) the double-stranded viral replicative intermediate of RNA viruses; (2) highly structured hairpin regions of single-stranded viral RNA viruses or mRNAs of both RNA and DNA viruses; (3) different forms of single-stranded RNA viruses which are turned into dsRNA by the action of RDRs (Stavolone et al. 2020). From the six *Arabidopsis* RDRs, RDR1, RDR2, and RDR6 have been shown to have an antiviral role, through the production of secondary virus-derived siRNAs (vsiRNAs) (Xie et al. 2001; Schwach et al. 2005; Wang et al. 2010; Stavolone et al. 2020).

Following RDR-mediated dsRNA synthesis, DCL proteins process dsRNAs into vsiRNAs (Li and Wang 2019). The main antiviral DCL is DCL4, which produces 21-nt vsiRNAs from dsRNAs of RNA viruses (Deleris et al. 2006; Qu et al. 2008; Stavolone et al. 2020). Nonetheless, in the absence of this protein, DCL2 produces 22-nt vsiRNAs, which can trigger antiviral silencing but are not as efficient as the DCL4-produced 21-nt vsiRNAs (Deleris et al. 2006; Qin et al. 2017; Stavolone et al. 2020). Moreover, DCL3 can produce 24-nt vsiRNAs during infection by DNA viruses, such as geminiviruses (Akbergenov et al. 2006; Hanley-Bowdoin et al. 2013), or when both DCL2 and DCL4 are not functional (Blevins et al. 2006; Stavolone et al. 2020). Finally, in the triple mutant *dcl2 dcl3 dcl4*, DCL1 produces 21-nt vsiRNAs (Blevins et al. 2006). Altogether, these different results show the functional hierarchy of the plant DCL proteins.

Once produced, vsiRNAs, just as siRNAs and miRNAs, are methylated and stabilized by HEN1 before being loaded into AGO proteins to form the RISC (Li and Wang 2019). In plants, the main antiviral AGOs are AGO1 (Morel et al. 2002) and AGO2 (Harvey et al. 2011). Other AGOs, mainly AGO5, AGO7 and AGO10, also have secondary roles in the antiviral response (Qu et al. 2008; Brosseau and Moffett 2015; Garcia-Ruiz et al. 2015; Stavolone et al. 2020). Once loaded into the AGOs, the vsiRNAs guide the RISC towards the target viral RNA to trigger either its cleavage or its translational repression. Moreover, DNA viruses can also be

silenced by the addition of repressive epigenetic marks on their genome through the action of AGO4 and the RdDM pathway (Li and Wang 2019).

Interestingly, the activation of the RNA silencing machinery in one particular cell can spread to other cells to amplify the silencing. This is achieved by both cell-to-cell and long-distance movement, via plasmodesmata and the vasculature respectively, of the silencing signal, which is usually a small RNA molecule (Ryabov et al. 2004; Maizel et al. 2020; Stavolone et al. 2020). Even though the mechanism by which this spread of the silencing happens is still not well understood, it is clear that siRNAs and DCL proteins play key roles (Qin et al. 2017). Recently, Devers. et al showed that the mobile silencing signal produced from three different origins (a transgene, an endogenous gene, and a viral genome) was constituted by AGO-free primary siRNA duplexes. Moreover, they proposed that the mobile siRNA duplexes undergo a selective process caused by their loading into the AGOs that are present in the cell they are travelling through, which in turn would lead to qualitative different silencing patterns (Devers et al. 2020).

In summary, plants have evolved specialized antiviral roles for the RNA silencing machinery taking advantage of the viral need for using the plant transcriptional machinery to transcribe and replicate their RNA/DNA genomes.

1.3.2 Viral Suppressors of RNA silencing (VSRs)

To counteract the antiviral activity of the RNA silencing machinery, viruses have evolved proteins that diminish or suppress the RNA silencing activity against them, known as VSRs. These proteins interfere with the molecular mechanisms carried out by the RNA silencing machinery at both the TGS and PTGS levels (Li and Wang 2019; Stavolone et al. 2020) (see Table 1 for the classification of the viruses mentioned throughout the thesis).

The first two VSR identified in plants were the 2b protein of CMV and the helper-component protease (HC-Pro) of Potato virus Y (PVY). The VSR function of 2b and HC-Pro was first observed in lines carrying silenced transgenes, particularly the nitrate reductase and the β-glucuronidase (GUS). When these plants were infected with CMV or transformed with a construct overexpressing HC-Pro, the silencing of the transgenes was lost, showing that the viral supressors reduced the endogenous silencing of the transgene (Béclin et al. 1998; Mallory et al. 2001; H.S. and Ding 2002). Since then, an increasing number of VSRs has been identified and their mechanism of action have been extensively studied (Stavolone et al. 2020).

One of the main strategies of VSRs is the sequestration of different types of RNAs, mainly siRNAs to avoid their loading into the RISC or long dsRNAs to avoid DCL cleavage (Stavolone et al. 2020). Several VSRs follow this strategy, including the widely studied P19 of Tombusviruses, 2b of CMV or HC-Pro of PVY. The VSR of Tombusviruses, the P19 protein, forms homodimers that bind to 21-nt siRNAs in the standard dsRNA structure. This binding is enhanced by the presence of 5' phosphate group and it is independent of the 2-nt 3' overhang (Vargason et al. 2003; Ye et al. 2003). Turnip mosaic virus (TuMV) HC-Pro and Beet yellows virus (BYV) P21 also bind to 21-nt siRNAs duplex but, opposite to P19, they require a 2-nt 3' overhang (Lakatos et al. 2006). Similarly, some VSRs function against the production of secondary vsiRNAs, by interfering with the activity of RDR6 and/or SGS3 (Stavolone et al. 2020). This is the case of, for example, Rice yellow stunt virus (RYSV) protein P6, which was the first VSR proposed to block RDR6 activity (Guo et al. 2013).

Another strategy followed by several VSRs occurs at the sRNA methylation step of the RNA silencing pathway. This strategy is followed by the VSRs from Oilseed rape mosaic virus (ORMV), P125 (Blevins et al. 2006; Malpica-Lopez et al. 2018), Tobacco mosaic virus (TMV), P122 (Csorba et al. 2007), and TCV, P38 (Vogler et al. 2007), which bind sRNAs before they are methylated by HEN1, causing their partial degradation at their 2 nt overhangs, which ultimately avoids their loading into the RISC (Blevins et al. 2006; Csorba et al. 2007; Malpica-Lopez et al. 2018).

VSRs are also capable of interacting with AGO proteins, being AGO1 one of their main targets (Stavolone et al. 2020). CMV 2b, Tomato ringspot virus (ToRSV) coat protein (CP), potato virus X (PVX) P25, TCV P38, the phosphoproteins Ps of Alfalfa dwarf virus (ADV) and Lettuce necrotic yellows virus and (LNYV), the P1 protein of Sweet potato mild mottle virus (SPMMV), or the P0 of poleroviruses, can all target AGO1 (Zhang et al. 2006; Chiu et al. 2010; Karran and Sanfacon 2014; Zhuo et al. 2014; Bejerman et al. 2016; Mann et al. 2016; Iki et al. 2017; Kenesi et al. 2017). Additionally, 2b can interact with other AGOs like AGO4 and P25 can also interact with AGO2, showing that these interactions might not be limited to AGO1 (González et al. 2010; Hamera et al. 2012). In the case of 2b, the physical interaction between the VSR and an AGO protein, takes place through the PAZ and PIWI domains of the AGO proteins (Zhang et al. 2006) (González et al. 2010; Hamera et al. 2012). Nevertheless, it must be noted that while 2b binding to long dsRNA precursors is indispensable for suppressing PTGS *in vivo*, its interaction with AGOs is not (Duan et al. 2012), highlighting that VSR-AGO interactions might not be functional.

Additionally, VSRs, specially of DNA viruses, can also suppress the DNA methylation-related silencing (Stavolone et al. 2020). The VSR of the geminiviruses Beet curly top virus (BCTV) and Tomato golden mosaic virus (TGMV), named C2 and AL2 respectively, interact and inactivate the adenosine kinase (ADK), an enzyme with a critical function for the S-adenosyl methionine dependent methylation and the methyl-cycle maintenance. This interaction causes a decrease on the cytosine methylation levels and inactivates the antiviral silencing activity (Buchmann et al. 2009).

Some VSRs can also inhibit the activity of miRNAs, as their share characteristics with siRNAs and they also regulate important components of the RNA silencing machinery (Stavolone et al. 2020). For instance, miRNA162, miRNA168 and miRNA403 regulate the expression of DCL1, AGO1 and AGO2, respectively, and they have been shown to be sequestered by P19 and 2b, with different affinities, as a strategy to avoid silencing (Pertermann et al. 2018). Furthermore, in some cases, the methylation of the miRNA/miRNA* duplex is also altered by VSRs, as it has been seen for TuMV HC-Pro, BYV P21 and Tomato bushy stunt virus (TBSV) P19 (Yu et al. 2006).

Overall, all these examples show how viruses have developed diverse strategies to efficiently manipulate the different steps of the host RNA silencing to favor their expression. This close interaction, which generates the well-known co-evolutionary arms race between plants and viruses, has been extensively studied, showing, on one hand, how complex plant RNA silencing is and, on the other hand, how viruses are masters of manipulation. Nevertheless, there are still aspects that remain unknown and require further analysis, which perfectly depicts the complexity of plant-virus interactions.

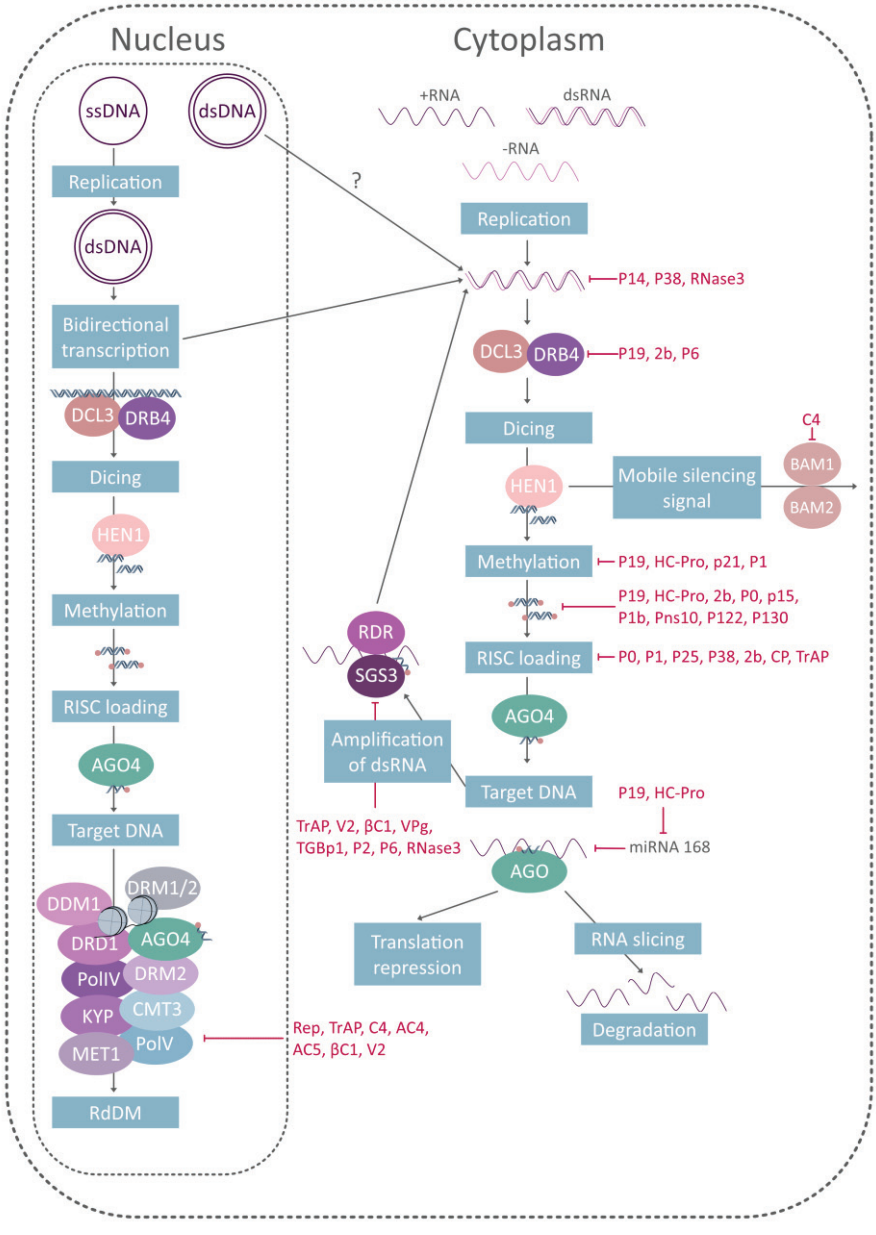


Figure 4. RNA silencing pathway and examples of VSRs which act at different steps of the pathway. Adapted from Li and Wang, 2019.

1.4 Roles of epigenetics and RNA silencing during biotic stress

As a result of their continuous exposure to pest, parasites and pathogens, plants have developed a complex immune system. One important component of the regulation of their immune system is its control by epigenetic modifications and the RNA silencing machinery (Huang et al. 2019). In this section, the roles of both epigenetics and RNA silencing are explained through the most relevant and illustrative studies published up to date.

1.4.1 Epigenetics

Epigenetics is characterized by being a fast, when considered in an evolutionary scale, and reversible molecular mechanism, allowing plants to cope with sudden new stresses as well as to overcome the recurrent ones (Ramirez-Prado et al. 2018; Alonso et al. 2019). Therefore, it is not surprising that both DNA methylation and histone modifications have been widely implicated in the response to different biotic stresses.

1.4.1.1 DNA methylation

Several studies have shown that DNA methylation and demethylation constitute essential regulatory mechanisms of the transcription of stress responsive genes (Lopez Sanchez et al. 2016; Halter et al. 2021), for instance, of NBS-LRR genes (Stokes et al. 2002; Kong et al. 2020). The genome of *Arabidopsis* encodes 149 NBS-LRR genes, which are a class of Toll/interleukin 1 (TIR) receptors that function during disease resistance and are located in heterochromatic clusters (Meyers et al. 2003). Some of these genes are regulated by epigenetic mechanisms and are affected directly or indirectly by both hyper and hypomethylation (Stokes et al. 2002; Lopez Sanchez et al. 2016), although other mechanisms have been proposed to play a role in the regulation of their transcription (Kong et al. 2020).

The role of DNA methylation has been widely studied during bacterial infections, particularly during *Pst* infection. One of the first evidence of the importance of DNA methylation during bacterial infections was the observation of decondensation of the chromocenters of infected plants, a change associated with a decrease on DNA methylation (Pavet et al. 2006). Later, the susceptibility of different RdDM mutants was tested, showing that while *ago4* (Agorio and Vera 2007) and *ros1* (Yu et al. 2013) are hypersensitive, *nrpe1* and *nrpd2* are more resistant to *Pst* infection (Lopez et al. 2011). Changes in DNA methylation under *Pst* infection, were confirmed by

performing whole genome bisulfite sequencing (WGBS), which showed that the infection caused both hypo- and hypermethylation events at the three methylation contexts (Dowen et al. 2012). Furthermore, the treatment with flg22 (a peptide from *Pst* flagellin that triggers the host immune response) causes the downregulation of key components of the RdDM pathway, including AGO4, AGO6 and different subunits of PolIV (Yu et al. 2013). The importance of DNA demethylation is exemplified in the regulation of the defense gene *RESISTANT METHYLATED GENE 1* (*RMG1*), which takes place through the regulation of two helitron-repeated sequences located in its promoter, that are targeted by ROS1 during infection (Yu et al. 2013) and regulated by the RdDM pathway (Halter et al. 2021). Altogether, these different works show that the regulation of DNA methylation during *Pst* infection constitutes part of a complex defense mechanism, that plays an important role in the defense response.

DNA methylation also plays a role in the response to other bacterial infection, for instance, *Agrobacterium tumefaciens*. In this case, DNA methylation inhibits the development of the tumors induced by the bacteria (Gohlke et al. 2013). The methylome of *A. tumefaciens*-induced tumors was characterized by an overall hypermethylation in the CHG and CHH contexts, which mainly affected protein-coding genes. In line with this, the triple mutant *dmr2 dmr3 cmt3* (*ddc*) and the *ago4* mutant, showed a greater tumor development (Gohlke et al. 2013).

DNA demethylation is also important in the response to the fungus. The triple mutant *ros1 dml2 dml3* (*rdd*) was hypersensitive to *Fusarium oxysporum* infection and showed repression of genes with putative or known stress-related functions. This increase on near-TE methylation could explain the decrease on the defense gene expression in the *rdd* plants and, therefore, their hypersensitivity (Le et al. 2014). *ros1* mutants are also hypersensitive to the infection by the obligate biotrophic oomycete *Hyaloperonospora arabidopsidis*, while demethylated mutants, including *nrpe1*, *drd1*, *cmt3* and *ddm1*, are more resistant to this pathogen (Lopez Sanchez et al. 2016). These opposite phenotypes are related to two important defense mechanisms, the induction of the SA-induced stress marker *PR1* and the deposition of callose, which are induced in *nrpe1* but inhibited in *ros1*. Moreover, almost half of the DEGs in *H. arabidopsidis* infected plants are affected by mutations in *ROS1* or *NRPE1*, showing that a considerable percentage of stress response genes are regulated by DNA methylation (Lopez Sanchez et al. 2016).

DNA methylation also plays a role in other types of biotic interactions, including nematodes, insects, viroids, and viruses. Regarding plant-parasitic nematodes, a

detailed study of the changes in the transcriptome, methylome and sRNA profile of *Arabidopsis* roots during its interaction with *Heterodera schachtii* was published recently (Hewezi et al. 2017). Overall, *H. schachtii* induced a dynamic demethylation of both genes and TEs, mainly in the CHG and CHH contexts, characterized by a lack of continuity between time points (5- and 10-days post-infection, dpi) but associated with the transcriptional reprogramming during infection (Hewezi et al. 2017).

In another interesting study, the changes on DNA methylation induced by the caterpillar *Pieris brassicae* on *Brassica rapa* were analyzed by performing methylation-sensitive amplified fragment length polymorphism (MSAP) (Doucet et al. 2016). Caterpillar feeding induced demethylation which was accompanied by a deterioration of important phenotypic traits for attractiveness towards pollinators, such as number of flowers, the volume of the inflorescences or the emission of volatile compounds. Interestingly, this study showed how, potentially, the demethylation caused by a herbivore can influence the interaction of the plant with its pollinators (Doucet et al. 2016).

Finally, changes in DNA methylation have been associated with the response to viral stress. This has been studied to greater detail for geminiviruses since their single-stranded DNA genomes replicate in the host nuclei and associate with cellular histones to create minichromosomes (Raja et al. 2008). As a defense mechanism, plants employ DNA methylation against the viral genome to restrict the propagation of the infection, as it effectively reduces the viral DNA replication, as well as viral gene expression (Zarreen and Chakraborty 2020). The importance of this epigenetic marks in the defense against these viruses was confirmed by infecting mutants in different epigenetic factors with two different geminiviruses, Cabbage leaf curly virus (CaLCuV) and BCTV, which also show stronger symptoms in methylation deficient and methyl cycle mutants (Raja et al. 2008; Rodriguez-Negrete et al. 2009). These evidences, reinforced the idea of a dual interaction of geminiviruses with the epigenetic machinery through: (1) the targeting of their genomes by repressive epigenetic marks by the endogenous host epigenetic pathways; and (2) the viral counterattack by means of their VRSs direct interaction with DNA methylation (Zarreen and Chakraborty 2020).

The defense response against RNA viruses and viroids also involves changes in DNA methylation. Viroids are plant-pathogenic, circular, single stranded noncoding RNAs (ncRNAs) characterized by their small size (240-400 nt). They fully depend on the molecular machinery of their host to fulfill their life cycle, as they do not code

for proteins (Castellano et al. 2015). In cucumber (*Cucumis sativus*) and *N. benthamiana*, the infection by Hop stunt viroid (HSVd) causes a dynamic misregulation of the DNA methylation levels on the 45S ribosomal RNA (rRNA) genes (Martinez et al. 2014; Castellano et al. 2015) and some selected TEs (Castellano et al. 2015). Similar to viroids, RNA viruses have also been associated with changes on the host DNA methylation levels. TMV-infected *N. benthamiana* plants show overall hypomethylation at the CG and CHG contexts localized to NBS-LRR genes conferring resistance to TMV, which were considerably hypomethylated in comparison to housekeeping components (Boyko et al. 2007). These results suggests that, during TMV infection, the changes on DNA methylation might serve as a mechanism to maintain genome stabilization, allowing the transcription and recombination of responsive genes, while preserving the status of housekeeping genes (Boyko et al. 2007). Changes in the methylome during CMV infection in *N. benthamiana* plants have been also studied (Wang et al. 2018). Interestingly, the changes in the methylome of the infected plants occurred mainly at the CHH context, and the differentially methylated regions (DMRs) associated with those changes were both hypo- and hypermethylated and located at gene bodies. Moreover, the presence of the DMRs correlated with changes in gene expression (predominantly related to defense response), with hypomethylated DMRs associated with up-regulated genes, and vice versa (Wang et al. 2018). Recently, the epigenetic changes induced by two different TuMV isolates with different degrees of adaptation to *Arabidopsis thaliana* were studied, showing that both isolates induced similar levels of cytosine methylation changes associated with TEs and stress-related genes (Correa et al. 2020).

In summary, all these studies reflect the global importance of DNA methylation during biotic stress response, but also how this mechanism of gene regulation acts differently depending on the specific stress. In general, changes on DNA methylation associated with the presence of a pathogen occur predominantly on the CHG and CHH contexts and are importantly related to the expression of responsive genes and TEs. Moreover, DNA methylation in plant defense is not only used to regulate their own responsive genes, but also to inhibit detrimental processes caused by the pathogen, for instance the formation of tumors induced by *A. tumefaciens*, and to even repress the expression of viral genomes, as it occurs during geminivirus infections, which highlights the versatility of this epigenetic mechanism. Nonetheless, only the most recent studies have started to explore the changes on DNA methylation during biotic stress at a genome-wide level and, very importantly, analyze how it affects gene expression. This is a crucial matter, as there is increasing

evidence showing that changes on DNA methylation do not always imply changes on gene expression (Zhong et al. 2021). Therefore, DNA methylation should be considered as one of the many potential mechanisms that regulate the transcriptional reprogramming that takes place during biotic stress response. In order to obtain a more accurate understanding of how this transcriptional reprogramming occurs, other mechanisms, for instance histone modifications or TF activity, should also be taken into consideration.

1.4.1.2 Histone modifications

Even though the role of histone modifications during biotic stress has been considerably less analyzed than the role of DNA methylation, these marks have also been associated with the regulation of the stress response (Ramirez-Prado et al. 2018). HISTONE DEACETYLASE 19 (HDA19), one of the best characterized HDACs, modulates the JA-dependent pathway, and exposure to JA or ET increases its expression (Zhou et al. 2005). Accordingly, *hda19* mutants show increased activity of the SA-dependent pathway and, therefore, higher expression of the SA-defense markers *PR1* and *PR2* which have higher levels of H3 acetylation on their promoters (Choi et al. 2012). In line with this regulatory role, *hda19* mutants showed an increased susceptibility to *Alternaria brassicicola* (Zhou et al. 2005), while the same mutants were more resistant to *Pst* infection, due to the increased basal expression of defense genes, such as *PR1* and *PR2* (Choi et al. 2012). This role does not seem to be shared with other HDAs, since HDA6 has been shown to participate in the regulation of defense genes through the JA pathway, as its expression is induced by this hormone. Hence, the loss of its activity leads to the upregulation of JA responsive genes, such as *PLANT DEFENSIN 1.2* (*PDF1.2*) or *ETHYLENE RESPONSIVE FACTOR 1* (*ERF1*). On the contrary, its expression is important for the downregulation of SA pathway defense genes, like *PR1*, *PR2* or *WRKY* TFs, showing an opposite role to HDA19 (Wu et al. 2008).

In addition to histone deacetylation, histone methylation also plays a role during defense response. For instance, mutants of the *Arabidopsis* H3K9me1/2 demethylase JMJ27 are hypersusceptible to *Pst*, as JMJ27 negatively regulates the TF *WRKY25*, which in turn inhibits the expression of *PR1* (Dutta et al. 2017). On the contrary, mutants for *LHP1-INTERACTING FACTOR 2* (*LIF2*), a ribonucleoprotein that interacts with PRC1, are more resistant to *Pst* infection but more susceptible to the necrotrophic fungus *Sclerotinia sclerotiorum*, as it implies a down-regulation of JA-related defense genes (Zhang et al. 2014). These results, similar to the ones from *hda19* mutants, indicate that H3K27me3 might negatively regulate SA-associated

response. Indeed, analysis of the genomic targets of LIF2 and LHP1 partially confirmed their involvement in the transcriptional regulation of stress-responsive genes (Molitor et al. 2016).

Active marks like the methylation of H3K4 and H3K36 have also been demonstrated to play roles during defense against different organisms. The Trithorax family member ATX1 is involved in the maintenance and the increase on *PR1* and *WRKY70* expression, acting as a node of convergence between the SA and the JA pathways. In line with the decrease of the SA-response, *atx1* mutants are more susceptible to the infection by *Pst* (Alvarez-Venegas et al. 2007). The H3K36 methyltransferase SDG8 is involved in the control of JA-related genes and the NBS-LRR gene *LAZARUS 5 (LAZ5)* and other stress-responsive genes (Berr et al. 2010) (Palma et al. 2010; Li et al. 2015b). Accordingly, *sdg8* mutants show reduced expression of *LAZ5* and an enhanced susceptibility to *A. brassicicola* and *Botrytis cinerea* (Berr et al. 2010). Similar to SDG8, another H3K36 methyltransferase, SDG25, controls the expression of defense genes, like *PR1*, *PDF1*.2 or *BIK1*, in response to both bacterial and fungal pathogens (Lee et al. 2016).

Histone modifications have also been associated with the response against viroids (Castellano et al. 2016). In cucumber plants infected with HSVd, the expression of the histone deacetylase HDA6 increases. This could constitute a potential host mechanism to counteract the interaction of HSVd with this protein (Castellano et al. 2016), which in turn could be causing the hypomethylation of certain loci that were previously described (Martinez et al. 2014). Moreover, overexpressing HDA6 lead to a decrease on the aforementioned hypomethylation (Castellano et al. 2016), showing that, indeed, the action of HDA6 is essential for the observed methylation changes during HSVd infection.

In addition to the changes on DNA methylation, the minichromosomes of geminiviruses also undergo histone modifications (Raja et al. 2008; Ceniceros-Ojeda et al. 2016; Rodriguez-Gandarilla et al. 2020). One of the epigenetic mechanisms that *Arabidopsis* employs to silence the genome of geminiviruses is the addition of H3K9 methylation, as it was seen during CaLCuV and BCTV infections (Raja et al. 2008). In line with this, *kyp2* mutants were hypersensitive to the infection by both geminiviruses (Raja et al. 2008). Nevertheless, their minichromosomes have also been associated with permissive histone marks, such as H3K4me3 (Raja et al. 2008; Ceniceros-Ojeda et al. 2016). Interestingly, in Pepper golden mosaic virus (PepGMV) infected pepper plants, the presence of this active histone mark and the repressive one H3K9 methylation was associated with symptomatic and recovered

plants, respectively (Ceniceros-Ojeda et al. 2016). In line with these results, in pepper plants superinfected with PepGMV, the viral genome showed high levels of H3K4me3 (Rodriguez-Gandarilla et al. 2020). Altogether, these studies show the importance of repressing the viral minichromosomes by the establishment of repressive histone marks, while avoiding the addition of active marks that enhance the transcription of the viral genes.

Overall, all these different examples illustrate the fundamental role of histone acetylation and methylation during biotic stress response. These studies have been particularly focused on their role as regulators of essential defense genes, such as *PR1/2*, *LAZ5*, *PDF1.2* and *WRKY* TFs, mainly through the regulation of the SA and JA pathways. In general, the effect of the histone marks on the defense response depends on which hormonal pathway they regulate and whether that pathway is beneficial or detrimental for the defense. Nevertheless, as these studies have been strongly based on analyzing the susceptibility of mutants, there is a lack of studies on the changes on their genome-wide profiles. This different approach has the potential to show how the stress affects the histone modifications of genes related to other fundamental processes that can be influenced by the stress, for instance development or photosynthesis. Furthermore, as with DNA methylation, the analysis of changes in histone marks should not be carried out isolated, and other mechanisms should be considered, especially when analyzing histone modifications that have been previously associated with other epigenetic marks, such as H3K9 and non-CG methylation. This approach will provide a more complete image of the changes on histone marks and their interplay with other regulatory mechanisms, which in turn will contribute to a better understanding of the overall status of the genome during biotic stress response.

In summary, all these different examples show the essential role of epigenetic regulation during biotic stress response, and how the different epigenetic mechanisms are modulated specifically to overcome each particular stress (Fig. 5). Nonetheless, the majority of these studied provided the analysis of one particular mechanism or mark, highlighting the need to take one step further and study their genome-wide profiles. Moreover, there are several interesting aspects that also require deeper understanding, such as how the specific enzymes are recruited to their target sites or the extent of the effects of epigenetic marks on gene expression. Addressing these questions is a difficult but fascinating task that would provide a much deeper understanding of the interplay between epigenetics and the biotic stress response.

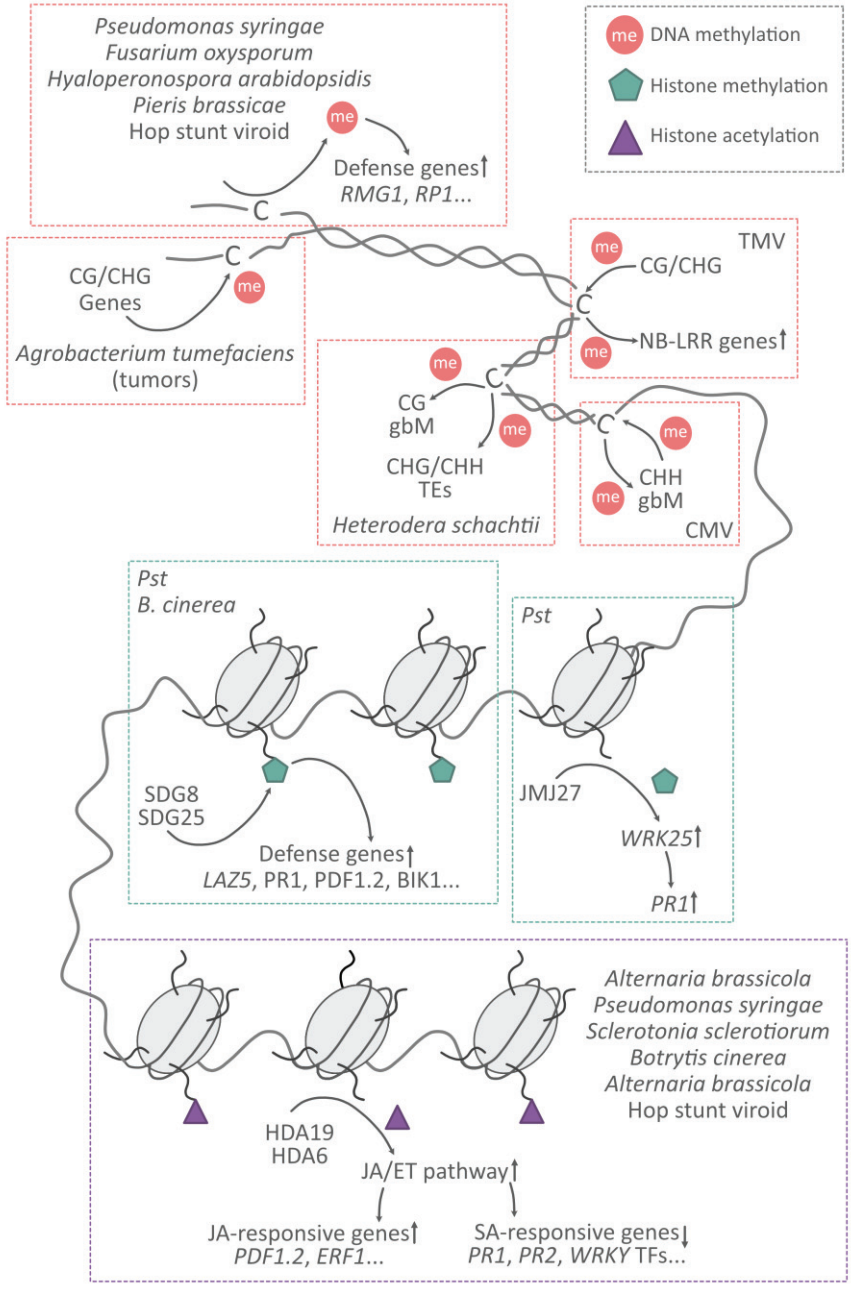


Figure 5. Summary of the main epigenetic changes associated with biotic stress response. Arrows point towards the chromatin represent gain of the marks, while arrows pointing from the chromatin represent loss of the marks.

1.4.2 RNA silencing

Over the past years, the role of the different sRNAs produced by the RNA silencing machinery during biotic interactions has been increasingly studied (Huang et al. 2019). Both endogenous (miRNAs, natsiRNAs, etc.) and pathogen-derived sRNAs change their accumulation dynamics during biotic stress and have different roles in the regulation of the response to stress.

1.4.2.1 miRNAs

miRNAs are well-known regulators of both development and the response to stress by modulating the expression of important genes performing each of these roles. miRNA accumulation and their regulatory role have been studied in multiple host-pathogen systems, showing an immense diversity of regulatory roles.

For example, in the *Pst*-infection context several miRNAs seem to play important roles. In 2006, the role of a plant miRNA in biotic defense was described for the first time (Navarro et al. 2006). In this study, the presence of flg22 induced the expression of *Arabidopsis* miR393, which negatively affects auxin signaling by the cleavage of auxin receptors mRNAs, specifically, the F-box proteins TIR1, AFB1 and AFB2, to provide resistance towards the bacteria (Navarro et al. 2006). Interestingly, miR393* also presents an antibacterial function, as it positively regulates defense by targeting *MEMB12* to promote *PR1* expression (Zhang et al. 2011b). miR393 is also induced in *Nicotiana tabacum* infected by an *A. tumefaciens* oncogenic strain which, together with miR167, represses the auxin signaling pathway enhancing tumor development (Pruss et al. 2008).

As a result of the extensive study of the *Arabidopsis*-*Pst* pathosystem, the defense roles of several other miRNAs have been described in this interaction. Among them, miR863-3p stands out because of its sequential roles: at early stages of infection, it targets two atypical receptor-like kinases (RLKs) which are negative regulators of the defense response, while at later infection stages, it targets SE, a positive regulator of the defense response that mediates the production of miRNAs (Niu et al. 2016). Other miRNAs that have been described to participate in the defense against *Pst* are: miR159, which is upregulated during later stages of the infection and downregulates the TFs MYB33 and MYB101 to promote the SA signaling pathway (Zhang et al. 2011a); miR160, which promotes callose deposition (Li et al. 2010); miR398 and miR773, which negatively regulates the deposition of callose (Li et al. 2010); miR167, which is upregulated during the infection and controls the auxin signaling pathway through the regulation of *ARF6* and *ARF8* to promote defense (Fahlgren et al. 2007; Zhang et al. 2011a); miR390, which is downregulated, as it enhances the

accumulation of ta-siRNAs that inhibits the expression of *ARF3* and *ARF4* (Zhang et al. 2011a); or miR472, which targets CC-NBS-LRRs (CNLs) resistance (R) proteins and, therefore, it negatively regulates defense responses (Boccara et al. 2014).

Furthermore, miRNAs also participate in the response to fungi. During treatment with *F. oxysporum*, *Arabidopsis* miR168 was the only miRNA that showed a constant upregulation among all of the induced ones, suggesting it might play an essential role during the fungal infection. This miRNA carries out a fundamental function in *Arabidopsis*, as it regulates the levels of AGO1. Therefore, the fungal-induced regulation of miR168 could potentially serve as a mechanism to ensure proper AGO1 levels during the infection (Baldrich et al. 2014). Furthermore, other miRNAs that participate in the defense response against fungi in *Arabidopsis* are: miR396, which is a negative regulator of the *GROWTH-REGULATING FACTOR* (*GRF*) family and represses defense responses (Soto-Suárez et al. 2017); miR773, which targets *METHYLTRANSFERASE 2* (*MET2*) and whose expression is detrimental for the defense response (Salvador-Guirao et al. 2018); miR858, which is a negative regulator of the defense, as it decreases the expression of flavonoid-specific *MYB* TFs, which have an antifungal function (Camargo-Ramírez et al. 2018); miR844, which regulates the expression of *CYTIDINEPHOSPHATE DIACYLGLYCEROL SYNTHASE 3*, an enzyme that plays a positive role during the stress (Lee et al. 2015); and miR400, which presents a negative defense function, as it regulates the expression of the defense genes *PENTATRICOPEPTIDE REPEAT PROTEIN* (*PRR*) (Park et al. 2014).

In *Arabidopsis*, the development of galls induced by the root-knot nematode *Meloidogyne javanica* relies on the expression of the regulatory module miR390/TAS3. miR390 regulates the expression of *TAS3* derived tasiRNAs, which, in turn, targets and degrades *ARFs* transcripts. When exposed to the nematode, *miR390* mutants developed a smaller number of galls, showing the importance of miR390 for *M. javanica* infection (Cabrera et al. 2016). Moreover, other miRNAs that are associated to the response against nematodes in *Arabidopsis* are: miR159, which is upregulated in the galls, and it has a negative effect for the defense, as it regulates the expression of the TF *MYB33*, necessary for the defense (Medina et al 2017); miR172, whose mature form is downregulated in the infected galls and roots and it is potentially linked to the auxin response during the stress (Hewezi et al. 2008; Cabrera et al. 2016; Jaubert-Possamai et al. 2019); miR319, which is downregulate in the infected galls and targets *MYB33* (Cabrera et al. 2016); miR396, which is downregulated in infected roots and acts as a key regulator of the reprogramming of

root cells (Hewezi et al. 2008; Hewezi et al. 2012); miR827, which is upregulated in the infected roots and acts as a negative regulator of the defense response (Hewezi et al. 2016); and miR858, which is upregulated during early stages of the infection and downregulated at later ones, and it regulates the expression of the TF *MYB83* (Piya et al. 2017).

miRNAs also play important roles in the defense against other biotic stresses, such as herbivores. Kettles et al. demonstrated that the reproduction of aphids is impaired in miRNAs mutants, including *dcl1*, *ago1*, *hen1*, *hst* and *se*. Interestingly, mutations in components of other sRNA pathways did not affect the production of progeny, showing that miRNAs are essential for aphid reproduction. In aphid-exposed *dcl1* mutants, the induction of two genes associated with camalexin biosynthesis led to an increased accumulation of camalexin, which provided the resistance phenotype of the mutant (Kettles et al. 2013). Moreover, several miRNAs act on the response against aphids (*Aphis gossypii*) in melon (*Cucumis melo*), including miR160 and miR167, that target *ARF2* and *ARF3*, respectively; miR393, that regulates *Transport Inhibitor Response 1* (*TIR1*); miR164, that targets NAC domain proteins; and miR395, that modulates the expression of ATP-sulfurylases (Sattar et al. 2012). Particularly, miR167 and miR393 negatively regulate the auxin response of the host plant to potentially promote defense, as auxin can interfere with the plant immunity (Sattar et al. 2016).

Together with the direct regulation of their genomes by the RNA silencing machinery, the defense response to viruses is also associated with miRNA activity. Especially, the activity of miRNAs regulating RNA silencing components, like AGO1 or AGO2, regulated by miR168 and miR403 respectively, seem to be important during the virus-plant interaction. Indeed, the enhanced accumulation of miR168 constitutes a general mechanism against viruses, as it has been described in several plant-virus systems: rice infected with Rice stripe virus (RSV) (Wu et al. 2015); *Arabidopsis* infected with CMV (Zhang et al. 2006), TMV (Csorba et al. 2007), TCV and Ribgrass mosaic virus (RMV) (Várallyay et al. 2010); *N. benthamiana* infected with Cymbidum ringspot virus (CymRSV), Crucifer-infecting tobacco mosaic virus (crTMV), Tobacco etch virus (TEV) and PVX (Várallyay et al. 2010); *Medicago truncatula* infected with Sunn-hemp mosaic virus (SHMV) (Várallyay et al. 2010); and *Solanum lycopersicum* infected with TMV and PVX (Várallyay et al. 2010).

The regulation of AGO1 mediated by miR168 seems to be involved in both cleavage and transcriptional repression, since in some pathosystems the levels of *AGO1*

mRNAs were also increased (Zhang et al. 2006; Csorba et al. 2007; Várallyay et al. 2010), while the AGO1 accumulation was either at mock levels or decreased (Várallyay et al. 2010). The increased accumulation of miR168 in CymRSV infected *N. benthamiana* plants is caused by the action of its VSR, the p19 protein, by a mechanism that has not been described yet (Várallyay et al. 2010). On the other hand, miR403 levels have been reported to decrease after *N. benthamiana* infection with Beet black scorch virus (BBSV) (Xu et al. 2016) and TMV (Diao et al. 2019), which in both cases increased the expression of *AGO2*.

Additionally, miRNAs regulate other components of the RNA silencing during viral infections. In rice, the expression of *RDR1* is indirectly regulated by a molecular cascade involving several MADS box proteins modulated by the monocot specific miRNA miR444 (Wang et al. 2016). Hence, miR444 has a positive regulatory role during RSV infection in rice (Wang et al. 2016).

In addition to miRNAs regulating RNA silencing components, other miRNAs with a different array of targets are affected by viral infections. CMV downregulates miR159 in *Arabidopsis*, which targets a group of regulatory genes known as GAMYB or GAMYB-*like*, R2R3 MYB domain TFs that are involved in GA signal transduction (Millar et al. 2019). Its downregulation leads to the upregulation of its target genes *MYB33* and *MYB65,* which is associated with the development of symptoms, without affecting the viral accumulation (Du et al. 2014). TMV infection in *N. tabacum* affects the levels of miR160, miR165/6, miR415, miR397, and miR535, all of which were previously associated with the response to different stresses. At an early stage of infection (5 dpi), when the virus is not detectable yet, these miRNAs were downregulated, while at later stages (15 and 22 dpi), they were upregulated, indicating that plants modulate the expression of miRNAs even at early times of viral accumulation (Bazzini et al. 2011).

Altogether, these different studies highlight the importance of the action of miRNAs during the defense response towards different biotic stresses, from herbivores to viruses (Fig. 6A). They also show the broad range of defense mechanisms that these miRNAs are modulating, such as the expression of defense-related TFs, the RNA silencing machinery or defense genes. Nevertheless, some of these studies are focused on particular miRNAs, or lack an in-depth analysis of the incorporation of these miRNAs into AGO proteins, which will determine their activity. Detailed analyses of the population of sRNAs loaded in different AGO proteins during stress would provide a complete scenario of the reprogramming of the miRNA population

during biotic stress response and, ultimately, would deepen our understanding of the molecular basis of these fascinating plant-pathogen interactions.

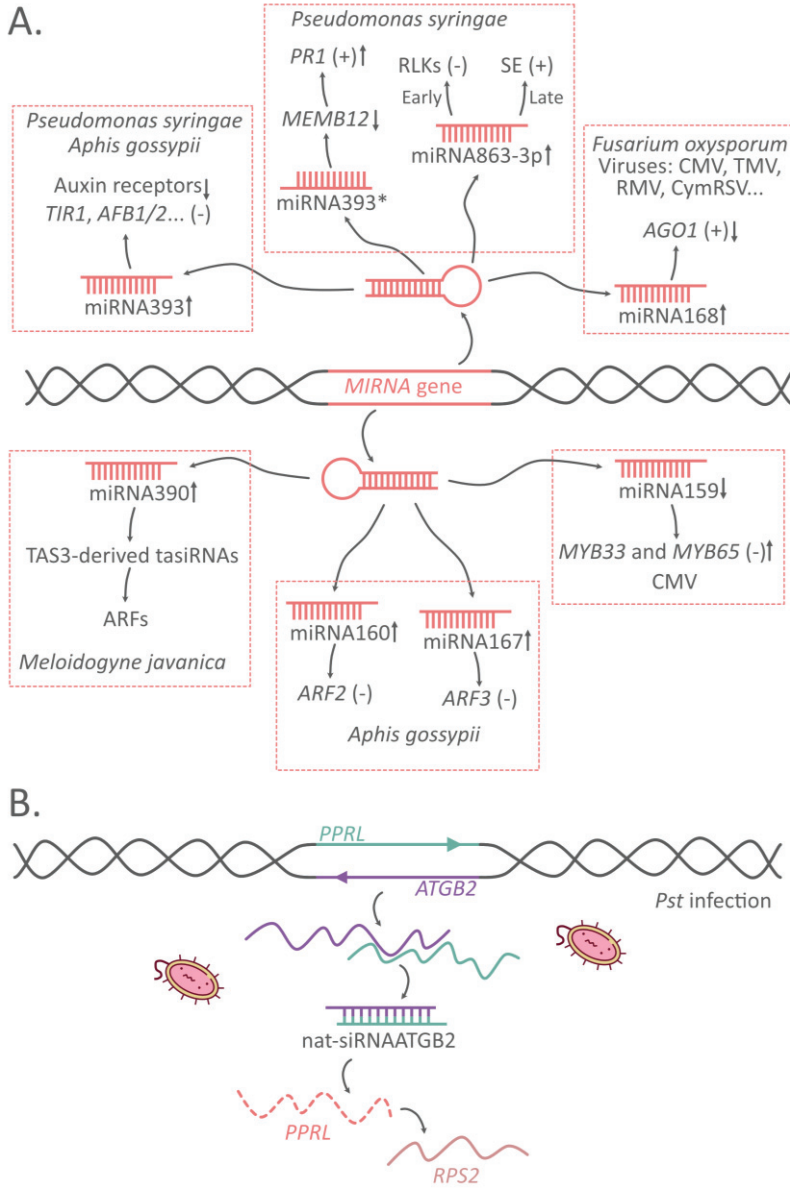


Figure 6. A. Summary of the main miRNAs involved in biotic stress response. (+) represents a positive role of the miRNA for the defense response, while (-) represents a negative role. B. Role of nat-siRNAATGB2 during *Pseudomonas syringae* infection.

1.4.2.2 Endogenous siRNAs

miRNAs are not the only sRNA class involved in the biotic stress response. Other classes of endogenous sRNAs have been shown to be active and relevant during infection by diverse pathogens, including natsiRNAs, phasiRNAs and other types of endogenous siRNAs.

The first siRNA associated with plant immunity was nat-siRNAATGB2 (Katiyar-Agarwal et al. 2006) which is promoted by *Pst* infection and produced from the overlap region of an anti-sense transcript pair constituted by a Rab2-like small GTP-binding protein gene (*ATGB2*) and a pentatricopeptide repeats (PPR) protein-like gene (*PPRL*). The production of nat-siRNAATGB2 depends on the NBS-LRR R proteins RESISTANT TO P. SYRINGAE 2 (RPS2) and NON-RACE-SPECIFIC DISEASE RESISTANCE 1 (NDR1), which are required for a series of resistance response, such as transcriptional reprogramming, production of ROS and induction of hypersensitive responses (HR). Once produced, nat-siRNAATGB2 down-regulates PPR, a potential negative regulator of *RPS2* (Fig. 6B). Therefore, nat-siRNAATGB2 promotes the defense response by down-regulating *PPRL* to allow the expression of *RPS2* (Katiyar-Agarwal et al. 2006).

Recently, the role of phasiRNAs on plant defense has been increasingly demonstrated, as they target key defense genes such as PPR or NBS-LRR genes (Liu et al. 2020). For example, *B. cinerea* infected *Arabidopsis* produces two tasiRNAs, *TAS1c*-siR483 and *TAS2*-siR453, that are transported to the fungi via exosome-like vesicles to silence pathogenic genes (Cai et al. 2018) (Fig. 7A). Moreover, the production of phasiRNAs is also favorable against the infection by the oomycete *Phytophtora capsici* (Hou et al. 2019). In this case, miR161 triggers the production of *PPR*-derived phasiRNAs, which are transported into the haustorium to target *P. capsica* genes, providing enhanced resistance to this pathogen (Fig. 7B) (Hou et al. 2019). The importance of miRNA-triggered phasiRNA production for plant defense has been described in several species, for instance: in Norway spruce (*Picea abies*), the miR482/2118 family target NBS-LRR genes to trigger the production of phasiRNAs (Xia et al. 2015); in legumes, miR1507 and miR1510 trigger the production of phasiRNAs that target R genes (Zhai et al. 2011; Fei et al. 2015; Zhao et al. 2015); in wheat (*Triticum aestivum*) and barley (*Hordeum vulgare*), miR9863 triggers the production of phasiRNAs that regulate the *MLA* genes, which are NBS-LRR genes (Liu et al. 2014a); and in *Solanaceae*, miR6019 and miR6027 trigger the production of phasiRNAs associated with R genes (Li et al. 2012).

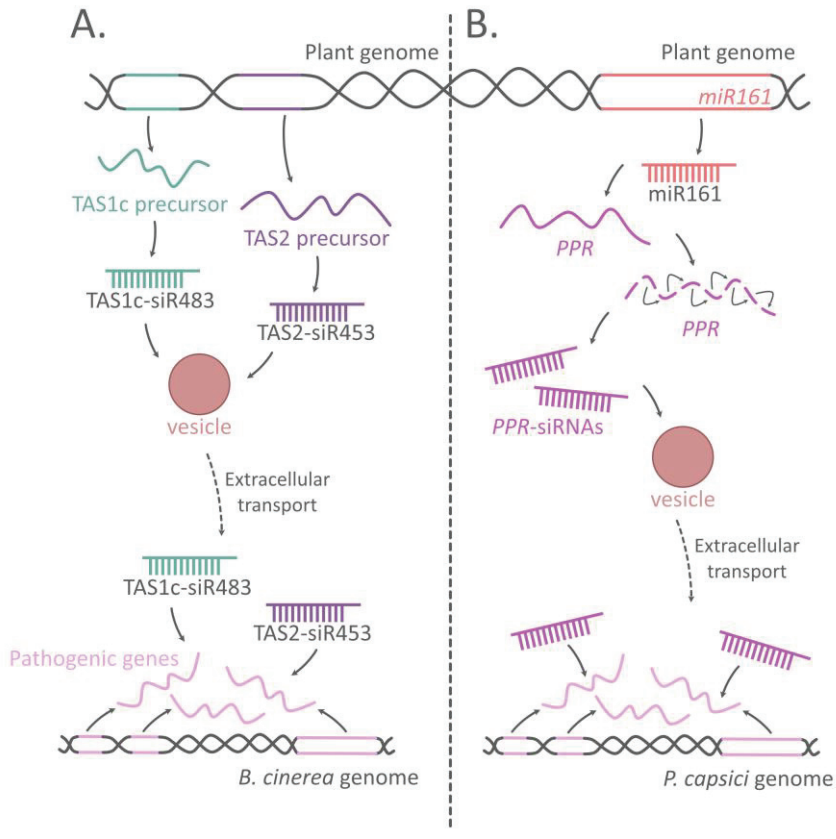


Figure 7. Role of endogenous siRNAs on silencing pathogenic genes. A. Role of TAS-derived siRNAs on the silencing of *Botrytis cinerea* genes. B. Role of miR161 on the production of PPR-derived secondary siRNAs and, ultimately, silencing of *Phytophtora capsici* genes.

Furthermore, a new class of bacteria-induced siRNAs was described during *Pst* infection (Katiyar-Agarwal et al. 2007). These novel siRNAs were named as long siRNAs (lsiRNAs) because, as their name suggests, they are longer than the usual siRNAs (30-40 nt). These lsiRNAs were produced by the action of DCL1 and DCL4, AGO7, RDR6 and PolIV and they presented common characteristics with siRNAs. Among the six lsiRNAs identified, lsiRNA-1 was exclusively detected during infection by *Pst* carrying the effector *avrRpt2*, which makes the bacteria avirulent. lsiRNA-1 was generated from the overlapping region of a NAT pair formed by two different loci, one that encodes a receptor-like protein kinase (RLK) and one that encodes a putative RNA-binding domain protein (RAP). The expression of lsiRNA-

1 led to the downregulation of its target *RAP*, whose expression is unfavorable for the response against bacteria (Fig. 8A) (Katiyar-Agarwal et al. 2007).

In 2014, Cao et al. described the first endogenous siRNAs produced by RDR1, named as virus-activated siRNAs (vasiRNAs) (Cao et al. 2014). Interestingly, these vasiRNAs were discovered on *Arabidopsis* plants infected with a mutant CMV defective on its VSR (CMV-Δ2b), and they were not produced when the plants were infected with the wild-type virus, showing that the presence of 2b is sufficient to inhibit their production. Generally, vasiRNAs were 21-nt long and they were generated from protein-coding genes, mapping exclusively to the mature mRNA. Moreover, their production relied on the activity of RDR1 and DCL4 and they were loaded into both AGO1 and AGO2. The accumulation of vasiRNAs was correlated with enhanced viral resistance, as seen in mutants for the gene *ETHYLENE-INSENSITIVE 5* (*EIN5*), which accumulated more vasiRNAs and were more resistant towards the infection (Fig. 8B). Accordingly, *rdr1* mutants, where the production of vasiRNAs is abolished, showed higher accumulation of CMV-Δ2b (Cao et al. 2014). Recently, the production of vasiRNAs was explored in turnip (*Brassica napa*), oilseed rape (*Brassica napus*) and *Arabidopsis* infected with Cauliflower mosaic virus (CaMV) (Leonetti et al. 2021). In these three plants, vasiRNAs were produced from 15 common genes, associated with stress response and photosynthesis, which were downregulated during infection and promoted resistance to the stress (Leonetti et al. 2021). Altogether, these studies show that, on the Brassicaceae family, the production of vasiRNAs is an important defense mechanism that promotes the downregulation of genes that are potentially detrimental for the defense response against different types of viruses.

All these different studies show that endogenous siRNAs play fundamental roles during defense response in plants. These studies have not only provided the knowledge about already described sRNAs, but also the discovery of new classes of siRNAs, as it occurred with lsiRNAs and vasiRNAs. Nevertheless, there are numerous questions regarding their implication on plant defense that need to be answered: do siRNAs generally participate on defense response or only specific ones have defensive roles? Do they participate in the response against other types of biotic stresses, such as herbivores? Are there any other new classes of siRNAs that have not yet been discovered? What are their effects of genes not associated to defense during the stress? Are they associated with other regulatory mechanisms? Providing the answers to those questions would importantly contribute to the better understanding of how the genome is regulated to cope with biotic stresses.

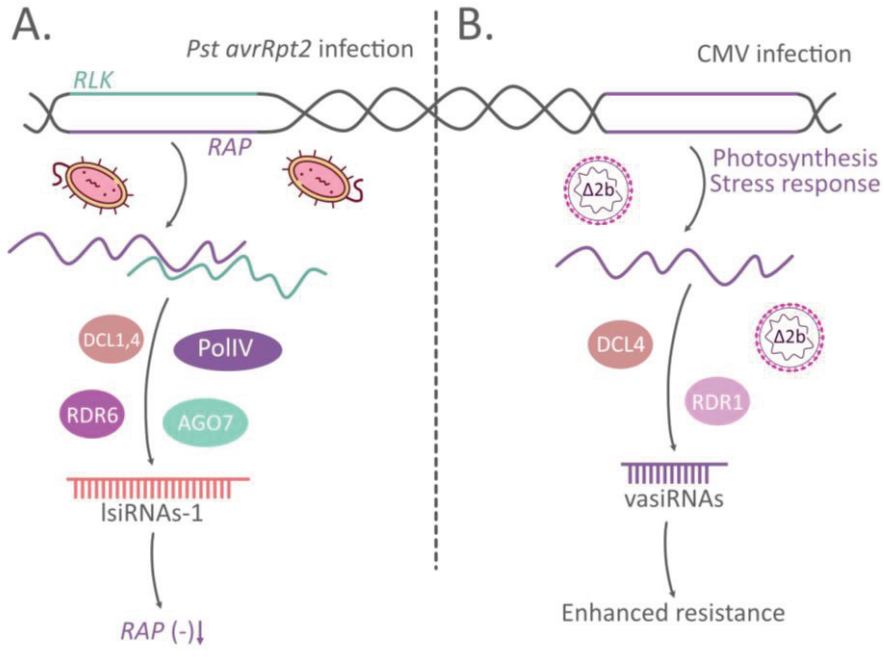


Figure 8. Role of biotic stress-derived siRNAs. A. Role of long siRNAs (lsiRNAs) during *Pseudomonas syringae* avRpt2 infection. B. Role of virus-activated siRNAs during CMV- Δ2b infection.

1.4.2.3 Pathogen-derived siRNAs

In the last years, the study of the effect of siRNAs derived from the genome of pathogens has gained increasing interest, unraveling a fascinating aspect of the interactions between pathogens and plants. Although this interaction remains poorly understood, several examples have been described, including siRNAs derived from fungi, parasitic plants, viroids, and viruses.

In 2013, sRNAs produced by a pathogenic fungus were demonstrated to act as effectors (molecules that suppress the host immunity), a function only assigned to proteins until that moment (Weiberg et al. 2013). In this study, *B. cinerea* sRNAs were proved to silence *Arabidopsis* genes involved in immunity by binding to the host AGO1. Specifically, the fungal sRNAs target mitogen activated protein kinases (*MPK1* and *MPK2*); peroxiredoxin (*PRXIIF*), a gene involved in oxidative stress; and a cell wall-associated kinase (*WAK*) (Fig. 9A). Furthermore, *Arabidopsis ago1* mutants showed a reduced susceptibility to the infection, while *B. cinerea dcl1 dcl2*

mutants did not produce sRNAs and, therefore, had a decreased pathogenicity, highlighting the importance of the fungal sRNAs for virulence (Weiberg et al. 2013).

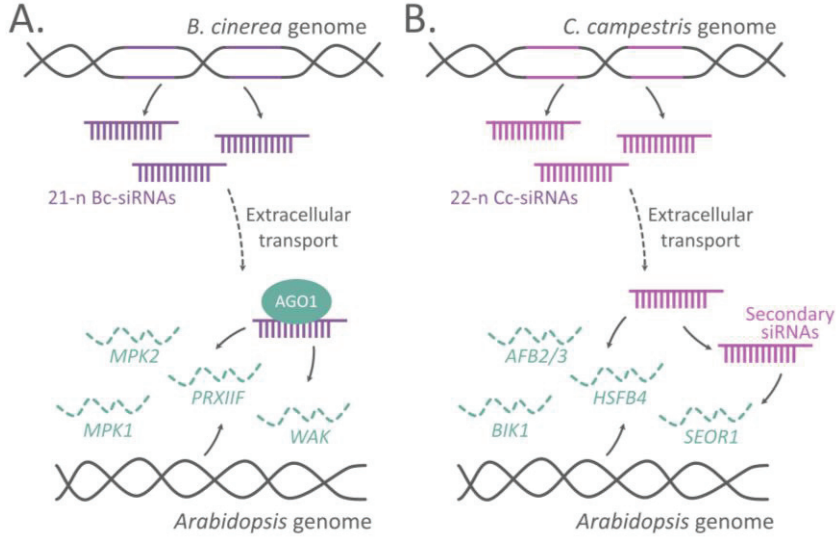


Figure 9. Roles of pathogen-derived siRNAs on the silencing of host genes. A. Role of 21-n *Botrytis cinerea*-derived siRNAs on silencing host genes. B. Role of 22-nt *Cuscuta campestris*-derived siRNAs on silencing host genes.

The parasitic weed *Cuscuta campestris*, also known as field dodder or golden dodder, produces 22-nt miRNAs during the parasitic relationship with *Arabidopsis* that target host genes. These 22-nt miRNAs target genes with different functions, such as a kinase localized in the plasma membrane (*BOTRYTIS-INDUCED KINASE 1*, *BIK1*), auxin receptors (*TIR1*, *AUXIN SIGNALING F BOX 2* (*AFB2*) and *AFB3*), a predicted transcriptional repressor from the root stem cells (*HEAT SHOCK TRANSCRIPTION FACTOR B4*, *HSFB4*), and a phloem protein related to photosynthesis regulation (*SIEVE ELEMENT OCCLUSION-RELATED 1*, *SEOR1*). In addition to cause mRNA cleavage, the *C. campestris* miRNAs lead to the production of secondary siRNAs which reinforce the silencing effect (Fig. 9B) (Shahid et al. 2018).

Interestingly, this phenomenon commonly occurs during viroid and viral infections (Ramesh et al. 2020). One of the first examples of this viral-mediated host silencing was described in CMV-infected plants. The silencing of CMV Y-satellite RNA (Y-RNA) produces 22-nt vsiRNAs that target a chlorophyll biosynthetic gene (*CHLI*) in different species, including tomato, *Arabidopsis* and several *Nicotiana* species. Interestingly, the targeting of the plant gene by the vsiRNAs leads to its down-

regulation and the appearance of the yellowing symptoms (Fig. 10B) (Shimura et al. 2011; Smith et al. 2011). Moreover, other cases where the presence of a vsiRNA is associated with the development of symptoms have been described. For instance, a vsiRNA derived from the RNA4 of the genome of Rice stripe virus (RSV), named vsiRNA-4A, targets and down-regulates the *N. benthamiana* gene *translational initiation factor 4A* (*eIF4A*). The decrease in expression of this gene causes the development of the typical symptomatology, including stunting, twisted leaves, and defective flowering (Shi et al. 2016).

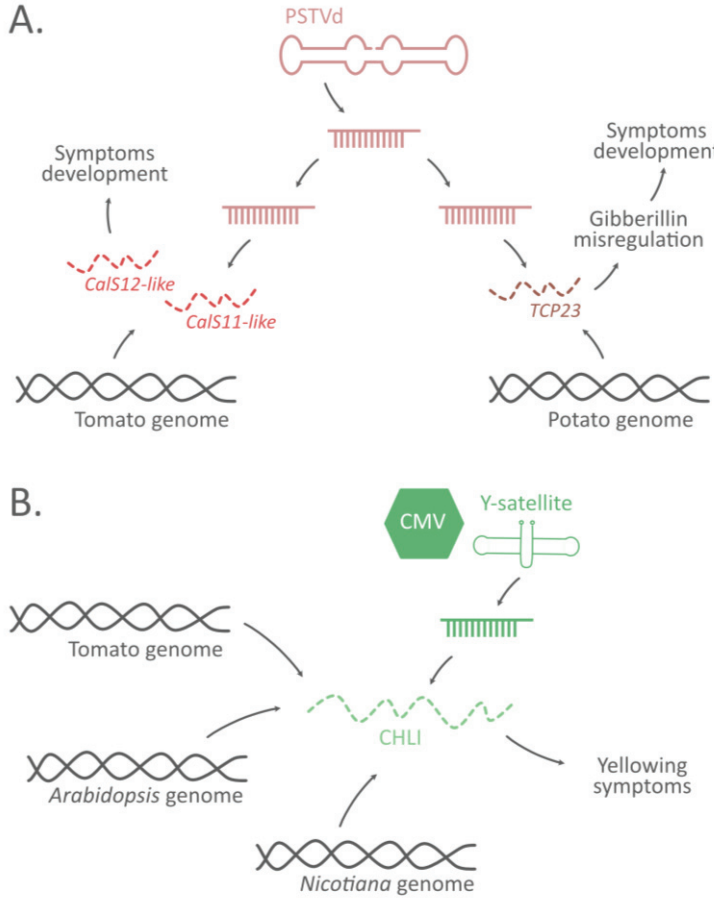


Figure 10. Role of pathogen-derived siRNAs on the silencing of host genes. A. Role of Potato spindle tuber viroid (PSTVd)-derived siRNAs on tomato and potato genes. B. Role of CMV-Y satellite-derived siRNAs on silencing the CHLI gene of tomato, *Arabidopsis* and *Nicotiana* plants.

Recently, this aspect of viral infections was explored at a genome-wide level during TuMV infection (Pitzalis et al. 2020). In this study, the authors performed sRNA-seq and degradome-seq to analyze the potential interactions between vsiRNAs and host mRNAs. Interestingly, they found 205 host mRNAs on their degradome analysis that are potentially targeted by 87 TuMV-derived siRNAs. Nonetheless, the targeting of the host genes by these vsiRNAs does not fully correlate with the transcriptomic changes, which suggests that this can be a mechanism to control, but not complete repress, the expression of host genes (Ramesh et al. 2020).

Other RNA pathogens, such as viroids, can produce siRNAs with activity against endogenous genes. It has been shown that a sRNA produced from the virulence modulating region of PSTVd down-regulates the expression of two tomato *callose synthase* genes (*CalS11-like* and *CalS12-like*). Moreover, the decrease of these two genes is related to the typical symptoms, as mutating the viroid to avoid the targeting of the genes remove the infected phenotype (Adkar-Purushothama et al. 2015). Recently, it has been shown that the same region of PSTVd produces sRNAs that targets the potato *TCP23* TF. The downregulation of this TF results in a misregulation of the gibberellin metabolism pathway, which leads to the stunting phenotype and the development of small and spindle-shaped tubers (Fig. 10A) (Bao et al. 2019).

All the above examples reflect the complex arms race between plants and pathogens and depict how intertwined the relation among them is, emphasizing the importance of a meticulous defense response that not only copes with the presence of the pathogen, but also avoids that it utilizes the responsive mechanisms, such as RNA silencing, on its own benefit. Nevertheless, there are still some remarkably interesting aspects about pathogen-derived siRNAs that have not been addressed yet, for instance, their potential genome-wide production, their impact on the endogenous siRNA population or their global exploitation of the plant RNA silencing machinery. This fascinating aspect of plant-pathogens interaction has the potential to not only provide further understanding of aspects related to the stress, such as the development of symptoms, but also to expand the knowledge about how RNA molecules can move from one organism to another.

.

2 Aims of the study

The overall aim of this study was to expand the current knowledge about the roles of epigenetic regulation and RNA silencing during biotic stress response in *Arabidopsis thaliana*, focusing on how these two molecular mechanisms modulate gene expression. To fulfil this goal, we aimed to:

I. Characterize the genome-wide changes in the epigenetic control and the transcriptomic reprogramming during aphid infestation.

II. Uncover the interaction of CMV-derived siRNAs with the endogenous RNA silencing machinery of *Arabidopsis thaliana*.

III. Explore the genome-wide changes on DNA methylation and repressive histone marks during CMV infection, determine the interplay between them and analyze their impact on gene expression and, ultimately, defense response.

3 Results and discussion

This section contains a brief description of the results obtained from the work presented in this thesis. The figures corresponding to the explained data are indicated throughout the text and they can be found in the compiled papers at the end of this PhD thesis.

3.1 Impact of biotic stress on the host RNA silencing

Plants are exposed to different stresses that can negatively impact their development and growth, being aphid infestation and CMV infection two very important biotic stress, from both an ecological and an economical point of view. Here, we have explored the impact of these two stresses on the activity of the RNA silencing machinery in *Arabidopsis thaliana* **(papers I, II and III)**.

In line with previous results (Donaire et al. 2009), CMV infected plants **(Fig. 1A, paper II)** accumulated high levels of CMV vsiRNAs **(Fig. 1B, paper II; Fig. 3B, paper III)**, which were mainly 21nt long (70.2%) **(Fig. 1D, paper II; Fig. 3B, paper III)**, as previously shown (Diaz-Pendon et al. 2007; Donaire et al. 2009; Herranz et al. 2015). Surprisingly, these vsiRNAs constituted almost half of the sequenced sRNA pool (43.1%) **(Fig. 1C, paper II)**, suggesting a considerable impact on the plant sRNAs homeostasis. This result led us to explore the changes on the endogenous sRNA populations, which, to our knowledge, have not been the focus of any previous study, contrary to the vsiRNAs (Donaire et al. 2009; Qiu et al. 2018), satellite-derived siRNAs (Du et al. 2007; Fang et al. 2015), or vasiRNAs (Cao et al. 2014). Indeed, infected plants showed an overall decreased accumulation of sRNAs, except miRNAs and 21 nt siRNAs derived from mRNAs and TEs, which were increasingly accumulated **(Fig. 1E, F and G, paper II; Fig. 2A, paper III)**. These results considerably contrast with the changes on the endogenous 24 nt sRNA populations both during CMV infection and aphid infestation, which are

characterized by a significant decrease at TEs **(Fig. 3C, D and E, paper I; Fig. 3A and C, paper III)**. Altogether, these results show that different biotic stresses lead to considerably different changes on the plant sRNA profiles. This is largely caused by the nature of the two studied stresses, as CMV is a virus and, therefore, it replicates inside the host cells interacting with the RNA silencing machinery, while aphids cause a subtle wounding during their feeding.

To understand the potential function of vsiRNAs, we explored their interaction with the plant AGOs **(Fig. 2B, paper II)**. First, we analyzed their 5' nucleotide as an evidence of their potential AGO loading (Mi et al. 2008). Overall, U was the predominant 5' nt of the vsiRNAs, followed by A, C and G, suggesting a potential loading into AGO1, AGO2 and AGO5 **(Fig. 2A, paper II)**. Moreover, their protein levels were increased during infection, with the exception of AGO7 **(Fig. 2C, paper II)**, which reflects their importance during the defense responsive, as previously shown (Harvey et al. 2011; Garcia-Ruiz et al. 2015). In the case of AGO2, its importance as an antiviral AGO was also confirmed at the transcriptional level, as this gene presents the highest fold-change of all the RNA silencing components in our RNA-seq dataset **(Fig. 2E and F, paper III).**

To characterize vsiRNA loading into antiviral AGOs, we performed AGO-IP followed by sRNA-seq **(Fig. 2D, paper II)**. In general, AGO1 and AGO2 showed the highest vsiRNAs loading **(Fig. 2E, paper II)**, highlighting once again their important antiviral role (Harvey et al. 2011; Watt et al. 2020), with minor contributions of AGO5 and AGO7. Furthermore, the 5' nt preference of the loaded vsiRNAs was as expected according previous reports (Mi et al. 2008) **(Fig. 2F, paper II)**. Overall, the loading of vsiRNAs displaced endogenous siRNAs in all AGOs, which could contribute to the development of the symptoms. This displacement could be particularly important in the case of miRNAs, whose lack of activity has been previously linked to viral infections (Bazzini et al. 2007).

Additionally, we explored the interference of CMV's VSR, the 2b protein, with the endogenous sRNA populations during viral infection, which was previously studied using 2b reporter constructs expressed through a transgene but not in natural infections (Hamera et al. 2012) **(Fig. 3A, paper II)**. Interestingly, 2b-IP followed by sRNA-seq revealed that 2b not only binds to vsiRNAs, but also to host siRNAs, with a preference for mRNA-derived sRNAs **(Fig. 3B, C and D, paper II)**. Then, we analyzed the expression of the genes from which these 2b-loaded mRNA-derived were being produced on public available data. Indeed, during CMV infection, these genes were downregulated **(Fig. 3E, paper II)**, suggesting that the loading of sRNAs

on 2b could potentially be functional. 2b-loaded sRNAs, in general, were 20- and 21-nt in length and did not present a 5' nt preference **(Fig. 3E and G, paper II)**. In summary, 2b participates in the viral infection by loading its own vsiRNAs, to avoid its silencing, and host sRNAs, to potentially interfere with host transcription, reflecting how VSRs have evolved to play key roles during the infection.

Previous studies described the targeting of plant mRNAs by pathogen-derived siRNAs, including fungi (Weiberg et al. 2013), parasitic plants (Shahid et al. 2018), viroids (Adkar-Purushothama et al. 2015) and viruses (Shimura et al. 2011; Smith et al. 2011). To explore this at a genome-wide level, we performed Parallel Analysis of RNA Ends (PARE) sequencing, which in plants is commonly used to identify the targets of miRNA activity. Our analysis identified 61 endogenous mRNAs with potential vsiRNA-induced cleavage **(Fig. 4A and B, paper II)**. The vsiRNAs that were potentially targeting these genes were distributed among all the studied AGOs, with a slight preference for AGO5 **(Fig. 4A, paper II)**. Moreover, these targeted mRNAs were enriched in molecular function Gene Ontology (GO) categories related to nucleic acid binding and biological function categories associated with photosynthesis, translation and production of metabolite precursors and energy **(Fig. 4C, paper II)**. This agrees with previous works that showed that a gene associated with chlorophyll production is targeted by a vsiRNA derived from the satellite RNA of CMV in *N. benthamiana* (Shimura et al. 2011; Smith et al. 2011). Interestingly, we saw that the targeting of endogenous mRNAs by vsiRNAs did not lead to the production of secondary siRNA, as the 21- and 22-nt mRNA-derived sRNAs from targeted genes were decreased in the infected plants **(Fig. 4E, paper II)**. Furthermore, we validated the targeting of two candidate genes, AT2G21210 (*RP1*) and AT4G36195 (a serine carboxypeptidase) by vsiRNAs **(Fig. 5A and B, paper II)** using transient expression of their targeted sequences fused to GFP in mock and infected *N. benthamiana* **(Fig. 5C, paper II)**. For both genes, the levels of GFP were downregulated on the infected plants when compared to the mock ones **(Fig. 5D and E, paper II)**. Moreover, mutations in the targeted sequenced significantly decreased the targeting of the vsiRNAs **(Fig. 5D and E, paper II)**. Overall, our results validated the ability of these vsiRNAs to effectively downregulated plant mRNAs. Altogether, our results suggest that the loading of vsiRNAs in the host RNA silencing machinery could act as a functional regulator of endogenous gene expression by mediating their silencing and potentially inhibiting secondary sRNA production.

In summary, our results show that the massive production of vsiRNAs alters the activity of the host RNA silencing machinery at different levels, disrupting the

endogenous sRNA population; hijacking the activity of host AGOs; decreasing the expression of plant genes through its VSR 2b; and inducing the cleavage of endogenous genes. Overall, our work provides further understanding on the complex interaction between RNA viruses and their hosts silencing machinery.

3.2 Impact of DNA methylation changes on the transcriptional reprogramming during biotic stress

Changes on DNA methylation have been associated with the response to a plethora of biotic stresses, from bacteria (Pavet et al. 2006; Dowen et al. 2012; Gohlke et al. 2013; Yu et al. 2013; Halter et al. 2021) to herbivores (Doucet et al. 2016). Nevertheless, most of these studies lacked a genome-wide perspective, as well as the study of the impact of DNA methylation changes on gene expression. In this doctoral thesis, we explored the genome-wide changes on DNA methylation caused by *Myzus persicae* infestation and CMV infection and we analyzed their effect on the plant transcriptional response by performing WGBS and RNA-seq, respectively **(paper I and III)**.

First, we analyzed the overall DNA methylation profiles of the stressed plants **(Fig. 4A, paper I; Fig. 4A, paper III)**. Interestingly, while aphid infestation did not cause dramatic differences on the levels of this mark between control and infested plants **(Fig. 4b, paper I)**, CMV infection led to significantly higher global levels all methylation context in TEs and the CG context in genes, at the two infection times analyzed, 10 and 20 dpi **(Fig. 4b, paper III)**. Despite these differences at the global level, both stresses were associated with a considerable number of localized methylation changes, and we were able to identify several DMRs, 2125 in the case of aphid infestation and 2768 and 4921 at 10 and 20 dpi, respectively, in the case of CMV infection **(see table 2 for more detailed information)**, being CHG the methylation context with the highest number of DMRs in both stresses. Next, we assessed the genomic location of these DMRs. In the case of aphid infestation, CG DMRs were predominantly found at genes, while CHG and CHH DMRs were mainly located at TEs **(Fig. 4C, paper I)**. In the case of CMV infection, both at 10 and 20 dpi, CG and CHH DMRs were mainly located genes, whereas the majority of the CHG DMRs were located at TEs **(Fig. 4C, paper III)**. Overall, these results show that, despite their different nature, both stresses cause different general changes on DNA methylation.

Interestingly, the aphid-related CHG and CHH DMRs were mainly located at TEs that lose 24 nt sRNAs **(Fig. 4D, paper I)** and were associated with both hypo- and hypermethylation **(Fig. 4E, paper I)**. Furthermore, hypomethylated CHH DMRs were associated with *Rath* elements and depleted from *Gypsy* elements **(Fig. 4G, paper I)**. On the other hand, CG DMRs had no significant changes, suggesting that they might be related to gbM **(Fig. 4E, paper I)**. Similarly, CMV-induced hypomethylated DMRs were associated with loss of 24-nt sRNAs **(Figure 4D, paper III)**. Loss of 24-nt during CMV infection was stronger at targets of the RdDM pathway as TEs, intergenic regions and rRNA repeats indicates **(Figure 3C, paper III**). Altogether, these results show that, during both stresses, changes on RdDM activity associated with the accumulation of sRNAs are partially translated into DNA methylation changes.

To explore the influence of the DMRs on gene expression, we analyzed the expression of genes located 2 kb upstream and downstream of them. During aphid infestation, upregulated genes were enriched on GO categories related to oxygen binding, translation regulator activity, nuclease and motor activity, fruit ripening and cell death **(Fig. 5A, paper I) (see Table 3 for the number of DEG in each experiment)**. Interestingly, we found several genes related to defense response close to hypomethylated CHH DMRs, including, *ERF022*, *AP2C1*, *ACS6* and *SYP122* **(Fig. 5C-F, paper I)**. To further explore the influence of CHH methylation on the expression of these genes, we analyzed their expression levels on public RNA-seq datasets of *ago4* and *polIV* mutants, where CHH methylation is considerably reduced. Indeed, DEG associated with DMRs were significantly enriched in these datasets, suggesting that they are regulated by the action of the RdDM pathway **(Fig. 5G and H, paper I)**. Interestingly, TF binding sites were enriched at DMRs of DEGs associated with loss of CHH methylation **(Supp. Fig. 5C and E, paper I)**. Additionally, several of the members of these TF families were differentially expressed under aphid infestation, indicating that the overlap between TF regulation and DNA methylation might play an important role in the orchestration of the aphid infestation transcriptional response.

Similar to aphid infestation, upregulated genes under CMV infection, at both time points, were associated with GO categories related to cell death and stress response **(Fig. 2C and D, paper III) (see Table 3 for the number of DEG in each experiment)**. Interestingly, the overlap between DMRs and DEGs was lower than expected **(see tables 4 and 5 for more detailed information)**, with 26 DEGs associated with DMRs at 10 dpi and 10 DEG associated with DMRs at 20 dpi **(Fig. 4E, paper III)**. Among these DMR-associated DEG, AT3G26220 and AT5G43750,

are illustrative examples of an upregulated gene associated with loss of DNA methylation and a downregulated gene associated with gain of DNA methylation, respectively **(Fig. 4F, paper III)**. Overall, these results suggest that changes on DNA methylation led to significant changes on gene expression during CMV infection.

Altogether, our analysis of the changes on DNA methylation during *Myzus persicae* infestation and CMV infection show how this mechanism is used to regulate particular genes whose expression is needed to overcome the stress and suggests that it might interact with other regulatory mechanisms such as TF-binding which might be an important part of the interplay between DNA methylation changes and the transcriptional changes observed under stress, as previously shown (O'Malley et al. 2016).

3.3 Roles of histone modifications during biotic stress

In plants, histone modifications have been indirectly associated with defense response mainly due to their regulatory role of the JA and SA pathways (Zhou et al. 2005; Alvarez-Venegas et al. 2007; Berr et al. 2010; Choi et al. 2012) and directly associated with the regulation of a defense gene activated by bacterial stress (Zervudacki et al. 2018). In our work, we wanted to explore the global changes on histone modifications during aphid infestation and CMV infection and analyze their correlation with gene expression and DNA methylation by performing Chromatin Immunoprecipitation (ChIP)-seq, RNA-seq, and WGB-seq.

Our screenings of sensitivity to aphid infestation and CMV infection in epigenetic mutants revealed that histone marks are fundamental for the defense in both scenarios, as *kyp* mutants are less susceptible to aphid infestation **(Fig. 6A and B paper I)** and *clf* mutants were more tolerant to CMV infection **(Fig. 1A, paper III)**, pointing to H3K9 methylation and H3K27 methylation as important regulatory marks, respectively. Interestingly, some of the defense genes upregulated under aphid feeding that were associated with hypomethylated CHH DMRs, particularly *SYP122*, *GER5* and *TCH4*, were also upregulated in *kyp* mutants not exposed to aphids **(Fig. 6C, paper I)**. This result suggests that H3K9me2 could regulate the transcription response to stress since, at least partially, the lower susceptibility of *kyp* to the aphid infestation could be due to the basal activation of defense genes. On the other hand, *clf* plants showed an enhanced tolerance towards CMV as their rosette radius, a phenotype trait that is severely affected by the infection on Col-0

plants, remained at mock levels **(Fig. 1B and C, paper III)**. These results suggested that repressive histone marks could be playing a key role during the defense response against *Myzus persicae* and CMV. To understand in detail how histone makes could contribute to the defense response against CMV infection we performed ChIP-seq of two well-known repressive marks, H3K9me2 and H3K27me3, to characterize their genome-wide changes at the two studied infection time points.

As expected, at both 10 and 20 dpi, changes on H3K9me2 were mainly found at TEs, while changes on H3K27me3 were predominantly located at genes **(Fig. 5A, paper III)**, a distribution that correlates the previously described roles of these two marks in *Arabidopsis* (Roudier et al. 2011). Overall, CMV infection led to a gain of H3K9me2 at TEs and a loss of both H3K27me3 and H3K9me2 at genes **(Fig. 5B, paper III**, **see table 4 for the number of identified peaks)**. At both time points, loss of H3K9me2 and gain of H3K27me3 were predominantly found on genes, while gain of H3K9me2 and loss of H3K27me3 were mainly seen at TEs **(Fig. 5C, paper III)**. Similar to DNA methylation, the number of DEGs associated with histone peaks was relatively low, only 3 DEGs were associated with changes on H3K9me2 and 3 DEGs associated by changes on H3K27me3 at 10 dpi, and 19 DEGs were associated with changes on H3K9me2 and 27 were associated with changes on H3K27me3 at 20 dpi **(Fig. 5D, paper III)**. Among these genes, AT5G35970 and AT4G01950 are examples of downregulated DEGs controlled by a gain of H3K9me2 and a gain of H3K27me3, respectively **(Fig. 5E, paper III)**. Altogether, these results show that CMV infection causes genome-wide changes on the pattern of H3K9me2 and H3K27me3 that might regulate the expression of specific genes.

Previous studies have described the correlation between DNA methylation and histones marks, including the interplay between CHG methylation and H3K9me2 (Du et al. 2015) and between DNA methylation and the recruitment of the H3K27 demethylase REF6 (Qiu et al. 2019). Hence, we explored the overlap between changes in DNA methylation and changes in histone marks in our datasets. To investigate this in the context of aphid infestation, we analyzed the histone environment of our identified DMRs on public datasets of different histone modifications. Interestingly, we observed that hypomethylated CHH DMRs were enriched on the permissive histone mark H3K18ac and depleted on the repressive marks H3K27me and H3K9me2 **(Fig. 4F, paper I)**, indicating that changes in CHH methylation took place at regions with a permissive chromatin status or potentially sensitive to active demethylation, since that histone environment resembles the genomic target regions of ROS1 (Tang et al. 2016).

The genome-wide distribution of the changes on DNA methylation, H3K9me2 and H3K27me3 during CMV infection showed that increased methylation and gain on H3K9me2 were correlated at centromeric and pericentromeric region, especially in the case of CHG methylation. Moreover, loss of H3K27me3 was associated with a significant increase of DNA methylation at non-CG contexts **(Fig. 6A and B, paper III)**. Finally, we explored the overlap between these two histone marks, as they are usually exclusive (Roudier et al. 2011). Interestingly, there was a considerable overlap between regions losing H3K27me3 and gaining H3K9me2 at both 10 and 20 dpi **(Fig. 6C, paper III)** , indicating that their accumulation is dynamic during stress and that regions losing H3K27me3 might be targets of the RdDM pathway which might attract H3K9me2 to maintain the repression of those regions. This is in line with recent publications indicating that H3K27me3 serves as a back-up silencing machinery for heterochromatic sequences experiencing demethylation (Rougée et al. 2021).

Overall, our work shows that, despite their different nature, both aphid infestation and CMV infection are associated with changes on histone modifications. In the case of CMV infections, these changes lead to a reorganization of the genome-wide patterns of H3K9me2 and H3K27me3, which are associated with changes in DNA methylation and have the potential to regulate gene expression.

4 Conclusions

The use of different high-throughput sequencing techniques coupled with analysis of mutant lines for different epigenetic and RNA silencing components, has allowed us to identify the overall genome-wide epigenetic changes taking place under two different biotic stresses: aphid infestation and viral infection.

Several biotic stresses have been previously associated with changes in the activity of the RNA silencing and epigenetic machineries and their potential connection to the orchestration of the transcriptional response against stress. My work showed that indeed, the epigenetic response towards aphid infestation is characterized by a localized loss of DNA methylation at TEs, which is accompanied by a loss of 24-nt sRNAs, reflecting that RdDM activity is partially responsible of demethylation under this stress. These results potentially explain the TE reactivation that takes place during this stress, as they suggest a decrease of RdDM activity that could allow their transcriptional reactivation. Furthermore, we showed that these changes on DNA methylation are involved in the transcriptional control of defense genes against aphid infestation. The other biotic stress analyzed in my thesis, infection by CMV, led to a global gain of DNA methylation and a reorganization of the repressive marks H3K9me2 and H3K27me3. These changes did not have a genome-wide impact on gene expression, but they affected the expression of genes associated with the defense response and physiological processes linked to the viral infection. Furthermore, these mechanisms might represent a mechanism of surveillance of genome stability in order to maintain TEs in a transcriptionally silenced state.

Additionally, we characterized the interference of CMV with the RNA silencing machinery and the impact of vsiRNA on the plant sRNA populations. Our results showed that vsiRNAs constituted almost half of the population of sRNAs, which led to an overall reduction on the endogenous sRNAs. Moreover, we demonstrated that vsiRNAs were loaded into plant AGOs and negatively regulated the expression of host genes. Both works focused on CMV interaction with the RNA silencing and

epigenetic machineries reflect the complexity of plant-virus interactions and the need to study in detail all the molecular mechanisms affected by biotic stresses to understand their full interference with their hosts biology.

Altogether, the work presented in this thesis indicates that the response to biotic stresses is associated with RNA silencing and epigenetic changes, including both DNA methylation and the distribution of histone marks. This thesis contributes to a better understanding of how the different epigenetic mechanisms and the RNA silencing machinery function during biotic stress response, their dynamism during stress and how they impact gene expression.

5 Future perspective

The work presented in this doctoral thesis has contributed to a better understanding of the roles of epigenetics and RNA silencing during two important biotic stresses, aphid infestation and CMV infection. Nevertheless, there are still several interesting aspects that remain to be studied in the future.

Our work has extensively described the role of DNA methylation during aphid infestation and pointed towards histone marks, particularly H3K9me2, as another fundamental regulatory mechanism. Therefore, it would be of highly interest to explore the changes of this repressive mark during this stress using ChIP-seq. Our data showed that *kyp* mutants were more resistant to aphids, potentially due to an increased basal expression of defense genes. Hence, one possibility would be that these genes are located in a chromatin environment regulated by H3K9me2. Furthermore, exploring other histone marks, for instance H3K27me3, would provide a more complete picture of the epigenetic changes that occur during aphid infestation. Finally, the role of other regulatory mechanisms, especially TFs, and their interplay with the epigenetic regulation is another important question that needs to be addressed. To answer this, it would be interesting to perform assay for transposase-accessible chromatin using sequencing (ATAC-seq) under aphid infestation and analyze the overall correlation with repressive epigenetic marks.

Another limitation of our study is the lack of different exposure times to aphid feeding. We focused on analyzing the epigenetic changes that took place after 72 hours of aphid exposure but including other time points could potentially provide further details. Moreover, this could also be closer to what crops encounter in fields, where the precise time of the duration of the stress might be unknown. Therefore, analyzing different time points and performing recurrent exposures could provide knowledge that could potentially be used in future applied projects. These analyses could also contribute to a better understanding of the components regulating the

overall changes observed at the RNA silencing and DNA methylation levels, for which we lack mechanistic information.

Despite our thorough analysis of the epigenetic changes associated with CMV infection, more details of these interaction could be investigated. Similar to the aphid infestation, other histone marks could be analyzed, for instance, histone acetylation, which has been previously associated with the defense response against several pathogens. This will not only provide further knowledge about the chromatin dynamics during CMV infection, but also about the effect of these marks on gene expression. Furthermore, the implication of TFs on the transcriptional reprogramming and their interplay with epigenetic marks should also be analyzed in detailed. We also lack mechanistic information of how the observed epigenetic changes take place, so the analysis of the responsiveness of genes from different epigenetic pathways should be explored in more detail in the future.

Once the mechanistic framework during CMV infection is defined, the next step would be to select those genes that present considerable changes and might be related to the defense response or to the development of the symptoms. Characterizing the performance of their mutants and overexpressing lines during CMV infection could be a first step towards the potential production of crops with improved resistance to CMV infection.

The recovery of the viral-infected plants constitutes another remarkably interesting aspect to be analyzed in the future. This would provide valuable information about the whole infection process, but also about potential ways of enhancing recovery on infected plants. Furthermore, the analysis of the combination of viral stress with other stress, for instance aphids, is another fascinating research field. These scenarios of multiple stresses represent a better picture of what plants undergo in natural environments, so understanding how they deal with concurrent stresses has the potential to provide extremely helpful knowledge for applied purposes.

On the other hand, our study of the genome-wide effects of vsiRNAs on the host RNA silencing machinery had shed light in this less understood aspect of viral infections. Nevertheless, there are still some details that could be analyzed in the future. First, even if we had already analyzed the main antiviral AGOs, the role of the rest of them should also be investigated, as they could also participate in this phenomenon. Moreover, our study has revealed that the VSR of CMV, the 2b protein, apart from loading its own vsiRNAs, also loads host mRNA-derived sRNAs, potentially silencing plant genes. This regulatory effect should be experimentally

confirmed, for instance, by analyzing the expression level of the potentially affected genes in CMV-Δ2b infected plants or 2b-overexpressing lines.

Overall, our work has established the basis of fascinating future research that, hopefully, will provide even further knowledge on the intriguing relationship between plants and their environment.

References

Adkar-Purushothama CR, Brosseau C, Giguere T, Sano T, Moffett P, Perreault JP. 2015. Small RNA Derived from the Virulence Modulating Region of the Potato spindle tuber viroid Silences callose synthase Genes of Tomato Plants. *Plant Cell* **27**: 2178-2194.

Agorio A, Vera P. 2007. ARGONAUTE4 is required for resistance to Pseudomonas syringae in Arabidopsis. *Plant Cell* **19**: 3778-3790.

Akbergenov R, Si-Ammour A, Blevins T, Amin I, Kutter C, Vanderschuren H, Zhang P, Gruissem W, Meins F, Jr., Hohn T et al. 2006. Molecular characterization of geminivirus-derived small RNAs in different plant species. *Nucleic Acids Res* **34**: 462-471.

Aldon D, Mbengue M, Mazars C, Galaud JP. 2018. Calcium Signalling in Plant Biotic Interactions. *Int J Mol Sci* **19**.

Alonso C, Ramos-Cruz D, Becker C. 2019. The role of plant epigenetics in biotic interactions. *New Phytol* **221**: 731-737.

Alvarez-Venegas R, Abdallat AA, Guo M, Alfano JR, Avramova Z. 2007. Epigenetic control of a transcription factor at the cross section of two antagonistic pathways. *Epigenetics* **2**: 106-113.

Annacondia ML, Mageroy MH, Martinez G. 2018. Stress response regulation by epigenetic mechanisms: changing of the guards. *Physiol Plant* **162**: 239-250.

Axtell MJ. 2013. Classification and Comparison of Small RNAs from Plants. *Annual Review of Plant Biology* **64**: 137-159.

Backer R, Naidoo S, van den Berg N. 2019. The NONEXPRESSOR OF PATHOGENESIS-RELATED GENES 1 (NPR1) and Related Family: Mechanistic Insights in Plant Disease Resistance. *Front Plant Sci* **10**: 102.

Baldrich P, Kakar K, Siré C, Moreno AB, Berger A, García-Chapa M, López-Moya JJ, Riechmann JL, San Segundo B. 2014. Small RNA profiling reveals regulation of Arabidopsis miR168 and heterochromatic siRNA415 in response to fungal elicitors. *BMC Genomics* **15**: 1625-1640.

Bansal R, Mian MAR, Mittapalli O, Michel AP. 2014. RNA-Seq reveals a xenobiotic stress response in the soybean aphid, Aphis glycines, when fed aphid-resistant soybean. *BMC Genomics* **15**: 972.

Bao S, Owens RA, Sun Q, Song H, Liu Y, Eamens AL, Feng H, Tian H, Wang MB, Zhang R. 2019. Silencing of transcription factor encoding gene StTCP23 by small RNAs derived from the virulence modulating region of potato spindle tuber viroid is associated with symptom development in potato. *PLoS Pathog* **15**: e1008110.

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

Bass C, Puinean AM, Zimmer CT, Denholm I, Field LM, Foster SP, Gutbrod O, Nauen R, Slater R, Williamson MS. 2014. The evolution of insecticide resistance in the peach potato aphid, Myzus persicae. *Insect Biochem Mol Biol* **51**: 41-51.

Baulcombe DC. 2002. RNA silencing. *Current Biology* **12**: 82-84.

Baulcombe DC. 2004. RNA silencing in plants. *Nature* **431**: 356-363.

Baulcombe DC, Dean C. 2014. Epigenetic regulation in plant responses to the environment. *Cold Spring Harb Perspect Biol* **6**: a019471.

Baumbusch LO, Thorstensen T, Krauss V, Fischer A, Naumman K, Assalkhou R, Schulz I, Reuter G, Aalen RB. 2001. The Arabidopsis thaliana genome contains at least 29 active genes encoding SET domain proteins that can be assigned to four evolutionary conserved classes. *Nucleic Acid Research* **29**: 4319-4333.

Bazzini AA, Hopp HE, Beachy RN, Asurmendi S. 2007. Infection and coaccumulation of tobacco mosaic virus proteins alter microRNA levels, correlating with symptom and plant development. *Proc Natl Acad Sci U S A* **104**: 12157-12162.

Bazzini AA, Manacorda CA, Tohge T, Conti G, Rodriguez MC, Nunes-Nesi A, Villanueva S, Fernie AR, Carrari F, Asurmendi S. 2011. Metabolic and miRNA profiling of TMV infected plants reveals biphasic temporal changes. *PLoS One* **6**: e28466.

Béclin C, Berthomé R, Palauqui J-C, Tepfer M, Vaucheret H. 1998. Infection of Tobacco or Arabidopsis Plants by CMV Counteracts Systemic Post-transcriptional Silencing of Nonviral (Trans)Genes. *Virology* **252**: 313-317.

Bejerman N, Mann KS, Dietzgen RG. 2016. Alfalfa dwarf cytorhabdovirus P protein is a local and systemic RNA silencing supressor which inhibits programmed RISC activity and prevents transitive amplification of RNA silencing. *Virus Res* **224**: 19-28.

Berr A, McCallum EJ, Alioua A, Heintz D, Heitz T, Shen WH. 2010. Arabidopsis histone methyltransferase SET DOMAIN GROUP8 mediates induction of the jasmonate/ethylene pathway genes in plant defense response to necrotrophic fungi. *Plant Physiol* **154**: 1403-1414.

Berry S, Dean C. 2015. Environmental perception and epigenetic memory: mechanistic insight through FLC. *Plant J* **83**: 133-148.

Bewick AJ, Schmitz RJ. 2017. Gene body DNA methylation in plants. *Curr Opin Plant Biol* **36**: 103-110.

Blackman RL. 1974. Life-cycle variation of Myzus persicae (Sulz.) (Horn., Aphididae) in different parts of the world, in relation to genotype and environment. *Bulletin of Entomological Research* **63**: 595-607.

Blevins T, Rajeswaran R, Shivaprasad PV, Beknazariants D, Si-Ammour A, Park HS, Vazquez F, Robertson D, Meins F, Jr., Hohn T et al. 2006. Four plant

Dicers mediate viral small RNA biogenesis and DNA virus induced silencing. *Nucleic Acids Res* **34**: 6233-6246.

Boccara M, Sarazin A, Thiebeauld O, Jay F, Voinnet O, Navarro L, Colot V. 2014. The Arabidopsis miR472-RDR6 silencing pathway modulates PAMP- and effector-triggered immunity through the post-transcriptional control of disease resistance genes. *PLoS Pathog* **10**: e1003883.

Bologna NG, Iselin R, Abriata LA, Sarazin A, Pumplin N, Jay F, Grentzinger T, Dal Peraro M, Voinnet O. 2018. Nucleo-cytosolic Shuttling of ARGONAUTE1 Prompts a Revised Model of the Plant MicroRNA Pathway. *Mol Cell* **69**: 709-719 e705.

Borges F, Martienssen RA. 2015. The expanding world of small RNAs in plants. *Nat Rev Mol Cell Biol* **16**: 727-741.

Boyko A, Kathiria P, Zemp FJ, Yao Y, Pogribny I, Kovalchuk I. 2007. Transgenerational changes in the genome stability and methylation in pathogen-infected plants: (virus-induced plant genome instability). *Nucleic Acids Res* **35**: 1714-1725.

Braendle C, Davis GK, Brisson JA, Stern DL. 2006. Wing dimorphism in aphids. *Heredity (Edinb)* **97**: 192-199.

Brodersen P, Voinnet O. 2006. The diversity of RNA silencing pathways in plants. *Trends Genet* **22**: 268-280.

Brosseau C, Moffett P. 2015. Functional and Genetic Analysis Identify a Role for Arabidopsis ARGONAUTE5 in Antiviral RNA Silencing. *Plant Cell* **27**: 1742-1754.

Buchmann RC, Asad S, Wolf JN, Mohannath G, Bisaro DM. 2009. Geminivirus AL2 and L2 proteins suppress transcriptional gene silencing and cause genome-wide reductions in cytosine methylation. *J Virol* **83**: 5005-5013.

Cabrera J, Barcala M, Garcia A, Rio-Machin A, Medina C, Jaubert-Possamai S, Favery B, Maizel A, Ruiz-Ferrer V, Fenoll C et al. 2016. Differentially expressed small RNAs in Arabidopsis galls formed by Meloidogyne javanica: a functional role for miR390 and its TAS3-derived tasiRNAs. *New Phytol* **209**: 1625-1640.

Cai Q, Qiao L, Wang M, He B, Lin F-M, Palmquist J, Huang S-D, Jin H. 2018. Plants send small RNAs in extracellular vesicles to fungal pathogen to silence virulence genes. *Science* **360**: 1126-1129.

Camargo-Ramírez R, Val-Torregrosa B, San Segundo B. 2018. MiR858-Mediated Regulation of Flavonoid-Specific MYB Transcription Factor Genes Controls Resistance to Pathogen Infection in Arabidopsis. *Plant and Cell Physiology* **59**: 190-204.

Canto T, Prior DAM, Hellward K-H, Oparka KJ, Palukaitis P. 1997. Characterization of Cucumber Mosaic Virus. *Virology*: 237-248.

Cao HH, Zhang ZF, Wang XF, Liu TX. 2018. Nutrition versus defense: Why Myzus persicae (green peach aphid) prefers and performs better on young leaves of cabbage. *PLoS One* **13**: e0196219.

Cao M, Du P, Wang X, Yu YQ, Qiu YH, Li W, Gal-On A, Zhou C, Li Y, Ding SW. 2014. Virus infection triggers widespread silencing of host genes by a distinct class of endogenous siRNAs in Arabidopsis. *Proc Natl Acad Sci U S A* **111**: 14613-14618.

Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, Jones RS, Zhang Y. 2002. Role of Histone H3 Lysine 27 Methylation in Polycomb-Group Silencing. *Science* **298**: 1039-1043.

Castellano M, Martinez G, Pallás V, Gómez G. 2015. Alterations in host DNA methylation in response to constitutive expression of Hop stunt viroid RNA in Nicotiana benthamiana plants. *Plant Pathology* **64**: 1247-1257.

Castellano M, Pallas V, Gomez G. 2016. A pathogenic long noncoding RNA redesigns the epigenetic landscape of the infected cells by subverting host Histone Deacetylase 6 activity. *New Phytol* **211**: 1311-1322.

Ceniceros-Ojeda EA, Rodríguez-Negrete EA, Rivera-Bustamante RF, Simon A. 2016. Two Populations of Viral Minichromosomes Are Present in a Geminivirus-Infected Plant Showing Symptom Remission (Recovery). *Journal of Virology* **90**: 3828-3838.

Chang YN, Zhu C, Jiang J, Zhang H, Zhu JK, Duan CG. 2020. Epigenetic regulation in plant abiotic stress responses. *J Integr Plant Biol* **62**: 563-580.

Chen B, Francki RIB. 1990. Cucumovirus transmission by the aphid Myzus persicae is determined solely by the viral coat protein. *Journal of General Virology* **71**: 939-944.

Chen Y, Singh A, Kaithakottil GG, Mathers TC, Gravino M, Mugford ST, van Oosterhout C, Swarbreck D, Hogenhout SA. 2020. An aphid RNA transcript migrates systemically within plants and is a virulence factor. *Proceedings of the National Academy of Sciences* **117**: 12763-12771.

Cheng K, Xu Y, Yang C, Ouellette L, Niu L, Zhou X, Chu L, Zhuang F, Liu J, Wu H et al. 2020. Histone tales: lysine methylation, a protagonist in Arabidopsis development. *J Exp Bot* **71**: 793-807.

Chiu MH, Chen IH, Baulcombe DC, Tsai CH. 2010. The silencing suppressor P25 of Potato virus X interacts with Argonaute1 and mediates its degradation through the proteasome pathway. *Mol Plant Pathol* **11**: 641-649.

Choi J, Lyons DB, Kim MY, Moore JD, Zilberman D. 2020. DNA Methylation and Histone H1 Jointly Repress Transposable Elements and Aberrant Intragenic Transcripts. *Mol Cell* **77**: 310-323 e317.

Choi SM, Song HR, Han SK, Han M, Kim CY, Park J, Lee YH, Jeon JS, Noh YS, Noh B. 2012. HDA19 is required for the repression of salicylic acid biosynthesis and salicylic acid-mediated defense responses in Arabidopsis. *Plant J* **71**: 135-146.

Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, Pradhan S, Nelson SF, Pellegrini M, Jacobsen SE. 2008. Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature* **452**: 215-219.

Copenhaver GP, Law JA, Vashisht AA, Wohlschlegel JA, Jacobsen SE. 2011. SHH1, a Homeodomain Protein Required for DNA Methylation, As Well As RDR2, RDM4, and Chromatin Remodeling Factors, Associate with RNA Polymerase IV. *PLoS Genetics* **7**.

Correa RL, Sanz-Carbonell A, Kogej Z, Muller SY, Ambros S, Lopez-Gomollon S, Gomez G, Baulcombe DC, Elena SF. 2020. Viral Fitness Determines the Magnitude of Transcriptomic and Epigenomic Reprograming of Defense Responses in Plants. *Mol Biol Evol* **37**: 1866-1881.

Costa S, Dean C. 2019. Storing memories: the distinct phases of Polycomb-mediated silencing of Arabidopsis FLC. *Biochem Soc Trans* **47**: 1187-1196.

Crevillen P, Yang H, Cui X, Greeff C, Trick M, Qiu Q, Cao X, Dean C. 2014. Epigenetic reprogramming that prevents transgenerational inheritance of the vernalized state. *Nature* **515**: 587-590.

Csorba T, Bovi A, Dalmay T, Burgyan J. 2007. The p122 subunit of Tobacco Mosaic Virus replicase is a potent silencing suppressor and compromises both small interfering RNA- and microRNA-mediated pathways. *J Virol* **81**: 11768-11780.

Cuerda-Gil D, Slotkin RK. 2016. Non-canonical RNA-directed DNA methylation. *Nat Plants* **2**: 16163.

Czermin B, Melfi R, McCabe D, Seitz V, Imhof A, Pirrotta V. 2002. Drosophila Enhancer of Zeste-ESC Complexes Have a Histone H3 Methyltransferase Activity that Marks Chromosomal Polycomb Sites. *Cell* **111**: 185-196.

De Vos M, Van Oosten VR, Van Poecke RMP, Van Pelt JA, Pozo MJ, Mueller MJ, Buchala AJ, Métraux J-P, Van Loon LC, Dicke M et al. 2005. Signal signature and transcriptome changes of Arabidopsis during pathogen and insect attack. *Molecular Plant-Microbe Interactions* **18**: 923-937.

Deleris A, Gallego-Bartolome J, Bao J, Kasschau KD, Carrington JC, Voinnet O. 2006. Hierarchical Action and Inhibition of Plant Dicer-Like Proteins in Antiviral Defense. *Science* **313**: 68-71.

Deleris A, Greenberg MV, Ausin I, Law RW, Moissiard G, Schubert D, Jacobsen SE. 2010. Involvement of a Jumonji-C domain-containing histone demethylase in DRM2-mediated maintenance of DNA methylation. *EMBO Rep* **11**: 950-955.

Devers EA, Brosnan CA, Sarazin A, Albertini D, Amsler AC, Brioudes F, Jullien PE, Lim P, Schott G, Voinnet O. 2020. Movement and differential consumption of short interfering RNA duplexes underlie mobile RNA interference. *Nat Plants* **6**: 789-799.

Diao P, Zhang Q, Sun H, Ma W, Cao A, Yu R, Wang J, Niu Y, Wuriyanghan H. 2019. miR403a and SA Are Involved in NbAGO2 Mediated Antiviral Defenses Against TMV Infection in Nicotiana benthamiana. *Genes (Basel)* **10**.

Diaz-Pendon JA, Li F, Li WX, Ding SW. 2007. Suppression of antiviral silencing by cucumber mosaic virus 2b protein in Arabidopsis is associated with

drastically reduced accumulation of three classes of viral small interfering RNAs. *Plant Cell* **19**: 2053-2063.
Doerfler W. 2008. In pursuit of the first recognized epigenetic signal--DNA methylation: a 1976 to 2008 synopsis. *Epigenetics* **3**: 125-133.
Donaire L, Wang Y, Gonzalez-Ibeas D, Mayer KF, Aranda MA, Llave C. 2009. Deep-sequencing of plant viral small RNAs reveals effective and widespread targeting of viral genomes. *Virology* **392**: 203-214.
Dong Z, Han M-H, Fedoroff N. 2008. The RNA-binding proteins HYL1 and SE promote accurate in vitro processing of pri-miRNA by DCL1. *Proceedings of the National Academy of Sciences* **105**: 9970-9975.
Doolittle SP. 1916. A new infectious mosaic disease of cucumber. *Phytopathology* **6**: 145-147.
Doucet D, Kellenberger RT, Schlüter PM, Schiestl FP. 2016. Herbivore-Induced DNA Demethylation Changes Floral Signalling and Attractiveness to Pollinators in Brassica rapa. *Plos One* **11**.
Dowen RH, Pelizzola M, Schmitz RJ, Lister R, Dowen JM, Nery JR, Dixon JE, Ecker JR. 2012. Widespread dynamic DNA methylation in response to biotic stress. *Proc Natl Acad Sci U S A* **109**: E2183-2191.
Du J, Johnson LM, Jacobsen SE, Patel DJ. 2015. DNA methylation pathways and their crosstalk with histone methylation. *Nat Rev Mol Cell Biol* **16**: 519-532.
Du QS, Duan CG, Zhang ZH, Fang YY, Fang RX, Xie Q, Guo HS. 2007. DCL4 targets Cucumber mosaic virus satellite RNA at novel secondary structures. *J Virol* **81**: 9142-9151.
Du Z, Chen A, Chen W, Westwood JH, Baulcombe DC, Carr JP. 2014. Using a viral vector to reveal the role of microRNA159 in disease symptom induction by a severe strain of Cucumber mosaic virus. *Plant Physiol* **164**: 1378-1388.
Duan CG, Fang YY, Zhou BJ, Zhao JH, Hou WN, Zhu H, Ding SW, Guo HS. 2012. Suppression of Arabidopsis ARGONAUTE1-mediated slicing, transgene-induced RNA silencing, and DNA methylation by distinct domains of the Cucumber mosaic virus 2b protein. *Plant Cell* **24**: 259-274.
Duan CG, Wang X, Xie S, Pan L, Miki D, Tang K, Hsu CC, Lei M, Zhong Y, Hou YJ et al. 2017. A pair of transposon-derived proteins function in a histone acetyltransferase complex for active DNA demethylation. *Cell Res* **27**: 226-240.
Duhlian L, Koramutla MK, Subramanian S, Chamola R, Bhattacharya R. 2020. Comparative transcriptomics revealed differential regulation of defense related genes in Brassica juncea leading to successful and unsuccessful infestation by aphid species. *Sci Rep* **10**: 10583.
Dutta A, Choudhary P, Caruana J, Raina R. 2017. JMJ27, an Arabidopsis H3K9 histone demethylase, modulates defense against Pseudomonas syringae and flowering time. *Plant J* **91**: 1015-1028.
Eamens A, Wang M-B, Smith NA, Waterhouse PM. 2008. RNA Silencing in Plants: Yesterday, Today, and Tomorrow. *Plant Physiology* **147**: 456-468.

Erdmann RM, Picard CL. 2020. RNA-directed DNA Methylation. *PLoS Genet* **16**: e1009034.

Fahlgren N, Howell MD, Kasschau KD, Chapman EJ, Sullivan CM, Cumbie JS, Givan SA, Law TF, Grant SR, Dangl JL et al. 2007. High-throughput sequencing of Arabidopsis microRNAs: evidence for frequent birth and death of MIRNA genes. *PLoS One* **2**: e219.

Fang Y-Y, Smith NA, Zhao J-H, Lee JRM, Guo H-S, Wang M-B. 2015. Cloning and Profiling of Small RNAs from Cucumber Mosaic Virus Satellite RNA. In *Plant Virology Protocols: New Approaches to Detect Viruses and Host Responses*, doi:10.1007/978-1-4939-1743-3_9 (ed. I Uyeda, C Masuta), pp. 99-109. Springer New York, New York, NY.

Fei Q, Li P, Teng C, Meyers BC. 2015. Secondary siRNAs from Medicago NB-LRRs modulated via miRNA-target interactions and their abundances. *Plant J* **83**: 451-465.

Finnegan EJ, Dennis ES. 1993. Isolation and identification by sequence homology of a putative cytosine methyltransferase from Arabidopsis thaliana. *Nucleic Acid Research* **21**: 2383-2388.

Finnegan EJ, Peacock WJ, Dennis ES. 1996. Reduced DNA methylation in Arabidopsis thaliana results in abnormal plant development. *Proc Natl Acad Sci USA* **93**: 8449-8454.

Frost JM, Kim MY, Park GT, Hsieh PH, Nakamura M, Lin SJH, Yoo H, Choi J, Ikeda Y, Kinoshita T et al. 2018. FACT complex is required for DNA demethylation at heterochromatin during reproduction in Arabidopsis. *Proc Natl Acad Sci U S A* **115**: E4720-E4729.

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

García-Arenal F, Palukaitis P. 2008. Cucumber Mosaic Virus. In *Encyclopedia of Virology*, pp. 614-619. Academic Press.

Garcia-Ruiz H, Carbonell A, Hoyer JS, Fahlgren N, Gilbert KB, Takeda A, Giampetruzzi A, Garcia Ruiz MT, McGinn MG, Lowery N et al. 2015. Roles and programming of Arabidopsis ARGONAUTE proteins during Turnip mosaic virus infection. *PLoS Pathog* **11**: e1004755.

Gaspar T, Franck T, Bisbis B, Kevers C, Jouve L, Hausman JF, Dommes J. 2002. Concepts in plant stress physiology. Application to plant tissue cultures. *Plant Growth Regulation* **37**: 263-285.

Gohlke J, Scholz CJ, Kneitz S, Weber D, Fuchs J, Hedrich R, Deeken R. 2013. DNA methylation mediated control of gene expression is critical for development of crown gall tumors. *PLoS Genet* **9**: e1003267.

González I, Martínez L, Rakitina DV, Lewsey MG, Atencio FA, Llave C, Kalinina NO, Carr JP, Palukaitis P, Canto T. 2010. Cucumber Mosaic Virus 2b Protein Subcellular Targets and Interactions: Their Significance to RNA Silencing Suppressor Activity. *Molecular Plant-Microbe Interactions®* **23**: 294-303.

Grabowska A, Bhat SS, Smoczynska A, Bielewicz D, Jarmolowski A, Kulinska ZS. 2020. Regulation of Plant microRNA Biogenesis. In *Plant microRNAs Shaping Development and Enviromental Responses*, pp. 3-24. Springer.

Greenberg MV, Deleris A, Hale CJ, Liu A, Feng S, Jacobsen SE. 2013. Interplay between active chromatin marks and RNA-directed DNA methylation in Arabidopsis thaliana. *PLoS Genet* **9**: e1003946.

Guo H, Song X, Xie C, Huo Y, Zhang F, Chen X, Geng Y, Fang R. 2013. Rice yellow stunt rhabdovirus protein 6 suppresses systemic RNA silencing by blocking RDR6-mediated secondary siRNA synthesis. *Mol Plant Microbe Interact* **26**: 927-936.

H.S. G, Ding SW. 2002. A viral protein inhibits the long range signaling activity of the gene silencing signal. *EMBO J* **21**: 398-407.

Halter T, Wang J, Amesefe D, Lastrucci E, Charvin M, Singla Rastogi M, Navarro L. 2021. The Arabidopsis active demethylase ROS1 cis-regulates defence genes by erasing DNA methylation at promoter-regulatory regions. *eLife* **10**.

Hamera S, Song X, Su L, Chen X, Fang R. 2012. Cucumber mosaic virus suppressor 2b binds to AGO4-related small RNAs and impairs AGO4 activities. *The Plant Journal* **69**: 104-115.

Hanley-Bowdoin L, Bejarano ER, Robertson D, Mansoor S. 2013. Geminiviruses: masters at redirecting and reprogramming plant processes. *Nat Rev Microbiol* **11**: 777-788.

Harvey JJ, Lewsey MG, Patel K, Westwood J, Heimstadt S, Carr JP, Baulcombe DC. 2011. An antiviral defense role of AGO2 in plants. *PLoS One* **6**: e14639.

Herranz MC, Navarro JA, Sommen E, Pallas V. 2015. Comparative analysis among the small RNA populations of source, sink and conductive tissues in two different plant-virus pathosystems. *BMC Genomics* **16**: 117.

Hewezi T, Howe P, Maier TR, Baum TJ. 2008. Arabidopsis Small RNAs and Their Targets During Cyst Nematode Parasitism. *Molecular Plant Microbe Interactions* **21**: 1622-1634.

Hewezi T, Lane T, Piya S, Rambani A, Rice JH, Staton M. 2017. Cyst Nematode Parasitism Induces Dynamic Changes in the Root Epigenome. *Plant Physiol* **174**: 405-420.

Hewezi T, Maier TR, Nettleton D, Baum TJ. 2012. The Arabidopsis microRNA396-GRF1/GRF3 regulatory module acts as a developmental regulator in the reprogramming of root cells during cyst nematode infection. *Plant Physiol* **159**: 321-335.

Hewezi T, Piya S, Qi M, Balasubramaniam M, Rice JH, Baum TJ. 2016. Arabidopsis miR827 mediates post-transcriptional gene silencing of its ubiquitin E3 ligase target gene in the syncytium of the cyst nematode Heterodera schachtii to enhance susceptibility. *Plant J* **88**: 179-192.

Hirochika H, Okamoto H, Kakutani T. 2000. Silencing of Retrotransposons in Arabidopsis and Reactivation by the ddm1 Mutation. *The Plant Cell* **12**: 357-368.

Hohn T, Vazquez F. 2011. RNA silencing pathways of plants: Silencing and its suppression by plant DNA viruses. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1809**: 588-600.

Holliday R. 1987. The Inheritance of Epigenetic Defects. **238**: 163-170.

Holliday R. 1994. Epigenetics: An overview. *Dev Genet* **15**: 453-457.

Hou Y, Zhai Y, Feng L, Karimi HZ, Rutter BD, Zeng L, Choi DS, Zhang B, Gu W, Chen X et al. 2019. A Phytophthora Effector Suppresses Trans-Kingdom RNAi to Promote Disease Susceptibility. *Cell Host Microbe* **25**: 153-165 e155.

Huang CY, Wang H, Hu P, Hamby R, Jin H. 2019. Small RNAs - Big Players in Plant-Microbe Interactions. *Cell Host Microbe* **26**: 173-182.

Huot B, Yao J, Montgomery BL, He SY. 2014. Growth-defense tradeoffs in plants: a balancing act to optimize fitness. *Mol Plant* **7**: 1267-1287.

Iki T, Schopp M-A, Voinnet O. 2017. Biochemical and genetic functional dissection of the P38 viral suppressor of RNA silencing. *RNA* **23**: 639-654.

Inagaki S, Miura-Kamio A, Nakamura Y, Lu F, Cui X, Cao X, Kimura H, Saze H, Kakutani T. 2010. Autocatalytic differentiation of epigenetic modifications within the Arabidopsis genome. *EMBO J* **29**: 3496-3506.

Iwasaki M, Paszkowski J. 2014. Epigenetic memory in plants. *EMBO J* **33**: 1987-1998.

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

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

Jacquemond M. 2012. Cucumber mosaic virus. *Adv Virus Res* **84**: 439-504.

Jagger IC. 1916. *Phytopathology* **6**: 148-151.

Jasencakova Z, Soppe WJJ, Meister A, Gernand D, Turner BM, Schubert I. 2003. Histone modifications in Arabidopsis– high methylation of H3 lysine 9 is dispensable for constitutive heterochromatin. *The Plant Journal* **33**: 471-480.

Jaubert-Possamai S, Noureddine Y, Favery B. 2019. MicroRNAs, New Players in the Plant-Nematode Interaction. *Front Plant Sci* **10**: 1180.

Ji L, Chen X. 2012. Regulation of small RNA stability: methylation and beyond. *Cell Res* **22**: 624-636.

Jiang D, Yang W, He Y, Amasino RM. 2007. Arabidopsis relatives of the human lysine-specific Demethylase1 repress the expression of FWA and FLOWERING LOCUS C and thus promote the floral transition. *Plant Cell* **19**: 2975-2987.

Johnson L, Mollah S, Garcia BA, Muratore TL, Shabanowitz J, Hunt DF, Jacobsen SE. 2004. Mass spectrometry analysis of Arabidopsis histone H3 reveals distinct combinations of post-translational modifications. *Nucleic Acids Res* **32**: 6511-6518.

Johnson LM, Cao X, Jacobsen SE. 2002. Interplay between Two Epigenetic Marks - DNA Methylation and Histone H3 Lysine 9 Methylation. *Current Biology* **12**: 1360-1367.

Johnson LM, Du J, Hale CJ, Bischof S, Feng S, Chodavarapu RK, Zhong X, Marson G, Pellegrini M, Segal DJ et al. 2014. SRA- and SET-domain-containing proteins link RNA polymerase V occupancy to DNA methylation. *Nature* **507**: 124-128.

Johnson LM, Law JA, Khattar A, Henderson IR, Jacobsen SE. 2008. SRA-domain proteins required for DRM2-mediated de novo DNA methylation. *PLoS Genet* **4**: e1000280.

Jones AL, Thomas CL, Maule AJ. 1998. De novo methylation and co-suppression induced by a cytoplasmically replicating plant RNA virus. *EMBO J* **17**: 6385-6393.

Kankel MK, Ramsey DE, Stokes TL, Flowers SK, Haag JR, Jeddeloh JA, Riddle NC, Verbsky ML, Richards EJ. 2003. Arabidopsis MET1 Cytosine Methyltransferase Mutants. *Genetics* **163**: 1109-1122.

Karran RA, Sanfacon H. 2014. Tomato ringspot virus coat protein binds to ARGONAUTE 1 and suppresses the translation repression of a reporter gene. *Mol Plant Microbe Interact* **27**: 933-943.

Katiyar-Agarwal S, Gao S, Vivian-Smith A, Jin H. 2007. A novel class of bacteria-induced small RNAs in Arabidopsis. *Genes Dev* **21**: 3123-3134.

Katiyar-Agarwal S, Jin H. 2010. Role of small RNAs in host-microbe interactions. *Annu Rev Phytopathol* **48**: 225-246.

Katiyar-Agarwal S, Morgan R, Dahlbeck D, Borsani O, Villegas A, Zhu J-K, Staskawicz BJ, Jin H. 2006. A pathogen-inducible endogenous siRNA in plant immunity. *PNAS* **103**: 18002-18007.

Kenesi E, Carbonell A, Lozsa R, Vertessy B, Lakatos L. 2017. A viral suppressor of RNA silencing inhibits ARGONAUTE 1 function by precluding target RNA binding to pre-assembled RISC. *Nucleic Acids Res* **45**: 7736-7750.

Kennedy JS, Ibbotson A, Booth CO. 1950. The distribution of aphid infestation in relation to leaf age. *Annals of Applied Biology* **37**: 651-679.

Kephalogianni TE, Tsitsipis JA, Margaritopoulos JT, Zintzaras E, Delon R, Blanco Martin I, Schwaer W. 2002. Variation in the life cycle and morphology of the tobacco host-race of Myzus persicae (Hemiptera: Aphididae) in relation to its geographical distribution. *Bull Entomol Res* **92**: 301-307.

Kettles GJ, Drurey C, Schoonbeek HJ, Maule AJ, Hogenhout SA. 2013. Resistance of Arabidopsis thaliana to the green peach aphid, Myzus persicae, involves camalexin and is regulated by microRNAs. *New Phytol* **198**: 1178-1190.

Khan SA, Li MZ, Wang SM, Yin HJ. 2018. Revisiting the Role of Plant Transcription Factors in the Battle against Abiotic Stress. *Int J Mol Sci* **19**.

Kim J, Kim JH, Richards EJ, Chung KM, Woo HR. 2014. Arabidopsis VIM Proteins Regulate Epigenetic Silencing by Modulating DNA Methylation and Histone Modification in Cooperation with MET1. *Molecular Plant* **7**: 1470-1485.

Kim JH, Jander G. 2007. Myzus persicae (green peach aphid) feeding on Arabidopsis induces the formation of a deterrent indole glucosinolate. *Plant J* **49**: 1008-1019.

Kim JM, To TK, Ishida J, Morosawa T, Kawashima M, Matsui A, Toyoda T, Kimura H, Shinozaki K, Seki M. 2008. Alterations of lysine modifications on the histone H3 N-tail under drought stress conditions in Arabidopsis thaliana. *Plant Cell Physiol* **49**: 1580-1588.

Kim JM, To TK, Matsui A, Tanoi K, Kobayashi NI, Matsuda F, Habu Y, Ogawa D, Sakamoto T, Matsunaga S et al. 2017. Acetate-mediated novel survival strategy against drought in plants. *Nat Plants* **3**: 17097.

Kinoshita Y, Saze H, Kinoshita T, Miura A, Soppe WJ, Koornneef M, Kakutani T. 2007. Control of FWA gene silencing in Arabidopsis thaliana by SINE-related direct repeats. *Plant J* **49**: 38-45.

Kohler C, Villar CB. 2008. Programming of gene expression by Polycomb group proteins. *Trends Cell Biol* **18**: 236-243.

Kong W, Xia X, Wang Q, Liu LW, Zhang S, Ding L, Liu A, La H. 2020. Impact of DNA Demethylases on the DNA Methylation and Transcription of Arabidopsis NLR Genes. *Front Genet* **11**: 460.

Kranner I, Minibayeva FV, Beckett RP, Seal CE. 2010. What is stress? Concepts, definitions and applications in seed science. *New Phytol* **188**: 655-673.

Krenz B, Bronikowski A, Lu X, Ziebell H, Thompson JR, Perry KL. 2015. Visual monitoring of Cucumber mosaic virus infection in Nicotiana benthamiana following transmission by the aphid vector Myzus persicae. *J Gen Virol* **96**: 2904-2912.

Kurihara Y, Takashi Y, Watanabe Y. 2006. The interaction between DCL1 and HYL1 is important for efficient and precise processing of pri-miRNA in plant microRNA biogenesis. *RNA* **12**: 206-212.

Kusnierczyk A, Winge P, Midelfart H, Armbruster WS, Rossiter JT, Bones AM. 2007. Transcriptional responses of Arabidopsis thaliana ecotypes with different glucosinolate profiles after attack by polyphagous Myzus persicae and oligophagous Brevicoryne brassicae. *J Exp Bot* **58**: 2537-2552.

Lakatos L, Csorba T, Pantaleo V, Chapman EJ, Carrington JC, Liu YP, Dolja VV, Calvino LF, Lopez-Moya JJ, Burgyan J. 2006. Small RNA binding is a common strategy to suppress RNA silencing by several viral suppressors. *EMBO J* **25**: 2768-2780.

Laloum T, Martin G, Duque P. 2018. Alternative Splicing Control of Abiotic Stress Responses. *Trends Plant Sci* **23**: 140-150.

Lanet E, Delannoy E, Sormani R, Floris M, Brodersen P, Crete P, Voinnet O, Robaglia C. 2009. Biochemical evidence for translational repression by Arabidopsis microRNAs. *Plant Cell* **21**: 1762-1768.

Law JA, Ausin I, Johnson LM, Vashisht AA, Zhu JK, Wohlschlegel JA, Jacobsen SE. 2010. A protein complex required for polymerase V transcripts and RNA- directed DNA methylation in Arabidopsis. *Curr Biol* **20**: 951-956.

Law JA, Du J, Hale CJ, Feng S, Krajewski K, Palanca AM, Strahl BD, Patel DJ, Jacobsen SE. 2013. Polymerase IV occupancy at RNA-directed DNA methylation sites requires SHH1. *Nature* **498**: 385-389.

Law JA, Jacobsen SE. 2010. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet* **11**: 204-220.

Le T-N, Schumann U, Smith NA, Tiwari S, Khang Au PC, Zhu Q-H, Taylor JM, Kazan K, Llewellyn DJ, Zhang R et al. 2014. DNA demethylases target promoter transposable elements to positively regulate stress responsive genes in Arabidopsis. *Genome Biology* **15**.

Lee HJ, Park YJ, Kwak KJ, Kim D, Park JH, Lim JY, Shin C, Yang KY, Kang H. 2015. MicroRNA844-Guided Downregulation of Cytidinephosphate Diacylglycerol Synthase3 (CDS3) mRNA Affects the Response of Arabidopsis thaliana to Bacteria and Fungi. *Mol Plant Microbe Interact* **28**: 892-900.

Lee K, Park OS, Seo PJ. 2018. JMJ30-mediated demethylation of H3K9me3 drives tissue identity changes to promote callus formation in Arabidopsis. *Plant J* **95**: 961-975.

Lee S, Fu F, Xu S, Lee SY, Yun D-J, Mengiste T. 2016. Global regulation of plant immunity by histone lysine methyl transferases. *The Plant Cell* doi:10.1105/tpc.16.00012.

Leonetti P, Ghasemzadeh A, Consiglio A, Gursinsky T, Behrens SE, Pantaleo V. 2021. Endogenous activated small interfering RNAs in virus-infected Brassicaceae crops show a common host gene-silencing pattern affecting photosynthesis and stress response. *New Phytol* **229**: 1650-1664.

Li C, Chen C, Gao L, Yang S, Nguyen V, Shi X, Siminovitch K, Kohalmi SE, Huang S, Wu K et al. 2015a. The Arabidopsis SWI2/SNF2 chromatin Remodeler BRAHMA regulates polycomb function during vegetative development and directly activates the flowering repressor gene SVP. *PLoS Genet* **11**: e1004944.

Li F, Pignatta D, Bendix C, Brunkard JO, Cohn MM, Tung J, Sun H, Kumar P, Baker B. 2012. MicroRNA regulation of plant innate immune receptors. *Proc Natl Acad Sci U S A* **109**: 1790-1795.

Li F, Wang A. 2019. RNA-Targeted Antiviral Immunity: More Than Just RNA Silencing. *Trends in Microbiology* **27**: 792-805.

Li Y, Mukherjee I, Thum KE, Tanurdzic M, Katari MS, Obertello M, Edwards MB, McCombie WR, Martienssen RA, Coruzzi GM. 2015b. The histone methyltransferase SDG8 mediates the epigenetic modification of light and carbon responsive genes in plants. *Genome Biol* **16**: 79.

Li Y, Zhang Q, Zhang J, Wu L, Qi Y, Zhou JM. 2010. Identification of microRNAs involved in pathogen-associated molecular pattern-triggered plant innate immunity. *Plant Physiol* **152**: 2222-2231.

Lindroth AM, Shultis D, Jasencakova Z, Fuchs J, Johnson L, Schubert D, Patnaik D, Pradhan S, Goodrich J, Schubert I et al. 2004. Dual histone H3 methylation marks at lysines 9 and 27 required for interaction with CHROMOMETHYLASE3. *EMBO J* **23**: 4286-4296.

Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR. 2008. Highly integrated single-base resolution maps of the epigenome in Arabidopsis. *Cell* **133**: 523-536.

Liu C, Lu F, Cui X, Cao X. 2010. Histone methylation in higher plants. *Annu Rev Plant Biol* **61**: 395-420.

Liu F, Quesada V, Crevillen P, Baurle I, Swiezewski S, Dean C. 2007. The Arabidopsis RNA-binding protein FCA requires a lysine-specific demethylase 1 homolog to downregulate FLC. *Mol Cell* **28**: 398-407.

Liu J, Cheng X, Liu D, Xu W, Wise R, Shen QH. 2014a. The miR9863 family regulates distinct Mla alleles in barley to attenuate NLR receptor-triggered disease resistance and cell-death signaling. *PLoS Genet* **10**: e1004755.

Liu R, Lang Z. 2020. The mechanism and function of active DNA demethylation in plants. *J Integr Plant Biol* **62**: 148-159.

Liu X, Luo M, Wu K. 2012a. Epigenetic interplay of histone modifications and DNA methylation mediated by HDA6. *Plant Signaling & Behaviour* **7**: 633-635.

Liu X, Yang S, Yu CW, Chen CY, Wu K. 2016. Histone Acetylation and Plant Development. *Enzymes* **40**: 173-199.

Liu X, Yu C-W, Duan J, Luo M, Wang K, Tian G, Cui Y, Wu K. 2012b. HDA6 Directly Interacts with DNA Methyltransferase MET1 and Maintains Transposable Element Silencing in Arabidopsis *Plant Physiology* **158**: 119-129.

Liu Y, Teng C, Xia R, Meyers BC. 2020. PhasiRNAs in Plants: Their Biogenesis, Genic Sources, and Roles in Stress Responses, Development, and Reproduction. *Plant Cell* **32**: 3059-3080.

Liu ZW, Shao CR, Zhang CJ, Zhou JX, Zhang SW, Li L, Chen S, Huang HW, Cai T, He XJ. 2014b. The SET domain proteins SUVH2 and SUVH9 are required for Pol V occupancy at RNA-directed DNA methylation loci. *PLoS Genet* **10**: e1003948.

Lopez A, Ramirez V, Garcia-Andrade J, Flors V, Vera P. 2011. The RNA silencing enzyme RNA polymerase v is required for plant immunity. *PLoS Genet* **7**: e1002434.

Lopez Sanchez A, Stassen JH, Furci L, Smith LM, Ton J. 2016. The role of DNA (de)methylation in immune responsiveness of Arabidopsis. *Plant J* **88**: 361-374.

Lu F, Cui X, Zhang S, Jenuwein T, Cao X. 2011. Arabidopsis REF6 is a histone H3 lysine 27 demethylase. *Nat Genet* **43**: 715-719.

Maizel A, Markmann K, Timmermans M, Wachter A. 2020. To move or not to move: roles and specificity of plant RNA mobility. *Curr Opin Plant Biol* **57**: 52-60.

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

Mallory AC, Ely L, Smith TH, Marathe R, Anandalakshmi R, Fagard M, Vaucheret H, Pruss G, Bowman L, Vance VB. 2001. HC-Pro Suppression of Transgene Silencing Eliminates the Small RNAs but Not Transgene Methylation or the Mobile Signal. *The Plant Cell* **13**: 571-583.

Malpica-Lopez N, Rajeswaran R, Beknazariants D, Seguin J, Golyaev V, Farinelli L, Pooggin MM. 2018. Revisiting the Roles of Tobamovirus Replicase Complex Proteins in Viral Replication and Silencing Suppression. *Mol Plant Microbe Interact* **31**: 125-144.

Mann KS, Johnson KN, Carroll BJ, Dietzgen RG. 2016. Cytorhabdovirus P protein suppresses RISC-mediated cleavage and RNA silencing amplification in planta. *Virology* **490**: 27-40.

Mari-Ordonez A, Marchais A, Etcheverry M, Martin A, Colot V, Voinnet O. 2013. Reconstructing de novo silencing of an active plant retrotransposon. *Nat Genet* **45**: 1029-1039.

Martinez G, Castellano M, Tortosa M, Pallas V, Gomez G. 2014. A pathogenic non-coding RNA induces changes in dynamic DNA methylation of ribosomal RNA genes in host plants. *Nucleic Acids Res* **42**: 1553-1562.

Matzke MA, Mosher RA. 2014. RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nat Rev Genet* **15**: 394-408.

McCue AD, Panda K, Nuthikattu S, Choudury SG, Thomas EN, Slotkin RK. 2015. ARGONAUTE 6 bridges transposable element mRNA-derived siRNAs to the establishment of DNA methylation. *EMBO J* **34**: 20-35.

Metzlaff M, O'Dell M, Cluster PD, Flavell RB. 1997. RNA-Mediated RNA Degradation and Chalcone Synthase A Silencing in Petunia. *Cell* **88**: 845-854.

Meyers BC, Kozik A, Griego A, Kuang H, Michelmore RW. 2003. Genome-wide analysis of NBS-LRR-encoding genes in Arabidopsis. *Plant Cell* **15**: 809-834.

Mi S, Cai T, Hu Y, Chen Y, Hodges E, Ni F, Wu L, Li S, Zhou H, Long C et al. 2008. Sorting of small RNAs into Arabidopsis argonaute complexes is directed by the 5' terminal nucleotide. *Cell* **133**: 116-127.

Millar AA, Lohe A, Wong G. 2019. Biology and Function of miR159 in Plants. *Plants (Basel)* **8**.

Mittler TE, Dadd RH. 1966. Food and Wing Determination in Myzus persicae (Homoptera, Aphidae). *Annals of the Entomological Society of America* **59**: 1162–1166.

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

Mochizuki T, Ohki ST. 2012. Cucumber mosaic virus: viral genes as virulence determinants. *Mol Plant Pathol* **13**: 217-225.

Molitor AM, Latrasse D, Zytnicki M, Andrey P, Houba-Herin N, Hachet M, Battail C, Del Prete S, Alberti A, Quesneville H et al. 2016. The Arabidopsis hnRNP-Q Protein LIF2 and the PRC1 Subunit LHP1 Function in Concert to Regulate the Transcription of Stress-Responsive Genes. *Plant Cell* **28**: 2197-2211.

Morel JB, Godon C, Mourrain P, Beclin C, Boutet S, Feuerbach F, Proux F, Vaucheret H. 2002. Fertile hypomorphic ARGONAUTE (ago1) mutants impaired in post-transcriptional gene silencing and virus resistance. *Plant Cell* **14**: 629-639.

Morgado L. 2020. Diversity and types of small RNA. In *Plant Small RNA*, doi:10.1016/b978-0-12-817112-7.00002-x, pp. 37-47.

Mossop DW, Francki RIB. 1977. Association of RNA 3 with aphid transmission of cucumber mosaic virus. *Virology* **81**: 177-181.

Napoli CA, Lemieux C, Jorgensen RA. 1990. Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. *The Plant Cell* **2**: 279-289.

Navarro L, Dunoyer P, Jay F, Arnold B, Dharmasiri N, Estelle M, Voinnet O, Jones JDG. 2006. A Plant miRNA Contributes to Antibacterial Resistance by Repressing Auxin Signaling. *Science* **312**: 436-439.

Ng J, Perry KL. 1999. Stability of the aphid transmission phenotype in cucumber mosaic virus. *Plant Pathology* **48**: 388-394.

Niu D, Lii YE, Chellappan P, Lei L, Peralta K, Jiang C, Guo J, Coaker G, Jin H. 2016. miRNA863-3p sequentially targets negative immune regulator ARLPKs and positive regulator SERRATE upon bacterial infection. *Nat Commun* **7**: 11324.

Noh B, Lee S-H, Kim H-J, Yi G, Shin E-A, Lee M, Jung K-J, Doyle MR, Amasino RM, Noh Y-S. 2004. Divergent Roles of a Pair of Homologous Jumonji/Zinc-Finger–Class Transcription Factor Proteins in the Regulation of Arabidopsis Flowering Time. *The Plant Cell* **16**: 2601-2613.

Nuthikattu S, McCue AD, Panda K, Fultz D, DeFraia C, Thomas EN, Slotkin RK. 2013. The initiation of epigenetic silencing of active transposable elements is triggered by RDR6 and 21-22 nucleotide small interfering RNAs. *Plant Physiol* **162**: 116-131.

O'Malley RC, Huang SC, Song L, Lewsey MG, Bartlett A, Nery JR, Galli M, Gallavotti A, Ecker JR. 2016. Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. *Cell* **165**: 1280-1292.

Özgökçe MS, Chi H, Atlıhan R, Kara H. 2018. Demography and population projection of Myzus persicae (Sulz.) (Hemiptera: Aphididae) on five pepper (Capsicum annuum L.) cultivars. *Phytoparasitica* **46**: 153-167.

Palma K, Thorgrimsen S, Malinovsky FG, Fiil BK, Nielsen HB, Brodersen P, Hofius D, Petersen M, Mundy J. 2010. Autoimmunity in Arabidopsis acd11 is mediated by epigenetic regulation of an immune receptor. *PLoS Pathog* **6**: e1001137.

Palukaitis P, García-Arenal F. 2003. Cucumoviruses. In *Advances in Virus Research*, Vol 62, pp. 241-323. Academic Press.

Panda K, Ji L, Neumann DA, Daron J, Schmitz RJ, Slotkin RK. 2016. Full-length autonomous transposable elements are preferentially targeted by expression-dependent forms of RNA-directed DNA methylation. *Genome Biol* **17**: 170.

Pandey R, Müller A, Napoli CA, Selinger DA, Pikaard CS, Richards EJ, Bender J, Mount DW, Jorgensen RA. 2002. Analysis of histone acetyltransferase and histone deacetylase families of Arabidopsis thaliana suggests functional diversification of chromatin modification among multicellular eukaryotes. *Nucleic Acid Research* **30**: 5036-5055.

Parent JS, Cahn J, Herridge RP, Grimanelli D, Martienssen RA. 2021. Small RNAs guide histone methylation in Arabidopsis embryos. *Genes Dev* **35**: 841-846.

Park MY, Wu G, Gonzalez-Sulser A, Vaucheret H, Poethig RS. 2005. Nuclear processing and export of microRNAs in Arabidopsis. *PNAS* **102**: 3691-3696.

Park YJ, Lee HJ, Kwak KJ, Lee K, Hong SW, Kang H. 2014. MicroRNA400-guided cleavage of Pentatricopeptide repeat protein mRNAs Renders Arabidopsis thaliana more susceptible to pathogenic bacteria and fungi. *Plant Cell Physiol* **55**: 1660-1668.

Pavet V, Quintero C, Cecchini NM, Rosa AL, Alvarez ME. 2006. Arabidopsis Displays Centromeric DNA Hypomethylation and Cytological Alterations of Heterochromatin Upon Attack by Pseudomonas syringae. *Mol Plant Microbe Interact* **19**: 577-587.

Perry KL, Zhang L, Shintaku MH, Palukaitis P. 1994. Mapping Determinants in Cucumber Mosaic Virus for Transmission by Aphis gossypii. *Virology* **205**: 591-595.

Pertermann R, Tamilarasan S, Gursinsky T, Gambino G, Schuck J, Weinholdt C, Lilie H, Grosse I, Golbik RP, Pantaleo V et al. 2018. A Viral Suppressor Modulates the Plant Immune Response Early in Infection by Regulating MicroRNA Activity. *mBio* **9**.

Pitzalis N, Amari K, Graindorge S, Pflieger D, Donaire L, Wassenegger M, Llave C, Heinlein M. 2020. Turnip mosaic virus in oilseed rape activates networks of sRNA-mediated interactions between viral and host genomes. *Commun Biol* **3**: 702.

Piya S, Kihm C, Rice JH, Baum TJ, Hewezi T. 2017. Cooperative Regulatory Functions of miR858 and MYB83 during Cyst Nematode Parasitism. *Plant Physiol* **174**: 1897-1912.

Pontier D, Picart C, Roudier F, Garcia D, Lahmy S, Azevedo J, Alart E, Laudie M, Karlowski WM, Cooke R et al. 2012. NERD, a plant-specific GW protein, defines an additional RNAi-dependent chromatin-based pathway in Arabidopsis. *Mol Cell* **48**: 121-132.

Pruss GJ, Nester EW, Vance V. 2008. Infiltration with Agrobacterium tumefaciens induces host defense and development-dependent responses in the infiltrated zone. *Molecular Plant-Microbe Interactions* **21**: 1528-1538.

Qi Y, He X, Wang XJ, Kohany O, Jurka J, Hannon GJ. 2006. Distinct catalytic and non-catalytic roles of ARGONAUTE4 in RNA-directed DNA methylation. *Nature* **443**: 1008-1012.

Qin C, Li B, Fan Y, Zhang X, Yu Z, Ryabov E, Zhao M, Wang H, Shi N, Zhang P et al. 2017. Roles of Dicer-Like Proteins 2 and 4 in Intra- and Intercellular Antiviral Silencing. *Plant Physiol* **174**: 1067-1081.

Qiu Q, Mei H, Deng X, He K, Wu B, Yao Q, Zhang J, Lu F, Ma J, Cao X. 2019. DNA methylation repels targeting of Arabidopsis REF6. *Nat Commun* **10**: 2063.

Qiu Y, Wu Y, Zhang Y, Xu W, Wang C, Zhu S. 2018. Profiling of small RNAs derived from cucumber mosaic virus in infected Nicotiana benthamiana plants by deep sequencing. *Virus Res* **252**: 1-7.

Qu F, Ye X, Morris TJ. 2008. Arabidopsis DRB4, AGO1, AGO7, and RDR6 participate in a DCL4-initiated antiviral RNA silencing pathway negatively regulated by DCL1. *PNAS* **105**: 14732–14737.

Quadrana L, Colot V. 2016. Plant Transgenerational Epigenetics. *Annu Rev Genet* **50**: 467-491.

Raja P, Sanville BC, Buchmann RC, Bisaro DM. 2008. Viral genome methylation as an epigenetic defense against geminiviruses. *J Virol* **82**: 8997-9007.

Ramesh SV, Yogindran S, Gnanasekaran P, Chakraborty S, Winter S, Pappu HR. 2020. Virus and Viroid-Derived Small RNAs as Modulators of Host Gene Expression: Molecular Insights Into Pathogenesis. *Front Microbiol* **11**: 614231.

Ramirez-Prado JS, Piquerez SJM, Bendahmane A, Hirt H, Raynaud C, Benhamed M. 2018. Modify the Histone to Win the Battle: Chromatin Dynamics in Plant-Pathogen Interactions. *Front Plant Sci* **9**: 355.

Rhee SJ, Watt LG, Bravo AC, Murphy AM, Carr JP. 2020. Effects of the cucumber mosaic virus 2a protein on aphid-plant interactions in Arabidopsis thaliana. *Mol Plant Pathol* **21**: 1248-1254.

Rodriguez-Gandarilla MG, Rodriguez-Negrete EA, Rivera-Bustamante RF. 2020. Superinfection by PHYVV Alters the Recovery Process in PepGMV-Infected Pepper Plants. *Viruses* **12**.

Rodriguez-Negrete EA, Carrillo-Tripp J, Rivera-Bustamante RF. 2009. RNA silencing against geminivirus: complementary action of posttranscriptional gene silencing and transcriptional gene silencing in host recovery. *J Virol* **83**: 1332-1340.

Roossinck MJ. 2001. Cucumber mosaic virus, a model for RNA virus evolution. *Molecular Plant Pathology* **2**: 59-63.

Rosa C, Kuo Y-W, Wuriyanghan H, Falk BW. 2018. RNA Interference Mechanisms and Applications in Plant Pathology. *Annual Review of Phytopathology* **56**: 581-610.

Roudier F, Ahmed I, Berard C, Sarazin A, Mary-Huard T, Cortijo S, Bouyer D, Caillieux E, Duvernois-Berthet E, Al-Shikhley L et al. 2011. Integrative epigenomic mapping defines four main chromatin states in Arabidopsis. *EMBO J* **30**: 1928-1938.

Rougée M, Quadrana L, Zervudacki J, Hure V, Colot V, Navarro L, Deleris A. 2021. Polycomb mutant partially suppresses DNA hypomethylation-associated phenotypes in Arabidopsis. *Life Sci Alliance* **4**.

Ryabov EV, van Wezel R, Walsh J, Hong Y. 2004. Cell-to-Cell, but not long-distance, spread of RNA silencing that is induced in individual epidermal cells. *J Virol* **78**: 3149-3154.

Saleh A, Alvarez-Venegas R, Yilmaz M, Le O, Hou G, Sadder M, Al-Abdallat A, Xia Y, Lu G, Ladunga I et al. 2008. The Highly Similar Arabidopsis Homologs of Trithorax ATX1 and ATX2 Encode Proteins with Divergent Biochemical Functions. *The Plant Cell* **20**: 568-579.

Salvador-Guirao R, Baldrich P, Weigel D, Rubio-Somoza I, San Segundo B. 2018. The MicroRNA miR773 Is Involved in the Arabidopsis Immune Response to Fungal Pathogens. *Mol Plant Microbe Interact* **31**: 249-259.

Sattar S, Addo-Quaye C, Thompson GA. 2016. miRNA-mediated auxin signalling repression during Vat-mediated aphid resistance in Cucumis melo. *Plant Cell Environ* **39**: 1216-1227.

Sattar S, Song Y, Anstead JA, Sunkar R, Thompson GA. 2012. Cucumis melo MicroRNA Expression Profile During Aphid Herbivory in a Resistant and Susceptible Interaction. *Molecular Plant Microbe Interactions* **25**: 839-848.

Saze H, Shiraishi A, Miura A, Kakutani T. 2008. Control of Genic DNA Methylation by a jmjC Domain-Containing Protein in Arabidopsis thaliana. *Science* **319**: 462-465.

Schwach F, Vaistij FE, Jones L, Baulcombe DC. 2005. An RNA-dependent RNA polymerase prevents meristem invasion by potato virus X and is required for the activity but not the production of a systemic silencing signal. *Plant Physiol* **138**: 1842-1852.

Scott H. 1963. *Virology* **20**: 103-106.

Shahbazian MD, Grunstein M. 2007. Functions of site-specific histone acetylation and deacetylation. *Annu Rev Biochem* **76**: 75-100.

Shahid S, Kim G, Johnson NR, Wafula E, Wang F, Coruh C, Bernal-Galeano V, Phifer T, dePamphilis CW, Westwood JH et al. 2018. MicroRNAs from the parasitic plant Cuscuta campestris target host messenger RNAs. *Nature* **553**: 82-85.

Shahzad Z, Moore JD, Zilberman D. 2021. Gene body methylation mediates epigenetic inheritance of plant traits. *bioRxiv* doi:10.1101/2021.03.15.435374.

Shi B, Lin L, Wang S, Guo Q, Zhou H, Rong L, Li J, Peng J, Lu Y, Zheng H et al. 2016. Identification and regulation of host genes related to Rice stripe virus symptom production. *New Phytol* **209**: 1106-1119.

Shimura H, Pantaleo V, Ishihara T, Myojo N, Inaba J, Sueda K, Burgyan J, Masuta C. 2011. A viral satellite RNA induces yellow symptoms on tobacco by targeting a gene involved in chlorophyll biosynthesis using the RNA silencing machinery. *PLoS Pathog* **7**: e1002021.

Sigman MJ, Slotkin RK. 2016. The First Rule of Plant Transposable Element Silencing: Location, Location, Location. *Plant Cell* **28**: 304-313.

Smith NA, Eamens AL, Wang MB. 2011. Viral small interfering RNAs target host genes to mediate disease symptoms in plants. *PLoS Pathog* **7**: e1002022.

Soppe WJJ, Jacobsen SE, Alonso-Blanco C, Jackson JP, Kakutani T, Koornneef M, Peeters AJM. 2000. The Late Flowering Phenotype of fwa Mutants Is Caused by Gain-of-Function Epigenetic Alleles of a Homeodomain Gene. *Molecular Cell* **6**: 791-802.

Soto-Suárez M, Baldrich P, Weigel D, Rubio-Somoza I, San Segundo B. 2017. The Arabidopsis miR396 mediates pathogen-associated molecular pattern-triggered immune responses against fungal pathogens. *Scientific Reports* **7**.

Spoel SH, Koornneef A, Claessens SM, Korzelius JP, Van Pelt JA, Mueller MJ, Buchala AJ, Metraux JP, Brown R, Kazan K et al. 2003. NPR1 modulates cross-talk between salicylate- and jasmonate-dependent defense pathways through a novel function in the cytosol. *Plant Cell* **15**: 760-770.

Stavolone L, Prigigallo MI, Cillo F. 2020. Plant viruses against RNA silencing-based defenses: Strategies and solutions. In *Applied Plant Biotechnology for Improving Resistance to Biotic Stress*, doi:10.1016/b978-0-12-816030-5.00011-2, pp. 225-250.

Stokes TL, Kunkel BN, Richards EJ. 2002. Epigenetic variation in Arabidopsis disease resistance. *Genes Dev* **16**: 171-182.

Stroud H, Do T, Du J, Zhong X, Feng S, Johnson L, Patel DJ, Jacobsen SE. 2014. Non-CG methylation patterns shape the epigenetic landscape in Arabidopsis. *Nat Struct Mol Biol* **21**: 64-72.

Stroud H, Greenberg MV, Feng S, Bernatavichute YV, Jacobsen SE. 2013. Comprehensive analysis of silencing mutants reveals complex regulation of the Arabidopsis methylome. *Cell* **152**: 352-364.

Sutter D, Doerfler W. 1980. Methylation of integrated adenovirus type 12 DNA sequences in transformed cells is inversely correlated with viral gene expression. *PNAS* **77**.

Sutter D, Westphal M, Doerfler W. 1978. Patterns of integration of viral DNA sequences in the genomes of adenovirus type 12-transformed hamster cells. *Cell* **14**: 569-585.

Tang K, Lang Z, Zhang H, Zhu JK. 2016. The DNA demethylase ROS1 targets genomic regions with distinct chromatin modifications. *Nat Plants* **2**: 16169.

Tang X, Hou A, Babu M, Nguyen V, Hurtado L, Lu Q, Reyes JC, Wang A, Keller WA, Harada JJ et al. 2008. The Arabidopsis BRAHMA chromatin-remodeling ATPase is involved in repression of seed maturation genes in leaves. *Plant Physiol* **147**: 1143-1157.

To TK, Kim J-M, Matsui A, Kurihara Y, Morosawa T, Ishida J, Tanaka M, Endo T, Kakutani T, Toyoda T et al. 2011. Arabidopsis HDA6 Regulates Locus-Directed Heterochromatin Silencing in Cooperation with MET1. *PLoS Genetics* **7**.

Tungadi T, Donnelly R, Qing L, Iqbal J, Murphy AM, Pate AE, Cunniffe NJ, Carr JP. 2020. Cucumber mosaic virus 2b proteins inhibit virus-induced aphid resistance in tobacco. *Mol Plant Pathol* **21**: 250-257.

Ueda M, Seki M. 2020. Histone Modifications Form Epigenetic Regulatory Networks to Regulate Abiotic Stress Response. *Plant Physiol* **182**: 15-26.

van der Krol AR, Mur LA, Mol JNM, Stuitje AR. 1990. Flavonoid genes in petunia - addition of a limited number of gene copies may lead to a suppression of gene expression. *The Plant Cell* **2**: 291-299.

Van Emden HF, Eastop VF, Hughes RD, Way MJ. 1969. The ecology of Myzus persicae. *Annual Review of Entomology* **14**.

van Loon LC, Rep M, Pieterse CMJ. 2006. Significance of Inducible Defense-related Proteins in Infected Plants. *Annual Review of Phytopathology* **44**: 135-162.

Várallyay É, Válóczi A, Ágyi Á, Burgyán J, Havelda Z. 2010. Plant virus-mediated induction of miR168 is associated with repression of ARGONAUTE1 accumulation. *The EMBO Journal* **29**: 3507-3519.

Vargason JM, Szittya G, Burgyán J, Hall TMT. 2003. Size Selective Recognition of siRNA by an RNA Silencing Suppressor. *Cell* **115**: 799-811.

Verma V, Ravindran P, Kumar PP. 2016. Plant hormone-mediated regulation of stress responses. *BMC Plant Biol* **16**: 86.

Vogler H, Akbergenov R, Shivaprasad PV, Dang V, Fasler M, Kwon MO, Zhanybekova S, Hohn T, Heinlein M. 2007. Modification of small RNAs associated with suppression of RNA silencing by tobamovirus replicase protein. *J Virol* **81**: 10379-10388.

Waddington CH. 1942. Canalization of development and the inheritance of acquired characters. *Nature* **150**: 563-565.

Waddington CH. 1956. Principles of Embriology. In *Quarterly Journal of Experimental Physiology and Cognate Medical Sciences*, Vol 10, pp. 475-477.

Wang C, Wang C, Xu W, Zou J, Qiu Y, Kong J, Yang Y, Zhang B, Zhu S. 2018. Epigenetic Changes in the Regulation of Nicotiana tabacum Response to Cucumber Mosaic Virus Infection and Symptom Recovery through Single-Base Resolution Methylomes. *Viruses* **10**.

Wang H, Jiao X, Kong X, Hamera S, Wu Y, Chen X, Fang R, Yan Y. 2016. A Signaling Cascade from miR444 to RDR1 in Rice Antiviral RNA Silencing Pathway. *Plant Physiol* **170**: 2365-2377.

Wang J, Mei J, Ren G. 2019. Plant microRNAs: Biogenesis, Homeostasis, and Degradation. *Frontiers in Plant Science* **10**.

Wang XB, Wu Q, Ito T, Cillo F, Li WX, Chen X, Yu JL, Ding SW. 2010. RNAi-mediated viral immunity requires amplification of virus-derived siRNAs in Arabidopsis thaliana. *Proc Natl Acad Sci U S A* **107**: 484-489.

Wang Y, Hu Q, Wu Z, Wang H, Han S, Jin Y, Zhou J, Zhang Z, Jiang J, Shen Y et al. 2017. HISTONE DEACETYLASE 6 represses pathogen defence responses in Arabidopsis thaliana. *Plant Cell Environ* **40**: 2972-2986.

Wassenegger M, Heimes S, Riedel L, Sänger HL. 1994. RNA-directed de novo methylation of genomic sequences in plants. *Cell* **76**: 567-576.

Watt LG, Crawshaw S, Rhee SJ, Murphy AM, Canto T, Carr JP. 2020. The cucumber mosaic virus 1a protein regulates interactions between the 2b protein and ARGONAUTE 1 while maintaining the silencing suppressor activity of the 2b protein. *PLoS Pathog* **16**: e1009125.

Weiberg A, Wang M, Lin F-M, Zhao H, Zhang Z, Kaloshian I, Huang H-D, Jin H. 2013. Fungal Small RNAs Suppress Plant Immunity by Hijacking Host RNA Interference Pathways. *Science* **342**: 118-123.

Whittaker C, Dean C. 2017. The FLC Locus: A Platform for Discoveries in Epigenetics and Adaptation. *Annu Rev Cell Dev Biol* **33**: 555-575.

Wierzbicki AT, Cocklin R, Mayampurath A, Lister R, Rowley MJ, Gregory BD, Ecker JR, Tang H, Pikaard CS. 2012. Spatial and functional relationships among Pol V-associated loci, Pol IV-dependent siRNAs, and cytosine methylation in the Arabidopsis epigenome. *Genes Dev* **26**: 1825-1836.

Wingard SA. 1928. Hosts and symptoms of ring spot, a virus disease of plants. *Journal of Agricultural Research* **37**: 127-153.

Wollmann H, Stroud H, Yelagandula R, Tarutani Y, Jiang D, Jing L, Jamge B, Takeuchi H, Holec S, Nie X et al. 2017. The histone H3 variant H3.3 regulates gene body DNA methylation in Arabidopsis thaliana. *Genome Biol* **18**: 94.

Wu C-T, Morris JR. 2001. Genes, Genetics, and Epigenetics: A Correspondence. *Science* **293**: 1103-1105.

Wu J, Yang Z, Wang Y, Zheng L, Ye R, Ji Y, Zhao S, Ji S, Liu R, Xu L et al. 2015. Viral-inducible Argonaute18 confers broad-spectrum virus resistance in rice by sequestering a host microRNA. *Elife* **4**.

Wu K, Zhang L, Zhou C, Yu CW, Chaikam V. 2008. HDA6 is required for jasmonate response, senescence and flowering in Arabidopsis. *J Exp Bot* **59**: 225-234.

Xia R, Xu J, Arikit S, Meyers BC. 2015. Extensive Families of miRNAs and PHAS Loci in Norway Spruce Demonstrate the Origins of Complex phasiRNA Networks in Seed Plants. *Mol Biol Evol* **32**: 2905-2918.

Xie Z, Fan B, Chen C, Chen Z. 2001. An important role of an inducible RNA-dependent RNA polymerase in plant antiviral defense. *Proc Natl Acad Sci U S A* **98**: 6516-6521.

Xu J, Liu D, Zhang Y, Wang Y, Han C, Li D, Yu JL, Wang XB. 2016. Improved Pathogenicity of a Beet Black Scorch Virus Variant by Low Temperature and Co-infection with Its Satellite RNA. *Front Microbiol* **7**: 1771.

Yang DL, Zhang G, Wang L, Li J, Xu D, Di C, Tang K, Yang L, Zeng L, Miki D et al. 2018. Four putative SWI2/SNF2 chromatin remodelers have dual roles in regulating DNA methylation in Arabidopsis. *Cell Discov* **4**: 55.

Yang L, Liu Z, Lu F, Dong A, Huang H. 2006. SERRATE is a novel nuclear regulator in primary microRNA processing in Arabidopsis. *Plant J* **47**: 841-850.

Yang R, Zheng Z, Chen Q, Yang L, Huang H, Miki D, Wu W, Zeng L, Liu J, Zhou JX et al. 2017. The developmental regulator PKL is required to maintain correct DNA methylation patterns at RNA-directed DNA methylation loci. *Genome Biol* **18**: 103.

Ye K, Malinina L, Patel DJ. 2003. Recognition of small interfering RNA by a viral suppressor of RNA silencing. *Nature* **426**: 874-878.

Ye R, Chen Z, Lian B, Rowley MJ, Xia N, Chai J, Li Y, He XJ, Wierzbicki AT, Qi Y. 2016. A Dicer-Independent Route for Biogenesis of siRNAs that Direct DNA Methylation in Arabidopsis. *Mol Cell* **61**: 222-235.

Yu A, Lepere G, Jay F, Wang J, Bapaume L, Wang Y, Abraham AL, Penterman J, Fischer RL, Voinnet O et al. 2013. Dynamics and biological relevance of DNA demethylation in Arabidopsis antibacterial defense. *Proc Natl Acad Sci U S A* **110**: 2389-2394.

Yu B, Chapman EJ, Yang Z, Carrington JC, Chen X. 2006. Transgenically expressed viral RNA silencing suppressors interfere with microRNA methylation in Arabidopsis. *FEBS Lett* **580**: 3117-3120.

Yu B, Yang Z, Minakhina S, Yang M, Padgett RW, Steward R, Chen X. 2005. Methylation as a Crucial Step in Plant microRNA Biogenesis. *Science* **307**: 932-935.

Zarreen F, Chakraborty S. 2020. Epigenetic regulation of geminivirus pathogenesis: a case of relentless recalibration of defence responses in plants. *J Exp Bot* **71**: 6890-6906.

Zemach A, Kim MY, Hsieh PH, Coleman-Derr D, Eshed-Williams L, Thao K, Harmer SL, Zilberman D. 2013. The Arabidopsis nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. *Cell* **153**: 193-205.

Zervudacki J, Yu A, Amesefe D, Wang J, Drouaud J, Navarro L, Deleris A. 2018. Transcriptional control and exploitation of an immune-responsive family of plant retrotransposons. *EMBO J* **37**.

Zhai J, Bischof S, Wang H, Feng S, Lee TF, Teng C, Chen X, Park SY, Liu L, Gallego-Bartolome J et al. 2015. A One Precursor One siRNA Model for Pol IV-Dependent siRNA Biogenesis. *Cell* **163**: 445-455.

Zhai J, Jeong DH, De Paoli E, Park S, Rosen BD, Li Y, Gonzalez AJ, Yan Z, Kitto SL, Grusak MA et al. 2011. MicroRNAs as master regulators of the plant NB-LRR defense gene family via the production of phased, trans-acting siRNAs. *Genes Dev* **25**: 2540-2553.

Zhang H, Lang Z, Zhu JK. 2018. Dynamics and function of DNA methylation in plants. *Nat Rev Mol Cell Biol* **19**: 489-506.

Zhang W, Gao S, Zhou X, Chellappan P, Chen Z, Zhou X, Zhang X, Fromuth N, Coutino G, Coffey M et al. 2011a. Bacteria-responsive microRNAs

regulate plant innate immunity by modulating plant hormone networks. *Plant Mol Biol* **75**: 93-105.

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

Zhang X, Clarenz O, Cokus S, Bernatavichute YV, Pellegrini M, Goodrich J, Jacobsen SE. 2007. Whole-genome analysis of histone H3 lysine 27 trimethylation in Arabidopsis. *PLoS Biol* **5**: e129.

Zhang X, Le Roux C, Del Prete S, Boutet-Mercey S, Perreau F, Balagué C, Roby D, Fagard M, Gaudin V. 2014. The hnRNP-Q Protein LIF2 Participates in the Plant Immune Response. *PLoS ONE* **9**.

Zhang X, Yuan YR, Pei Y, Lin SS, Tuschl T, Patel DJ, Chua NH. 2006. Cucumber mosaic virus-encoded 2b suppressor inhibits Arabidopsis Argonaute1 cleavage activity to counter plant defense. *Genes & Development* **20**: 3255-3268.

Zhang X, Zhao H, Gao S, Wang WC, Katiyar-Agarwal S, Huang HD, Raikhel N, Jin H. 2011b. Arabidopsis Argonaute 2 regulates innate immunity via miRNA393(*)-mediated silencing of a Golgi-localized SNARE gene, MEMB12. *Mol Cell* **42**: 356-366.

Zhang Y, Fu Y, Wang Q, Liu X, Li Q, Chen J. 2020. Transcriptome analysis reveals rapid defence responses in wheat induced by phytotoxic aphid Schizaphis graminum feeding. *BMC Genomics* **21**: 339.

Zhao M, Meyers BC, Cai C, Xu W, Ma J. 2015. Evolutionary patterns and coevolutionary consequences of MIRNA genes and microRNA targets triggered by multiple mechanisms of genomic duplications in soybean. *Plant Cell* **27**: 546-562.

Zhao T, Zhan Z, Jiang D. 2019. Histone modifications and their regulatory roles in plant development and environmental memory. *J Genet Genomics* **46**: 467-476.

Zhong X, Du J, Hale CJ, Gallego-Bartolome J, Feng S, Vashisht AA, Chory J, Wohlschlegel JA, Patel DJ, Jacobsen SE. 2014. Molecular mechanism of action of plant DRM de novo DNA methyltransferases. *Cell* **157**: 1050-1060.

Zhong X, Hale CJ, Law JA, Johnson LM, Feng S, Tu A, Jacobsen SE. 2012. DDR complex facilitates global association of RNA polymerase V to promoters and evolutionarily young transposons. *Nat Struct Mol Biol* **19**: 870-875.

Zhong Z, Feng S, Duttke SH, Potok ME, Zhang Y, Gallego-Bartolome J, Liu W, Jacobsen SE. 2021. DNA methylation-linked chromatin accessibility affects genomic architecture in Arabidopsis. *Proc Natl Acad Sci U S A* **118**.

Zhou C, Zhang L, Duan J, Miki B, Wu K. 2005. HISTONE DEACETYLASE19 Is Involved in Jasmonic Acid and Ethylene Signaling of Pathogen Response in Arabidopsis. *The Plant Cell* **17**: 1196-1204.

Zhuo T, Li YY, Xiang HY, Wu ZY, Wang XB, Wang Y, Zhang YL, Li DW, Yu JL, Han CG. 2014. Amino acid sequence motifs essential for P0-mediated

suppression of RNA silencing in an isolate of potato leafroll virus from Inner Mongolia. *Mol Plant Microbe Interact* **27**: 515-527.

Zilberman D, Coleman-Derr D, Ballinger T, Henikoff S. 2008. Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks. *Nature* **456**: 125-129.

Zust T, Agrawal AA. 2017. Trade-Offs Between Plant Growth and Defense Against Insect Herbivory: An Emerging Mechanistic Synthesis. *Annu Rev Plant Biol* **68**: 513-534.

Popular science summary

Opposite to animals, plants are not able to fight, hide or run from stressful situations. They have to deal with unfavorable scenarios from the inside, and one of the most efficient ways to do so is by "turning on and off" specific genes. There are several mechanism to regulate the expression of genes, of which two were the focus of this thesis, epigenetics and RNA silencing. Epigenetics refers to those reversible changes on the DNA molecules that do not affect its sequence, and it comprises two different types of changes. On one hand, histone modifications, which consists of the addition (or removal) of different molecules to the proteins that support the DNA, the histones. On the other hand, DNA methylation, which is the addition (or removal) of a methyl group (a small molecule constituted by one atom of carbon and three hydrogen atoms, $-CH_3$). Each of these different modifications contribute to either the activation or the repression of genes, in other words, to "turn on" or "turn off" those genes. On the other hand, RNA silencing is a mechanism by which RNA molecules are cleaved into smaller fragments to avoid the production of proteins. This is achieved by the action of two important molecules: a group of proteins called Argonautes (AGO) and a special type of RNAs, known as small RNAs (sRNAs), which bind to the AGO proteins to perform their function. This mechanism of slicing RNAs is particularly important during viral infections, as it is a key mechanism to avoid the replication of the virus.

We studied these two regulatory mechanisms in two different stresses, aphid infestation and Cucumber mosaic virus (CMV) infection, both economically important, as they can affect a broad range of crops. Our study of the epigenetic changes during aphid infestation revealed that DNA methylation is an important regulatory mechanism, as it is involved in the activation of genes that have a defense role and, therefore, are necessary during this stressful situation. On the other hand, our work about the changes associated with CMV infection showed how complex the interaction of the components of the plant cell and the virus is. For example, we

demonstrated that when the plant slices the RNA of the virus to avoid its replication by the action of the RNA silencing machinery, the small RNAs that are produced are, in turn, used to avoid the expression of plant genes. Moreover, our study of epigenetic changes occurring during CMV infection revealed that both DNA methylation and histone modifications are modulated in a way that allows the activation of defense genes and the repression of those that are not necessary to cope with the stress.

In summary, the work presented in this thesis provided detailed knowledge about how plants use epigenetics and RNA silencing to regulate the expression of their genes in order to deal with different biotic stresses.

Populärvetenskaplig sammanfattning

Till skillnad från djur kan inte växter fly, gömma sig eller springa från stressfulla situationer. De måste hantera ogynnsamma situationer från inuti sig själva och ett av dem mest effektiva sätten att hantera sådana situationer är att slå "av eller på" specifika gener. Det finns flertalet mekanismer för att reglera genuttryck, av vilka två stycken är i fokus i denna doktorsavhandling, epigenetik och 'RNA silencing'. Epigenetik refererar till dem reversibla förändringar på DNA molekyler som inte påverkar dess sekvens, och innefattar två olika typer av förändringar. Å ena sidan, histon-modifikation,som innefattar addition (eller borttagande) av olika molekyler till proteinet som stödjer DNA:t, histonerna. Å andra sidan, DNA metylering, vilket är additionen (eller borttagandet) av en metylgrupp (en liten molekyl som består av en kolatom och 3 väteatomer, $-CH_3$). Dessa olika metoder bidrar till att antingen aktivera eller undertrycka gener, med andra ord, att "slå på" eller "slå av" dessa gener. Å andra sidan, 'RNA silencing' är en mekanism med vilken RNA molekyler delas till mindre fragment för att undvika att protein produceras. Detta uppnås genom handlingen av två viktiga molekyler, en grupp av protein som kallas Argonautes (AGO) och en speciell typ av RNAs som kallas 'small RNA' (sRNA), som binder AGO proteinet för att utföra deras handling. Den här mekanismen med att klyva RNA är särskilt viktig under virala infektioner eftersom det är en nyckelmekanism för att förhindra att viruset replikerar sig.

Vi har studerat dessa två reglerande mekanismer med två olika stressorer, aphid-infektion och 'cucumber mosaic virus' (CMV) infektion, båda ekonomiskt viktiga då de kan påverka ett stort omfång av grödor. Vår studie på dem epigenetiska förändringarna av aphid-infektion avslöjade att DNA metylering är en viktig reglerande mekanism, eftersom det involverar aktivering av gener som har en defensiv roll och, därför, är nödvändiga under den stressiga situation. Å andra sidan, vårt arbete om förändringarna associerade med CMV infektion visar hur komplext interaktionerna av komponenterna av växtcellen och viruset är. Till exempel så demonstrerar vi att när växten klyver RNA:t av viruset för att förhindra dess

replikation genom handlingen av ‘RNA silencing’ maskineriet, så de ‘small RNA’ som tillverkas används i sin tur för att förhindra växtens genuttryck. Dessutom, vår studie av epigenetiska förändringar under CMV-infektion avslöjade att både DNA metylering och histon-modifikation är modulerade på ett sådant sätt att det tillåts aktivering av defensiva gener och kan undertrycka de som inte är nödvändiga för att hantera stressen.

Sammanfattningsvis, det arbete som presenteras i den här avhandlingen förser detaljerad kunskap om hur växter använder epigenetik och ‘RNA silencing’ för att reglera dess genuttryck för att hantera olika biotiska stressorer.

Resumen de divulgación

Al contrario que los animales, las plantas no pueden luchar, esconderse o pelear contra las situaciones estresantes. Tienen que hacerles frente desde su interior y una de las maneras más eficaces de conseguirlo es “apagando” y “encendiendo” genes específicos. Existen varios mecanismos por los cuales se regula la expresión de genes, de los cuales dos fueron el foco de esta tesis doctoral: la epigenética y el silenciamiento génico por ARN. La epigenética es el estudio de aquellos cambios reversibles de la molécula de ADN que no afectan su secuencia y abarca dos tipos de cambios. Por una parte, las modificaciones de las histonas, que consisten en la adición (o extracción) de diferentes moléculas en las proteínas que sostienen el ADN, llamadas histonas. Por otra parte, la metilación del ADN, que es la adición (o extracción) de un grupo metilo (una molécula pequeña formada por un átomo de carbono y tres átomos de hidrógeno, -CH3). Cada una de estas diferentes modificaciones contribuye a la activación o la represión de genes o, en otras palabras, a “encender” o “apagar” esos genes. Por otra parte, el silenciamiento génico por ARN es un mecanismo por el cual moléculas de ARN son divididas en fragmentos más pequeños para evitar la producción de proteínas. Esto se consigue mediante la acción de dos importantes moléculas: un grupo de proteínas llamadas Argonautas (AGO) y un tipo especial de ARN, conocidos como ARN pequeños, que se unen a las proteínas AGO para ejecutar su función. Este mecanismo de corte de ARN es particularmente importante durante infecciones virales, ya que es un mecanismo clave para evitar la replicación del virus.

En esta tesis, estudiamos estos dos mecanismos de regulación durante dos estreses diferentes, la infestación de áfidos y la infección por el virus mosaico del pepino (CMV), ambos importantes a nivel económico ya que afectan a un amplio rango de cultivos. Nuestro estudio de los cambios epigenéticos durante la infestación de áfidos reveló que la metilación del ADN es un mecanismo regulador importante, ya que está involucrada en la activación de genes que cumplen una función defensiva

y, por tanto, son necesarios durante esta situación estresante. Por otra parte, nuestro trabajo sobre los cambios asociados a la infección por CMV demostró cuán compleja es la interacción entre los componentes de la célula vegetal y el virus. Por ejemplo, demostramos que cuando la planta corta el ARN del virus para evitar su replicación mediante la acción de la maquinaria del silenciamiento por ARN, los ARN pequeños que se producen se utilizan, a su vez, para evitar la expresión de genes de la planta. Además, nuestro trabajo sobre los cambios epigenéticos durante la infección por CMV reveló que tanto la metilación del ADN como las modificaciones de las histonas son moduladas de manera que permitan la activación de los genes de defensa y la represión de aquellos que no son necesarios para combatir el estrés.

En resumen, el trabajo presentado en esta tesis doctoral aportó conocimiento detallado sobre cómo las plantas usan la epigenética y el silenciamiento génico por ARN para regular la expresión de sus genes para lidiar con diferentes estreses bióticos.

Acknowledgements

The pages of this thesis represent more than experiments and data, they represent my biggest adventure (at least so far). They contain long days at the lab, failed experiments, frustration, a few tears… But also, lots of laughs (and gossips), countless fikas, crazy parties, new friends, new hobbies, new cities… And the most important, so much personal growth. I feel so lucky for having had this opportunity, which would have ever happened without the help and the support of so many people.

First of all, I would like to thank my supervisor Germán, as without him I would have not written these pages. Thank you for giving this opportunity, for your mentoring and your endless patience with me. I know it was not always easy, but I am grateful for everything I have learnt from you. I really hope that this experience was as enriching to you as it was for me.

I would like to give a huge thank you to Christina and Anders, for being my co-supervisors, and Charles and Eugene, for being my external examiners. Your participation on my evaluations and your feedback and advice were extremely helpful.

I would also like to write a special thank you to Claudia and her group, who were a constant support throughout all my thesis and were always there when I needed help or advice.

I would like to thank everyone at the Plant Biology department, current and former members, for always trying to make the best working environment possible. Even if the liquid nitrogen tank was not always full or the gel room was a bit messy, I never felt alone working at the department as I always knew I would get help if I needed it.

I would like to give a special thank you to Gesa, Anna and Anis, but unfortunately, I do not think I am allowed to add glitter on these pages, so words will have to do the work. I will forever be grateful for your company and support during this journey, you made it so much more enjoyable (especially the teaching part). I will never forget all the teaching anecdotes, the gossips, the parties, the fikas… I would not be here without you.

I will be always grateful to my former supervisors María Jesús, Mónica and Luis, my academic journey started with them, and it would have never reached this point without their support. Also, I would like to thank Chus, for teaching me so many protocols with a never-ending patience when I was just an undergrad student.

I would like to specially thank María and Fran, for your friendship, inside and outside the lab. I would never forget those days when we used to sing Paulina Rubio's songs in the lab, the coffee breaks and the nights out. And, of course, I would like to also thank Jesús, who played a big role on those nights out and brought me some many happy memories.

An enormous thank you to my friend Tiago, who supports me unconditionally and understands me in a way that not many people do. We have grown so much together, and I truly hope we keep on doing, hopefully in closer cities so we can visit each other more often. I am also very grateful to you for introducing Sara to me. The times you were in Uppsala were always wonderful, your warm personality and your fun plans made those days so special.

A warm thank you to Jose, whose office was always open for me, independently if I wanted to gossip or to complain. Your wise advice and your funny company made the hard days so much easier.

A huge thank you to my friend Silvia, who arrived at the department (and my life) in the right moment. Your support and advice were always there when I needed them, and I know they will always be. I also know that no matter where I am in the world, you will take that plane just to see me. I hope we keep adding pushpins in our world maps together. I would not have finished this thesis without you.

A giant *gracies* to my friend Abra, the first person I met when I arrived in Uppsala. It feels like yesterday when you were waiting for me with an Asturian flag at the bus stop. I will never be able to thank you enough for everything you have done for me these past, from translating my letters to listening to all my dramas.

A massive thank you to my friend Sandra, who, no matter the distance, always has my back and understands me so well. We have been together in this journey since the beginning of it when we were just first-year Biology students. We went from surviving exams, boring lab practices, exhausting field trips… to dealing with statistical analysis, failed experiments, paper rejections… And I hope that whenever I am preparing my retirement papers, you are there by my side to patiently listen to my complains about how complicated they are.

I kept the most special thank you for my partner, Rickard. You became my biggest support through this, especially during this past year that was so intense. I will never be able to thank you for all the support and love you give me every day. Everything was worthy just because it led me to meet you. *Jag älskar dig*.

The most eternal thank you to my family. To my siblings, Paloma and José, and sister-in-law, Noemí, for their unconditional support and, especially, to my parents, Pepa and Charly, for everything they have sacrificed so I could fulfill my goals. I have become the person I am because you always believed on me and always encouraged me to aim higher, not matter how hard it would be.

Finally, I would like to thank that young Luz who fell in love with Biology at high school, worked hard at university and did not quit looking for a PhD position until she finally got one and, once she got it, she did her best. Look, you made it!

Appendix

Table 1. Classification of the cited virus.

	Full name	Genus	Family
CMV	Cucumber mosaic virus	Cucumovirus	Bromoviridaae
CaMV	Cauliflower mosaic virus	Caulimovirus	Caulimoviridae
TMV	Tobacco mosaic virus	Tobamovirus	Virgaviridae
ORMV	Oilseed rape mosaic virus	Tobamovirus	Virgaviridae
RMV	Ribgrass mosaic virus	Tobamovirus	Virgaviridae
TuMV	Turnip mosaic virus	Potyvirus	Potyviridae
TEV	Tobacco etch virus	Potyvirus	Potyviridae
PVY	Potato virus Y	Potyvirus	Potyviridae
PSbMV	Pea seed-borne mosaic virus	Potyvirus	Potyviridae
SPMMV	Sweet potato mild mottle virus	Ipomovirus	Potyviridae
TCV	Tobacco crinkle virus	Carmovirus	Tombusviridae
TBSV	Tomato bushy stunt virus	Tombusvirus	Tombusviridae
CymRSV	Cymbidium ringspot virus	Tombusvirus	Tombusviridae
BBSV	Beet black scorch virus	Necrovirus	Tombusviridae
PVX	Potato virus X	Potexvirus	Alphalflexiviridaea
BYV	Beet yellow virus	Closterovirus	Closteroviridae
ToRSV	Tomato ringspot virus	Nepovirus	Secoviridae
TRSV	Tobacco ringspot virus	Nepovirus	Secoviridae
ADV	Alfalfa dwarf virus	Cytorhabdovirus	Rhabdoviridae
LNYV	Lettuce necrotic yellows virus	Cytorhabdovirus	Rhabdoviridae
RYSV	Rice yellow stunt virus	Nucleorhabdovirus	Rhabdoviridae
BCTV	Beet curly top virus	Curtovirus	Geminiviridae
TGMV	Tomato golden mosaic virus	Begomovirus	Geminiviridae
CaLCuV	Cabbage leaf curly virus	Begomovirus	Geminiviridae
PepGMV	Pepper golden mosaic virus	Begomovirus	Geminiviridae

Table 2. Number of DMRs per context during aphid infestation and CMV infection.

	CG	CHG	CHH
Aphid infestation	691	1123	311
CMV (10 dpi)	296	1978	494
CMV (20 dpi)	1342	3022	557

Table 3. Number of DEGs during aphid infestation and CMV infection.

	Upregulated	Downregulated
Aphid infestation	265	2
CMV (10 dpi)	427	460
CMV (20 dpi)	309	414

Table 4. Number of H3K9me2 and H3K27me3 peaks at 10 and 20 dpi.

	H3K9me2		H3K27me3	
	Gain	Loss	Gain	Loss
Peaks (10 dpi)	892	66	117	223
Peaks (20 dpi)	3053	258	382	1451

I

New Phytologist

Research

Aphid feeding induces the relaxation of epigenetic control and the associated regulation of the defense response in *Arabidopsis*

Maria Luz Annacondia[1]*, Dimitrije Markovic[2,3]*, Juan Luis Reig-Valiente[1], Vassilis Scaltsoyiannes[1,4], Corné M. J. Pieterse[5], Velemir Ninkovic[6], R. Keith Slotkin[7,8] and German Martinez[1]

[1]Department of Plant Biology, Uppsala BioCenter, Swedish University of Agricultural Sciences and Linnean Center for Plant Biology, Uppsala 75007, Sweden; [2]Department of Crop Production Ecology, Swedish University of Agricultural Sciences, Uppsala 75007, Sweden; [3]Faculty of Agriculture, University of Banja Luka, Banja Luka 78000, Bosnia and Herzegovina; [4]Institut de Biologie de Moléculaire des Plantes, UPR 2357 du CNRS, Strasbourg University, Strasbourg 67000, France; [5]Department of Biology, Science4Life, Utrecht University, Utrecht 3584 CS, the Netherlands; [6]Department of Ecology, Swedish University of Agricultural Sciences, Uppsala 75007, Sweden; [7]Donald Danforth Plant Science Center, St Louis, MO 63132, USA; [8]Division of Biological Sciences, University of Missouri-Columbia, Columbia, MO 65021, USA

Author for correspondence:
German Martinez
Email: german.martinez.arias@slu.se

Received: *11 November 2020*
Accepted: *14 January 2021*

New Phytologist (2021)
doi: 10.1111/nph.17226

Key words: aphids, *Arabidopsis*, defense response, epigenetics, transcriptional regulation, transposable elements.

Summary

- Environmentally induced changes in the epigenome help individuals to quickly adapt to fluctuations in the conditions of their habitats.
- We explored those changes in *Arabidopsis thaliana* plants subjected to multiple biotic and abiotic stresses, and identified transposable element (TE) activation in plants infested with the green peach aphid, *Myzus persicae*. We performed a genome-wide analysis mRNA expression, small RNA accumulation and DNA methylation
- Our results demonstrate that aphid feeding induces loss of methylation of hundreds of loci, mainly TEs. This loss of methylation has the potential to regulate gene expression and we found evidence that it is involved in the control of plant immunity genes. Accordingly, mutant plants deficient in DNA and H3K9 methylation (*kyp*) showed increased resistance to *M. persicae* infestation.
- Collectively, our results show that changes in DNA methylation play a significant role in the regulation of the plant transcriptional response and induction of defense response against aphid feeding.

Introduction

While adaptation to long-term environmental changes involves genetic variation, fluctuating stresses are normally coped with through the modulation of the transcription machinery (Lamke & Baurle, 2017). Several mechanisms govern the transcriptional response during stress, including transcription factors (TFs) and epigenetic regulation (Gutzat & Mittelsten, 2012). In eukaryotic organisms, epigenetic modifications of chromatin and DNA are the core of genome stability regulation through the control of transposable element (TE) expression and transposition (Law & Jacobsen, 2010). Epigenetic modifications consist of covalent and reversible marks that are deposited on both the DNA and the histones. DNA methylation constitutes a vital and widespread mark in plant genomes, where it can happened in three different sequence combinations: the symmetric contexts CG and CHG, and the asymmetric CHH (where H can be A, C or T) (Law & Jacobsen, 2010). This mark is established by the action of small RNAs (sRNAs) through a pathway named RNA-directed DNA methylation (RdDM) and can be actively removed from any context by the action of DNA glycosylases (Matzke & Mosher, 2014; Zhang *et al.*, 2018). The modifications that occur in the tails of histones can be active or repressive marks. For example, H3K4 mono-, di- and tri-methylation (H3K4me1, H3K4me2 and H3K4me3) are associated with highly transcribed genes (Zhang *et al.*, 2009), H3K27 tri-methylation (H3K27me3) is mainly found in silenced genes (Zhang *et al.*, 2007) and H3K9 di-methylation (H3K9me2) is rarely seen in genes while being predominantly present in TEs, where it correlates with the presence of DNA methylation, leading to transcriptional silencing and the formation of heterochromatin (Zhou *et al.*, 2010).

Transposable elements are a source of new mutations and genetic/genomic variation and of new regulatory regions for genes (Kidwell & Lisch, 1997; Lisch, 2009). Several agricultural traits like orange, maize and apple color or pepper pungency are regulated by TEs inserted in new locations, creating new expression patterns for the gene(s) in the vicinity of the insertion (Dooner *et al.*, 1991; Butelli *et al.*, 2012; Tanaka *et al.*, 2019; Zhang *et al.*, 2019). These TE domestication events are especially important for plant interaction with their environment (Annacondia *et al.*, 2018). Different abiotic and biotic stresses (including drought, salinity, heat, cold, UV radiation, chemical agents

*These authors contributed equally to this work.

and viral, viroid, bacterial and fungal infections) show examples of TE domestication events that influence gene expression and/or induce changes in the epigenetic regulation of repeats (Annacondia *et al.*, 2018; Mozgova *et al.*, 2019). Defense genes are interesting examples of the interaction between epigenetic regulation and gene regulation and evolution, as most nucleotide-binding site and leucine-rich repeat domain protein (NBS-LRR) genes accumulate in heterochromatic clusters populated by TEs (Meyers *et al.*, 2003). As an example of the role of epigenetic regulation in their transcriptional control, several defense genes, such as *RECOGNITION OF PERONOSPORA PARASITICA 7* (*RPP7*), *RPP4* and *RESISTANCE METHYLATED GENE 1* (*RMG1*), are transcriptionally regulated by domesticated TEs (Tsuchiya & Eulgem, 2013; Yu *et al.*, 2013; Zervudacki *et al.*, 2018). Additionally, mutants of different DNA methylation, RdDM and small RNA pathways regulate immunity to bacterial and fungal infection (Agorio & Vera, 2007; Lopez *et al.*, 2011; Dowen *et al.*, 2012; Yu *et al.*, 2013). Intriguingly, some biotic stresses can induce tolerance towards the pathogen in the subsequent generation (Boyko *et al.*, 2007; De Vos & Jander, 2009; Boyko *et al.*, 2010; Kathiria *et al.*, 2010; Luna *et al.*, 2012; Slaughter *et al.*, 2012), a phenomenon that could be explained by changes in the methylation status of the DNA or chromatin rather than by spontaneous mutagenesis and reversion (Boyko & Kovalchuk, 2011; Luna & Ton, 2012; Annacondia & Martinez, 2019).

The relationship between pathogens and host plants involves an interaction between both genomes and leads to events of coevolution. An example of this interaction takes place between plants and insects. Both groups interact in different ways and have influenced each other during evolution (e.g. the appearance in land plants of entomophily (Darwin, 1899) or carnivory (Renner & Specht, 2013) or the artificial selection of insects that evolve resistance to plants with defense genes (Bown *et al.*, 1997)). Plant–insect interactions are classified as mutualistic, antagonistic or commensalistic. Although they are basic for the ecological equilibrium, some of them can be a threat to the agricultural ecosystems and, hence, to food production. Herbivory insects represent *c.* 50% of the total insect species (Schoonhoven *et al.*, 2005) and are considered a threat to plant productivity. They are among the stresses that induce parental transmission of acquired resistance to the next generation, pointing to a potential role of epigenetic regulation of plant defense (Rasmann *et al.*, 2012). Nevertheless, how this epigenetic response is established during insect infestation is poorly characterized.

Here, we report that epigenetic control is an important part of the *Arabidopsis thaliana* defense response against the infestation by the green peach aphid *Myzus persicae*. Our analysis of DNA methylation, mRNA and sRNA changes induced in plants exposed to aphid feeding shows that the response of the plant is characterized by a transcriptional reprogramming and methylation changes in TEs. These TEs are normally associated with repressive/heterochromatic marks and are dependent on the RdDM pathway for their silencing. Along with this, we find that upon infestation, certain differentially methylated regions (DMRs) are associated with infestation-responsive genes and TF binding sites. Finally, we find that mutant plants deficient in epigenetic silencing show increased resistance to *M. persicae* infestation. Together, our data uncover a novel role for plant epigenetic control in the induction of the transcriptional response to aphid feeding.

Materials and Methods

Plant and insect material

Arabidopsis thaliana (Columbia wild-type Col-0, *ddm1-2, ddc, nrpd1a-4* and *kyp-6*) were sown into potting soil (P-Jord, Hasselfors Garden, Örebro, Sweden). At the four-leaf stage, seedlings were selected by uniformity and carefully replanted into plastic pots ($9 \times 9 \times 7$ cm) with one plant per pot at temperature 20–22°C and 70% relative humidity. Plants were grown under a 16 h : 8 h, light : dark photoperiod. The light was provided by FQ, 80 W, Hoconstant lumix (Osram, Munich, Germany) with a light intensity of 220 µmol photons $m^{-2}\,s^{-1}$. Green peach aphid *M. persicae* (Sulzer) was reared in cultures on potted rapeseed plants *Brassica napus* L. under the same climate conditions as the test plants but in different climate chambers.

Aphid settling test

An aphid no-choice settling test (Ninkovic *et al.*, 2002) was used to investigate aphid behavioral response to different *Arabidopsis* mutants. One randomly chosen leaf was placed inside a transparent polystyrene tube (diameter 1.5 cm, length 5 cm). The lower end of the tube was plugged with a plastic sponge through which the leaf entered via a slit. Ten wingless second- to fourth-instar larvae of *M. persicae* were placed inside the tube. The upper end of the tube was sealed with nylon net. A leaf of each treatment plant placed inside the tube represented a replicate. The number of aphids that settled on the leaf was recorded after 2 h, which is sufficient time for aphids to settle and reach the phloem (Prado & Tjallingii, 1997).

Tissue collection for sRNA, RNA and bisulfite sequencing

Five-week-old plants were infested with 40 wingless second- to fourth-instar larvae of *M. persicae* and covered with a net cage. After 72 h, all aphids were carefully removed using a brush and all the leaves from the *Arabidopsis* rosette (between eight and 10 leaves) were sampled into Falcon tubes and placed in liquid nitrogen for nucleic acid extraction. Four plants were pooled on each bioreplicate. Frozen plant tissue was stored at −70°C before being used for RNA and DNA extraction. The same tissue was used for sRNA, mRNA and genome-wide bisulfite sequencing.

DNA and RNA extraction

Total RNA was extracted using TRIzol reagent (Life Technologies) following the manufacturer's instructions. mRNA for RNA sequencing was obtained by purification with the NEB mRNA isolation kit (New England Biolabs, Ipswich, MA, USA). RNA for sRNA library preparation was enriched with the mirVana

miRNA Isolation Kit (Life Technologies). Genomic DNA was extracted using the DNeasy Plant Mini Kit (Qiagen).

Small RNA, RNA sequencing and analysis

Small RNA libraries were produced using the TruSeq Small RNA Sample Preparation Kit (Illumina Inc., San Diego, CA, USA). Each library was barcoded and sequenced in one lane of an Illumina HiSeq 2000. RNA libraries were produced using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (New England Biolabs). Each library was barcoded and sequenced in one lane of an Illumina HiSeq 2500. The resulting sequences were de-multiplexed, adapter trimmed, and filtered on length and quality. Three bioreplicates were sequenced for sRNA analysis. sRNAs were matched to the *Arabidopsis* genome. Library size was normalized by calculating reads per million of 18–28 nt genome-matched sRNAs. sRNA alignments were performed using BOWTIE (Langmead *et al.*, 2009) with the parameters –t –v2, which allow two mismatches to the alignments. For gene expression analysis, two bioreplicates from each treatment were sequenced. RNA-sequencing paired reads were aligned to the Arabidopsis TAIR10 genome using BOWTIE2 (Langmead & Salzberg, 2012) with default parameters. HTSEQ-COUNTS (Anders *et al.*, 2014) was used to count reads per gene with the parameters --mode union --stranded no --minequal 10 and --nonunique none. For TE expression analysis, RNA sequencing paired reads were aligned to the Arabidopsis TAIR10 genome using STAR (Dobin *et al.*, 2013), allowing mapping to at most 100 'best' matching loci with the following parameters, --outMultimapperOrder Random --outSAMmultNmax -1 --outFilterMultimapNmax 100, used previously for TE analysis (Warman *et al.*, 2020). HTSEQ-COUNTS was used to count reads per TE with the parameters --mode union --stranded no --minequal 0 and --nonunique all. Count tables obtained were used in DESEQ2 (Love *et al.*, 2014) to infer significant expression with fit type set to parametric. Volcano plots were created using GGPLOT2 (Wickham, 2009). All these tools were used through the Galaxy platform (Afgan *et al.*, 2018).

RT–qPCR

For quantitative reverse transcription polymerase chain reaction (RT-qPCR) analysis, total RNA was DNAseI-treated and reverse-transcribed using the RevertAid RT Reverse Transcription Kit (Thermo Fisher Scientific, Waltham, MA, USA). qPCR was performed with 5× HOT FIREPol EvaGreen qPCR Mix Plus (Solis Biodyne, Tartu, Estonia) using three technical replicates from three biological replicates each. qPCR was performed on a CFX Connect Real-Time Detection System and the results analyzed on the CFX MANAGER software package (BioRad, Hercules, CA, USA). The relative expression values for all experiments were calculated based on the expression of the control housekeeping gene AT4G05320 (UBIQUITIN 10). Relative expression was calculated using the 'delta-delta method' formula $2^{-[\Delta CP\ sample - \Delta CP\ control]}$, where 2 represents perfect PCR efficiency. Statistical significance was calculated using unpaired *t*-tests. RT-qPCR primers are shown in Supporting Information Table S1.

Bisulfite library preparation and sequencing analysis

Bisulfite libraries were produced from genomic DNA at Novogene (Beijing, China) and sequenced as paired-end 150 bp fragments in an Illumina Novaseq 6000. Two bioreplicates from each treatment were sequenced. Raw reads were trimmed using TRIMGALORE 0.6.1 for removal of adapters and 10 bases from 5′ ends. Clean reads were mapped to the reference *Arabidopsis* genome TAIR 10 using BISMARK (Krueger & Andrews, 2011), allowing one mismatch per 25 nt seed. Forward and reverse reads were mapped independently. Alignments at the same position were removed using deduplicate_bismark script, including alignments of reads 1 and 2 together. Conversion rates of cytosines were obtained using bismark_methylation_extractor; the first seven bases from the 5′ end and 13 from the 3′ end of each read were ignored. The mean conversion rate based on the cytosine methylation levels in the chloroplast genome for the four samples was 99.76%, and the estimated false-positive methylation rates were 0.24% (Fig. S4e; see later). Tile values for genomic DNA methylation were obtained using the Circos: Interval to Tiles pipeline in the Galaxy platform (Afgan *et al.*, 2018). Circular plots were obtained using J-CIRCOS (An *et al.*, 2015).

DMR identification

The DMR analysis was carried on with the R package DMRCALLER (Catoni *et al.*, 2018); biological replicates from control and infected samples were pooled and compared between treatments. In order to compare both pools, the genome was divided in equal bins of 50 bp size. The DMRs were then computed by performing Fisher's exact test between the number of methylated reads and the total number of reads in both conditions for each bin. The obtained *P*-values were then adjusted for multiple testing using Benjamini and Hochberg's (Benjamini & Hochberg, 1995) method to control the false discovery. Bins with fewer than three cytosines in the specified context or < 0.25 difference in methylation proportion between the two conditions or an average number of reads lower than 8 were discarded. Finally, bins that were at < 300 bp were joined.

Microarray analysis

Microarray analysis was performed for the datasets indicated in Table S2 and retrieved from the NCBI Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/). The expression values were normalized by quantifying the ratio of the mean value for each treatment to the mean value of its respective control. A heat map for the analysis of microarray data was produced using HEATMAPPER (Babicki *et al.*, 2016).

Transcription factor binding site prediction

Transcription factor binding site prediction was performed using the plant transcription factor database (http://planttfdb.cbi.pku.edu.cn/). The prediction tool was used against the nucleotide sequences of the CHH DMRs indicated.

Gene ontology (GO) term analysis

Gene ontology term analysis was carried out using the *GO annotation search, functional categorization and download* tool from the TAIR website (www.arabidopsis.org). In the different analysis, the whole genome categorization was compared to the categorization for the specific group of genes selected for the analysis. Bubble graphs were produced in Microsoft Excel. A biomaps graph was obtained using VirtualPlant 1.3 (http://virtualplant.bio.nyu.edu/cgi-bin/vpweb/).

Results

Analysis of TE activation under different stresses identifies *M. persicae* as a potential inducer of epigenetic changes

To identify stresses that alter the epigenetic regulation in *A. thaliana*, we performed an analysis of TE expression from ATH1 microarray datasets, which have been widely used by the community. The ATH1 microarray contains 1155 TE probes used to track changes in transcript abundance influenced by epigenetic reprogramming (Slotkin *et al.*, 2009). We investigated TE expression under different stresses, including abiotic (heavy metal presence, exposure to heat, cold, spaceflight or UV light among others) and biotic (viral, oomycete, bacterial and insect infection/infestation) (Fig. 1a; Table S2). We found that, in general, these stresses can induce a modest reactivation of TEs, although this response is dependent on the specific stress (Fig. 1a,b). Biotic stress seemed to activate TE expression more consistently than the abiotic stresses analyzed here (Fig. 1a,b). This analysis identified that among the stresses inducing TE reactivation, *M. persicae* infestation after 72 h induced the highest TE transcription. *Myzus persicae* is a major agricultural pest to a large variety of plants that include stone fruits, potato and horticultural crops (Louis & Shah, 2013). A high number of TEs (533 TEs, 46.1% of all the TEs represented in the ATH1 microarray; Fig. 1b) showed evidence of transcriptional activation when plants were under attack from *M. persicae* as compared with control plants. This reactivation included > 40% of all the DNA transposons and retrotransposons represented in the ATH1 microarray, is overrepresented by *Gypsy* and *Copia* retrotransposons and TIR DNA transposons, and is significantly enriched in upregulated members of the *MuDR* nonTIR transposon class ($P = 0.039$, Fisher's exact test, TEs upregulated more than two-fold; Fig. 1c). Analysis of the reactivation indicated that TE activation takes place at 48 h and increases by 72 h post-infestation (pi) (Figs 1b, S1; average fold-change values for retrotransposons at 72 hpi are 3.64- and 4.4-fold for DNA transposons). Other cases of large-scale TE activation are seen when DNA methylation, histone modification and/or heterochromatin formation are lost (Lippman *et al.*, 2003; Lippman *et al.*, 2004; Zilberman *et al.*, 2007; Panda *et al.*, 2016). Together, these results indicate that *M. persicae* infestation results in TE activation, potentially as a result of a large-scale change in the epigenome.

Transcriptional response to aphid feeding in *Arabidopsis* is characterized by transcription factor activity

The extent of TE reactivation observed in our previous analysis could be biased by the presence of TE probes on the ATH1 microarray. To monitor the transcriptional changes under aphid infestation, we repeated the experiment described in De Vos *et al.* (2005) (*Arabidopsis plants* infested with *M. persicae* for 72 h, for details see the Materials and Methods section) and prepared and sequenced high-throughput mRNA libraries (Table S3). First, we focused on understanding the genic transcriptional changes taking place in our libraries. Principal component analysis of gene expression in mRNA libraries generated from control and infested tissue demonstrated that biological replicates clustered together (Fig. S2a). Our differential expression analysis identified 267 genes that were significantly differentially expressed (adjusted $P < 0.05$), with almost all of these being upregulated (265 genes; Fig. 2a; Table S4). Differentially expressed genes contained a significant overrepresentation of mobile mRNAs (24.34% of differentially expressed genes; two-tailed $P < 0.0001$ calculated with a χ^2 test with Yates correction) (Thieme *et al.*, 2015) (Fig. S2c). As expected, the analysis of the GO categories for significantly upregulated genes indicated that these genes were associated with the response to stress or environmental stimuli (Figs 2b,c, S2B).

We further analyzed the molecular functions of these stress-responsive genes by checking the GO term enrichment according to molecular function (Fig. 2d). This revealed an overrepresentation of DNA-binding/transcription factor categories (GO terms 'nucleic acid binding', 'DNA-binding transcription factor activity' and 'DNA binding' were significantly enriched with $P < 0.000\,01$, calculated with Fisher's exact test; Fig. 2d), indicating that these transcriptional regulators are an important part of the response to aphid feeding. Several well-studied TFs showed a strong upregulation (higher than 1.5 $\log_2$fold-change) including members of the WRKY and ERF families (Fig. 2e), which have previously been associated with the response to aphid feeding (Gao *et al.*, 2010). Furthermore, we identified the overexpression of a single component of the epigenetic regulatory pathways that was overexpressed under aphid attack, *HIKESHI-LIKE PROTEIN1* (*HLP1,* significantly overexpressed), a promoter-binding protein that promotes chromatin acetylation (Sharma *et al.*, 2019) (Fig. 2f). In summary, the transcriptional response against aphids showed an overrepresentation of TF activity.

Aphid infestation induces transcriptional activation of TEs

Our previous analysis of ATH1 public datasets indicated a potential reactivation of TEs during aphid infestation. However, the TE probes on the ATH1 array do not represent the genomic distribution of TEs, and favor *Helitron* elements that resemble genes. Accordingly, we explored TE transcriptional and post-transcriptional regulation by performing RNA and sRNA sequencing, which target (respectively) mRNAs and sRNAs derived from Pol II and Pol IV activity (Fig. 3). Principal component analysis of TE expression in mRNA libraries generated from control and infested tissue demonstrated that biological replicates

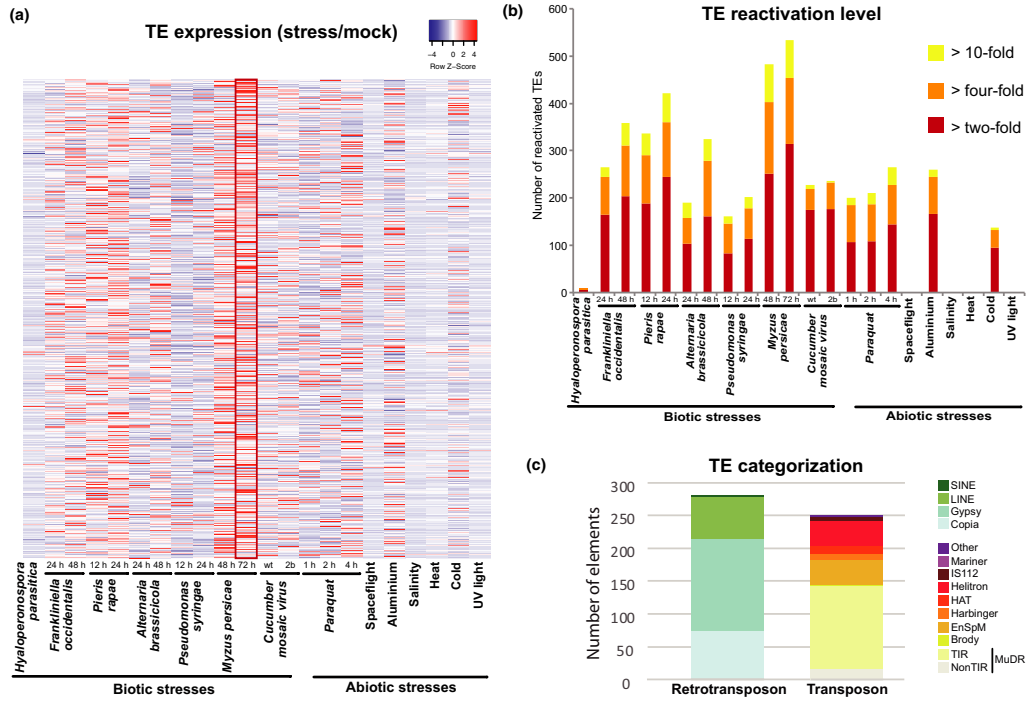


Fig. 1 Aphid infestation induces transposable element (TE) reactivation. (a) Analysis of *Arabidopsis* TE expression in the ATH1 microarray under several stresses. Heat map of the expression values of the indicated treatment relative to its respective control. In experiments with several bioreplicates, the mean values between bioreplicates were used. (b) Number of TEs reactivated in the analyzed stresses grouped by fold categories. (c) Percentage of reactivated TEs belonging to different categories in the ATH1 microarray.

clustered together (Fig. S3a). Analysis of RNA sequencing indicated that 71 TEs suffered a transcriptional reactivation upon aphid infestation (Fig. 3a; Table S5). Reactivated TEs included several members of the ATDNA12T3 family, which cluster in the centromeric regions of chromosomes 3, 4 and 5 and other TEs known to be reactivated under other stresses like ATCOPIA78/ONSEN (Fig. 3a). The DNA TE superfamily (to which ATDNA12T3 TEs belong) was indeed significantly overrepresented in the reactivated TEs population (Fisher's exact test, $P < 0.000\,01$; Fig. 3b).

Next, the analysis of our sRNA sequencing revealed dramatic differences taking place almost exclusively at 24 nt TE-derived sRNAs (Figs 3c,d, S3b–i). Principal component analysis of sRNA libraries generated from control and infested tissue demonstrated that biological replicates clustered together (Fig. S3b). Loss of 24 nt sRNAs was significant at both total sRNAs and TE-derived sRNA populations (*P*-value calculated trough an unpaired *t*-test; Fig. 3c,d). This loss of 24 nt sRNAs was slightly more pronounced in long transposons of almost all TE families (Figs 3e, S3c). Long retrotransposons are located in centromeric and pericentromeric regions, which are the genomic habitats of *Gypsy* and *Copia/LINE* elements, respectively (Underwood *et al.*, 2017). The subgroup of RNA sequencing-identified reactivated TEs also experienced changes at the sRNA level with significant increased levels of 21 nt sRNAs and significant loss of 24 nt sRNAs (Fig. 3f). Indeed, 21 nt sRNAs showed a significant increase of their accumulation levels between control and *M. persicae* sRNA libraries (*P*-value calculated trough a paired *t*-test; Fig. 3g), which are dependent on Pol II and, subsequently, their overaccumulation is a common signature of TE transcriptional reactivation in *Arabidopsis* epigenetic mutants like *ddm1* or *met1* (McCue *et al.*, 2012). In summary, our RNA- and sRNA-sequencing data indicated that during aphid infestation plants reduced the activity of the RdDM pathway, leading to the transcriptional reactivation of centromeric TEs.

Differential methylation of the *Arabidopsis* genome upon aphid infestation

The transcriptional changes observed and the loss of TE-derived 24 nt sRNAs lead us to analyze the levels of DNA methylation. Genomic DNA was isolated, treated with sodium bisulfite and

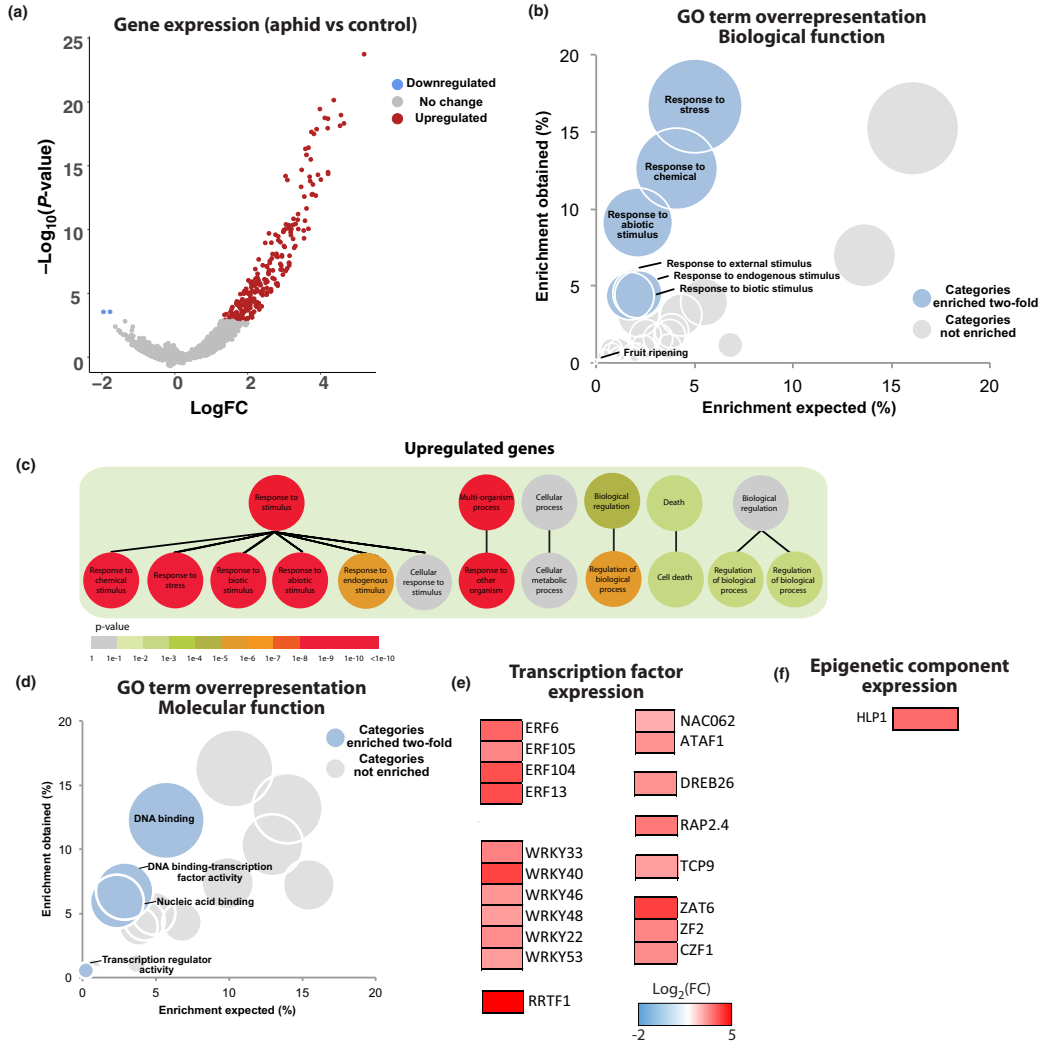


Fig. 2 Aphid feeding-induced changes in gene expression. (a) Volcano plot depicting *Arabidopsis* gene expression in the comparison between aphid-infested and control samples. Red dots indicate genes with significant upregulation. (b) Bubble graph depicting the gene ontology (GO) term overrepresentation test for upregulated genes grouped by biological function. Bubbles in blue show GO categories upregulated two-fold or more. (c) Biomaps of upregulated genes. The colors indicate the statistical significance of the overrepresentation as indicated in the legend. (d) Bubble graph depicting the GO term overrepresentation test for upregulated genes grouped by molecular function. Bubbles in blue show GO categories upregulated two-fold or more. (e) Examples of different transcription factors showing significant upregulation during aphid infestation. (f) A single epigenetic component is upregulated upon aphid infestation. FC, fold-change.

sequenced at 26.4 times average coverage (Table S3; mean conversion rate based on the cytosine methylation levels in the chloroplast genome for the four samples was 99.76%; Fig. S4e). The data were plotted as a heat map on all five chromosomes comparing the control and aphid-infested samples (Fig. 4a). These data revealed a strong enrichment of DNA methylation in the pericentromeric heterochromatin, as expected from somatic tissues. A global analysis of the methylation level at genes and

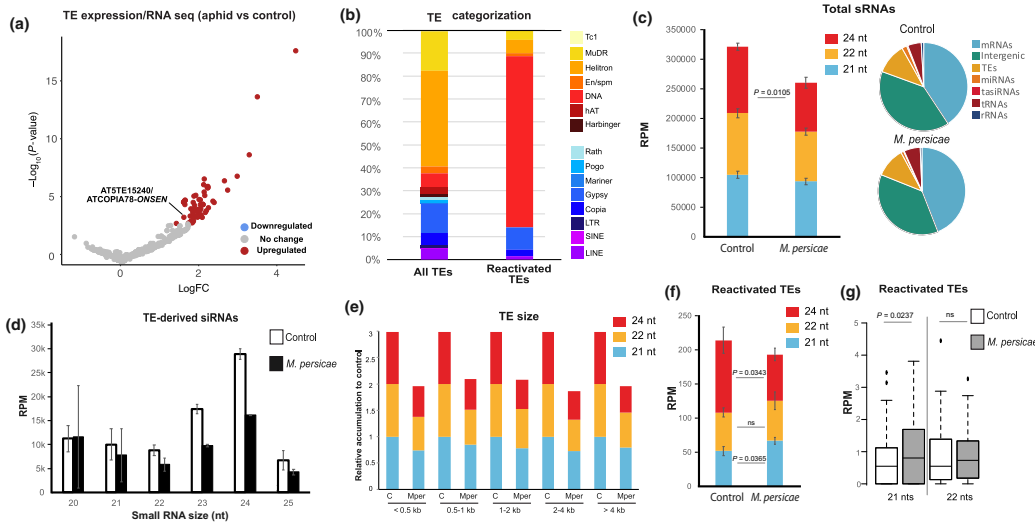


Fig. 3 Changes induced in transposable element (TE) expression by aphid feeding. (a) Volcano plot depicting *Arabidopsis* TE mRNA-seq expression in the comparison between aphid-infested and control RNA samples. Red dots indicate genes with significant upregulation. (b) Categorization of reactivated TEs (right column) compared with categorization of all TEs in the *Arabidopsis* genome (left column). (c) Accumulation of 21, 22 and 24 nt small RNAs (sRNAs) in control and aphid-infested samples from total sRNAs mapping to the *Arabidopsis* chromosomes and normalized to reads per million (RPM). Error bars indicate standard deviation (SD) of three bioreplicates. Pie charts indicate the categorization of total sRNAs from 18 to 28 nt for the categories indicated. The *P*-value was calculated using an unpaired *t*-test. (d) TE-derived sRNA profiles of control and stressed samples normalized to RPM. Error bars indicate SD of three bioreplicates. (e) Relative accumulation of 21, 22 and 24 nt sRNAs in control (C) and aphid-infested samples (Mper) for TEs of different sizes. Values shown are relative to control, where accumulation values for each sRNA category were set to 1. (f) Accumulation of 21, 22 and 24 nt sRNAs in control (C) and aphid-infested samples (Mper) for reactivated TEs. Values are shown in RPM. Error bars indicate the SD of three bioreplicates. The *P*-value was calculated using an unpaired *t*-test. (g) Box plot of 21 and 22 nt sRNA accumulation values per TE member for reactivated TEs in control and *Myzus persicae* samples. Whiskers extend to 5th and 95th percentiles. *P*-values were calculated using a paired *t*-test. FC, fold-change.

TEs for each methylation context revealed that, overall, no dramatic differences existed between the control and aphid-infested samples in any of the profiles for each methylation context (Fig. 4b). This is expected, as aphids cause very subtle wounding as a result of their feeding strategy.

To identify specific regions in the genome harboring differential methylation upon aphid feeding, we determined DMRs (Catoni *et al.*, 2018). This analysis revealed the presence of 2125 statistically significant DMRs for all the DNA methylation contexts and associated both with hypo- and hypermethylation (false discovery rate < 0.05; Figs S4a, 4e; Table S6). The CHG context had the greatest amount of DMRs (1123) followed by CG (691) and CHH (311). Furthermore, while CG DMRs were both present at genes and TEs, most of the CHG and CHH DMRs were associated with TEs (Fig. 4c). TEs located at DMRs were mostly the same TEs that lose 24 nt sRNAs (Fig. 4d). DMRs in the CG context have low CHG and CHH methylation values and the changes observed in these contexts during aphid feeding were not significant (Fig. 4e), pointing to their association with gene body methylation (Fig. 4c). On the other hand, DMRs in the CHG and CHH contexts are highly dynamic and experienced significant changes in different methylation contexts (especially in the CHG and CHH contexts) in the regions that experienced hypo- and hypermethylation (Fig. 4e). Owing to the tight association between CHG and CHH methylation with H3K9me2 (Du *et al.*, 2015), this might indicate that a strong reorganization of heterochromatin takes place in these regions upon aphid feeding.

The relative low number of DMRs identified and the lack of overall changes in the global profiles of DNA methylation indicated that methylation changes could take place only in specific regions. To test if DMRs were associated with particular histone marks, we retrieved public datasets of different histone modification coverage in *Arabidopsis* somatic tissues (Luo *et al.*, 2013) and checked the enrichment of those histone marks in our identified DMRs. Hypomethylated DMRs in the CHH context showed enrichment in the permissive mark H3K18ac ($P = 0.0174$, calculated using an unpaired *t*-test) while simultaneously showing low amounts of the repressive marks H3K27me3 and H3K9me2 (although these changes were not statistically significant) when compared with hypermethylated DMRs (Figs 4f, S4b–d). This indicated that removal of CHH methylation during aphid infestation only took place at regions of the genome that had a high level of permissive histone marks and a low level of repressive

New Phytologist

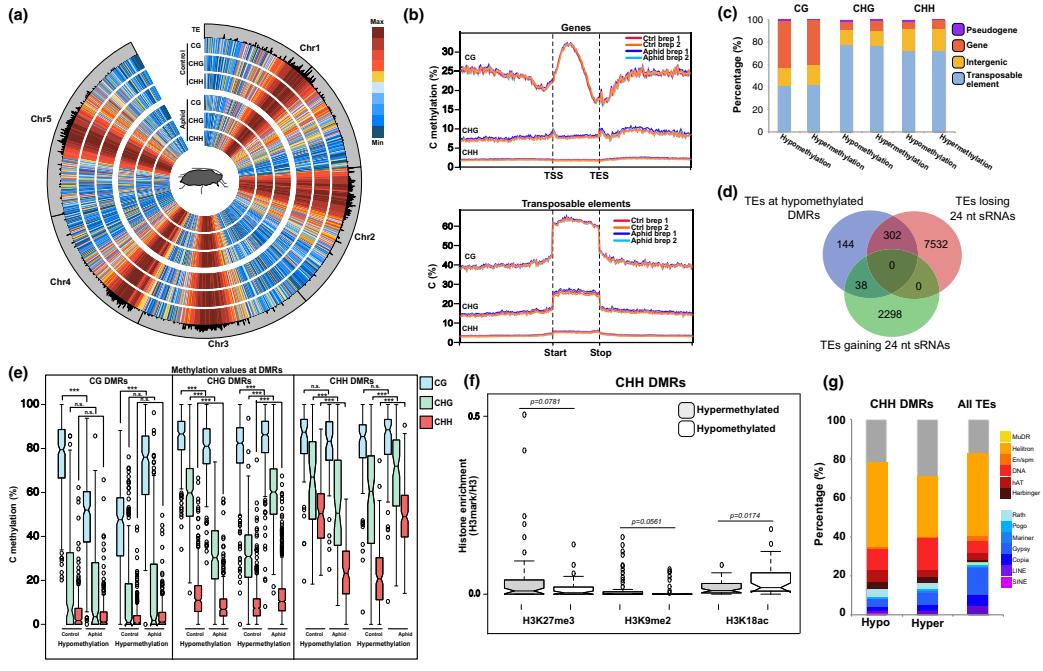


Fig. 4 DNA methylation changes induced by aphid feeding. (a) *Arabidopsis* genome-wide methylation levels for each of the cytosine (C) methylation contexts (CG, CHG and CHH) in control and aphid-infested samples. (b) DNA methylation coverage for genes and transposable elements (TEs) for each C methylation context. (c) Hypermethylation and hypomethylation differentially methylated regions (DMRs) identified for each C methylation context. (d) DMR colocalization with different genomic entities. (e) C methylation values at hypermethylation and hypomethylation DMRs for each methylation context. Asterisks indicate the different levels of significance between the comparisons (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$). The *P*-value was calculated using an unpaired *t*-test. (f) H3K27me3, H3K9me2 and H3K18ac enrichment relative to H3 for hypermethylated and hypomethylated DMRs. Whiskers extend to 5th and 95th percentiles. *P*-values were calculated using an unpaired *t*-test. (g) Categorization of TEs colocalizing with CHH hypermethylation and hypomethylation DMRs in comparison to all the TEs in the TAIR10 *Arabidopsis* genome.

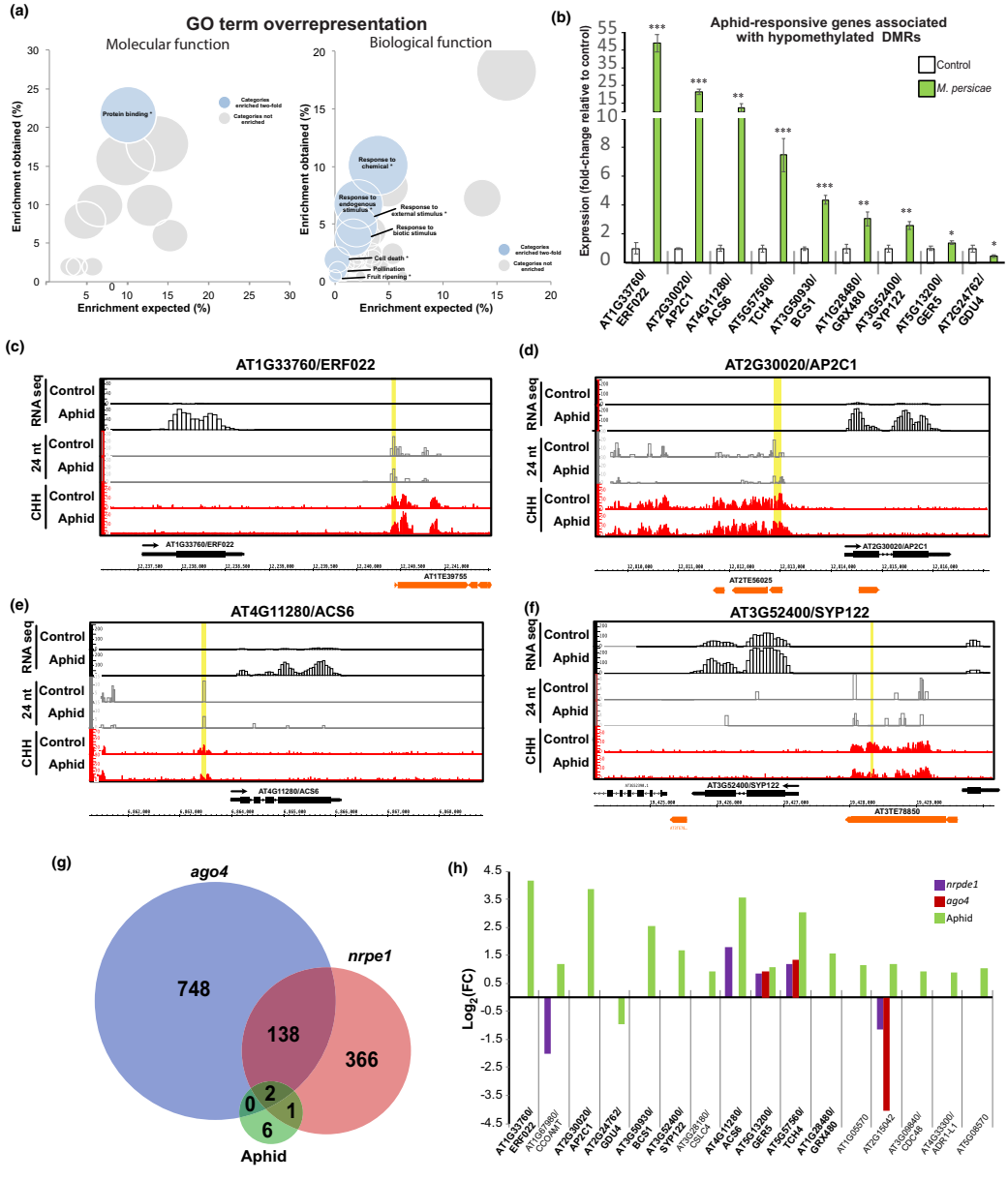
histone marks. Furthermore, hypomethylated CHH DMRs showed an enrichment in *Rath* elements and significant depletion of *Gypsy* elements (two-tailed $P < 0.05$ calculated by a Fisher exact test compared with presence of those elements against the whole genome; Fig. 4g). Therefore, upon aphid feeding, very localized methylation changes take place, mainly associated with epigenetic labile TE regions.

Stress-induced changes in methylation are associated with expression changes in defense-associated genes

Changes in TE methylation can influence the expression of neighboring genes (Wang *et al.*, 2013). To test if the identified DMRs could influence gene expression during aphid feeding, we obtained the list of neighbor genes within a 4 kb window (2 kb

Fig. 5 Transcriptional changes associated with differentially methylated regions (DMRs). (a) Bubble graph depicting the gene ontology (GO) term overrepresentation test for upregulated *Arabidopsis* genes grouped by molecular (left panel) or biological function (right panel). Bubbles in blue show GO categories enriched two-fold or more. Statistically significant categories ($P < 0.05$ determined by Fisher's exact test) are indicated with an asterisk. (b) Quantitative reverse transcription polymerase chain reaction (RT-qPCR) analysis of mRNA expression in control and aphid samples for RNA sequencing-identified aphid-responsive genes. Error bars represent the SD values for the three bioreplicates analyzed. The *P*-value is the result of a standard *t*-test with two tails and unequal variance. Asterisks indicate the different levels of significance between the comparisons (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$). (c–f) Examples of upregulated genes associated with DMRs and confirmed by RT-qPCR. The localization of statistically significant DMRs is highlighted in yellow. (g) Venn diagram depicting overlap between differentially and significantly expressed genes in *polv (nrpe1)*, *ago4* and DMR-associated genes in aphid-infested samples corroborated by RT-qPCR. (h) Expression of DMR-associated differentially expressed genes in *polv (nrpe1)*, *ago4* and aphid-infested samples. Only values with significant difference in expression ($P < 0.05$) in RNA-sequencing data are shown. Genes highlighted in bold were confirmed as significantly overexpressed in aphid samples through RT-qPCR. FC, fold-change.

upstream and downstream) for each DMR. This strategy identified 1010 genes associated with hypermethylated DMRs and 661 genes associated with hypomethylated DMRs (Table S7). As hypomethylation is expected to affect gene expression we focused our analysis on this category. Genes located in the proximities of hypomethylated DMRs were associated with oxygen binding,



translation regulator activity, nuclease and motor activity, and fruit ripening and cell death when associated by biological function (> 1.5-fold upregulation; Fig. S6a,b; see later). When the GO categories were restrained to genes that showed a significant change of expression ($P < 0.05$, 16 genes), we obtained an enrichment of genes with protein-binding activity functions when grouped by molecular function (fruit ripening, cell death, pollination) (> two-fold upregulation and $P = 0.0075$ calculated with Fisher's exact test), and response to endogenous, chemical, external and biotic stimulus when grouped by biological function (> two-fold upregulation categories with significant enrichment; $P < 0.05$ values are indicated in the figure with an asterisk; Fig. 5a). We further confirmed the significant change in expression of nine of those 16 genes by RT-qPCR (Fig. 5b).

We identified several significantly overexpressed genes ($P < 0.05$) located in the proximity of CHH hypomethylated DMRs that were related to plant defense (Figs 5c–f, S5; Table S8). These genes included *AP2C1*, a PP2C-type phosphatase that modulates innate immunity (Schweighofer *et al.*, 2007); *ACS6*, a 1-aminocyclopropane-1-carboxylic acid synthase that is a rate-limiting enzyme that catalyzes the committing step of ethylene biosynthesis (Joo *et al.*, 2008); *SYP122*, a Qa-SNARE protein that drives vesicle fusion and is important for cell growth and expansion and pathogen defense (Waghmare *et al.*, 2018); *GER5*, a stress-responsive glucosyltransferase, rab-like GTPase activator and myotubularin domain protein involved in ABA-mediated stress responses (Baron *et al.*, 2014); and the ethylene response factor *ERF022*, which belongs to the IIIa subgroup of the ERF subfamily which is associated with the response to stress (Nakano *et al.*, 2006).

Next, to explore the potential epigenetic regulation of these genes, we analyzed their expression in epigenetic mutants (not exposed to aphid feeding). We used RNA-sequencing public datasets from Pol V and AGO4 mutants (Zhu *et al.*, 2013; Rowley *et al.*, 2017). Pol V and AGO4 are components of the RdDM pathway that produces sRNAs to target genomic regions and introduces DNA methylation (Matzke & Mosher, 2014). Pol V produces long noncoding transcripts that guide Pol IV-derived 24 nt sRNAs loaded into AGO4 to chromatin (Wierzbicki *et al.*, 2009). Mutations in AGO4 or PolV impair RdDM-dependent methylation, especially in the CHH context, and 82% of loci regulated by Pol V or Pol IV are also regulated by AGO4/AGO6 (Duan *et al.*, 2015). Differentially expressed genes associated with DMRs and confirmed by RT-qPCR were significantly enriched in the portion of genes regulated by the RdDM pathway components AGO4 and/or Pol V (33.3% overlap, two-tailed $P < 0.0001$ calculated by a χ^2 test with Yates correction; Fig. 5g, h). Although some genes showed a similar expression pattern between the RdDM mutants and the aphid-infested samples (e.g. *GER5*, *ACS6*; Fig. 5h) others showed opposing patterns of expression between the aphid-infested samples and the RdDM mutants (notably *CCOAMT* and *AT2G15042*). This different expression pattern led us to question whether the expression of these genes could be regulated by TFs that were not overexpressed in the RdDM mutants. The analysis of TF-binding motifs present in the DMRs of differentially expressed genes showed that several TF-binding motifs were highly enriched, including B3 binding domain-containing TFs such as B3/ARF, AP2/B3 and B3 (20.65-, 11.8- and 8.6-fold enrichment, respectively; Fig. S6c,e). Several TFs of the B3 subfamily belonging to the ERF/AP2 TF family were differentially expressed in the aphid-infested samples, while they did not show this pattern of expression in RdDM mutants (Table S4; Fig. S6d). This indicated that differential expression of TFs probably leads to the observed differences in the expression pattern between aphid-infested samples and RdDM mutants. Overall, our data indicate that DNA methylation changes are associated with gene expression changes, probably in combination with TF-induced expression.

Epigenetic mutants show enhanced defense against aphids

Finally, we tested whether different *Arabidopsis* mutants defective in epigenetic regulation were resistant to aphid infestation. For this, we analyzed aphid no-choice settling where 10 aphids were transferred to a random caged leaf (Fig. 6a). We performed this test in different mutants, including the histone remodeler *DDM1*, the triple mutant defective in maintenance of nonCG methylation *ddc* (*drm1 drm2 cmt3*), the main subunit of the principal factor of the RdDM pathway RNA Pol IV (*nrpd1*), the main subunit of the other principal factor of the RdDM pathway RNA Pol V (*nrpe1*), the main ARGONAUTE protein introducing methylation in the DNA through the canonical RdDM pathway ARGONAUTE 4 (*ago4*), and the H3K9me2 methyltransferase *KYP* (Fig. 6b). All these mutants are known to affect DNA methylation/histone modifications genome-wide and a preliminary analysis of CHH methylation changes in our identified DMRs indicated that, indeed, all of them affect CHH methylation levels in CHH DMRs and in DMRs associated with differentially expressed genes (Fig. S7b,c). Our aphid-settling analysis indicated that, from these components, mutations in *nrpd1* (the largest subunit of Pol IV) and *kyp* show a reduced number of aphids settled, although only *kyp* had a significant decrease (Fig. 6b). Interestingly, we observed the natural overexpression of aphid-resistance genes in *kyp* (Fig. 6c). Furthermore, this resistance is not connected to the leaf phenotype for each of the mutants analyzed here (Fig. S7a).

This indicated that, first, heterochromatin maintenance (regulated by DDM1) and maintenance of nonCG methylation (*ddc*) were not fundamental to elicit a defense response against aphid feeding. Second, our result indicated that the roles of KYP in the regulation of H3K9me2 and CHG methylation (Jackson *et al.*, 2002) and/or its uncharacterized role in the maintenance of CHH methylation (Stroud *et al.*, 2013) were an important part of the defense response against aphid infestation. This result correlates with our observed increase in transcription of centromeric TEs and reduction of sRNAs in centromeric and pericentromeric regions (rich in H3K9me2), and the observed changes in CHH and CHG methylation (tightly associated with H3K9me2). KYP has been previously associated with the regulation of the defense against geminiviruses (Raja *et al.*, 2008; Castillo-Gonzalez *et al.*, 2015; Sun *et al.*, 2015) and the maintenance of β-aminobutyric

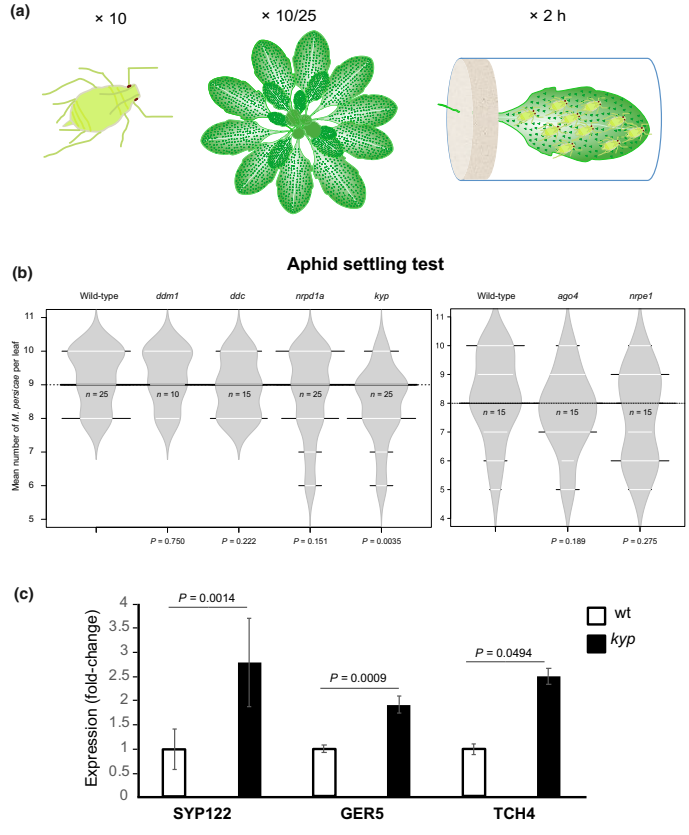


Fig. 6 Epigenetic mutants are resistant to aphid settlement. (a) Depiction of the aphid settlement experiment carried out in our analysis. In brief, 10 aphids were moved to a single caged leaf (attached to the plant) from 10/25 individual *Arabidopsis* plants. (b) Aphid settlement test in different epigenetic mutants. The *P*-values shown were calculated using an unpaired *t*-test. *n*, number of individuals analyzed. (c) Quantitative reverse transcription polymerase chain reaction (RT-qPCR) analysis of mRNA expression in wild-type and *kyp* for RNA sequencing-identified aphid-responsive genes. Error bars represent the SD values for the three bioreplicates analyzed. The *P*-value is the result of a standard *t*-test with two tails and unequal variance.

acid (BABA)-induced priming of the salicylic acid (SA)-dependent defense response (Luna *et al.*, 2014). In summary, our proof-of-concept analysis indicates that mutants in different layers of epigenetic regulation do indeed show enhanced resistance against aphid settlement.

Discussion

Organisms monitor environmental conditions and adapt their development according to them. Plants have developed elegant mechanisms of gene regulation adapted to their sessile nature. One such mechanism is epigenetic regulation, which could maintain modified transcriptional states through cell division and be reversible once the trigger condition disappears. Although it has been widely proposed that epigenetic regulation is an important part of the stress response, we lack a comprehensive knowledge of the genomic loci that are susceptible to those epigenetic changes and their variability between stresses. Here, we demonstrated that aphid feeding induced changes in the epigenetic regulation of the plant genome and that these changes correlated with the transcriptional response. Our data suggest that these epigenetic changes were taking place mainly in TEs. We hypothesize that these changes could be important for recruiting TFs that in turn affect the expression of a specific set of defense genes. This will explain why, despite having a relatively high number of DMRs (Fig. 4), only a very small subset enriched in specific TF-binding motifs was associated with transcriptional changes (Figs 5, S5, S6). An alternative hypothesis to this is that DNA methylation

changes took place downstream of TF binding, a situation that has been described in human dendritic cells (Pacis *et al.*, 2019). Nevertheless, the presence in our analysis of a high number of DMRs without effects at the transcriptional level argues against this hypothesis. As our analysis focused only on a single data point, it is also plausible that the low correlation between the transcriptional and DNA methylation changes could be a result of a temporal separation between both events, as previously described under phosphate starvation in rice (Secco *et al.*, 2015).

Despite their subtle wounding strategy, aphid feeding activates hormonal signals that trigger the reprogramming of the plant transcriptome (Moran *et al.*, 2002; De Vos *et al.*, 2005; Couldridge *et al.*, 2007; Kusnierczyk *et al.*, 2007; Gao *et al.*, 2010). In our study, the transcriptional changes identified by RNA sequencing showed enrichment in genes associated with TF-related activities (Fig. 2). These TFs include AR2/ERF and WRKY TFs, which have previously been associated with the transcriptional response against aphid infestation (Foyer *et al.*, 2015; Kloth *et al.*, 2016). Our analysis of the transcriptional and post-transcriptional regulation of TEs during aphid infestation indicated that TEs were reactivated during aphid feeding, although to a lower extent than expected from our initial study of similar experiments analyzed with the ATH1 microarray data (Fig. 3). One of the reasons for this divergence in the number of reactivated TEs between both analyses could be a result of the nature of the RNA used in the two experiments, that is, total RNA in De Vos *et al.* (2005) against purified mRNA in our study. This aphid-reactivated TE group included the reactivation of the Copia retrotransposon ONSEN, which is known to activate and transpose in *Arabidopsis* plants exposed to different stresses (Cavrak *et al.*, 2014; Matsunaga *et al.*, 2015; Ito *et al.*, 2016). We further detected that TEs experienced a decrease in the activity of the RdDM pathway translated in a loss of 24 nt sRNAs, which led to their transcriptional reactivation. Several of these TEs have a centromeric localization, which correlated with the transposon families losing the majority of 24 nt sRNAs. Furthermore, reactivated TEs accumulated higher levels of 21 and 22 nt sRNAs, which is a signature of TE transcriptional reactivation in epigenetic mutants like *met1* or *ddm1* (McCue *et al.*, 2012).

The changes of TE activity at the transcriptional level prompted us to profile the genome-wide methylation changes under aphid infestation (Fig. 4). Our genome-wide analysis of DNA methylation changes induced by aphid feeding showed that methylation changes happened primarily at genes (in the CG context) and TEs (in the CHG and CHH contexts). CHH hypomethylated DMRs took place only at epigenetically labile regions characterized by low levels of the repressive histone marks H3K27me3 and H3K9me2 and high levels of the transcriptionally permissive mark H3K18ac. As expected, CHH hypomethylated DMRs were predominantly depleted of *Gypsy* TEs, which are long centromeric elements with relative low influence on gene expression (Lermontova *et al.*, 2015). An analysis of the presence of genes in a 4 kb window for CHH hypomethylated DMRs showed the potential transcriptional changes associated with these DMRs. Between differentially expressed genes associated with DRMs, we found several genes related to the defense response at different levels, such as *AP2C1* (Schweighofer *et al.*, 2007), *ACS6* (Joo *et al.*, 2008), *SYP122* (Waghmare *et al.*, 2018), *GER5* (Baron *et al.*, 2014) and the ethylene response factor *ERF022* (Nakano *et al.*, 2006) (Figs 5, S5). A percentage (33.3%) of the differentially expressed genes associated with DMRs were also differentially expressed in *nrpe1* and/or *ago4* mutants, indicating an influence of the RdDM pathway in the regulation of this response (exemplified by *SYP122* in the data showed in Fig. 5f). Together with this observation, we found that DMRs associated with differentially expressed genes showed an enrichment in binding motifs for certain families of TFs, including the AP2-ERF/B3, which has seven members significantly upregulated upon aphid infestation (Fig. S6d). These TFs showed a modest upregulation in the *nrpe1* mutant and none in an *ago4* mutant, which could be one of the reasons why the transcriptional response differed between aphid-infested samples and RdDM mutants. While aphid feeding induced the expression of several TFs, RdDM mutants lack the presence of aphid-induced TFs that would stimulate the defense transcriptional response. As a proof-of-concept, we tested whether *Arabidopsis* mutants defective in DNA and histone methylation had a differential susceptibility to aphid infestation (Fig. 6). Our analysis indicated that mutations in *Pol IV* and *KYP* show increased resistance to aphid settling, confirming the importance of epigenetic regulation in the response against aphids. In *Arabidopsis,* defense genes are located in pericentromeric regions which are densely populated by TEs (Meyers *et al.*, 2003). KYP and Pol IV have a known role in the repression of TEs, so we speculate that their lack of function can also facilitate the transcription of genes located in the proximities of TEs. Our data also indicate that *kyp* has a natural reactivation of some of the aphid-responsive genes. In *kyp* and *nrpd1* mutants, the enhanced activation of defense genes (via transcription or binding of TFs) will explain the increased defense against aphid feeding. Indeed, most of the differentially expressed genes with a proximal CHH DMR identified in our analysis had a TE in the proximities of their regulatory regions (Fig. 5). We hypothesize that *kyp* might show increased resistance compared with other mutants as a result of its reduced methylation level genome-wide, but also as a result of its reduced level of the repressive histone mark H3K9me2. Chromatin marks are known to modulate transcription through influence over TF binding sites (Wu *et al.*, 2019). In *kyp* the low levels of H3K9me2 might allow for a more favorable environment for aphid-responsive TF binding.

It is tempting to speculate that together with the downregulation of the epigenetic silencing at DMRs, the observed overexpression of mobile mRNAs and decrease of 24nt sRNAs would trigger transcriptional or post-transcriptional changes on gene expression at distal tissues, other than leaves, including the precursors of the reproductive structures. Some herbivore insects, like *Pieris rapae*, are known to trigger a defense phenotype in the next generation (Rasmann *et al.*, 2012). TE silencing is reinforced in the shoot apical meristem (SAM) by the RdDM pathway, which leads to the correct transmission of the right epigenetic states for TEs during vegetative growth (Baubec *et al.*, 2014). A potential lack of mobile 24 nt (Molnar *et al.*, 2010) or

21 nt (Dunoyer *et al.*, 2010) TE-derived small interfering RNAs in the SAM or the reproductive structures could lead to epigenetic states that could be inherited. Further analysis of the effect of localized stresses on distal tissues and their offspring could shed light onto the existence of such an elegant overlapping of pathways potentially regulating transgenerational inheritance.

In summary, the evidence presented in our work indicates that changes in epigenetic control were associated with the defense response against aphid infestation in *A. thaliana*. Intriguingly, this response is more complex than previously thought and may involve the interplay between epigenetic and transcriptional regulation. Our work exemplifies the importance of epigenetic regulation in the stress response and the epigenetic plasticity of plant genomes subjected to stress.

Acknowledgements

Research in the Martinez group is supported by SLU, the Carl Tryggers Foundation and the Swedish Research Council (VR 2016-05410). Sequencing was performed by the SNP&SEQ Technology Platform in Uppsala. The facility is part of the National Genomics Infrastructure (NGI) Sweden and Science for Life Laboratory. The SNP&SEQ Platform is also supported by the Swedish Research Council and the Knut and Alice Wallenberg Foundation.

Author contributions

RKS, VN and GM designed the experiment. CMP contributed material. MLA, DM, VS and GM performed the experiments. JLR-V and GM carried out bioinformatic processing of the data. GM did the data analysis. GM wrote the manuscript. All authors interpreted the data and thoroughly checked the manuscript. MLA and DM contributed equally to this work.

ORCID

Maria Luz Annacondia https://orcid.org/0000-0001-7998-8362
Dimitrije Markovic https://orcid.org/0000-0003-3030-4155
German Martinez https://orcid.org/0000-0002-5215-0866
Velemir Ninkovic https://orcid.org/0000-0001-9276-7169
Corné M. J. Pieterse https://orcid.org/0000-0002-5473-4646
Juan Luis Reig-Valiente https://orcid.org/0000-0001-9894-0196
Vassilis Scaltsoyiannes https://orcid.org/0000-0003-4155-9540
R. Keith Slotkin https://orcid.org/0000-0001-9582-3533

Data availability

All raw and processed sequencing data generated in this study have been submitted to the NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) under accession no. GSE144181. Reviewers can access the data using the reviewer token: qvqdmcwefbolhkf.

References

Afgan E, Baker D, Batut B, van den Beek M, Bouvier D, Cech M, Chilton J, Clements D, Coraor N, Gruning BA *et al.* 2018. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Research* **46**: W537–W544.

Agorio A, Vera P. 2007. ARGONAUTE4 is required for resistance to *Pseudomonas syringae* in *Arabidopsis*. *Plant Cell* **19**: 3778–3790.

An J, Lai J, Sajjanhar A, Batra J, Wang C, Nelson CC. 2015. J-Circos: an interactive Circos plotter. *Bioinformatics* **31**: 1463–1465.

Anders S, Pyl PT, Huber W. 2014. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**: 166–169.

Annacondia ML, Mageroy MH, Martinez G. 2018. Stress response regulation by epigenetic mechanisms: changing of the guards. *Physiologia Plantarum* **162**: 239–250.

Annacondia ML, Martinez G. 2019. Chapter 12 – Plant models of transgenerational epigenetic inheritance. In: Tollefsbol TO, ed. *Transgenerational epigenetics, 2nd edn*. Cambridge, UK: Academic Press, 263–282.

Babicki S, Arndt D, Marcu A, Liang Y, Grant JR, Maciejewski A, Wishart DS. 2016. Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Research* **44**: W147–W153.

Baron KN, Schroeder DF, Stasolla C. 2014. GEm-Related 5 (GER5), an ABA and stress-responsive GRAM domain protein regulating seed development and inflorescence architecture. *Plant Science* **223**: 153–166.

Baubec T, Finke A, Mittelsten Scheid O, Pecinka A. 2014. Meristem-specific expression of epigenetic regulators safeguards transposon silencing in *Arabidopsis*. *EMBO Reports* **15**: 446–452.

Benjamini Y, Hochberg Y. 1995. Controlling the False Discovery Rate – a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B–Statistical Methodology* **57**: 289–300.

Bown DP, Wilkinson HS, Gatehouse JA. 1997. Differentially regulated inhibitor-sensitive and insensitive protease genes from the phytophagous insect pest, *Helicoverpa armigera*, are members of complex multigene families. *Insect Biochemistry and Molecular Biology* **27**: 625–638.

Boyko A, Blevins T, Yao Y, Golubov A, Bilichak A, Ilnytskyy Y, Hollunder J, Meins F Jr, Kovalchuk I. 2010. Transgenerational adaptation of *Arabidopsis* to stress requires DNA methylation and the function of Dicer-like proteins. *PLoS ONE* **5**: e9514.

Boyko A, Kathiria P, Zemp FJ, Yao Y, Pogribny I, Kovalchuk I. 2007. Transgenerational changes in the genome stability and methylation in pathogen-infected plants: (virus-induced plant genome instability). *Nucleic Acids Research* **35**: 1714–1725.

Boyko A, Kovalchuk I. 2011. Genetic and epigenetic effects of plant-pathogen interactions: an evolutionary perspective. *Molecular Plant* **4**: 1014–1023.

Butelli E, Licciardello C, Zhang Y, Liu J, Mackay S, Bailey P, Reforgiato-Recupero G, Martin C. 2012. Retrotransposons control fruit-specific, cold-dependent accumulation of anthocyanins in blood oranges. *Plant Cell* **24**: 1242–1255.

Castillo-Gonzalez C, Liu X, Huang C, Zhao C, Ma Z, Hu T, Sun F, Zhou Y, Zhou X, Wang XJ *et al.* 2015. Geminivirus-encoded TrAP suppressor inhibits the histone methyltransferase SUVH4/KYP to counter host defense. *eLife* **4**: e06671.

Catoni M, Tsang JM, Greco AP, Zabet NR. 2018. DMRcaller: a versatile R/Bioconductor package for detection and visualization of differentially methylated regions in CpG and non-CpG contexts. *Nucleic Acids Research* **46**: e114.

Cavrak VV, Lettner N, Jamge S, Kosarewicz A, Bayer LM, Scheid OM. 2014. How a retrotransposon exploits the plant's heat stress response for its activation. *PLoS Genetics* **10**: e1004115.

Coulldridge C, Newbury HJ, Ford-Lloyd B, Bale J, Pritchard J. 2007. Exploring plant responses to aphid feeding using a full Arabidopsis microarray reveals a small number of genes with significantly altered expression. *Bulletin of Entomological Research* **97**: 523–532.

Darwin C. 1899. *The various contrivances by which orchids are fertilized by insects.* London: John Murray.

De Vos M, Jander G. 2009. *Myzus persicae* (green peach aphid) salivary components induce defence responses in *Arabidopsis thaliana*. *Plant, Cell & Environment* **32**: 1548–1560.

De Vos M, Van Oosten VR, Van Poecke RM, Van Pelt JA, Pozo MJ, Mueller MJ, Buchala AJ, Metraux JP, Van Loon LC, Dicke M *et al.* 2005. Signal signature and transcriptome changes of *Arabidopsis* during pathogen and insect attack. *Molecular Plant–Microbe Interactions* **18**: 923–937.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15–21.

Dooner HK, Robbins TP, Jorgensen RA. 1991. Genetic and developmental control of anthocyanin biosynthesis. *Annual Review of Genetics* **25**: 173–199.

Dowen RH, Pelizzola M, Schmitz RJ, Lister R, Dowen JM, Nery JR, Dixon JE, Ecker JR. 2012. Widespread dynamic DNA methylation in response to biotic stress. *Proceedings of the National Academy of Sciences, USA* **109**: E2183–E2191.

Du J, Johnson LM, Jacobsen SE, Patel DJ. 2015. DNA methylation pathways and their crosstalk with histone methylation. *Nature Reviews Molecular Cell Biology* **16**: 519–532.

Duan CG, Zhang H, Tang K, Zhu X, Qian W, Hou YJ, Wang B, Lang Z, Zhao Y, Wang X *et al.* 2015. Specific but interdependent functions for *Arabidopsis* AGO4 and AGO6 in RNA-directed DNA methylation. *EMBO Journal* **34**: 581–592.

Dunoyer P, Schott G, Himber C, Meyer D, Takeda A, Carrington JC, Voinnet O. 2010. Small RNA duplexes function as mobile silencing signals between plant cells. *Science* **328**: 912–916.

Foyer CH, Verrall SR, Hancock RD. 2015. Systematic analysis of phloem-feeding insect-induced transcriptional reprogramming in *Arabidopsis* highlights common features and reveals distinct responses to specialist and generalist insects. *Journal of Experimental Botany* **66**: 495–512.

Gao LL, Kamphuis LG, Kakar K, Edwards OR, Udvardi MK, Singh KB. 2010. Identification of potential early regulators of aphid resistance in *Medicago truncatula* via transcription factor expression profiling. *New Phytologist* **186**: 980–994.

Gutzat R, Mittelsten SO. 2012. Epigenetic responses to stress: triple defense? *Current Opinion in Plant Biology* **15**: 568–573.

Ito H, Kim JM, Matsunaga W, Saze H, Matsui A, Endo TA, Harukawa Y, Takagi H, Yaegashi H, Masuta Y *et al.* 2016. A stress-activated transposon in *Arabidopsis* induces transgenerational abscisic acid insensitivity. *Scientific Reports* **6**: 23181.

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

Joo S, Liu Y, Lueth A, Zhang S. 2008. MAPK phosphorylation-induced stabilization of ACS6 protein is mediated by the non-catalytic C-terminal domain, which also contains the cis-determinant for rapid degradation by the 26S proteasome pathway. *The Plant Journal* **54**: 129–140.

Kathiria P, Sidler C, Golubov A, Kalischuk M, Kawchuk LM, Kovalchuk I. 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**: 1859–1870.

Kidwell MG, Lisch D. 1997. Transposable elements as sources of variation in animals and plants. *Proceedings of the National Academy of Sciences, USA* **94**: 7704–7711.

Kloth KJ, Wiegers GL, Busscher-Lange J, van Haarst JC, Kruijer W, Bouwmeester HJ, Dicke M, Jongsma MA. 2016. AtWRKY22 promotes susceptibility to aphids and modulates salicylic acid and jasmonic acid signalling. *Journal of Experimental Botany* **67**: 3383–3396.

Krueger F, Andrews SR. 2011. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**: 1571–1572.

Kusnierczyk A, Winge P, Midelfart H, Armbruster WS, Rossiter JT, Bones AM. 2007. Transcriptional responses of *Arabidopsis thaliana* ecotypes with different glucosinolate profiles after attack by polyphagous *Myzus persicae* and oligophagous *Brevicoryne brassicae*. *Journal of Experimental Botany* **58**: 2537–2552.

Lamke J, Baurle I. 2017. Epigenetic and chromatin-based mechanisms in environmental stress adaptation and stress memory in plants. *Genome Biology* **18**: 124.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**: 357–359.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* **10**: R25.

Law JA, Jacobsen SE. 2010. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature Reviews Genetics* **11**: 204–220.

Lermontova I, Sandmann M, Mascher M, Schmit AC, Chaboute ME. 2015. Centromeric chromatin and its dynamics in plants. *The Plant Journal* **83**: 4–17.

Lippman Z, Gendrel AV, Black M, Vaughn MW, Dedhia N, McCombie WR, Lavine K, Mittal V, May B, Kasschau KD *et al.* 2004. Role of transposable elements in heterochromatin and epigenetic control. *Nature* **430**: 471–476.

Lippman Z, May B, Yordan C, Singer T, Martienssen R. 2003. Distinct mechanisms determine transposon inheritance and methylation via small interfering RNA and histone modification. *PLoS Biology* **1**: E67.

Lisch D. 2009. Epigenetic regulation of transposable elements in plants. *Annual Review of Plant Biology* **60**: 43–66.

Lopez A, Ramirez V, Garcia-Andrade J, Flors V, Vera P. 2011. The RNA silencing enzyme RNA polymerase v is required for plant immunity. *PLoS Genetics* 7: e1002434.

Louis J, Shah J. 2013. *Arabidopsis thaliana-Myzus persicae* interaction: Shaping the understanding of plant defense against phloem-feeding aphids. *Frontiers in Plant Science* **4**: 213.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**: 550.

Luna E, Bruce TJ, Roberts MR, Flors V, Ton J. 2012. Next-generation systemic acquired resistance. *Plant Physiology* **158**: 844–853.

Luna E, Lopez A, Kooiman J, Ton J. 2014. Role of NPR1 and KYP in long-lasting induced resistance by beta-aminobutyric acid. *Frontiers in Plant Science* **5**: 184.

Luna E, Ton J. 2012. The epigenetic machinery controlling transgenerational systemic acquired resistance. *Plant Signaling & Behavior* **7**: 615–618.

Luo C, Sidote DJ, Zhang Y, Kerstetter RA, Michael TP, Lam E. 2013. Integrative analysis of chromatin states in *Arabidopsis* identified potential regulatory mechanisms for natural antisense transcript production. *The Plant Journal* **73**: 77–90.

Matsunaga W, Ohama N, Tanabe N, Masuta Y, Masuda S, Mitani N, Yamaguchi-Shinozaki K, Ma JF, Kato A, Ito H. 2015. A small RNA mediated regulation of a stress-activated retrotransposon and the tissue specific transposition during the reproductive period in *Arabidopsis*. *Frontiers in Plant Science* **6**: 48.

Matzke MA, Mosher RA. 2014. RNA-directed DNA methylation: An epigenetic pathway of increasing complexity. *Nature Reviews Genetics* **15**: 394–408.

McCue AD, Nuthikattu S, Reeder SH, Slotkin RK. 2012. Gene expression and stress response mediated by the epigenetic regulation of a transposable element small RNA. *PLOS Genetics* **8**: e1002474.

Meyers BC, Kozik A, Griego A, Kuang H, Michelmore RW. 2003. Genome-wide analysis of NBS-LRR-encoding genes in *Arabidopsis*. *Plant Cell* **15**: 809–834.

Molnar A, Melnyk CW, Bassett A, Hardcastle TJ, Dunn R, Baulcombe DC. 2010. Small silencing RNAs in plants are mobile and direct epigenetic modification in recipient cells. *Science* **328**: 872–875.

Moran PJ, Cheng Y, Cassell JL, Thompson GA. 2002. Gene expression profiling of *Arabidopsis thaliana* in compatible plant-aphid interactions. *Archives of Insect Biochemistry and Physiology* **51**: 182–203.

Mozgova I, Mikulski P, Pecinka A, Farrona S. 2019. Epigenetic Mechanisms of Abiotic Stress Response and Memory in Plants. In: Alvarez-Venegas R, De-la-Peña C, Casas-Mollano JA, eds. *Epigenetics in plants of agronomic importance: fundamentals and applications: transcriptional regulation and chromatin remodelling in plants*. Cham, Switzerland: Springer International, 1–64.

Nakano T, Suzuki K, Fujimura T, Shinshi H. 2006. Genome-wide analysis of the ERF gene family in *Arabidopsis* and rice. *Plant Physiology* **140**: 411–432.

Ninkovic V, Olsson U, Pettersson J. 2002. Mixing barley cultivars affects aphid host plant acceptance in field experiments. *Entomologia Experimentalis et Applicata* **102**: 177–182.

Pacis A, Mailhot-Leonard F, Tailleux L, Randolph HE, Yotova V, Dumaine A, Grenier JC, Barreiro LB. 2019. Gene activation precedes DNA demethylation in response to infection in human dendritic cells. *Proceedings of the National Academy of Sciences, USA* **116**: 6938–6943.

Panda K, Ji L, Neumann DA, Daron J, Schmitz RJ, Slotkin RK. 2016. Full-length autonomous transposable elements are preferentially targeted by expression-dependent forms of RNA-directed DNA methylation. *Genome Biology* **17**: 170.

Prado E, Tjallingii WF. 1997. Effects of previous plant infestation on sieve element acceptance by two aphids. *Entomologia Experimentalis et Applicata* **82**: 189–200.

Raja P, Sanville BC, Buchmann RC, Bisaro DM. 2008. Viral genome methylation as an epigenetic defense against geminiviruses. *Journal of Virology* **82**: 8997–9007.

Rasmann S, De Vos M, Casteel CL, Tian D, Halitschke R, Sun JY, Agrawal AA, Felton GW, Jander G. 2012. Herbivory in the previous generation primes plants for enhanced insect resistance. *Plant Physiology* **158**: 854–863.

Renner T, Specht CD. 2013. Inside the trap: gland morphologies, digestive enzymes, and the evolution of plant carnivory in the *Caryophyllales*. *Current Opinion in Plant Biology* **16**: 436–442.

Rowley MJ, Rothi MH, Bohmdorfer G, Kucinski J, Wierzbicki AT. 2017. Long-range control of gene expression via RNA-directed DNA methylation. *PLoS Genetics* **13**: e1006749.

Schoonhoven LM, van Loon JJA, Dicke M. 2005. *Insect-plant biology.* Oxford, UK: OUP Oxford.

Schweighofer A, Kazanaviciute V, Scheikl E, Teige M, Doczi R, Hirt H, Schwanninger M, Kant M, Schuurink R, Mauch F *et al.* 2007. The PP2C-type phosphatase AP2C1, which negatively regulates MPK4 and MPK6, modulates innate immunity, jasmonic acid, and ethylene levels in *Arabidopsis*. *Plant Cell* **19**: 2213–2224.

Secco D, Wang C, Shou HX, Schultz MD, Chiarenza S, Nussaume L, Ecker JR, Whelan J, Lister R. 2015. Stress induced gene expression drives transient DNA methylation changes at adjacent repetitive elements. *eLife* **4**: e09343.

Sharma M, Banday ZZ, Shukla BN, Laxmi A. 2019. Glucose-regulated HLP1 acts as a key molecule in governing thermomemory. *Plant Physiology* **180**: 1081–1100.

Slaughter A, Daniel X, Flors V, Luna E, Hohn B, Mauch-Mani B. 2012. Descendants of primed *Arabidopsis* plants exhibit resistance to biotic stress. *Plant Physiology* **158**: 835–843.

Slotkin RK, Vaughn M, Borges F, Tanurdzic M, Becker JD, Feijo JA, Martienssen RA. 2009. Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. *Cell* **136**: 461–472.

Stroud H, Greenberg MV, Feng S, Bernatavichute YV, Jacobsen SE. 2013. Comprehensive analysis of silencing mutants reveals complex regulation of the Arabidopsis methylome. *Cell* **152**: 352–364.

Sun YW, Tee CS, Ma YH, Wang G, Yao XM, Ye J. 2015. Attenuation of histone methyltransferase KRYPTONITE-mediated transcriptional gene silencing by Geminivirus. *Scientific Reports* **5**: 16476.

Tanaka Y, Asano T, Kanemitsu Y, Goto T, Yoshida Y, Yasuba K, Misawa Y, Nakatani S, Kobata K. 2019. Positional differences of intronic transposons in pAMT affect the pungency level in chili pepper through altered splicing efficiency. *The Plant Journal* **100**: 693–705.

Thieme CJ, Rojas-Triana M, Stecyk E, Schudoma C, Zhang W, Yang L, Minambres M, Walther D, Schulze WX, Paz-Ares J *et al.* 2015. Endogenous *Arabidopsis* messenger RNAs transported to distant tissues. *Nature Plants* **1**: 15025.

Tsuchiya T, Eulgem T. 2013. An alternative polyadenylation mechanism coopted to the Arabidopsis RPP7 gene through intronic retrotransposon domestication. *Proceedings of the National Academy of Sciences, USA* **110**: E3535–E3543.

Underwood CJ, Henderson IR, Martienssen RA. 2017. Genetic and epigenetic variation of transposable elements in *Arabidopsis*. *Current Opinion in Plant Biology* **36**: 135–141.

Waghmare S, Lileikyte E, Karnik R, Goodman JK, Blatt MR, Jones AME. 2018. SNAREs SYP121 and SYP122 mediate the secretion of distinct cargo subsets. *Plant Physiology* **178**: 1679–1688.

Wang X, Weigel D, Smith LM. 2013. Transposon variants and their effects on gene expression in *Arabidopsis*. *PLoS Genetics* **9**: e1003255.

Warman C, Panda K, Vejlupkova Z, Hokin S, Unger-Wallace E, Cole RA, Chettoor AM, Jiang D, Vollbrecht E, Evans MMS *et al.* 2020. High expression in maize pollen correlates with genetic contributions to pollen fitness as well as with coordinated transcription from neighboring transposable elements. *PLoS Genetics* **16**: e1008462.

Wickham H. 2009. *ggplot2: elegant graphics for data analysis.* New York, NY, USA: Springer.

Wierzbicki AT, Ream TS, Haag JR, Pikaard CS. 2009. RNA polymerase V transcription guides ARGONAUTE4 to chromatin. *Nature Genetics* **41**: 630–634.

Wu Z, Tang J, Zhuo J, Tian Y, Zhao F, Li Z, Yan Y, Yang R. 2019. Chromatin signature and transcription factor binding provide a predictive basis for understanding plant gene expression. *Plant and Cell Physiology* **60**: 1471–1486.

Yu A, Lepere G, Jay F, Wang J, Bapaume L, Wang Y, Abraham AL, Penterman J, Fischer RL, Voinnet O *et al.* 2013. Dynamics and biological relevance of DNA demethylation in *Arabidopsis* antibacterial defense. *Proceedings of the National Academy of Sciences, USA* **110**: 2389–2394.

Zervudacki J, Yu A, Amesefe D, Wang J, Drouaud J, Navarro L, Deleris A. 2018. Transcriptional control and exploitation of an immune-responsive family of plant retrotransposons. *EMBO Journal* **37**: e98482.

Zhang H, Lang Z, Zhu JK. 2018. Dynamics and function of DNA methylation in plants. *Nature Reviews Molecular Cell Biology* **19**: 489–506.

Zhang L, Hu J, Han X, Li J, Gao Y, Richards CM, Zhang C, Tian Y, Liu G, Gul H *et al.* 2019. A high-quality apple genome assembly reveals the association of a retrotransposon and red fruit colour. *Nature Communications* **10**: 1494.

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

Zhang X, Clarenz O, Cokus S, Bernatavichute YV, Pellegrini M, Goodrich J, Jacobsen SE. 2007. Whole-genome analysis of histone H3 lysine 27 trimethylation in *Arabidopsis*. *PLoS Biology* **5**: e129.

Zhou J, Wang X, He K, Charron J-BF, Elling AA, Deng XW. 2010. Genome-wide profiling of histone H3 lysine 9 acetylation and dimethylation in *Arabidopsis* reveals correlation between multiple histone marks and gene expression. *Plant Molecular Biology* **72**: 585–595.

Zhu Y, Rowley MJ, Bohmdorfer G, Wierzbicki AT. 2013. A SWI/SNF chromatin-remodeling complex acts in noncoding RNA-mediated transcriptional silencing. *Molecular Cell* **49**: 298–309.

Zilberman D, Gehring M, Tran RK, Ballinger T, Henikoff S. 2007. Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. *Nature Genetics* **39**: 61–69.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Expression of TEs represented in the ATH1 microarray in *Myzus persicae*-infested samples.

Fig. S2 Analysis of differential gene expression in mRNA sequencing libraries.

Fig. S3 Analysis of differential TE expression in mRNA sequencing libraries and global characterization of sRNA sequencing libraries.

Fig. S4 Characterization of DMR regions and genome-wide bisulfite DNA libraries.

Fig. S5 Examples of genes associated with DMRs.

Fig. S6 Characterization of differentially expressed genes and TFs associated with DMRs.

Fig. S7 Phenotypic and epigenetic characterization of mutants analyzed in the aphid settling test.

Table S1 RT-qPCR primers used in this study.

Table S2 Public data used in this analysis.

Table S3 High-throughput sequencing data produced in this analysis.

Table S4 Significantly differentially expressed genes upon aphid infestation.

Table S5 Significantly differentially expressed TEs upon aphid infestation.

Table S6 Differentially methylated regions identified.

Table S7 Genes identified in the proximity of DMRs.

Table S8 Differentially expressed genes associated with hypomethylated DMRs.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.

New Phytologist

***New Phytologist* Supporting Information**

Article title: Aphid feeding induces the relaxation of epigenetic control and the associated regulation of the defense response in *Arabidopsis*

Authors: Maria Luz Annacondia, Dimitrije Markovic, Juan Luis Reig-Valiente, Vassilis Scaltsoyiannes, Corné M.J. Pieterse, Velemir Ninkovic, R. Keith Slotkin and German Martinez.

Article acceptance date: 14 January 2021

The following Supporting Information is available for this article:

Figure S1. Expression of TEs represented in the ATH1 microarray in *Myzus persicae*-infested samples.

Figure S2. Analysis of differential gene expression in mRNA sequencing libraries.

Figure S3. Analysis of differential TE expression in mRNA sequencing libraries and global characterization of sRNA sequencing libraries.

Figure S4. Characterization of DMR regions and genome-wide bisulfite DNA libraries.

Figure S5. Examples of genes associated with DMRs.

Figure S6. Characterization of differentially expressed genes and TFs associated with DMRs.

Figure S7. Phenotypic and epigenetic characterization of mutants analyzed in the aphid settling test.

Table S1. RT-qPCR primers used in this study.

Table S2. Public data used in this analysis

Table S3. High-throughput sequencing data produced in this analysis.

Table S4. Significantly differentially expressed genes upon aphid infestation.

Table S5. Significantly differentially expressed TEs upon aphid infestation.

Table S6. Differentially methylated regions identified**.**

Table S7. Genes identified in the proximities of DMRs.

Table S8. Differentially expressed genes associated with hypomethylated DMRs.

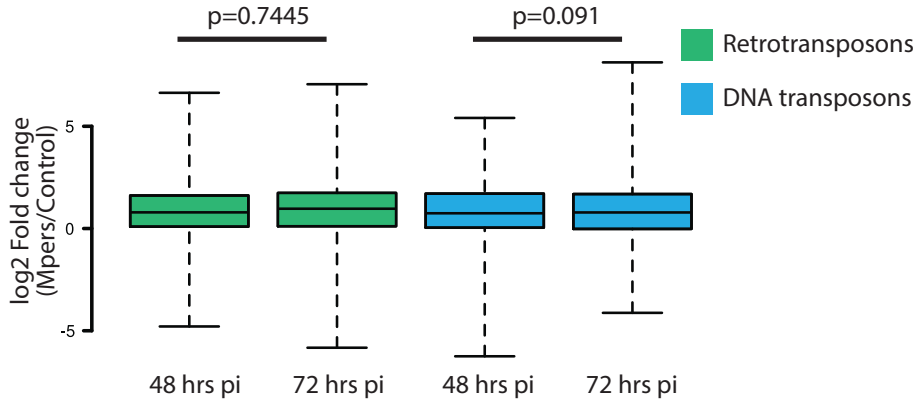


Fig. S1 Expression of TEs represented in the ATH1 microarray in *Myzus persicae*-infested samples. Expression change (log2 fold change) of TEs belonging to different categories in the *Myzus persicae* ATH1 datasets at 48 and 72 hrs pi. Whiskers extend to the maximum and minimum values. p-values were calculated using an unpaired t-test.

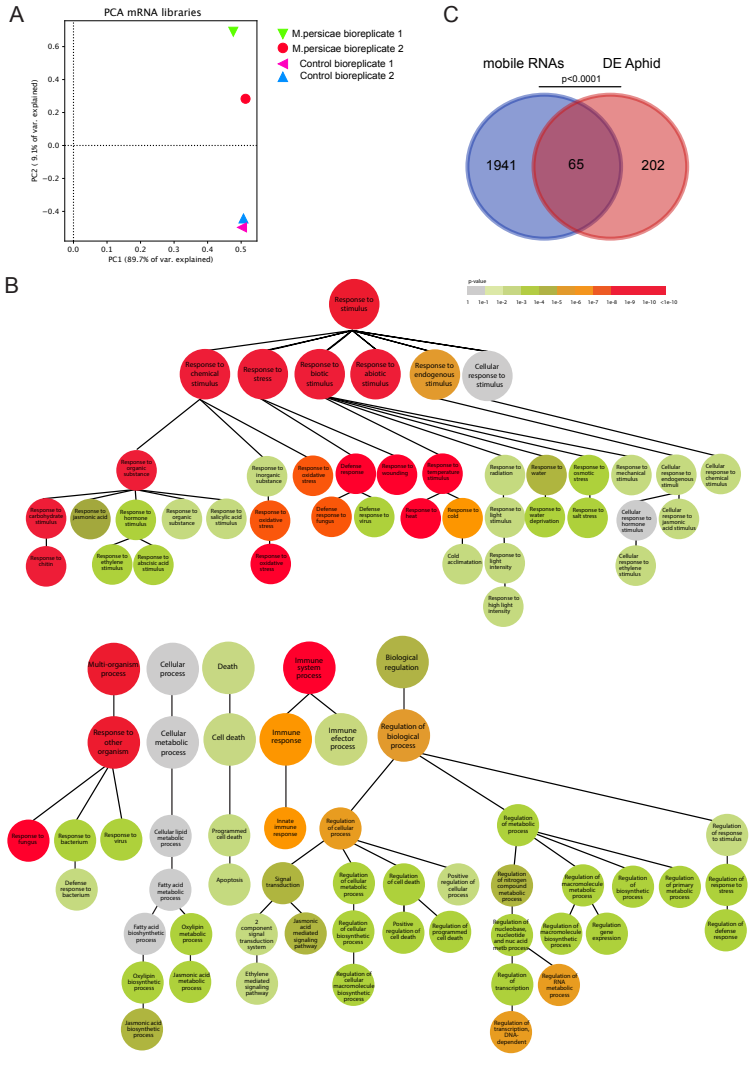


Fig. S2 Analysis of differential gene expression in mRNA sequencing libraries. A. Principal component analysis illustrating the relationships of the mRNA-seq libraries generated from control and aphid-infestated tissues based on gene levels. **B.** Complete biomaps of upregulated genes. **C.** Venn diagram showing the overlap between the whole mobile mRNAs identified in Arabidopsis and the differentially expressed genes during aphid infestation.

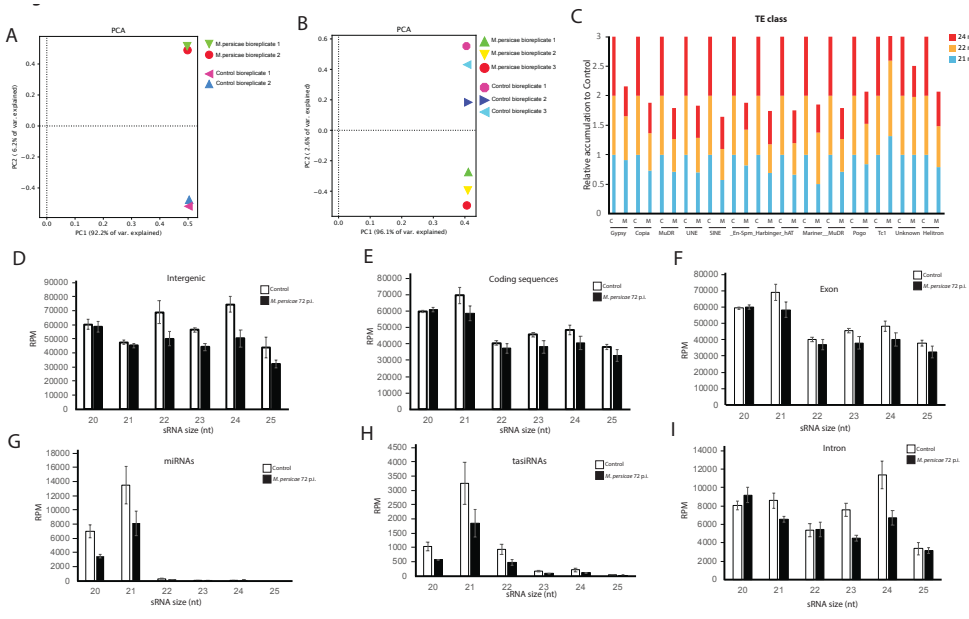


Fig. S3 Analysis of differential TE expression in mRNA sequencing libraries and global characterization of sRNA sequencing libraries. **A.** Principal component analysis illustrating the relationships of the mRNA-seq libraries generated from control and aphid-infestated tissues based on TE levels. **B.** Principal component analysis illustrating the relationships of the sRNA-seq libraries generated from control and aphid-infestated tissues based on genome-wide bins. **C.** Relative accumulation of 21, 22 and 24 nt sRNAs in aphid infested samples (Mper) relative to control (C) for various TE families. Accumulation values in the control sample were set to one. **D-I.** Global sRNAs profiles of control and stressed samples mapping to intergenic regions (**D**), coding sequences (**E**), miRNAs (**F**), tasiRNAs (**G**), exons (**H**) and introns (**I**). Error bars indicate standard deviation of 3 bioreplicates.

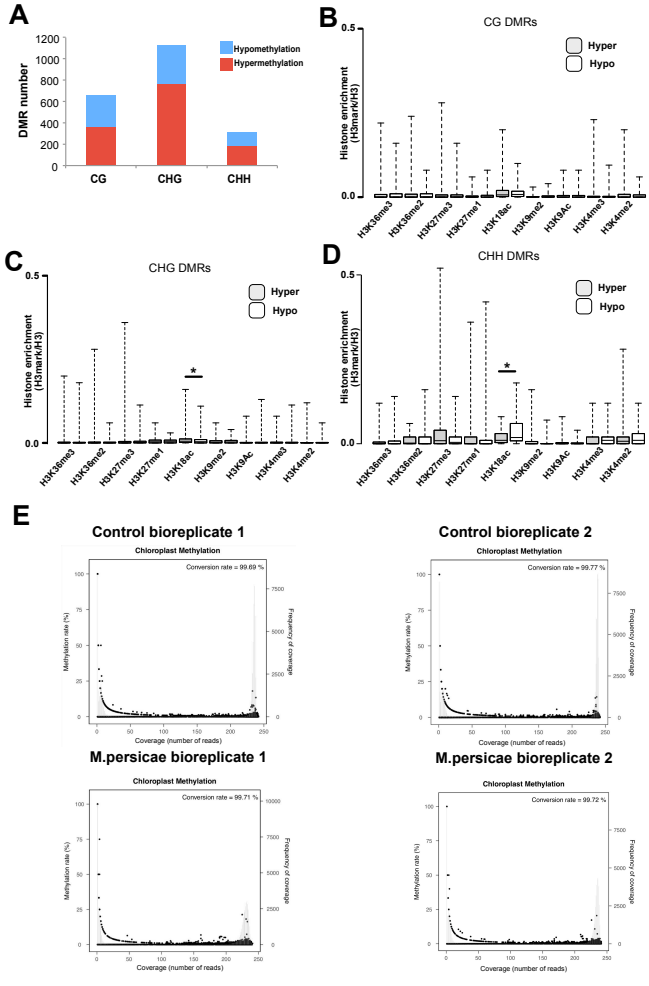


Fig. S4 Characterization of DMR regions and genome-wide bisulfite DNA libraries. A. Number of hypomethylated (blue) and hypermethylated (red) DMRs present in aphid infested samples for the different DNA methylation contexts. **B-D.** Histone mark enrichment relative to H3 for hypermethylated and hypomethylated DMRs in the CG (**B**), CHG (**C**) and CHH (**D**) contexts. P-values were calculated using an unpaired t-test. Whiskers extend to minimum and maximum values. **E.** Coverage and methylation percentage of cytosines in the chloroplast genome for each bioreplicate analyzed in control and aphid-infested samples.

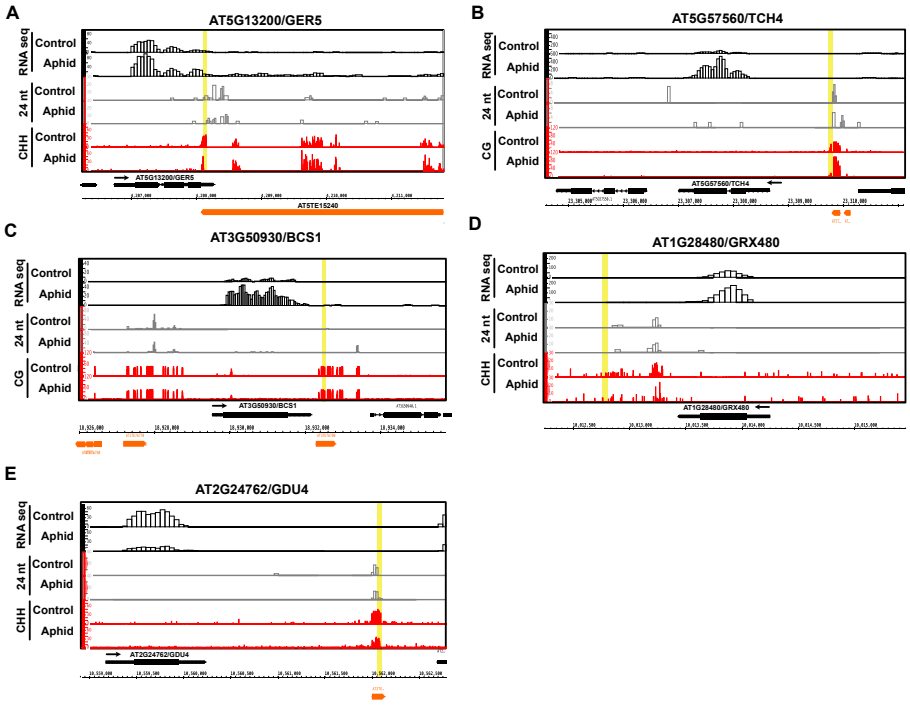


Fig. S5 Examples of genes associated with DMRs. A-E. Examples of upregulated genes associated with CHH DMRs confirmed by RT-qPCR. The localization of statistically significant DMRs is highlighted in yellow.

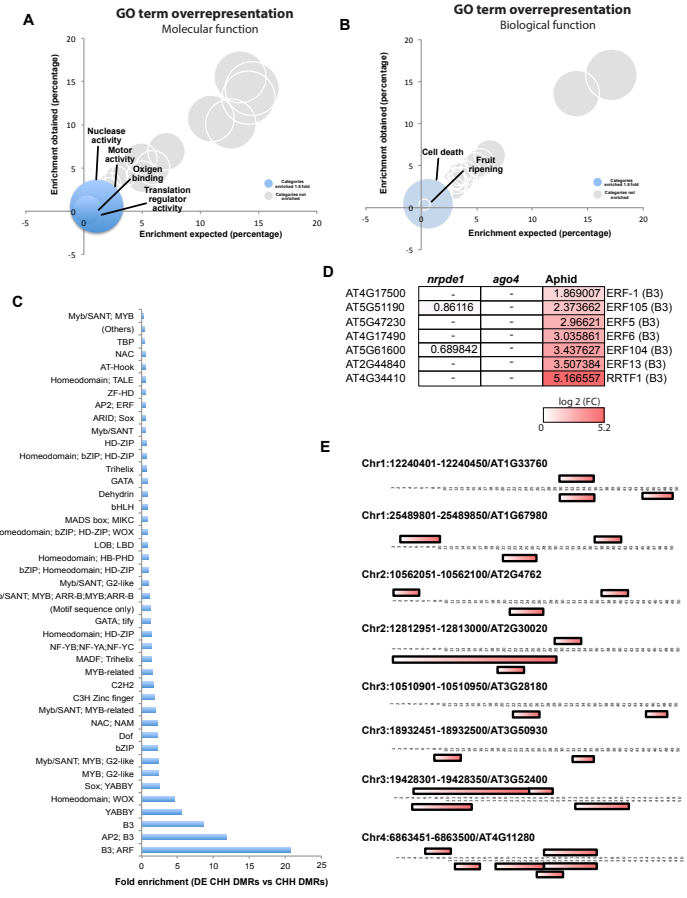


Fig. S6 Characterization of differentially expressed genes and TFs associated with DMRs. A-B. Bubble graph depicting the GO term overrepresentation test for all genes associated with DMRs grouped by molecular (**A**) or biological (**B**) function. Bubbles in blue show GO categories enriched 1.5 fold or more. C. Fold enrichment of transcription factor binding sites at CHH DMRs harboring a differentially expressed gene vs all CHH DMRs. **D.** Examples of different ERF/AP2 transcription factors belonging to the B3 subfamily showing upregulation during aphid infestation in *nrpde1*, *ago4* and aphid infested RNA sequencing libraries. Only values of significant differentially expressed genes is shown. **E.** B3, AP2 and ERF binding sites located at CHH DMRs associated with differentially expressed genes.

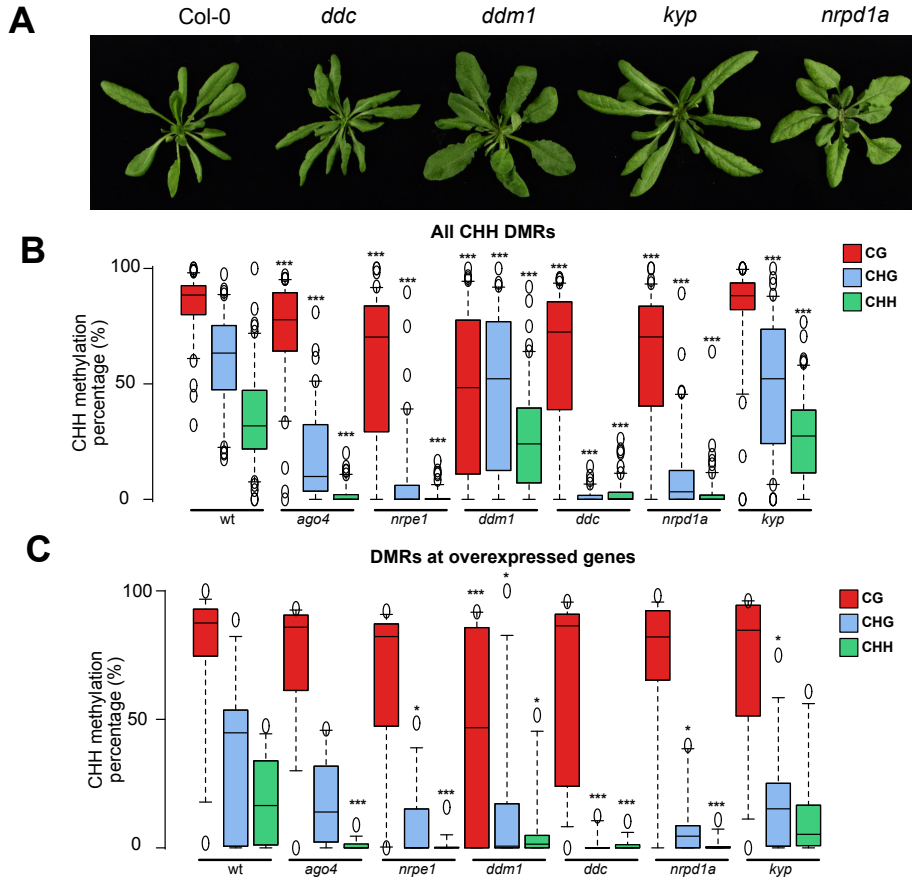


Fig. S7 Phenotypic and epigenetic characterization of mutants analyzed in the aphid settling test. A. Representative pictures showing the leaf phenotype of the mutants used in the aphid settling test. **B.-C** DNA methylation values for all DMRs (B) and DMRs in the proximities of genes with significant change of expression (C). P-values were calculated using a paired t-test. Whiskers extend to 5th and 95th percentile. Asterisks indicate the different levels of significance between the comparisons (*<0.05, **<0.01, ***<0.001).

Acta Universitatis Agriculturae Sueciae

Doctoral Thesis No. 2021:58

The development and survival of plants rely on their ability to adapt to environmental changes and stressful scenarios. Therefore, plants have developed an incredible genome plasticity that allows them to quickly reprogram their transcription to adapt to new conditions. In this study, we characterize the genome-wide changes on sRNAs, DNA methylation and repressive histone marks during two important biotic stresses, aphid infestation and CMV infection. Moreover, we analyze their interplay with gene expression and their impact on the regulation of the defense response. This doctoral thesis provides a better understanding of the role of epigenetic regulation and the RNA silencing during biotic stress in *Arabidopsis thaliana*.

María Luz Annacondia López received her graduate education at the Department of Plant Biology, SLU, Uppsala. She obtained her B.Sc. and M.Sc. degrees in Biology from the University of Oviedo, Spain.

Acta Universitatis Agriculturae Sueciae presents doctoral theses from the Swedish University of Agricultural Sciences (SLU).

SLU generates knowledge for the sustainable use of biological natural resources. Research, education, extension, as well as environmental monitoring and assessment are used to achieve this goal.

Online publication of thesis summary: http://pub.epsilon.slu.se/

ISSN 1652-6880
ISBN (print version) 978-91-7760-794-6
ISBN (electronic version) 978-91-7760-795-3