

Small RNAs guide histone methylation in *Arabidopsis* embryos

Jean-Sébastien Parent,^{1,2} Jonathan Cahn,¹
Rowan P. Herridge,^{1,5} Daniel Grimanelli,^{3,4}
and Robert A. Martienssen¹

¹Howard Hughes Medical Institute, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA; ²Agriculture and Agri-Food Canada, Ottawa Research and Development Center, Ottawa Ontario K1A 0C6, Canada; ³Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA; ⁴Epigenetic Regulations and Seed Development, UMR232, Institut de Recherche pour le Développement (IRD), Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD), Université de Montpellier (UM), 34394 Montpellier, France

Epigenetic reprogramming occurs during gametogenesis as well as during embryogenesis to reset the genome for early development. In flowering plants, many heterochromatic marks are maintained in sperm, but asymmetric DNA methylation is mostly lost. Asymmetric DNA methylation is dependent on small RNA but the re-establishment of silencing in embryo is not well understood. Here we demonstrate that small RNAs direct the histone H3 lysine 9 dimethylation during *Arabidopsis thaliana* embryonic development, together with asymmetric DNA methylation. This de novo silencing mechanism depends on the catalytic domain of SUVH9, a Su(VAR)3-9 homolog thought to be catalytically inactive.

Supplemental material is available for this article.

Received August 19, 2020; revised version accepted April 16, 2021.

Constitutive heterochromatin in many eukaryotes is characterized by the presence of methyl groups on the cytosine bases of DNA (mC) and by di- or trimethylation of lysine 9 of histone H3 (H3K9me2/3) (Du et al. 2015). Maintenance of these marks through cell divisions is essential to keep transposable elements (TEs) silenced and ensure genome stability (Allshire and Madhani 2018). In plants, DNA methylation is found in three functionally distinct sequence contexts, namely, symmetric CG, CHG, and asymmetric CHH (where H is either A, C, or T), with the non-CG contexts tightly associated with H3K9me2 in heterochromatin (Stroud et al. 2014). In *Arabidopsis thaliana*, this association relies on the activity of three H3K9 methyltransferases (KMTs), SU(VAR)3-9 HOMO-

LOG4 (SUVH4), SUVH5, and SUVH6 that bind methylated cytosines in different contexts through their SRA domain (Du et al. 2015). The methylated histones are then recognized by the DNA methyltransferases (DMTs) CHROMOMETHYLASE2 (CMT2) and CMT3 that perpetuate CHH and CHG methylation, respectively (Stroud et al. 2014). The interdependence between the two marks is thought to be responsible for the propagation of H3K9me2 through cell generations.

During the flowering plant male germline development, CG and CHG methylation are retained but CHH methylation is lost from the paternal genome, first during meiosis and then through the microspore and mature sperm cells (Calarco et al. 2012; Ibarra et al. 2012; Ingouff et al. 2017; Walker et al. 2017). Following fertilization, this methylation must be restored de novo and has been proposed to be guided by small RNA (sRNA) molecules (Calarco et al. 2012; Ibarra et al. 2012). In *Arabidopsis*, these sRNA molecules are 24 nt long and produced by RNA POLYMERASE IV (POLIV) (Matzke et al. 2015). They are loaded onto a specific clade of ARGONAUTE proteins (mainly AGO4, AGO6, and AGO9), which can then detect homology in noncoding transcripts produced by RNA POLYMERASE V (POLV). POLV is recruited to methylated DNA by SUVH2 and SUVH9, the latter having specific affinity for CHH methylation bestowed by its SRA domain (Johnson et al. 2014; Liu et al. 2014; Ingouff et al. 2017). The recognition of POLV transcripts by AGO proteins ultimately recruits DOMAIN REARRANGED METHYLTRANSFERASE1 (DRM1) and DRM2 proteins that methylate cytosines in all three sequence contexts, a process known as RNA-directed DNA methylation (RdDM) (Matzke et al. 2015).

However, the current model for RdDM implies pre-existing DNA methylation and so cannot account for purely de novo silencing. One possible explanation is that sRNAs targeting TEs instead guide KMTs as they do in fission yeast, *Caenorhabditis elegans*, and *Drosophila* (Castel and Martienssen 2013), all of which lack these DMTs. H3K9me2 could then recruit POLIV via the SAWADEE HOMEODOMAIN HOMOLOG1 (SHH1) (Law et al. 2011) and maintain silencing through RdDM. Here we investigated the genetic dependencies of methylated DNA and histone H3K9 in mature embryos of *Arabidopsis thaliana*. We observe that H3K9 methylation at small RNA targets is independent of the maintenance KMTs SUVH4/5/6 and does not correlate perfectly with non-CG methylation at the same loci. We propose a role for a new family of KMT that is directly guided by small RNA molecules for de novo silencing.

Results and Discussion

sRNA molecules guide CHH and CHG DNA methylation to specific targets in embryos

It was recently shown that RdDM activity peaks during embryogenesis in *Arabidopsis* as well as in soybeans and

[**Keywords:** plant embryogenesis; epigenetic reprogramming; small RNA molecules; DNA methylation; histone methylation]

⁵Present address: Department of Biochemistry, University of Otago, Dunedin 9016, New Zealand.

Corresponding author: martiens@cshl.edu

Article published online ahead of print. Article and publication date are online at <http://www.genesdev.org/cgi/doi/10.1101/gad.343871.120>. Freely available online through the *Genes & Development* Open Access option.

© 2021 Parent et al. This article, published in *Genes & Development*, is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

chickpeas (An et al. 2017; Bouyer et al. 2017; Kawakatsu et al. 2017; Rajkumar et al. 2020), which is otherwise consistent with a role in epigenetic reprogramming. We therefore examined both the repressive mC and H3K9me2 modifications in *Arabidopsis* embryos. First, whole-genome bisulfite sequencing (WGBS) was performed on dissected mature embryos. We tested an *ago4 ago6 ago9* triple mutant, given that the three proteins are present in the cell nucleus of different embryonic tissues throughout development (Supplemental Fig. S1) and all are required for 24-nt sRNA activity (Liu et al. 2018). We also tested a *polIV* mutant devoid of 24-nt sRNAs, as well as *drm1 drm2*, and *svh2 svh9* double mutants to compare them with the *svh4 svh5 svh6* triple mutant with impaired H3K9me2 maintenance (Stroud et al. 2014). Methylation ratios were separated by context, and CHH was used to call differentially methylated regions (DMRs).

We observed that sRNA-dependent DMRs in embryos (defined as hypomethylated in *ago4 ago6 ago9*) mapped mostly to DNA transposons (Fig. 1A) and were largely excluded from sRNA-independent DMRs (defined as hypomethylated in *svh4 svh5 svh6*), comprising mostly Copia and Gypsy elements (retrotransposons). As in vegetative tissues (Stroud et al. 2013), CHG methylation at sRNA targets is reduced in mutant embryos but not entirely lost (Fig. 1B). CHH methylation is also exclusive to each target group with *svh4 svh5 svh6* mutants being unaffected at sRNA targets while losing all CHH methylation at sRNA-independent targets. Finally,

the two pathways are able to repress their respective targets without influencing one another, as exemplified by neighboring TEs at many different locations in the genome (Fig. 1C).

sRNA molecules target H3K9me2 independently of SUVH4, SUVH5, and SUVH6 in embryos

Next, we performed chromatin immunoprecipitation followed by sequencing (ChIP-seq) on dissected mature embryos. The genotypes used in WGBS experiments were again tested in duplicates to ascertain reproducibility (Supplemental Fig. S2). We used previously defined DMRs to map H3K9me2 levels and showed that RdDM mutants have lower H3K9me2 levels at sRNA-dependent regions but wild-type levels at sRNA-independent regions (Fig. 2A). We noted that the *drm1 drm2* mutant appeared less affected, suggesting sRNA-dependent H3K9me2 did not fully depend on mC in embryos (Fig. 2A, top panel; Supplemental Fig. S3). In contrast, *svh4 svh5 svh6* H3K9me2 levels were not decreased compared with wild type at sRNA-dependent regions (Fig. 2A, top panel). This observation argues that neither SUVH4, SUVH5, nor SUVH6 is responsible for depositing H3K9me2 at these loci.

This prompted us to use immunofluorescence to test whether the global levels of H3K9me2 were perceptibly changed in these mutants during embryo development, something not previously reported. We identified a developmental window between the zygote and the 64-cell stage where H3K9me2 immunofluorescence at heterochromatic foci is sharply reduced in *ago4 ago6 ago9* and *svh2 svh9* mutants (Fig. 2B; Supplemental Fig. S4), as it is in *svh4 svh5 svh6* mutants at these same embryonic stages. This was unexpected because, as small RNA-dependent regions are scattered throughout the genome, loss of RdDM does not lead to loss of immunofluorescent signal in somatic cells (Johnson et al. 2008; Jing et al. 2016). However, *ago4 ago6 ago9* and *svh2 svh9* mutants recover H3K9me2 signal at later stages unlike *svh4 svh5 svh6* (Fig. 2C; Supplemental Table S1), in agreement with our ChIP-seq results that show massive loss of H3K9me2 signal in the pericentromeric regions only in the latter.

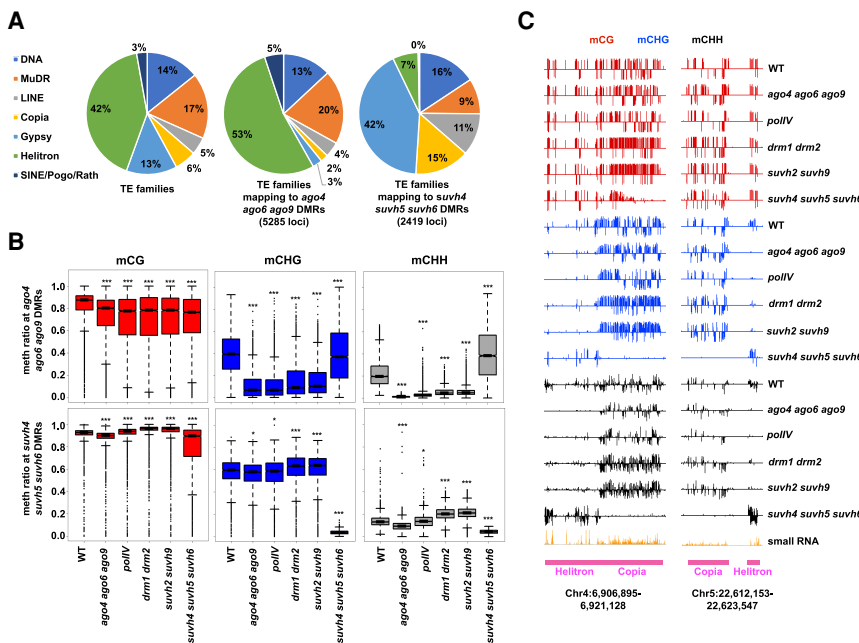


Figure 1. Small RNA-dependent and -independent DNA methylation in embryos. (A) Hypomethylated DMRs were extracted for *ago4 ago6 ago9* and *svh4 svh5 svh6* mutants and mapped on TE features and broken down in families. Distribution of all transposable elements from the *Arabidopsis thaliana* genome is presented at the left. (B) Distributions representing methylation ratios of the different contexts are identified above in the different genetic backgrounds indicated below. Ratios were averaged within the genomic regions identified in the DMR analysis. Differences between distributions and the first WT distribution were calculated using a Kolmogorov–Smirnov test, and asterisks represent the significance of the difference. (*) $P \leq 1.0 \times 10^{-4}$, (**) $P \leq 1.0 \times 10^{-11}$, (***) $P \leq 1.0 \times 10^{-15}$. (C) Methylation plot of representative genomic regions with coordinates indicated below and containing TEs. Methylation context is color-coded as indicated above, and genotypes are indicated at the right. Small RNA track was generated using published embryo data (Ye et al. 2016).

developmental window between the zygote and the 64-cell stage where H3K9me2 immunofluorescence at heterochromatic foci is sharply reduced in *ago4 ago6 ago9* and *svh2 svh9* mutants (Fig. 2B; Supplemental Fig. S4), as it is in *svh4 svh5 svh6* mutants at these same embryonic stages. This was unexpected because, as small RNA-dependent regions are scattered throughout the genome, loss of RdDM does not lead to loss of immunofluorescent signal in somatic cells (Johnson et al. 2008; Jing et al. 2016). However, *ago4 ago6 ago9* and *svh2 svh9* mutants recover H3K9me2 signal at later stages unlike *svh4 svh5 svh6* (Fig. 2C; Supplemental Table S1), in agreement with our ChIP-seq results that show massive loss of H3K9me2 signal in the pericentromeric regions only in the latter.

SUVH9 catalytic site is required for epigenetic reprogramming

Our results therefore pointed to the existence of one or more KMTs acting in concert with sRNA molecules in embryos to impose a silencing mark. We rationalized that SUVH2 and SUVH9 were the most likely candidates and, like G9a in human, might both recruit

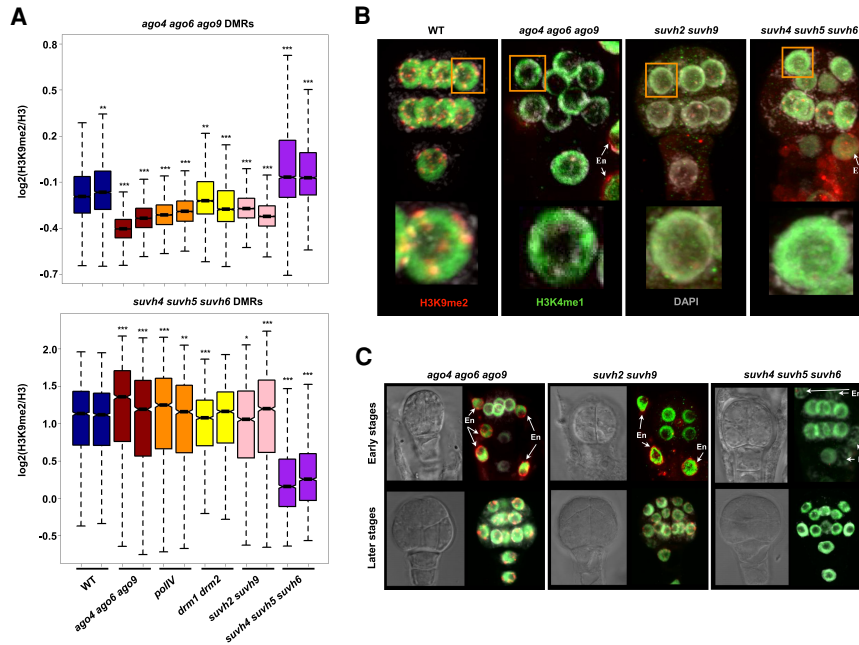


Figure 2. Small RNA-dependent and -independent H3K9me2 in embryos. (A) Distributions of normalized H3K9me2 ratios [$\log_2(\text{H3K9me2}/\text{H3})$] were calculated over the genomic regions defined by the DMR analysis. Biological replicates are presented for the genomic backgrounds indicated below. Differences between distributions and the first WT distribution were calculated using a Kolmogorov–Smirnov test, and asterisks represent the significance of the difference. (*) $P \leq 1.0 \times 10^{-4}$, (**) $P \leq 1.0 \times 10^{-11}$, (***) $P \leq 1.0 \times 10^{-15}$. (B,C) Immunofluorescence imaging of H3K9me2 (red), H3K4me1 (green), and chromatin (gray) in representative embryos of indicated background. Early embryos from the four-cell to the 16-cell stage are presented. “En” and white arrows point at endosperm cells. (C) DIC images are presented at the left, early stages are presented at the top, and later stages are below.

DMTs and deposit methyl groups on lysine 9 (Dong et al. 2008). However, these two proteins are not thought to be active KMTs due mainly to their lack of a post-SET domain (Johnson et al. 2008, 2014). Despite this short truncation, they do have a SET domain with the catalytic tyrosine residue conserved, and the *suvh2 suvh9* mutant has sharply reduced H3K9me2 foci in early embryogenesis (Fig. 2B) and loss of sRNA-dependent K9 methylation in mature embryos (Fig. 2A, top panel). We therefore attempted to complement *suvh2 suvh9* plants with a catalytic site mutant transgene for both *SUVH2* and *SUVH9*, along with wild-type controls. These proteins were tagged with mCitrine in order to confirm their presence in the embryonic cells. We were unable to recover embryos with fluorescent signal in the case of *SUVH2*, probably due to low gene expression (Johnson et al. 2008). However, we did recover lines with a visible nuclear signal in the cotyledons of mature embryonic cells for the different *SUVH9* transgenes (Fig. 3A). We also expressed *SUVH9* with a mutated SRA domain as reported previously (Johnson et al. 2008) but were unable to see a positive mCitrine signal in any transformants. Overexpression revealed that the S252F mutation made *SUVH9* unstable, proscribing its use in our experiments (Supplemental Fig. S5).

We tested whether the new constructs could rescue the imprinted *SUPPRESSOR OF DRM2 AND CMT3* (*SDC*) gene that is activated by complete loss of epigenetic silencing in the *suvh2 suvh4 suvh9* mutant, causing developmental phenotypes (Johnson et al. 2008). Since the activated *SDC* allele requires functional de novo silencing

activity to restore silencing (Henderson and Jacobsen 2008), crossing *suvh2 suvh4 suvh9* to *suvh2 suvh9* results in 100% F1 plants with the *SDC* phenotype, while crossing *suvh2 suvh4 suvh9* to WT yields normal progeny (Fig. 3B). When the triple mutant was crossed with two different lines of *suvh2 suvh9* complemented with the wild-type transgene, the offspring showed a wild-type phenotype, confirming the rescue of *SUVH9* function. However, when the catalytic tyrosine Y636 is mutated, the rescue is abolished in two independent lines (Fig. 3B; Supplemental Table S2). This suggests that catalytic activity is essential for the reprogramming of the *SDC* allele by the RdDM pathway.

We then performed ChIP-seq and WGBS on developing seeds at the mature embryo stage from *suvh2 suvh9* mutants complemented by mutated and wild-type *SUVH9*. Using MACS2 (Zhang et al. 2008), we identified 203 regions with H3K9me2 peaks in seeds complemented with *SUVH9*(WT) that did not show enrichment in the control (Fig. 3C). The majority of regions identified overlapped with DNA transposons of the MuDR (44%) and Helitron (19%) families, in line with the target preference of

this silencing pathway. H3K9me2 levels at these loci were also increased in plants complemented with the Y636F version but to a significantly lower extent (Fig. 3C). Similarly, DNA methylation in the CHH context at the same loci is significantly increased in both complemented lines but to a lower extent with the mutant version (Fig. 3D). This suggests that the catalytic residue also has a direct or indirect influence on the ability to recruit DMTs. Importantly both WT and catalytic mutants were also able to restore non-CG methylation to the *SDC* locus (Supplemental Fig. S6), but only the WT form was able to restore silencing (Fig. 3B; Supplemental Table S2), further demonstrating the importance of *SUVH9* catalytic activity for epigenetic reprogramming.

The *SUVH9* family appeared early in the evolution of flowering plants, as suggested by the presence of a single gene in the basal angiosperm *Amborella trichopoda* (Li et al. 2015). This copy already displays the truncation of the C-terminal post-SET domain that is essential for completing the catalytic site of other KMTs (Johnson et al. 2014). This supported the notion that *SUVH9* lost catalytic function early in its evolution and served only as an adaptor between DNA methylation and POLV (Johnson et al. 2008, 2014; Liu et al. 2014). However, all the *SUVH9* orthologs we have examined possess the catalytic tyrosine as well as many other structurally important residues of the SET domain. Such a level of sequence conservation in an inactive SET domain seems unlikely, and indeed, mutating the catalytic residue abolished *SUVH9* function (Fig. 3B,C). The crystal structure of *SUVH9* has

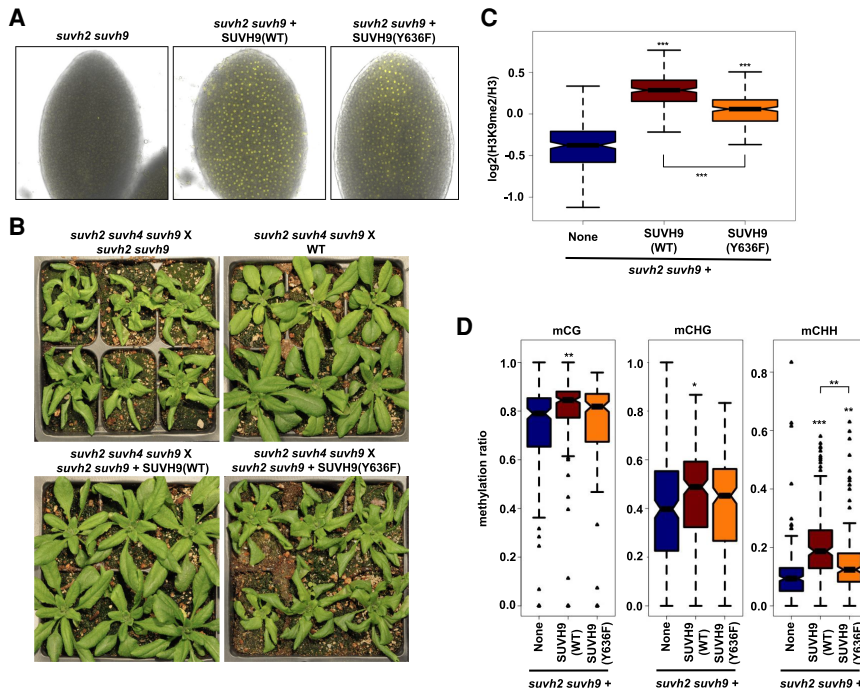


Figure 3. Functional complementation of SUVH9. (A) Confocal imaging of mCitrine signal and transmitted light on mature embryo cotyledons of the genotype indicated above. (B) Pictures of 4-wk-old F1 plants resulting from the crosses indicated above. (C) Distributions of normalized H3K9me2 ratios [$\log_2(\text{H3K9me2}/\text{H3})$] were calculated over the genomic regions identified by the MACS2 broad peak analysis. (D) Distributions representing methylation ratios of the different contexts are identified above in the different genetic backgrounds indicated below. Ratios were calculated over the genomic regions identified by the MACS2 broad peak analysis. Distances between distributions were calculated using a Kolmogorov–Smirnov test, and asterisks represent the significance of the difference. (*) $P \leq 1.0 \times 10^{-6}$, (**) $P \leq 1.0 \times 10^{-10}$, (***) $P \leq 1.0 \times 10^{-15}$. The difference is between the distribution identified and the equivalent control unless otherwise represented.

revealed that the substrate binding pocket is too widely opened and incomplete, explaining the lack of in vitro activity in recombinant protein (Johnson et al. 2008, 2014). This is reminiscent of other well-known KMTs; namely, CURLY LEAF (CLF) and MEDEA (MEA), two E(z) homologs in plants. E(z) homologs do not possess a cysteine-rich post-SET domain and yet are known to catalyze the methylation of lysine 27 (Schmitges et al. 2011; Jacob et al. 2014). Like their animal counterparts, these proteins need additional subunits to be active and are therefore purified as complexes for in vitro assays (Schmitges et al. 2011; Jacob et al. 2014). We suspect that SUVH9 also requires interactors to complete its catalytic site in embryonic cells.

Our results reveal the existence of sRNA-dependent H3K9me2 in embryos, independent of canonical KMTs SUVH4, SUVH5, and SUVH6 (Fig. 2A, top panel) and partially independent of the DMTs DRM1 and DRM2 (Supplemental Fig. S3). Interestingly, it was recently shown that targeting of some RdDM components to a naive locus triggers silencing even in the absence of DRM1 and DRM2 (Gallego-Bartolomé et al. 2019). This activity, which is dependent on SUVH2/9, appears to be widespread in early embryos but limited to certain targets, particularly DNA transposons along chromosome arms, by the mature embryo stage. It is also noteworthy that the pericentromeric H3K9me2 in endosperm of early seed is not dependent on SUVH2/9 but only SUVH4/5/6 (Fig. 2B,C; Supplemen-

tal Fig. S4), similar to what is observed in leaves. This is consistent with a need for epigenetic reprogramming in early embryogenesis where asymmetric DNA methylation levels are low. Pericentromeric regions in *suvh2 suvh9* mutants are nonetheless able to recover normal H3K9me2 levels later in development (Fig. 2C, bottom panel), presumably through the persistence of symmetric DNA methylation and the activity of SUVH4/5/6 (To et al. 2020).

Flowering plants have therefore evolved a specialized clade of Su (Var) 3-9 homologs, which we show are associated with small RNA-dependent transcriptional silencing in the developing seed. Interestingly, this pathway appeared in angiosperms at the same time as the appearance of dosage imbalance and imprinting in the endosperm, due to double fertilization (Wang and Köhler 2017). Remarkably, both SUVH9 and H3K9me2 are important for imprinting and chromosome dosage response in the endosperm (Jiang et al. 2017). Indeed, the sRNA-dependent DMRs we have identified in the embryo also gain H3K9me2 in the unbalanced 4n endosperm (Supplemental Fig. S7), implicating sRNA-dependent H3K9me2 in triploid seed abortion. Our study therefore links the emergence of SUVH9

homologs to the appearance of interploidy hybridization barriers in endosperm development, an important driver of plant speciation. It has previously been speculated that the small RNA silencing pathway could be a facilitator of genome duplications that are frequent in the history of flowering plants (Matzke et al. 2015). A mechanism that acts to control both TE expansion during genome duplication and the establishment of interploidy barriers would undoubtedly have played a crucial role in Darwin's "abominable" diversification of angiosperms (Cibrian-Jaramillo and Martienssen 2009).

Materials and methods

Plant material and growth conditions

Plants were grown under long-day conditions at 22°C. Seeds were sown directly on soil after stratification in distilled water for 3 d at 4°C. For developing seed isolation, the mature fruits were collected from the primary inflorescence after 7–8 wk of growth and cut open, and fresh seeds were harvested and frozen. For embryo isolation, siliques were cut open with insulin syringe needles; whole seeds were transferred on microscope slide in a drop of PBS; and embryos were squeezed out, applying pressure with a microslide. Embryos were then manually separated from the other tissues and washed three times in PBS before freezing in liquid nitrogen. We combined the alleles *ago4-5*, *ago6-2*, and *ago9-2* (used in Stroud et al. 2013) to generate the triple *ago* mutant. *nrd1a-3* was described before (Herr et al. 2005). *drm1 drm2* (Stroud et al. 2013), *suvh2 suvh9* (Johnson et al. 2008),

svvh4 svvh5 svvh6 (Stroud et al. 2014), and *svvh2 svvh4 svvh9* (Johnson et al. 2008) were kindly provided by Steve E. Jacobsen.

Bisulfite sequencing and DNA methylation analysis

DNA was isolated from 2000–3000 mature embryos using the Quick-DNA microprep Plus kit (Zymo Research), according to the manufacturer's instructions. DNA (150 ng) was sheared using a Bioruptor to sizes ranging from 200–600 bp. The fragments were ligated with NEBNext methylated adaptors (New England Biolabs) and AMPure bead-purified (Beckman Coulter) before being treated with the EZ DNA Methylation-Lightning kit (Zymo Research), according to the manufacturer's instructions. Converted DNA was eluted with 20 μ L of buffer, and 10 μ L was used for amplification with uracil-tolerant polymerase (KAPA Biosystems) with NEBNext index containing primers.

For developing seeds, DNA was isolated from 250–500 young seeds at the mature embryo stage using the plant/seed DNA miniprep kit (Zymo Research). Five nanograms of DNA was converted and processed into sequencing libraries with the Zymo Pico Methyl-seq library preparation kit, according to the manufacturer's instructions. A single sample was made and analyzed for each genotype.

Libraries from each genotype were pooled and sequenced on a NextSeq 500 platform (Illumina) with single-end 75-nt runs at high output. Adaptor sequences were removed using Trimmomatic (Bolger et al. 2014); trimmed reads were then aligned on the *Arabidopsis* TAIR10 genome using Bismark (Krueger and Andrews 2011). Alignment files were then analyzed with the Methpipe programs (Song et al. 2013) for quality assessment, and biological replicates (embryos) were compared to assess reproducibility. The replicates were then merged to give more statistical power to the DMR analysis. In short, only cytosines covered more than five times were considered; the cytosines were separated in three contexts (CG, CHG, and CHH), and the CHH methylation was used to define hypomethylated DMRs in mutant backgrounds. The algorithm (Song et al. 2013) used a hidden Markov model to identify methylated regions in each data set, these regions were then compared between data sets (mutants compared with WT) to identify regions with a minimum of 10 methylated cytosines, at least five of which were hypomethylated in the mutant. DMRs hypomethylated in *ago4 ago6 ago9* compared with wild type were defined as small RNA dependent, and DMRs hypomethylated in *svvh4 svvh5 svvh6* compared with wild type were defined as small RNA independent. Regions appearing in both sets were discarded; all distributions were compared with the WT distribution using the Kolmogorov–Smirnov test with R.

ChIP-seq and analysis

The detailed ChIP protocol is in the Supplemental Material. For sequencing, DNA fragments were made into libraries using the NEBNext Ultra II FS kit (New England Biolabs). Libraries were pooled and sequenced on a NextSeq 500 platform (Illumina) with paired-end 150-nt runs at high output (developing seeds) or single-end 75-nt runs at high output (mature embryos). Adaptor sequences were removed using Trimmomatic (Bolger et al. 2014); trimmed reads were then aligned on the *Arabidopsis* TAIR10 genome using Bowtie2 (Langmead and Salzberg 2012). Duplicated reads are removed using the Picard Tool suite, and coverage was calculated and mapped to DMRs using Bedtools. All distributions were compared with the first WT sample distribution using the Kolmogorov–Smirnov test with R.

Immunofluorescence

Cytological technique was essentially described elsewhere (Ingouff et al. 2017). Young siliques of different stages were harvested and incubated in fixating solution for 5 min under vacuum and then for 2 h with rotation. Developing seeds were then dissected, embedded in acrylamide, digested, and incubated with antibodies anti-H3K9me2 (ab1220) and anti-H3K4me1 (Abcam ab8895) and secondary antibodies coupled with Alexa Fluor 488 and 564 (Thermo Fisher Scientific) together with DAPI and ProLong Gold antifade mountant (Thermo Fisher Scientific). Imaging was performed using either a Zeiss LSM780 or LSM880 with AiryScan detector.

Cloning

SUVH2 and SUVH9 genomic sequences were amplified and then were cloned into pDONR-221 and pDONR-207 (Thermo Fisher Scientific) using the BP enzyme mix according to the manufacturer's instructions. The mCitrine fragment was then added before the start codon using the HiFi DNA assembly master mix (New England Biolabs). PCR mutagenesis was achieved, amplifying the entry clone with mutation-containing primers and the KOD Xtreme hot start DNA polymerase (MilliporeSigma). Entry clones were transferred to the binary vector pMDC99 (Curtis and Grossniklaus 2003). Binary vectors were transformed into *Agrobacterium tumefaciens* strain GV3101 and used to transform *svvh2 svvh9* plants by floral dip (Clough and Bent 1998). All primers used are in Supplemental Table S3.

Accession numbers

The data sets generated in this study are available at Gene Expression Omnibus (GSE147245).

Competing interest statement

The authors declare no competing interests.

Acknowledgments

We thank E. Ernst and B. Berube for help with computational analysis; M. Ingouff for help with microscopy; S. Bhattacharjee, U. Ramu, and M. Callahan for technical help; and T. Mulligan for plant care. We thank F. Borges, B. Roche, and J.-F. Couture for fruitful discussions, as well as I. R. Henderson for critical reading of this manuscript. This work was supported by the Howard Hughes Medical Institute and grants from the National Institutes of Health (R01 GM067014) and the National Science Foundation Plant Genome Research Program (to R.A.M.). D.G. was supported by grants from the Agence Nationale de la Recherche (REMETH; ANR-15-CE12-0012), Chromobreed (ANR-18-CE92-0041), and Marie Skłodowska-Curie Actions (REP-658900-2). We acknowledge assistance from the Cold Spring Harbor Laboratory Shared Resources, which are funded in part by the Cancer Center (Support Grant 5PP30CA045508).

Author contributions: J.-S.P., D.G., J.C., and R.A.M. conceived the study. J.-S.P., R.P.H., D.G., J.C., and R.A.M. designed the research. J.-S.P., R.P.H., J.C., and D.G. performed the experiments and subsequent analysis. J.-S.P., D.G., J.C., and R.A.M. wrote the paper.

References

- Allshire RC, Madhani HD. 2018. Ten principles of heterochromatin formation and function. *Nat Rev Mol Cell Biol* **19**: 229–244. doi:10.1038/nrm.2017.119
- An YC, Goettl W, Han Q, Bartels A, Liu Z, Xiao W. 2017. Dynamic changes of genome-wide DNA methylation during soybean seed development. *Sci Rep* **8**: 7882. doi:10.1038/s41598-016-0028-x
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120. doi:10.1093/bioinformatics/btu170
- Bouyer D, Kramdi A, Kassam M, Heese M, Schnittger A, Roudier F, Colot V. 2017. DNA methylation dynamics during early plant life. *Genome Biol* **18**: 179. doi:10.1186/s13059-017-1313-0
- Calarco JP, Borges F, Donoghue MTA, Van Ex F, Jullien PE, Lopes T, Gardner R, Berger F, Feijó JA, Becker JD, et al. 2012. Reprogramming of DNA methylation in pollen guides epigenetic inheritance via small RNA. *Cell* **151**: 194–205. doi:10.1016/j.cell.2012.09.001
- Castel SE, Martienssen Ra. 2013. RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. *Nat Rev Genet* **14**: 100–112. doi:10.1038/nrg3355
- Cibrian-Jaramillo A, Martienssen RA. 2009. Darwin's 'abominable mystery': the role of RNA interference in the evolution of flowering plants. *Cold Spring Harb Symp Quant Biol* **74**: 267–273. doi:10.1101/sqb.2009.74.051

- Clough SJ, Bent AF. 1998. Floral dip: a simplified method for agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J* **16**: 735–743. doi:10.1046/j.1365-313x.1998.00343.x
- Curtis MD, Grossniklaus U. 2003. A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol* **133**: 462–469. doi:10.1104/pp.103.027979
- Dong KB, Maksakova IA, Mohn F, Leung D, Appanah R, Lee S, Yang HW, Lam LL, Mager DL, Schübeler D, et al. 2008. DNA methylation in ES cells requires the lysine methyltransferase G9a but not its catalytic activity. *EMBO J* **27**: 2691–2701. doi:10.1038/emboj.2008.193
- Du J, Johnson LM, Jacobsen SE, Patel DJ. 2015. DNA methylation pathways and their crosstalk with histone methylation. *Nat Rev Mol Cell Biol* **16**: 519–532. doi:10.1038/nrm4043
- Gallego-Bartolomé J, Liu W, Kuo PH, Feng S, Ghoshal B, Gardiner J, Zhao JMC, Park SY, Chory J, Jacobsen SE. 2019. Co-targeting RNA polymerases IV and V promotes efficient de novo DNA methylation in *Arabidopsis*. *Cell* **176**: 1068–1082.e19. doi:10.1016/j.cell.2019.01.029
- Henderson IR, Jacobsen SE. 2008. Tandem repeats upstream of the *Arabidopsis* endogene *SDC* recruit non-CG DNA methylation and initiate siRNA spreading. *Genes Dev* **22**: 1597–1606. doi:10.1101/gad.1667808
- Herr AJ, Jensen MB, Dalmay T, Baulcombe DC. 2005. RNA polymerase IV directs silencing of endogenous DNA. *Science* **308**: 118–120. doi:10.1126/science.1106910
- Ibarra CA, Feng X, Schoft VK, Hsieh T-F, Uzawa R, Rodrigues JA, Zemach A, Chumak N, Machlicova A, Nishimura T, et al. 2012. Active DNA demethylation in plant companion cells reinforces transposon methylation in gametes. *Science* **337**: 1360–1364. doi:10.1126/science.1224839
- Ingouff M, Selles B, Michaud C, Vu TM, Berger F, Schorn AJ, Autran D, Durme MV, Nowack MK, Martienssen RA, et al. 2017. Live-cell analysis of DNA methylation during sexual reproduction in *Arabidopsis* reveals context and sex-specific dynamics controlled by noncanonical RdDM. *Genes Dev* **31**: 72–83. doi:10.1101/gad.289397.116
- Jacob Y, Bergamin E, Donoghue Ma, Mongeon V, LeBlanc C, Voigt P, Underwood CJ, Brunzelle JS, Michaels SD, Reinberg D, et al. 2014. Selective methylation of histone H3 variant H3.1 regulates heterochromatin replication. *Science* **343**: 1249–1253. doi:10.1126/science.1248357
- Jiang H, Moreno-Romero J, Santos-González J, De Jaeger G, Gevaert K, Van De Slijke E, Köhler C. 2017. Ectopic application of the repressive histone modification H3K9me2 establishes post-zygotic reproductive isolation in *Arabidopsis thaliana*. *Genes Dev* **31**: 1272–1287. doi:10.1101/gad.299347.117
- Jing Y, Sun H, Yuan W, Wang Y, Li Q, Liu Y, Li Y, Qian W. 2016. SUVH2 and SUVH9 couple two essential steps for transcriptional gene silencing in *Arabidopsis*. *Mol Plant* **9**: 1156–1167. doi:10.1016/j.molp.2016.05.006
- Johnson LM, Law JA, Khattar A, Henderson IR, Jacobsen SE. 2008. SRA-domain proteins required for DRM2-mediated de novo DNA methylation. *PLoS Genet* **4**: e1000280. doi:10.1371/journal.pgen.1000280
- Johnson LM, Du J, Hale CJ, Bischof S, Feng S, Chodavarapu RK, Zhong X, Marson G, Pellegrini M, Segal DJ, et al. 2014. SRA- and SET-domain-containing proteins link RNA polymerase V occupancy to DNA methylation. *Nature* **507**: 124–128. doi:10.1038/nature12931
- Kawakatsu T, Nery JR, Castanon R, Ecker JR. 2017. Dynamic DNA methylation reconfiguration during seed development and germination. *Genome Biol* **18**: 171. doi:10.1186/s13059-017-1251-x
- Krueger F, Andrews SR. 2011. Bismark: a flexible aligner and methylation caller for bisulfite-seq applications. *Bioinformatics* **27**: 1571–1572. doi:10.1093/bioinformatics/btr167
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357–359. doi:10.1038/nmeth.1923
- Law JA, Vashisht AA, Wohlschlegel JA, Jacobsen SE. 2011. SHH1, a homeodomain protein required for DNA methylation, as well as RDR2, RDM4, and chromatin remodeling factors, associate with RNA polymerase IV. *PLoS Genet* **7**: e1002195. doi: 10.1371/journal.pgen.1002195
- Li S, Liu L, Li S, Gao L, Zhao Y, Kim YJ, Chen X. 2015. SUVH1, a Su(var)3-9 family member, promotes the expression of genes targeted by DNA methylation. *Nucleic Acids Res* **44**: gkv958. doi:10.1093/nar/gkv958
- Liu ZW, Shao CR, Zhang CJ, Zhou JX, Zhang SW, Li L, Chen S, Huang HW, Cai T, He XJ. 2014. The SET domain proteins SUVH2 and SUVH9 are required for Pol V occupancy at RNA-directed DNA methylation loci. *PLoS Genet* **10**: e1003948. doi: 10.1371/journal.pgen.1003948
- Liu W, Duttke SH, Hetzel J, Groth M, Feng S, Gallego-Bartolomé J, Zhong Z, Kuo HY, Wang Z, Zhai J, et al. 2018. RNA-directed DNA methylation involves co-transcriptional small-RNA-guided slicing of polymerase V transcripts in *Arabidopsis*. *Nat Plants* **4**: 181–188. doi:10.1038/s41477-017-0100-y
- Matzke MA, Kanno T, Matzke AJM. 2015. RNA-directed DNA methylation: the evolution of a complex epigenetic pathway in flowering plants. *Annu Rev Plant Biol* **66**: 243–267. doi:10.1146/annurev-arplant-043014-114633
- Rajkumar MS, Gupta K, Khemka NK, Garg R, Jain M. 2020. DNA methylation reprogramming during seed development and its functional relevance in seed size/weight determination in chickpea. *Commun Biol* **3**: 340. doi:10.1038/s42003-020-1059-1
- Schmitges FW, Prusty AB, Faty M, Stützer A, Lingaraju GM, Aiwazian J, Sack R, Hess D, Li L, Zhou S, et al. 2011. Histone methylation by PRC2 is inhibited by active chromatin marks. *Mol Cell* **42**: 330–341. doi:10.1016/j.molcel.2011.03.025
- Song Q, Decato B, Hong EE, Zhou M, Fang F, Qu J, Garvin T, Kessler M, Zhou J, Smith AD. 2013. A reference methylome database and analysis pipeline to facilitate integrative and comparative epigenomics. *PLoS One* **8**: e81148. doi:10.1371/journal.pone.0081148
- Stroud H, Greenberg MVC, Feng S, Bernatavichute YV, Jacobsen SE. 2013. Comprehensive analysis of silencing mutants reveals complex regulation of the *Arabidopsis* methylome. *Cell* **152**: 352–364. doi:10.1016/j.cell.2012.10.054
- Stroud H, Do T, Du J, Zhong X, Feng S, Johnson L, Patel DJ, Jacobsen SE. 2014. Non-CG methylation patterns shape the epigenetic landscape in *Arabidopsis*. *Nat Struct Mol Biol* **21**: 64–72. doi:10.1038/nsmb.2735
- To TK, Nishizawa Y, Inagaki S, Tarutani Y, Tominaga S, Toyoda A, Fujiyama A, Berger F, Kakutani T. 2020. RNA interference-independent reprogramming of DNA methylation in *Arabidopsis*. *Nat Plants* **6**: 1455–1467. doi: 10.1038/s41477-020-00810-z
- Walker J, Gao H, Zhang J, Aldridge B, Vickers M, Higgins JD, Feng X. 2017. Sexual-lineage-specific DNA methylation regulates meiosis in *Arabidopsis*. *Nat Genet* **50**: 130–137. doi:10.1038/s41588-017-0008-5
- Wang G, Köhler C. 2017. Epigenetic processes in flowering plant reproduction. *J Exp Bot* **68**: 797–807.
- Ye R, Chen Z, Lian B, Rowley MJ, Xia N, Chai J, Li Y, He X-J, Wierzbicki AT, Qi Y. 2016. A Dicer-independent route for biogenesis of siRNAs that direct DNA methylation in *Arabidopsis*. *Mol Cell* **61**: 222–235. doi:10.1016/j.molcel.2015.11.015
- Zhang Y, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, Nussbaum C, Myers RM, Brown M, Li W, et al. 2008. Model-based analysis of ChIP-seq (MACS). *Genome Biol* **9**: R137. doi: 10.1186/gb-2008-9-9-r137