SRA- and SET-domain-containing proteins link RNA polymerase V occupancy to DNA methylation

Lianna M. Johnson, Jiamu Du, Christopher J. Hale, Sylvain Bischof, Suhua Feng, Ramakrishna K. Chodavarapu, Xuehua Zhong, Giuseppe Marson, Matteo Pellegrini, David J. Segal, Dinshaw J. Patel & Steven E. Jacobsen

RNA-directed DNA methylation in Arabidopsis thaliana depends on the upstream synthesis of 24-nucleotide small interfering RNAs (siRNAs) by RNA POLYMERASE IV (Pol IV) and downstream synthesis of non-coding transcripts by Pol V. Pol V transcripts are thought to interact with siRNAs which then recruit DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2) to methylate DNA. The SU(VAR)-3-9 homologues SUVH2 and SUVH9 act in this downstream step but the mechanism of their action is unknown. Here, we show that genome-wide Pol V association with chromatin redundantly requires SUVH2 and SUVH9. Although SUVH2 and SUVH9 resemble histone methyltransferases, a crystal structure reveals that SUVH9 lacks a peptide-substrate binding cleft and lacks a properly formed S-adenosyl methionine (SAM)-binding pocket necessary for normal catalysis, consistent with a lack of methyltransferase activity for these proteins. SUVH2 and SUVH9 both contain SRA (SET- and RING-ASSOCIATED) domains capable of binding methylated DNA, suggesting that they function to recruit Pol V through DNA methylation. Consistent with this model, mutation of DNA METHYLTRANSFERASE 1 (MET1) causes loss of DNA methylation, a nearly complete loss of Pol V at its normal locations, and redistribution of Pol V to sites that become hypermethylated. Furthermore, tethering SUVH2 with a zinc finger to an unmethylated site is sufficient to recruit Pol V and establish DNA methylation and gene silencing. These results indicate that Pol V is recruited to DNA methylation through the methyl-DNA binding SUVH2 and SUVH9 proteins, and our mechanistic findings suggest a means for selectively targeting regions of plant genomes for epigenetic silencing.

To gain insights into the function of SUVH2/SUVH9, we solved the crystal structure of an amino-terminally truncated SUVH9 construct, which contains all the known functional domains (the SRA, pre-SET and SET domains) (Fig. 1a, Extended Data Table 1). The structure of SUVH9 is composed of three segments: a two-helix bundle towards the N terminus (residues 138–194), the SRA domain (residues 195–379) and the pre-SET/SET domains (residues 380–637). There are extensive inter-domain interactions that can stabilize the overall architecture of the protein (Fig. 1a and Extended Data Fig. 1b–g).

The SRA domain of SUVH9 resembles those of UHRF1 and SUVH5, which all contain the known functional domains (the SRA, pre-SET and SET domains) (Fig. 1b). Based on the SUVH5 SRA–mCHH (mCHH, methylated CHH, where H refers to A, T or C) DNA complex structure, we modelled a mCHH DNA into SUVH9 (Extended Data Fig. 2a). The DNA could be positioned in the nucleic-acid-binding cleft of the SRA domain without significant steric clashes and the proposed flipped-out 5-methylcytosine (5mC) base readily inserts into the binding pocket of the SRA domain.

Although SUVH9 contains histone methyltransferase pre-SET and SET domains similar to Dim5, G9a and GLP14–16, it shows neither detectable histone methyltransferase activity nor binding capacity for the SAM cofactor in vitro. SUVH9 and SUVH2 lack post-SET domains that are normally critical for cofactor- and peptide-substrate-binding; as well as catalysis. With the structure of human H3K9 methyltransferase GLP (Fig. 1c, d), the putative S-adenosylhomocysteine (SAH)-binding pocket and peptide-binding channel of SUVH9 are very open and incompletely formed (Fig. 1e, f) and cannot stably retain either a bound SAH molecule or the peptide substrate, especially in the absence of the stabilizing role of the post-SET domain. In summary, the SUVH9 structure is consistent with the demonstrated ability of SUVH9 to bind to methylated DNA, and supports the view that SUVH9 (and SUVH2) proteins encode inactive methyltransferase homologues.

We sought to further characterize the role of SUVH2 and SUVH9 in the RNA-directed DNA methylation (RdDM) pathway by determining their effect on siRNAs genome-wide. Most siRNAs are eliminated in Pol IV (nrpe1) mutants, whereas only some siRNAs are reduced in Pol V (nrpe1) mutants. We found that the suvh2 suvh9 double mutant reduced siRNA abundance at siRNA clusters that were dependent on both Pol IV and Pol V, but not at Pol V-independent clusters (Fig. 2a, b). Thus, the suvh2 suvh9 double mutant siRNA phenotype closely resembles that of a Pol V mutant. We next sought to determine whether SUVH2/SUVH9 might be involved in the production of non-coding transcripts by Pol V. At two characterized sites (IGN5 and P6) we found that suvh2 suvh9 reduced Pol V transcripts to the same extent as nrpe1 (Fig. 2c). Similar results were observed at the AtSN1 locus in suvh2 suvh9 (Fig. 2d). Test whether SUVH2/SUVH9 are required for Pol V chromatin occupancy, we used chromatin immunoprecipitation (ChiP) of a Flag-tagged NRPE1 (the largest subunit of Pol V). We observed only background levels of Pol V binding at IGN5 and IGN2 in suvh2 suvh9 compared to a sixfold enrichment in wild type (WT; Fig. 2d). We further analysed our ChiPs by next generation sequencing (ChiP-seq) and found that binding of Pol V at all previously identified sites was significantly decreased or eliminated in suvh2 suvh9 (Fig. 2e, f).

To determine the effect of SUVH2/SUVH9 on DNA methylation at defined Pol V binding sites, we used whole-genome bisulphite sequencing (BS-seq). As in nrpe1, CHH methylation at Pol V binding sites was eliminated in suvh2 suvh9 (Fig. 3a). We also analysed BS-seq of the single mutants suvh2 and suvh9 to determine whether SUVH2/SUVH9 act redundantly at all sites or have non-overlapping sites where they function. We found that suvh2 had a stronger effect than suvh9 at Pol V sites as well as at differentially methylated regions (DMRs) defined in either suvh2 or suvh9 single mutants, or in the suvh2 suvh9 double mutant (Extended Data Fig. 4a–c). These results indicate that SUVH2 and SUVH9 act redundantly throughout the genome to control RdDM.

Our results suggest that a reinforcing loop exists between DNA methylation and Pol V binding via SUVH2/SUVH9. To further test this model we used a mutation in the maintenance methyltransferase MET1 that eliminates CG methylation genome-wide and also reduces CHG and
To test directly whether SUVH2/SUVH9 may be sufficient to recruit Pol V, we used a zinc finger (ZF) to tether SUVH2 to an unmethylated epiallele of FWA, fwa-4. FWA is normally silenced owing to DNA methylation of tandem repeats in its promoter23. FWA epialleles have lost this methylation, leading to ectopic expression of FWA and a heritable late flowering phenotype23,24. Surprisingly, siRNAs are still observed in fwa epialleles, but are unable to direct DNA methylation25. We proposed that Pol V transcripts may be missing in fwa epialleles and that recruitment of Pol V by SUVH2 might therefore stimulate FWA methylation and silencing. To test this we transformed a ZF–SUVH2 fusion protein construct into the fwa-4 epiallele (Extended Data Fig. 10a). As negative controls, we also transformed a haemagglutinin (HA)-tagged SUVH2 line without the ZF (HA–SUVH2) as well as a construct in which the ZF was fused to KRYPTONITE (KYP/SUVH4) (a SUVH protein not required for RdDM).

Approximately 75% of the fwa-4 plants transformed with the ZF–SUVH2 (T1 generation) flowered early as compared to the parental fwa-4 line, suggesting silencing of FWA (Fig. 4a). The fwa-4 line transformed with HA–SUVH2 or ZF–KYP flowered at the same time as the fwa-4 parent, showing that the effect was specific to the ZF–SUVH2 fusion. Flowering time was measured in the T2 generation confirming these observations (Fig. 4b). The presence of the control ZF–KYP at FWA was shown by ChIP (Extended Data Fig. 10b); however, we were unable to detect the ZF–SUVH2, most likely owing to its instability or low abundance.

We next used bisulphite sequencing to determine whether FWA gene silencing was associated with DNA methylation. In wild type, DNA methylation was detected over a large region, whereas in both fwa-4 and transformants with ZF–KYP or HA–SUVH2, this region was devoid of DNA methylation (see ZF–KYP; Fig. 4c). In three independent ZF–SUVH2 T1 lines, DNA methylation was observed immediately around the Zn finger binding sites in all cytosine sequence contexts (Extended Data Fig. 10c). We analysed one of the lines (ZF–SUVH2–2) in the T2 and T3 generations using BS-seq and found that methylation extended approximately 150 base pairs in either direction from the binding sites and did not expand significantly between generations (Fig. 4c). FWA methylation and gene silencing were maintained in T2 plants that had segregated away the ZF–SUVH2 transgene (Extended Data Fig. 10d), indicating that targeting by SUVH2 is capable of inducing DNA methylation and gene silencing that can be maintained in the absence of the initial trigger.

To determine whether Pol V was present at FWA, we used NRPE1 ChIP. As expected, we found enrichment of Pol V at two known Pol V sites, IGN5 and IGN22, in the wild-type and fwa-4 lines, but not in nrpe1 mutant plants (Fig. 4d). At FWA, we found enrichment of Pol V
Figure 2 | SUVH2 and SUVH9 are required for Pol V-dependent siRNA production, chromatin binding and transcription. a, Boxplot (whiskers extend to 1.5 interquartile range (IQR)) of reads per kilobase per million reads (RPKM) values for 24-nucleotide siRNAs at previously defined siRNA clusters dependent on Pol IV (NRPD1), but not Pol V (NRPE1). 24-nt counts allow for up to 100 identical reads to be counted at any given position. * indicates a significant decrease (\( P < 2.2 \times 10^{-16} \), Mann–Whitney U-test). b, Similar to a for clusters defined as dependent on Pol IV and Pol V. c, Quantitative PCR with reverse transcription (RT–qPCR) of IGN22 and P6 relative to ACTIN7 and normalized to Columbia-0 (WT). Mean ± standard deviation (s.d.) of two biological replicates. d, Quantitative PCR (qPCR) of IGN22 and IGN5 from Flag ChIP shown as enrichment of IP/input relative to ACTIN7 in NRPE1–Flag/WT and NRPE1–Flag/suvh2 suvh9 lines. Mean ± s.d. of two biological replicates. e, Heat map of NRPE1 enrichment at defined NRPE1 sites determined by Flag ChIP-seq in either NRPE1–Flag/WT or NRPE1–Flag/suvh2 suvh9, with Flag ChIP in WT as negative control. f, Box plot of NRPE1 enrichment sites shown as Fig. 2e for NRPE1–Flag/WT and NRPE1–Flag/suvh2 suvh9.

in the wild type at both the promoter and transcript regions, but not in nrpe1 or fwa-4 (Fig. 4d). However, in the ZF–SUVH2-transformed fwa-4 plants we could now see enrichment of Pol V at FWA (Fig. 4d), indicating that SUVH2 is sufficient to recruit Pol V.

To look for a direct interaction between SUVH2/SUVH9 and Pol V, we queried several IP–mass spectrometry data sets from purifications of NRPE1–Flag, but failed to find any SUVH2/SUVH9 peptides. However, we did identify SUVH2 peptides in two independent mass spectrometry data sets from DRD1 purifications (Supplementary Table 2). DRD1 is a component of the DDR complex (also containing DMS3 and RDM1) which interacts with Pol V and is required for Pol V occupancy throughout the genome. The number of SUVH2 peptides observed was lower than those from the DMS3 and RDM1 proteins and also lower than the level of peptides of most Pol V complex components, indicating that the interaction between SUVH2 and DRD1 is weaker or more transient than the interaction between the DDR components or between DDR and Pol V. To confirm the interaction, we performed co-immunoprecipitation experiments with HA-tagged SUVH2 and Flag-tagged DRD1 in transgenic Arabidopsis plants and were able to detect HA–SUVH2 in a Flag–DRD1 pull-down (Fig. 4e). We were also able to detect Flag–DRD1 in a pull-down using HA–SUVH2 expressed in leaves of Nicotiana benthamiana and purified on HA magnetic beads (Extended Data 10e). These results confirm the IP–mass spectrometry observations and are consistent with a model in which SUVH2 acts indirectly via a transient interaction with DRD1 to recruit Pol V.

Because RdDM is triggered at genomic locations that synthesize both Pol IV-dependent siRNAs and Pol V-dependent non-coding transcripts, it is critical to understand the mechanisms that direct Pol IV and Pol V to chromatin. Furthermore, because RdDM functions to silence transposons that do not necessarily share consensus DNA sequences, a strong prediction is that Pol IV and Pol V will use epigenetic information in their targeting. We propose that SUVH2/SUVH9 serve as critical components of this targeting system by recruiting Pol V to DNA methylation during transformation of an unmethylated FWA gene into Arabidopsis. An interesting question is how an originally unmethylated DNA is first recognized as an RdDM target if SUVH2/SUVH9 DNA methylation-mediated Pol V recruitment cannot initially occur. One possibility is that shortly after the transgene is integrated into the genome, a permissive chromatin environment may allow RNA Pol V (or another polymerase) to produce surveillance transcripts that might serve to initiate RdDM through siRNA-mediated recruitment of DRM2. After this initial round of DNA methylation, SUVH2/SUVH9 could then bind methylated DNA...
and recruit Pol V for subsequent rounds of transcription. In this way, a self-reinforcing loop would be established in which pre-existing DNA methylation is required for the maintenance of RNA-directed DNA methylation.

METHODS SUMMARY

All biological materials and extended methods are described in detail in the Methods. siRNA libraries were made from flowers of the indicated lines as described previously. ChIP-seq, BS-seq, co-immunoprecipitation, protein purification and crystallization, and data analysis are described in the Methods. Supplementary Table 1 provides crystallography data. Supplementary Table 2 provides IP-mass spectrometry data and Supplementary Table 3 provides primers used in this study.

Online Content Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 29 May; accepted 2 December 2013.
Published online 22 January 2014.

1. Pélissier, T. & Wassernegger, M. A DNA target of 30 bp is sufficient for RNA-directed DNA methylation. RNA 6, 55–65 (2000).
2. Aufsatz, W., Mette, M. F., van der Winden, J., Matzke, A. J. & Matzke, M. RNA-directed DNA methylation in Arabidopsis. Proc. Natl Acad. Sci. USA 99 (Suppl. 4), 16499–16506 (2002).
3. Law, J. A. & Jacobsen, S. E. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nature Rev. Genet. 11, 204–220 (2010).
4. Pontier, D. et al. Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases IV in Arabidopsis. Genes Dev. 19, 2030–2040 (2005).
5. Pikkaard, C. S., Haag, J. R., Ream, T. & Wierzbicki, A. T. Roles of RNA polymerase IV in gene silencing. Trends Plant Sci. 13, 390–397 (2008).
6. Law, J. A. et al. A protein complex required for polymerase V transcripts and RNA-directed DNA methylation. Proc. Natl Acad. Sci. USA 108, 951–956 (2010).
BioSequencing Core Facility. This work was supported by the Abby Rockefeller Mauze Trust and the Maloris and STARR foundations to D.J.P., and NIH grant GM60398 to S.E.J. C.J.H. is supported by the Damon Runyon post-doctoral fellowship, S.B. is supported by a post-doctoral fellowship of the Swiss National Science Foundation, S.F. is a Special Fellow of the Leukemia & Lymphoma Society, and X.Z. is supported by Ruth L. Kirschstein National Research Service grant F32GM096483-01. S.E.J. is an Investigator of the Howard Hughes Medical Institute.

Author Contributions Protein purification and crystallography were designed and performed by J.D., G.M. and D.J.P., siRNA libraries were made by S.B., whole-genome bisulphite sequencing was performed and analysed by S.F., L.M.J., C.H., R.K.C. and M.P., Pol V transcription assays were performed by X.Z., ChIP-seq was performed and analysed by L.M.J. and C.H., cloning and flowering time experiments were performed by L.M.J., the Zn finger was designed by D.J.S. and S.E.J., pull-down experiments were performed by L.M.J. and S.B., the manuscript was written by L.M.J., J.D., D.J.P. and S.E.J., and S.E.J. participated in all experimental design.

Author Information The coordinates and structure factors of SUVH9 have been deposited in the RCSB Protein Data Bank with the accession code 4NJ5. Sequencing data have been deposited at GEO (GSE52041). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.E.J. (jacobsen@ucla.edu) or D.J.P. (pateld@mskcc.org).
METHODS

Biological material. All plants used in this study were of the Arabidopsis thaliana Columbia-0 (Col-0) accession, with WT referring to the parental strain. The nrp1–12 T-DNA (SALK_033852), mett1-3 and the suvh2 suvh9 double mutant lines were described previously13,15,16. NRP1–Flag transgenic plants were crossed to the suvh2 suvh9 double mutant and homozygous F2 plants were identified. The fwa-4 ep allele was isolated from a mett1 segregating population.

Experimental procedures in Arabidopsis. siRNA libraries were made from flowers of the indicated lines as described previously19 and Pol V transcription assays were performed as described previously20. ChiPs were performed as described previously21 except that ground tissue was cross-linked with formaldehyde in the following buffer: 10 mM HEPEs pH 8.0, 1% sucrose, 5 mM KCl, 5 mM MgCl2, 5 mM EDTA, and 0.6% Triton X-100. Libraries for NRP1E ChiP-seq were generated using the Ovation Ultralow DR Multiplex System (NuGen). Bisulphite sequencing followed by PCR amplification and cloning of FWA fragments was done using EZ DNA Methylation-Gold kit (Zymo Research) as described previously22. BS-seq libraries were generated as previously reported23 and all libraries were sequenced using the HiSeq 2000 platform following manufacturer instructions (Illumina) at a length of 50 bp. Supplementary Table 3 provides primers used in this study.

Co-immunoprecipitation experiments. Nicotiana benthamiana was infiltrated with plJ322 (ZF–HA–SUVH2 in JP7468) and leaves were collected after 3 days. Co-immunoprecipitation experiments were performed starting with 2 g of flower extracts from plants expressing HA–SUVH2, Flag–DRD1 or T2 plants expressing both HA–SUVH2 and Flag–DRD1. Extracts were made as described above and 100 μl of Flag-magnetic beads (Sigma) were added and incubated with rotation at 4 °C for 45 min. Washes and western blots were performed as described above.

Data analysis. Bisulphite-seq (BS-seq), small RNA-seq and ChiP-seq data were aligned to the TAIR10 version of the Arabidopsis thaliana reference genome using BS-seeker for BS-seq data and Bowtie for the small RNA-seq and ChiP-seq data. For BS-seq up to 2 mismatches were allowed, whereas for endogenous NRP1E ChiP and small RNA sequencing only 1 mismatch was allowed. For NRP1E–Flag ChiP, only perfect matches were used and the TAIR8 version of the genome was used to match an earlier report24. For BS-seq and ChiP-seq, only uniquely mapping reads were used. For small RNA-seq, reads were normalized as previously described25.

The small RNA clusters used in Fig. 2a, b were previously defined as “pol-iv only” and “shhl/drm2/pol-v” dependent clusters, respectively26. For NRP1E–Flag ChiP, enrichment over NRP1E sites was normalized to Flag ChiP in Columbia (negative control). For ChiP-seq using the endogenous Pol V antibody, the read density over Pol V peaks was calculated using the reads per kilobase per million reads (RPKM) metric. For the purposes of Figs 2e, f, 3a, b and d and Extended Data Fig. 4b, Pol V sites were used as previously defined27. For the purposes of Fig. 3e, Pol V enrichment was compared to mett1 hypermethylated CHH sites28. To ensure these sites were comparable in genomic character, all sites were expanded to be 1,000 bp in length, defined by extending the borders of sites ≥ 500 bp from the previously defined midpoint. For mett1 hypermethylated CHH sites, we applied a size filter to the previously described sites29 so that we retained only sites larger than 100 bp. This was in an effort to filter out sites of spurious CHH methylation and retain bona fide hypermethylation sites in the mett1 mutant.

DMR definitions. Differentially methylated regions (DMRs) were defined for the suvh2 suvh9, suvh2 and suvh9 libraries by comparing methylation profiles to a suite of wild-type libraries as previously described30. The peptide containing 6 Zn fingers was designed as described previously31,32 and cloned into pUC57 with XhoI sites at both ends (Geneviz). This XhoI fragment was excised and inserted into the unique XhoI site located upstream of the HA–SUVH2 construct33, which encodes an HA–SUVH2 construct34. This ZF–SUHV2 fragment was then recombined into JP7268 and introduced into fwa-4 using agrobacterium-mediated transformation. The ZF was also fused to the N terminus of the KRYPTONITE H3K9 methyltransferase and transformed into fwa-4.

Protein preparation. The N-terminal truncated SUVH9 (residues 134–650) was cloned into a 1PlasticHT B vector (Invitrogen), which fuses a hexa-histidine tag followed by a tobacco etch virus protease (TEV) cleavage site to the N terminus of the target gene. The plasmid was transformed into E. coli strain DH10Bac (Invitrogen) to generate the bacmid. Bacterial virus was generated by transfecting S9 cells with the bacmid following standard Bac-to-Bac protocol (Invitrogen). The collected virus was subsequently used to infect the suspended H5 cell for recombinant protein expression. The recombinant protein was first purified using nickel affinity chromatography column (GE Healthcare). The hexa-histidine tag was cleared by TEV protease. The target protein was further purified using a Q Sepharose column and a Superdex 200 gel filtration column (GE Healthcare). The purified protein was concentrated to 15 mg ml−1 and stored at −80 °C.

Crystallization. Before crystal screening, the SUVH9 protein was mixed with the putative cofactor S-adenosyl-L-homocysteine (SAH) in a molar ratio of 1:3 or with SAH and a histone H3(1–15) peptide in a molar ratio of 1:3:3 at 4 °C for 2 h. Crystallization of the SUVH9 was carried out at 20 °C using the hanging drop vapour diffusion method by mixing 1 μl protein sample at a concentration of 11.5 mg ml−1 and 1 μl reservoir solution and equilibrating against 500 μl reservoir solution. SUVH9 was crystallized in the condition of 0.2 M potassium thiocyanate, 0.1 M Bis-Tris propane, pH 7.5, and 20% PEG 3350 in free form, as well as in the presence of SAH or in the presence of SAH and H3(1–15) peptide. Square-shaped crystals appeared in 2 weeks. All the crystals were dipped into reservoir solution supplemented with 15% glycerol and flash cooled into liquid nitrogen for diffraction data collection. The diffraction data were collected at beamline X29A, National Synchrontron Light Source (NSLS) at Brookhaven National Laboratory (BNL), New York at the zinc peak wavelength. The data were indexed, integrated and further scaled with the program HKL200035. The statistics of the diffraction data are summarized in Supplementary Table 1.

For investigating the molecular mechanism for recognition of mCHH containing DNA by SUVH9, we have extensively screened SUVH9 in complex with DNAs as a function of length and termini. We used DNA lengths from 8 bp to 20 bp, with a blunt end or with one or two nucleotides overhanging at either 5’ ends or 3’ ends, all of which containing a central mCHH site. Unfortunately, we were not able to get any hits on the SUVH9–DNA complex. The methylated and unmethylated oligonucleotides were purchased from Keck Oligonucleotide Synthesis Facility at Yale University and Life Technologies, respectively.

Structure determination and refinement. The structure of SUVH9 in the free state was solved using single–wavelength anomalous dispersion method implemented in the program Phenix36. The model building was carried out using the program Coot37 and structural refinement using the program Phenix38. Throughout the refinement, a free R factor was calculated using 5% random chosen reflections. The stereochemistry of the structural model was assessed using the program Procheck39. The statistics of the refinement and structure models are shown in Supplementary Table 1. Although we have also collected data for the crystals of SUVH9 in the presence of SAH and in the presence of SAH and H3(1–15) peptide, we observed no electron density for either the SAH moiety or the peptide. Our data are indicative that crystals of SUVH9 in presence of SAH and in the presence of SAH and H3(1–15) peptide are indeed the crystal of free form SUVH9. All the molecular graphics were generated with the program PyMol (DeLano Scientific LLC). The sequence alignment was conducted using the program ClustalX240 and illustrated using the ESPript server41.

29. Taric, M. et al. Erasure of CpG methylation in Arabidopsis alters patterns of histone H3 methylation in heterochromatin. Proc. Natl Acad. Sci. USA 100, 8823–8827 (2003).
30. El-Shami, M. et al. Reiterated WG/GW motifs form functionally and evolutionarily conserved ARGONAUTE-binding platforms in RNA-related components. Genes Dev. 21, 2539–2544 (2007).
31. Berntavichute, Y. V., Zhang, X., Cokus, S., Pellegrini, M. & Jacobsen, S. E. Genome-wide association of histone H3 lysine nine methylation with CHG DNA methylation in Arabidopsis thaliana. PLoS Genet. 9, e1003246 (2013).
32. Kinoshi, Y. et al. Control of FWA gene silencing in Arabidopsis thaliana by SINE-related direct repeats. Plant J. 49, 38–45 (2007).
33. Kob, A. F. et al. Site-directed genome modification: nucleic acid and protein modules for targeted integration and gene correction. Trends Biotechnol. 23, 399–406 (2005).
34. Segal, D. J. et al. Evaluation of a modular strategy for the construction of novel polyadecylnzinc finger DNA-binding proteins. Biochemistry 42, 2137–2148 (2003).
35. Ochowinski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. 276, 307–326 (1997).
36. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D 66, 213–221 (2010).
38. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).

39. Laskowski, R. A., Macarthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.*** **26**, 283–291 (1993).

40. Larkin, M. A. et al. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).

41. Gouet, P., Courcelle, E., Stuart, D. I. & Metoz, F. ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**, 305–308 (1999).
Extended Data Figure 1 | Interdomain interactions of SUVH9. a, Colour-coded schematic representation of full length SUVH9 and the N-terminally truncated construct used for crystallization. The hydrophobic interactions and charged interactions within the two-helix bundle shown in two alternate views rotated by 180 degree. Residues involved in inter-helix hydrophobic interactions are highlighted in yellow. b, The N-terminal part of the first α-helix forms charged and hydrogen bonding interactions with the SRA domain and the SET domain. The interacting residues are shown in stick representation and the hydrogen-bonding interactions are shown with dashed red lines. d, The C-terminal part of the first α-helix exhibits extensive hydrophobic interactions with the SRA domain and the pre-SET/SET domains. The tip of a long loop from the SET domain covers over the first helix and forms hydrophobic interactions with it. e, The second α-helix forms some interactions with the SRA domain. f, The SRA domain forms a hydrophobic core that interacts with the pre-SET/SET domains. g, A long insertion loop of SUVH9 SET domain (highlighted in magenta) is enriched with hydrophobic residues and forms extensive hydrophobic interactions with the two-helix bundle, the pre-SET and SET domains.
Extended Data Figure 2 | SUVH9 SRA and pre-SET/SET domains. a, A model positioning the mCHH DNA to the active site of SUVH9 SRA domain following superposition of the structures of the SUVH5 SRA–mCHH complex (PDB code 3Q0F) and SUVH9 in the free state. SUVH9 domains are depicted with the same colour-coding as in Fig. 1a and the modelled DNA is coloured in yellow. The DNA fits well into the SRA domain without significant steric clashes. Some surrounding residues on the second z-helix of the two-helix bundle, which can potentially be involved in the binding to the DNA, are highlighted in a stick representation. b, A stereo view of the superposition of the structure of SUVH9 in the free state and the structure of human GLP catalytic fragment complexed with SAH (PDB code 2IGQ). The GLP pre-SET and SET domains are coloured in silver and its post-SET domain is coloured in cyan. The zinc-binding motif of GLP post-SET domain and SET domain, the bound SAH molecule, and the corresponding Thr 597 of SUVH9 are highlighted in a stick representation.
Extended Data Figure 3 | Structure-based sequence alignment of SUVH family proteins from Arabidopsis. The secondary structural elements of SUVH9 are labelled on the top of the sequence alignment. The domain boundaries are marked on the top and depicted with colour-coding as in Fig. 1a. Conserved residues involved in the interaction with flipped 5mC base and the DNA backbone available from the published SUVH5–DNA complex structures are highlighted with cyan circles and blue hexagons, respectively. The insertions in the SET domains are highlighted with a purple box. The zinc-coordinating Cys residues are highlighted with black stars in the SET domain and grey stars in the post-SET domain. Two-tyrosine residues that are conserved and normally important for enzymatic activity are highlighted with red dots.
Extended Data Figure 4 | SUVH2 and SUVH9 act redundantly genomewide. a, Metaplots of CHH methylation over DMRs identified in the various SUVH mutants. b, Metaplots of CHH methylation over Pol V binding sites. c, Venn diagram detailing the overlaps between CHH hypo-methylated regions in SUVH mutants.
Extended Data Figure 5 | Pol V occupancy in WT versus met1. Chromosome 1 showing Pol V ChIP in WT versus met1 as mapped over TAIR10 (green genes, red transposable elements (TEs)).
Extended Data Figure 6 | Screen shot of Pol V binding in WT versus met1. An example of reduced Pol V binding in met1 at sites that become hypomethylated.
Extended Data Figure 7 | Screen shot of Pol V binding in WT versus met1. Reduction in Pol V binding in a met1 hypomethylated site.
Extended Data Figure 8 | Screen shot of Pol V binding at a hyper-CHH methylated site in WT versus met1. An example of Pol V being redistributed to regions that gain methylation in met1.
Extended Data Figure 9 | Pol V binding at hyper-CHH methylated site that is also transcribed. Strong Pol V binding was detected at regions in the genome that not only retained high levels of non-CG methylation, but also were transcriptionally activated in \textit{met1}.
Extended Data Figure 10 | ZF–SUVH2 construct stably recruits Pol V to FWA through a direct interaction with DRD1. **a**, Top, diagram of SUVH2 with Zn finger (ZF) inserted immediately before the HA tag. Bottom, schematic of FWA gene showing the two small and two large repeats (blue arrows), the regions amplified by PCR (promoter and transcript, green lines), the start and direction of transcription (red arrow), and the sites bound by the ZF (indicated by two orange arrows). **b**, Flag-ChIP in WT versus ZF–KYP (Flag-tagged) showing enrichment at FWA in both the promoter and transcript region (see above). **c**, Per cent methylation at each C in the FWA repeat region from three individual T1 plants. Per cent methylation was determined from 20–25 clones of bisulphite-treated DNA. **d**, BS-seq of FWA from a Basta-resistant line containing the ZF–SUVH2 transgene and two Basta-sensitive siblings which had lost the ZF–SUVH2 transgene. **e**, Pull-down of DRD1–Flag with ZF–SUVH2. Input, DRD1–Flag extract from Arabidopsis; Beads-mock, elution from DRD1–Flag pull-down using HA-magnetic beads pre-bound with Nicotiana benthamiana extract; Beads–ZF–SUVH2, elution from DRD1–Flag pull-down using HA-magnetic beads pre-bound with Nicotiana benthamiana ZF–SUVH2 extract. Top, Flag blot; bottom, HA blot.
CORRIGENDUM

doi:10.1038/nature21398

Corrigendum: SRA- and SET-domain-containing proteins link RNA polymerase V occupancy to DNA methylation

Lianna M. Johnson, Jiamu Du, Christopher J. Hale, Sylvain Bischof, Suhua Feng, Ramakrishna K. Chodavarapu, Xuehua Zhong, Giuseppe Marson, Matteo Pellegrini, David J. Segal, Dinshaw J. Patel & Steven E. Jacobsen

Nature 507, 124–128 (2014); doi:10.1038/nature12931

In this Letter, we showed that tethering SUVH2 with a zinc finger to the FWA gene was sufficient to cause DNA methylation and gene silencing. We have discovered that we inadvertently used the SUVH9 protein, not the SUVH2 protein, in this zinc finger fusion. SUHV2 and SUVH9 are redundant homologues that have similar roles in RNA-directed DNA methylation. The mistake affects Fig. 4a–d and Extended Data Fig. 10a, c and d of the original Letter, in which the labels ‘ZF–SUVH2’ should be ‘ZF–SUVH9’; the data in these figures are correct. In the abstract, ‘SUVH2’ should be ‘SUVH9’ in the following sentence, to read: “Furthermore, tethering SUVH9 with a zinc finger to an unmethylated site is sufficient to recruit Pol V and establish DNA methylation and gene silencing.” The results and conclusions are not affected. These errors have not been corrected online.