



HOKKAIDO UNIVERSITY

Title	Analysis of Transcriptional and Transpositional Regulation of a Heat-Activated Retrotransposon in <i>Arabidopsis thaliana</i>
Author(s)	Niu, Xiaoying
Degree Grantor	北海道大学
Degree Name	博士(生命科学)
Dissertation Number	甲第16283号
Issue Date	2025-03-25
DOI	https://doi.org/10.14943/doctoral.k16283
Doc URL	https://hdl.handle.net/2115/97692
Type	doctoral thesis
File Information	Xiaoying_Niu.pdf



Analysis of Transcriptional and Transpositional Regulation of a Heat-Activated Retrotransposon in *Arabidopsis thaliana*

(シロイヌナズナにおける熱活性型レトロトランスポゾンの制御機構の解明)

Xiaoying Niu

Biosystems Science Course

Graduate School of Life Science

Hokkaido University

2025.03

CONTENTS

Abstract

Abbreviations List

General Introduction

Chapter 1: Analysis of Transpositional Regulation of *ONSEN* on RdDM Pathway

Introduction

Materials

Results

1. The Transpositional Regulation of *ONSEN* by Downstream Factors of RdDM.
2. Transcriptional Regulation of *ONSEN* by Downstream RdDM Factors.
3. The Correlation Between DNA Methylation and *ONSEN* Silencing.
4. Chromatin Condensation is Independent of *ONSEN* Activation.
5. siRNA Plays a Crucial Role in the Silencing Mechanism of *ONSEN*.

* Figure and Table (legends)

Discussion

Chapter 2: Analysis of a Novel Mechanism of Retrotransposon Repression by Histone Deacetylases

Introduction

Materials

Results

1. TSA Inhibits *ONSEN* Transcriptional Activity.
2. Loss of HDA19 Leads to Decreased *ONSEN* Expression.
3. *ONSEN* Downregulation in *hda19* Is Independent of the HSFA2 Pathway.

4. *ONSEN* Transcription Levels and DNA Methylation in *hda19*.

5. Analysis of Histone H3 Acetylation and Methylation in *hda19*.

* Figure and Table (legends)

Discussion

Chapter 3: Investigating a Novel Transposable Regulator of *ONSEN*

Introduction

Materials

Results

1. Transgenerational Transposition of *ONSEN* in EMS Mutants.

2. Characteristics of *Screen 1*.

3. Analysis of Transposition in SALK Line Mutants.

4. Classical Mapping Approach to Investigate *Screen 1*.

5. Transpositional Study of *ONSEN* in Further Selected SALK Mutants.

* Figure and Table (legends)

Discussion

General Discussion

Figure and Table legends

Methods

Growing conditions

Inhibitor medium

Heat-stress treatment

Quantitative analysis

Southern blot analysis

DNA methylation analysis

Formaldehyde-Assisted Isolation of Regulatory Elements (FAIRE)

Cytology

Northern blot analysis

RNA sequencing (RNA-seq)

Whole Genome Bisulfite Sequencing (WGBS)

Chromatin Immunoprecipitation (ChIP)

Next-Generation Sequencing (NGS)

Re-sequencing (Whole Genome Resequencing; WGRS)

Transposon Display Sequencing (TED-seq)

Sequence data analysis

Primer List

References

Acknowledgment

Abstract

Transposon transposition can induce mutations in the genome, potentially causing adverse effects. Therefore, mechanisms that control transposon transposition are essential for maintaining genomic stability in the host. While most transposons are regulated at the transcriptional level through epigenetic modifications such as DNA methylation and histone modifications, the details of post-transcriptional control mechanisms for transposon transposition remain poorly understood. In *Arabidopsis thaliana*, the retrotransposon *ONSEN* is activated by heat stress (HS) at 37°C. In mutants deficient in the RNA-directed DNA methylation (RdDM) pathway, HS triggers transgenerational transposition of *ONSEN*. However, since *ONSEN* transcription levels do not strongly correlate with transposition frequency, RdDM components may have distinct roles in suppressing both transcription and transposition, or other regulatory pathways may act independently of RdDM. This study uses *ONSEN* as a representative model to explore the regulatory mechanisms governing transposons, offering new insights into retrotransposon regulation.

Abbreviations List

5-Azac	5-aza-2'-deoxycytidine
AGO4	Argonaute 4
CMT2	Chromomethylase 2
CMT3	Chromomethylase 3
DAPI	4',6-Diamidino-2-phenylindole
DCL3	Dicer-like 3
DDM1	Decreased DNA Methylation I
DDR	DRD1-DMS3-RDM1
DMS3-1	defective in meristem silencing 3-1
DMRs	Differentially Methylated Regions
DNMT	DNA Methyltransferase
DRD1-6	defective in RNA-directed DNA methylation 1-6
DRM2	Domains Rearranged Methyltransferase 2
dsRNA	double-stranded RNA
EMS	Ethyl Methanesulfonate
GFP	Green Fluorescent Protein
HAT	Histone Acetyltransferase
HDAC	Histone Deacetylase
HS	Heat-stress
Hsfa2	Heat Shock Factor A2
IGN	Intergenic non-coding RNA
lncRNA	Long non-coding RNA
MET1	Methyltransferase 1
MORC6	Microrchidia 6

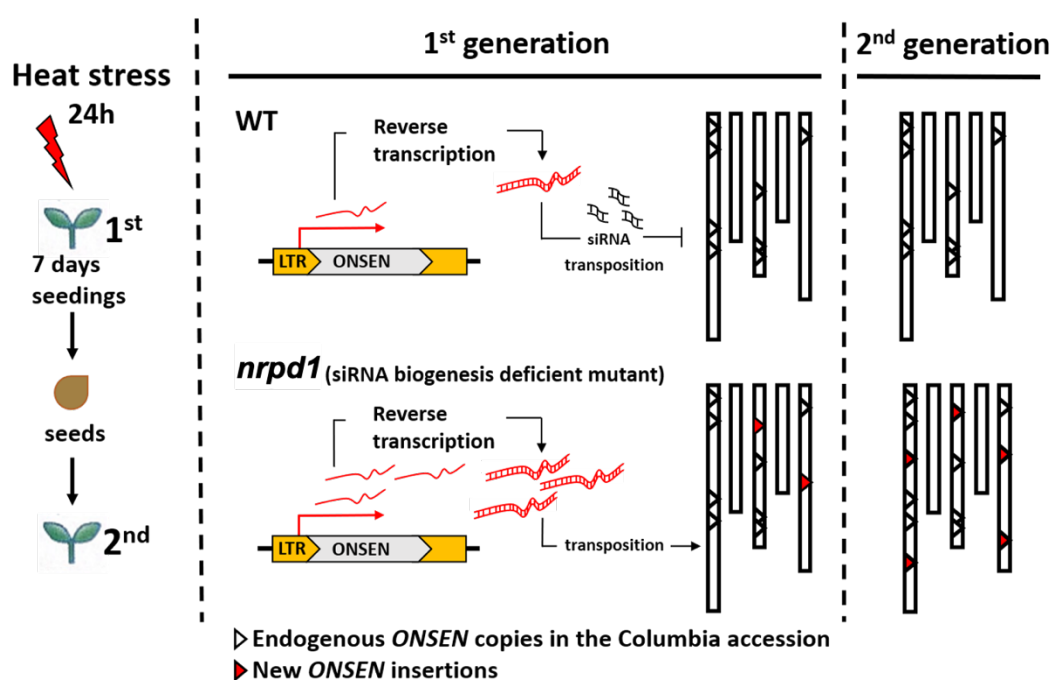
NILs	Near-Isogenic Lines
NS	Non-stress
Pol IV	RNA polymerase IV
Pol V	RNA polymerase V
PTM	post-translational modifications
RdDM	RNA-directed DNA Methylation
RDM1-4	RNA-directed DNA methylation 1-4
RDR2	RNA-dependent RNA polymerase 2
siRNA	small interfering RNA
SNP	Single Nucleotide Polymorphism
SUVH2	SET domain-containing SU(VAR)3-9 homologs 2
SUVH9	SET domain-containing SU(VAR)3-9 homologs 9
TE	Transposable Element
TGS	Transcriptional Gene Silencing
TSA	Trichostatin A
TSS	Transcription Start Site

General Introduction

Transposons are essential components of eukaryotic genomes, consisting of DNA fragments capable of moving within the genome. Barbara McClintock pioneered the discovery of transposons, identifying "mutable loci" in the maize genome in the late 1940s, laying the foundation for transposon research. Since then, extensive studies have revealed the widespread presence and diversity of transposons across eukaryotic organisms (Arkhipova IR and Neutral Theory., 2018). In the model plant *Arabidopsis thaliana*, transposable elements (TEs) account for about 21% of the genome, with about 32,000 TE copies (Quesneville H., 2020). TEs are primarily classified by their transposition mechanisms, structures, and functions into retrotransposons (Class I), DNA transposons (Class II), and other categories (Wicker T et al., 2007). Class I retrotransposons transpose via reverse transcription of an RNA intermediate. They are further divided into long terminal repeat (LTR) retrotransposons and non-LTR retrotransposons. LTR retrotransposons initiate transposition from a promoter within the 5' LTR, where RNA polymerase II (Pol II) transcribes the element into RNA. The RNA is then reverse-transcribed into cDNA in the cytoplasm and re-imported into the nucleus, where it integrates into the host genome (Havecker E.R et al., 2004). Compared to animals, LTR retrotransposons are more prevalent in plants, with the *copia* and *gypsy* superfamilies being the most prominent (Levin H et al., 2011). In contrast, Class II transposons employ a 'cut-and-paste' mechanism. They typically encode transposases within autonomous copies, which catalyze the excision and reintegration of the transposon at different genomic locations (Feschotte C and Pritham EJ., 2007).

Interestingly, numerous studies have shown that environmental stresses, such as high temperature, salinity, and drought, can activate TEs (Casacuberta E and González J.,

2013; Ito H., 2022). In the host organism, TE activity is often associated with alterations in genetic material, and when it occurs in germ cells, it can be transmitted to the next generation (Grandbastien M.A et al., 1997; Oliver K.R and Greene, 2009). However, TE transposition has the potential to induce genetic disorders and threaten the genetic stability and adaptability of populations. As a result, TE activity is often tightly regulated by silencing mechanisms (Belancio VP et al., 2008; Bourgeois Y et al., 2019). *ONSEN*, a *Ty1/copia*-type retrotransposon, is typically activated under HS (Ito H et al., 2011). Previous research has demonstrated that *ONSEN* activity is suppressed by the RNA-directed DNA methylation (RdDM) pathway, preventing transposition in the Col wild type (Ito H et al., 2011). However, in mutants deficient in RdDM factors, such as *nrpd1*, both the transcriptional and transpositional silencing of *ONSEN* are alleviated. (Fig S1, Ito H et al., 2011; Hayashi Y et al., 2020). The RdDM pathway is considered a key defense mechanism in plants, playing a central role in regulating TE activity alongside histone modifications (Lippman Z et al., 2004; Slotkin R and Martienssen R, 2007; Mhiri C et al., 2022).



Supplementary Figure 1: Diagram of the Transposition Mechanism of *ONSEN*.

Under heat stress (HS) at 37°C, *ONSEN* transcription is activated, followed by reverse transcription into cDNA. Due to mechanisms that inhibit transposable element (TE) activity, such as the RdDM pathway in plants, *ONSEN* cannot complete transposition in Col. However, in the *nprdl* mutant, where the RdDM pathway is compromised, *ONSEN* successfully inserts into the chromosome and is inherited by the next generation.

Epigenetics refers to heritable mechanisms that regulate gene expression without altering the DNA sequence. These mechanisms include chemical modifications, changes in chromatin structure, and RNA-mediated regulation, all of which influence gene expression and function (Han H et al., 2021). Among these, DNA methylation and histone modification are crucial molecular processes driving epigenetic regulation (Jaenisch R and Bird A., 2003; Bird A., 2007). DNA methylation, in particular, suppresses gene transcription (Zilberman D et al., 2007). In plants, RdDM is a DNA methylation mechanism dependent on the small interfering RNA (siRNA) pathway, which promotes transcriptional gene silencing (TGS) (Wakasa Y et al., 2018; Pontes O et al., 2006;). The RdDM process is critical in regulating plant development, stress responses, and gene transcription (Matzke M and Mosher R et al., 2014). It operates through both canonical and non-canonical pathways. In the canonical RdDM pathway, RNA polymerase IV (Pol IV) and RNA polymerase V (Pol V) are essential. Pol IV generates 24-nucleotide small interfering RNAs (24-nt siRNAs), while Pol V produces long non-coding RNAs (lncRNAs) (Onodera Y et al., 2005; Wierzbicki A.T et al., 2008). At RdDM target loci, Pol IV transcribes non-coding RNAs, which then associate with RNA-dependent RNA polymerase 2 (RDR2) to form double-stranded

RNAs (dsRNAs) (Haag J et al., 2012). Dicer-like 3 (DCL3) then processes the dsRNAs into 24-nt siRNAs. These 24-nt siRNAs are loaded onto ARGONAUTE 4 (AGO4) in the cytoplasm and subsequently transported back into the nucleus (Qi Y et al., 2006; Xie Z et al., 2004). Downstream, SUVH2 and SUVH9, which are SET domain-containing homologs of SU(VAR)3-9, bind to methylated DNA and recruit Pol V to transcribe lncRNAs at target loci. This recruitment involves interactions with the DDR complex (DRD1-DMS3-RDM1; defective in RNA-directed DNA methylation 1 [DRD1], defective in meristem silencing 3 [DMS3], and RNA-directed DNA methylation 1 [RDM1]) as well as with the MORC protein family (Microorchidia) (Liu ZW et al., 2014; Law J.A. et al., 2010). Finally, AGO4-siRNA complexes guide Pol V-derived scaffold RNAs to induce de novo DNA methylation via DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2) (Haag J and Pikaard C., 2011; Matzke M et al., 2014; Wierzbicki A.T et al., 2009). In contrast, the non-canonical pathway is initiated by RNA polymerase II (Pol II) and RNA-dependent RNA polymerase 6 (RDR6). DICER-LIKE 2 (DCL2) or DCL4 processes the dsRNAs into siRNAs, which are loaded onto AGO6 and guide DRM2 to induce DNA methylation at target loci (Wu L et al., 2012).

Methylation at the fifth carbon position of cytosine (5-methylcytosine, 5-mC) is a key epigenetic modification for regulating gene expression. In plants, DNA methylation occurs across all cytosine sequence contexts, including CG, CHG, and CHH (where H represents A, T, or C) (Zhang H et al., 2018; Shim S et al., 2021). After de novo methylation is established via the RdDM pathway, its maintenance is primarily governed by methyltransferase 1 (MET1) and chromomethylase 2 and 3 (CMT2 and CMT3). Specifically, MET1 ensures the maintenance of CG methylation, while CMT3 preserves CHG methylation. Asymmetric CHH methylation is maintained

cooperatively by CMT2, DRM1, and DRM2 (Zubko E et al., 2012; Cao X et al., 2002; Stroud H et al., 2014). Additionally, CHH methylation partially depends on the presence of CG and CHG methylation (Cokus S et al., 2008). Beyond the RdDM pathway, Decreased DNA Methylation I (DDM1) also plays an essential role in DNA methylation (Jeddeloh J et al., 1999). DDM1 is a chromatin remodeler responsible for DNA and histone methylation in both CG and non-CG contexts (Akinmusola RY et al., 2023). Methylation damage caused by DDM1 mutations primarily affects TEs located in heterochromatin regions, with a lesser impact on genes (Lyons DB and Zilberman D., 2017). DNA methylation is intricately linked with histone methylation, as histone methylation can guide DNA methylation patterns and serve as templates for certain histone modifications following DNA replication (Cedar H and Bergman Y., 2009). Okitsu CY and Hsieh CL (2007) proposed that di-methylation of lysine 4 on histone H3 (H3K4me2) is associated with DNA methylation, while tri-methylation of lysine 4 on histone H3 (H3K4me3) shows weaker correlation with DNA methylation. Furthermore, CMT2 and CMT3 localize to specific chromatin regions by recognizing and binding to di-methylated lysine 9 on histone H3 (H3K9me2), facilitating DNA methylation in those regions (Du J et al., 2012).

Histones are essential proteins found in the nuclei of all eukaryotic cells, where they bind with DNA to form nucleosome structures. Each nucleosome core particle consists of an octamer of histones, including two copies each of H3, H4, H2A, and H2B, wrapped by approximately 147 base pairs of DNA (Kornberg RD and Lorch Y., 1999). Chromatin is formed by an array of nucleosomes and its structure plays a crucial role in regulating DNA activity (Luger K et al., 1997). The function and structure of chromatin are primarily regulated through histone post-translational modifications (PTMs) (Bannister AJ and Kouzarides T, 2011). Histone modifications include

acetylation, methylation, ubiquitination, phosphorylation, etc. Methylation and acetylation, in particular, often occur on the amino (N)-terminal ends of histones, which extend from the nucleosome and interact with adjacent nucleosomes (Nguyen K.T et al., 2018; Kouzarides T, 2007). Modifications to these N-terminal tails can alter nucleosome interactions, influencing chromatin structure and function (Rothbart B and Strahl B, 2014).

Histone acetylation typically occurs on lysine residues in histones. The positively charged N-terminal tails of histones interact with the negatively charged DNA, promoting tight wrapping of DNA around histones to form nucleosomes. Acetylation involves the addition of an acetyl group, which carries a negative charge, to lysine residues. This neutralizes the positive charge of lysine, weakening the electrostatic attraction between histones and DNA. As a result, the relaxed nucleosome structure increases DNA accessibility, facilitating gene transcription (Grunstein M., 1997; Shahbazian MD and Grunstein M., 2007; Luger K et al., 2012). Histone acetylation and deacetylation exist in dynamic equilibrium. Acetylation is mediated by histone acetyltransferases (HATs), while histone deacetylases (HDACs) catalyze deacetylation (Bannister AJ and Kouzarides T., 2007; Gujral P et al., 2020). Generally, acetylation is associated with chromatin relaxation and gene activation. Histone methylation regulates gene expression in more complex ways and can either activate or repress transcription. Methylation occurs on various amino acid residues, primarily lysine, in mono-, di-, or tri-methylated forms. Common methylation sites include lysine K4, K9, K27, and K36 on histone H3 (Hu H and Du J., 2022; Du J et al., 2015). The impact of methylation depends on the specific site and degree of modification. For example, trimethylation of lysine 4 (H3K4me3) and lysine 36 (H3K36me3) on histone H3 is generally associated with transcriptional activation. In contrast, di-methylation of

lysine 9 (H3K9me2) and trimethylation of lysine 27 (H3K27me3) are markers of transcriptional repression (Kang H et al., 2022).

The transcriptional and transpositional regulation mechanisms of LTR retrotransposons remain partially unexplored. *ONSEN*, a typical LTR retrotransposon, serves as a valuable research model due to its ability to be artificially activated under controlled conditions. Previous studies have demonstrated that *ONSEN* exhibits varying degrees of transcriptional or transpositional activity in mutants of specific RdDM pathway factors (Ito H et al., 2011; Hayashi Y et al., 2020; Takehira K et al., 2021). However, since *ONSEN*'s transcription levels do not correlate with its transposition frequency, it is possible that RdDM components independently regulate its transcription and transposition, or an unidentified pathway outside of RdDM may be involved.

This study aims to investigate the transcriptional and transpositional regulation of *ONSEN* from three perspectives: (1) the relationship between transcription and transposition, (2) the role of epigenetic modification mechanisms, and (3) the identification of novel regulatory mechanisms. The goal is to elucidate the underlying regulatory processes governing *ONSEN*'s activity, contributing to a broader understanding of retrotransposon regulation. The study is divided into three chapters, each addressing one of these aspects: **Chapter One** focuses on the role of downstream factors in the RdDM pathway in regulating *ONSEN* transcription and transposition. Using mutants, it investigates how these factors contribute to the silencing of *ONSEN* at both the transcriptional and transpositional levels. **Chapter Two** examines the relationship between histone modifications and *ONSEN* transcriptional regulation. Through the application of histone deacetylase (HDAC) inhibitors, it explores how histone modifications impact *ONSEN* regulation. **Chapter Three** investigates new regulatory mechanisms using EMS-induced mutants to uncover novel pathways

involved in transposon regulation. This chapter aims to provide insights into previously unknown regulatory processes beyond the RdDM pathway.

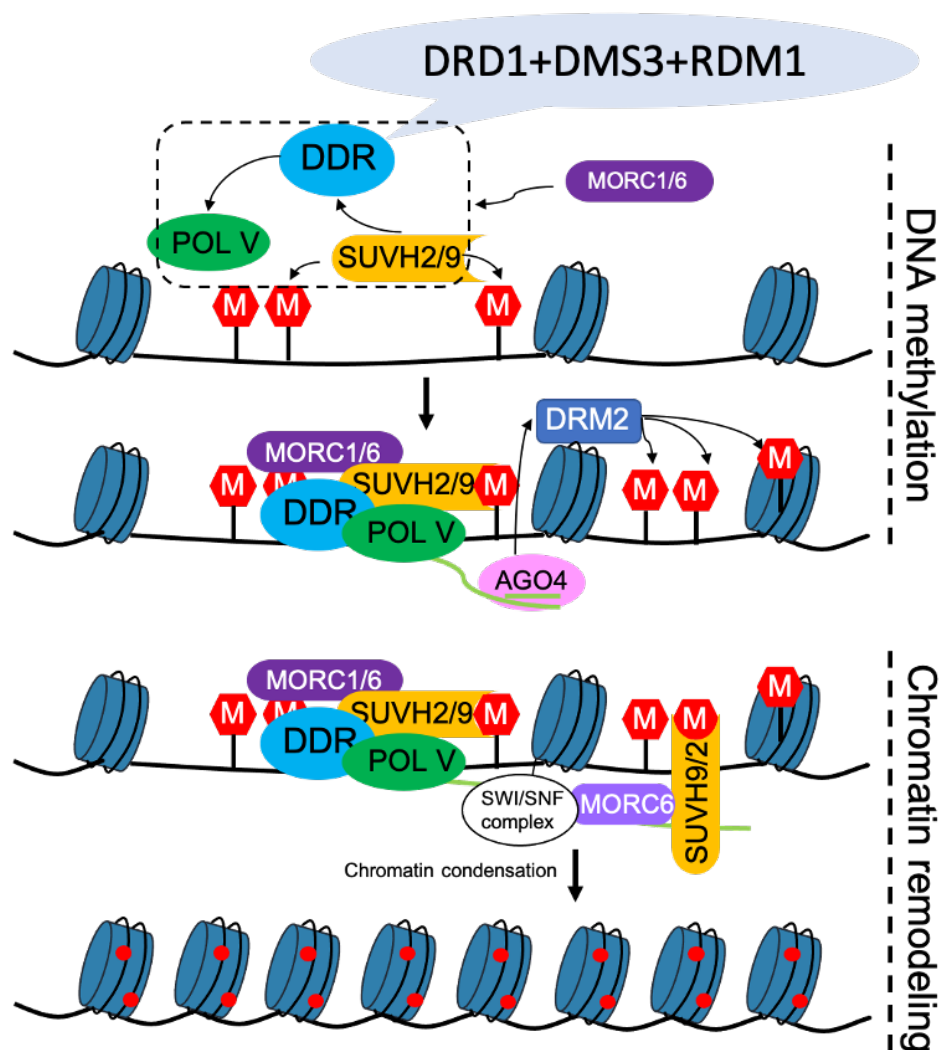
Chapter 1: Analysis of Transpositional Regulation of *ONSEN* on RdDM Pathway

Introduction

The nrpd1 and nrpe1 mutants are loss-of-function mutants for Pol IV and Pol V, respectively, which are key components of the RdDM pathway. Pol IV and Pol V are essential for the synthesis of siRNAs and long non-coding RNAs (lncRNAs) (Onodera Y et al., 2005). The downstream factors of the RdDM pathway primarily function through Pol V, contributing to the establishment of DNA methylation and chromatin condensation (Matzke M et al., 2009). Among these downstream components, the DDR complex—comprising DRD1, DMS3, and RDM1—plays a critical role in RdDM. Additionally, members of the SUVH protein family, including SUVH2, SUVH9, and MORC6, along with the closely related SUVH4, SUVH5, and SUVH6, are involved in facilitating transcriptional gene silencing and chromatin modification within the canonical RdDM pathway (Fig S2).

The DDR (DRD1-DMS3-RDM1) complex is essential for supporting Pol V-dependent transcription in the RdDM pathway. It functions upstream of Pol V by regulating its activity and stabilizing Pol V binding to chromatin (Law J.A et al., 2010). The DDR complex consists of DRD1, DMS3, and RDM1. The absence of any of these components disrupts RdDM activity (Zhong X et al., 2012). DMS3 and RDM1 form a stable DR complex, which subsequently associates with DRD1 (Wongpalee S.P. et al., 2019). DRD1 is a plant-specific protein similar to the chromatin remodeling protein SWI2/SNF2, facilitating the recruitment of Pol V to intergenic regions where it helps form repressive heterochromatin, thereby silencing adjacent transposons (Huetzel B et al., 2007; Kanno T et al., 2005; Wierzbicki et al., 2008). The structure of DMS3

resembles the hinge domain found in SMC (Structural Maintenance of Chromosomes) proteins, suggesting its involvement in the production of secondary siRNAs. Proper levels of DMS3 are critical for DDR complex assembly and RdDM activity, as excess DMS3 can impair RdDM function (Zhong S et al., 2019). RDM1, a small plant-specific protein, plays a key role in linking Pol II, DRM2, and AGO4 to guide DNA methylation at siRNA-complementary sites (Sasaki T et al., 2014). Downstream in the RdDM pathway, RDM1 can bind to single-stranded methylated DNA and may also interact with single-stranded DNA at Pol V transcription sites, facilitating the recruitment of Pol V to RdDM targets near the peri-nuclear region (Gao Z et al., 2010).



Supplementary Figure 2: Diagram of the Working Mechanism of RdDM Downstream Factors.

The DDR complex, composed of DRD1, DMS3, and RDM1, acts upstream of POL V to regulate its activity and stabilize its binding to chromatin. SUVH2 and SUVH9 bind to preexisting methylated DNA and interact with the DDR complex. This promotes the production of noncoding RNAs and induces de novo DNA methylation.

Arabidopsis has ten different SU(VAR)3-9 HOMOLOG (SUVH) family members, with KRYPTONITE (KYP or SUVH4), SUVH5, and SUVH6 serving as the primary histone H3K9 methyltransferases (Jackson J et al., 2002; Naumann K et al., 2005). Although SUVH2 and SUVH9 lack histone methyltransferase activity, their SET and RING-associated (SAR) domains are closely linked to DNA methylation activity (Johnson L.M et al., 2008). MORC6 is an additional factor required for downstream gene silencing via DNA methylation. While it plays a lesser role in establishing methylation, MORC6 is crucial for regulating chromatin structure and gene silencing across various eukaryotes (Moissiard G et al., 2012; Brabbs TR et al., 2013). SUVH2 and SUVH9 bind to pre-existing methylated DNA, interact with the DDR complex, and promote the recruitment of Pol V to target loci, facilitating the production of long non-coding RNAs. MORC1/6 enhances the function of the DDR-SUVH complex. In terms of chromatin regulation, MORC6 is recruited by SUVH2 and SUVH9, supporting transposable element (TE) silencing through SWI/SNF-dependent chromatin remodeling (Liu ZW et al., 2014; Jing Y et al., 2016).

This chapter examines the transcriptional and transpositional regulation of *ONSEN* by downstream factors of the RdDM pathway. The study primarily highlights the

suppressive roles of DDR complexes and SUVH2 in regulating *ONSEN* activity within the RdDM pathway. It further clarifies the independent functions of RdDM factors in inhibiting both the transcription and transposition of retrotransposons, providing new insights into the distinct regulatory mechanisms governing retrotransposon activity.

Material

All mutants used in this study were developed from the *Arabidopsis thaliana* accession Columbia (Col). The *drd1-6* (CS69758), *dms3-1* (CS69174), and *rdm1-4* (CS69759) mutants were sourced from Ohio State University, Columbus, OH, USA. The *drd1-6* mutant was crossed with either the *dms3-1* or *rdm1-4* mutants to generate the *drd1-6/dms3-1* and *drd1-6/rdm1-4* lines, respectively. Subsequently, the *drd1-6/dms3-1* line was crossed with the *drd1-6/rdm1-4* line to generate the *drd1-6/dms3-1/rdm1-4* triple mutant. The *suvh2* (SALK_079574) and *suvh9* (SALK_048033) mutants were obtained from the ABRC stock center, and the *nrpd1* (C366150) mutant was also acquired from Ohio State University. Crosses between *suvh2* and either *suvh9* or *nrpd1* mutants produced the *suvh2/9* and *suvh2/nrpd1* double mutants, respectively. The *morc6* mutant (GK_599B06) was provided by Professor Yoko Ikeda at the University of Okayama. The *suvh456* triple mutant was constructed from *kyp-4* (SALK_044606), *suvh5-2* (Ebbs ML and Bender J, 2006), and *suvh6-1* (Ebbs ML et al., 2005) and was provided by Professor Taiko To from Tokyo Institute of Technology.

Results

1. The Transpositional Regulation of *ONSEN* by Downstream Factors of RdDM

1.1 Loss of Downstream RdDM Pathway Factors Increases *ONSEN* Copy Number Under Heat Stress

ONSEN can be activated by HS at 37°C. Once transcribed, *ONSEN* is reverse-transcribed into complementary DNA (cDNA), which reflects its copy number. To assess the role of downstream factors in the RdDM pathway in silencing *ONSEN* transposition, I measured the *ONSEN* copy number in *Arabidopsis* Col and the mutants *drd1-6*, *dms3-1*, *rdm1-4*, *suvh2*, *suvh9*, *morc6*, *nrpe1*, and *nrpd1* after 24 hours of HS. The results indicate that after 24 hours of exposure to elevated temperatures, the *ONSEN* copy numbers were significantly higher in the *drd1-6*, *dms3-1*, *suvh2*, *nrpe1*, and *nrpd1* mutants compared to the Col (Fig. 1). These findings suggest that DRD1, DMS3, SUVH2, NRPE1, and NRPD1 play key roles in suppressing *ONSEN* activation under heat stress conditions.

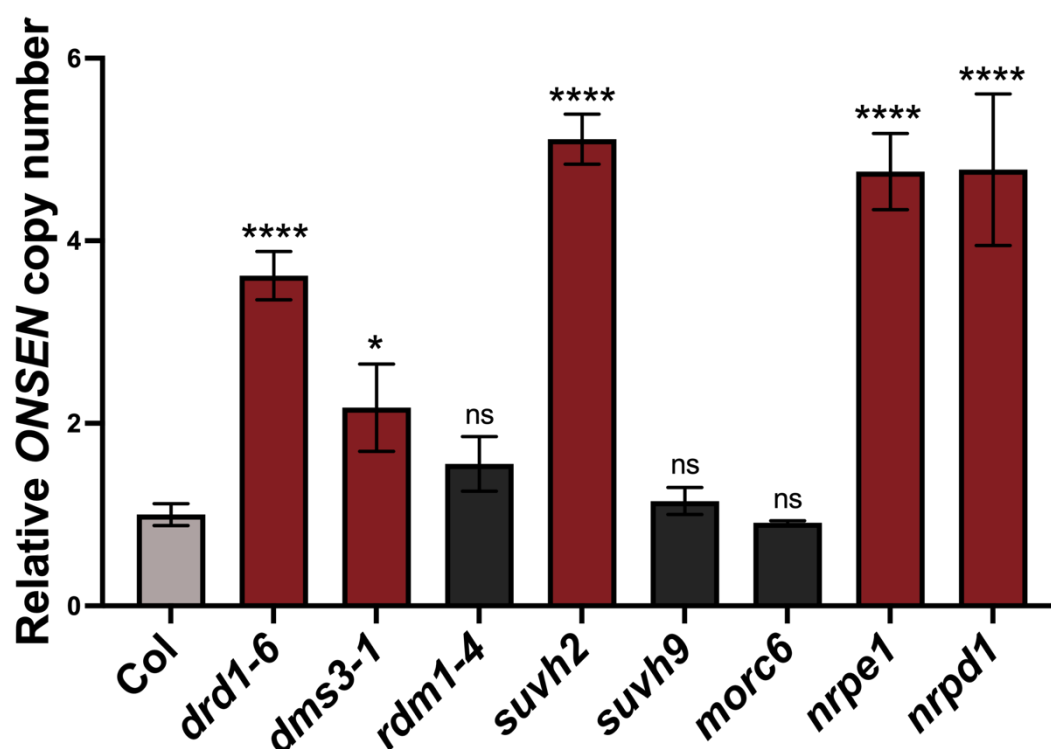


Figure 1: *ONSEN* Copy Number in RdDM-Related Mutants.

The bar graph represents the relative *ONSEN* copy number under heat stress (HS) conditions in Col, and the RdDM-related mutants *drd1-6*, *dms3-1*, *rdm1-4*, *suvh2*, *suvh9*,

morc6, *nrpe1*, and *nrpd1*. Asterisks indicate significant differences compared to Col (one-way ANOVA and Dunnett's test, *. $P < 0.05$; **. $P < 0.01$; ***. $P < 0.001$; ****. $P < 0.0001$). "ns" denotes no significant differences compared to Col (one-way ANOVA and Dunnett's test, $P > 0.05$).

1.2 Duration of Heat Stress Influences *ONSEN* Activity

Before analyzing *ONSEN* transposition, I first investigated how the duration of HS affects *ONSEN* activity by measuring its transcriptional levels in Col and *nrpd1* mutants after 24-hour and 48-hour heat treatments. The results show that after 48 hours of HS, the transcriptional level of *ONSEN* in Col and *nrpd1* was significantly higher than after 24 hours of HS (Fig. 2A), indicating that the duration of HS influences *ONSEN* transcription. Additionally, qPCR analysis revealed that *ONSEN* copy numbers in *nrpd1* increased after 48 hours of HS compared to 24 hours (Fig. 2B). To further examine *ONSEN* transposition, I performed Southern blotting under the two HS conditions. As expected, *ONSEN* transposition occurred in the *nrpd1* mutants after both 24-hour and 48-hour treatments, with a higher transposition frequency observed at 48 hours (Fig. 2C, D). These results demonstrate that the duration of HS significantly impacts *ONSEN* activity, affecting its transcriptional level, copy number, and transposition frequency.

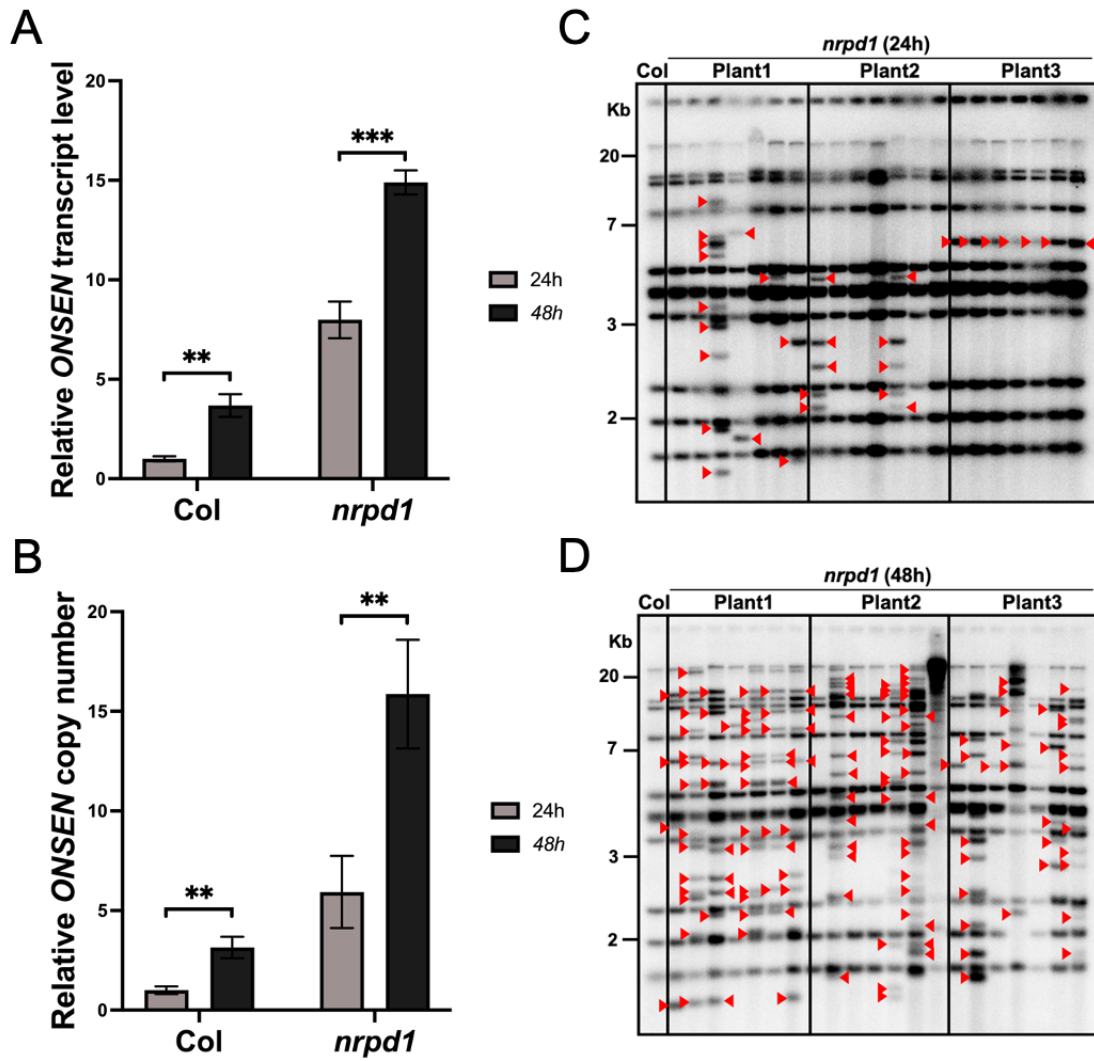


Figure 2: Impact of Heat Stress Duration on *ONSEN* Activation.

(A) and (B) Relative transcription levels (A) and relative copy numbers (B) of *ONSEN* in Col and *nrpd1*, measured by RT-qPCR and qPCR under heat stress (HS) conditions for 24 and 48 hours. Asterisks indicate significant differences between the 24-hour and 48-hour HS treatments (Student's t-test, *. $P < 0.05$; **. $P < 0.01$; ***. $P < 0.001$).

(C) and (D) Southern blot analysis of *ONSEN* in the progeny of *nrpd1* subjected to HS for 24 hours (C) and 48 hours (D). Seven individuals from each of the three parent plants subjected to heat stress were analyzed. The leftmost panel in each figure shows Col under non-stress (NS) conditions. Red arrows indicate new *ONSEN* insertions.

1.3 RdDM Downstream Factors Are Essential for Silencing *ONSEN* Transposition

Since 48 hours of HS significantly enhances *ONSEN* activity, I conducted 48-hour heat treatments on the *drd1-6*, *dms3-1*, *rdm1-4*, *suvh2*, *suvh9*, and *morc6* mutants. DNA was then extracted from their progeny and analyzed using Southern blotting. The results revealed that *ONSEN* transposition occurred in the *drd1-6*, *suvh2*, and *nrpe1* mutants (Fig 3). These findings, along with previous research by Hayashi Y et al. (2020), which reported *ONSEN* transposition in *nrpe1* mutants, suggest that DRD1, SUVH2, and Pol V are essential for silencing *ONSEN* transposition. Interestingly, although the *dms3-1* mutant showed a high *ONSEN* copy number, the "free-floating *ONSEN*" did not appear to have successfully integrated into the chromosomes (Fig 1).

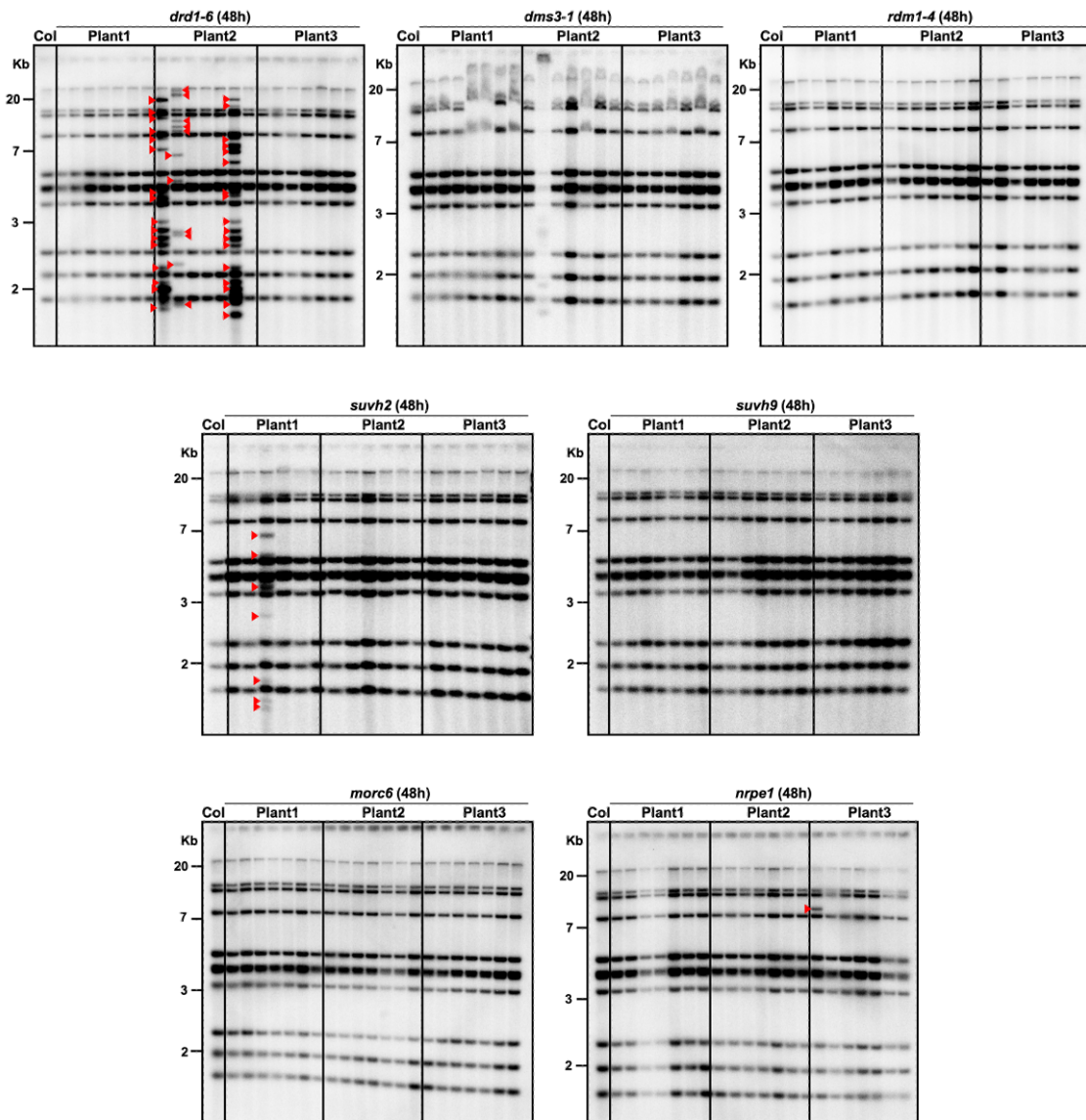


Figure 3: Southern Blotting of RdDM Downstream Factors.

Transposition analysis of *ONSEN* under heat stress (HS) conditions for 48 hours in *drd1-6*, *dms3-1*, *rdm1-4*, *suvh2*, *suvh9*, *morc6*, and *nrpe1*. Seven individuals from each of the three parent plants subjected to heat stress were analyzed. The leftmost panel in each figure shows Col under non-stress (NS) conditions. Red arrows indicate new *ONSEN* insertions.

2. Transcriptional Regulation of *ONSEN* by Downstream RdDM Factors

Given the critical roles of DRD1 and SUVH2 in regulating *ONSEN* transposition, these two factors became the focus of further investigation. In the following sections, I will explore the regulatory roles of the DDR complex and SUVH proteins in controlling both the transcription and transposition of *ONSEN*.

2.1 Regulation of *ONSEN* Transcription and Transposition by DDR Mutants

DRD1, a key component of the DDR complex, plays a critical role in the RdDM pathway. Under 24-hour HS conditions, the transcript levels of *ONSEN* were significantly increased in the *dms3-1*, *rdm1-4*, and *nprdl* mutants compared to Col, while *ONSEN* transcripts were only slightly increased in *drd1-6* ($P = 0.0530$) (Fig 4A). To further investigate the role of the DDR complex in regulating *ONSEN* transcription, I measured *ONSEN* expression levels in DDR complex mutants under HS conditions. After 48 hours of HS, *ONSEN* transcript levels were not significantly higher in *drd1-6/dms3-1*, *drd1-6/rdm1-4*, and DDR mutants than in *drd1-6* (Fig 4B). This indicates that DRD1-6, DMS3-1, and DDM1-4 have redundant roles in the transcriptional regulation of *ONSEN*. *ONSEN* transcription in DDR mutants reached levels comparable to those in *nprdl* (Fig 4B), indicating that the DDR complex contributes to the transcriptional repression of *ONSEN* in the RdDM pathway. I analyzed *ONSEN* transposition activity using Southern blotting in the *drd1-6/dms3-1*, *drd1-6/rdm1-4*, and DDR triple mutants. Although *ONSEN* transposition was not detected under 24-hour HS (Fig S3), transposition was observed in the *drd1-6/dms3-1* double mutants and DDR triple mutants after 48 hours of HS (Fig 4C). These results suggest that the DDR complex, particularly through the action of DRD1, plays a critical role in suppressing *ONSEN* transposition. In summary, while the absence of individual DDR complex

components or the entire complex can relieve *ONSEN*'s transcriptional silencing, DRD1 plays a crucial role in repressing *ONSEN* transposition.

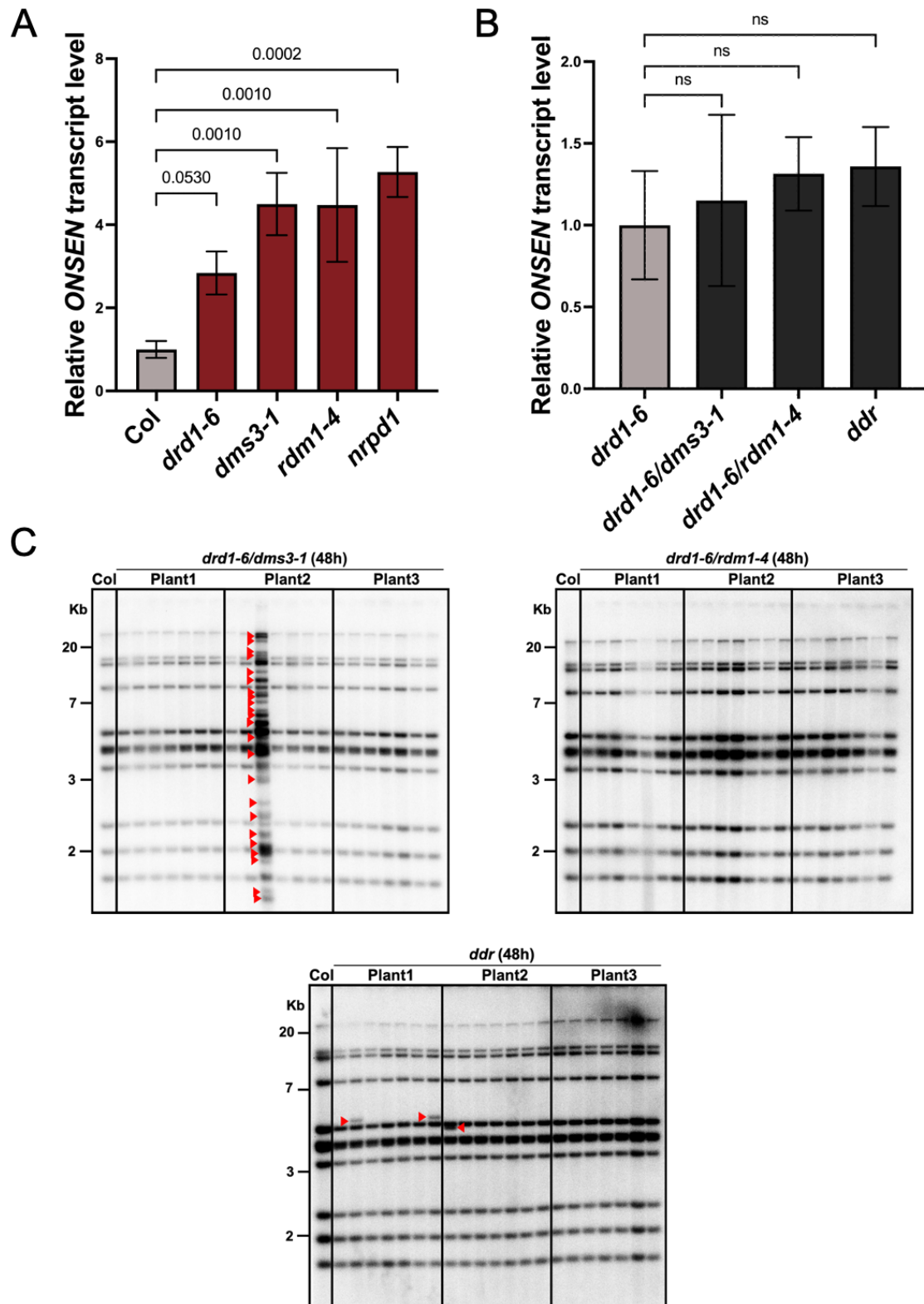
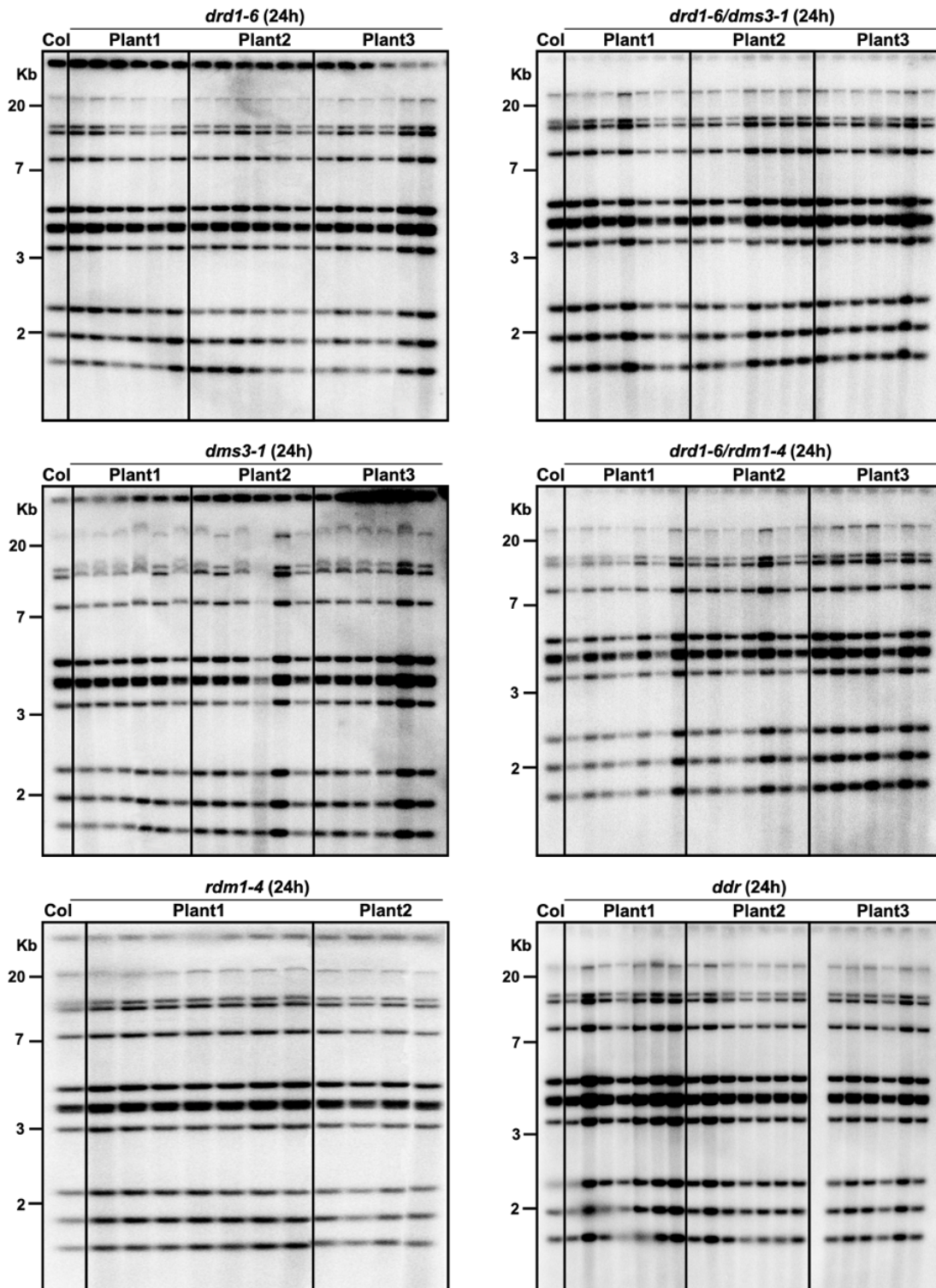


Figure 4: *ONSEN* Activation in DDR Complex-Associated Mutants Subjected to Heat Stress (HS).

(A) Relative transcription level of *ONSEN* in *drd1-6*, *dms3-1*, and *rdm1-4* after 24 hours of HS. Significant differences between groups were determined by one-way ANOVA and Dunnett's test (P values indicated).

(B) Relative transcription levels of *ONSEN* in *drd1-6*, *drd1-6/dms3-1*, *drd1-6/rdm1-4*, and *ddr* after 48 hours of HS. "ns" indicates no significant differences between the two groups (one-way ANOVA and Dunnett's test, $P > 0.05$).

(D) Southern blot analysis of *ONSEN* in the next-generation progeny of *drd1-6/dms3-1*, *drd1-6/rdm1-4*, and *ddr* under 48 hours of HS. Seven individuals from each of the three parent plants subjected to heat stress were analyzed. The leftmost panel in each figure shows the Col pattern under non-stress (NS) conditions. Red arrows indicate new *ONSEN* insertions.



Supplementary Figure 3: No *ONSEN* Insertion Observed in DDR Complex Mutants After 24 Hours of HS.

Southern blot analysis of *ONSEN* in the next-generation DDR complex mutants under 24h of HS. Seven individuals from each of the three parent plants subjected to heat

stress were analyzed. The leftmost panel in each figure shows the Col pattern under non-stress (NS) conditions.

2.2 Regulation of *ONSEN* Transcription and Transposition by SUVH Family Proteins (SUVH 2 and SUVH 9)

To investigate the role of SUVH family proteins in *ONSEN* regulation, I analyzed *ONSEN* transcript levels in the *suvh2* and *suvh9* mutant. After 24 hours of HS, the transcript level of *ONSEN* in *suvh2* was significantly higher than in Col, indicating that the loss of *suvh2* relieves transcriptional silencing of *ONSEN* (Fig 5A). The *ONSEN* transcript level in *suvh9* remained comparable to that in Col (Fig 5A). Since SUVH2 and SUVH9 exhibit partial redundancy in the RdDM pathway (Kuhlmann M and Mette M.F, 2012), I further explored their cooperative role by examining *ONSEN* transcript levels in the *suvh2* mutant and the *suvh2/9* double mutant. After 48 hours of HS, the transcript levels of *ONSEN* in the *suvh2/9* double mutant were not higher than those in the *suvh2* single mutant (Fig 5B). These results suggest that while the loss of SUVH2 relieves the transcriptional repression of *ONSEN*, SUVH9 does not appear to play an essential role in the RdDM-mediated suppression of *ONSEN* transcription.

ONSEN transposition was also observed in the *suvh2/9* double mutants after 48 hours of HS, consistent with the results seen in the *suvh2* single mutant (Fig 5C, and see Fig 3). This further confirms that SUVH2 is a critical factor in suppressing both *ONSEN* transcription and transposition. However, the transposition frequency of *ONSEN* in the *suvh2* mutant was notably lower than that observed in *nrpd1* (see Fig 2E, and also Fig 3). The simultaneous loss of SUVH2 and SUVH9 did not result in an increased transposition frequency of *ONSEN* (Fig 5C, and see Fig 3), suggesting that SUVH2 and

SUVH9 do not act cooperatively to suppress *ONSEN* transposition. Additionally, consistent with the results in DDR complex mutants, no new *ONSEN* insertions were detected in *suvh2* mutants after 24 hours of HS (Fig S4). Although *ONSEN* cDNA accumulated in large amount in *suvh2* following HS, its integration into new genomic locations remained limited. To further investigate *ONSEN* activity, I analyzed *suvh2/nrpd1* double mutants. After 24 hours of heat stress (HS), there were an average of 18 new *ONSEN* insertions in the *suvh2/nrpd1* double mutant (calculated as 55 total insertions divided by 3 biological replicates), and around 10 insertions in the *nrpd1* single mutant (31 total insertions divided by 3 replicates) (see Fig. 5D and Fig. 2C). However, this increase at 24 hours was not statistically significant. In contrast, after 48 hours of HS, the *nrpd1* single mutant showed a significant increase, with an average of 46 new *ONSEN* insertions (137 total insertions divided by 3 replicates). Despite the higher *ONSEN* copy number in the *suvh2/nrpd1* double mutants compared to the *suvh2* or *nrpd1* single mutants (Fig 5E), these results suggest that SUVH2 and NRPD1 together primarily influence *ONSEN* transcriptional regulation.

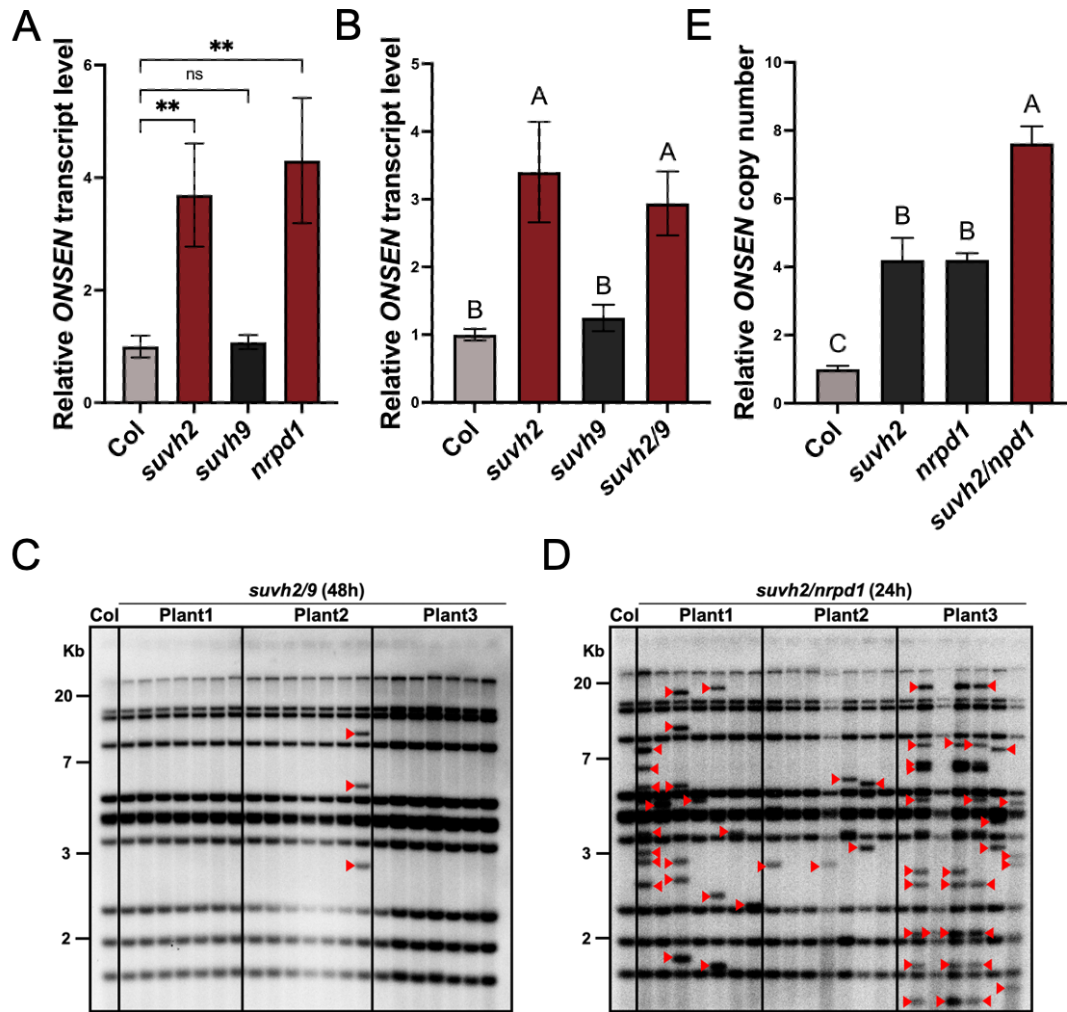


Figure 5: ONSEN Transposition is Infrequent in *suvh2* Mutants.

(A) Relative transcription level of *ONSEN* in *suvh2* and *suvh9* after 24 hours of heat stress (HS). (one-way ANOVA and Dunnett's test, *. $P < 0.05$; **. $P < 0.01$). "ns" indicates no significant differences between the two groups (one-way ANOVA and Dunnett's test, $P > 0.05$).

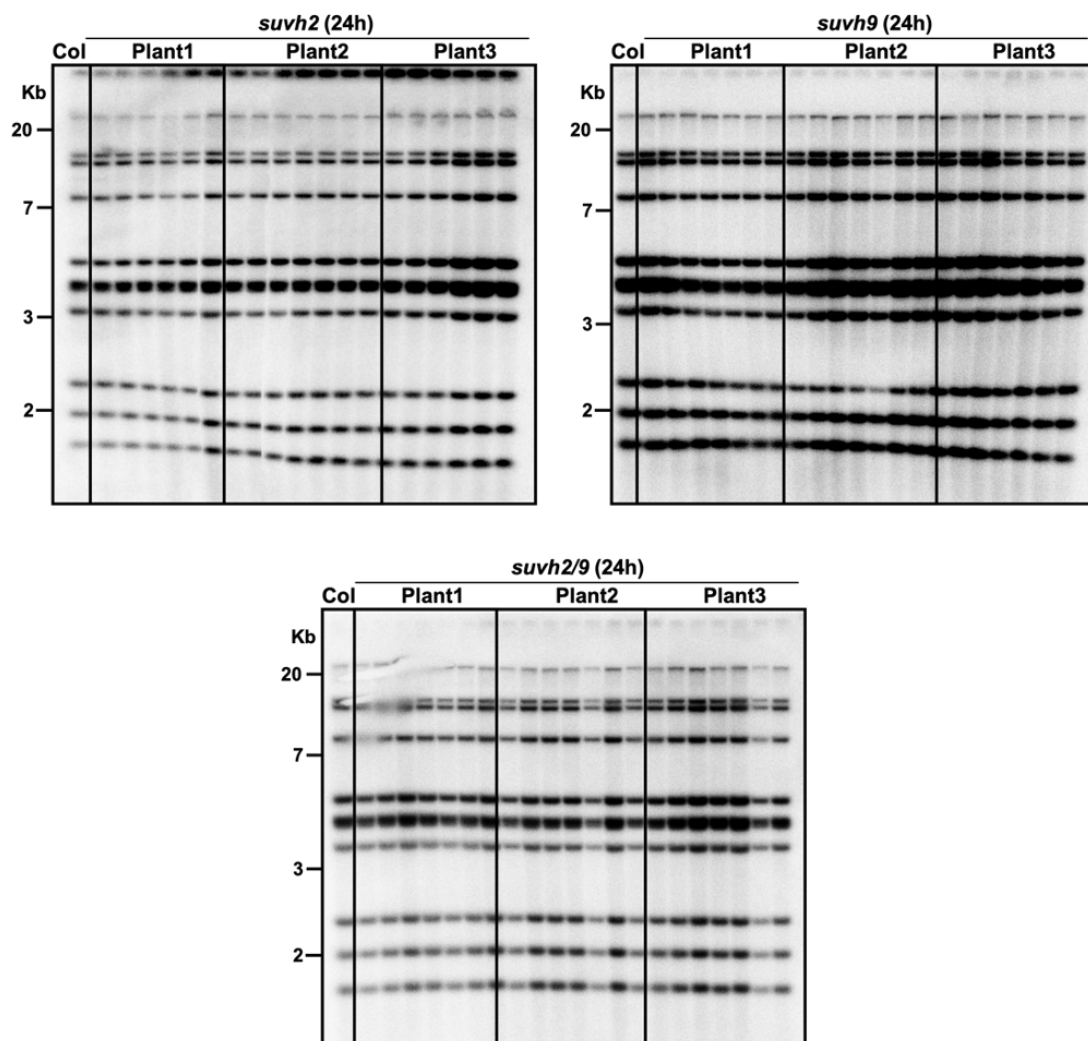
(B) Relative transcription levels of *ONSEN* in *suvh2*, *suvh9*, and *suvh2/9* after 48 hours of HS. Different letters indicate significant differences between samples (one-way ANOVA and Tukey test, $P < 0.05$).

(C) Southern blot analysis of *ONSEN* in the next-generation *suvh2/9* mutants after 48 hours of HS. Seven individuals from each of the three parent plants subjected to heat

stress were analyzed. The leftmost panel shows the Col under non-stress (NS) conditions. Red arrows indicate new *ONSEN* insertions.

(D) Southern blot analysis of the progeny of *suvh2/nrpd1* mutants after 24 hours of HS to examine the transgenerational transposition of *ONSEN*. The leftmost lane represents the Col under NS conditions. Three homozygous mutant plants were subjected to HS for 24 hours, and seven progenies from each plant were analyzed for *ONSEN* transposition. Red triangles indicate new *ONSEN* insertions.

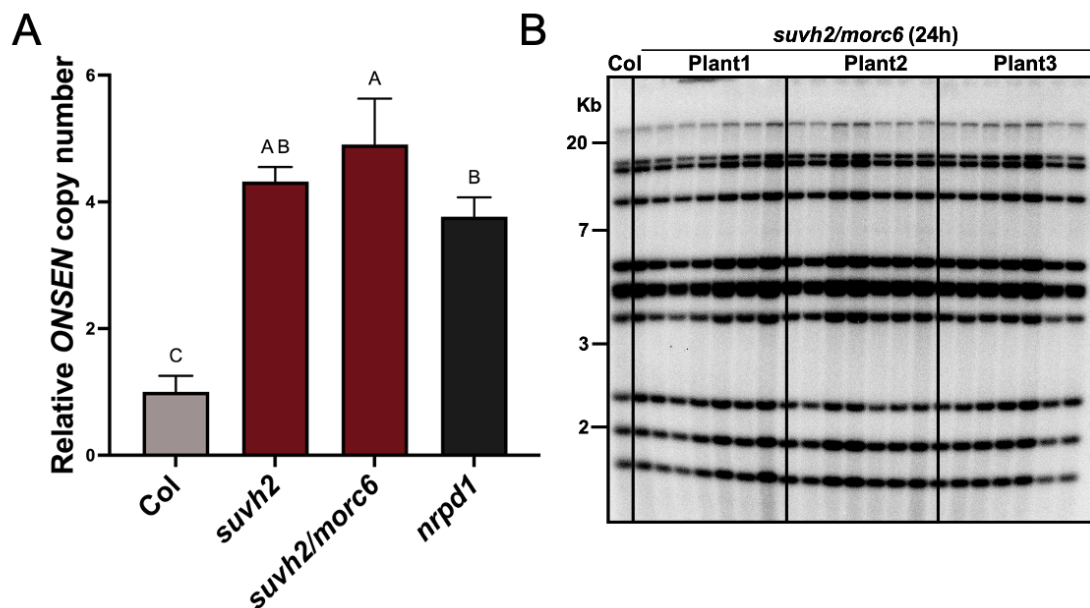
(E) The bar graph shows *ONSEN*'s relative copy number after 24 hours of HS in Col, *suvh2*, *nrpd1*, and *suvh2/nrpd1*. Different letters indicate significant differences between samples (one-way ANOVA and Tukey test, $P < 0.05$).



Supplementary Figure 4: No *ONSEN* Transposition Observed in *suvh2* and *suvh2/9* Mutants After 24 Hours of Heat Stress.

Southern blot analysis of *ONSEN* in the next-generation *suvh2*, *suvh9*, and *suvh2/9* under 24h of HS. Seven individuals from each of the three parent plants subjected to heat stress were analyzed. The leftmost panel in each figure shows the Col pattern under non-stress (NS) conditions.

In addition, MORC6 is recruited by SUVH2 and SUVH9 to enhance TE silencing through SWI/SNF-dependent chromatin remodeling (Liu ZW et al., 2014; Jing Y et al., 2016). However, in the *suvh2/morc6* double mutant, the *ONSEN* copy number did not increase significantly compared to *suvh2*, and no *ONSEN* transposition was observed (Fig S5A, B). These findings suggests that the suppressive effect of SUVH2 on *ONSEN* may function independently of MORC6.



Supplementary Figure 5: MORC6 Is Not Required for SUVH2-Mediated

Regulation of *ONSEN*.

(A) Relative copy numbers of *ONSEN* in *suvh2/morc6* mutants under heat stress (HS). Different letters indicate significant differences between samples (one-way ANOVA and Tukey test, $P < 0.05$).

(B) The next generation of *suvh2/morc6* mutants after 24 hours of HS was analyzed by Southern blot to examine the transgenerational transposition of *ONSEN*. The leftmost lane represents the Col under non-stress (NS) conditions. Three homozygous mutant plants were subjected to HS for 24 hours, and seven progenies from each plant were analyzed for *ONSEN* transposition.

3. The Correlation Between DNA Methylation and *ONSEN* Silencing

DNA methylation levels are strongly correlated with the transcriptional regulation of genes and TEs. In the RdDM pathway, each factor plays a distinct role in establishing and maintaining DNA methylation, contributing to the suppression of TE activity. Therefore, analyzing methylation levels is essential to understanding how the RdDM pathway silences *ONSEN* and regulates its activity.

3.1 Dynamic Equilibrium of *ONSEN* Methylation Levels During Heat Stress

Previous studies have identified eight *ONSEN* copies in *A.thaliana*, each exhibiting varying levels of activity. Among these, *At1g11265*, *At3g61330*, and *At5g61305* show the highest activity (Cavrak V.V et al., 2014). To investigate the dynamics of *ONSEN* methylation, I analyzed the DNA methylation levels of *At1g11265* in Col and *nrpd1* mutants under non-stress (NS) and HS conditions (24 and 48 hours) using bisulfite sequencing (BS-seq). In both Col and *nrpd1* mutant, the levels of CG and CHG methylation in the *At1g11265* region after HS (24 hours and 48 hours) were comparable

to those observed under the non-stress (NS) condition (Fig 6A). The CHH methylation levels initially increased after 24 hours of HS but gradually returned to levels comparable to non-stress (NS) conditions after 48 hours (Fig 6B). These results suggest that *ONSEN* methylation levels undergo dynamic equilibrium during heat treatment, with methylation patterns fluctuating over the course of the stress response.

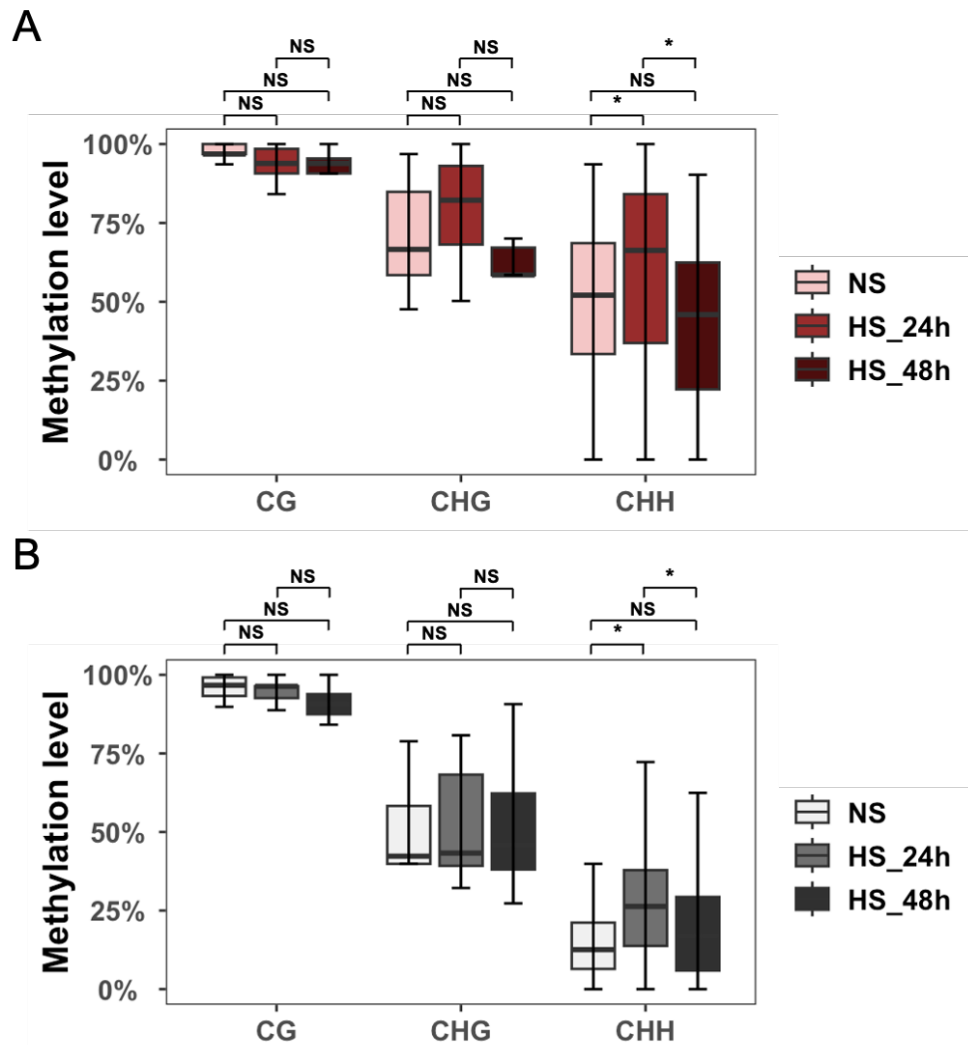


Figure 6: Relationship Between Duration of Heat Treatment and *ONSEN* DNA Methylation.

(A) and (B) Boxplots showing CG, CHG, and CHH (H = A, T, C) methylation levels of *ONSEN* (*At1g11265*) in *Col* (A) and *nprdl* (B) mutants under non-stress (NS), 24-hour,

or 48-hour heat stress (HS) conditions. Significant differences between groups are marked with asterisks (Wilcoxon test and Benjamini-Hochberg method, $P < 0.05$). "ns" indicates no significant differences between the two groups (Wilcoxon test and Benjamini-Hochberg method, $P > 0.05$).

3.2 DNA Hypomethylation Influences *ONSEN* Transcription but Does Not Directly Affect Transposition

To explore how DNA methylation impacts *ONSEN* activity, I analyzed the methylation levels of three different *ONSEN* copies (*At1g11265*, *At3g6133*, and *At5g13205*) in Col, *drd1-6*, DDR complex mutants (*ddr*), and *nRPD1* after 48 hours of HS. Compared to the Col, CHH methylation levels in these *ONSEN* copies were significantly reduced in *drd1-6*, *ddr*, and *nRPD1* mutants (Fig 7A). In the *drd1-6* mutant, CG methylation levels were significantly reduced in *At1g11265* and *At5g13205*, while in *ddr* and *nRPD1* mutants, CG methylation levels of *At3g6133* were also markedly decreased (Fig 7A). In *drd1-6*, *ddr*, and *nRPD1* mutants, the CHH methylation levels of *At1g11265*, *At3g6133*, and *At5g13205* were significantly reduced (Figure 7A). Although DNA methylation levels showed slight change across different *ONSEN* copies, the overall DNA methylation pattern remained similar (Fig 7A). I hypothesize that these differences in methylation levels of *At1g11265*, *At3g6133*, and *At5g13205* contribute to their varying transcriptional activities. In addition, impaired methylation in the *drd1-6* and *ddr* mutants might account for the widespread upregulation of *ONSEN* transcription (Fig 7A). However, despite the reduction in methylation levels, the transposition frequency of *ONSEN* in *drd1-6* and *ddr* mutants remained much lower than in the *nRPD1* mutant (see Fig 2D, Fig 3). These findings suggest that while

hypomethylation promotes *ONSEN* transcription, it does not relieve transpositional silencing. This conclusion is further supported by methylation studies involving *suvh2*.

Consistent with previous results, HS does not alter the DNA methylation levels of the *ONSEN* region in *suvh2* mutants (Fig S6). However, compared to Col, the CHH methylation levels of *ONSEN* are significantly reduced in the *suvh2* mutant (Fig 7B), indicating that SUVH2 is crucial for establishing CHH DNA methylation at *ONSEN*. This finding aligns with previous observations involving the single-copy SINE element *AtSN1* (Johnson L.M et al., 2008). The simultaneous loss of SUVH2 and SUVH9 also results in decreased CHH methylation levels (Fig 7B), suggesting that these proteins may cooperate in regulating CHH methylation. Interestingly, in the *suvh9*, the CHH methylation level of *ONSEN* is higher than in the Col (Fig 7B). Since the *ONSEN* promoter is located within the long terminal repeat (LTR) region, which primarily consists of CHH methylation sites, increased CHH methylation in *suvh9* may suppress *ONSEN* transcription (see Fig 5A). On the other hand, the DNA methylation pattern of *ONSEN* in *suvh2* closely mirrors that of *nrpd1* (Fig 7B), further supporting the idea that reduced DNA methylation alleviates transcriptional silencing of *ONSEN*. However, this reduction in methylation does not appear to significantly affect *ONSEN* transposition.

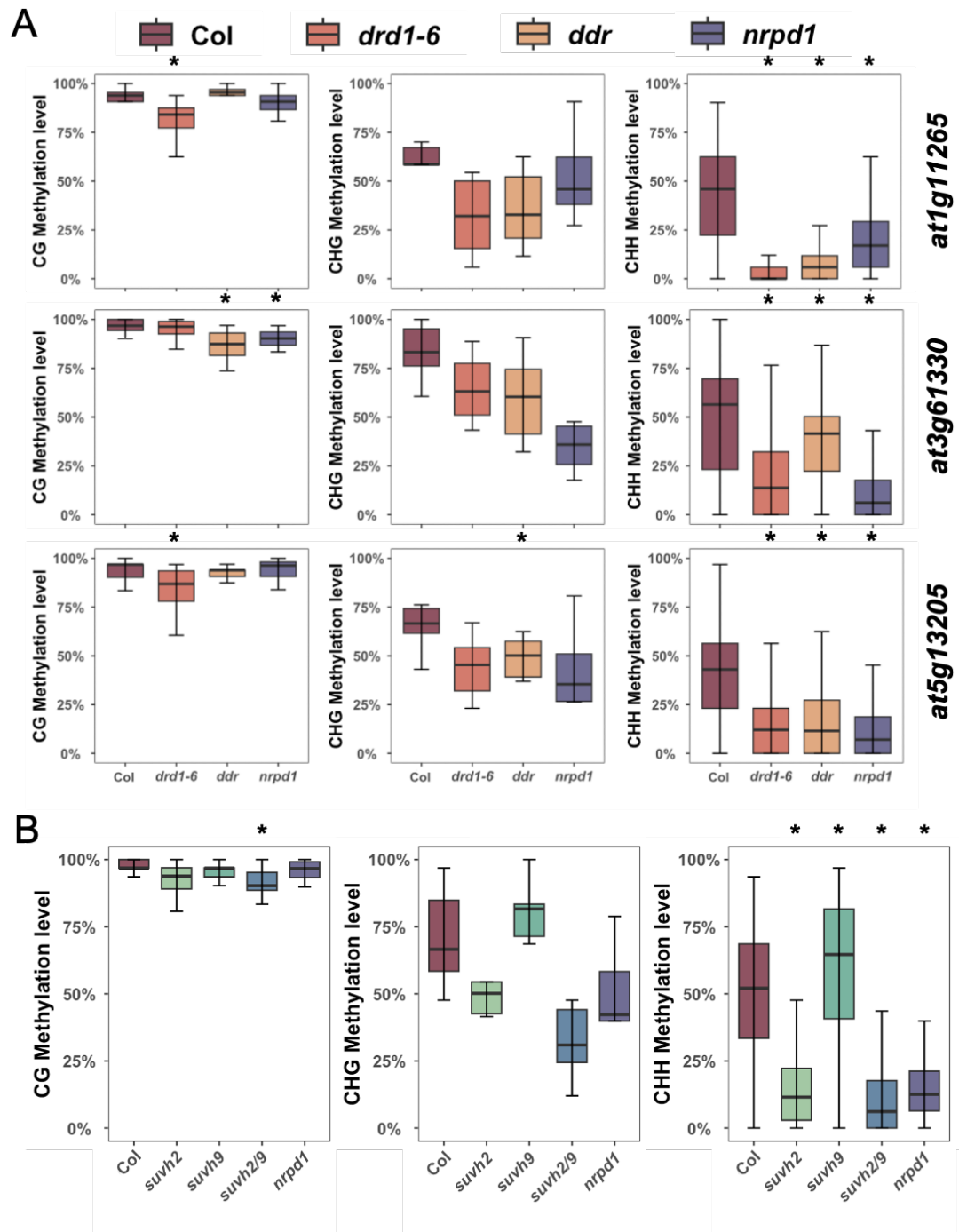
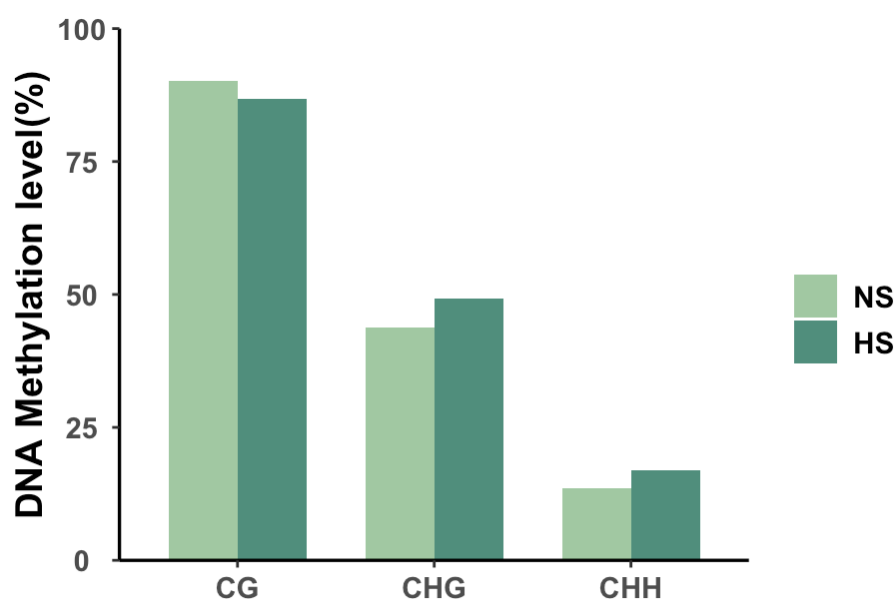


Figure 7: DRD1-6 and SUVH2 Are Essential for Establishing DNA Methylation on *ONSEN*.

(A) DNA methylation levels (CG, CHG, and CHH (H = A, T, C)) of *ONSEN* (*At1g11265*, *At3g61330*, and *At5g13205*) in *Col*, *drd1*, *ddr*, and *nrpd1* after 48 hours of heat stress (HS).

(B) Boxplots showing CG, CHG, and CHH (H = A, T, C) methylation levels of *ONSEN* (*Atlg11265*) in Col, *suvh2*, *suvh9*, *suvh2/9*, and *nrpd1* under non-stress (NS) conditions. Asterisks indicate significant differences in DNA methylation levels compared to *ONSEN* in Col (Wilcoxon test and Benjamini-Hochberg method, $P < 0.05$).

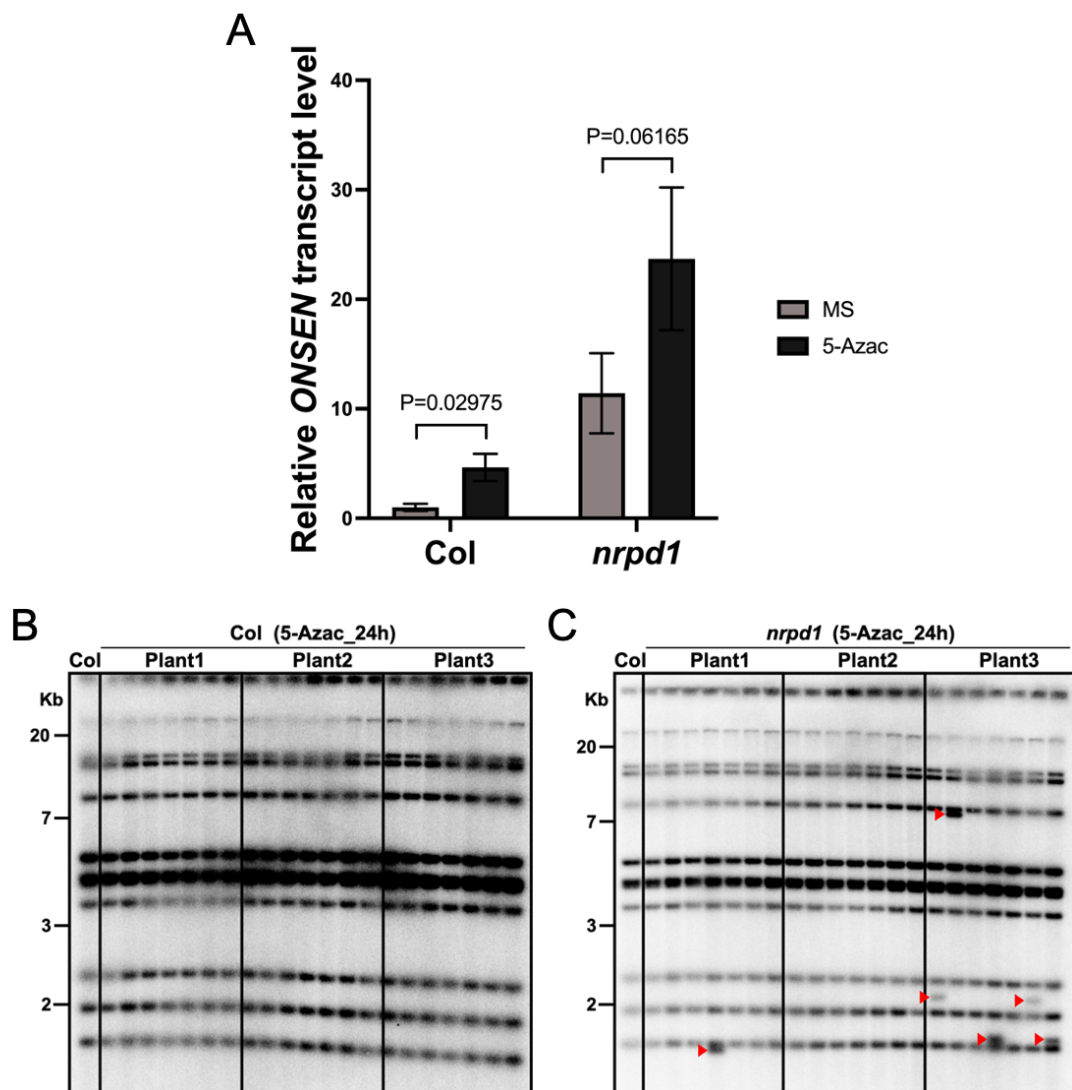


Supplementary Figure 6: HS Does Not Affect *ONSEN* DNA Methylation Levels in *suvh2*.

Average levels of CG, CHG, and CHH (where H = A, T, or C) methylation in the *ONSEN* (*Atlg11265*) region in *suvh2* under non-stress (NS) and HS conditions.

Additionally, 5-aza-2'-deoxycytidine (5-Azac) is a DNA methylation inhibitor that irreversibly binds to ZCyt residues in DNA, leading to a rapid loss of DNA (cytosine-C5) methyltransferase (Dnmt) activity (Christman J., 2002). Treatment with 5-Azac significantly increases the transcription levels of *ONSEN*, although only a slight increase is observed in *nrpd1* mutants (Fig S7A). However, removing DNA

methylation does not induce *ONSEN* transcription in Col, nor does it result in a higher transposition frequency in *nrpd1* (Fig S7 B, C). These findings suggest that while DNA methylation suppresses *ONSEN* transcription, it is not directly associated with regulating *ONSEN* transposition.



Supplementary Figure 7: Effect of 5-AzaC on *ONSEN* Activity.

(A) Transcription levels of *ONSEN* in Col and *nrpd1* grown on MS and 5-AzaC mediums under heat stress (HS). Significant differences between groups were determined by Student's t-test (P values indicated).

(B) and (C) The transgenerational transposition of *ONSEN* examined in Col (B) and

nrpd1 (C) treated with 5-AzaC under HS conditions. The leftmost panel in each figure shows the Col pattern under non-stress (NS) conditions. Red arrows indicate new *ONSEN* insertions.

4. Chromatin Condensation is Independent of *ONSEN* Activation

Chromatin suppression typically induces gene silencing through mechanisms independent of DNA methylation (Mittelsten S.O et al., 2002). In an open chromatin state, which facilitates the binding of transcription factors and RNA polymerase, promoting transcription. To assess the chromatin openness of the *ONSEN* long terminal repeat (LTR) region (Fig 8A), I performed FAIRE-qPCR. FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) technology is a method designed to isolate DNA regions that are not associated with histones, based on variations in the cross-linking efficiency between DNA and nucleosomes. A lower cross-linking efficiency (input (%)) indicates a more open chromatin structure. In both *suvh2* and *suvh2/9* mutants, the level of chromatin opening in the *ONSEN* LTR is identical to that in Col (Fig 8B). Since the *ONSEN* promoter resides within the LTR region, the absence of SUVH2 does not promote chromatin opening at the *ONSEN* LTR. It is important to note that HS activates *ONSEN* and leads to the production of extrachromosomal cDNA, complicating the FAIRE-based detection of the *ONSEN* region after HS treatment.

Next, DAPI staining was used to macroscopically assess heterochromatin condensation in mutant cell nuclei. Heterochromatin states were categorized into condensed and dispersed types (Fig 8C). The nuclei of Col, *drd1-6*, DDR complex mutants (*ddr*), *suvh2*, *suvh9*, *suvh2/9*, *nrpd1*, and *suvh2/nrpd1* were examined under both non-stress (NS) and 48h HS conditions. The results showed that HS induces heterochromatin

decondensation, leading to more dispersed nuclei (Fig 8D, E). However, the proportion of dispersed nuclei was similar across all mutants, regardless of HS treatment (Fig 8D, E). In summary, while HS promotes chromatin decondensation, the extent of decondensation does not correlate with the frequency of *ONSEN* transposition. These findings suggest that heterochromatin decondensation may facilitate *ONSEN* transcriptional activation but does not directly impact *ONSEN* transposition.

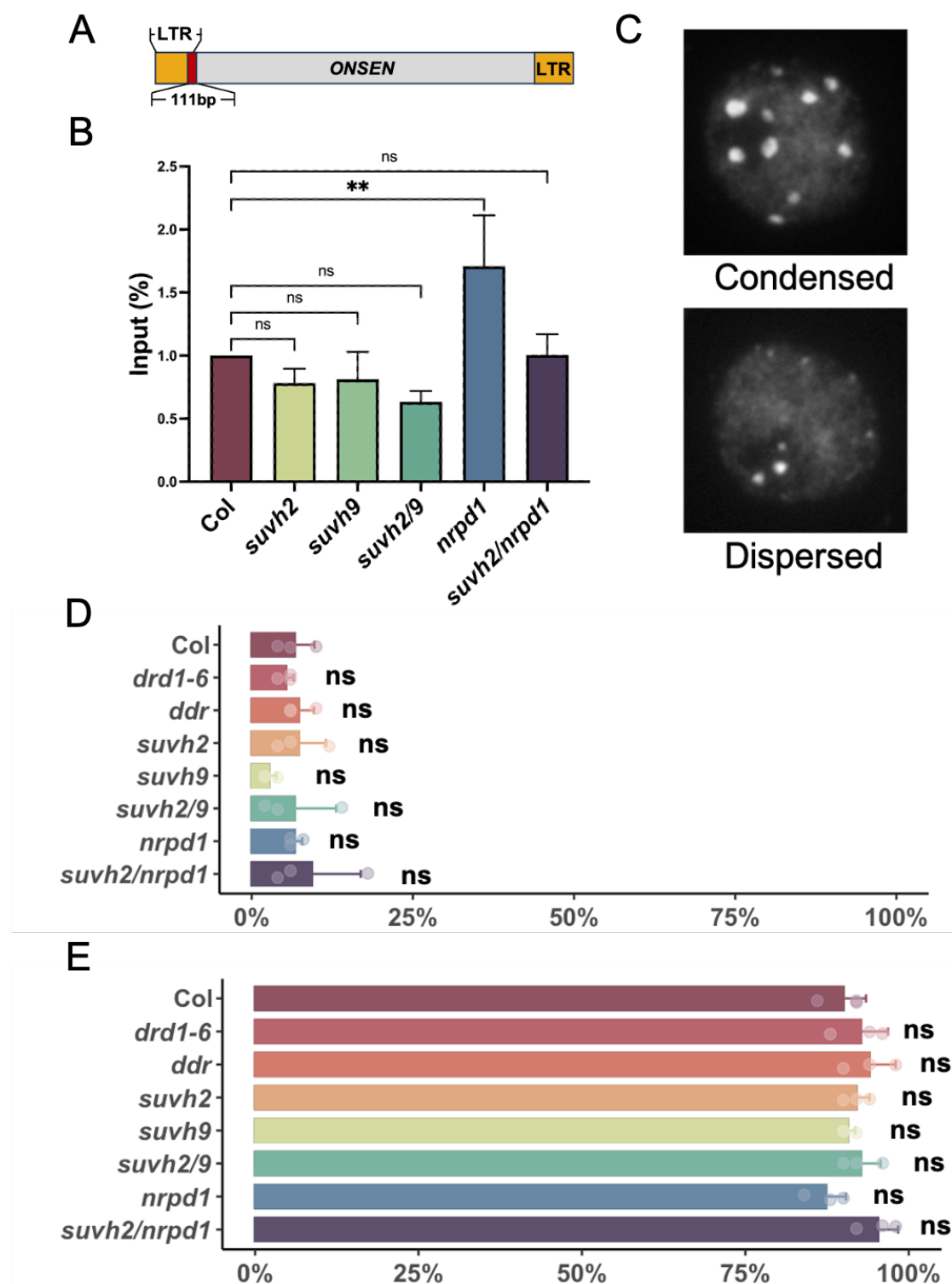


Figure 8: Analysis of Open Chromatin.

(A) Map of the size (111 bp) and location of the FAIRE-qPCR amplified region.

(B) FAIRE-qPCR results demonstrating the level of open chromatin in the LTR region

of *ONSEN* in Col, *suvh2*, *suvh9*, *suvh2/suvh9*, *nrpd1*, and *suvh2/nrpd1*. Asterisks indicate significant differences between the two groups (one-way ANOVA and Dunnett's test, *. $P < 0.05$; **. $P < 0.01$). "ns" indicates no significant differences between the two groups (one-way ANOVA and Dunnett's test, $P > 0.05$).

(C) Representative images of nuclei stained with DAPI.

(D) and (E) Percentage of chromatin in dispersed states for Col, *drd1-6*, *ddr*, *suvh2*, *suvh9*, *suvh2/9*, *nrpd1*, and *suvh2/nrpd1* under non-stress (NS) (D) and 48 hours of heat stress (HS) (E) conditions ($n = 150$). "ns" denotes no significant differences compared to Col (one-way ANOVA and Dunnett's test, $P > 0.05$).

5. siRNA Plays a Crucial Role in the Silencing Mechanism of *ONSEN*

5.1 Gradual Increase of siRNA Over Time

In the RdDM pathway, Pol IV is essential for siRNA synthesis. To explore the role of siRNA in *ONSEN* silencing, I analyzed the accumulation of *ONSEN*-related siRNA in various mutants. Specifically, I examined siRNA levels in the *ONSEN* LTR region in Col under both non-stress (NS) and HS conditions. The results showed that siRNA was present in the *ONSEN* LTR region under both non-stress (NS) and HS conditions. However, siRNA did not accumulate immediately following HS. Instead, its levels gradually increased during the recovery phase at 21°C (Fig 9A).

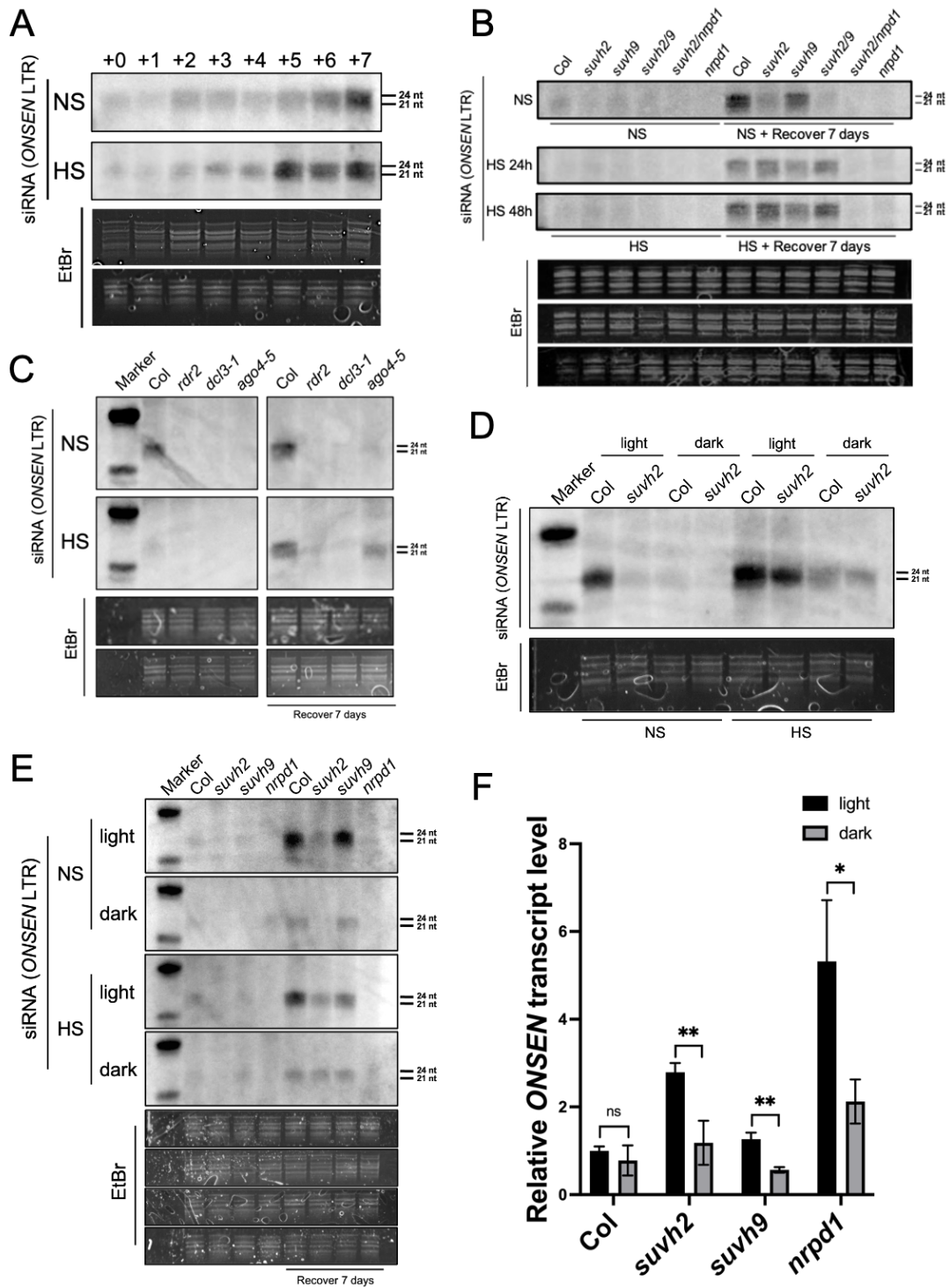


Figure 9: Analysis of siRNA Accumulation from the *ONSEN* LTR.

(A) siRNA accumulation levels in Col under non-stress (NS) and heat stress (HS) conditions. The numbers above indicate the number of days the plants recovered at ambient conditions (21°C) after being subjected to non-stress (NS) or HS at 37°C for

24 hours (+0), with recovery for one day (+1), two days (+2), and up to seven days (+7).

(B) siRNA accumulation levels in Col, *suvh2*, *suvh9*, *suvh2/9*, *nrpd1*, and *suvh2/nrpd1* in 7-week-old seedlings treated with non-stress (NS, 15-day-old seedlings) or HS (24 h or 48 h) and recovered at 21°C for seven days (+7) after treatment.

(C) siRNA accumulation in Col, *rdr2*, *dcl3-1*, and *ago4-5* in 7-week-old seedlings subjected to non-stress (NS) or HS and recovered at 21°C for seven days (+7) after treatment.

(D) siRNA accumulation in 7-week-old Col and *suvh2* seedlings after treatment with non-stress (NS) or HS and recovery for seven days (+7) at 21°C, grown under light or dark conditions during the non-stress (NS) or HS treatment and recovery.

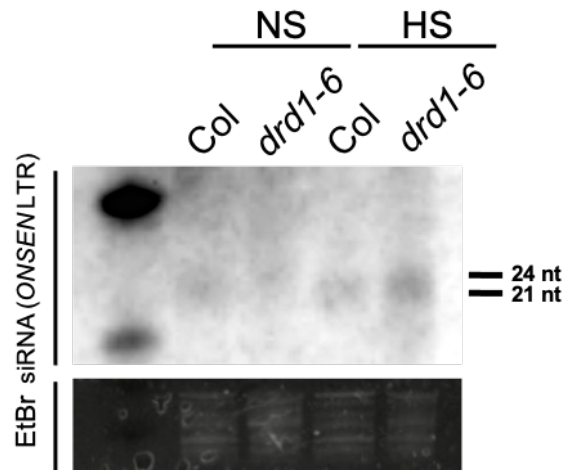
(E) siRNA accumulation in Col, *suvh2*, *suvh9*, and *nrpd1* in 7-week-old seedlings treated with non-stress (NS) or HS (24 h or 48 h) and recovered at 21°C for seven days (+7) after treatment, grown under light or dark conditions during the non-stress (NS) or HS treatment and recovery.

(F) Transcription levels of *ONSEN* in Col, *suvh2*, *suvh9*, and *nrpd1* under HS treatment in light and dark conditions. Asterisks denote significant differences between groups (Student's t-test, *, $P < 0.05$; **, $P < 0.01$). "ns" indicates no significant differences between groups (Student's t-test, $P \geq 0.05$).

5.2 siRNA Suppresses *ONSEN* Transposition During Heat Stress and Silences Transcription Under Normal Conditions

Since siRNA does not accumulate immediately after HS, the siRNA levels measured seven days after HS and recovery at 21°C were used for subsequent analyses. As expected, no siRNA production was detected in the *nrpd1* mutant (Fig 9B). Under non-

stress (NS) conditions, siRNA accumulation was detected in the *suvh9* mutant, while the *suvh2* and *suvh2/9* did not produce siRNA (Fig 9B). Interestingly, in Col, *suvh2*, *suvh9*, and *suvh2/9* lines, siRNA accumulation was observed during the 7-day recovery period following 24 and 48 hours of HS treatment (Fig 9B). This indicated that the absence of SUVH2 and/or SUVH9 does not completely block siRNA generation under heat condition. Following HS, siRNA levels in *suvh2* mutants were comparable to those in Col (Fig 9B), suggesting that the siRNA produced by HS phase plays a crucial role in suppressing *ONSEN* transposition in *suvh2*. In addition, the same pattern of siRNA accumulation was observed in the *drd1-6* mutant (Fig S8), indicating that SUVH2 and DRD1 relay on the siRNA pathway to suppress *ONSEN* transposition effectively.

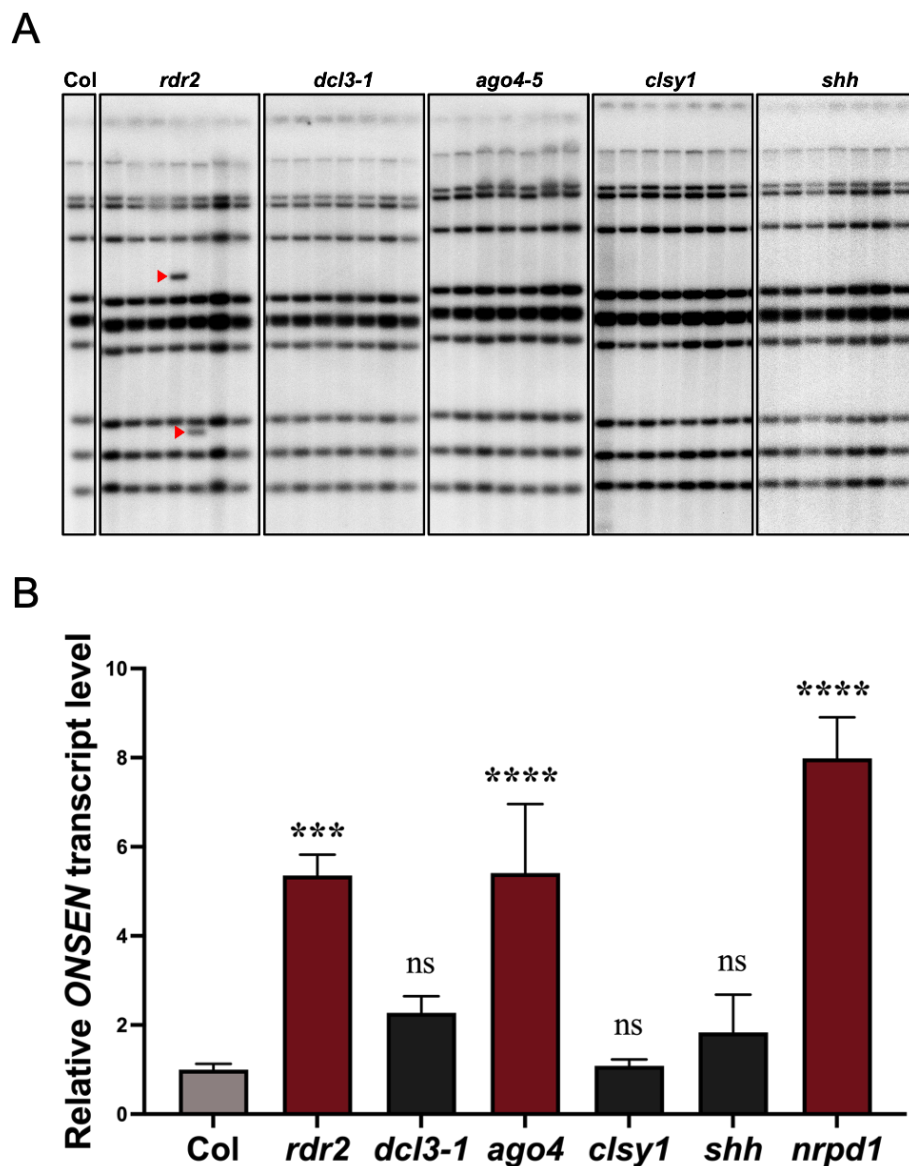


Supplementary Figure 8: Accumulation of siRNA in *drd1-6*.

Accumulation levels of siRNAs in 7-week-old Col and *drd1-6* seedlings after treatment with non-stress (NS) or heat stress (HS) and recovery for seven days (+7) under normal conditions (21°C).

Similar results were observed in the *rdr2* and *ago4-5* mutants (Fig 9C, Fig S9A, B). In *rdr2*, where siRNA accumulation is absent, *ONSEN* exhibited high transposition

activity (Fig 9C, Fig S9A, B). In contrast, in *ago4-5*, although *ONSEN* was highly expressed, transposition did not occur. This lack of transposition correlates with siRNA accumulation that was detected only under HS conditions (Fig 9C, Fig S9A, B). Matsunaga et al. (2012) proposed that functional redundancy among DCL proteins in Arabidopsis might allow other DCL proteins to compensate for the loss of DCL3, suppressing the transgenerational transposition of *ONSEN*. This hypothesis could explain why neither siRNA nor *ONSEN* transposition was observed in the *dcl3-1* mutant (Fig 9C, Fig S9A, B).



Supplementary Figure 9: Activity of *ONSEN* in RdDM Mutants.

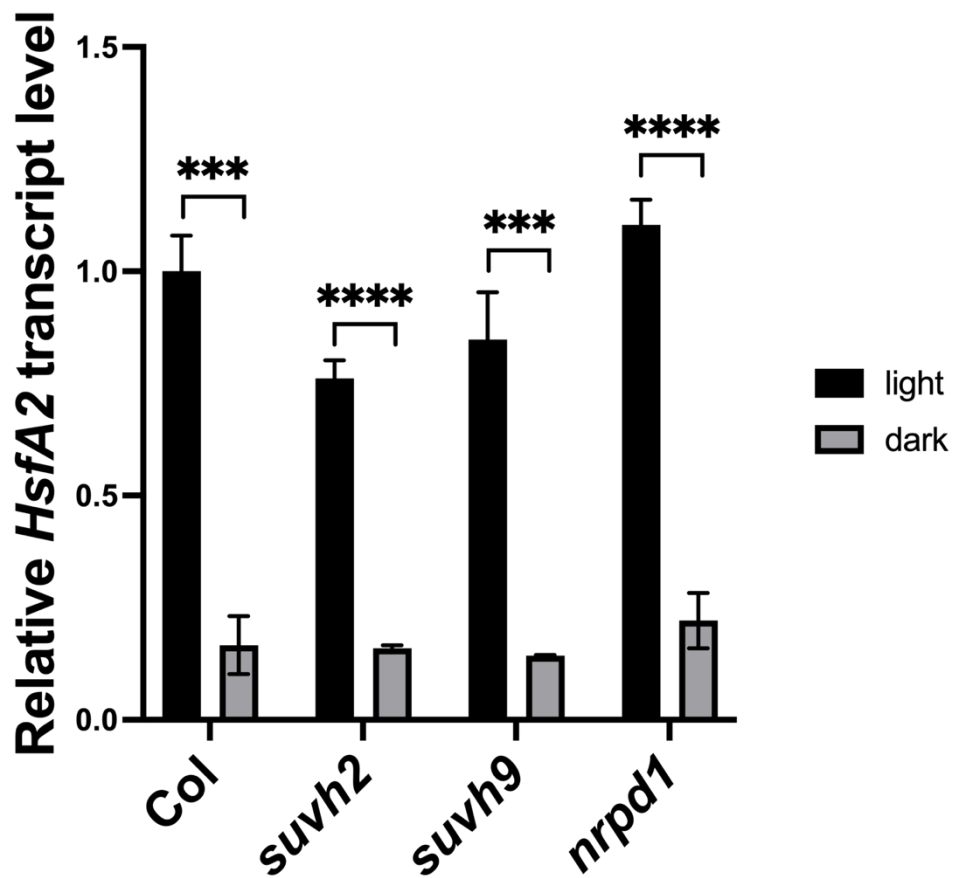
(A) Relative transcription levels of *ONSEN* in Col, *rdr2*, *dcl3-1*, *ago4-5*, *clsyl*, *shh*, and *nprdl* after 24 hours of heat stress (HS). Asterisks indicate significant differences compared to Col (one-way ANOVA and Dunnett's test, *. $P < 0.05$; **. $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$); "ns" denotes no significant differences (one-way ANOVA and Dunnett's test, $P > 0.05$).

(B) Southern blot analysis showing new *ONSEN* insertions in *rdr2*, *dcl3-1*, *ago4-5*, *clsyl*, *shh*, and *nprdl* after 24 hours of HS. The leftmost panel shows the Col pattern under non-stress (NS) conditions. Red arrows indicate new *ONSEN* insertions.

5.3 The Impact of Light on siRNA Accumulation

To further explore siRNA production, I investigated its accumulation under both light and dark conditions. In the absence of light, no siRNA was detected in *suvh2* mutants without HS; however, siRNA accumulated following HS treatment (Fig 9D). Interestingly, siRNA accumulation was lower under dark conditions compared to light conditions (Fig 9D, E). These results suggest that light influences the amount of siRNA accumulation but does not alter the function of RdDM factors in siRNA production. qRT-PCR analysis revealed that *ONSEN* transcription levels decreased under dark conditions in *suvh2*, *suvh9*, and *nprdl* mutants, whereas no significant change was observed in Col (Fig 9F). *ONSEN* expression depends on the heat response pathway, with transcription activated by the binding of heat shock factor A2 (HsfA2) to the LTR region (Cavrak V.V. et al., 2014). Further investigation into *hsfa2* expression under light and dark conditions showed that *hsfa2* transcription was significantly reduced in Col, *suvh2*, *suvh9*, and *nprdl* under dark conditions (Fig S10). These results suggest that

light influences the RdDM-dependent transcriptional silencing of *ONSEN*, potentially through the *HsfA2* pathway.



Supplementary Figure 10: Light Influences the Transcriptional Levels of *HsfA2*.

The bar graph illustrates the transcription levels of *HsfA2* in Col, *suvh2*, *suvh9*, and *nrpd1* under heat stress (HS) treatment in light and dark conditions. Asterisks denote significant differences between groups (Student's t-test, *. $P < 0.05$; **. $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$).

Discussion

The RdDM plays a crucial role in plant responses to biotic and abiotic stresses (Erdmann R.M and Picard C.L., 2020). Downstream components of the RdDM pathway,

including the DDR complex and SUVH family proteins, play significant roles in this process (Zhong X et al., 2012; Johnson L.M et al., 2014). This study focuses on the inhibitory effects of the DDR complex and SUVH2 and SUVH9 on *ONSEN* activity. Measurement of *ONSEN* copy numbers in RdDM downstream factor mutants revealed that the absence of DRD1-6, DMS3-1, and SUVH2 results in substantial accumulation of *ONSEN* cDNA (Fig 1). Since the frequency of *ONSEN* transposition is correlated with the duration of HS, 48 hours of HS led to higher *ONSEN* activity than 24 hours (Fig 2). Notably, under 48 hours of HS, *ONSEN* transposition was only observed in the *drd1-6* and *suvh2* mutants among the RdDM downstream factor mutants (Fig 3). These results indicate that DRD1-6 and SUVH2 are essential components of the RdDM pathway, acting to inhibit *ONSEN* transposition effectively.

DRD1-6 is an essential component of the DDR complex. To further explore the relationship between DRD1-6 and the DDR complex in the RdDM process, I examined various DDR complex mutants. The results showed that under HS conditions, *ONSEN* transcription levels and copy numbers significantly increase in these mutants (Fig 4A, Fig 1). However, the increase in *ONSEN* copy number did not directly correlate with increased transposition, as *ONSEN* transposition was only observed in the *drd1-6*, *drd1-6/dms3-1*, and DDR triple mutants after 48 hours of HS (Fig 4C, Fig S3). Interestingly, although the *ONSEN* transcript levels and copy numbers in the DDR complex mutants were comparable to those in *nrrpd1*, the transposition frequency of *ONSEN* in the DDR mutants was notably lower than in *nrrpd1*. Furthermore, no *ONSEN* transposition was detected in *dms3-1* and *rdm1-4*, despite their high *ONSEN* transcription levels (Fig 2, Fig 4). These findings suggest that the components of the DDR complex play distinct roles in the transcriptional and transpositional suppression of *ONSEN*, with DRD1-6 likely playing a key role in the DDR complex-mediated suppression of *ONSEN*

transposition (Fig 10).

In the RdDM pathway, DRD1 forms the DDR complex by the DR complex and Pol V (Law J.A. et al., 2010; Wierzbicki A.T. et al., 2008). Previous studies suggest that the stable formation of the DR complex by DMS3-1 and RDM1-4 is a prerequisite for assembling the DDR complex (Wongpalee S.P. et al., 2019). However, direct interactions between DMS3-1, RDM1-4, and Pol V have not been reported (Gao, Z. et al., 2010). These findings suggest that DRD1-6 may act as a bridge connecting the DR complex to Pol V (Wongpalee S.P et al., 2019). Therefore, the absence of DRD1 could lead to a partial loss of Pol V function, ultimately releasing the transpositional activity of *ONSEN*.

Some studies suggest that SUVH2 and SUVH9 have partially redundant roles in transcriptional silencing. In the *suvh2/9* double mutant, the expression level of *AtSNI*—a SINE retrotransposon widely used as a model for studying RdDM and transcriptional silencing—significantly increased, whereas no changes were observed in the single mutants of *suvh2* or *suvh9* (Liu ZW et al., 2014). However, the expression pattern of *ONSEN* contrasts with *AtSNI*, suggesting that the roles of SUVH2 and SUVH9 in *ONSEN* transcriptional repression may differ. It is possible that SUVH9 is not involved in *ONSEN* transcriptional repression or plays only a minor role. In studies of transposon regulation, *ONSEN* transposition was observed in the *suvh2* mutant after 48 hours of HS but not after 24 hours (Fig 3, Fig S4). However, despite *ONSEN* transposition occurring in *suvh2*, the transposition frequency remained much lower than that observed in *nrpd1*, and the *ONSEN* copy number in *suvh2* did not increase significantly compare to *nrpd1*(Fig 1). Given that SUVH2 and SUVH9 have partially non-redundant functions in the RdDM pathway (Kuhlmann M and Mette M.F., 2012), the transcription levels of *ONSEN* do not directly correlate with its transposition frequency. Although

ONSEN transposition occurred in the *suvh2/9* double mutant, the transposition frequency was not significantly higher than in the *suvh2* single mutant (Fig 5C). These results indicate that suppression of *ONSEN* transposition is primarily mediated by SUVH2, while SUVH9 appears to play little to no role in this process.

Extensive studies have shown that the loss of the DDR complex leads to reduced DNA methylation levels (Kanno T et al., 2004; Kanno T et al., 2008; Sasaki T et al., 2014). BS-seq results from this study confirm that DRD1-6 and the DDR complex are crucial for maintaining *ONSEN* DNA methylation (Fig 7A). Under heat shock conditions, transcription levels of *ONSEN* increase in mutants with impaired non-CG methylation such as *cmt2*, *cmt2/3*, and *drm1/2* (Nozawa K et al., 2021). These findings align with the results of this study, which demonstrate increased *ONSEN* transcription levels in the *drd1-6* and DDR complex mutants with defective non-CG methylation (Fig 7A, Fig 4A). Similar trends have been observed in studies of SUVH proteins. SUVH2 and SUVH9 lack the SET domain, which is necessary for histone methyltransferase activity, they bind to methylated DNA through their SET and RING-associated (SAR) domains and participate in the RdDM pathway (Johnson L.M., 2008). A study on SDC indicated that loss of *suvh2* or *suvh9* results in the reduced CHH methylation (Johnson L.M., 2008). Consistent with these findings, this study found that DNA methylation levels of the *ONSEN* LTR region were indeed reduced in SUVH2-deficient plants, with a significant decrease in CHH methylation (Fig 7B). Interestingly, in the *suvh9* mutant, the *ONSEN* region exhibited higher methylation levels, which may account for the continued transcriptional silencing of *ONSEN* in *suvh9* (Fig 7B, Fig 4A). In addition, the methylation patterns of *ONSEN* in *drd1-6* or *suvh2* closely resembled those in *nrpd1* (Fig 7A, 7B). However, despite these similarities, the transposition frequency of *ONSEN* in *drd1-6* and *suvh2* mutants was much lower than in *nrpd1* (Fig 2D, Fig 3). In

conclusion, while *ONSEN* transcriptional activity is influenced by DNA methylation levels, its transposition is largely regulated by factors other than DNA methylation.

Heterochromatin is enriched with transposable elements (Marsano R.M. and Dimitri P., 2022). Although chromatin accessibility is associated with the formation of DNA methylation (Zhong Z et al., 2021), studies suggest that the role of chromatin remodeling factors may be independent of the DNA methylation. For example, Morpheus Molecule 1 (MOM1) silences repetitive heterochromatic sequences through a mechanism that does not rely on DNA methylation (Amedeo P et al., 2000). Similarly, MORC1 and MORC6 contribute to the RdDM pathway by regulating heterochromatin de-condensation rather than directly through DNA methylation (Moissiard G et al., 2012). Jing Y et al. (2016) identified that SUVH9 as a critical link in MORC-mediated chromatin remodeling (Jing Y et al., 2016). However, the absence of SUVH2 and/or SUVH9 does not result in chromatin opening in the *ONSEN* LTR region (Fig 8B), suggesting that SUVH9 plays a limited role in *ONSEN* silencing. DAPI staining results further show that HS induces heterochromatin de-condensation, but this effect is not unique to the *nrpd1* mutant (Fig 8D, E). While open chromatin may be necessary for transposon activity, identifying open chromatin in the *ONSEN* promoter region under HS remains challenging due to the complexity introduced by *ONSEN*'s extrachromosomal cDNA. In addition, the chromatin state at *ONSEN* insertion sites may influence insertion events. However, *ONSEN* in the *nrpd1* mutant displays random insertion patterns (Ito H et al., 2016), making it difficult to investigate open chromatin at random insertion sites under HS condition.

siRNA is closely associated with gene silencing (Hamilton A et al., 2002). In the RdDM pathway, Pol IV transcribes TEs and repeats sequence genes into primary transcripts. These transcripts are subsequently processed by RDR2, DCL3, and AGO4 to produce

siRNA (Ferrafiat L et al., 2019). In *nrpd1* mutants under HS, *ONSEN* frequently undergoes transposition (Ito H et al., 2011; Ito H et al., 2016). Northern blot analysis shows that *ONSEN* siRNA accumulates in the *suvh2* and *suvh2/9* mutants only under HS conditions (Fig 9B). Conversely, the absence of SUVH9 has no effect on siRNA accumulation under either non-stress (NS) or HS conditions (Fig 9B). This indicates that SUVH2 plays a role in regulating the synthesis of *ONSEN* siRNA. Further research reveals that in RdDM mutants with high *ONSEN* transcription levels, siRNA accumulation occurs only under HS. Therefore, siRNA produced under HS suppresses *ONSEN* transposition. Wierzbicki et al. (2008) proposed that transposon silencing depends on the synergistic action of Pol V transcription and siRNA production. DRD1 influences the localization of NRPD1b (NRPE1), indirectly affecting Pol IV function and impairing siRNA production (Pontes O et al., 2006). In the *drd1-6* mutant, *ONSEN* siRNA is also produced exclusively under HS condition (Fig S8). This suggests that partial dysfunction of Pol V and Pol IV disrupts siRNA synthesis, impairing the complete RdDM process. In addition, this study found that while dark conditions reduce siRNA levels (Fig 9D, E). It has been proposed that *ONSEN* is activated by high light and oxidative stress activation pathway dependent on HsfA2 transcription factor (Matsunaga W et al., 2015). Dark conditions impact the transcription of *HsfA2*, which in turn affects *ONSEN* transcription levels (Fig S10, 9F). Since transcription of *ONSEN* is strongly activated by high light stress and HS in the siRNA-deficient plant *nrpd1*, it implies that the transcriptional activation of *ONSEN* may be regulated by siRNA-mediated mechanisms (Matsunaga W et al., 2015, Fig 9F).

In summary, the results of this study indicate that the DDR complex and SUVH2 play crucial roles in silencing of *ONSEN*, highlighting the importance of RdDM factors in regulating *ONSEN* activity. Given the differences in complex components involved in

the DDR-related regulation of *ONSEN* transcription and transposition, I propose that these regulatory mechanisms should be discussed separately when addressing the transcriptional and transpositional control of retrotransposons. RdDM-mediated DNA methylation is associated with the suppression of *ONSEN* transcription, but it does not directly affect the silencing of *ONSEN* transposition. The evidence suggests that SUVH2 and DRD1 contribute to the transcriptional silencing of *ONSEN* through a siRNA-dependent pathway (Fig 10). This study proposes that plants, when exposed to adverse environmental stresses, adapt their protective mechanisms to mitigate potential genetic damage caused by these stresses.

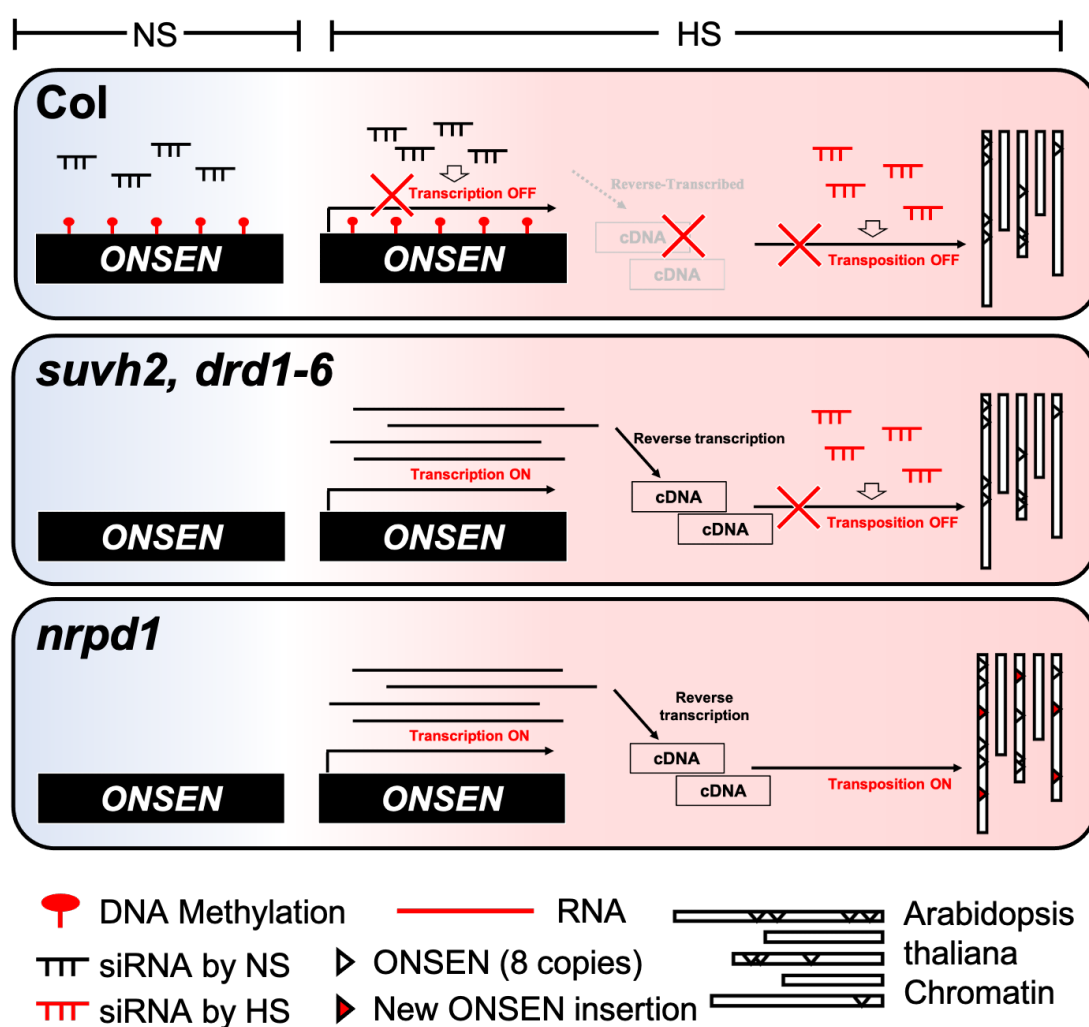


Figure 10: Model of SUVH2 and DRD1-6 in *ONSEN* Regulation.

In Col, siRNA is produced through the RdDM pathway, which inhibits *ONSEN* transcription. In the *drd1-6* or *suvh2* mutant, the RdDM pathway is disrupted, allowing heat stress (HS) to activate *ONSEN* transcription. Under HS, plants bypass the DRD1-6 or SUVH2-dependent pathway to produce new siRNAs. The siRNAs generated during HS prevent *ONSEN* from transposing into other chromosomal sites. In *nRPD1*, siRNA synthesis is impaired under HS conditions, leading to a significant reduction in the silencing of *ONSEN* transposition.

Chapter 2: Analysis of a Novel Mechanism of Retrotransposon

Repression by Histone Deacetylases

Introduction

Histone acetyltransferases (HATs) and histone deacetylases (HDACs) collaboratively regulate the dynamic equilibrium of histone acetylation within the cell. HDACs remove acetyl groups from histones, which tightens chromatin structure and makes it more difficult for transcription factors to bind to DNA, thereby silencing gene expression (Du Q et al., 2022). In *Arabidopsis*, 18 types of HDACs have been identified. Based on their catalytic domains, HDACs can be classified into three classes: class I (RPD3-like family), class II (SIR2-like family), and class IV (HDA11). The RPD3 superfamily of HDACs consists of 12 types, which are further divided into three classes: the first class includes HDA6, HDA7, HDA9, and HDA19; the second class includes HDA5, HDA15, and HDA18; and the third class includes HDA2, as well as the unclassified HDA8, HDA14, HDA10, and HDA17 (Ritu P et al., 2002; Núñez-Álvarez Y and Suelves M. 2022; Guo Q et al., 2023). Extensive research has shown that members of the RPD3 family of HDACs are involved in various biological processes, including flowering time regulation, abscisic acid (ABA) signaling, and light signaling (Mehdi S et al., 2016; Wang Z et al., 2014; Liu X et al., 2014; Nguyen HN et al., 2013). Notably, HDA19 plays a key role in modulating plant developmental and stress response pathways (Zheng Y et al., 2023). The absence of HDA19 frequently affects seedlings, flowers, and seed development (Gorham SR et al., 2018). Additionally, *hda19* mutants exhibit enhanced tolerance to salt, drought, and heat stress (Ueda M et al., 2017; Ueda M et al., 2018).

Trichostatin A (TSA) is a commonly used HDAC inhibitor for studying the epigenetic

regulatory mechanisms of cells. Typically, inhibiting HDAC activity is expected to reverse transcriptional silencing and promote gene expression by increasing histone acetylation (Wójcikowska B et al., 2018; Tanaka M et al., 2008). However, the results of this study showed that TSA treatment led to a reduction in *ONSEN* transcription in the plants, contradicting the conventional view that histone acetylation activates gene transcription. This study aims to further explore the relationship between histone acetylation, methylation, and the transcriptional regulation of *ONSEN* by utilizing HDAC inhibitors and HDAC mutants.

Material

Arabidopsis thaliana accession Columbia (Col) was used as the background for all the mutants. The *hda6* (*axe1-5*, *C66153-1*), *hda9* (*GABI_305G0*), *hda5* (*SALK_007503*), *hda14* (*SALK_097005*), *hda15* (*SALK_004027*), and *hda18* (*SALK_006938*) mutants were provided by Professor Atsushi Kato at Hokkaido University. The *hda7-1* and *hda7-2* mutants were obtained from Professor Hidetoshi Saze at the Okinawa Institute of Science and Technology (OIST). The *hda19-3* (*SALK_139445*) mutant was acquired from the ABRC stock center. The *nrpd1* (*C366150*) mutants were sourced from Ohio State University, Columbus, OH, USA. The *ddm1-1* (Vongs A et al., 1993) mutant was provided by Professor Tetsuji Kakutani at the University of Tokyo. The *hda19-3* mutant was crossed with *nrpd1* and *ddm1* mutants to generate the *hda19/nrpd1* and *hda19/ddm1* double mutants.

Results

1. TSA Inhibits *ONSEN* Transcriptional Activity

1.1 TSA Reduces *ONSEN* Transcription Levels Under Heat Stress

As an HDAC inhibitor, TSA is known to suppress HDAC activity, leading to increased histone acetylation levels, in theory, enhanced gene transcription (Xu CR et al., 2005). However, contrary to expectations, *ONSEN* transcription was suppressed in TSA-treated Col plants (Fig 11A). Furthermore, RNA-seq analysis revealed that *ONSEN* transcription levels were lower in the TSA-treated group compared to the control stress group (CS, indicating no TSA treatment) under heat stress (HS) (Fig 11B).

1.2 Effects of TSA on Gene and Transposon Transcriptional Activity

Under non-stress (NS) conditions, TSA treatment led to the downregulation of 14.9% of genes, while only 1.7% of TEs were downregulated (Fig 11C). However, after combined HS and TSA treatment, the number of downregulated TEs increased from 9 to 58, including *ONSEN* (Fig 11C). Wójcikowska et al. also showed in their study that downregulated transcripts were more frequently observed than upregulated transcripts in TSA-treated *A.thaliana* (Wójcikowska et al., 2024). These results suggest that TSA does not generally increase the transcriptional activity of all genes and TEs. Instead, under abiotic stress conditions, TSA appears to silence the transcription of certain TEs, including *ONSEN*.

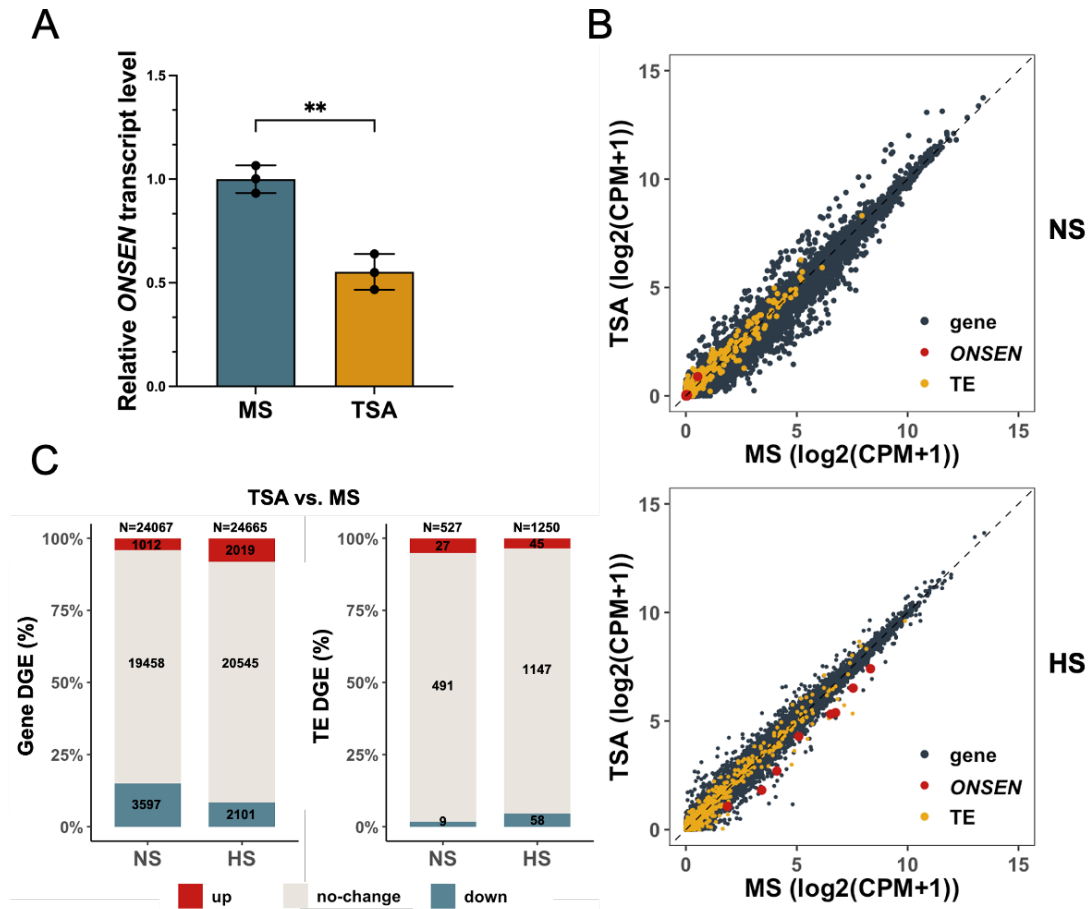


Figure 11: *ONSEN* Transcription Levels Decrease After Treatment with the HDAC Inhibitor TSA Under Heat Stress (HS).

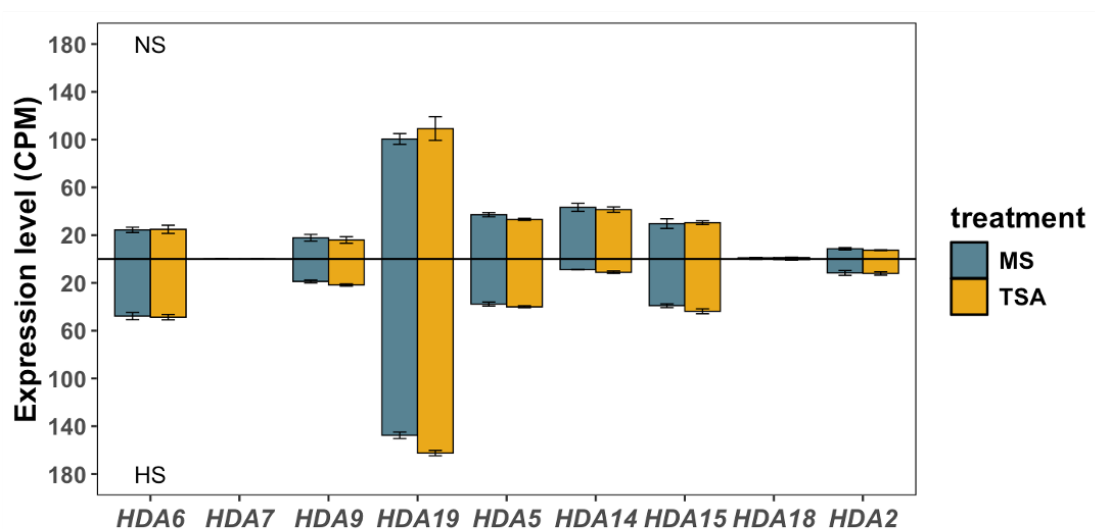
(A) Transcription levels of *ONSEN* in Col grown on MS and TSA media under HS conditions. Asterisks indicate significant differences between the two groups (Student's t-test, **, $P < 0.01$).

(B) Scatterplots showing the count per million reads (CPM) of genes and transposable elements (TEs) in MS medium versus TSA medium under non-stress (top) and heat stress (bottom) conditions.

(C) Differential gene expression (DGE) analysis of genes (left) and transposable elements (TEs, right) in TSA medium versus MS medium under non-stress (NS) and heat stress (HS) conditions. Up (LogFC (Fold Change) > 0 , $P < 0.05$); No-change (LogFC = 0, $P < 0.05$); Down (LogFC < 0 , $P < 0.05$).

1.3 TSA Does Not Affect Transcription Levels of HDAC Genes

RNA-seq results revealed the expression levels of HDAC genes following TSA treatment. Among the HDAC-related genes, *HDA19* exhibited high expression levels, while *HDA7* and *HDA18* showed minimal expression (Fig S11). Under HS, transcription of *HDA6* and *HDA19* was further activated compared to non-stress (non-stress (NS)) conditions (Fig S11). However, no significant differences were observed in the transcription levels of HDAC genes between control stress (CS) and TSA-treated samples under either NS or HS conditions. This indicates that TSA does not significantly alter the transcription of HDAC genes but prevents HDAC from functioning by chelating zinc ions in the active site of histone deacetylase (Fig. S1; Vladimir Beljanski et al., 2009).



Supplementary Figure 11: Effect of TSA on HDAC Expression Levels.

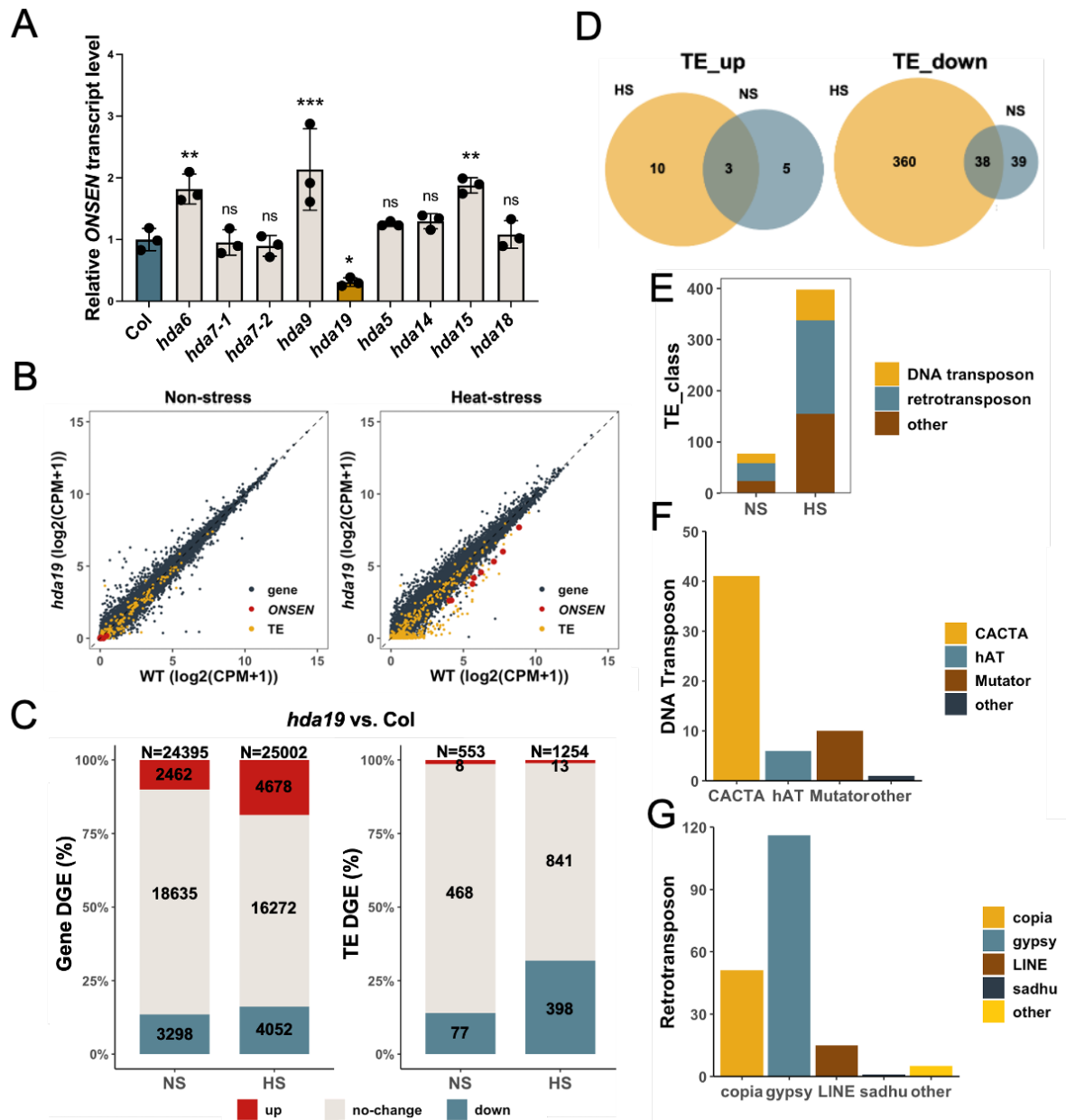
Symmetrical bar graphs illustrate the count per million reads (CPM) from RNA-seq analysis of various HDAC protein genes (*hda6*, *hda7*, *hda9*, *hda19*, *hda5*, *hda14*, *hda15*, *hda18*, and *hda2*) under non-stress (NS) and heat stress (HS) conditions in MS and TSA media.

2. Loss of *HDA19* Leads to Decreased *ONSEN* Expression

2.1 *ONSEN* Expression is Suppressed in *hda19*

To determine which specific HDAC regulates *ONSEN* expression, qRT-PCR was performed to analyze *ONSEN* transcription levels in *hda6*, *hda7-1*, *hda7-2*, *hda9*, *hda19*, *hda5*, *hda14*, *hda15*, and *hda18* under HS conditions. The results showed that *ONSEN* transcription levels significantly increased in *hda6* and *hda15*, while lower transcription levels were observed in *hda19* compared to Col (Fig 12A). Similarly, *ROMANIAT5-2*, a transposon associated with HS response, exhibited significantly reduced expression in *hda19* (Fig S12). Given the distinct role of HDA19 compared to other HDACs, subsequent studies focused on HDA19. RNA-seq results confirmed that all eight copies of *ONSEN* were downregulated in *hda19* under HS conditions (Fig 12B).

Figure 12: Most TEs Were Down-Regulated in *hda19* Under Heat Stress (HS).



(A) Relative transcription levels of *ONSEN* in Col, *hda6*, *hda7-1*, *hda7-2*, *hda9*, *hda19*, *hda5*, *hda14*, *hda15*, and *hda18* after 24 hours of HS. Asterisks indicate significant differences between the two groups (one-way ANOVA and Dunnett's test, *. $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$). "ns" indicates no significant differences between the two groups (Student's t-test, $P > 0.05$).

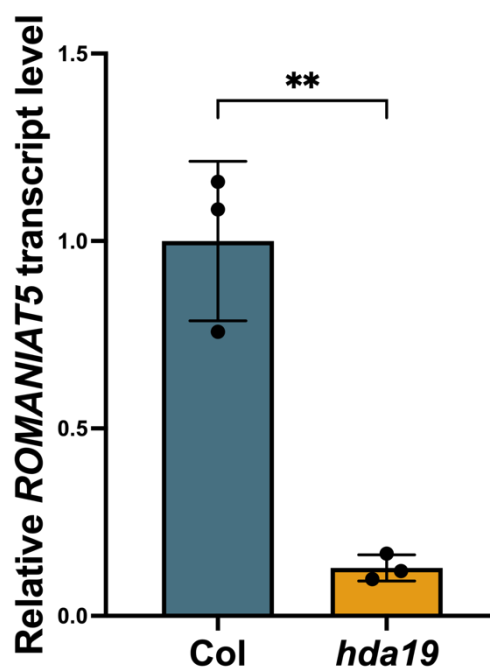
(B) Count per million reads (CPM) of genes and TEs in Col versus *hda19* under non-stress (NS) (left) and heat stress (HS) (right) conditions.

(C) Differential gene expression (DGE) of genes (left) and TEs (right) in *hda19* versus Col under non-stress (NS) and HS conditions. Up (LogFC (Fold Change) > 0, $P < 0.05$); No-change (LogFC = 0, $P < 0.05$); Down (LogFC < 0, $P < 0.05$).

(D) Venn diagrams showing the overlap of up-regulated TEs (left) and down-regulated TEs (right) in *hda19* under non-stress (NS) and HS conditions.

(E) Classification of down-regulated TE families in *hda19* under non-stress (NS) and HS conditions.

(F) Classification of down-regulated TE super-families in *hda19* under HS, including DNA transposons (top) and retrotransposons (bottom).



Supplementary Figure 12: Transcription of *ROMANIAT5* is Downregulated in *hda19*.

Relative transcription levels of *ROMANIAT5* in Col and *hda19* under 24 h of HS. Asterisks indicate significant differences between the two groups (Student's t-test, **).

$P < 0.01$).

2.2 Broad Suppression of Multiple Transposons in *hda19* Under Heat Stress

RNA-seq analysis revealed that under HS conditions, 398 transposable elements (TEs) exhibited decreased transcription levels in *hda19* (Fig 12C). In contrast, only a small number of TEs showed increased transcription levels under both non-stress (NS) and HS conditions in *hda19* (Fig 12C). An intersection analysis of upregulated (Fig. 12D, left) and downregulated (Fig. 12D, right) TEs under non-stress (NS) and HS conditions revealed that only 38 TEs were commonly downregulated under both conditions, representing just 9% of the HS-downregulated TEs ($n = 398$) (Fig. 12D). This suggests that HS induces the downregulation of additional TEs in *hda19*. Further analysis showed considerable number of the TEs downregulated by HS in *hda19* belonged to the retrotransposon type (Fig 12E), with the gypsy family of retrotransposons being the most prevalent (Fig 12G). Among DNA-type TEs, the *CACTA* family dominated the group downregulated by HS in *hda19* (Fig 12F).

3. *ONSEN* Downregulation in *hda19* Is Independent of the HSFA2 Pathway

To understand how TEs are downregulated in *hda19* under HS, I first investigated the transcriptional activation pathway of *ONSEN*. The heat-responsive transcription factor heat shock factor A2 (HsfA2) binds to the long terminal repeat (LTR) region of *ONSEN* to activate its expression (Cavrak V.V. et al., 2014). qRT-PCR was used to measure the transcription levels of *hsfa2* in Col and *hda19* under HS conditions. The results showed that *HsfA2* transcription levels in *hda19* were not significantly lower than those in Col. This suggests that *HsfA2* levels are maintained even in the absence of HDA19, and the repression of *ONSEN* could be attributed to other mechanisms, such as histone

deacetylation or chromatin remodeling directly affecting the accessibility or activity of the *ONSEN* LTR region, rather than a decrease in *HsfA2* expression itself (Fig 13).

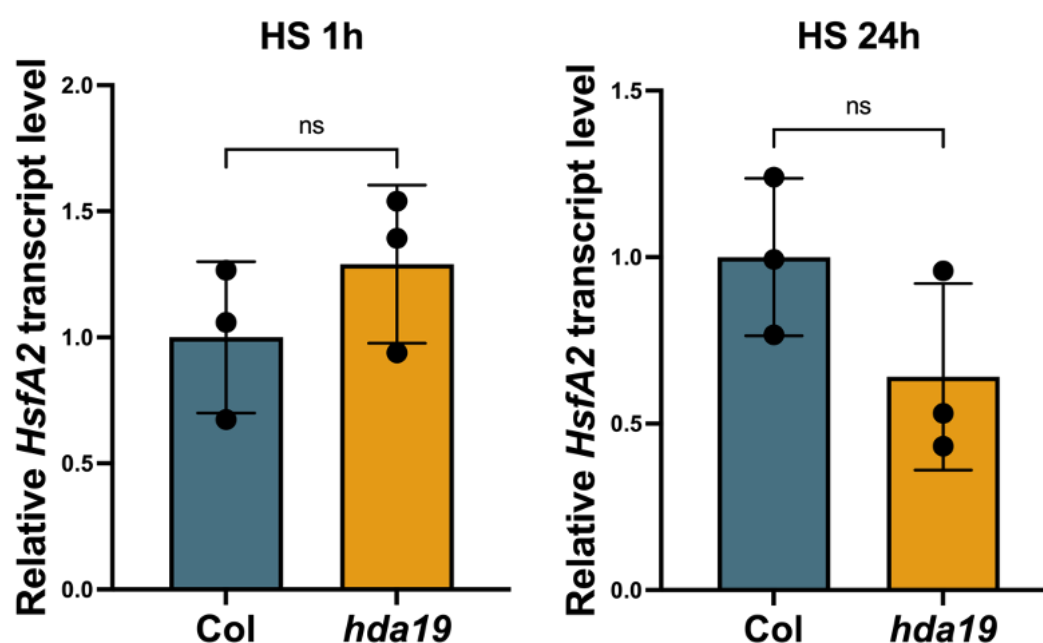


Figure 13: Reduction in *ONSEN* Transcript Levels in *hda19* Is Independent of the Heat Shock Pathway.

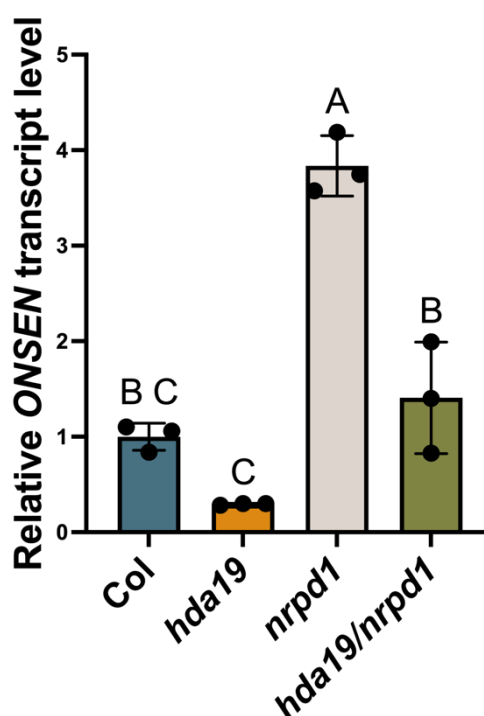
Relative transcription levels of *HsfA2* in Col and *hda19* under 1 hour (left) or 24 hours (right) of HS. "ns" indicates no significant differences between the two groups (Student's t-test, $P > 0.05$).

4. *ONSEN* Transcription Levels and DNA Methylation in *hda19*

4.1 Decreased *ONSEN* Transcription in *hda19* is Associated with CHG and CHH Hypermethylation

Previous studies have shown that *ONSEN* exhibits high transcriptional activity and hypo-methylation in *nrepd1* (Takehira K et al., 2021). To explore the regulation of

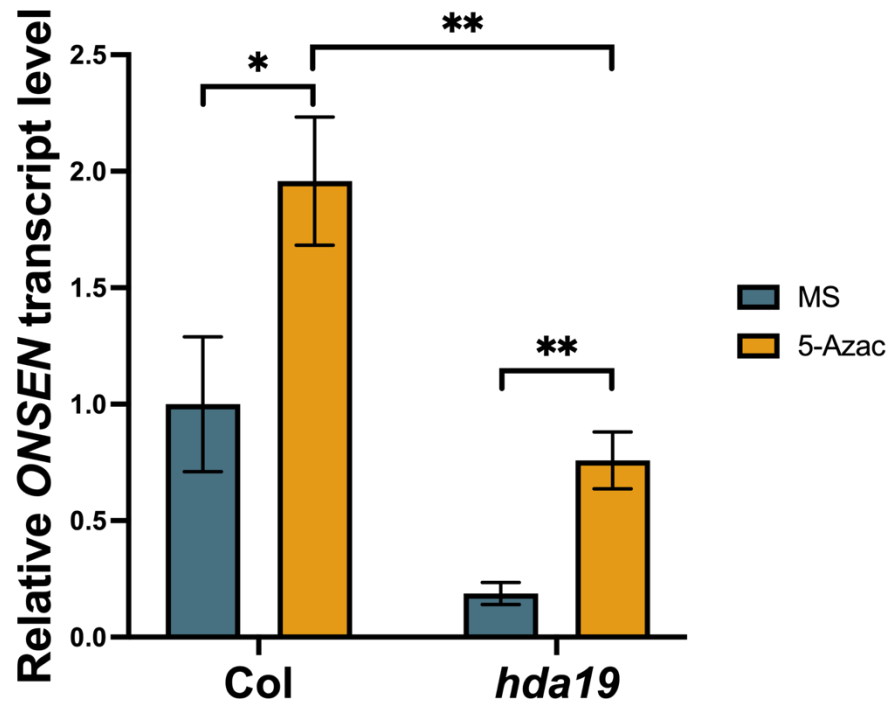
ONSEN transcription by HDA19 through DNA methylation, we measured *ONSEN* transcription levels in the *hda19/nrpd1* double mutant. The results showed that *ONSEN* transcription levels in the *hda19/nrpd1* double mutant were significantly lower than in the *nrpd1* mutant, but significantly higher than in the *hda19* mutant (S13). Additionally, 5-Aza-2'-deoxycytidine (5-Azac), a DNA methylation inhibitor, was used to assess the relationship between *ONSEN* transcription and DNA methylation in *hda19*. Following 5-Azac treatment, *ONSEN* transcription levels significantly increased, suggesting that transcriptional silencing of *ONSEN* in *hda19* may be linked to DNA methylation (Fig S14).



Supplementary Figure 13: Downregulation of *ONSEN* Transcription in *hda19/nrpd1* under Heat Stress

Relative transcription levels of *ONSEN* in Col, *hda19*, *nrpd1*, and *hda19/nrpd1* after 24 hours of HS using RT-qPCR. Significant differences between groups are marked with

asterisks Different letters indicate significant differences between samples (one-way ANOVA and Tukey test, $P < 0.05$).



Supplementary Figure 14: Impact of 5-AzaC Treatment on *ONSEN* Transcription in *hda19* under Heat Stress

Transcription levels of *ONSEN* in *Arabidopsis thaliana* (Col) and *hda19* grown on MS and 5-AzaC media under heat stress (HS). Significant differences between groups are marked with asterisks (Student's t-test, *. $P < 0.05$; **. $P < 0.01$).

Whole-genome bisulfite sequencing (WGBS) was conducted on *hda19*. The results revealed that under non-stress (NS) conditions, average CG, CHG, and CHH methylation levels across the eight *ONSEN* copies in *hda19* were generally higher than in Col (Fig 14A). Notably, CHG and CHH methylation levels increased significantly in the highest transcriptional activity copies, *At1g11265*, *At3g61330*, and *At5g61305* (Fig

14A). IGV screenshots showed that in *hda19*, CHG and CHH methylation levels were consistently higher in the *At3g61330* and *At5g61305* copies compared to Col (Fig 14B). These findings suggest that in the absence of HDA19, increased CHG and CHH methylation in the *ONSEN* region may suppress *ONSEN* transcription.

Analysis of differentially methylated regions (DMRs) revealed that *hda19* exhibited 25,125 hypermethylated regions across CG, CHG, and CHH contexts, accounting for 86.7% of the total DMRs ($n = 28,975$). In contrast, hypomethylated regions made up only 13.3% (Fig 14C). It is worth noting that CHG hypo-DMRs were extremely rare in *hda19*, representing only 0.05% of total CHG-DMRs, while CHG hyper-DMRs accounted for 99.5% of total CHG-DMRs (Fig 14C). CHH hyper-DMRs were less frequent compared to CG and CHG, comprising 63.4% of the total CHH-DMRs (Fig 14C). Overall, HDA19 deficiency leads to extensive hypermethylation, particularly in CG and CHG regions. In the *ONSEN* region, statistical analysis of DMRs showed the presence of only CHG and CHH hyper-DMRs (Fig 14D). CG hyper-DMRs and hypo-DMRs were found in four of eight *ONSEN* copies, representing 50% (Fig 14D). These results indicate a significant increase in CHG and CHH methylation in the *ONSEN* region of *hda19*, particularly in the highly active copies *At1g11265*, *At3g61330*, and *At5g61305* (Fig 14D). Thus, the reduced *ONSEN* transcription in *hda19* is likely associated with increased CHG and CHH methylation. Additionally, since *ONSEN* produces substantial cDNA under HS conditions (Hayashi Y et al., 2020), which contaminates genomic DNA and cannot be separated, WGBS analysis under HS conditions may be skewed by cDNA methylation. IGV screenshots indicate that CG, CHG, and CHH methylation levels appeared reduced under HS conditions in both Col and *hda19* (S15 Fig). However, due to the effect of cDNA methylation, the WGBS results under HS are not discussed further in this study.

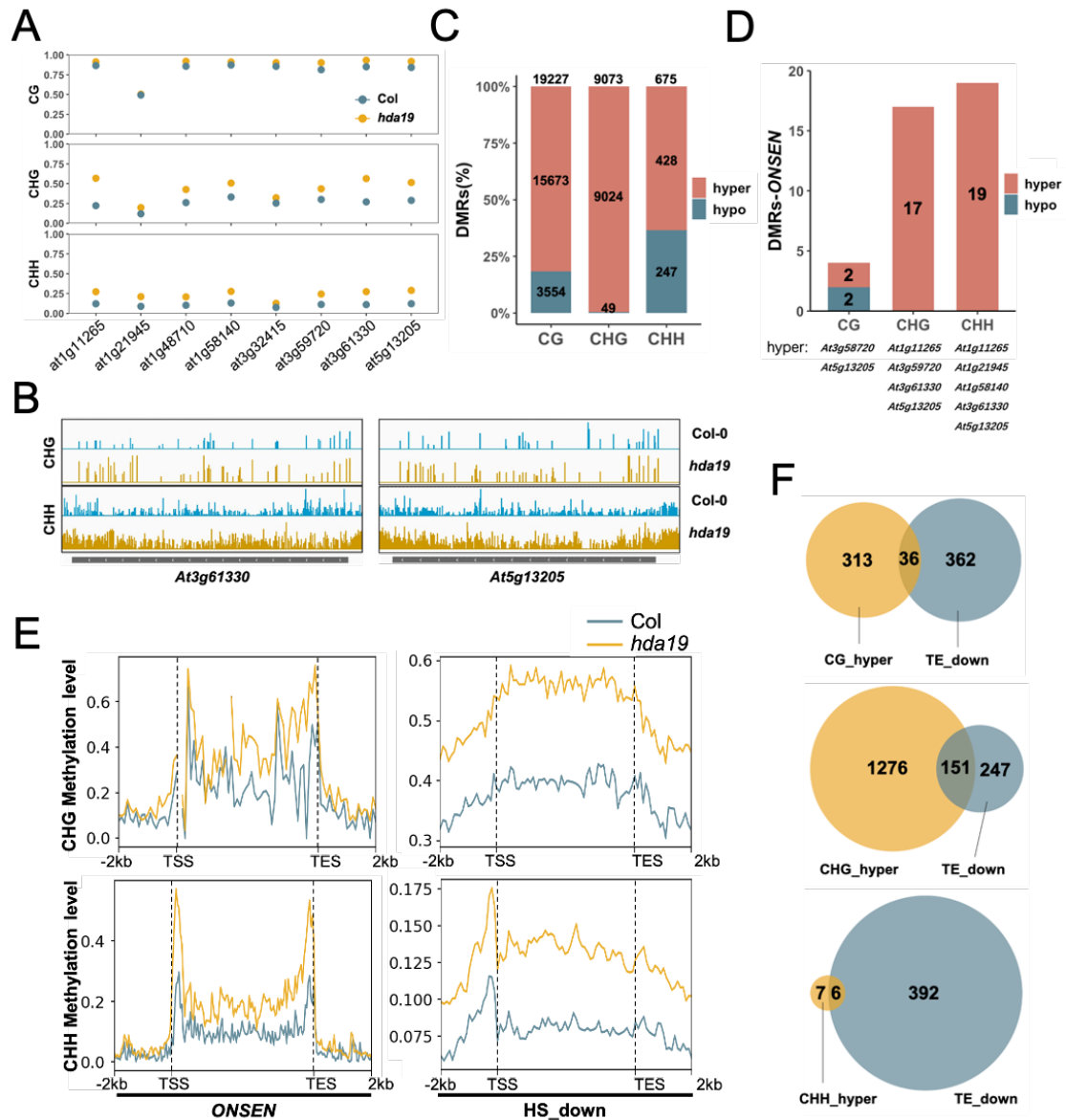


Figure 14: HDA19 Deficiency Results in Elevated CHG and CHH Methylation Levels.

(A) Average CG, CHG, and CHH methylation levels across the eight *ONSEN* copies (*At1g11265*, *At1g21945*, *At1g48710*, *At1g58140*, *At3g32415*, *At3g59720*, *At3g61330*, and *At5g13205*).

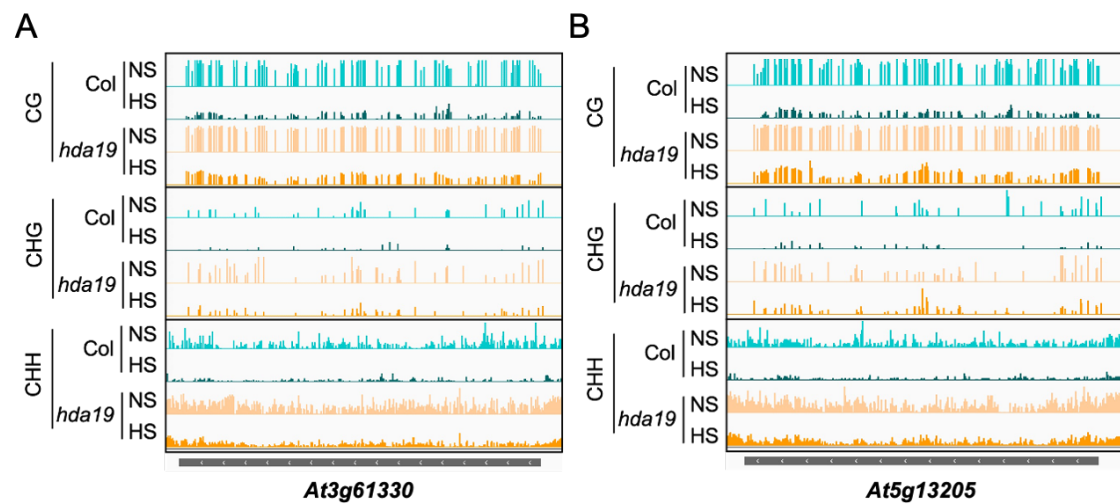
(B) Integrative Genomics Viewer (IGV) screenshots showing CHG and CHH methylation within *ONSEN* (*At3g61330* and *At5g13205*) in Col and *hda19*.

(C) Distribution of hypermethylated (hyper-DMRs) and hypomethylated regions (hypo-DMRs) across CG, CHG, and CHH contexts in *hda19*.

(D) CG, CHG, and CHH DMRs in the *ONSEN* regions of *hda19*.

(E) Average CHG and CHH methylation of the eight *ONSEN* copies (left) and down-regulated transposable elements (TEs) (right) in Col and *hda19*. "-/+2kb" indicates the upstream 2000 bp of the start site and the downstream 2000 bp of the end site.

(F) Number of CG (top), CHG (middle), and CHH (bottom) hyper-DMRs that overlap between down-regulated TEs (TE_down).



Supplementary Figure 15: WGBS Results of *hda19* Under HS.

Integrative Genomics Viewer (IGV) screenshots displaying the CG, CHG, and CHH methylation levels within *ONSEN* (*At3g61330* and *At5g13205*) in Col and *hda19* under non-stress (NS) and heat stress (HS) conditions.

High levels of CHG and CHH methylation were also observed in the TE region (n=398) that were transcriptionally downregulated in *hda19* under HS conditions (Fig 14E). To further investigate CHG and CHH methylation in these downregulated TEs (n=398)

under HS conditions in *hda19*, I compared the list of DMRs with the list of downregulated TEs. Hyper-DMRs were found in only a subset of the downregulated TEs under HS conditions (Fig 14F). In contrast, the detection rate of CHG hyper-DMRs was relatively high among the downregulated TEs but only accounted for 37.9% of the total (Fig 14F). This suggests that while CHG and CHH methylation may contribute to the transcriptional silencing of some TEs in *hda19*, not all downregulated TEs exhibit high levels of CHG and CHH methylation.

4.2 The Relationship Between DDM1 and HDA19

The intricate relationship between RdDM and DDM1 creates a feedback loop. DDM1-dependent silencing and Pol IV-mediated siRNA silencing overlap, collectively suppressing aberrant 5S rDNA transcription and contribute to the formation of 5S rDNA heterochromatin (Blevins T et al., 2009). The requirement for DDM1 is positively correlated with heterochromatin formation, which in turn, inhibits the RdDM process (Zemach A et al., 2013). On the other hand, RdDM and DDM1 together mediate nearly all transposon methylation, and the loss of DDM1 leads to strong activation of certain TE transcription (Zemach A et al., 2013; Lippman Z et al., 2004). To investigate whether the establishment of high DNA methylation in HDA19 is dependent on the DDM1 pathway, I generated *hda19/ddm1* double mutants. qRT-PCR results showed that *ONSEN* transcription levels were significantly higher in *ddm1* than in *hda19* (Fig 15A). Although *ONSEN* transcription levels were reduced in *hda19/ddm1* compared to *ddm1*, they remained significantly higher than in *hda19* (Fig 15A). The loss of DDM1 function in *hda19* restores some *ONSEN* transcriptional activity (Fig 15A). These results indicate that the absence of DDM1 can release the transcriptional silencing of *ONSEN*, even in the *hda19* background.

Next, BS-seq was performed to verify the DNA methylation levels in the *ONSEN* region. The loss of DDM1 results in an increase in CHH methylation in the *ONSEN* LTR region (Fig. 15B). CHH methylation level at *ONSEN* in *ddm1* were even higher than that in *hda19* (Fig. 15B). CHH methylation in the *ONSEN* region showed similar levels in the *ddm1* and *hda19/ddm1* double mutants (Fig 15B). However, despite the high CHH methylation in *ddm1*, *ONSEN* transcription is not silenced, suggesting that the increase in CHH methylation in the *ddm1* background is not sufficient to silence *ONSEN* (Fig. 15A, B). The transcriptional regulation of *ONSEN* by DDM1 may involve pathways other than DNA methylation. These results suggest that the establishment of high DNA methylation in *hda19* is independent of the DDM1 pathway.

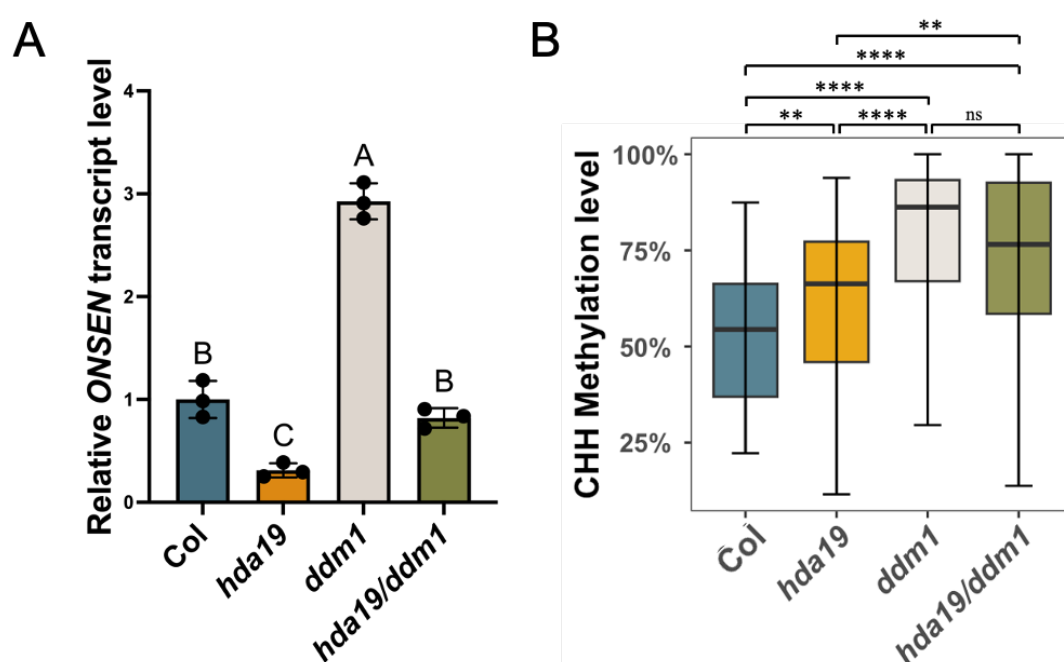


Figure 15: Reducing DNA Methylation Recovers *ONSEN* Transcription Levels in the *hda19* Background.

(A) Relative transcription levels of *ONSEN* in Col, *hda19*, *ddm1*, and *hda19/ddm1* after 24 hours of HS. Different letters indicate significant differences between samples (one-

way ANOVA and Tukey test, $P < 0.05$).

(B) Methylation levels of CHH (H = A, T, C) in the *ONSEN* LTR locus (*Atlg11265*) in Col, *hda19*, *ddm1*, and *hda19/ddm1* mutants under non-stress conditions. Significant differences between groups are marked with asterisks (Wilcoxon test and Benjamini-Hochberg method *. $P < 0.05$; **. $P < 0.01$; ***. $P < 0.001$; ****. $P < 0.0001$), while "ns" denotes no significant differences (Wilcoxon test and Benjamini-Hochberg method $P > 0.05$).

5. Analysis of Histone H3 Acetylation and Methylation in *hda19*

5.1 Loss of Histone H3 Acetylation in *hda19*

Since HDA19 plays a distinct role in regulating TE transcription compared to other HDAC proteins (see Fig 12A), I investigated changes in histone modifications in the *hda19*. ChIP-seq was used to analyze H3K9K27Ac and H3K9K14Ac levels in Col and *hda19*. For the eight *ONSEN* copies, H3Ac (H3K9K27Ac and H3K9K14Ac) levels under non-stress (NS) were similar between Col and *hda19* (Fig 16A, B left). However, under HS conditions, Col exhibited higher H3Ac accumulation than *hda19* (Fig 16A, B left). Moreover, in the regions of downregulated TEs ($n=398$) in *hda19*, H3Ac levels were consistently lower than in Col under both non-stress (NS) and HS conditions (Fig 16A, B, right).

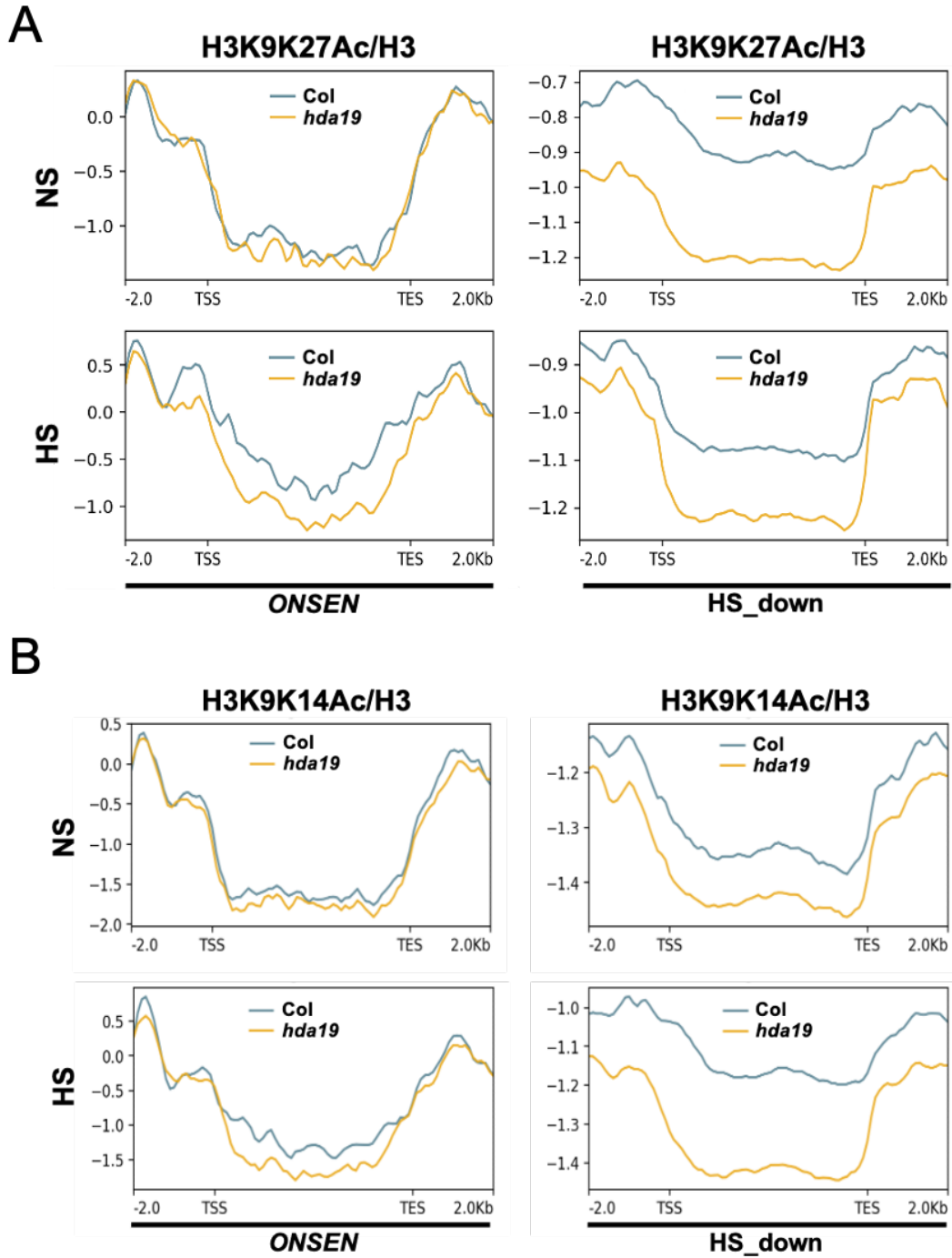


Figure 16: The *hda19* Mutation Slightly Impairs H3 Acetylation Establishment.

(A)-(B) Metaplots showing average H3K9K27Ac (A) and H3K9K14 (B) levels of the eight *ONSEN* copies (left) and down-regulated transposable elements (TEs) (right) in Col and *hda19* under non-stress (NS) and HS conditions. H3 was used as the normalization control.

5.2 *ONSEN* Transcription Suppression Linked to H3K4me3 Loss

Histone H3 tri-methylation at lysine 4 (H3K4me3) is generally associated with transcriptional activation, while di-methylation at lysine 9 (H3K9me2) is linked to transcriptional repression (Cadavid IC et al., 2023). Although H3K4me3 levels in the *ONSEN* region are similar between Col and *hda19* under normal conditions, HS impairs the establishment of H3K4me3 in the *ONSEN* region of *hda19* (Fig 17A). In heat-stressed *hda19*, transcriptionally downregulated TE regions (n=398) also show lower H3K4me3 levels compared to Col under both normal and HS conditions (Fig 17A). These results suggest that the loss of HDA19 may impair the establishment of H3K4me3, thereby contributing to the transcriptional repression of most TEs, including *ONSEN*.

On the other hand, in *hda19*, the transcriptional repression marks H3K9me2 are slightly higher in the regions of downregulated TE (n=398) compared to Col under non-stress (NS) conditions (Fig 17B). H3K9me2 is known to recruit CMT3 and CMT2, enzymes that catalyze the transfer of CHG and CHH methylation to the corresponding DNA regions (Du J et al., 2015). Therefore, the elevated CHG and CHH methylation levels in *ONSEN* in *hda19* may be linked to the high levels of H3K9me2 (Fig 17B, and see Fig 14E). However, after HS, the accumulation of H3K9me2 in Col was higher than in *hda19*, particularly in the transcription start site (TSS) region of *ONSEN* (Fig 17B). This suggests that H3K9me2 may not play a major role in regulating *ONSEN* transcription. In summary, the results indicate that the loss of HDA19 leads to improper establishment of H3K4me3, which may indirectly restrict the transcription of certain transposable elements.

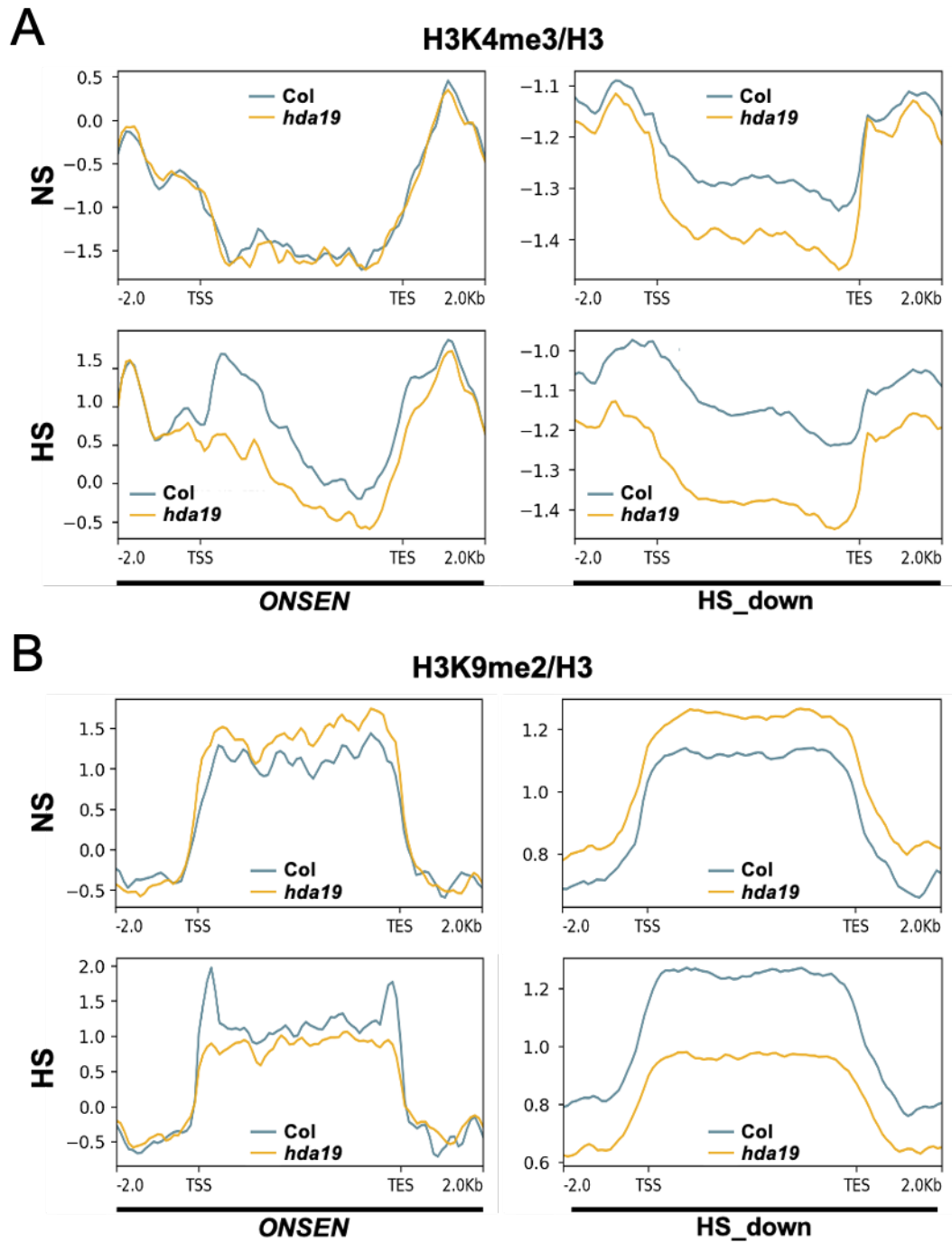


Figure 17: *ONSEN* Transcription Activity Is Associated with H3K4me3, Not H3K9me2.

(A) and (B) Metaplots showing average H3K4me3 (A) and H3K9me2 (B) levels across the eight *ONSEN* copies (left) and down-regulated TE (right) in Col and *hda19* under non-stress (NS) and heat stress (HS) conditions. H3 was used as the normalization control.

5.3 Effects of Heat Stress on Histone Acetylation and Methylation in *hda19*

I compared the levels of H3Ac and H3K4me3 under non-stress (NS) and HS conditions in both Col and *hda19*. The results showed that HS in Col led to an increased accumulation of H3Ac and H3K4me3, suggesting that enhanced *ONSEN* transcription might be linked to the establishment of these histone modifications (Fig 18). However, in *hda19*, HS does not result in higher levels of H3Ac and H3K4me3, indicating that the absence of HDA19 may disrupt the establishment of these modifications, which may indirectly lead to the repression of certain TE transcripts (Fig 18). In conclusion, the results indicate that HDA19 is involved in the establishment of H3Ac and H3K4me3 under HS, and these modifications likely facilitate the activation of *ONSEN* transcription.

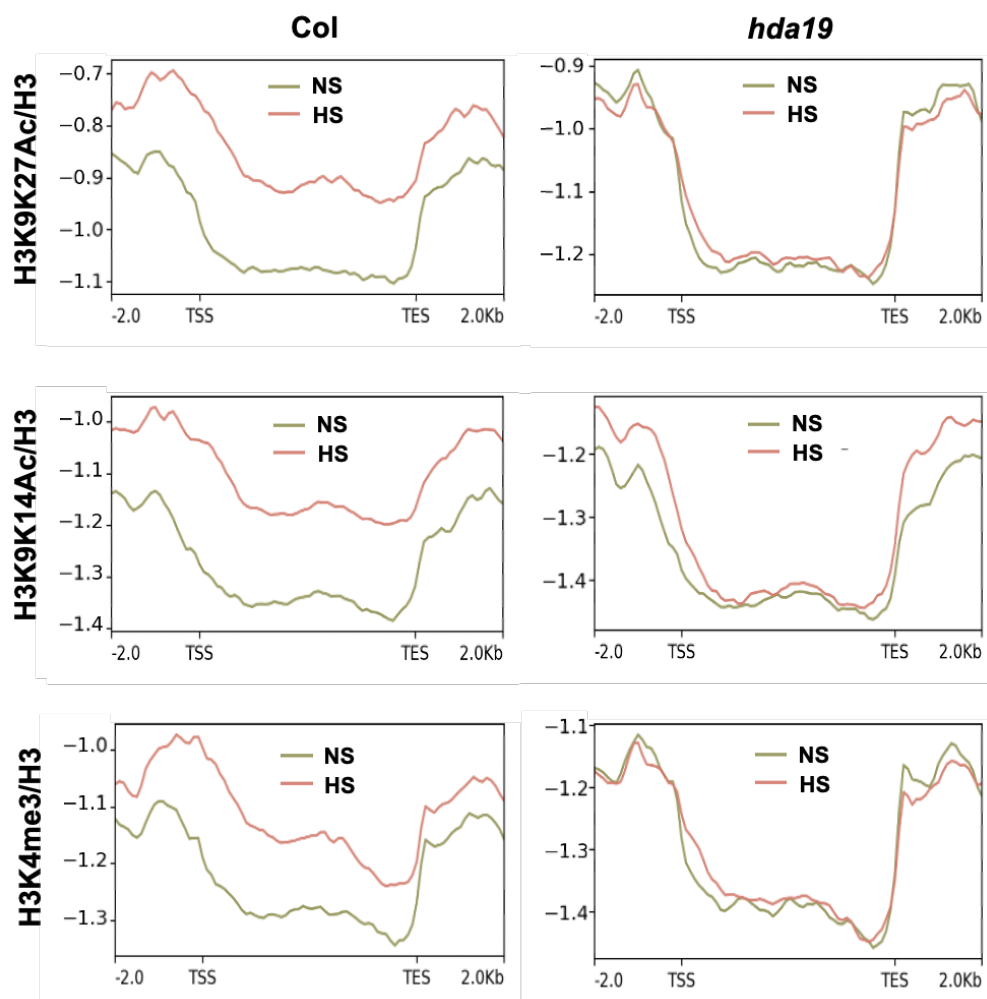


Figure 18: Impact of Heat Stress on Histone Modifications.

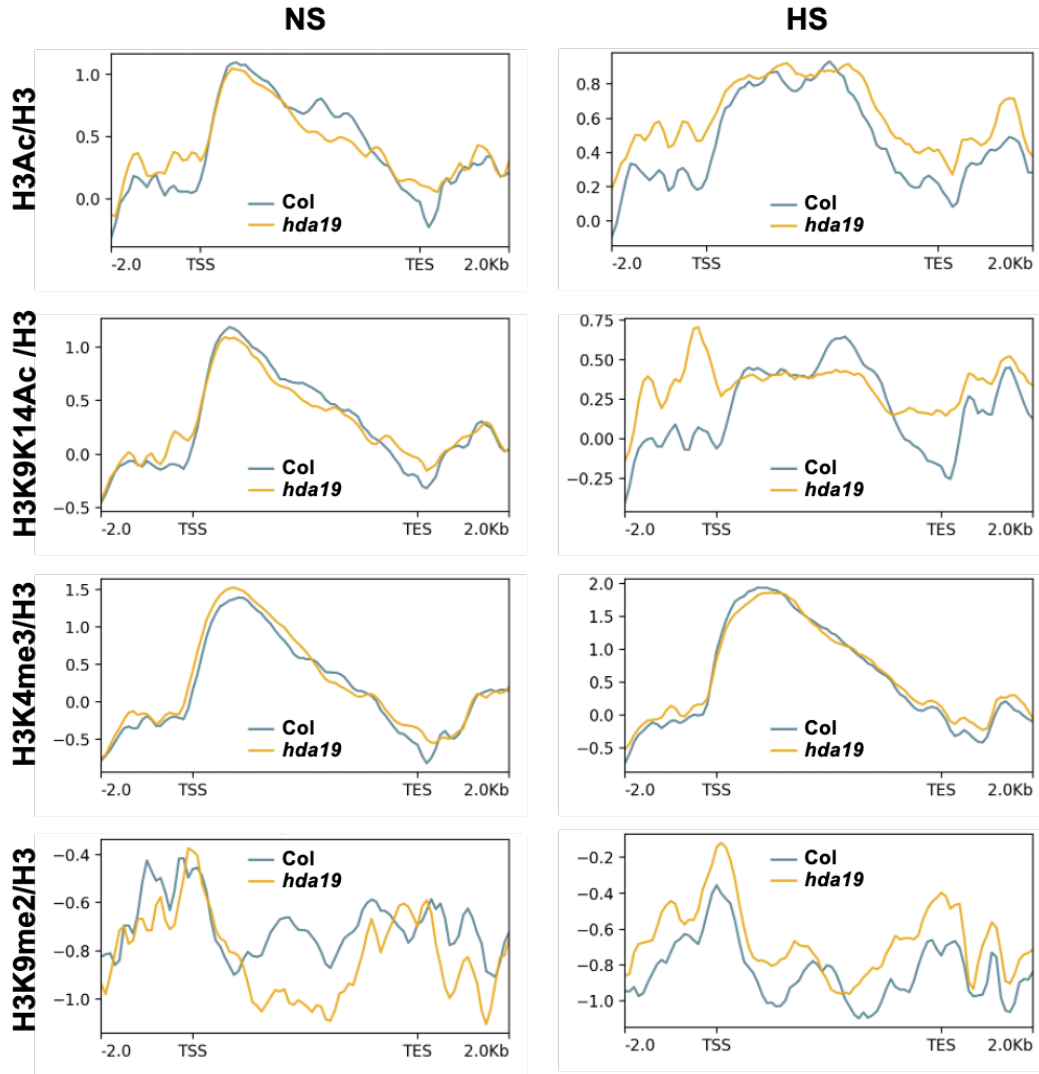
(E) Metaplots displaying the accumulation of H3K9K27Ac (top), H3K9K14Ac (middle), and H3K4me3 (bottom) signals across the down-regulated TE under non-stress (NS) and HS conditions in Col (left) and *hda19* (right), normalized to H3.

Discussion

Histone acetylation plays a crucial role in various cellular processes, including gene and transposon regulation, cell cycle progression, and even cancer development. This study aims to elucidate novel mechanisms of transposon transcription regulation associated with histone deacetylases (HDACs). Initially, a decrease in *ONSEN* transcription was observed in Col treated with the HDAC inhibitor TSA (Fig 11A, B). Given the diverse members of the HDAC family, *ONSEN* transcriptional levels were measured in various HDAC mutants. The results revealed that the absence of HDA19 under HS conditions led to a significant reduction in the transcriptional levels of approximately 398 TEs, including *ONSEN* (Fig 12A-C). Although histone acetylation is generally thought to enhance gene and TE transcription, HDA19 appears to exert a unique regulatory effect on TE transcription, suggesting a novel HDA19-dependent mechanism for transposon transcription regulation. Furthermore, compared to TSA treatment, the absence of HDA19 results in a greater degree of TE transcriptional silencing (Fig 11C, 12C). This indicates that TSA treatment and HDA19-induced transcriptional regulation may involve distinct mechanisms, highlighting the complexity of histone acetylation in regulating transposon activity.

HDA19 is a critical epigenetic regulatory factor, and extensive research has demonstrated its significant role in developmental processes and stress responses

in *Arabidopsis thaliana*. HDA19 has been implicated in the regulation genes associated with the Jasmonic Acid and ethylene signaling pathway, which are involved in pathogen responses (Zhou C et al., 2005). In addition, it plays a role in seed germination, salt stress responses, and gene expression induced by abscisic acid and salt stress (Chen LT et al., 2010). Interestingly, HDA19 exhibits functional differences in response to abiotic stress. Research by Mehdi S et al. (2016) showed that removing HDA19 enhances plant tolerance to salt stress. Similarly, a report by Ueda M et al. (2018) indicated that *hda19* mutants exhibit higher tolerance to drought and heat stress compared to Col plants. Given that the role of HDA19 in transposon silencing mechanisms remains unclear, this chapter focuses on elucidating the transcriptional regulation of TEs by the histone deacetylase family member HDA19. The activation of *ONSEN* transcription by heat shock is often associated with the heat stress response transcription factor HsfA2 (Cavrak V.V. et al., 2014). Although HDA19 is involved in various abiotic stress processes, the transcriptional levels of *HsfA2* in *hda19* mutants are similar to those in Col (Fig 13). In addition, *hda19* did not affect the levels of H3K9K27Ac, H3K9K14Ac, and H3K4me3 on genes related to the heat response pathway (Fig S15). Thus, HDA19 controls *ONSEN* following its activation by HsfA2.



Supplementary Figure 16: Effects of HDA19 on Histone Modifications in Heat Response Genes.

(A)-(D) Metaplots displaying the accumulation of H3K9K27Ac (A), H3K9K14Ac (B), H3K4me3 (C), and H3K9me2 (D) signals across the down-regulated TE under non-stress (NS) (left) and HS (right) conditions in heat response genes, normalized to H3. Heat response genes, including *HsfA2* (*At2g26150*), *HsfA3* (*At5g03720*), *HsfA7a* (*At3g51910*), *HsfA7b* (*At3g63350*), *HsfB1* (*At4g36990*), *HsfB2a* (*At4g11660*), *HsfB3* (*At2g41690*), *DREB2a* (*At1g06770*), and *MBF1c* (*At3g24500*).

The transcriptional activity of *ONSEN* is often linked to DNA methylation levels. In *hda19* mutants, increased CHG and CHH methylation levels were observed in the *ONSEN* region (Fig 14A, B, D, E). Differentially methylated regions (DMRs) were also identified in the three most transcriptionally active copies of *ONSEN* in *hda19* mutants, with hyper-DMRs specifically occurring in CHG and CHH contexts (Fig 14D). These results indicate that the absence of HDA19 leads to extensive DNA methylation in the *ONSEN* region, particularly in CHG and CHH contexts, which suppresses *ONSEN* transcription. In addition, *hda19* mutants displayed a large number of hyper-DMRs in CG, CHG, and CHH contexts, with relatively few hypo-DMRs (Fig 14C), indicating widespread increases in DNA methylation in the absence of HDA19. However, not all TE regions with downregulated transcription under HS (n=398) in *hda19* exhibited hyper-DMRs (Fig 14F), suggesting that only a portion of TE downregulation is associated with increased DNA methylation levels. Given the complexity of transposon (TE) silencing mechanisms in plants, HDA19 may regulate TE activity through pathways that are independent of DNA methylation. In summary, the reduced transcriptional activity of *ONSEN* in *hda19* is primarily due to increased DNA methylation (Fig 19A).

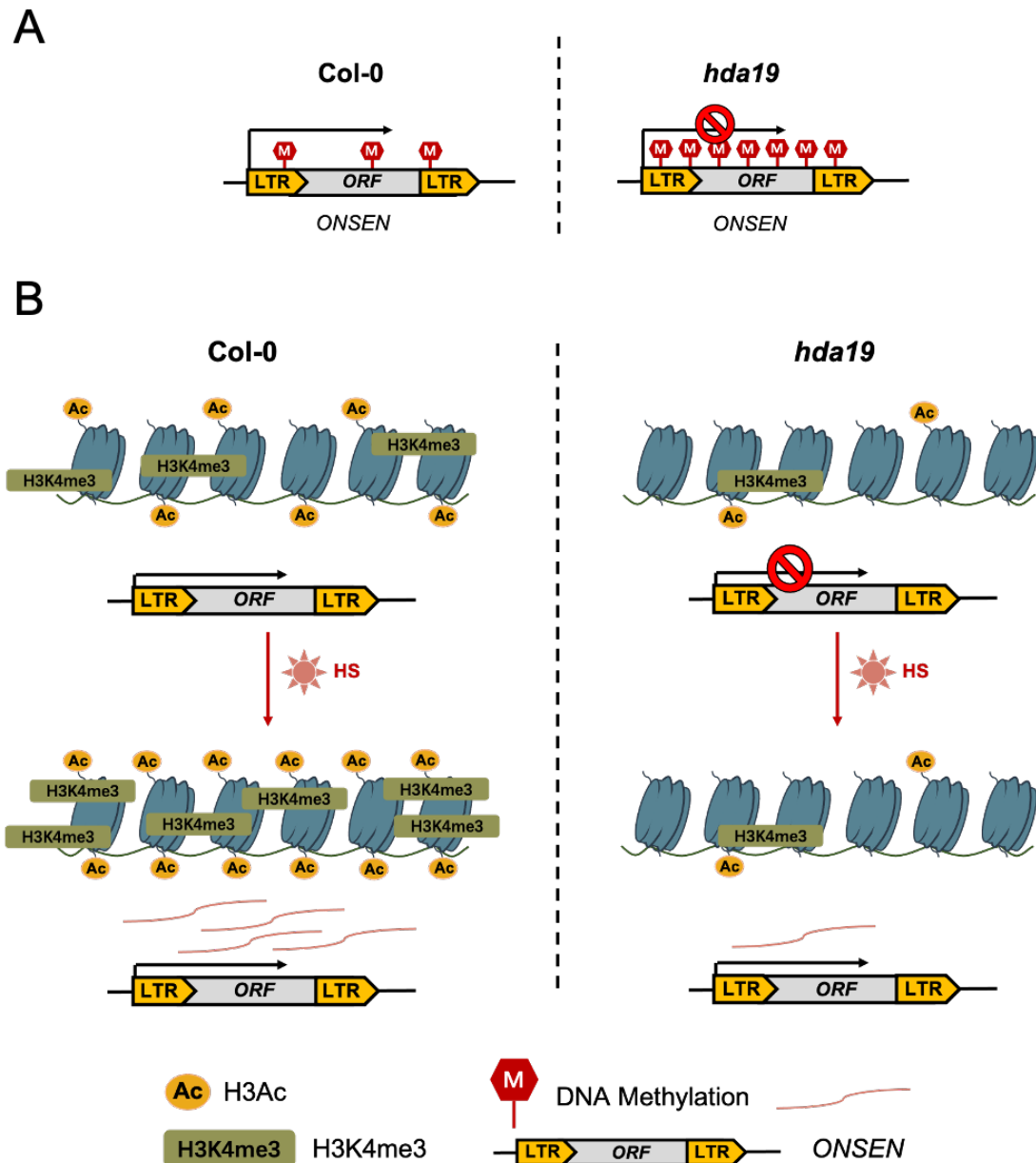


Figure 19: Model of HDA19 in *ONSEN* Regulation.

(A) Deletion of HDA19 increases DNA methylation levels in the *ONSEN* region. High DNA methylation levels in *hda19* result in silencing of *ONSEN* transcription.

(B) In *hda19*, low levels of H3Ac and H3K4me3 accumulation with high accumulation of H3K9me2, lead to the transcriptional silencing of *ONSEN*. In the Col, HS increases the accumulation of H3Ac and H3K4me3, activating *ONSEN* transcription. The

absence of HDA19 prevents the establishment of H3Ac, H3K9K14Ac, and H3K4me3 under HS, leading to reduced *ONSEN* transcription in *hda19*.

Next, I investigated the DNA methylation pathways associated with HDA19. Although the transcriptional level of *ONSEN* in *hda19/ddm1* was lower than that of *ddm1*, it was significantly higher than in *hda19*, suggesting that deletion of DDM1 in *hda19* partially restored *ONSEN* transcriptional activity (Fig 15A). However, the deletion of DDM1 did not reduce the DNA methylation level in the LTR region of *ONSEN* (Fig 15B), suggesting that DDM1 may regulate *ONSEN* transcription via HDA19 through a DNA methylation-independent mechanism.

Several studies have indicated that HDA19 is involved in histone modification processes, particularly histone acetylation. Research by Zhou Y et al. (2013) emphasized the crucial role of HDA19 in regulating H3K14Ac levels and gene activity. Their study found that *hda19* mutants show a significant increase in H3K14Ac levels on genes involved in seed development, resulting in ectopic expression of seed maturation genes in seedlings. Similarly, Zhou J et al. (2010) reported that increased transcription levels of specific genes in *hda19* mutants are associated with the accumulation of H3K9Ac. In addition, the absence of HDA19 enhances the expression of seed dormancy-related genes, which correlates with increased H3Ac levels and decreased H3K9me2 levels (Zhou Y et al., 2020). However, my study found that in *hda19* mutants, H3Ac and H3K4me3 levels were reduced at the *ONSEN* locus (Fig 17-18). This suggests that HDA19 may have distinct roles in regulating the activity of genes and TEs. While HDA19 appears to be critical for gene regulation related to plant growth and stress responses, a substantial number of TEs, rather than genes, are

transcriptionally silenced in *hda19* mutants (Fig 12C). Therefore, it can be concluded that HDA19's deacetylation function is typically targeted toward specific genes, particularly those involved in development and stress responses. Additionally, my research showed that HS influences histone modification levels. In the *hda19* mutant, HS did not induce changes in H3Ac and H3K4me3 levels, suggesting that HDA19 may play a role in establishing these modifications under HS in some TEs (Fig 19B).

In summary, the results of this chapter suggest that the role of HDA19 in regulating TE transcription differs from traditional perspectives. The loss of HDA19 under HS leads to the silencing of many TEs, including *ONSEN*. This silencing is likely associated with increased DNA methylation in CHG and CHH contexts, and the loss of histone modifications such as H3Ac and H3K4me3.

Chapter 3: Investigating a Novel Transposable Element Regulator of *ONSEN*

Introduction

Ethyl Methanesulfonate (EMS) is a widely used chemical mutagen that induces base substitutions and gene mutations, particularly in model plants such as *Arabidopsis thaliana* (Chen L et al., 2023). Although the involvement of the RdDM pathway in the transcription and transpositional regulation of *ONSEN* has been established, no clear correlation has been observed between *ONSEN* transcript level and transposon frequency. This suggests the existence of an unknown critical factor that regulates *ONSEN* activity. To explore this, a forward genetic approach was employed to identify novel genes involved in the regulation of *ONSEN*.

Materials

LTR: GFP transgenic plants were generated by inserting a sequence into the LTR of *ONSEN* (*AT5G13205*) (Fig 20A; Matsunaga W et al., 2015). The Col ecotype was used as the background for all SALK_line mutants (Tables 2 and 3). All SALK_line mutants were obtained from the ABRC stock center. The *nrpd1* (*C366150*) mutants were sourced from Ohio State University, Columbus, OH, USA. The *Arabidopsis thaliana* ecotype Ler-0 was provided by Professor Atsushi Kato from Hokkaido University.

Results

1. Transgenerational Transposition of *ONSEN* in EMS Mutants

Using EMS to induce mutations in *Arabidopsis thaliana* is a common approach in genetic research to uncover novel gene functions. *ONSEN* activity can be categorized into transcriptional and transpositional activity. Traditionally, techniques such as qRT-

PCR and Southern blot are used to measure *ONSEN* transcription and transposition, but these methods are time-consuming and costly. To streamline the process, a simpler method was employed in this experiment to preliminarily assess *ONSEN*'s transcriptional activity.

To visualize *ONSEN* transcription, LTR: GFP transgenic plants were used as the material for EMS screening. These transgenic plants have the GFP (green fluorescent protein) gene inserted downstream of the LTR (long terminal repeat) sequence in the *ONSEN* promoter region (Fig 20A), allowing for the detection of *ONSEN* transcription activity via microscopy. Previous studies have demonstrated a correlation between the transcription level of *ONSEN* and the intensity of GFP fluorescence (Matsunaga W et al., 2015).

First, approximately 5000 LTR: GFP transgenic plants were treated with EMS. Since *ONSEN* transcription requires HS, the second generation of EMS-treated plants was subjected to HS treatment. Fluorescence microscopy was used to screen for individuals displaying GFP fluorescence, indicating increased *ONSEN* transcription. The selected individuals were self-crossed and re-screened for GFP fluorescence to confirm stable inheritance of the EMS-induced mutations. Subsequently, individuals with reproducible GFP fluorescence were subjected to Southern blot analysis to check for *ONSEN* transposition. Ultimately, 11 mutants (*screen 1-11*) exhibited GFP fluorescence and detectable *ONSEN* transposition (Fig 20B). Further analysis confirmed reproducible *ONSEN* transposition in *screen 1*, *screen 2*, *screen 3*, *screen 7*, *screen 8*, *screen 9*, *screen 10*, and *screen 11* (Fig 20C). In summary, these preliminary results indicate that EMS treatment successfully generated new mutants (*screens 1-3, 7-11*) with novel genetic variations associated with *ONSEN* transposition, demonstrating the practical applicability of this genetic screening approach.

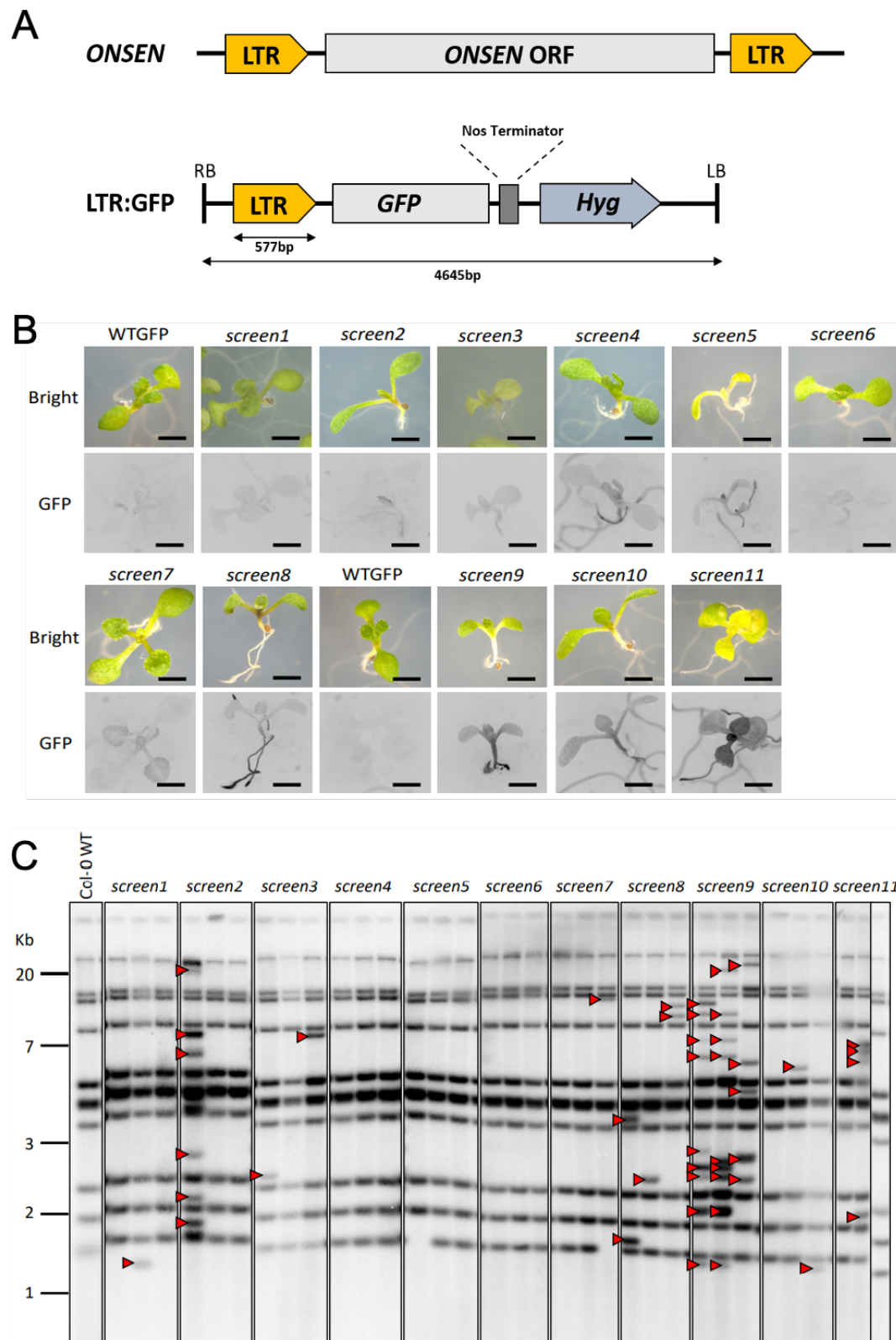


Figure 20: GFP Screening Identified New Mutations Affecting *ONSEN* Transposon Regulation.

(A) Diagram of the construct for *ONSEN* (top) and LTR (bottom).

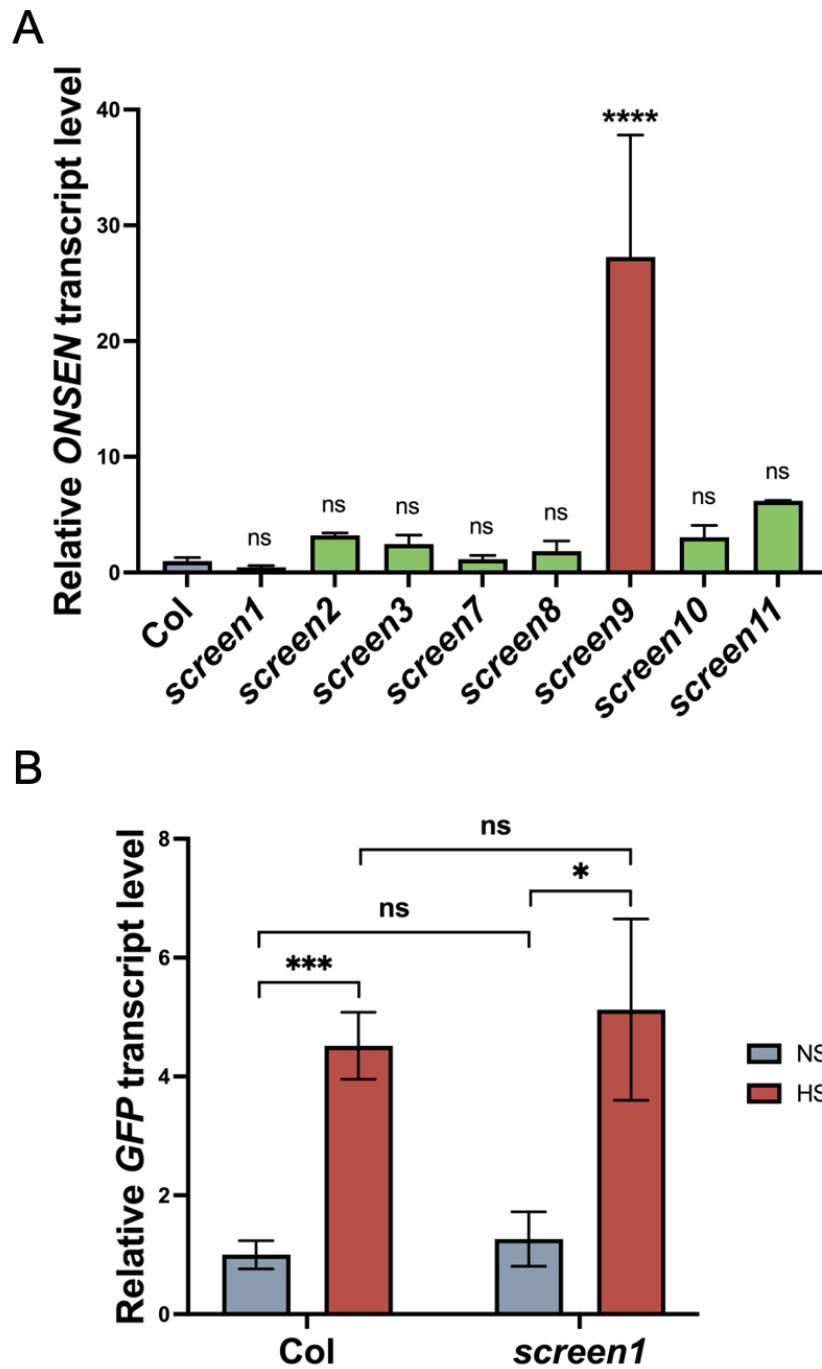
(B) Fluorescence microscopy images of 11 selected stable mutants (*screen 1-11*) showing GFP fluorescence. Images were taken after heat stress (HS) treatment.

(C) Transgenerational transposition of *ONSEN* examined in the 11 GFP-expressing screen lines shown in (B). The leftmost panel in each figure shows the Col pattern under non-stress (NS) conditions. Red arrows indicate new *ONSEN* insertions.

2. Characteristics of *Screen 1*

2.1 *ONSEN* Transposition Observed in *Screen 1* Despite Low Transcription Levels

Next, qRT-PCR was used to measure the transcription levels of *ONSEN* in the 11 mutants. In *screen 1*, *screen 3*, *screen 7*, *screen 8*, and *screen 10*, *ONSEN* transcription levels were comparable to those in Col (Fig S17 A). Previous studies, along with results from Chapter 1, indicate that mutants exhibiting *ONSEN* transposition typically show relatively high transcription levels of *ONSEN* (Fig 1-5; Hayashi Y et al., 2020). Since the aim of this study is to identify new regulatory factors of *ONSEN* and to avoid mutants with both high *ONSEN* transcription and transposition, I focused on *screen 1* (Fig 20C, Fig S16A). Although *screen 1* exhibited low *ONSEN* transcription and low cDNA levels, transposition still occurred in its progeny (Fig 20C; Fig 21A, B; Fig S17B). The low transcriptional level of *ONSEN* in *screen 1* suggests that this mutant is not involved in the transcriptional regulation of *ONSEN*. However, the occurrence of *ONSEN* transposition indicates that *screen 1* is likely involved in the transposition regulating mechanism of *ONSEN*.



Supplementary Figure 17: Low Transcriptional Levels of *ONSEN* and *GFP* in *Screen 1*.

(A) Relative transcription levels of *ONSEN* in *screens 1–11* after 24 hours of heat stress (HS). Asterisks indicate significant differences compared to Col (one-way ANOVA and Dunnett's test, *. P < 0.05; **. P < 0.01; ***. P < 0.001); "ns" denotes no significant differences (one-way ANOVA and Dunnett's test, P > 0.05).

(B) Transcriptional levels of GFP in Col and *screen 1* under non-stress (NS) and heat stress (HS) conditions. Significant differences between groups are marked with asterisks (Student's t-test, *, $P < 0.05$; ***, $P < 0.001$). "ns" denotes no significant differences (Student's t-test, $P > 0.05$).

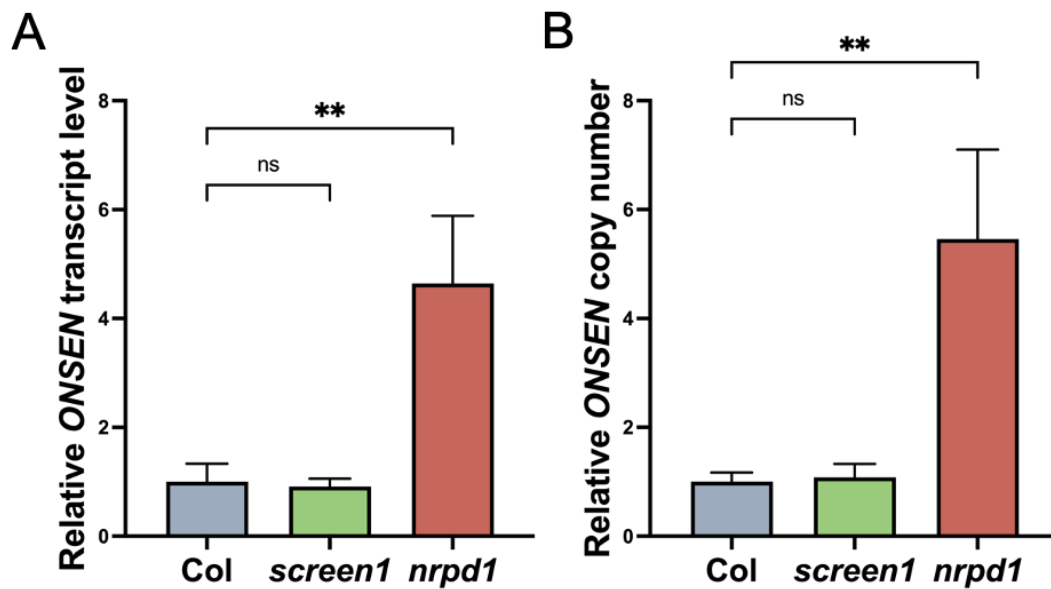


Figure 21: *ONSEN* Exhibits Reduced Transcriptional Activity in *Screen 1*.

(A) and (B) Relative transcription levels (A) and relative copy numbers (B) of *ONSEN* in Col, *screen 1*, and *nrpd1* detected by RT-qPCR and qPCR under HS conditions. Asterisks indicate significant differences between the two groups (one-way ANOVA and Dunnett's test, *, $P < 0.05$; **, $P < 0.01$).

2.2 The Gene Responsible for *Screen 1* is Recessive

To investigate the gene responsible for *screen 1*, a cross was made between *screen 1* and Col, and the resulting F1 heterozygous individuals were subjected to heat stress (HS) treatment. The progeny (F2) of these heat-treated F1 individuals were then

analyzed for *ONSEN* transposition. The results showed no evidence of *ONSEN* transposition in the F2 generation, indicating that no transposition occurred in the F1 heterozygous individuals. This suggests that the gene responsible for *screen 1* is recessive (Fig 22). Since *screen 1* is recessive, only homozygous individuals are expected to exhibit *ONSEN* transposition.

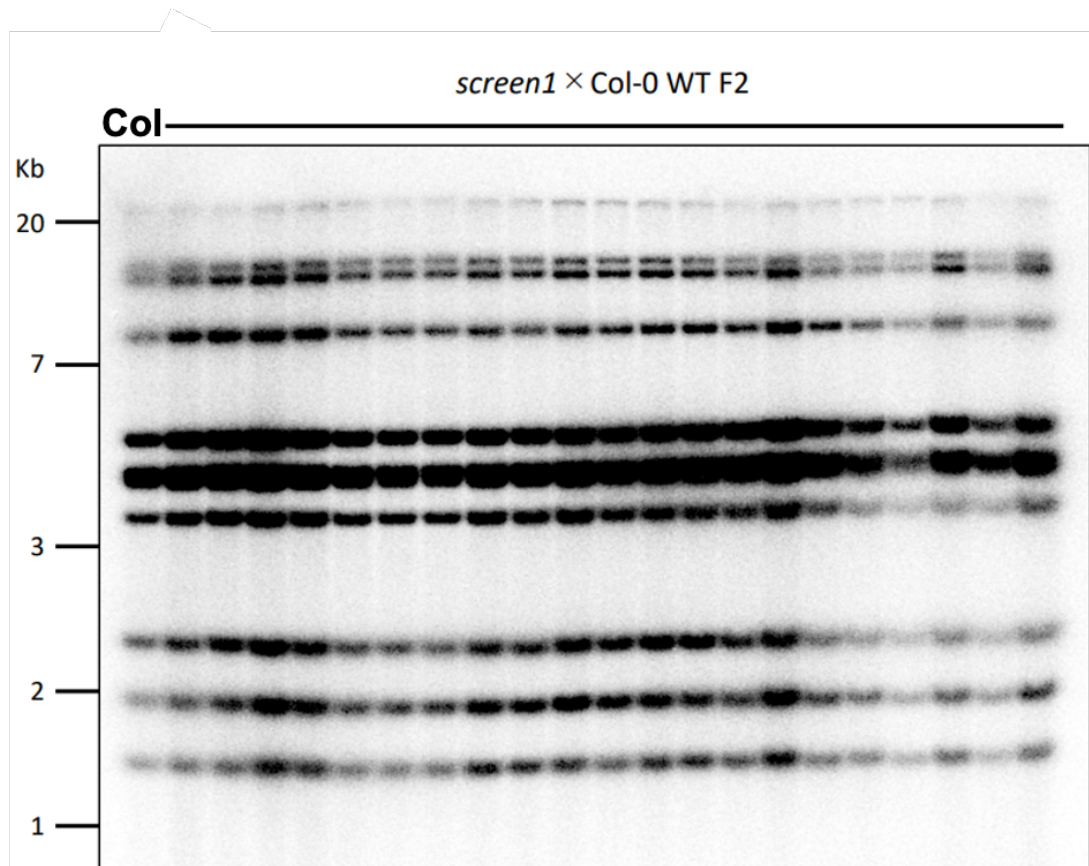


Figure 22: *Screen 1* Is a Recessive Gene.

Southern blot analysis shows that the F1 progeny from the cross between *screen 1* and Col do not exhibit new *ONSEN* insertions. Seven individuals from the three parent plants subjected to heat stress (HS) were analyzed. The leftmost panel in the figure shows the Col pattern under non-stress (NS) conditions.

2.3 NGS Findings for *Screen 1*

To identify the gene responsible for *screen 1*, the Mut-Map method was employed. Mut-Map is commonly used to map recessive genes associated with mutant phenotypes. The process involves crossing the mutant of interest with its original Col strain to obtain F1 plants, which are then self-crossed to produce F2 progeny. In the F2 generation, mutant individuals displaying significant phenotypic changes (selected F2) are chosen, and their DNA is pooled for sequencing. The sequencing results are then compared with the original Col strain to calculate the Single Nucleotide Polymorphism (SNP) index, defined as the ratio of mutant SNP reads to the total number of reads. In the F2 generation, most SNPs segregate in a 1:1 ratio. However, since the target gene is recessive, SNPs associated with the mutant phenotype will be present in all selected mutants, resulting in an SNP index close to 1. For SNPs unrelated to the phenotype, the segregation ratio in the selected mutants will be approximately 1:2, yielding an SNP index close to 0.5. Therefore, regions with an SNP index close to 1 are likely to contain the gene responsible for the target phenotype (Abe A et al., 2012).

In this study, *screen 1* was crossed with Col, and the F2 generation was subjected to HS treatment. Southern blot analysis was performed on F3 progeny to detect *ONSEN* transposition. The F2 individuals were identified based on the F3 progeny that exhibited *ONSEN* transposition. Next-Generation Sequencing (NGS) was then conducted on 40 selected F2 individuals. Mapping results indicated that the nucleotide substitution rate in the 3.5-10 megabase pair (Mbp) region on chromosome 3 was approximately 70-100%, suggesting that the gene responsible for *screen 1* is likely located within this region (Fig 23).

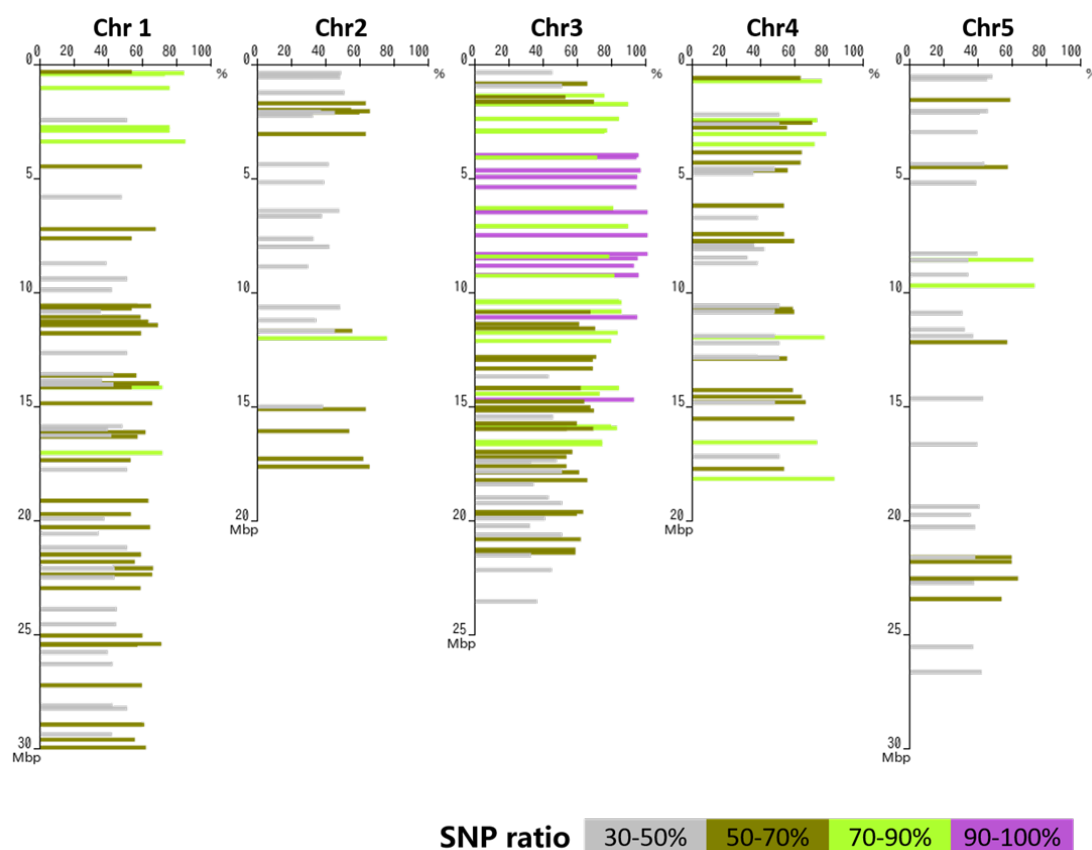


Figure 23: Screen 1 Genome Mapping Results Using MUT-MAP.

The SNP index was determined by comparing the *screen 1* sequence with Col. The SNP-index plot displays chromosomal positions and base substitution rates using a color gradient: grey indicates a substitution rate of 30-50%, green represents 50-70%, pale green denotes 70-90%, and purple signifies 90-100%.

3. Analysis of Transposition in SALK Line Mutants

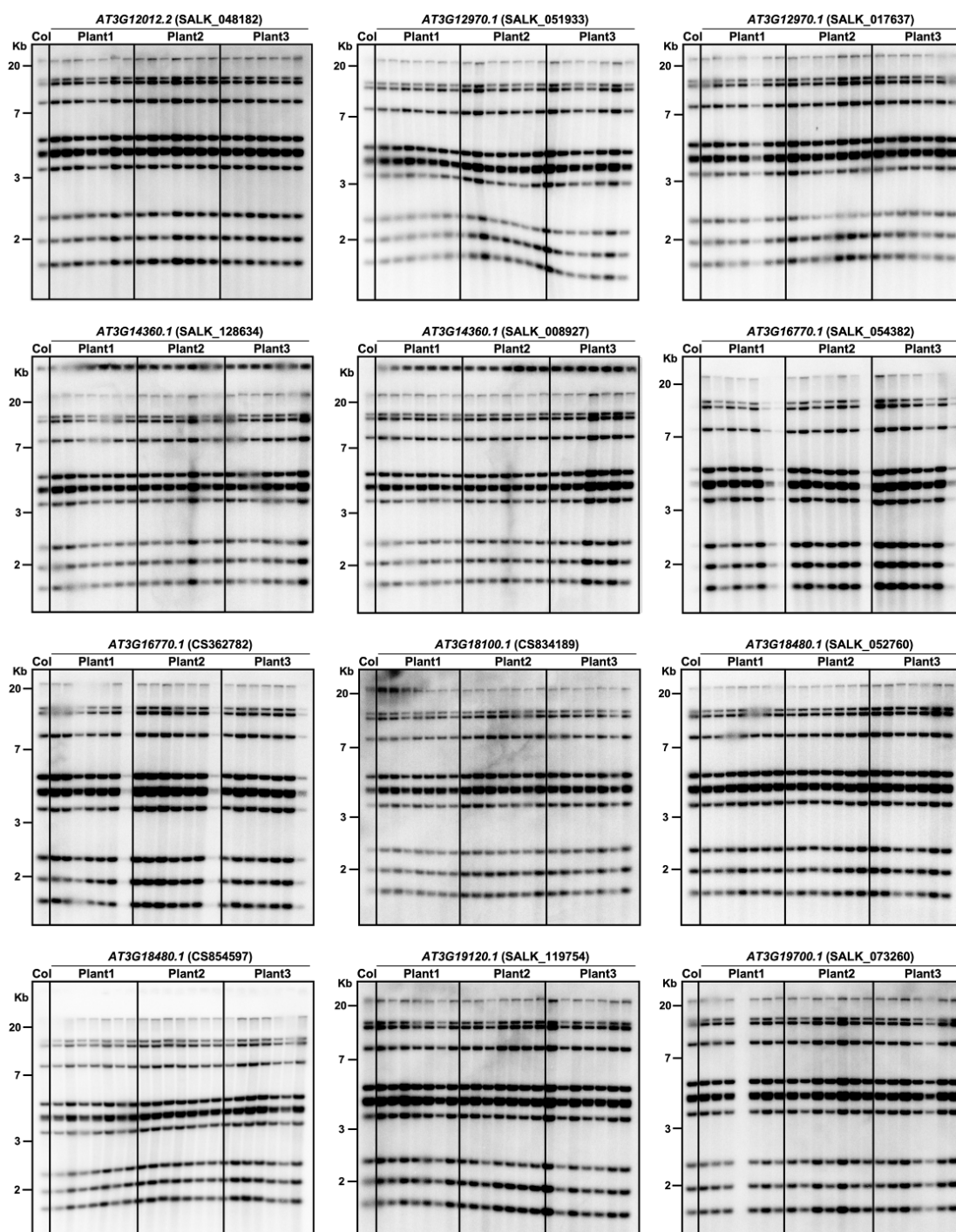
To identify the gene responsible for *screen 1*, candidate genes with amino acid sequence changes (non-synonymous mutation) were selected from the 3.5-10 Mbp region on chromosome 3, based on Single Nucleotide Polymorphisms (SNPs) identified from the NGS results (Table 2). Corresponding SALK mutants for these candidate genes were then subjected to *ONSEN* transposition analysis (Table 2). However, Southern blot

analysis revealed no new *ONSEN* insertions in any of the SALK line progenies (Fig 24), suggesting that the selected SALK lines do not contain mutants related to the gene responsible for *screen 1*.

List of candidate gene and SALK mutants for *screen1* (1st)

	Candidate gene	SALK_line	position (Mbp)	base substitution	annotation
1	AT3G12020.2	SALK_048182	3.8	G > A	microtubule motor
2	AT3G12970.1	SALK_051933	4.1	C > T	serine/arginine repetitive matrix-like protein;(source:Araport11)
3		SALK_017637			
4	AT3G14360.1	SALK_128634	4.8	G > A	Lipid droplet-associated triacylglycerol lipase (TAG) involved in pollen tube growth.
5		SALK_008927			
6	AT3G16770.1	SALK_054382	5.7	G > A	Encodes a member of the ERF (ethylene response factor) subfamily B-2 of the plant specific ERF/AP2 transcription factor family (RAP2.3).
7		CS362782			
8	AT3G18100.1	CS834189	6.2	G > A	Member of the R2R3 transcription factor gene family.
9	AT3G18480.1	SALK_052760	6.3	G > A	This gene is predicted to encode a protein that functions as a Golgi apparatus structural component, known as a golgin in mammals and yeast.
10		CS854597			
11	AT3G19120.1	SALK_119754	6.6	C > T	PIF / Ping-Pong family of plant transposase;(source:Araport11)
12	AT3G19700.1	SALK_073260	6.8	G > A	Encodes leucine rich repeat (LRR) kinase.
13		SALK_078025			
14	AT3G22170.1	SALK_002711C	7.8	C > T	A component of the PHYA signaling network, mediates the FR-HIR response to far-red light in concert with FAR1.
15	AT3G23200.1	CS907248	8.3	C > T	Uncharacterized protein family (UPF0497);(source:Araport11)
16		SALK_017299			
17	AT3G24480.1	SALK_040784	8.9	C > T	Leucine-rich repeat (LRR) family protein;(source:Araport11)
18	AT3G24530.1	SALK_016578	8.9	G > A	AAA-type ATPase family protein / ankyrin repeat family protein;(source:Araport11)
19	AT3G24890.2	CS813672	9.1	C > T	member of VAMP72 Gene Family
20		SALK_064166			

Table 1: List of candidate gene and SALK mutants for *screen 1* (1st screening).



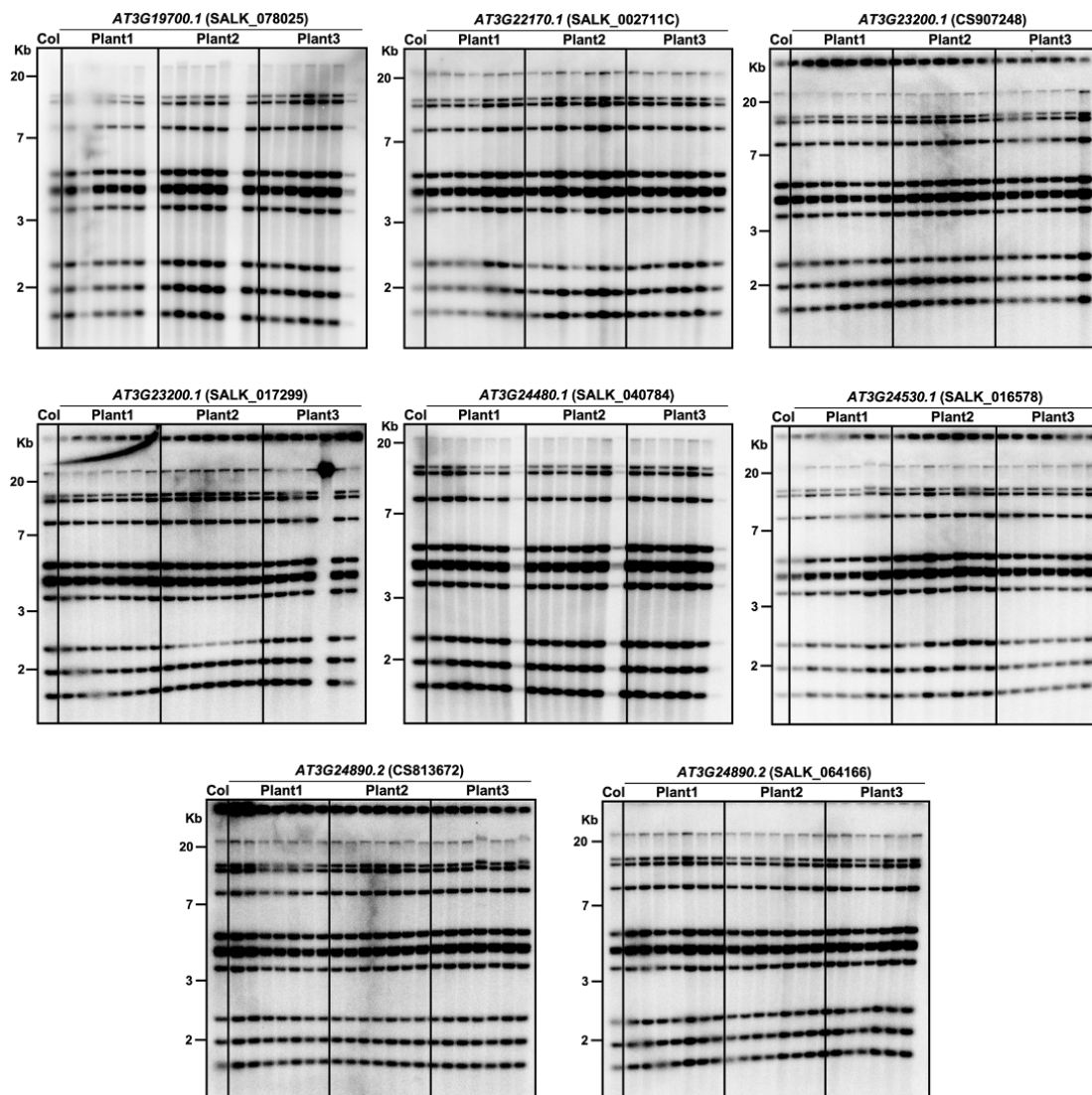


Figure 24: Southern Blot Analysis of *ONSEN* in the Progeny of Heat-Stressed SALK Lines (Mutant List Shown in Table 2).

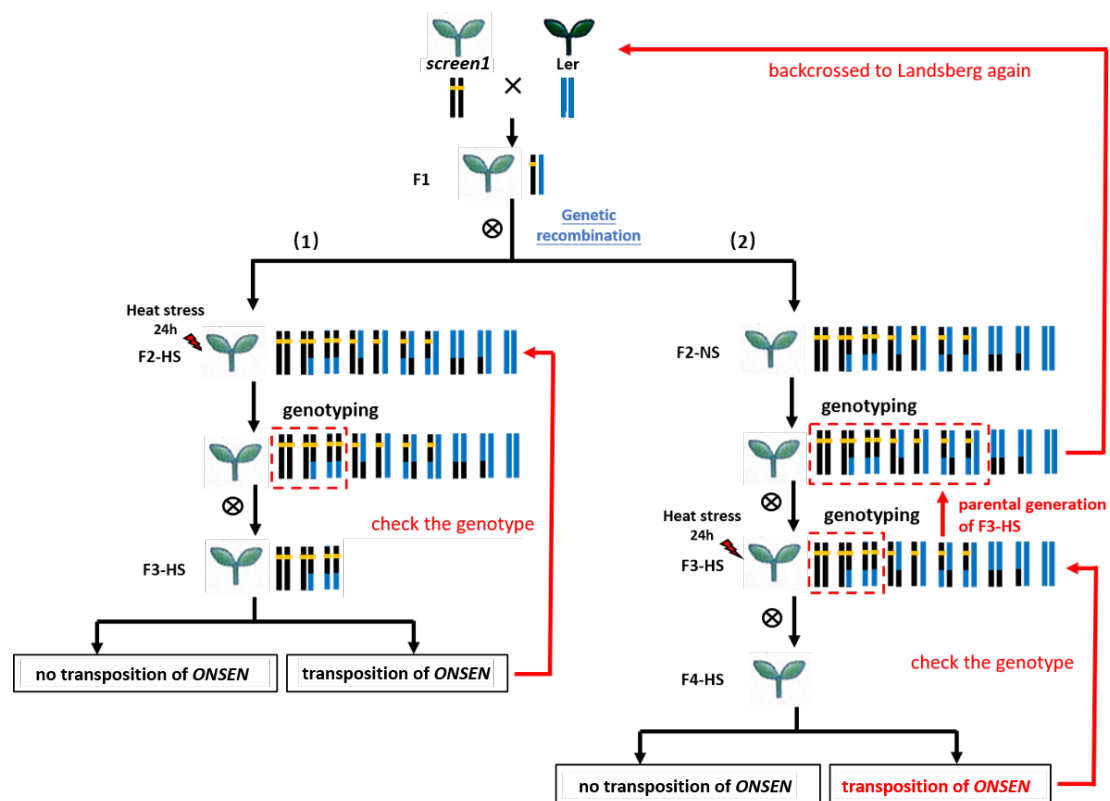
Seven individuals from each of the three parent plants subjected to heat stress (HS) were analyzed. The leftmost band in each image represents Col under non-stress (NS) conditions.

4. Classical Mapping Approach to Investigate *Screen 1*

4.1 Basics and Workflow of Classical Genetic Mapping

Classical genetic mapping is a traditional technique used to determine the location of genes. This approach is particularly useful in cases where complete genome sequence information is unavailable, making it practical for localizing and studying gene function. The method involves using genetic markers in breeding experiments to generate progeny with mixed genetic backgrounds. By analyzing the linkage between genetic markers (such as SNPs or microsatellite markers) and traits of interest, the gene's position on the chromosome can be inferred. The degree of linkage between markers and traits allows researchers to deduce the gene's relative position on the chromosome.

Since the SALK lines did not reveal the gene responsible for *screen 1*, I adopted a classical mapping to narrow down the region containing the *screen 1* gene. *A.thaliana* has many ecotypes with DNA sequence polymorphisms and Landsberg (Ler) strain is commonly used alongside the Col strain in experimental research. Since screen 1 was in the Col background, *screen 1* was crossed with Ler (F0), resulting in F1 progeny heterozygous for both Col and Ler alleles. When the F1 plants were self-crossed, chromatid crossover during meiosis, producing F2 progeny with a mixture of homozygous and heterozygous combinations of Col and Ler alleles (Fig S18). To identify the ecotype at specific chromosomal locations, DNA markers based on SNPs and microsatellites were used in the mapping process. These markers, designed as PCR primers, amplified target regions of both Col and Ler. Due to size differences in the amplified fragments between the Col and Ler ecotypes, the PCR product sizes could be easily distinguished (Fig S19).



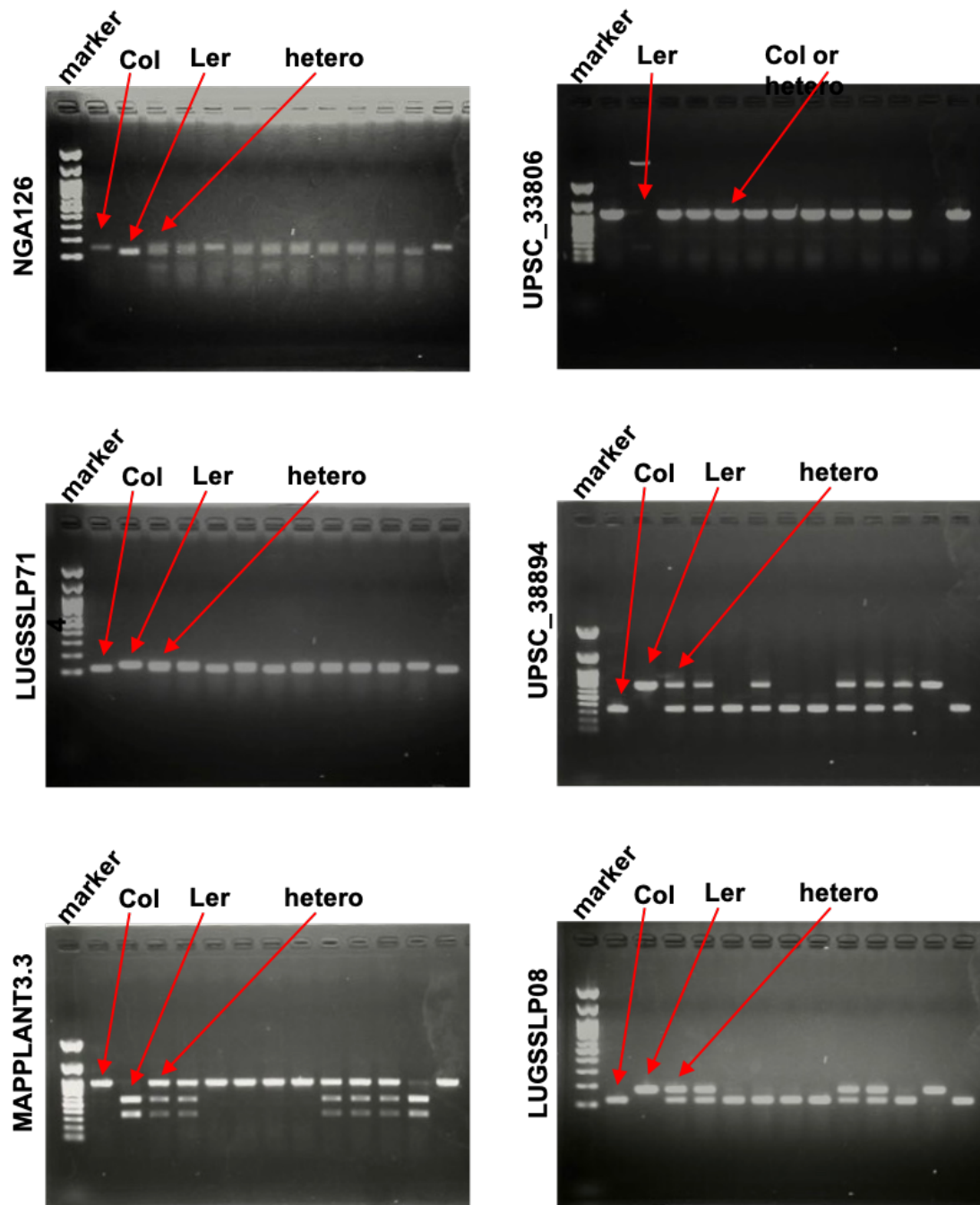
Supplementary Figure 18: Schematic Diagram of the Working Principle of Classical Mapping.

The diagram illustrates the mapping process, with black representing the Col genome, blue representing the Ler genome, and yellow indicating the position of *screen 1*.

1. **Hybridization:** *Screen 1* was crossed with the Arabidopsis wild type *Ler*.
2. **PCR Screening:** Individuals were screened using PCR to identify those with the Col genotype in the 3.1–10.7 Mbp region.
3. ***ONSEN* Transposition Analysis:** Individuals showing *ONSEN* transposition were selected for further analysis.
4. **Iterative Screening:** The process was repeated until the Col genotype region was narrowed down.

Note: *ONSEN* transposition occurs exclusively after heat stress (HS), which may

introduce additional mutations. Only lines not subjected to HS were used for backcrossing to avoid these complications. To confirm *ONSEN* transposition, two rounds of genetic screening were conducted (right).



Supplementary Figure 19: PCR Amplification Using DNA Markers.

The figure displays electrophoretic gel images of PCR amplification using six DNA

markers designed to differentiate between the Col and Ler genotypes.

DNA markers were selected from the region expected to contain the *screen 1* gene, specifically within the 3.1-10.7 Mbp range on chromosome 3 (Fig 25). Since *screen 1* is exclusively present in the Col background, PCR amplification using these DNA markers was performed on F2 individuals from the cross between *screen 1* and Ler. Near-isogenic lines (NILs) with the Col-type genotype within the 3.1-10.7 Mbp range were selected. These NILs were then subjected to *ONSEN* transposon analysis, and those in which *ONSEN* transposition occurred were identified (Fig S18). The NILs showing *ONSEN* transposition were used for the next round of cross-selection, progressively narrowing down the Col-type region where *ONSEN* transposition occurred (Fig S18). However, it is important to note that *ONSEN* transposes only after HS treatment, which may induce unknown mutations in plants. Therefore, lines not subjected to HS were used for backcrossing. Before performing *ONSEN* transposition analysis on the NILs, two rounds of selfing and genetic screening were conducted to stabilize the NILs (Fig S18).

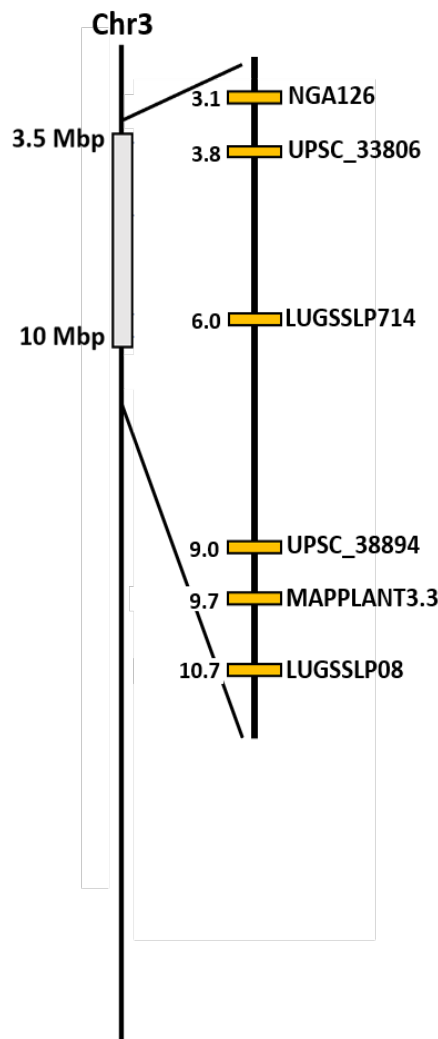


Figure 25: Locations of the DNA Markers Used for Classical Mapping. The figure displays the names and positions of DNA markers on chromosome 3 used in classical mapping.

4.2 Detection of *ONSEN* Transposition in Four NILs

After extensive crossing and screening, *ONSEN* transposition was observed in four NILs, designated as A, B, C, and D (Fig 26A, B). The transcription levels and copy number of *ONSEN* in these four NILs were comparable to those in Col, suggesting that the gene responsible for *screen 1* is present in these selected lines. The results indicated

that *screen 1* is likely located within the region of chromosome 3, between 3.8 Mbp and 9.7 Mbp.

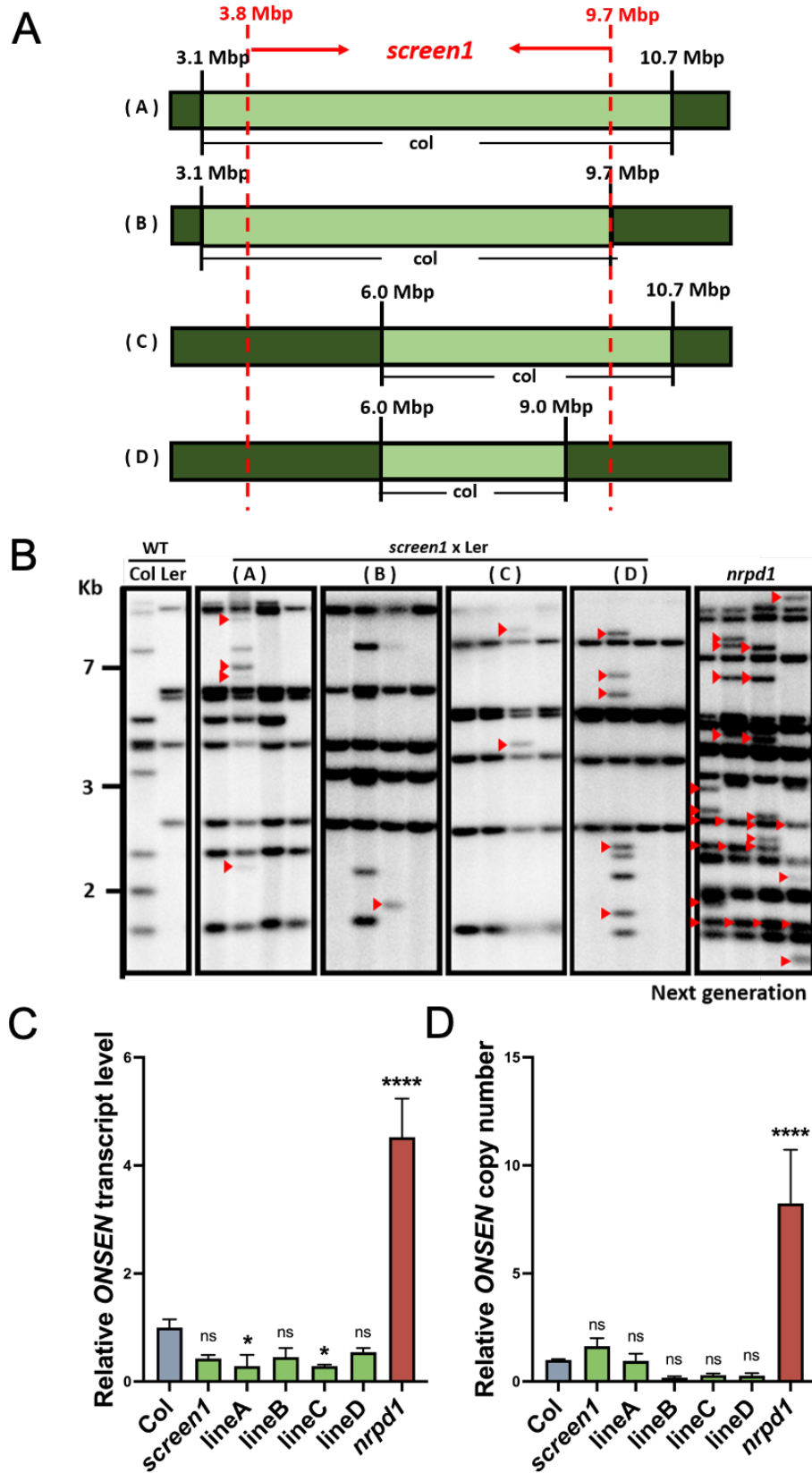


Figure 26: Narrowing Down the Region of *Screen 1* Using Classical Mapping.

(A) Genome structure of the Near-Isogenic Lines (NILs) for *screen 1*. After multiple rounds of screening and testing, *ONSEN* transposition was detected in four NILs (A-D). The diagrams show the gene patterns inherited by Col and Ler in NILs A-D.

(B) Southern blot showing new *ONSEN* insertions in NILs A-D as depicted in (A). The leftmost panel in the figure shows Col under non-stress (NS) conditions. Red arrows indicate new *ONSEN* insertions.

(C) and (D) Relative transcription levels (C) and relative copy numbers (D) of *ONSEN* in Col, *screen 1*, NILs (A-D), and *nprp1*, detected by RT-qPCR and qPCR under heat stress (HS) conditions. Asterisks indicate significant differences compared to Col (one-way ANOVA and Dunnett's test, *. $P < 0.05$; **. $P < 0.01$; ***. $P < 0.001$; ****. $P < 0.0001$); "ns" denotes no significant differences (one-way ANOVA and Dunnett's test, $P > 0.05$).

4.3 Whole-Genome Resequencing Identifies *Screen 1* Within the 5.3 to 8.9 Mbp Region

Due to the limited availability of DNA markers, further narrowing of the *screen 1* region using classical mapping became challenging. To overcome this limitation, whole-genome resequencing was performed on Col, Ler, *screen 1*, and NILs-D. SNP identification and GVCF (Genome Variant Call Format) variant detection revealed that within the 5.3-8.9 Mbp range on chromosome 3, NILs-D exhibited the same genotype as Col and *screen 1*. In contrast, regions below 5.3 Mbp and above 8.9 Mbp were more like Ler (Fig 27). These resequencing results strongly suggest that *screen 1* is located within the 5.3-8.9 Mbp. However, the absence of suitable DNA markers within this

range hindered further rounds of crossing and screening.

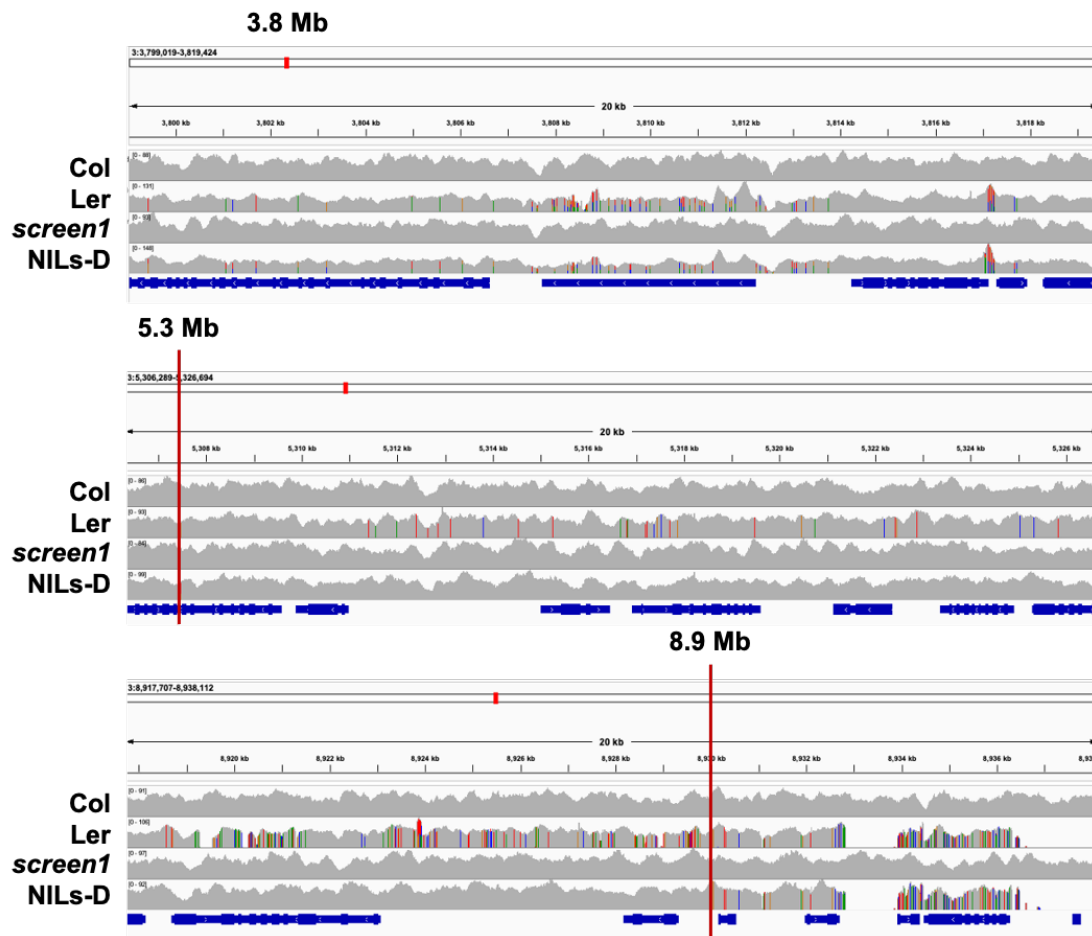


Figure 27: Resequencing Results of *Screen 1*.

Resequencing results indicate that the region containing *screen 1* has been narrowed down to 5.3-8.9 Mbp. IGV screenshots display the resequencing data for Col, Ler, *screen 1*, and NILs-D.

5. Transpositional Study of *ONSEN* in Further Selected SALK Mutants

5.1 Analysis of *ONSEN* Transposition in Additional SALK Mutants

Following the identification of the *screen 1* region within the 5.3-8.9 Mbp range from resequencing data, candidate genes in this region were selected for further analysis.

Corresponding SALK mutants for these candidate genes were obtained from TAIR (Table 2). These SALK mutants were subjected to Southern blot analysis to detect potential new *ONSEN* insertions. However, none of the SALK lines exhibited detectable new *ONSEN* transpositions (Fig 28).

List of candidate gene and SALK mutants for screen1 (2ed)

	Candidate gene	SALK_line	position (Mbp)	base substitution	annotation
1	AT3G16340.1	SALK_100847	5.5	C-T	Encodes a p-coumaryl alcohol exporter involved in lignin biosynthesis.
2		SALK_102381			
3	AT3G16910.1	SALK_133808	5.7	G-A	Encodes a peroxisomal protein with acetyl-CoA synthetase activity that is responsible for the activation of acetate for entry into the glyoxylate cycle.
4		SALK_055243			
5	AT3G18020.1	SALK_019722	6.2	G-A	Pentatricopeptide repeat (PPR) superfamily protein;(source:Araport11)
6	AT3G02875.1	CS827124	6.3	G-A	Hydrolyzes amino acid conjugates of the plant growth regulator indole-3-acetic acid (IAA), including IAA-Leu and IAA-Phe. Uses Mg and Co ions as cofactors.
7	AT3G18550.2	CS857231	6.4	G-A	Encodes a TCP transcription factor, closely related to teosinte branched1, arrests axillary bud development and prevents axillary bud outgrowth. Role in flowering control.
8		SALK_091920			
9	AT3G16400.1	SALK_072600	5.6	T--	Encodes a nitrile-specifier protein NSP1 responsible for constitutive and herbivore-induced simple nitrile formation in rosette leaves.
10	AT3G18730.1	SALK_057328	6.5	G-A	Encodes a protein with an important role in cell division control and plant morphogenesis and may also play a role in genome maintenance.
11	AT3G19960.2	SALK_074538C	6.9	G-A	member of Myosin-like proteins
12	AT3G20950.1	CS875061	7.3	G-A	member of CYP705A
#1	AT3G16240.1		5.5	G-A	Delta tonoplast intrinsic protein, functions as a water channel and ammonium (NH ₃) transporter.
#2	AT3G22555.1		8	C-T	pseudogene of putative DNA methyltransferase;(source:Araport11)

Table 2: List of candidate gene and SALK mutants for *screen 1* (2nd screening).

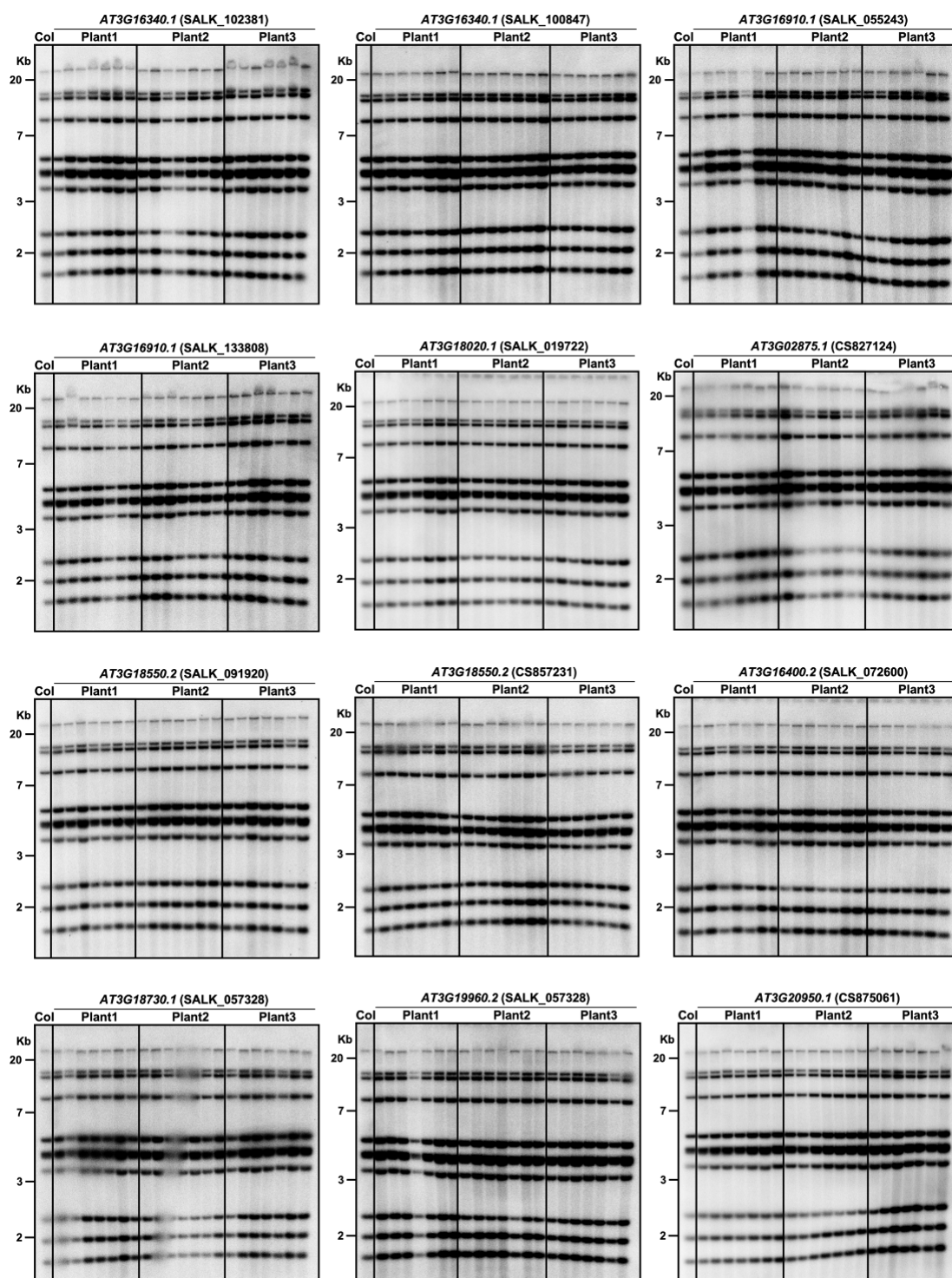


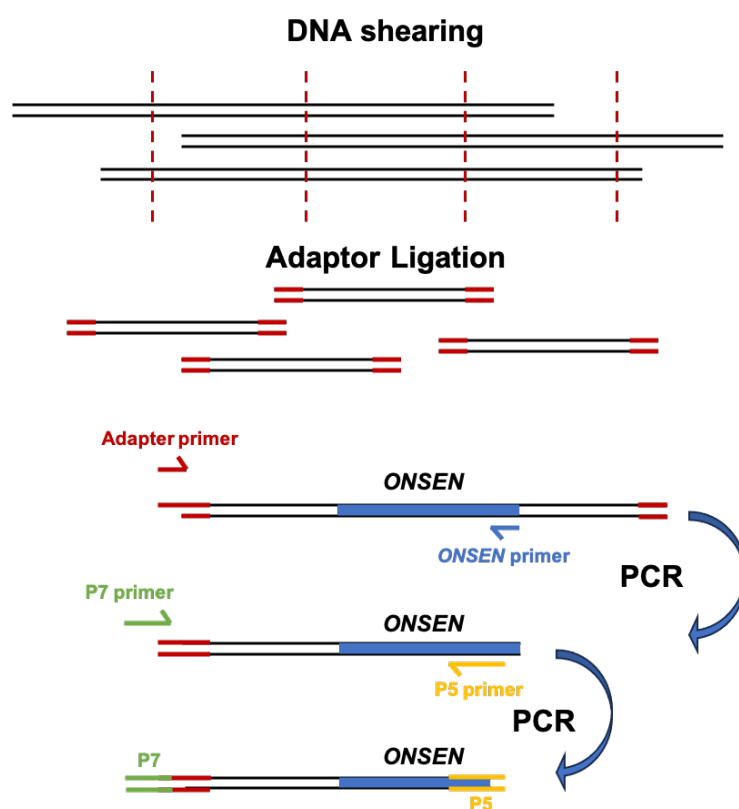
Figure 28: Southern Blot Analysis of *ONSEN* in the Progeny of Heat-Stressed SALK Lines (Mutant List Shown in Table 3).

Seven individuals from each of the three parent plants subjected to HS were analyzed.

The leftmost band in each image represents the Col pattern under non-stress (NS) conditions.

5.2 Overview of TED-seq Methodology

Considering that Southern blotting is not suitable for testing a large number of samples at once, I employed a new method called transposon display sequencing (TED-seq) to explore *ONSEN* transposition. TED-seq is a sequencing technique specifically developed for studying TE and can detect the insertion frequency of specific TEs with high sensitivity (Quadrant L et al., 2021). In TED-seq, genomic DNA is first fragmented by sonication, and adapter sequences are ligated to the ends of these fragments. PCR is then performed using primers based on the adapter and *ONSEN* sequences, which amplified fragments containing *ONSEN* sequences. This resulting product is sequenced allowing the determination the *ONSEN* insertions in the genome (Fig S20).



Supplementary Figure 20: Schematic Diagram of the TED-Seq Working Principle.

DNA was fragmented, and adaptor sequences were ligated to the ends of these fragments. The fragments were then selectively amplified by PCR to enrich for those containing the *ONSEN* sequence. The resulting PCR products were sequenced and mapped against the reference genome to identify the insertion sites of *ONSEN*.

5.3 *At3g16340.1* and *At3g23200.1* Identified as Candidate Genes for *Screen 1* Based on TED-seq Results

TED-seq experiments were conducted on all SALK mutants of the *screen 1* candidate genes within the 5.3-8.9 Mbp range, as well as on *screen 1* (Table 3). Six plants were subjected to HS treatment for each mutant. After HS, DNA was extracted from 1,000 randomly selected progeny of these six plants and used for TED-seq. The results showed 957, 541, and 930 *ONSEN* transpositions in three *screen 1* individuals, with transposition frequencies of approximately 0.96%, 0.54%, and 0.93%, respectively (Fig 29). This suggests that only about one in 100 *screen 1* individuals may exhibit *ONSEN* transposition under HS. Similarly, the SALK mutant *At3g16340.1* (*SALK_100847*), which is close to *screen 1*, also showed about 1% new *ONSEN* insertions (Fig 29). In addition, two lines of *At3g23200.1* (*SALK_017299* and *CS907248*) exhibited relatively high *ONSEN* transposition frequencies of 0.77% and 0.88%, respectively, compared to other SALK lines (Fig 29). These findings suggest that the absence of *At3g16340.1* and *At3g23200.1* may induce *ONSEN* transposition activity. However, no *ONSEN* transposition was detected in *At3g16340.1* and *At3g23200.1* using Southern blot analysis (Fig 30). Overall, TED-seq results confirm the low *ONSEN* transposition frequency in *screen 1* and suggest that *At3g16340.1* and *At3g23200.1* may be candidate

genes for *screen 1*.

List mutants for TED-seq

	mutant
1	AT3G16340.1(SALK_100847)
2	AT3G16340.1(SALK_102381)
3	AT3G16400.1(SALK_072600)
4	AT3G16770.1 (CS362782)
5	AT3G16770.1 (SALK_054382)
6	AT3G16910.1(SALK_055243)
7	AT3G16910.1(SALK_133808)
8	AT3G18020.1(SALK_019722)
9	AT3G18100 (SALK_009892)
10	AT3G18100 (SALK_038539)
11	AT3G18100.1 (CS834189)
12	AT3G18480.1 (CS854597)
13	AT3G18480.1 (SALK_052760)
14	AT3G18550.2(CS857231)
15	AT3G18550.2(SALK_091920)
16	AT3G18730.1(SALK_057328)
17	AT3G19120 (SALK_059906)
18	AT3G19120 (SALK_132948C)
19	AT3G19120 (SALK_133157)
20	AT3G19120.1 (SALK_119754)
21	AT3G19700.1 (SALK_073260)
22	AT3G19700.1 (SALK_078025)
23	AT3G19960.2(SALK_074538C)
24	AT3G20950.1(CS875061)
25	AT3G22170.1 (SALK_002711C)
26	AT3G23200.1 (CS907248)
27	AT3G23200.1 (SALK_017299)
28	AT3G24480.1 (SALK_040784)
29	AT3G24530.1 (SALK_016578)
30	screen1
31	screen1
32	screen1

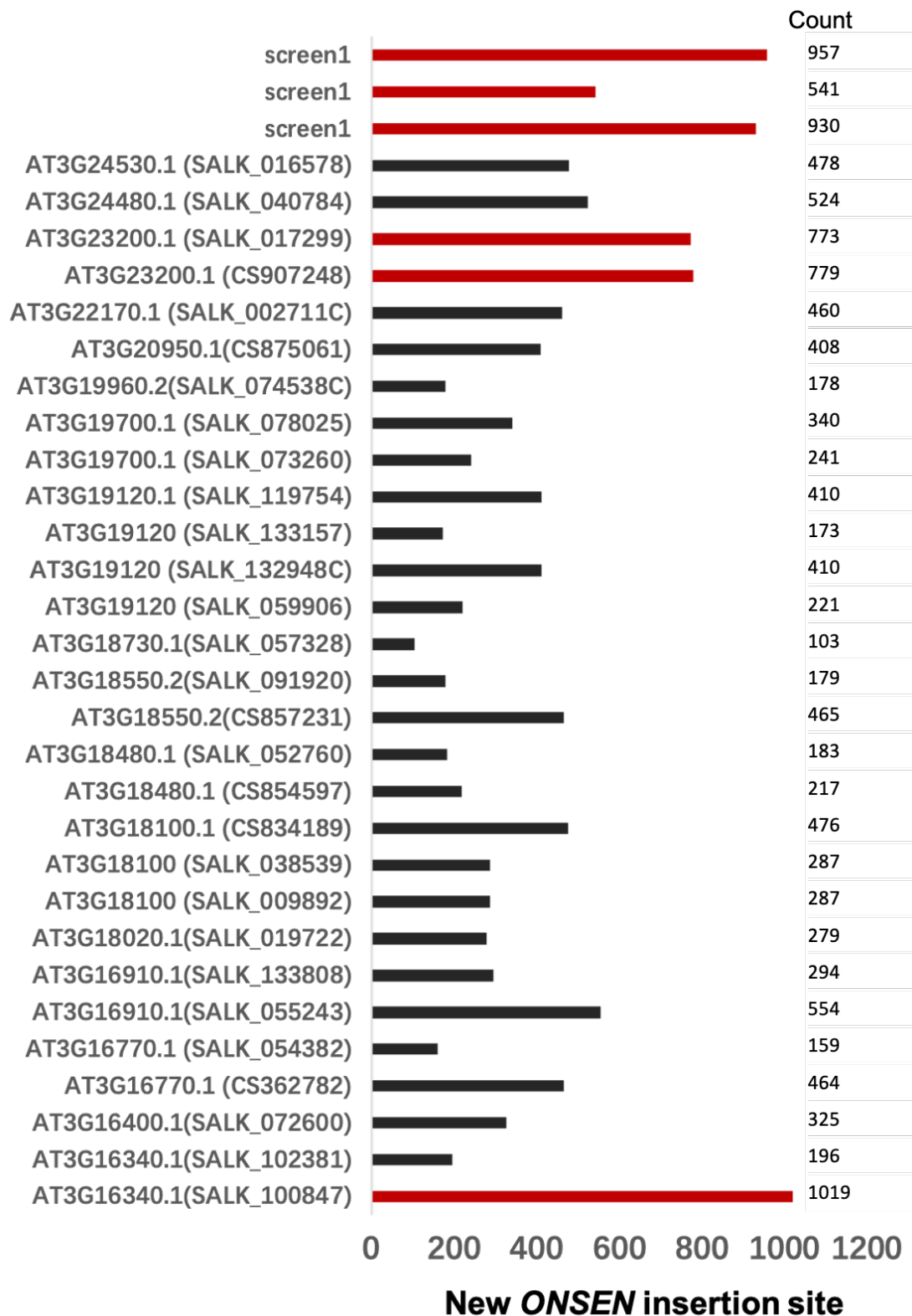


Table 3: List of mutants used for TED-seq analysis.

Figure 29: TED-Seq Results Revealed New Insertions of *ONSEN*.

The bar chart displays the new insertion sites of *ONSEN* found in *screen 1* and all *screen*

I-related SALK lines (mutant list shown in Table 4) within the 5.3-8.9 Mbp region on chromosome 3. One thousand individuals from each of the six parent plants subjected to heat stress were analyzed.

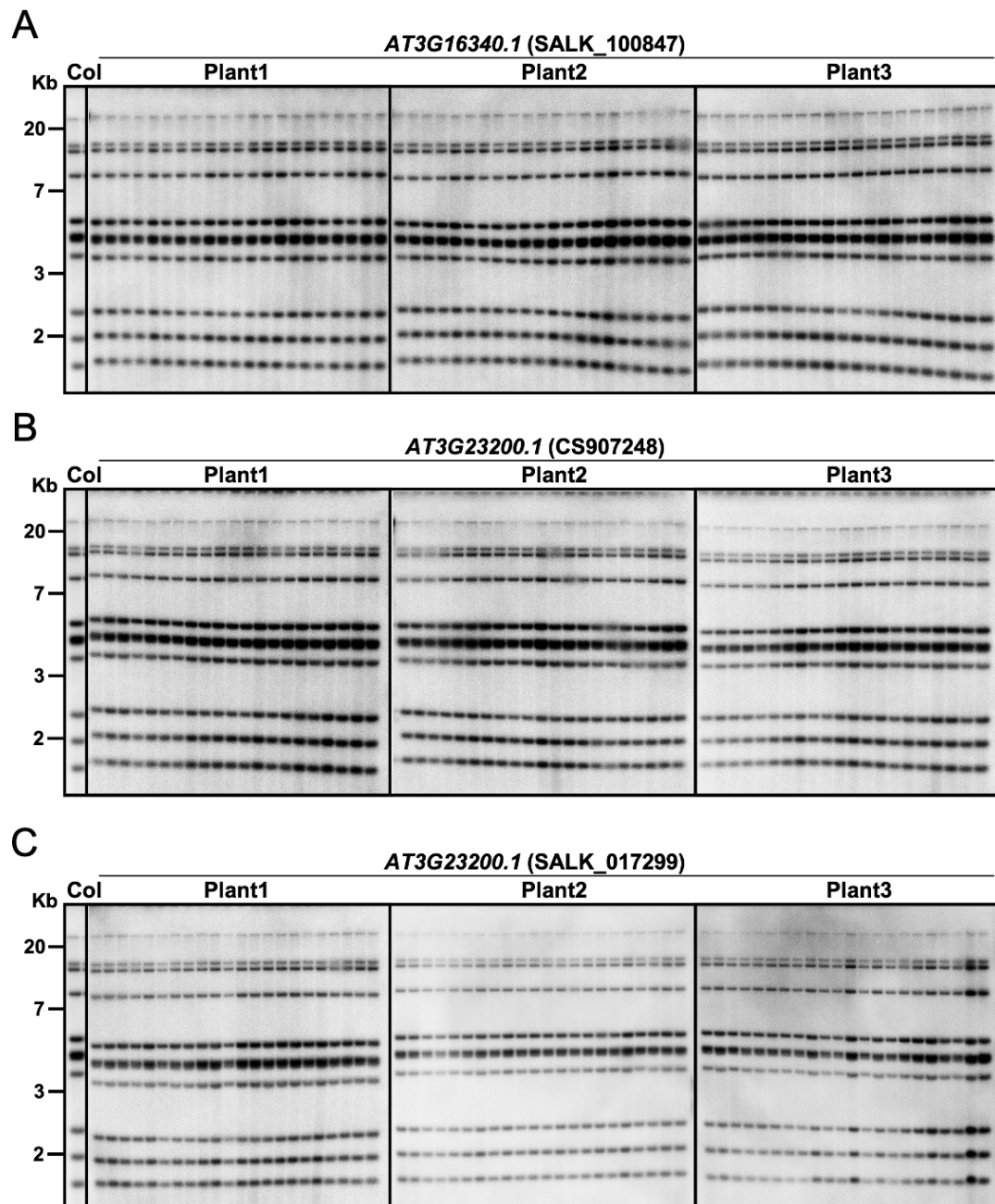


Figure 30: Investigation of *ONSEN* Transpositions in *At3g16340.1* and

At3g23200.1.

The next generation of *At3g16340.1* (SALK_100847), *At3g23200.1* (SALK_017299), and *At3g23200.1* (CS907248) after 48 hours of heat stress (HS) was analyzed by Southern blot to examine the transgenerational transposition of *ONSEN*. The leftmost panel in each figure shows Col under non-stress (NS) conditions. Twenty-one individuals from each of the three parent plants subjected to 48 hours of HS were analyzed.

Discussion

Since there is no direct correlation between *ONSEN* transcription levels and transposition frequency, it is suggested that unknown factors might play a crucial role in regulating *ONSEN*. In this chapter, seven stable mutants that disrupt *ONSEN* transposition silencing were identified from several unknown mutants obtained through EMS mutagenesis using GFP fluorescence and Southern blot transposition tests (Fig 20B, C). Previous research, as well as the first chapter, have thoroughly described the indispensable role of the RdDM pathway in regulating *ONSEN* transcription and transposition (Ito H et al., 2011; Hayashi Y et al., 2020). There are also examples of EMS-induced mutants with high *ONSEN* transcription. For instance, Takehira K et al. (2021) reported that EMS-induced mutations in DRD1-6 led to increased *ONSEN* transcription levels. Since individuals with high transcription and transposition frequencies are likely to be RdDM mutants, I focused on the *screen 1* mutant among the seven potentially related to *ONSEN* transposition regulation (Fig 20B, C). In *screen 1*, new *ONSEN* insertions were observed, even though *ONSEN*'s transcription level and copy number were equivalent to those in Col (Fig 20C, Fig 21). *Screen 1* is likely

involved in directly regulating *ONSEN* transposition, bypassing transcriptional control.

Based on the mapping results of *screen 1*, several SALK mutants of candidate genes were ordered from TAIR. When selecting SALK mutants, efforts were made to ensure that the T-DNA insertion sites were as close as possible to the positions of the identified SNPs. In total, 32 SALK mutants were ordered, many of which had not been previously studied. Due to failure to germinate, unsuccessful T-DNA insertion, or not leading to gene loss-of-function, only 20 SALK mutants were available for analysis (Table 2). The 3.8-9.7 Mbp region on chromosome 3, where *screen 1* is located, contains many genes that still need elucidation (Fig 23). Additionally, the low transposition frequency of *ONSEN* in *screen 1* presents a challenge for analyzing *ONSEN* transposition (Fig 20C, Fig 29). Therefore, although no *ONSEN* insertions were observed in these 20 SALK mutants, the possibility that the gene responsible for *screen 1* is among the selected candidate genes cannot be entirely ruled out.

During cell division, homologous chromosomes undergo crossover, which can lead to the exchange of genetic material in the same region. This process allows classical mapping to refine and narrow down the region for *screen 1*. After approximately 4-5 rounds of backcrossing, *ONSEN* transposition was observed in four NILs (Fig 26). At this point, the *screen 1* gene is expected to be located between 3.8 Mbp and 9.7 Mbp. However, classical mapping has limitations, including the need for large sample sizes and precise phenotypic data, with over 100 plants being screened in this study. Classical mapping is relatively less precise and efficient than modern high-throughput sequencing technologies. Re-sequencing results indicated that the region containing *screen 1* was narrowed down to 5.3–8.9 Mbp. Unfortunately, I still did not identify *screen 1*-responsive mutants among the new SALK mutants selected from the re-sequencing data (Table 3). The low frequency of *ONSEN* transposition was observed in

screen 1. Additionally, it cannot be ruled out that some SALK mutants may lack mutations. Therefore, these results do not exclude the possibility that some of the selected SALK mutants may contain mutations in candidate genes related to *screen 1*.

Quadrana L et al. (2021) reported that TED-seq can detect TE insertions with frequencies as low as 1 in 1,000 in DNA samples. I conducted TED-seq experiments on all selected SALK mutants within the 5.3–8.9 Mbp regions of *screen 1* (Table 4). As anticipated, the *ONSEN* transposition frequency in *screen 1* was approximately 1%, indicating that the mutation associated with *SCREEN 1* only partially relieved *ONSEN*'s transposition silencing (Fig 29). This also explains why new *ONSEN* insertions were not observed in all SALK mutants. However, it is also possible that none of these SALK mutants represent the responsible gene for *screen 1*, and therefore, the transposition silencing of *ONSEN* was not thoroughly relieved.

From TAIR, it is known that *At3g16340.1* is a protein-coding gene encoding a p-coumaric alcohol exporter involved in lignin biosynthesis, while *At3g23200.1* belongs to the uncharacterized protein family UPF0497. In the TED-seq results, *ONSEN* transposition frequencies in *At3g16340.1* and *At3g23200.1* were similar to those in *screen 1* (Fig 29), suggesting that these genes are likely candidates for the responsible gene of *screen 1*. However, since Southern blot analysis showed no *ONSEN* transposition in *At3g16340.1* and *At3g23200.1*, a definitive conclusion about the responsible gene for *screen 1* cannot yet be made (Fig 30). Additionally, *At3g16240.1* and *At3g22555.1* could not be explored due to the lack of available SALK mutants during the second selection (Table 3). Overall, the investigation into *screen 1* is ongoing, with future research focusing on *At3g16340.1*, *At3g23200.1*, *At3g16240.1*, and *At3g22555.1*. If the current research approach does not fully identify the responsible gene for *screen 1*, it may be necessary to explore multi-gene regulation or

non-gene regulation. Should the responsible gene be successfully identified, its function will be elucidated, and the mechanism of *ONSEN* transposition regulation will be clarified. If the *screen 1* gene encodes a known protein, co-immunoprecipitation (Co-IP) will identify interacting factors, determine which factors interact with the protein, and how it inhibits *ONSEN* transposition. If the gene is unknown, crystallographic techniques will be employed to predict and identify the protein's function based on nucleotide or nucleic acid sequence patterns. This research aims to reveal how *screen 1* participates in the regulation of *ONSEN* transposition.

General Discussion

This paper offers a comprehensive examination of TE regulation from both epigenetic and unknown pathway perspectives, with a focus on the roles of the RdDM pathway and histone modifications in regulating *ONSEN*. I discussed the crucial roles of DRD1-6 and SUVH2 in RdDM activity and explored the effects of DNA methylation levels and siRNA levels impact *ONSEN* transcription and transposition regulation. My findings underscore the essential role of siRNA in silencing TE transposition, as plants generate new siRNAs to suppress TE activity in response to environmental stress. Additionally, this study emphasizes the need to distinguish between transcriptional and transpositional regulation when investigating TE regulation pathways. The research on histone modifications suggests that HDA19 may play a positive role in regulating TE transcription in response to environmental stress. The absence of HDA19 leads to increased DNA methylation, histone acetylation, and methylation in several TEs, including *ONSEN*, offering new insights into the role of HDACs in regulating TE activity. In the final chapter, I identified an unknown factor involved in *ONSEN* transposition regulation, termed *screen 1*, and explored its genomic location. While the specific genes responsible for *screen 1* remain unidentified, future research will

continue to focus on this discovery. Identifying *screen 1* holds the potential for groundbreaking insights into the mechanisms governing TE transposition. This study contributes novel perspectives and research methodologies for understanding TE regulation in response to environmental stress.

Methods

Growing conditions

The plants were grown either on 0.5× Murashige and Skoog (MS) basal medium or in a potting mix consisting of soil and vermiculite in a 7:1 ratio. The growth environment was set to continuous light at 21 °C, with a pre-sowing cold treatment of 2 days at 4 °C in the dark. Prior to sowing on the MS medium, seeds were sterilized using a 1:1 solution of sodium hypochlorite and 0.04% Triton X-100.

Inhibitor medium

TSA: After autoclaving 300 ml of MS medium at 120°C for 20 minutes, the medium was allowed to cool to approximately 60°C. Then, 6 µl of 50 mM Trichostatin A (Tokyo Chemical Industry, T2477) was added. The medium was aliquoted into 29 ml portions and used to prepare solid media.

5-Azac: Solid MS medium was aliquoted into 32 ml portions per Petri dish, and 150 ml of liquid MS medium was prepared. In 100 ml of this liquid medium, 1 mg of 5-azaCdR (Tokyo Chemical Industry, A2232) was dissolved to create a 10 µg/ml solution, which was filter-sterilized. The sterilized 10 µg/ml 5-azaCdR solution was mixed 1:1 with the MS liquid medium to achieve the final concentration to 5 µg/ml. This 5 µg/ml 5-azaCdR liquid medium was applied in 8 ml portions to the MS solid medium, resulting in a final concentration of 1 µg/ml on the solid medium. The Petri dishes were sealed with tape and wrapped in plastic, and incubated at 4°C for 15 hours. After incubation, the liquid medium was removed from the surface of the solid medium.

Heat-stress treatment

Seven-day-old seedlings, grown under non-stress conditions (continuous light at 21°C),

were subjected to heat stress at 37°C for either 24 or 48 hours. After the HS treatment, seedlings were sampled immediately for analysis. For Southern blot analysis, seedlings were allowed to recover at 21°C, then potted in potting media approximately two weeks later. For Northern blot analysis, treated seedlings were allowed to recover at 21°C for a designated recovery period before being sampled directly for further analysis.

Quantitative analysis

For quantitative analysis, quantitative PCR (qPCR) was conducted. DNA extraction was performed using the Nucleon PhytoPure DNA extraction kit (Cytiva, Tokyo, Japan). For quantitative reverse transcription PCR (RT-qPCR), RNA was extracted using TRI Reagent (Sigma-Aldrich, St. Louis, MI, USA; T9424). The isolated total RNA was treated with RQ1 RNase-Free DNase (Promega, Madison, WI, USA; M6101) to eliminate any residual DNA. Complementary DNA (cDNA) was synthesized from the RNA using the ReverTra Ace® qPCR RT Kit (TOYOBO, Osaka, Japan). qPCR was performed on an Applied Biosystems 7300 Real-Time PCR System using Luna Universal qPCR Master Mix (New England BioLabs, Ipswich, MA, USA). Gene expression levels were calculated using the Ct method, with 18S rRNA gene expression used as an internal control.

Southern blot analysis

DNA was extracted from 3-week-old whole plant tissue using the Nucleon PhytoPure DNA Extraction Kit (Cytiva, Tokyo, Japan). A total of 2.4 µg of extracted DNA was treated overnight with the restriction enzyme EcoRV at 37°C and subsequently purified by ethanol precipitation. The DNA was separated on a 1% agarose gel for 24 hours at 20 V and transferred overnight onto a Biotodyne B Nylon Membrane (Thermo Fisher Scientific, Waltham, Massachusetts, USA). A 577 bp fragment of the predicted *ONSEN*

open reading frame (ORF) was used as the probe, which was generated by PCR using TaKaRa Ex Taq (TaKaRa, Kusatsu, Japan). The PCR primers used are listed in Table 5. Hybridization signals were detected using a high SDS hybridization buffer with radioisotope-labeled probes, following the Megaprime DNA Labeling System (Cytiva, Tokyo, Japan).

DNA methylation analysis

DNA for bisulfite sequencing was extracted from 7-day-old seedlings using the Nucleon PhytoPure DNA extraction kit (Cytiva, Tokyo, Japan). The extracted DNA was then subjected to bisulfite conversion and desulfation using the MethylCode Bisulfite Conversion Kit (Invitrogen, Waltham, Massachusetts, USA). Target DNA fragments were amplified by PCR, with the following conditions: 40 cycles of 98°C for 10 seconds, 55°C for 30 seconds, and 72°C for 1.5 minutes. PCR amplification was performed using the specific primers listed in Table 5. The resulting PCR products were cloned, and 24 individual clones were sequenced. The sequencing data were analyzed using MEGA-X software, and methylation levels were assessed using the Cymate website.

Formaldehyde-Assisted Isolation of Regulatory Elements (FAIRE)

Seven-day-old seedlings (0.3 g) were cross-linked in a 1% formaldehyde solution (Thermo Fisher Scientific, Waltham, Massachusetts, USA) under vacuum for 10 minutes (2 minutes initially, followed by 8 minutes). Cross-linking was stopped by adding glycine to a final concentration of 0.125 M and incubating for 5 minutes under vacuum. The samples were then ground into a fine powder using liquid nitrogen, and separation buffer 1 (0.4 M sucrose, 10 mM Tris-HCl, pH 8, 10 mM MgCl₂, 5 mM β -mercaptoethanol, and protease inhibitor [Roche, Basel, Switzerland]) was added. The mixture was filtered through Miracloth (Millipore, Burlington, Massachusetts, USA) to

remove cell debris. After centrifugation at 2000 g for 20 minutes at 4°C, the resulting nuclei pellets were resuspended in isolation buffer 2 (0.25 M sucrose, 10 mM Tris-HCl, pH 8, 10 mM MgCl₂, 1% Triton X-100, 5 mM β-mercaptoethanol, and protease inhibitor [Roche, Basel, Switzerland]). Stratification was maintained by adding isolation buffer 3 (0.25 M sucrose, 10 mM Tris-HCl, pH 8, 10 mM MgCl₂, 1% Triton X-100, 5 mM β-mercaptoethanol, and a protease inhibitor [Roche, Basel, Switzerland]), and the mixture was centrifuged at 13,000 rpm for 60 minutes at 4°C. Finally, the precipitates were resuspended in 0.3 mL of SDS nuclear lysis buffer (50 mM Tris-HCl, pH 8, 10 mM EDTA, 1% SDS, and protease inhibitor cocktail).

Chromatin was sheared using a Covaris M220 (Covaris, Woburn, Massachusetts, USA) for 5 minutes under the following settings: 75.0 watts, 10.0% duty factor, and 200 cycles per burst. The resulting suspensions were centrifuged at 13,000 rpm for 10 minutes at 4°C, and the supernatant was collected. DNA was purified with the phenol-chloroform method. DNA fragment abundance, relative to the input DNA, was quantified by qPCR, using the appropriate primers listed in Supplementary Table 1. For input DNA, 2 μL of supernatant was mixed with 198 μL of extraction buffer (10 mM Tris-HCl, pH 8, 0.3 M NaCl, 5 mM EDTA, 0.5% SDS) and incubated overnight at 65°C to reverse formaldehyde cross-linking. The de-crosslinked DNA was treated with RNase at 37°C for 30 minutes and subsequently with proteinase K at 45°C for 1 hour. After treatment, the DNA was purified using the phenol-chloroform method. The input DNA served as an internal control for quantifying the non-crosslinked fraction of the condensed genomic region (NDR).

Cytology

The leaves of 7-day-old plants were fixed overnight in a 3:1 solution of ethanol and

acetic acid. Following fixation, the leaves were treated with enzymes, including boehmotoxin (Yakult Pharmaceutical Industries), pectinase (Kyowa Chemical Products, Tokyo, Japan), and cytochrome enzymes (Sigma-Aldrich, St. Louis, MO, USA) at 1% (v/v) in citrate buffer, and incubated at 37°C for 2 hours. The treated leaves were then placed on a slide and gently torn with a needle on a heating plate at 45°C for 30 seconds to disperse the tissue in 45% acetic acid. After the acetic acid evaporated, a fresh fixative (ethanol and acetic acid, 3:1) was added, and the slides were air-dried. For staining, 4',6-diamidino-2-phenylindole (DAPI) (Vector Laboratories, Newark, CA, USA) was applied to visualize the nuclei. The cells in the images captured by fluorescence microscopy were classified into two states: condensed and dispersed.

Northern blot analysis

Total RNA was extracted using TRI reagent (Sigma-Aldrich, St. Louis, Missouri, USA). To separate small-molecular-weight RNA, 20% PEG (#6000) was added to the total RNA, and the mixture was incubated on ice for 1 hour. Following incubation, the mixture was centrifuged at 12,000 rpm for 10 minutes at 4°C, and the supernatant was collected. NaOAc and 100% ethanol were added to the supernatant to precipitate low-molecular-weight RNA. The siRNA was purified by ethanol precipitation and dissolved in 100% formamide. RNA samples (1400 ng) were denatured before electrophoresis on a 15% PAGE gel in 0.5× TBE buffer. Electrophoresis was performed initially at 50 V for 20 minutes to synchronize dye migration, followed by 200 V at 500 mA for 3 hours. The RNA was then transferred onto a Hybond-N⁺ hybridization membrane (Sigma-Aldrich, St. Louis, Missouri, USA) overnight and hybridized with a DIG-labeled RNA probe at 40°C overnight. PCR products for the Northern probe were generated using TaKaRa Ex Taq (TaKaRa, Kusatsu, Japan) (94°C for 30 seconds, 55°C for 30 seconds, 72°C for 1 minute, for 30 cycles). PCR primers for the probe were generated using

TaKaRa Ex Taq (TaKaRa, Kusatsu, Japan) with the following thermal profile: 94°C for 30 seconds, 55°C for 30 seconds, and 72°C for 1 minute over 30 cycles. The primers used for probe generation are listed in Table 1. DIG-labeled RNA probes were synthesized with the DIG RNA Labeling Mix (Roche, Basel, Switzerland) using synthetic oligonucleotides containing the T7 RNA polymerase promoter and the MRGA shortscript kit (Thermo Fisher Scientific, Waltham, MA, USA). Hybridization signals were detected using anti-DIG antibodies (Roche, Basel, Switzerland) and CDP-Star (Roche, Basel, Switzerland), and visualized with the LAS3000 imaging system.

RNA sequencing (RNA-seq)

Total RNA was extracted using TRI reagent (Sigma-Aldrich, St. Louis, Missouri, USA). For RNA-sequencing, library preparation and sequencing were carried out at the OIST Sequencing Center (SQC). PolyA mRNA purification was performed using the NEBNext® Poly(A) mRNA Magnetic Isolation Module, and 150-bp paired-end (150 PE) library was prepared with NEBNext® Ultra™ II Directional RNA Library Prep Kit for Illumina, following manufacturers' instructions. Sequencing was conducted on a NovaSeq6000 system (SP flowcell).

Whole Genome Bisulfite Sequencing (WGBS)

DNA for bisulfite sequencing was extracted from 7-day-old seedlings using Nucleon PhytoPure DNA extraction kit (Cytiva, Tokyo, Japan). Bisulfite conversion was carried out using the EZ DNA Methylation-Gold Kit (ZYMO RESEARCH, Irvine, CA, USA). Library preparation was performed with the Accel-NGS Methyl-Seq DNA Library Kit (Swift BIOSCIENCES, MI, USA). Sequencing was conducted on the NovaSeq 6000 platform with 150 bp paired-end (150 PE) conditions.

Chromatin Immunoprecipitation (ChIP)

Seven-day-old seedlings (2 g) were cross-linked in a 1% formaldehyde solution (Thermo Fisher Scientific, Waltham, MA, USA) under vacuum for 10 minutes (2 minutes + 8 minutes). Cross-linking was halted by adding glycine to a final concentration of 0.125 M and incubating for 5 minutes under vacuum. Samples were ground into a powder using liquid nitrogen, and separation buffer 1 (0.4 M sucrose, 10 mM Tris-HCl, pH 8, 10 mM MgCl₂, 5 mM β -mercaptoethanol, and protease inhibitor [Roche, Basel, Switzerland]) was added. The mixture was filtered through Miracloth (Millipore, Burlington, Massachusetts, USA) to remove debris. After centrifugation at 2000 g for 20 minutes at 4°C, the nuclei pellets were resuspended in isolation buffer 2 (0.25 M sucrose, 10 mM Tris-HCl, pH 8, 10 mM MgCl₂, 1% Triton X-100, 5 mM β -mercaptoethanol, and protease inhibitor [Roche, Basel, Switzerland]). The pellets were centrifuged at 13,000 rpm for 20 minutes at 4°C, and the nuclei were resuspended again in isolation buffer 2. Buffer 3 (0.25 M sucrose, 10 mM Tris-HCl, pH 8, 10 mM MgCl₂, 1% Triton X-100, 5 mM β -mercaptoethanol, and protease inhibitor [Roche, Basel, Switzerland]) was added to maintain stratification. The mixture was centrifuged at 13,000 rpm for 60 minutes at 4°C. The precipitates were then resuspended in 0.3 mL of SDS nuclear lysis buffer (50 mM Tris-HCl, pH 8, 10 mM EDTA, 1% SDS, and protease inhibitor cocktail). Chromatin was sheared using a Covaris M220 (Covaris, Woburn, Massachusetts, USA) for 5 minutes (settings: 75 watts, 10% duty factor, 200 cycles). Suspensions were centrifuged at 13,000 rpm for 10 minutes at 4°C to collect the supernatant. Five histone modification-related antibodies were used in the experiment: Anti-H3 antibody (ab1791, abcam), Anti-acetylated H3 antibody (MAB10010-100, 21016), anti-H3K9me₂ (ab1220, abcam), and anti-H3K4me₃ (07-473, merck). Each chromatin solution was incubated with the respective antibody at 4°C overnight with constant rotation. Set aside 10% of the IP input and store it at -20°C

until reverse cross-linking is performed. Each antibody was preincubated with 30µl of Dynabeads Protein G (Invitrogen, DB10004) for 2 hours. The chromatin-antibody complexes were washed sequentially with low-salt buffer (150 mM NaCl, 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl, pH 8), high-salt buffer (500 mM NaCl, 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl, pH 8), LiCl buffer (0.25 M LiCl, 1% Nonidet P-40, 1% sodium deoxycholate, 1 mM EDTA, 10 mM Tris-HCl, pH 8), and TE buffer (10 mM Tris-HCl, pH 8, 1 mM EDTA). The DNA-protein complexes were eluted with Direct Elution buffer (10 mM Tris-HCl, pH 8, 5 mM EDTA, 0.3 M NaCl, 0.5% SDS). Input samples were adjusted to the same volume using Direct Elution buffer. Cross-linking was reversed by incubating the samples overnight at 65°C. The DNA was treated with RNase (37°C for 30 minutes) and proteinase K (45°C for 1 hour) before being purified via phenol-chloroform extraction. For ChIP-qPCR, the purified DNA was analyzed using qPCR on an Applied Biosystems 7300 Real-Time PCR System with Luna Universal qPCR Master Mix (New England BioLabs, Ipswich, MA, USA). For ChIP-seq, purified DNA was prepared for library construction, followed by sequencing on the HiSeq X Ten platform (Macrogen, Japan).

Next-Generation Sequencing (NGS)

To identify the causative gene for screen 1, I performed a cross between screen 1 and Col to generate F1 heterozygous plants. Self-pollination of the F1 generation yielded an F2 segregating population. After subjecting the F2 plants to heat stress (HS), self-pollination was carried out to obtain the F3 generation. Southern blot analysis was employed to investigate the *ONSEN* transposition in the F3 generation, serving as a transgenerational transposition analysis for the F2 generation. F2 individuals displaying intergenerational *ONSEN* transposition were selected for further analysis. DNA was extracted from a pooled sample of 40 selected F2 individuals using the Nucleon

PhytoPure DNA Extraction Kit (Cytiva, Tokyo, Japan). Library construction was performed using the NEBNext Ultra DNA Library Prep Kit for Illumina (E7645), and sequencing was carried out on the NextSeq500 platform.

Re-sequencing (Whole Genome Resequencing; WGRS)

DNA was extracted from 7-day-old seedlings using the Nucleon PhytoPure DNA extraction kit (Cytiva, Tokyo, Japan). The extracted DNA was sent to BGI (China) for library construction and sequencing. Libraries were sequenced on the DNBseq™ platform under 150 paired-end (150PE) conditions.

Transposon Display Sequencing (TED-seq)

DNA was extracted from 1,000 seven-day-old seedlings using the Nucleon PhytoPure DNA extraction kit (Cytiva, Tokyo, Japan). To obtain DNA fragments ranging from 200-600 bp, 1000 ng of DNA was processed in a Covaris M220 sonicator for 175 seconds (settings: watts 75.0, percent 18.6, count 200). End repair and A-tailing were performed on the DNA fragments using NEBNext Ultra II End Prep Enzyme Mix (NEB, E7645), followed by ligation with TED-seq forked-adapters (final concentration of 0.8 μ M), NEBNext Ultra II Ligation Master Mix, and NEBNext Ligation Enhancer. Size selection of adapter-ligated DNA was performed using AMPure XP Beads (0.9X, Beckman Coulter, A63880). The selected DNA was used as a template for the selective amplification of TE-containing fragments using NEBNext Ultra II Q5 Master Mix PCR and primers TEDseq_outter_for and TE_outter_rev. PCR conditions were as follows: 98°C for 30 seconds; 20 cycles of 98°C for 10 seconds and 61°C for 75 seconds; with a final extension at 61°C for 5 minutes. The amplified DNA was subjected to a second round of PCR using nested primers containing custom Illumina P5 and barcoded P7 adapters (P7_Primer_index#, P5_TE_primer). The second PCR conditions were: 98°C

for 30 seconds, 6 cycles of 98°C for 10 seconds and 61°C for 75 seconds, followed by 10 cycles of 98°C for 10 seconds and 72°C for 75 seconds, with a final extension at 72°C for 5 minutes. Finally, the DNA library was purified, size-selected again with AMPure XP Beads (0.9X), and resuspended in 30 µl of 10 mM TrisHCl. Sequencing of the libraries was performed using the Hiseq X Ten (Macrogen, Japan).

Sequence data analysis

For RNA sequencing, low-quality reads were filtered using Trimmomatic-0.39 with the following parameters: TRAILING:20 SLIDINGWINDOW:4:20 MINLEN:36. The filtered reads were then mapped to the Arabidopsis reference genome (TAIR10) using HISAT2-2.2.1. Gene transcription levels were quantified using featureCounts (subread-2.0.3). Differentially expressed genes were identified in RStudio using the edgeR 3.20.9 package, and log2FoldChange (log2FC) and p-values were extracted. Data normalization was performed using counts per million (CPM), and visualization was performed using the ggplot2 package in RStudio.

For WGBS: Low-quality reads were trimmed by Trimmomatic-0.39 with the following parameters: HEADCROP:10 SLIDINGWINDOW:4:20 MINLEN:25. The filtered reads were mapped to the Arabidopsis reference genome (TAIR10) using BSMAP-2.90 (parameters: -w 100 -v 2 -n 1 -r 1). Methylation levels of cytosines were calculated using methratio.py in BSMAP-2.90. The MethyKit package in RStudio was used to detect differentially methylated regions (DMRs; q-value < 0.01), with the threshold for methylation differences set to 40% for mCG, 20% for mCHG, and 10% for mCHH. BEDtools (v2.21.1) was used to identify overlapping regions between the DMRs list and *ONSEN* or TE. Metaplots were generated using DeepTools (v3.5.1), with sequencing depth calculated for every 100-bp consecutive bin across the entire genome.

For ChIP-seq: Low-quality reads were filtered using Trimmomatic-0.39 with the following parameters: LEADING:3 TRAILING:3 SLIDINGWINDOW:4:20 MINLEN:30. The filtered data were mapped to the *Arabidopsis* TAIR10 genome using Bowtie2 (v2.5.1). Data from three replicates were merged using samtools-1.16.1. The accumulation levels of H3 were used to normalize the levels of H3Ac, H3K4me3, and H3K9me2 using the bamCompare tool from DeepTools (v3.5.1). The parameters for converting to bigWig format were --normalizeUsing CPM, --binSize 20, and --extendReads 200. Metaplots were generated using computeMatrix and plotHeatmap in DeepTools (v3.5.1) for every 100-bp consecutive bin.

For NGS: Use Mitsucal for result analysis. When searching for causative genes in screen 1, set the conditions as follows: Replacement number: blank, Replacement rate: 70-100, QV replacement rate: 70-100, Control replacement number: blank, only amino acid sequence mutations: checked, Nucleotide replacement type: blank.

For Re-seq (WGRS): Low-quality reads were filtered using Trimmomatic-0.39 with the following parameters: LEADING:20 TRAILING:20 SLIDINGWINDOW:4:20 MINLEN:36. The filtered data were mapped to the *Arabidopsis* TAIR10 genome using BWA-0.7.17. After removing duplicate sequences with GATK-4.2.6.0, variant detection was performed using GVCFS. SNP information obtained from SelectVariants of GATK-4.2.6.0 was then filtered using VariantFiltration with the following criteria: -filter-expression "QD < 2.0" --filter-name "LowQD" --filter-expression "MQ < 40.0" --filter-name "MQ40.0" --filter-expression "FS > 60.0" --filter-name "FS60.0" --filter-expression "SOR > 3.0" --filter-name "SOR3.0". Bcftools-1.15.1 was used to extract SNP information from a specific region (5.3-8.9 Mbp) on chromosome 3 where screen 1 is located. This SNP information was then annotated using ANNOVAR.

For TED-seq: TED-seq data were analyzed using shell scripts from Professor Leandro Quadrana at Université Paris-Saclay (Quadrana L et al., 2021). All reads were mapped to the 100bp extremity of *ONSEN* using bowtie2-2.5.1. Paired-end reads with only one mate mapped were extracted using FilterSamReads.jar from picard-tools-1.119, and reads with discordant mappings were selected. These discordant reads were then re-mapped to the reference genome (TAIR10) using bowtie2-2.5.1. Nearby informative reads were clustered using bedtools-2.25.0, and these clusters of reads were used to identify insertions.

Primer List

Table 4: List of primers used in this study.

Experiment	Name	Sequence (5'-3')
qRT-PCR, qPCR	COPIA78-4129F_RT	CCACAAGAGGAACCAACGAA
	COPIA78-4300R_RT	TTCGATCATGGAAGACCGG
	HsfA2_qPCR_F	TGGGATTCTCATAAGTTCTCAAC A
	HsfA2_qPCR_R	TGGATCAATCTTTCTGAATCCAT
	ROMANIAT5-2_qF	TTGTCACTTGAACGAATGTATTG CACC
	ROMANIAT5-2_qR	GCCAACAAGAATATATGAAGATA ATGC
	GFP_qPCR_F	TGGCTTTGATGCCGTTCTTTTG
	GFP_qPCR_R	CGATTTCAAGGAGGACGGAAAC AT
	18Sr-FW	CGTCCCTGCCCTTTGTACAC
	18Sr-RV	CGAACACTTCACCGGATCATT
Southern blotting	ONSEN-F	TAATGTTCCCTTCCAAGTCCC
	ONSEN-R	GCTTGTAATGACCCAAGAAGT
	ONSEN LTR F	GGTTGAAGGGTCAAAGAGTAAA T
	ONSEN LTR R	CCTCCAAACTACAAAATATCTAA AA
Genotyping	drd1-6-F	AGCTAAGGGATGGAACTAGG
	drd1-6-R	CGAGATGCTCCAACAAGCGAG
	DMS3-1 WT Forward	ACGGGGATCTTCAATCTTGG
	dms3-1 mut Forward	ACGGGGATCTTCAATCTTGA
	DMS3-1 Reverse	TAGTGCATGAACTCTGAGCCA
	RDM1-4 WT-n Forward	ACAACGATGGGATCAATCGC
	rdm1-4 mut-n Forward	ACAACGATGGGATCAATCGT
	RDM1-4-n Reverse	TTAGCAGCCACAGCGAGAAA
	SUVH2(SALK_079574)-F	CCTCAACCTCATGCCGGATT
	SUVH2(SALK_079574)-R	GTTCAACCACAGTCCCCGAT
	SUVH9(SALK_048033)-F	AGCTCTTCTTCCACCAACCC
	SUVH9(SALK_048033)-R	CCGTCACCTTCACTGATCCC

atmorec6-3 F	GGAAAGCTGGAAGCTATAATGATG
atmorec6-3 R	GATGACATCTGCCCCAAGTCTC
SDE4-3021F	TGTTCAATTCAGTTACAAGTCGTG
SDE4-4293R	TTCACCTTTTTTGATCCCTTGATC
nrpe1-11 Forward	TGAGCAGGTGATTACGGAT
nrpe1-11 Reverse	TGCGTTTGTCGTGGATATGAC
LBa1	TGGTTCACGTAGTGGGCCATCG
GABI-KAT LB O8409	ATATTGACCATCATACTCATTGC
hda19-3 F	TTATTTCTATGACCCTGAGGTC
hda19-3 R	AGTCATAACCCTGTCAGTAGC
HDA6W-52	AAATGGAAAGTAGGAGATATTAC
HDA6M-52	AAATGGAAAGTAGGAGATATTAT
5g63110-52	TTGGGTTGCTTCAACTTATCAGTC
DDM1-1-F1	ATTTGCTGATGACCAGGTCCT
DDM1-1-R1	CATAAACCAATCTCATGAGGC
AT3G12020-Fw	GGAGGCAGTCATCTGCACAC
AT3G12020-Rv	TGGGCTGTCATGTTCTTGCC
AT3G12970-Fw	GGCCAACATGGTAAAGCTGTC
AT3G12970-Rv	AATAGAACCCGACCCGAAGC
AT3G14360-Fw	CGACCAAATGACTACTGCAAC
AT3G14360-Rv	CATCAGCATCAAACGGCTCAG
AT3G16770-Fw	CACATTGCTTTGCTCAATACCG
AT3G16770-Rv	TCTCGATAGAGCAAACGGATCA
AT3G18100-Fw	CACTGATGCCGCCAATGTTA
AT3G18100-Rv	CAGAGACCATACAAGCAAACCA
AT3G18480-Fw	ACGCGTTTTTTGGGATTGGTT
AT3G18480-Rv	ACCAATGCAATCTCCCCGAA
SALK_119754-Fw	AATACCTCGCCGGTGTGAGA
SALK_119754-Rv	CGGCTCGGTTAATTTGCATAGG
AT3G19700-Fw	ATGCAAACGTGTTGGTGTGC
AT3G19700-Rv	GGCTAGTGTTTCACCGCTTC
AT3G22170-Fw	TTCTTGCTGGTTTTGCCTGC

	AT3G22170-Rv	GACGTTTGT TTTGCCTGAGACT
	AT3G23200-Fw	TCGAGAAGACGTCCAAGAAGC
	AT3G23200-Rv	TGGGATTGGATCTATACGGTTC
	AT3G24480-Fw	CTTCCCACGTGAAGCCCAC
	AT3G24480-Rv	TATCTGGGAAGATTCCGGCG
	AT3G24530-Fw	ACAACCGTTCTTAGCTGCCA
	AT3G24530-Rv	TGGATCAGTAACATTGGCTACC
	AT3G24890-Fw	CTCAGGCATGTCGCTGGTTT
	AT3G24890-Rv	AAGAGGCAAAGCAAGCTAAAG AC
	pDAPL1	AATGGATAAATAGCCTTGCTTCC
	WiSCL2	ATGTGTGAGTAGTCCCAGATAA G
BS-seq	AT1g11265 ONSEN F	AAGAGATATTTAYTGTGAGAT
	AT3G61330 BS F1	GAYTTATGTAAAAGTGTA AAAA GTTGAATYTG
	At5g13205_BS_F3	AA YATTAGAGGAATAGTATGTT
	COPIA78 BS R2	ATTCTCACATCATCTARCTTCTC TCCRTT
FAIRE-qPCR	FAIRE_LTR_qpcr_F	ACACTTWAACACTTTCTCCATT ACC
	FAIRE_LTR_qpcr_R	YGGGACWTGGAAGGGAACAT
ChIP-qPCR	COPIA78-4129F_RT	CCACAAGAGGAACCAACGAA
	COPIA78-4300R_RT	TTCGATCATGGAAGACCGG
Northern blotting	LTR_northern_F(T7)	TAATACGACTCACTATAGGGCC AAATCCAAGTTCTAGAGT
	LTR_northern_R	TTGCAACATTTGAGAGGGGGA
Classical mapping	NGA126_Fw	CAAGAGCAATATCAAGAGCAG C
	NGA126_Rv	GAAAAAACGCTACTTTCGTGG
	UPSC_3-3806-F	ATGAGGGGGAAAATCAAACC
	UPSC_3-3806-R	AGTGTTGCTGCTTCTTCGT
	LUGSSLP714 F	CGTGTCCATAGATTTAGGTTTGT ATCA
	LUGSSLP714 R	TCGTTTCGTATAGAACAAGATT CGT
	UPSC_3-8894-F	CCTTCGTTGGGTCAAACAGT

	UPSC_3-8894-R	CGGCAAAATGGTGAGAAGTT
	MAPPLANT3.3-F	TCCGTATCCACATCCAAACA
	MAPPLANT3.3-R	GCGGTCCACTCAGTGAATCCG
	LUGSSLP08-F	ATTTGCACTCAAGAAAAACAAA GA
	LUGSSLP08-R	AACGTATAACGTAATGAAAGTA GGTGA
TED-seq	P7_adapter_up	GTGACTGGAGTTCAGACGTGTG TTGAAAGTTAAAC*T
	P7_adapter_bottom	pGTTTAACTTTCAACAdc
	TEDseq_outter_for	GTGACTGGAGTTCAGACGTG
	P7_primer_index	CAAGCAGAAGACGGCATAACGA GATCGATGTGTGACTGGAGTTC AGACGTG
	TE_outter_rev	ATAAGATCTAGATTACATTATT CA
	TED_P5_COPIA78	AATGATACGGCGACCACCGAGA TCTACACTCTTTCCCTACACGAC GCTCTCCGATCTAATTGGATTT GACCATAATT
	TED_P5_COPIA78_1	AATGATACGGCGACCACCGAGA TCTACACTCTTTCCCTACACGAC GCTCTCCGATCTCACTTGAATT TGGCCATAATT
	TED_P5_COPIA78_2	AATGATACGGCGACCACCGAGA TCTACACTCTTTCCCTACACGAC GCTCTCCGATCTGCACTTGGAT TTGACCATAATT
	TED_P5_COPIA78_3	AATGATACGGCGACCACCGAGA TCTACACTCTTTCCCTACACGAC GCTCTCCGATCTTGCACTTAGA TTTGGCCATAATT
	TED_P5_COPIA78_4	AATGATACGGCGACCACCGAGA TCTACACTCTTTCCCTACACGAC GCTCTCCGATCTATGCAATTG GATTTGACCATAATT

References

Abe A, Kosugi S, Yoshida K et al. Genome sequencing reveals agronomically important loci in rice using MutMap. *Nature Biotechnology*, 174-178, 30(2) (2012) doi:10.1038/nbt.2095npg

Akinmusola RY, Wilkins CA, Doughty J. DDM1-Mediated TE Silencing in Plants. *Plants (Basel)*. 2023 Jan 18;12(3):437. Doi: 10.3390/plants12030437.

Amedeo P., Habu, Y., Afsar, K., Mittelsten Scheid, O., and Paszkowski, J. (2000). Disruption of the plant gene MOM releases transcriptional silencing of methylated genes. *Nature* 405(6783), 203-206. doi: 10.1038/35012108.

Arkhipova IR. Neutral Theory, Transposable Elements, and Eukaryotic Genome Evolution. *Mol Biol Evol*. 2018 Jun 1;35(6):1332-1337. doi: 10.1093/molbev/msy083

Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. *Cell Res*. 2011 Mar;21(3):381-95. doi: 10.1038/cr.2011.22.

Belancio VP, Hedges DJ, Deininger P. Mammalian non-LTR retrotransposons: for better or worse, in sickness and in health. *Genome Res*. 2008 Mar;18(3):343-58. doi: 10.1101/gr.5558208.

Bird A. Perceptions of epigenetics. *Nature* 447, 396–398 (2007). <https://doi.org/10.1038/nature05913>

Blevins T, Pontes O, Pikaard CS, Meins F Jr. Heterochromatic siRNAs and DDM1 independently silence aberrant 5S rDNA transcripts in Arabidopsis. *PLoS One*. 2009 Jun 16;4(6):e5932. doi: 10.1371/journal.pone.0005932.

Bourgeois Y, Boissinot S. On the Population Dynamics of Junk: A Review on the Population Genomics of Transposable Elements. *Genes (Basel)*. 2019 May

30;10(6):419. doi: 10.3390/genes10060419.

Brabbs TR, He Z, Hogg K, Kamenski A, Li Y, Paszkiewicz KH, Moore KA, O'Toole P, Graham IA, Jones L. The stochastic silencing phenotype of *Arabidopsis morc6* mutants reveals a role in efficient RNA-directed DNA methylation. *Plant J.* 2013 Sep;75(5):836-46. doi: 10.1111/tpj.12246.

Cadavid IC, Balbinott N, Margis R. Beyond transcription factors: more regulatory layers affecting soybean gene expression under abiotic stress. *Genet Mol Biol.* 2023 Jan 23;46(1 Suppl 1):e20220166. doi: 10.1590/1678-4685-GMB-2022-0166.

Cao X, Jacobsen SE. Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. *Proc Natl Acad Sci U S A.* 2002 Dec 10;99 Suppl 4(Suppl 4):16491-8. doi: 10.1073/pnas.162371599.

Casacuberta E, González J. The impact of transposable elements in environmental adaptation. *Mol Ecol.* 2013 Mar;22(6):1503-17. doi: 10.1111/mec.12170.

Cavrak, V. V., Lettner, N., Jamge, S. Kosarewicz, A., Bayer, M. L., Scheid, M. O. (2014). How a retrotransposon exploits the plant's heat stress response for its activation. *PLOS Genet.* 10, e1004115. doi: 10.1371/journal.pgen.1004115.

Cedar, H., Bergman, Y. Linking DNA methylation and histone modification: patterns and paradigms. *Nat Rev Genet* 10, 295–304 (2009). <https://doi.org/10.1038/nrg2540>

Chen L, Duan L, Sun M, Yang Z, Li H, Hu K, Yang H, Liu L. Current trends and insights on EMS mutagenesis application to studies on plant abiotic stress tolerance and development. *Front Plant Sci.* 2023 Jan 4;13:1052569. doi: 10.3389/fpls.2022.1052569.

Chen LT, Wu K. Role of histone deacetylases HDA6 and HDA19 in ABA and

abiotic stress response. *Plant Signal Behav.* 2010 Oct;5(10):1318-20. doi: 10.4161/psb.5.10.13168.

Christman, J. 5-Azacytidine and 5-aza-2'-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy. *Oncogene* 21, 5483–5495 (2002). <https://doi.org/10.1038/sj.onc.1205699>

Cokus, S., Feng, S., Zhang, X. et al. Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature* 452, 215–219 (2008). <https://doi.org/10.1038/nature06745>

Du, J., Johnson, L., Jacobsen, S. et al. DNA methylation pathways and their crosstalk with histone methylation. *Nat Rev Mol Cell Biol* 16, 519–532 (2015). <https://doi.org/10.1038/nrm4043>

Du J, Zhong X, Bernatavichute YV, Stroud H, Feng S, Caro E, Vashisht AA, Terragni J, Chin HG, Tu A, Hetzel J, Wohlschlegel JA, Pradhan S, Patel DJ, Jacobsen SE. Dual binding of chromomethylase domains to H3K9me₂-containing nucleosomes directs DNA methylation in plants. *Cell*. 2012 Sep 28;151(1):167-80. doi: 10.1016/j.cell.2012.07.034.

Du, Q., Fang, Y., Jiang, J. et al. Characterization of histone deacetylases and their roles in response to abiotic and PAMPs stresses in *Sorghum bicolor*. *BMC Genomics* 23, 28 (2022).

Ebbs ML, Bender J. Locus-specific control of DNA methylation by the *Arabidopsis* SUVH5 histone methyltransferase. *Plant Cell*. 2006 May;18(5):1166-76. doi: 10.1105/tpc.106.041400. Epub 2006 Mar 31. PMID: 16582009; PMCID: PMC1456864.

Ebbs ML, Bartee L, Bender J. H3 lysine 9 methylation is maintained on a transcribed inverted repeat by combined action of SUVH6 and SUVH4 methyltransferases. *Mol Cell Biol.* 2005 Dec;25(23):10507-15. doi: 10.1128/MCB.25.23.10507-10515.2005. PMID: 16287862; PMCID: PMC1291251.

Feschotte C, Pritham EJ. DNA transposons and the evolution of eukaryotic genomes. *Annu Rev Genet.* 2007;41:331-68. doi: 10.1146/annurev.genet.40.110405.090448.

Gao, Z., Liu, H. L., Daxinger, L., et al. (2010). An RNA polymerase II- and AGO4-associated protein acts in RNA-directed DNA methylation. *Nature.* 465, 106–109. doi: 10.1038/nature09025.

Gujral, P., Mahajan, V., Lissaman, A.C. et al. Histone acetylation and the role of histone deacetylases in normal cyclic endometrium. *Reprod Biol Endocrinol* 18, 84 (2020). <https://doi.org/10.1186/s12958-020-00637-5>

Erdmann, R.M., and Picard, C.L. (2020). RNA-directed DNA Methylation. *PLoS Genet* 16(10), e1009034. doi: 10.1371/journal.pgen.1009034.

Ferrafiat, L., Pflieger, D., Singh, J., Thieme, M., Bohrer, M., Himber, C., et al. (2019). The NRPD1 N-terminus contains a Pol IV-specific motif that is critical for genome surveillance in Arabidopsis. *Nucleic Acids Res* 47(17), 9037-9052. doi: 10.1093/nar/gkz618.

Guo Q, Jing Y, Gao Y, Liu Y, Fang X, Lin R. The PIF1/PIF3-MED25-HDA19 transcriptional repression complex regulates phytochrome signaling in Arabidopsis. *New Phytol.* 2023 Nov;240(3):1097-1115. doi: 10.1111/nph.19205.

Gorham SR, Weiner AI, Yamadi M, Krogan NT. HISTONE DEACETYLASE 19

and the flowering time gene *FD* maintain reproductive meristem identity in an age-dependent manner. *J Exp Bot*. 2018 Sep 14;69(20):4757-4771. doi: 10.1093/jxb/ery239.

Grandbastien, M.A., Lucas, H., Morel, J.B., Mhiri, C., Vernhettes, S., and Casacuberta, J.M. (1997). The expression of the tobacco *Tnt1* retrotransposon is linked to plant defense responses. *Genetica* 100(1-3), 241-252.

Grunstein, M. Histone acetylation in chromatin structure and transcription. *Nature* 389, 349–352 (1997). <https://doi.org/10.1038/38664>

Kanno T, Mette MF, Kreil DP, Aufsatz W, Matzke M, Matzke AJ. Involvement of putative SNF2 chromatin remodeling protein DRD1 in RNA-directed DNA methylation. *Curr Biol*. 2004 May 4;14(9):801-5. doi: 10.1016/j.cub.2004.04.037. PMID: 15120073.

Kuhlmann, M., and Mette, M.F. (2012). Developmentally non-redundant SET domain proteins SUVH2 and SUVH9 are required for transcriptional gene silencing in *Arabidopsis thaliana*. *Plant Mol Biol* 79(6), 623-633. doi: 10.1007/s11103-012-9934-x.

Haag, J., Pikaard, C. Multisubunit RNA polymerases IV and V: purveyors of non-coding RNA for plant gene silencing. *Nat Rev Mol Cell Biol* 12, 483–492 (2011). <https://doi.org/10.1038/nrm3152>

Haag J, Ream T, Marasco M et al. In Vitro Transcription Activities of Pol IV, Pol V, and RDR2 Reveal Coupling of Pol IV and RDR2 for dsRNA Synthesis in Plant RNA Silencing. *Molecular Cell*, 811-818, 48(5) (2012) doi:10.1016/j.molcel.2012.09.02

Hamilton, A., Voinnet, O., Chappell, L., and Baulcombe, D. (2002). Two classes of short interfering RNA in RNA silencing. *EMBO J* 21(17), 4671-4679. doi: 10.1093/emboj/cdf464.

Han H, Feng F, Li H. Research advances on epigenetics and cancer metabolism. *Zhejiang Da Xue Xue Bao Yi Xue Ban*. 2021 Feb 25;50(1):1-16. English. doi: 10.3724/zdxbyxb-2021-0053.

Havecker, E.R., Gao, X., and Voytas, D.F. (2004). The diversity of LTR retrotransposons. *Genome Biol* 5(6), 225. doi: 10.1186/gb-2004-5-6-225.

Hayashi, Y., Takehira, K., Nozawa, K., Suzuki, T., Masuta, Y., Kato, A., et al. (2020). ONSEN shows different transposition activities in RdDM pathway mutants. *Genes Genet Syst* 95(4), 183-190. doi: 10.1266/ggs.20-00019.

Huettel B, Kanno T, Daxinger L, Bucher E, van der Winden J, Matzke AJ, Matzke M. RNA-directed DNA methylation mediated by DRD1 and Pol IVb: a versatile pathway for transcriptional gene silencing in plants. *Biochim Biophys Acta*. 2007 May-Jun;1769(5-6):358-74. doi: 10.1016/j.bbaexp.2007.03.001.

Hu H, Du J. Structure and mechanism of histone methylation dynamics in Arabidopsis. *Curr Opin Plant Biol*. 2022 Jun;67:102211. doi: 10.1016/j.pbi.2022.102211. Epub

Ito, H. Environmental stress and transposons in plants. *Genes Genet. Syst*. 2022. Advanced online doi: 10.1266/ggs.22-00045

Ito H, Gaubert H, Bucher E, Mirouze M, Vaillant I, Paszkowski J. An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. *Nature*. 2011 Apr 7;472(7341):115-9. doi: 10.1038/nature09861.

Ito, H., Kim, J.M., Matsunaga, W., Saze, H., Matsui, A., Endo, T.A., et al. (2016). A Stress-Activated Transposon in Arabidopsis Induces Transgenerational Absciscic Acid Insensitivity. *Scientific Reports* 6. doi: Artn 2318110.1038/Srep23181.

Jackson, J., Lindroth, A., Cao, X. et al. Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* 416, 556–560 (2002). <https://doi.org/10.1038/nature731>

Jaenisch, R., Bird, A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet* 33 (Suppl 3), 245–254 (2003). <https://doi.org/10.1038/ng1089>

Jeddeloh, J., Stokes, T. & Richards, E. Maintenance of genomic methylation requires a SWI2/SNF2-like protein. *Nat Genet* 22, 94–97 (1999). <https://doi.org/10.1038/8803>

Jing Y, Sun H, Yuan W, Wang Y, Li Q, Liu Y, Li Y, Qian W. SUVH2 and SUVH9 Couple Two Essential Steps for Transcriptional Gene Silencing in Arabidopsis. *Mol Plant*. 2016 Aug 1;9(8):1156-1167. doi: 10.1016/j.molp.2016.05.006.

Johnson, L.M., Du, J., Hale, C.J., Bischof, S., Feng, S., Chodavarapu, R.K., et al. (2014). SRA- and SET-domain-containing proteins link RNA polymerase V occupancy to DNA methylation. *Nature* 507(7490), 124-128. doi: 10.1038/nature12931.

Johnson, L.M., Law, J.A., Khattar, A., Henderson, I.R., and Jacobsen, S.E. (2008). SRA-domain proteins required for DRM2-mediated de novo DNA methylation. *PLoS Genet* 4(11), e1000280. doi: 10.1371/journal.pgen.1000280.

Kang H, Fan T, Wu J, Zhu Y, Shen WH. Histone modification and chromatin remodeling in plant response to pathogens. *Front Plant Sci*. 2022 Oct 3;13:986940. doi: 10.3389/fpls.2022.986940.

Kanno, T., Aufsatz, W., Jaligot, E., et al. (2005). A SNF2-like protein facilitates dynamic control of DNA methylation. *EMBO Rep*. 6, 649–655. doi:

10.1038/sj.embor.7400446.

Kanno, T., Bucher, E., Daxinger, L., et al. (2008). A structural-maintenance-of-chromosomes hinge domain-containing protein is required for RNA-directed DNA methylation. *Nat. Genet.* 40, 670–675. doi: 10.1038/ng.119.

Kanno, T., Mette, M. F., Kreil, D. P., et al. (2004). Involvement of putative SNF2 chromatin remodeling protein DRD1 in RNA-directed DNA methylation. *Curr. Biol.* 14, 801–805. doi: 10.1016/j.cub.2004.04.037.

Kornberg RD, Lorch Y. Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell.* 1999 Aug 6;98(3):285-94. doi: 10.1016/s0092-8674(00)81958-3.

Kouzarides T. Chromatin modifications and their function. *Cell.* 2007 Feb 23;128(4):693-705. doi: 10.1016/j.cell.2007.02.005.

Kuhlmann, M., and Mette, M.F. (2012). Developmentally non-redundant SET domain proteins SUVH2 and SUVH9 are required for transcriptional gene silencing in *Arabidopsis thaliana*. *Plant Mol Biol* 79(6), 623-633. doi: 10.1007/s11103-012-9934-x.

Law, J.A., Ausin, I., Johnson, L.M., Vashisht, A.A., Zhu, J.K., Wohlschlegel, J.A., et al. (2010). A protein complex required for polymerase V transcripts and RNA-directed DNA methylation in *Arabidopsis*. *Curr Biol* 20(10), 951-956. doi: 10.1016/j.cub.2010.03.062.

Levin, H., Moran, J. Dynamic interactions between transposable elements and their hosts. *Nat Rev Genet* 12, 615–627 (2011). <https://doi.org/10.1038/nrg3030>

Lippman, Z., Gendrel, A.V., Black, M., Vaughn, M.W., Dedhia, N., McCombie, W.R., et al. (2004). Role of transposable elements in heterochromatin and epigenetic

control. *Nature* 430(6998), 471-476. doi: 10.1038/nature02651.

Liu X, Yang S, Zhao M, Luo M, Yu CW, Chen CY, Tai R, Wu K. Transcriptional repression by histone deacetylases in plants. *Mol Plant*. 2014 May;7(5):764-72. doi: 10.1093/mp/ssu033

Liu ZW, Shao CR, Zhang CJ, Zhou JX, Zhang SW, Li L, Chen S, Huang HW, Cai T, He XJ. The SET domain proteins SUVH2 and SUVH9 are required for Pol V occupancy at RNA-directed DNA methylation loci. *PLoS Genet*. 2014 Jan;10(1):e1003948. doi: 10.1371/journal.pgen.1003948.

Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*. 1997 Sep 18;389(6648):251-60. doi: 10.1038/38444.

Luger, K., Dechassa, M. & Tremethick, D. New insights into nucleosome and chromatin structure: an ordered state or a disordered affair?. *Nat Rev Mol Cell Biol* 13, 436–447 (2012). <https://doi.org/10.1038/nrm3382>

Lyons DB, Zilberman D. DDM1 and Lsh remodelers allow methylation of DNA wrapped in nucleosomes. *Elife*. 2017 Nov 15;6:e30674. doi: 10.7554/eLife.30674.

Marsano, R.M., and Dimitri, P. (2022). Constitutive Heterochromatin in Eukaryo

Matsunaga W, Ohama N, Tanabe N et al. 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(FEB-48) (2015) doi: 10.3389/fpls.2015.00048

Matzke, M. A., Mosher, R. A. Rebecca A. Mosher. (2014). RNA-directed DNA methylation: An epigenetic pathway of increasing complexity. *Nat. Rev. Genet.* 15,

394–408. doi: 10.1038/nrg3683

Matzke, M., Kanno, T., Daxinger, L., Huettel, B., Matzke, A. J. (2009). RNA-mediated chromatin-based silencing in plants. *Curr. Opin. Cell Biol.* 21, 367–376. doi: 10.1016/j.ceb.2009.01.025.

Mehdi S, Derkacheva M, Ramström M, Kraleman L, Bergquist J, Hennig L. The WD40 Domain Protein MSI1 Functions in a Histone Deacetylase Complex to Fine-Tune Absciscic Acid Signaling. *Plant Cell*. 2016 Jan;28(1):42-54. doi: 10.1105/tpc.15.00763.

Mhiri, C., Borges, F., and Grandbastien, M.A. (2022). Specificities and Dynamics of Transposable Elements in Land Plants. *Biology (Basel)* 11(4). doi: 10.3390/biology11040488.

Mittelsten S.O., Probst, A.V., Afsar, K., and Paszkowski, J. (2002). Two regulatory levels of transcriptional gene silencing in Arabidopsis. *Proc Natl Acad Sci U S A* 99(21), 13659-13662. doi: 10.1073/pnas.202380499.

Moissiard G, Cokus SJ, Cary J, Feng S, Billi AC, Stroud H, Husmann D, Zhan Y, Lajoie BR, McCord RP, Hale CJ, Feng W, Michaels SD, Frand AR, Pellegrini M, Dekker J, Kim JK, Jacobsen SE. MORC family ATPases required for heterochromatin condensation and gene silencing. *Science*. 2012 Jun 15;336(6087):1448-51. doi: 10.1126/science.1221472.

Naumann, K., Fischer, A., Hofmann, I., Krauss, V., Phalke, S., Irmeler, K., et al. (2005). Pivotal role of AtSUVH2 in heterochromatic histone methylation and gene silencing in Arabidopsis. *EMBO J* 24(7), 1418-1429. doi: 7600604 [pii]

Nguyen HN, Kim JH, Jeong CY, Hong SW, Lee H. Inhibition of histone

deacetylation alters Arabidopsis root growth in response to auxin via PIN1 degradation. *Plant Cell Rep.* 2013 Oct;32(10):1625-36. doi: 10.1007/s00299-013-1474-6.

Nguyen, K.T., Mun, S.H., Lee, C.S. et al. Control of protein degradation by N-terminal acetylation and the N-end rule pathway. *Exp Mol Med* 50, 1–8 (2018). <https://doi.org/10.1038/s12276-018-0097-y>

Nozawa, K., Chen, J., Jiang, J., et al. (2021). DNA methyltransferase CHROMOMETHYLASE3 prevents ONSEN transposon silencing under heat stress. *PLOS Genet.* 17, e1009710. doi: 10.1371/journal.pgen.1009710.

Núñez-Álvarez Y, Suelves M. HDAC11: a multifaceted histone deacetylase with proficient fatty deacylase activity and its roles in physiological processes. *FEBS J.* 2022 May;289(10):2771-2792. doi: 10.1111/febs.15895.

Okitsu CY, Hsieh CL. DNA methylation dictates histone H3K4 methylation. *Mol Cell Biol.* 2007 Apr;27(7):2746-57. doi: 10.1128/MCB.02291-06. Epub 2007 Jan 22. PMID: 17242185; PMCID: PMC1899905.

Oliver, K.R., and Greene, W.K. (2009). Transposable elements: powerful facilitators of evolution. *Bioessays* 31(7), 703-714. doi: 10.1002/bies.200800219

Onodera, Y., Haag, J.R., Ream, T., Costa Nunes, P., Pontes, O., and Pikaard, C.S. (2005). Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell* 120(5), 613-622. doi: 10.1016/j.cell.2005.02.007.

Pontes, O., Li, C. F., Costa Nunes, P. et al. (2006). The Arabidopsis chromatin-modifying nuclear siRNA pathway involves a nucleolar RNA processing center. *Cell.* 126, 79–92. doi: 10.1016/j.cell.2006.05.031.

Quadrana L, Silveira AB, Caillieux E, Colot V. Detection of Transposable Element Insertions in Arabidopsis Using Sequence Capture. *Methods Mol Biol.* 2021;2250:141-155. doi: 10.1007/978-1-0716-1134-0_14.

Quesneville, H. Twenty years of transposable element analysis in the Arabidopsis thaliana genome. *Mobile DNA* 11, 28 (2020). <https://doi.org/10.1186/s13100-020-00223-x>

Qi Y, He X, Wang XJ, Kohany O, Jurka J, et al. (2006) Distinct catalytic and non-catalytic roles of ARGONAUTE4 in RNA-directed DNA methylation. *Nature.* 26;443(7114):1008-12(2006) doi: 10.1038/nature05198.

Ritu Pandey, Andreas Müller, Carolyn A. Napoli, David A. Selinger, Craig S. Pikaard, Eric J. Richards, Judith Bender, David W. Mount, Richard A. Jorgensen, Analysis of histone acetyltransferase and histone deacetylase families of Arabidopsis thaliana suggests functional diversification of chromatin modification among multicellular eukaryotes, *Nucleic Acids Research*, Volume 30, Issue 23, 1 December 2002, Pages 5036–5055, <https://doi.org/10.1093/nar/gkf660>

Rothbart SB, Strahl BD. Interpreting the language of histone and DNA modifications. *Biochim Biophys Acta.* 2014 Aug;1839(8):627-43. doi: 10.1016/j.bbagr.2014.03.001.

Sasaki T, Lorković ZJ, Liang SC, Matzke AJ, Matzke M. The ability to form homodimers is essential for RDM1 to function in RNA-directed DNA methylation. *PLoS One.* 2014 Feb 3;9(2):e88190. doi: 10.1371/journal.pone.0088190.

Shahbazian MD, Grunstein M. Functions of site-specific histone acetylation and deacetylation. *Annu Rev Biochem.* 2007;76:75-100. doi: 10.1146/annurev.biochem.76.052705.162114.

Shim S, Lee HG, Seo PJ. MET1-Dependent DNA Methylation Represses Light Signaling and Influences Plant Regeneration in Arabidopsis. *Mol Cells*. 2021 Oct 31;44(10):746-757. doi: 10.14348/molcells.2021.0160.

Slotkin, R.K., and Martienssen, R. (2007). Transposable elements and the epigenetic regulation of the genome. *Nature Reviews Genetics* 8(4), 272-285. doi: 10.1038/nrg2072.

Stroud, H., Do, T., Du, J. et al. Non-CG methylation patterns shape the epigenetic landscape in Arabidopsis. *Nat Struct Mol Biol* 21, 64–72 (2014). <https://doi.org/10.1038/nsmb.2735>

Tanaka M, Kikuchi A, Kamada H. The Arabidopsis histone deacetylases HDA6 and HDA19 contribute to the repression of embryonic properties after germination. *Plant Physiol*. 2008 Jan;146(1):149-61. doi: 10.1104/pp.107.111674.

Takehira, K., Hayashi, Y., Nozawa, K., et al. (2021). DRD1, a SWI/SNF-like chromatin remodeling protein, regulates a heat-activated transposon in Arabidopsis thaliana. *Genes Genet. Syst.* 96, 151–158. doi: 10.1266/ggs.21-00005.

Ueda M, Matsui A, Tanaka M, Nakamura T, Abe T, Sako K, Sasaki T, Kim JM, Ito A, Nishino N, Shimada H, Yoshida M, Seki M. The Distinct Roles of Class I and II RPD3-Like Histone Deacetylases in Salinity Stress Response. *Plant Physiol*. 2017 Dec;175(4):1760-1773. doi: 10.1104/pp.17.01332. Epub 2017 Oct 10.

Ueda M, Matsui A, Nakamura T, Abe T, Sunaoshi Y, Shimada H, Seki M. Versatility of HDA19-deficiency in increasing the tolerance of Arabidopsis to different environmental stresses. *Plant Signal Behav*. 2018;13(8):e1475808. doi: 10.1080/15592324.2018.1475808.

Vladimir Beljanski, Trichostatin A, Editor(s): S.J. Enna, David B. Bylund,xPharm: The Comprehensive Pharmacology Reference, Elsevier, 2009, Pages 1-4, ISBN 9780080552323, <https://doi.org/10.1016/B978-008055232-3.64490-3>.

Vongs A, Kakutani T, Martienssen RA, Richards EJ. Arabidopsis thaliana DNA methylation mutants. Science. 1993 Jun 25;260(5116):1926-8. doi: 10.1126/science.8316832. PMID: 8316832.

Wakasa, Y., Kawakatsu, T., Harada, T., Takaiwa, F. (2018). Transgene-independent heredity of RdDM-mediated transcriptional gene silencing of endogenous genes in rice. Plant Biotechnol. J.. 16, 2007–2015. doi: 10.1111/pbi.12934.

Wang Z, Cao H, Chen F and Liu Y. (2014). The roles of histone acetylation in seed performance and plant development. Plant Physiology and Biochemistry. Volume 84(125-133), <https://doi.org/10.1016/j.plaphy.2014.09.010>.

Wójcikowska B, Botor M, Morończyk J, Wójcik AM, Nodzyński T, Karcz J, Gaj MD. Trichostatin A Triggers an Embryogenic Transition in Arabidopsis Explants via an Auxin-Related Pathway. Front Plant Sci. 2018 Sep 13;9:1353. doi: 10.3389/fpls.2018.01353.

Wójcikowska B, Chwiałkowska K, Nowak K, Citerne S, Morończyk J, Wójcik AM, Kiwior-Wesołowska A, Francikowski J, Kwaśniewski M, Gaj MD. Transcriptomic profiling reveals histone acetylation-regulated genes involved in somatic embryogenesis in Arabidopsis thaliana. BMC Genomics. 2024 Aug 15;25(1):788. doi: 10.1186/s12864-024-10623-5.

Matsunaga W, Kobayashi A, Kato A, Ito H. The effects of heat induction and the siRNA biogenesis pathway on the transgenerational transposition of ONSEN, a copia-like retrotransposon in Arabidopsis thaliana. Plant Cell Physiol. 2012 May;53(5):824-

33. doi: 10.1093/pcp/pcr179. Epub 2011 Dec 14. PMID: 22173101.

Matsunaga W, Ohama N, Tanabe N, Masuta Y, Masuda S, Mitani N, Yamaguchi-Shinozaki K, Ma JF, Kato A, Ito H. A small RNA mediated regulation of a stress-activated retrotransposon and the tissue specific transposition during the reproductive period in Arabidopsis. *Front Plant Sci.* 2015 Feb 9;6:48. doi: 10.3389/fpls.2015.00048. PMID: 25709612; PMCID: PMC4321352.

Wicker, T., Sabot, F., Hua-Van, A. et al. A unified classification system for eukaryotic transposable elements. *Nat Rev Genet* 8, 973–982 (2007). <https://doi.org/10.1038/nrg2165>

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

Wierzbicki, A. T., Ream, T. S., Haag, J. R., Pikaard, C. S. (2009). RNA polymerase V transcription guides ARGONAUTE4 to chromatin. *Nat. Genet.* 41, 630–634. doi: 10.1038/ng.365.

Wongpalee, S. P., Liu, S., Gallego-Bartolomé, J., et al. (2019). CryoEM structures of Arabidopsis DDR complexes involved in RNA-directed DNA methylation. *Nat. Commun.* 10, article number: 3916. doi: 10.1038/s41467-019-11759-9.

Wu, L., Mao, L. & Qi, Y. Roles of DICER-LIKE and ARGONAUTE proteins in TAS-derived small interfering RNA-triggered DNA methylation. *Plant Physiol.* 160,990–999 (2012) doi.org/10.1104/pp.112.200279

Xie Z, Johansen L, Gustafson A et al. Genetic and functional diversification of small RNA pathways in plants. *PLoS Biology*, 2(5) (2004) doi:

10.1371/journal.pbio.0020104

Xu CR, Liu C, Wang YL, Li LC, Chen WQ, Xu ZH, Bai SN. Histone acetylation affects expression of cellular patterning genes in the Arabidopsis root epidermis. *Proc Natl Acad Sci U S A*. 2005 Oct 4;102(40):14469-74. doi: 10.1073/pnas.0503143102.

Zhang H, Lang Z, Zhu JK. Dynamics and function of DNA methylation in plants. *Nat Rev Mol Cell Biol*. 2018 Aug;19(8):489-506. doi: 10.1038/s41580-018-0016-z.

Zemach A, Kim MY, Hsieh PH, Coleman-Derr D, Eshed-Williams L, Thao K, Harmer SL, Zilberman D. The Arabidopsis nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. *Cell*. 2013 Mar 28;153(1):193-205. doi: 10.1016/j.cell.2013.02.033.

Zheng Y, Li Z, Cui X, Yang Z, Bao C, Pan L, Liu X, Chatel-Innocenti G, Vanacker H, Noctor G, Dard A, Reichheld JP, Issakidis-Bourguet E, Zhou DX. S-Nitrosylation of the histone deacetylase HDA19 stimulates its activity to enhance plant stress tolerance in Arabidopsis. *Plant J*. 2023 May;114(4):836-854. doi: 10.1111/tpj.16174.

Zilberman, D., Gehring, M., Tran, R. et al. Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. *Nat Genet* 39, 61–69 (2007). <https://doi.org/10.1038/ng1929>

Zhou C, Zhang L, Duan J, Miki B, Wu K. HISTONE DEACETYLASE19 is involved in jasmonic acid and ethylene signaling of pathogen response in Arabidopsis. *Plant Cell*. 2005 Apr;17(4):1196-204. doi: 10.1105/tpc.104.028514.

Zhou J, Wang X, He K, Charron JB, Elling AA, Deng XW. Genome-wide profiling of histone H3 lysine 9 acetylation and dimethylation in Arabidopsis reveals correlation between multiple histone marks and gene expression. *Plant Mol Biol*. 2010

Apr;72(6):585-95. doi: 10.1007/s11103-009-9594-7.

Zhou Y, Tan B, Luo M, Li Y, Liu C, Chen C, Yu CW, Yang S, Dong S, Ruan J, Yuan L, Zhang Z, Zhao L, Li C, Chen H, Cui Y, Wu K, Huang S. HISTONE DEACETYLASE19 interacts with HSL1 and participates in the repression of seed maturation genes in Arabidopsis seedlings. *Plant Cell*. 2013 Jan;25(1):134-48. doi: 10.1105/tpc.112.096313.

Zhou Y, Yang P, Zhang F, Luo X, Xie J. Histone deacetylase HDA19 interacts with histone methyltransferase SUVH5 to regulate seed dormancy in Arabidopsis. *Plant Biol (Stuttg)*. 2020 Nov;22(6):1062-1071. doi: 10.1111/plb.13158.

Zhong S, Xu Y, Yu C, Zhang X, Li L, Ge H, Ren G, Wang Y, Ma J, Zheng Y, Zheng B. Anaphase-promoting complex/cyclosome regulates RdDM activity by degrading DMS3 in Arabidopsis. *Proc Natl Acad Sci U S A*. 2019 Feb 26;116(9):3899-3908. doi: 10.1073/pnas.1816652116.

Zhong, X., Hale, C. J., Law, J. A., et al. (2012). DDR complex facilitates global association of RNA polymerase V to promoters and evolutionarily young transposons. *Nat. Struct. Mol. Biol.* 19, 870–875. doi: 10.1038/nsmb.2354.

Zhong, Z., Feng, S., Duttke, S.H., Potok, M.E., Zhang, Y., Gallego-Bartolome, J., et al. (2021). DNA methylation-linked chromatin accessibility affects genomic architecture in Arabidopsis. *Proc Natl Acad Sci U S A* 118(5). doi: 10.1073/pnas.2023347118.

Zubko E, Gentry M, Kunova A, Meyer P. De novo DNA methylation activity of methyltransferase 1 (MET1) partially restores body methylation in Arabidopsis thaliana. *Plant J*. 2012 Sep;71(6):1029-37. doi: 10.1111/j.1365-313X.2012.05051.x.

Acknowledgement

The completion of this thesis is attributed to the support and encouragement of many people.

First and foremost, I would like to extend my heartfelt gratitude to Professors Hidetaka Ito and Atsushi Kato from the Laboratory of Structure and Cell function III at Hokkaido University, as well as Professors Tomomichi Fujita and Yukako Chiba, for their invaluable guidance and support throughout this research.

I also want to thank Prof. Yoko Ikeda from Institute of Plant Science and Resources, Okayama University, Prof. Xuehua Zhong from the Washington University in St. Louis; Prof. Tetsuji Kakutani and Soichi Inagaki from The University of Tokyo, Prof. Yijun Meng from Hangzhou Normal University, and Prof. Leandro Quadrana and the lab members from Université Paris-Saclay for their assistance in our collaborative research. Additionally, I greatly appreciate the guidance in bioinformatics data analysis from Prof. Hidetoshi Saze at the Okinawa Institute of Science and Technology, Takamasa Suzuki at Chubu University, Sayuri Tsukahara at The University of Tokyo, and Hiroki Takagi at Ishikawa Prefectural University. I would also like to extend my special thanks to my undergraduate mentor, Professor Qin Zhu, for your continued support.

Furthermore, I would like to thank all the members of the Laboratory of Structure and Cell function III for their help and encouragement while I was preparing this thesis. Lastly, I would like to thank my parents, friends, and Gaheris, who continue to support me.

This work was supported by the following grants: the Japan Society for the Promotion of Science (JSPS) Scientific Research Fund for International Joint Research Acceleration (International Joint Research Strengthening (A) (Principal Investigator:

Hidetaka Ito, Project Number 21KK0263)); the JSPS Scientific Research Fund for Basic Research (C) (Principal Investigator: Hidetaka Ito, Project Number 21K06008); the Hokkaido University Ambitious Doctoral Fellowship; and the JSPS Special Research Fellowship. H.S. was supported by MEXT Grant-in-Aid for Transformative Research Areas (A) JP20H05913 and by OIST. I thank OIST SQC for Illumina RNA-seq. Computations were partially performed on the NIG supercomputer at ROIS National Institute of Genetics. I am grateful to these funding organizations for their support of this research.