www.sciencemag.org/content/362/6419/1182/suppl/DC1

Supplementary Materials for

A DNA methylation reader complex that enhances gene transcription

C. Jake Harris*, Marion Scheibe*, Somsakul Pop Wongpalee, Wanlu Liu, Evan M. Cornett, Robert M. Vaughan, Xueqin Li, Wei Chen, Yan Xue, Zhenhui Zhong, Linda Yen, William D. Barshop, Shima Rayatpisheh, Javier Gallego-Bartolome, Martin Groth, Zonghua Wang, James A. Wohlschlegel, Jiamu Du, Scott B. Rothbart, Falk Butter†, Steven E. Jacobsen†

*These authors contributed equally to this work.
†Corresponding author. Email: f.butter@imb-mainz.de (F.B.); jacobsen@ucla.edu (S.E.J.)

Published 7 December 2018, *Science* **362**, 1182 (2018)
DOI: 10.1126/science.aar7854

This PDF file includes:

Materials and Methods:

Plant materials:

Unopened floral bud tissue samples were grown in 16-hour light, 8-hour dark cycle. Seedling tissue samples were grown under constant light. *Arabidopsis thaliana* Col-0 ecotype T-DNA mutants used; *SUVH1* (AT5G04940, SALK_003675), *SUVH3* (AT1G73100, SAIL_401b_D01), *DNAJ1* (AT5G64360, SALK_021078, aka *eip9-1*, described in (*32*)), *DNAJ2* (AT2G01710, SALK_057430), *NRPE1* (AT2G40030, SALK_029919, *nrpe1-11*), *DRM1/2* (AT5G15380, *drm1-2*; AT5G14620, *drm2-2*), *NRPD1* (AT1G63020, *nrpd1-4*). Transgenic plant materials were generated using the floral dip method (*33*) with *Agrobacterium* strain AGLO carrying the binary plasmid.

Comparative Interactomics:

Step One - Nuclear Extraction: Samples, processed in batches of 2 grams of unopened floral bud tissue per genotype, were snap frozen in liquid nitrogen and disrupted in 35 ml homogenizer tubes (Retsch) using a TissueLyser (Qiagen) for 2 minutes at 30 Hz. Material was mixed with 25 ml of ice cold Honda Buffer [2.5% Ficoll 400, 5% Dextran T40, 0.4 M Sucrose, 25 mM Tris pH 7.4, 10 mM $MgCl_2$, 0.035% β-mercaptoethanol, EDTA-free Protease Inhibitor *Cocktail* (Roche cOmplete)] and filtered through 2 layers of Miracloth (Sigma-Aldrich*)* followed by a 40 µM cell strainer (Falcon). Triton X-100 (Sigma-Aldrich) was added to a final concentration of 0.5% and samples were incubated on ice for 15 minutes followed by centrifugation at 2000xg for 20 minutes at 4ºC. Pellets were resuspended in Honda Buffer supplemented with 0.1% Triton X-100 and spun at 2000xg for 10 minutes at 4ºC. Pellets were re-suspended in 2.5 ml of Extraction Buffer [2M

Hexylene Glycol, 20 mM PIPES-KOH pH 7.0, 20 mM $MgCl_2$, 5mM β-mercaptoethanol, EDTA-free Protease Inhibitor Cocktail (Roche cOmplete), 2.5 mM benzamidine hydrochloride, 0.5 mM phenylmethylsulfonyl fluoride] and overlaid on top of 4 ml of 30% and 4 ml of 80% percoll gradients [30/80% Percoll (Sigma-Aldrich), 0.5M hexylene glycol, 5 mM PIPES-KOH pH 7.0, 10 mM $MgCl_2$, 1% Triton X-100, 5 mM β-mercaptoethanol, EDTA-free Protease Inhibitor Cocktail (Roche cOmplete), 1 mM benzamidine hydrochloride, 0.2 mM phenylmethylsulfonyl fluoride] in 15 ml round bottom Oak Ridge glass test tubes (Thermo) and spun in a swing rotor at 2000xg for 30 minutes at 4ºC. Nuclear extract from between the 30% and 80% percoll gradient layers was resuspended in 4 ml of Gradient Buffer [0.5M Hexylene Glycol, 5 mM PIPES-KOH pH 7.0, 10 mM $MgCl_2$, 1% Triton X-100, 5 mM β-mercaptoethanol, EDTA-free Protease Inhibitor Cocktail (Roche cOmplete), 1 mM benzamidine hydrochloride, 0.2 mM phenylmethylsulfonyl fluoride] and overlaid onto 4 ml of 30% percoll gradient and spun in a swing rotor at 2000xg for 10 minutes at 4ºC. We repeated resuspension of the pellet in Gradient Buffer and underlay with 30% percoll gradient buffer until the green hue was no longer visible (nuclei appear whitish brown). Samples were resuspended in 1 ml of Gradient Buffer in a 1.5 ml Eppendorf tube and spun at 1,000xg for 10 minutes at 4ºC. Nuclei were lysed by resuspending nuclear pellet volume in two volumes of Buffer C+ [420 mM NaCl, 20 mM Hepes KOH pH 7.9, 20% (v/v) glycerol, 2 mM $MgCl_2$, 0.2 mM EDTA, 0.1% IGEPAL CA-630, EDTA-free Protease Inhibitor Cocktail (Roche cOmplete), 0.5 mM DTT] and rotated at 4ºC for 1 hour. Samples were centrifuged in 15 ml round bottom Oak Ridge glass test tubes (Thermo) at 38,000xg for 30 minutes at 4ºC to separate nuclear protein extract from insoluble chromatin. Glycerol to a final concentration of 10% (v/v) was added, and extracts were snap frozen in liquid nitrogen for storage.

Step Two - DNA-pulldown: To design probes, we chose sequences based on naturally methylated *Arabidopsis* loci where methylation in CG, CHG and CHH contexts have a particularly well-established role in affecting transcription(*34–37*). Long probes, 75nt in length, were used to maximize protein binding surface and to more accurately reflect the in vivo chromosomal context (see Fig. S2). DNA oligonucleotides were chemically synthesized, HPLC purified and checked by mass spectrometry (Metabion). 2.5 nmol single-stranded forward and reverse oligonucleotides were annealed, phosphorylated with 25 units PNK (Thermo) and ligated overnight by 20 units T4 ligase (Thermo). Ligation success was monitored by agarose gel electrophoresis. DNA strands were purified by chloroform-phenol extraction using 25:24:1 (v/v) Chloroform-Phenol-Isoamyl Alcohol (Sigma-Aldrich) and precipitated with ethanol (Sigma-Aldrich). After resolubilization in water, the DNA strands were incubated with desthiobiotin-dATP (Jena Bioscience) and 30 units Klenow fragment (Thermo) at 37°C overnight. Unreacted biotinylated nucleotides were removed by a G50 spin column (GE Healthcare) following manufacturer's instruction. 25 µg of biotinylated oligonucleotides were coupled to 750 µg of Streptavidin Dynabeads MyOne C1 (Thermo) at room temperature for 30 min. Unbound DNA was removed by 3 washes and the DNA-coupled beads were incubated with 200 µg of *A. thaliana* nuclear extract in PBB buffer (150 mM NaCl, 50 mM Tris/HCl pH 7.6, 5 mM $MgCl_2$, 0.5% Igepal CA-630, EDTA-free Protease Inhibitor Cocktail (Roche cOmplete)). This step was performed in quadruplicates for each DNA bait. After incubation at 4°C with slight agitation for 1.5 hours, beads were washed three times and then heated in 1x LDS buffer (Thermo) at 80°C for 10 min.

Step Three - Mass spectrometry sample preparation: Samples were loaded on a NuPAGE 4-12% Bis-Tris protein gel (Thermo) and run at 180V for 8 min. The gel was stained with coomassie protein stain. Each gel lane was sliced, minced and transferred to an Eppendorf tube. The gel pieces

were destained with 50% ethanol in 25mM ammonium bicarbonate buffer, reduced with 10 mM DTT (Sigma-Aldrich) at 56°C for 1 hour, alkylated with iodoacetamide and digested with mass spectrometry grade trypsin (Sigma-Aldrich) at 37°C overnight. The tryptic peptides were desalted using a self-made C_{18}-StageTip and stored until MS measurement.

Step Four - Mass spectrometry acquisition: The sample was analyzed by nanoflow liquid chromatography on an EASY-nLC 1000 system (Thermo) coupled online to a Q Exactive Plus mass spectrometer (Thermo). Peptides were separated on a C18-reversed phase column (15 cm long, 75 µm inner diameter, packed in-house with ReproSil-Pur C18-AQ 1.9 µm resin (Dr. Maisch GmbH)). We used a 107 min gradient from 2% to 60% acetonitrile in 0.5% formic acid at a flow of 200 nl/min. The Q Exactive Plus was operated with a Top10 MS/MS data-dependent acquisition method using HCD fragmentation.

Step Five - MS data analysis: The raw files were processed with MaxQuant(*38*) (version 1.5.2.8) using the TAIR10 peptide database (35,386 entries). Carbamidomethylation was set as fixed modification while methionine oxidation and protein N-acetylation were considered as variable modifications. The search was performed with an initial mass tolerance of 7 ppm mass accuracy for the precursor ion and 20 ppm for the MS/MS spectra in the HCD fragmentation mode. Search results were filtered with a false discovery rate of 0.01 at the peptide and protein level. The match between run option and LFQ quantitation were activated. Prior to statistical analysis known contaminants and reverse hits were removed.

Fluorescence polarization (FP)

6xHis-MBP-SUVH1 (pD454) WT or Y277A were transformed into BL21(DE3) *E. coli*. Starter cultures were grown overnight in 120 mL LB media (Caisson) with 100 µg/mL ampicillin at 37ºC

with shaking at 250 RPM. The next morning, six 4 L baffled flask containing 2 L of LB media were inoculated with 20 mL of starter culture. Cultures were grown at 37ºC with shaking at 160 RPM until their OD_{600} was between 0.6 and 0.8. The temperature was lowered to 16ºC, and IPTG was added to a final concentration of 0.5 mM. Cultures were incubated overnight at 16ºC with shaking at 160 RPM. Bacteria were harvested by centrifugation and pellets were resuspended in 150 mL of lysis buffer (50 mM Tris pH 8.0, 250 mM NaCl, 20 mM imidazole, 1 mM DTT, 1 mM PMSF). A dash of lyophilized lysozyme (VWR Life Science) was added, mixed, and incubated on ice for an hour. Next, the cells were lysed using an APV laboratory homogenizer (9,000 psi for 5 min). Lysates were cleared by centrifugation at 38,000 g for 30 min at 4 ºC. Cleared lysates were loaded onto a HisPrep FF 16/10 column (GE Healthcare) and washed with 10 column volumes of buffer (50 mM Tris pH 8.0, 250 mM NaCl, 20 mM imidazole). Bound protein was eluted in 25 mM HEPES pH 7.5, 250 mM NaCl, 250 mM imidazole. Protein was concentrated using a spin concentrator (Amicon, 30K MWCO), and injected onto Superdex 200 column (GE Healthcare) equilibrated in 25 mM HEPES pH 7.5, 100 mM NaCl. Fractions were checked for purity by SDS-PAGE followed by coomassie blue staining.

Fluorescence polarization assays were based on those previously described(*39*) in black, flat bottom 384-well plates (Corning, 3575). Binding assays were performed in 25 mM HEPES pH 7.5, 66 mM NaCl, 0.05% NP-40 with 10 nM FAM-labeled DNA oligonucleotide (CG: sense FAM-5'-CCATG(5mC)GCTGAC-3', antisense 5'-GTCAG(5mC)GCATGG-3'; CHG: sense FAM-5'-CCATG(5mC)TGTGAC-3', antisense 5'-GTCA(5mC)AGCATGG-3'; CHH: sense FAM-5'-CCATG(5mC)TTTGAC-3', antisense 5'-GTCAAAG(5mC)ATGG-3'). Equimolar amounts of sense and antisense oligonucleotides were annealed in 10 mM Tris pH 8.0, 100 mM NaCl, 1 mM

EDTA by heating to 95 ºC for 10 min followed by a slow cooling to room temperature. SUVH1 WT or Y277A were serially diluted 2-fold and the final assay volume was 25 µL per well. Binding was measured at room temperature on a BioTek Synergy Neo plate reader. Polarization (P) was converted to anisotropy (A) using the formula A = 2P/(3-P). Data were plotted as fraction bound by setting the highest anisotropy measured equal to 1 for each curve. Data were plotted in GraphPad Prism, and dissociation constants (K_d) were fit using a non-linear regression of a one-site binder with Hill slope.

We note that SUVH1^{Y277A} interacts with unmethylated DNA more weakly than SUVH1 WT (Fig. 1B). One potential explanation for this is that the Y277 amino acid also helps to coordinate the interaction with DNA, which is consistent with our previous work on SUVH5 showing that unmethylated DNA can insert into the SRA binding pocket(*40*).

Microscale Thermophoresis (MST):

An N-terminal deletion construct of *Arabidopsis thaliana* SUVH3 (residues 135-669), which contains all the functional domains, including the two-helix bundle, SRA, pre-SET/SET/post-SET domains, was cloned in to a self-modified pFast-Bac-MBP vector to generate the His-MBP tagged target protein. Protein expression was conducted using the standard Bac-to-Bac baculovirus expression system in Sf9 insect cells. The recombinant expressed proteins were purified with a HisTrap column (GE Healthcare) and further purified by Heparin column and Superdex G200 column (GE Healthcare). All the mutant proteins were expressed and purified using the same protocol as their wild type counterparts.

The His-MBP tagged SUVH3 proteins were labeled with a Monolith NT Protein Labeling Kit (NanoTemper Technologies). The tested methylated DNA oligos were ordered from Shanghai Generay Biotech Company (CG: sense 5'-GAGTACT(5mC)GTCAGTTC-3', antisense, 3'-CTCATGAGCAGTCAAG -5'; CHG: sense 5'-GAGTACT(5mC)AGCAGTTC-3', antisense 3'-CTCATGAGTCGTCAAG-5'; CHH: 5'-GAGTACT(5mC)ATCAGTTC-3', antisense 3'-CTCATGAGTAGTCAAG-5'). The MST-based *in vitro* binding assay were conducted on a Monolith NT.115 instrument (NanoTemper Technologies) with blue/red filters as previously described(*41*). The fluorescence is situated to ~ 850 units. DNA oligos were annealed together with a PCR machine and half-and-half diluted in 16 steps with ddH_2O, resulting in a concentration gradient covering from 100 μM to 2 nM. The labeled protein samples were diluted in a buffer of 150 mM NaCl, 20 mM HEPES pH 7.5, and 1% Tween-20. The mixed protein-DNA samples were loaded into Monolith standard-treated capillaries (NanoTemper Technologies) at room temperature. All the binding experiments were performed with three repeats for each measurement and analyzed using the program NanoTemper analysis software (NanoTemper Technologies).

Plasmids:

For the *pSUVH1::*SUVH1-3xFLAG, *pSUVH3::*SUVH3-3xFLAG, *pDNAJ1::*DNAJ1-3xFLAG and *pDNAJ2::*DNAJ2-3xFLAG lines used for IP-MS and ChIP, genes were amplified from genomic template, including 1.5kb upstream from the 5'UTR (or until the next gene annotation) to the second last codon (to not include the stop codon) and cloned into pENTR /D-TOPO vectors (Thermo Fisher) and then transferred into a pEG302 based binary destination vector including a C-terminal 3xFLAG epitope tag via a Gateway LR Clonase II reaction (Invitrogen). For the co-immunoprecipitation experiments in *N. benthamiana*, these entry vectors were transferred into a

pEG302 based binary destination vector including a C-terminal 9xMYC epitope tag via a Gateway LR Clonase II reaction (Invitrogen). For the expression of SUVH1/3 DNAJ1/2 in the same bacterial cell, MBP-FLAG-SUVH1-6xHIS, MBP-MYC-SUVH3-6xHIS, MBP-HA-DNAJ1-V5, MBP-STag-DNAJ2-V5, constructs were synthesized by GenScript, using ATUM codon optimized CDS, and cloned into the dual expression vectors pETDuet-1 and pACYCDuet-1 (Novagen) to generate pACYCDuet-1_DNAJ1_SUVH1 and pETDuet-1_DNAJ2_SUVH3 constructs. In a second step of cloning, we used In-Fusion (Takara) to remove the 6xHIS tag from the C-terminus of MBP-FLAG-SUVH1 in the pACYCDuet-1 vector. For the Yeast Two-Hybrid experiments, the ATUM codon optimized CDS were cloned into pENTR /D-TOPO vectors, and then transferred into pDEST22/32 destination vectors (Invitrogen). For the ZF108 fusion constructs, the ATUM codon optimized CDS in pENTR /D-TOPO vectors were transferred into pMDC123 destination vectors encoding N-terminal ZF108-3xFLAG tag driven by the *Ubiquitin10* promoter (*24*, *42*). *UBQ10::*ZF108-YPET was previously described (*42*). *UBQ10::*DNAJ1 was generated by transferring the ATUM codon optimized DNAJ1 CDS into a pMDC123 destination vector encoding an N-terminal 3xFLAG tag driven by the *Ubiquitin10* promoter. For the *N. benthamiana* transient expression reporter constructs, the ZF108 binding site (5'-AGCCCATACATCTTTCCG-3') or a scrambled version (5'-CAATTCAGTCCGCCATTC-3'), was inserted 159bp upstream from the Luciferase start codon driven by the minimal 35S promoter, while on the same pGreen0800 based plasmid, Renilla was driven by the full 35S promoter. For the mammalian cell transfections, ATUM codon optimized CDS of DNAJ1 and DNAJ2 were cloned into the pECE72 based pBXG1 plasmid encoding an N-terminal Gal4 binding domain driven by the SV40 promoter, the VP16 and reporter construct were previously described (*43*).

ChIP-seq:

For all ChIP-seq experiments, 2-4 of grams of 12-14 days post germination seedling material was used per sample. For the ChIPs of *pSUVH1::*SUVH1-3xFLAG in *suvh1*, *pSUVH3::*SUVH3-3xFLAG in *suvh3*, *pDNAJ1::*DNAJ1-3xFLAG in *dnaj1* and *pDNAJ2::*DNAJ2-3xFLAG in *dnaj2*, input material was from T2 pools derived from 5-8 similarly expressing independent T1 lines alongside WT (non-transgenic) controls. For *pSUVH1::*SUVH1-3xFLAG in *suvh1,* vs. *pSUVH1::*SUVH1^{Y277A}-3xFLAG in *suvh1,* T2 material derived from two independent complementing or non-complementing lines (respectively, see Fig. S11) for each was used as starting material. For ChIPs of *pSUVH1::*SUVH1-3xFLAG in *suvh1* vs. RdDM backgrounds (*drm1/2, nrpd1-4, nrpe1-11)* we generated two technical replicates for each sample by splitting the chromatin into two before the addition of antibody. For the *pUBQ10::*ZF108-3xFLAG-DNAJ1 ChIP, as the plants displayed severe morphological defects, 0.6g of pooled adult T1 tissue from 6 week old plants was used as starting material. For the above described ChIPs, anti-FLAG M2 (F1804, Sigma-Aldrich) was used as antibody. For the H3K9me2 ChIP, anti-H3K9me2 = ab1220 (abcam) was used as antibody. ChIPs were performed as described previously with minor modifications(*22*). Samples were crosslinked *in vitro*, sheared using a Bioruptor Plus (Diagenode). Libraries were generated with NuGen Ovation Ultra Low System V2 kits following the manufacturer's instructions.

ATAC-seq:

ATAC-seq protocol was based on methods previously described in(*44*, *45*) using 1g of 12-14 days post germination seedling material from two biological replicates each of WT or *suvh1/3* mutant genotypes as input. Briefly, seedling tissue was chopped on glass using a razor blade in 5 ml of

ice-cold Lysis buffer (15 mM Tris-Cl pH 7.5, 0.5 mM spermine-4HCl, 80 mM KCl, 20 mM NaCl, 0.25 % Triton X-100, 5 mM β-mercaptoethanol*),* filtered through 40µM strainer (Falcon) into 5ml Sucrose buffer (20 mM Tris-Cl pH 8.0, 1.7M Sucrose, 5 mM $MgCl_2$, 0.25 % Triton X-100, 2 mM EDTA, 5mM β-mercaptoethanol), and spun at 20,000xg, for 30 mins at 4^{o}C. Pellets were resuspended in 1 ml of Lysis buffer and stained with DAPI (0.5 µg/ml final concentration) and 50,000 nuclei were collected by FACS into 0.5 ml of Lysis buffer. Transposition reaction and library amplification were performed as described in(*44*, *45*).

Whole Genome Bisulfite Sequencing (WGBS):

WGBS libraries were generated using the same protocol as described in(*46*). 1-2 bunches of unopened floral bud tissue were collected from individual plants, two plants per genotype, were used as starting material.

RNA-seq:

For RNA-seq of WT, *suvh1, suvh3, suvh1/3, dnaj1, dnaj2, dnaj1/2* mutants, 2-3 bunches of unopened floral bud per plant were collected, three plants per genotype. For RNA-seq of transgenic plants, we used leaf tissue from primary T1 transformants. For *UBQ10::*ZF108-DNAJ1, six biological replicates from individual T1 phenotype positive plants were collected. Three individual non-expressing and no phenotype T1 *UBQ10::*ZF108-DNAJ1 plants from the same tray, and three WT plants grown alongside, were used as controls. For *UBQ10::*ZF108-YPET and *UBQ10::*DNAJ1, four biological replicates of each from individual T1 expressing plants were harvested alongside four WT controls. Total RNA was extracted using TRIzol and Direct-zol RNA miniprep kit (Zymo, R2050) including in-column DNase treatment. Libraries were generated using

the TruSeq Stranded mRNA Library Prep Kit (Illumina), with ~2 µg RNA as input, following the manufacturer's instructions.

qRT-PCRs

For the *suvh1* complementation experiment in Fig. S11, RNA was extracted from unopened floral bud tissue, using TRIzol and Direct-zol RNA miniprep kit (Zymo, R2050) including in-column DNase treatment. For qRT-PCR of ROS1 expression in Fig. S21, tissue was harvested from 7 day old seedling tissue grown on plates as described by(*11*). cDNA was generated using the SuperScriptIII First-Strand Synthesis SuperMix kit (Thermo, 18080400) using oligo dT primed synthesis. qRT–PCRs were performed with iQ SYBR Green Mastermix (BioRad) using an Agilent Technologies Mx3005p qPCR System (Stratagene).

Bioinformatic Analysis:

Libraries were sequenced on illumina HiSeq 2000 or HiSeq 4000 machines. WGBS, RNA-seq, ChIP-seq and ATAC-seq mapping was performed as described previously (*24*, *45*, *46*). Briefly, RNA-seq reads were aligned using TopHat2(*47*) to the TAIR10 genome. For browser tracks, reads were normalised by mapped library size and binned into 25bp bins. For WGBS, reads were aligned using BSMAP(*48*). As there were few methylation differences between samples, we merged the biological replicates (2 per genotype) for downstream analysis to increase sequencing depth. Methylation metaplots were generated using the ComplexHeatmap package in R(*49*). Differentially methylated regions were defined in(*50*) (libraries: *cmt2* - GSM981002, *cmt3* - GSM981003, *drm1/2* - GSM981015, *met1* - GSM981031). As *drm1/2* hypo-CHH DMRs are largely a subset of *met1* hypo-CG DMRs, *met1* hypo-CG DMRs that overlapped with *drm1/2*

hypo-CHH DMRs were excluded from the *met1* hypo-CG set in Fig. 2A and Fig. S5B. ATAC-seq reads were aligned using Bowtie1(*51*), allowing 2 mismatches with unique mapping. ChIP-seq reads were aligned using Bowtie2, PCR duplicates were removed using SAMtools and MACS(*52*) was used for peak calling. As there were few differences between the technical replicates in the SUVH1 in RdDM mutant background ChIP-seqs, we merged the technical replicates for downstream analysis to increase sequencing depth. NGS.plot(*53*) and DeepTools(*54*) were used for generated metaplots with ChIP-seq and ATAC-seq data, and methylation metaplots were generated in R. For fold change analysis in (Fig. 3E, Fig. 4C, Fig. 4E, Fig. S8C, and Fig. S20), genes with no expression in any of the genotypes being compared were excluded. For proximal promoter analysis (in Fig. 3E, Fig. 4C-D, Fig. S20 and Fig. S25) scripts to define region proximity were previously described (*22*, *55*). In Figure 4E and Fig. S24, Bedtools intersect was used to define the overlap between ZF108-DNAJ1 peaks (MACS) and TAIR10 defined gene coordinates (from TSS to TTS), requiring at least 1bp overlap. FPKM +1 was used to plot WT FPKM and $\log_2$ fold change in Fig. 4E. Differentially expressed genes were determined using the DEseq(*56*) package in R.

Random Forest Regression for SUVH1 binding in vivo

We adapted the methods described in(*57*) for *Arabidopsis*, dividing the genome into 100bp bins and used CG bias (C+G%), context density (CG%, CHG%, CHH%) and WT methylation level (mCG, mCHG, mCHH) as feature predictors for each window. The SUVH1 binding response was defined as: if the fold change enrichment from ChIP-seq of SUVH1/WT >=1.5, then the bin is defined as 1 (binding); if the fold change of SUVH1/WT <1.5, then the bin is defined as 0 (no binding). We used 10,000 randomly sampled 100bp bins as a training set for the Random Forest

regressor, allowing 100 trees at a maximum depth of 10. We performed 5 permutations of this training to determine the relative importance of each predictor as depicted in Figure 2D. To determine the overall predictive power, we applied the model and features to predict another 10,000 randomly sampled 100bp bins that were excluded from the training set. The area under receiver-operating characteristic (ROC) curve (AUC) is shown in Figure 2E.

Immunoprecipitation Mass Spectrometry (IP-MS) and Western blots:

Starting material for the *pSUVH1::*SUVH1-3xFLAG in *suvh1, pSUVH3::*SUVH3-3xFLAG in *suvh3*, *pDNAJ1::*DNAJ1-3xFLAG in *dnaj1* and *pDNAJ2::*DNAJ2-3xFLAG in *dnaj2* IP-MSs was 8 grams of 12-14 days post germination seedling material from T2 pools derived from 5-8 similarly expressing independent T1 lines. The IP-MS method was performed as described in (*58*). For western on ChIP input fraction, anti-FLAG-HRP (Sigma-Aldrich, A8592) with ECL, and H3 loading control was detected using Rabbit anti-H3 primary (Abcam, ab1791) followed by IRDye 800CW Goat anti-Rabbit (LI-COR, 925-32211) secondary and visualized on a LI-COR Odyssey CLx.

Histone Methyltransferase (HMT) assays:

ATUM codon optimized SUVH1 and SUVH3 were cloned into a modified pD545 vector encoding N-terminal 6xHIS-MBP. These pD545:6xHIS-MBP-SUVH1/3 constructs were transformed into *Escherichia coli* strain BL21 (DE3) cells (NEB, C2527H). Cells were grown in LB medium to an OD_{600} of 0.4-0.6 at 37ºC, then protein expression was induced with 0.1mM isopropyl β-D-1-thiogalactopyranoside, followed by overnight incubation at 16ºC. Cells were lysed by sonication in lysis buffer (20 mM HEPES pH 7.5, 500 mM NaCl, 50 mM imidazole, 10% glycerol, 2 mM

DTT, 1 mM benzamidine, 1 mM PMSF and 1x cOmplete protease inhibitor). Clear lysate was incubated with Ni-NTA agarose resin (QIAgen, 30230). The resin was washed extensively with the lysis buffer before eluting with the buffer containing 500 mM imidazole. Eluted protein were dialyze against a storage buffer (20 mM HEPES, 150 mM NaCl, 2 mM DTT and 10% glycerol). HMT assays were performed as described in(*18*) using 4 µl of 5 µM enzyme with 25 mg/ml biotin conjugated H3K9me0/1/2 ($H3_{1-21}$, Millipore) peptides and 1 µl of 0.1 µM tritiated SAM[^{3}H] (Perkin Elmer) per assay.

***Nicotiana benthamiana* transient expressions for dual Luciferase assay:**

Agrobacterium cells were grown at 28ºC in LB supplemented with appropriate antibiotic to an OD_{600} of 1.5, then resuspended in (0.01M $MgCl_2$, 0.01 2-(*N*-morpholino)ethanesulfonic acid (MES) pH 5.6, 150 µM acetosyringone) to a final OD_{600} of 0.3, left for 2 hours at room temperature and then infiltrated into *N. benthamiana*. Three-week old *N. benthamiana* plant leaves were co-infiltrated with C58C1 *Agrobacterium* strains carrying the ZF108 target reporter mixed with ZF108-fused effector on one side of the leaf, and the ZF108 scrambled reporter mixed with the same ZF108-fused effector on the other side of the leaf (at a ratio of 1:5 reporter:effector fusion). Five independent leaf biological replicates were taken per combination and the experiment was repeated on at least 3 separate occasions. Tissue was collected 2 days post inoculation. Luciferase and Renilla levels were quantified using the Dual-Luciferase Reporter Assay System (Promega, E1910) according to manufacturer's instructions and a SynergyMx plate reader (BioTek).

Transfections for dual Luciferase assay:

Mouse neuroblastoma N2a cells (Flp-In host cell line) were provided by Dr. James Wohlschlegel and maintained at 37°C, 5% CO_2 and in humidity. Cells were plated in 6-well plates at $5x10^5$ cells/well in DMEM (Gibco) containing 10% FBS (Omega Scientific) and allowed to adhere to the plate overnight. The following morning, cells were co-transfected with 2ug of plasmids containing a 5xGal4 binding site luciferase reporter (pG5E4TluxpGL3basic (*43*)), an internal control (pRL-TK; Promega) and an Gal4 binding domain-fused effector (pBXG1 (*43*)) at a ratio 1:0.1:2, respectively, using BioT (Bioland Scientific) according to the manufacturer's protocol. Medium was changed after 5-7 hours to reduce cytotoxicity. Transfected cells were allowed to grow for 48 hours prior to harvesting. Luciferase and Renilla levels were quantified using the Dual-Luciferase Reporter Assay System (Promega, E1910) according to manufacturer's instructions and a SynergyMx plate reader (Biotek).

Co-immunoprecipitation in *N. benthamiana*:

Nicotiana benthamiana leaves were collected 72 hours after infiltration with *Agrobacterium* AGLO harboring corresponding expression plasmids as indicated. Expressions of all proteins were driven under their native promoters. As the stabilities of DNAJ1 and DNAJ2 when infiltrated alone were low, 10g of material was used as input as compared to ~1g for SUVH1/3 alone or co-infiltrated samples, and 25 μM of MG-132 was infiltrated into leaves expressing DNAJ1-9xMyc alone 12 hours before sample collection. Co-IPs were performed as previously described(*59*). Specifically, tissues were ground in liquid nitrogen and resuspended in 10 ml of Lysis Buffer (LB) (50 mM Tris pH7.5, 150 mM NaCl, 5 mM $MgCl_2$, 10% glycerol, 0.1% NP-40, 0.5 mM dithiothreitol (DTT), 1 mM phenylmethanesulfonylfluoride (PMSF), 1 tablet of protease inhibitor cocktail (Roche)). Lysates were cleared by filtration through miracloth followed by centrifugation

at 13,000 rpm for 10 min at 4˚C. Supernatants were incubated with 100 µl M2 FLAG magnetic beads (50% slurry; Sigma-Aldrich) for 1 hour at 4˚C with rotation. M2 FLAG magnetic beads were pre-blocked with LB with 5% bovine serum albumin (BSA) for 15 min at room temperature, followed by 3 times of wash with 1 ml LB. After incubation, beads were washed for 5 times with 1 ml LB with rotation at 4˚C. Elution was performed by incubating beads in 60 µl 2X SDS buffer (100 mM Tris-HCl pH 6.8, 200 mM DTT, 4% SDS, 0.2% bromophenol blue, 20% glycerol) at 80˚C for 20 min. Co-purification of SUVH1/SUVH3-9xMyc (with DNAJ1-3xFLAG) and DNAJ2-9xMyc (with SUVH1/SUVH3-3xFLAG) were detected with HRP-conjugated-anti-Myc 9E10 antibody (Santa Cruz Biotechnology). Elution of SUVH1/SUVH3-3xFLAG and DNAJ2-3xFLAG were detected with HRP-conjugate-anti-FLAG M2 antibody (Sigma-Aldrich). Western blots were developed with ECL Plus Western Blotting Detection System (GE Healthcare).

Bacterial co-expression with salt stringency washes:

BL21 (DE3) cells were transformed with the dual pETDuet-1:MBP-STag-DNAJ2-V5:MBP-MYC-SUVH3-6xHIS and pACYCDuet-1:MBP-HA-DNAJ1-V5:MBP-FLAG-SUVH1 vectors. Cells were grown in LB medium to an OD_{600} of 0.4-0.6 at 37ºC, then protein expression was induced with 0.1 mM isopropyl β-D-1-thiogalactopyranoside, followed by overnight incubation at 16ºC. For the SUVH1 immunoprecipitation, cells were pelleted the resuspended in 20x volume of (20 mM HEPES pH 8, 150 mM NaCl, 10% glycerol, 1.5 g/mL lysozyme, 1 mM PMSF and 1x cOmplete protease inhibitor) and sonicated. Lysate was mixed and incubated with anti-FLAG M2 magnetic beads (Sigma-Aldrich, M8823). Samples were then split and washed three times with buffer containing (20 mM HEPES pH 8, 10% glycerol, 1 mM PMSF and 1x cOmplete protease inhibitor) supplemented with 150 mM, 200 mM, 300 mM, 400 mM or 500 mM NaCl. Protein was

eluted with (20 mM HEPES pH 8, 150 mM NaCl, 10% glycerol, 1 mM PMSF and 1x cOmplete protease inhibitor, 150 ng/µL 3xFLAG peptide (Sigma-Aldrich, F4799)). For the SUVH3 pulldown, we followed the same protocol as above, but supplemented all buffers with 20 mM Imidazole and 0.05% Tween20 (as recommended by the Ni-NTA magnetic agarose bead manufacturer protocol), incubated the lysate with Ni-NTA magnetic agarose beads (Qiagen, 36111) and eluted with 300 mM Imidazole.

Western blots were visualized on a LI-COR Odyssey CLx machine. For primary antibodies, we used: Anti-FLAG M2 mouse (Sigma-Aldrich, F1803), Anti-HA rabbit (Cell Signaling, C29F4), Anti-c-MYC mouse (Sigma-Aldrich, 9E10), Anti-STAG mouse (Sigma-Aldrich, SAB2702204) and Anti-GAPDH goat (Life Span Biosciences, LS-C51084-40). For secondary antibodies we used (LI-COR IRDye): Anti-MOUSE 680CW, Anti-RABBIT 800CW and Anti-GOAT 680CW.

Yeast One- and Two-Hybrid:

Yeast two-hybrid was performed using the ProQuest Two-Hybrid System (Invitrogen) according to manufacturer's instructions. For Yeast One-Hybrid, we followed same procedure as with yeast two-hybrid using only pDEST32 based plasmids for MaV203 transformation onto SD-Leucine selection media.

Primer List:

name	Sequence in 5' to 3' direction	notes
qRT-PCR Primers		
AT5G17540 F	AAGCCGGAGTTAGTTTCTCC	used for qRT-PCR
AT5G17540 R	ATCCACAGCCAGTTTCCG	

AT3G26450 F	CAAGGAGAGGAGAGAGATAGACG	used for qRT-PCR
AT3G26450 R	CTCTTGAGGAGTTGCATGTAGC	
AT1G52040 F	GGGAAAAGAGAAACAAGAGACC	used for qRT-PCR, *same primers used in (*17*)
AT1G52040 R	GAACACAAGAGCAGTGACGA	
AT1G54040 F	AGGTATGGCCTGATCTCAAT	used for qRT-PCR, *same primers used in (*17*)
AT1G54040 R	GACAGTGGCAGCAGTATAGG	
AT2G21140 F	CCACGATGGCCTTGAGGC	used for qRT-PCR
AT2G21140 R	GAGGCTTGTAGATGGGAACCG	
AT3G02780 IPP2 F	GTATGAGTTGCTTCTCCAGCAAAG	used for qRT-PCR
AT3G02780 IPP2 R	GAGGATGGCTGCAACAAGTGT	
AT2G36490 ROS1 F	CAGGCTTGCTTTTGGAAAGGGTACG	used for qRT-PCR *same primers used in (*11*)
AT2G36490 ROS1 R	GTGCTCTCTCACTCTTAACCATAAGCT	
AT4G05320 UBQ10 F	GATCTTTGCCGGAAAACAATTGGAGG	used for qRT-PCR
AT4G05320 UBQ10 R	CGACTTGTCATTAGAAAGAAAGAGAT	

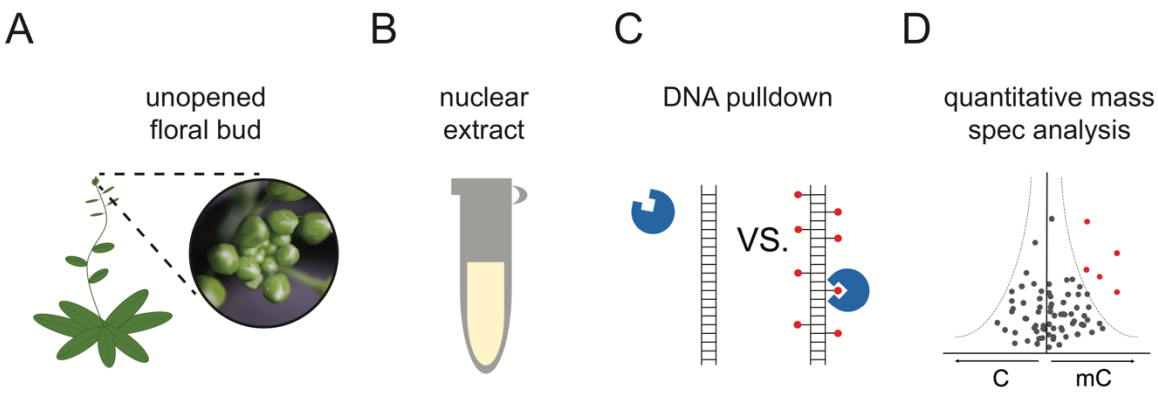

Fig. S1. Experimental setup for quantitative interactomic analysis. (A) Tissue is collected from unopened floral buds. **(B)** Extracts from isolated nuclei are prepared. **(C)** Nuclear extracts are incubated with either methylated or un-methylated biotinylated double stranded DNA oligonucleotides. **(D)** Interacting proteins are affinity purified and subjected to high-resolution mass spectrometry and label-free comparative analysis.

FWA F	TTCTTGCGCCGCTCTTTATCCCATTCAACATTCATACGAGCAC CGCTTTACGGTTTTTGCTTTTCGACATTGGTCGAAG	FWA unmeth
FWA R	AACTTCGACCAATGTCGAAAAGCAAAAACCGTAAAGCGGTGC TCGTATGAATGTTGAATGGGATAAAGAGCGGCGCAAG	
FWA CG F	TTCTTG(M)GCCGCTCTTTATCCCATTCAACATTCATA(M)GAG CAC(M)GCTTTA(M)GGTTTTTGCTTTT(M)GACATTGGT(M)GAA G	FWA mCG
FWA CG R	AACTT(M)GACCAATGT(M)GAAAAGCAAAAAC(M)GTAAAG(M) GGTGCT(M)GTATGAATGTTGAATGGGATAAAGAG(M)GG(M)G CAAG	
FWA CHH F	TT(M)TTGCGCCG(M)T(M)TTTAT(M)(M)(M)ATT(M)AA(M)ATT(M) ATACGAG(M)A(M)CG(M)TTTACGGTTTTTG(M)TTTTCGA(M)AT TGGTCGAAG	FWA mCHH
FWA CHH R	AA(M)TTCGA(M)(M)AATGTCGAAAAG(M)AAAAACCGTAAAGC GGTG(M)TCGTATGAATGTTGAATGGGATAAAGAGCGGCG(M) AAG	
SDC F	TTCACGTCAGTTATAAAGATAAGATTTCACAGTACACATCAGT TATAAAGATAAGATTTCACAGTACACGTCAGTTATA	SDC unmeth
SDC R	AATATAACTGACGTGTACTGTGAAATCTTATCTTTATAACTGAT GTGTACTGTGAAATCTTATCTTTATAACTGACGTG	
SDC CHG F	TTCACGT(M)AGTTATAAAGATAAGATTTCA(M)AGTACACAT(M) AGTTATAAAGATAAGATTTCA(M)AGTACACGT(M)AGTTATA	SDC mCHG
SDC CHG R	AATATAA(M)TGACGTGTA(M)TGTGAAATCTTATCTTTATAA(M) TGATGTGTA(M)TGTGAAATCTTATCTTTATAA(M)TGACGTG	
SUP F 1	AAGAGACAGACAGACATAGATATATCTTAGATTTTTCCCAGG GGAAAAAAGTGAGTTCATCTTTTCAGATCTTACAAA	SUP unmeth for CHG
SUP R 1	TTTTTGTAAGATCTGAAAAGATGAACTCACTTTTTTCCCCTGG GAAAAATCTAAGATATATCTATGTCTGTCTGTCTC	
SUP CHG F	AAGAGA(M)AGA(M)AGACATAGATATATCTTAGATTTTTCC(M) AGGGGAAAAAAGTGAGTTCATCTTTT(M)AGATCTTACAAA	SUP mCHG
SUP CHG R	TTTTTGTAAGAT(M)TGAAAAGATGAACTCACTTTTTTCCC(M)T GGGAAAAATCTAAGATATATCTATGT(M)TGT(M)TGTCTC	
MEA F	TTGTGGAATCGTTAATGACCACGGTTAAATGGCGGGATCCAA AATCCGGTTAGATTTCACAATGTCATATTACGTCATAT	MEA unmeth
MEA R	AAATATGACGTAATATGACATTGTGAAATCTAACCGGATTTTG GATCCCGCCATTTAACCGTGGTCATTAACGATTCCAC	
MEA CG F	TTGTGGAAT(M)GTTAATGACCA(M)GGTTAAATGG(M)GGGATC CAAAATC(M)GGTTAGATTTCACAATGTCATATTA(M)GTCATAT	MEA mCG
MEA CG R	AAATATGA(M)GTAATATGACATTGTGAAATCTAAC(M)GGATTT TGGATCC(M)GCCATTTAAC(M)GTGGTCATTAA(M)GATTCCAC	
SUP F 2	TTCACACACACCTCTCTCTCATCTCTATATCTCTCTCTCTCTCT CT	SUP unmeth control for CHH
SUP R 2	AAAGAGAGAGAGAGAGAGATATAGAGATGAGAGAGAGGTGT GTGTG	
SUP CHH F	TT(M)A(M)A(M)A(M)A(M)(M)T(M)T(M)T(M)T(M)AT(M)T(M)TATAT (M)T(M)T(M)T(M)T(M)T(M)T(M)T(M)T	SUP mCHH (annealed with **SUP R 2**)

Fig. S2. Oligonucleotide sequences used for DNA pulldown of nuclear extract. Sequences are from the *Arabidopsis* genome, based on naturally methylated loci where methylation in CG, CHG and CHH contexts have a particularly well-established role in affecting transcription(*34–37*). (M) = 5mC. Sequences are listed 5’ to 3’.

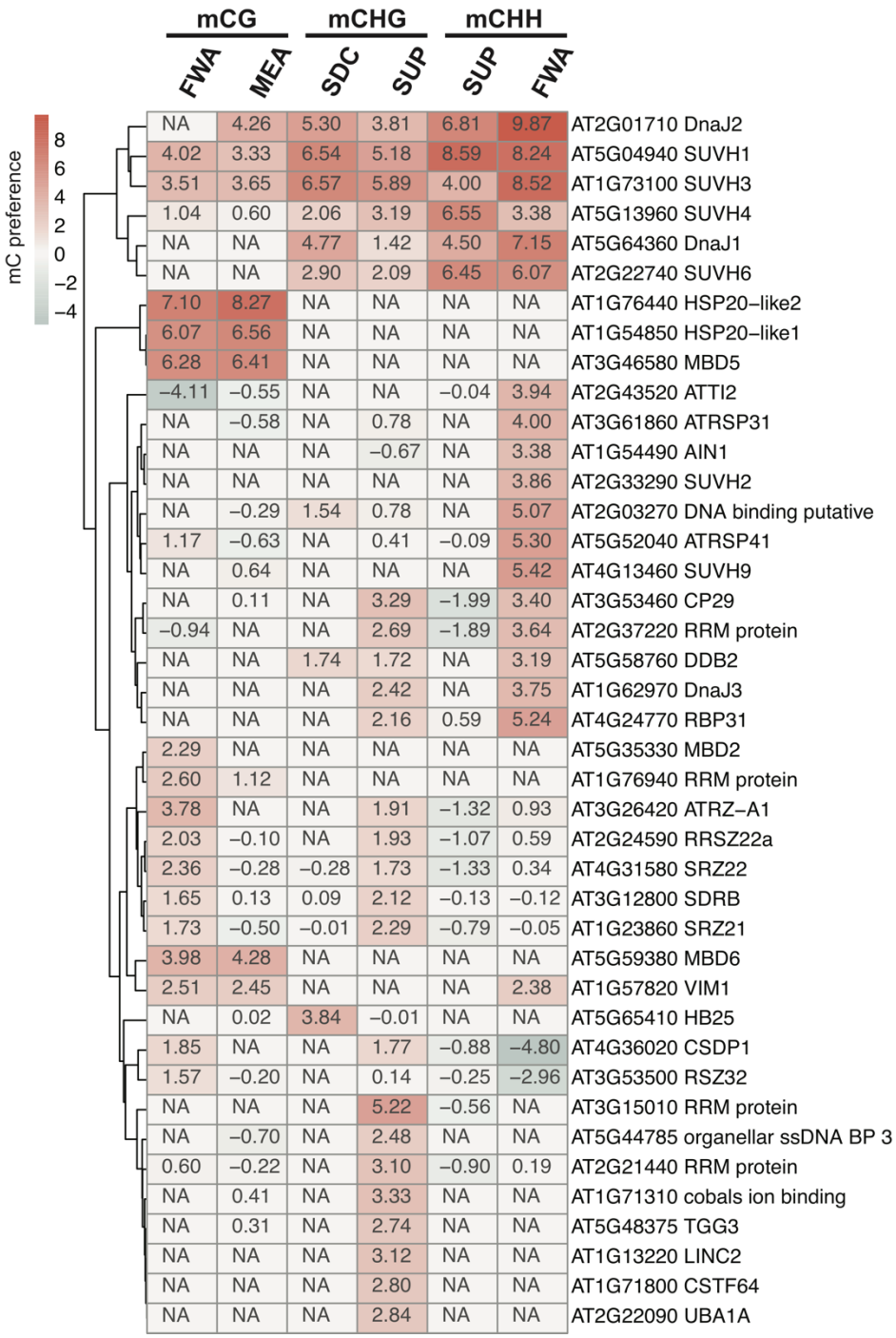

Fig. S3. Heatmap of proteins with significant methyl-cytosine (mC) binding preference. All proteins identified as significant methyl-cytosine binders when enriched in at least one methyl vs. non-methyl C pulldown (see Primer List).

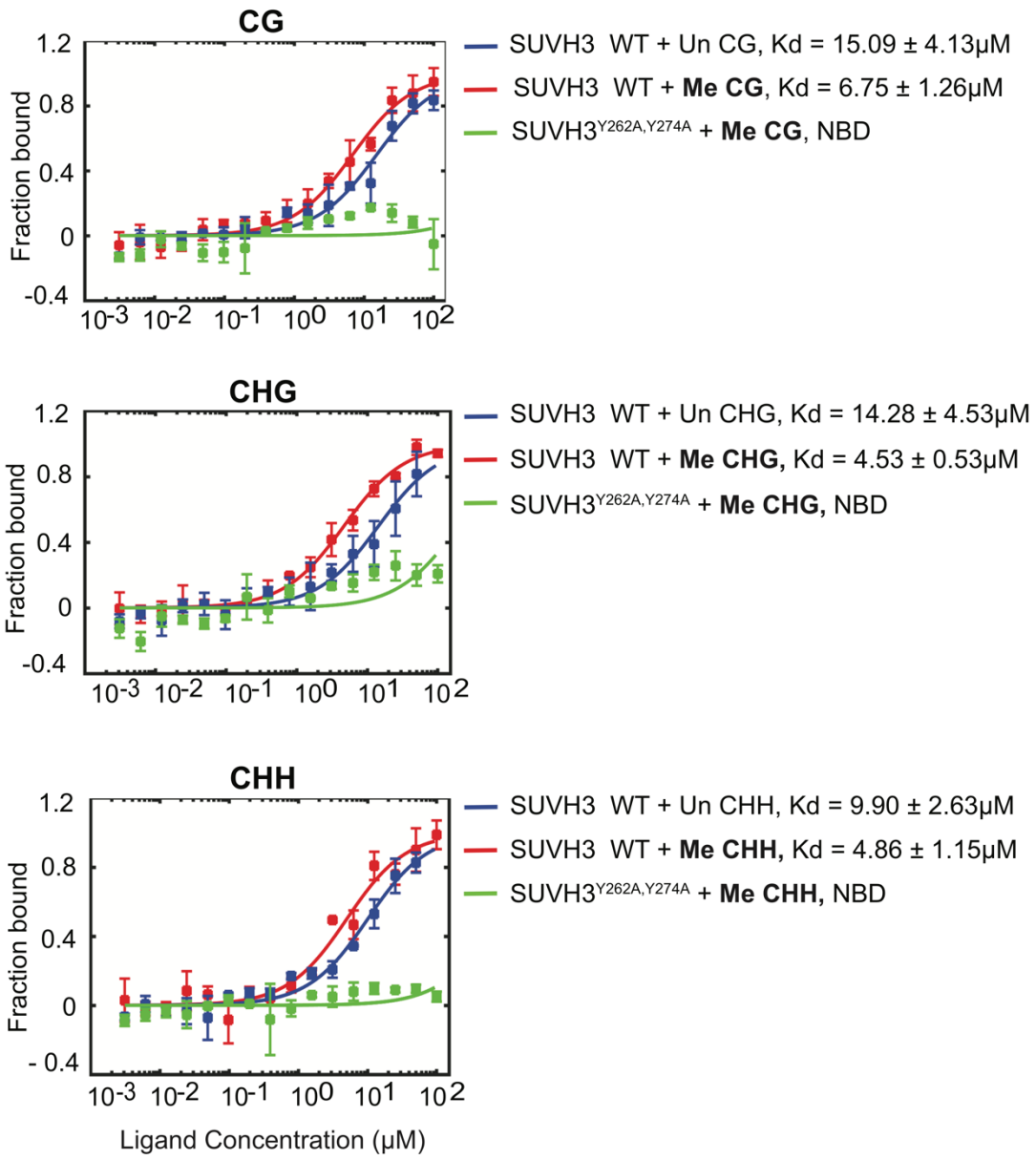

Fig. S4. SUVH3 methyl-binding assessed by microscale thermophoresis (MST). MST binding assays to quantify the interaction of SUVH3 with methylated (Me) or unmethylated (Un) probes in CG, CHG and CHH contexts. SUVH3^{Y262A,Y274A} encodes two amino acid changes in the SRA domain predicted to abrogate methyl-binding(*18*). Data points are shown as means +/- SD (n=3). Curves indicate calculated fits, and binding affinities (K_d) are listed. NBD indicates 'no binding detected'.

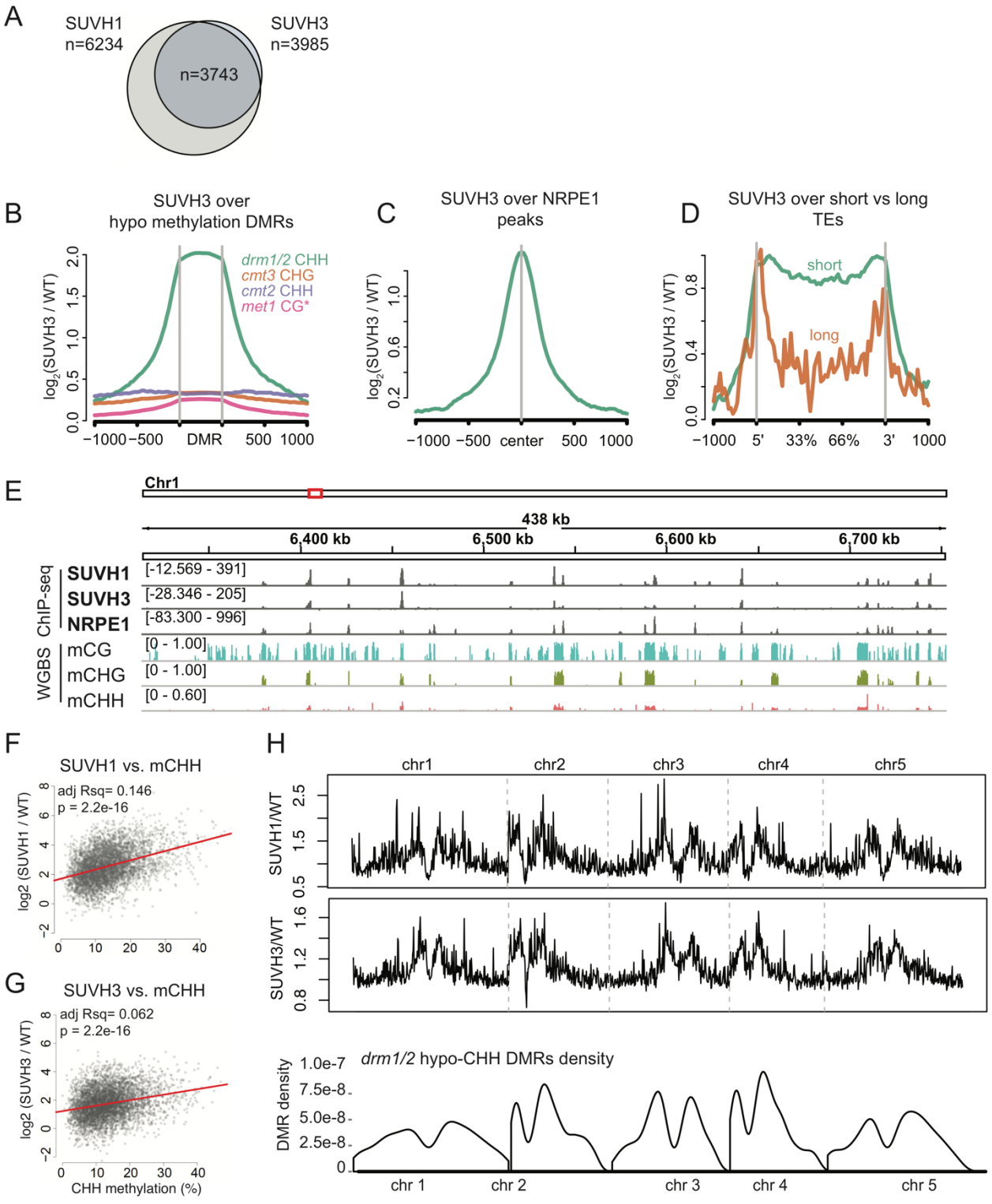

Fig. S5. SUVH3 is colocalized with SUVH1 and recruited to RdDM loci. (A) Overlap between SUVH1 and SUVH3 ChIP-seq peaks. **(B)** SUVH3 enrichment at loci defined by loss of

methylation - 'hypo' differentially methylated regions (DMRs) - in mutant genotypes indicated. The DRM1/2 methyltransferases are responsible for mCHH at RdDM target sites, while mCG, mCHG and heterochromatic mCHH are maintained by MET1, CMT3 and CMT2, respectively(*1*). * indicates that *met1* hypo CG DMRs that overlap with *drm1/2* hypo CHH DMRs are removed. **(C)** SUVH3 enrichment at NRPE1 peaks. **(D)** SUVH3 enrichment at NRPE1 associated short vs. long TEs. **(E)** Representative browser track showing ChIP-seq co-localization of SUVH1, SUVH3 (y-axis = normalized, anti-FLAG [SUVH1/3-FLAG transgenic minus WT]) and NRPE1 (y-axis = normalized, WT [anti-NRPE1 minus input]) at methylated loci (WGBS data from WT, y-axis = methylation fraction). **(F)** Scatterplot showing SUVH1 enrichment (y-axis) vs. wild type CHH methylation level (x-axis) at *drm12* hypo-CHH DMRs. **(G)** Scatterplot showing SUVH3 enrichment (y-axis) vs. wild type CHH methylation level (x-axis) at *drm12* hypo-CHH DMRs. **(H)** Upper panel: SUVH1 enrichment, Middle panel: SUVH3 enrichment, Lower panel: *drm1/2* hypo-CHH DMR density, in 100kb bins across all five chromosomes of the *A. thaliana* genome.

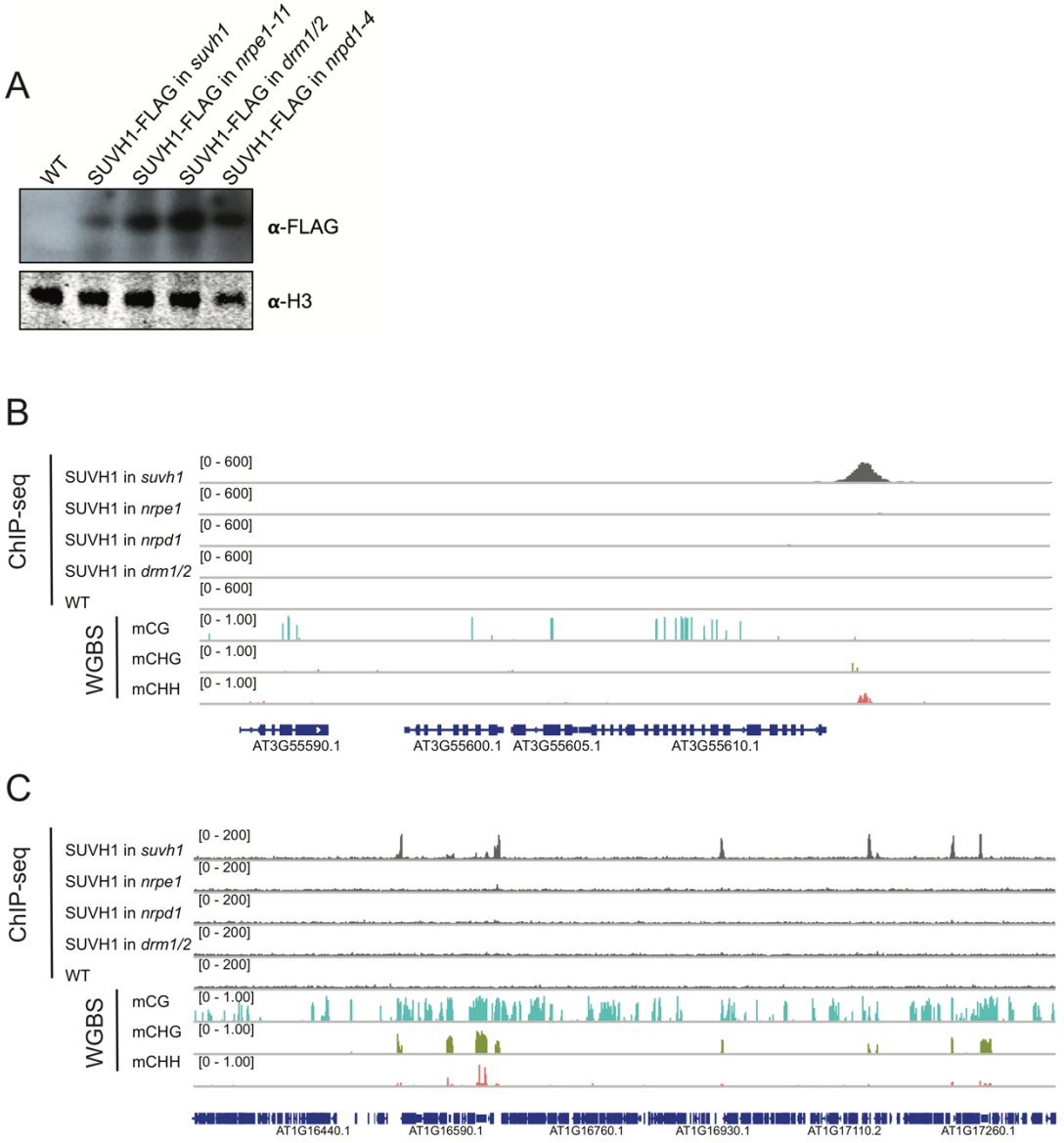

Fig. S6. SUVH1 enrichment is lost in RdDM mutant backgrounds. (A) Western blot on the input fractions used for the SUVH1 ChIP-seqs shown in (B) and (C) and Fig. 2F, showing that SUVH1 stability is not compromised in the *nrpe1-11*, *drm1/2* or *nrpd1-4* backgrounds. H3 is used as a loading control. **(B)** Representative browser track showing loss of SUVH1 occupancy at a highly CHH methylated locus in RdDM mutant backgrounds. ChIP-seq y-axis = normalized reads; WGBS data is from WT, y-axis = methylation fraction. **(C)** Representative browser track as in (B) at a more zoomed out chromosomal viewpoint.

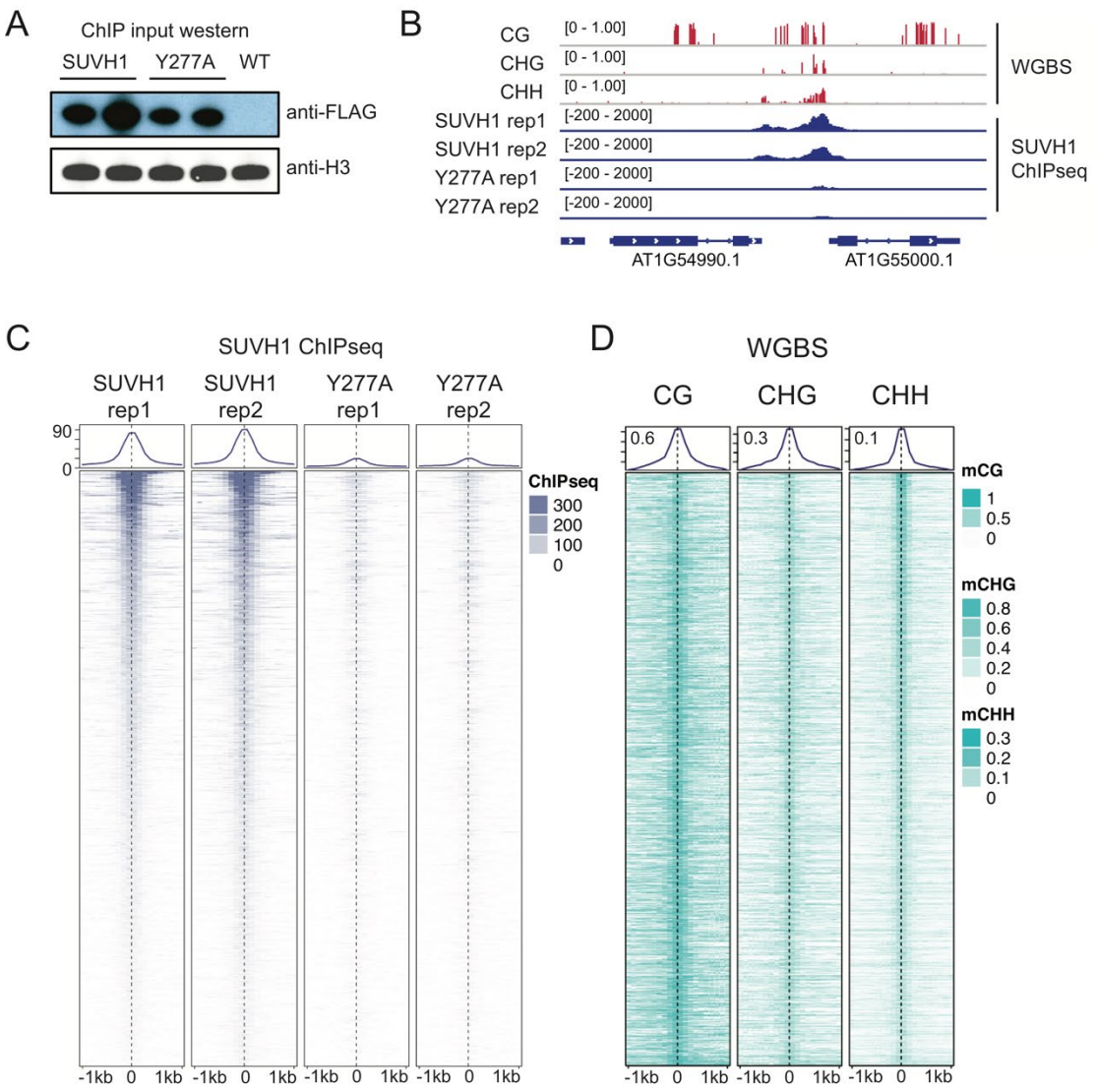

Fig. S7. SUVH1^{Y277A} shows reduced recruitment to in vivo. (A) Western blot on the input fractions used for SUVH1 vs. SUVH1^{Y277A} ChIP-seqs shown in (B) and (C) and Fig. 2G. H3 is used as a loading control. **(B)** Representative browser track showing SUVH1 vs. SUVH1^{Y277A} occupancy at a highly CHH methylated locus. **(C)** ChIP-seq metaplot for SUVH1 vs. SUVH1^{Y277A} replicates at SUVH1 peaks. (normalized, anti-FLAG [transgenic minus WT]). **(D)** Methylation metaplot in WT at SUVH1 peaks.

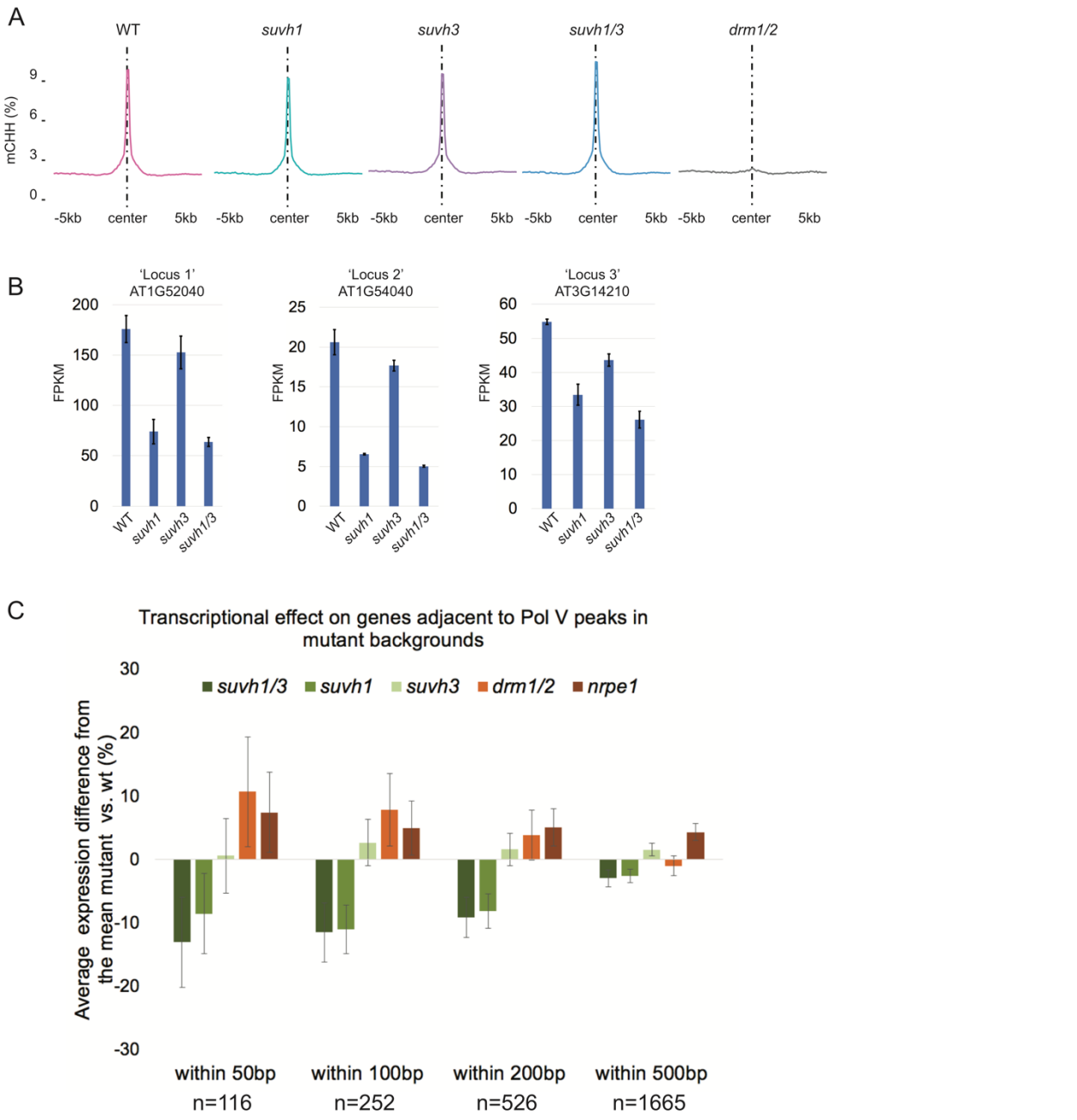

Fig. S8. SUVH1/3 are required for the expression of RdDM proximal genes but not methylation maintenance. (A) CHH methylation levels in genotypes indicated over SUVH1 defined ChIP-seq peaks. **(B)** RNA-seq expression in the genotypes indicated at the three genic loci defined by(*17*) as requiring *SUVH1* for expression. FPKM = fragments per kilobase per million mapped reads. Error bars represent s.e.m. from 3 biological replicates. **(C)** Average difference

from the mean (mutant FPKM – [mutant FPKM + wt FPKM]/2) (*100 for percentage) for mutants indicated (*nrpe1* and *drm1/2* RNA-seq data from (*20*, *60*)), at genes with Pol V (NRPE1) peaks within 50, 100, 200 or 500bp from the TSS. Genes with no expression in both genotypes were excluded. *drm1/2* and *nrpe1* are RdDM mutants that are used as controls for comparison. Error bars depict s.e.m.

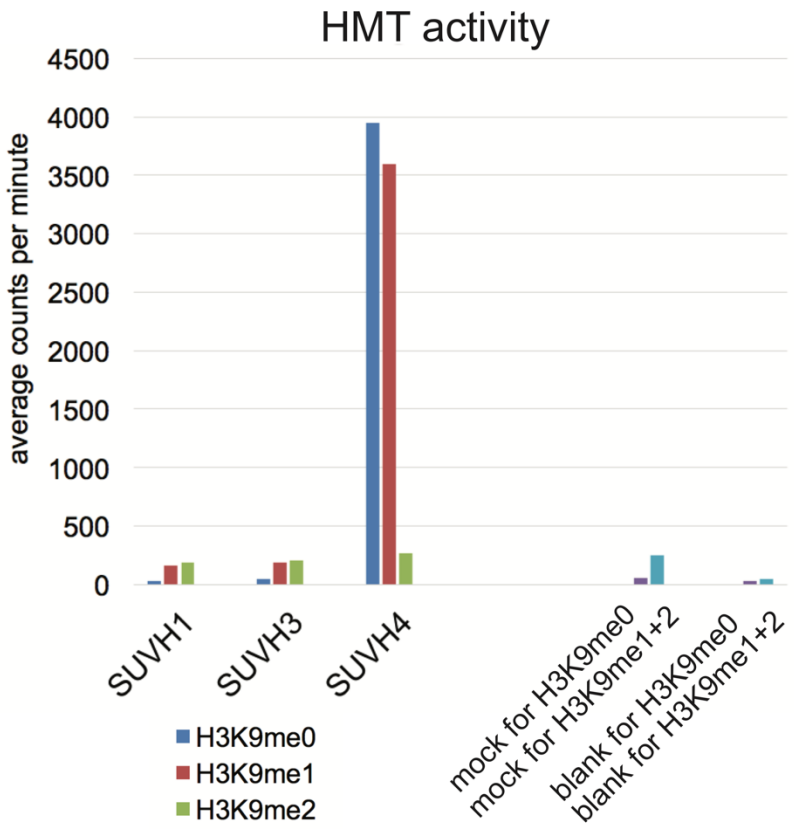

Fig. S9. Full length SUVH1 and SUVH3 do not display HMT activity *in vitro*. HMT assay using SAM[^{3}H], for the peptide substrate and enzyme combinations indicated. mock = no enzyme added. blank = no enzyme and no peptide.

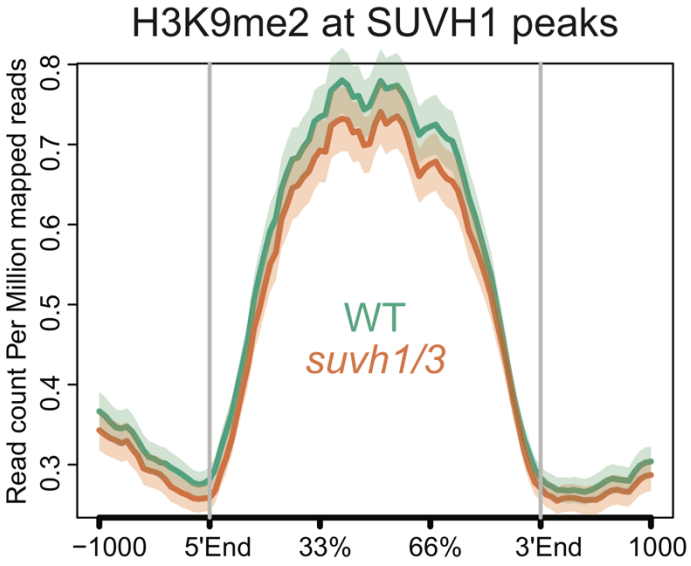

Fig. S10. H3K9me2 levels are unchanged in *suvh1/3* mutants over SUVH1 peaks. Metaplot of H3K9me2 ChIP-seq reads in wild type (WT) vs. the *suvh1/3* mutant at SUVH1 peaks.

A

EFFECT	SUVH4 AA CHANGE	EQUIVALENT AA CHANGE IN SUVH1
METHYL BINDING NULL	Y219A	**Y277A**
HMT CATALYTIC HYPOMORPH	Y475F	**Y524F**
HMT CATALYTIC NULL	Y593F	**Y638F**

B

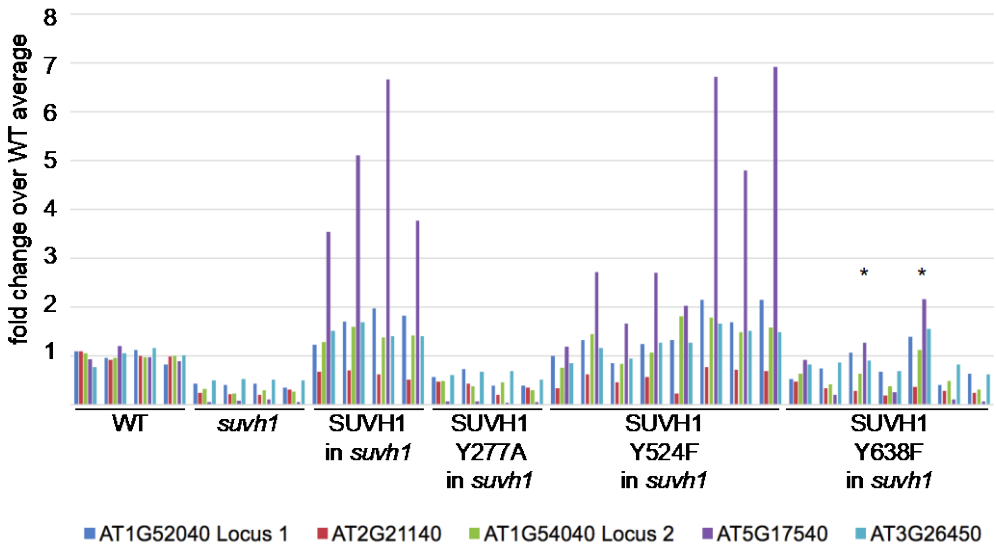

Fig. S11. SET domain HMT catalytic activity is non-essential for SUVH1 function *in vivo*. **(A)** Table showing amino acid changes defined in(*18*) and equivalent amino acid changes in SUVH1 based on sequence alignment. **(B)** qRT-PCR at 5 SUVH1 target loci (genes where mRNA levels are decreased in *suvh1* but are rescued by re-introduction of SUVH1) for individual plants. Wild type (WT) and *suvh1* genotypes are shown as controls, 4 plants each. Four representative T1 plants of SUVH1 (wild type copy) in *suvh1*, and SUVH1 Y277A are shown (all T1s of these genotypes behaved similarly). 8 independent T1 plants for Y524F and 7 independent T1 plants for Y638F are shown. All plants were grown side by side and unopened floral bud tissue was extracted for RNA isolation at the same time. IPP2 was used for normalization (see primer list). * indicates two Y638F T1 plants that partially complement. 'Locus 1' and 'Locus 2' are the same SUVH1 target loci identified in (*17*).

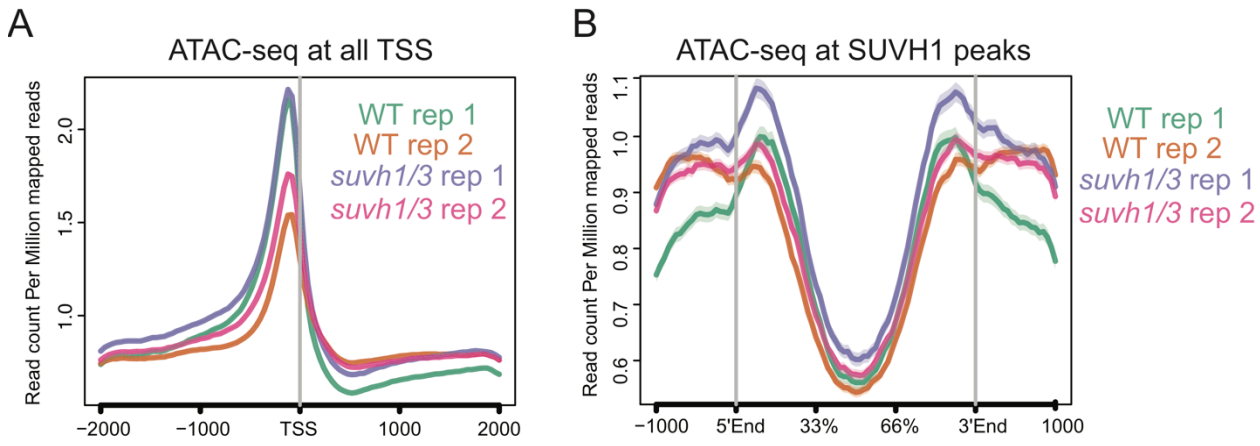

Fig. S12. ATAC-seq levels are unchanged in *suvh1/3* mutants over SUVH1 peaks. (A) ATAC-seq signal in two biological replicates of WT and *suvh1/3* over all gene TSS. **(B)** ATAC-seq signal of same libraries shown in (A) over SUVH1 peaks.

		SUVH1 (AT5G04940)		SUVH3 (AT1G73100)		DNAJ1 (AT5G64360)		DNAJ2 (AT2G01710)	
		NSAF	Uniq Pept	NSAF	Uniq Pept	NSAF	Uniq Pept	NSAF	Uniq Pept
IP SUVH1	rep. 1	1152.3	51	126.2	14	390.0	17	989.1	24
	rep. 2	1328.6	47	212.3	15	551.1	17	1370.2	25
IP SUVH3	rep. 1 †	91.9	10	1311.0	71	68.7	7	280.9	16
	rep. 2 †	117.0	13.5	1416.7	73.5	76.1	9	316.9	18.5
IP DNAJ1	rep. 1	347.9	20	278.7	19	1185.4	27	539.6	15
	rep. 2 †	288.8	24	182.9	23.5	797.2	39	315.1	15.5
IP DNAJ2	rep. 1 †	147.3	4	227.9	7	19.4	0.5	749.0	7.5
	rep. 2 †	131.9	5	96.3	4	17.7	0.5	748.1	10
IP WT control	rep. 1	0.0	0	0.0	0	0.0	0	0.0	0
	rep. 2	0.0	0	0.0	0	0.0	0	0.0	0
	rep. 3	0.0	0	0.0	0	0.0	0	0.0	0
	rep. 4	0.0	0	0.0	0	0.0	0	0.0	0

Fig. S13. Table of IP-MS results from SUVH1, SUVH3, DNAJ1, DNAJ2 tagged lines and four non-transgenic WT controls. Immunoprecipitation mass spectrometry (IP-MS) from tagged lines. Only proteins present in each of the four transgenic but not in WT pulldowns are presented in the table. Same data as depicted in Fig. 3A, with the biological replicates shown separately. NSAF = normalized spectral abundance factor. Uniq pept = unique peptides. † = average values from 2 technical replicates.

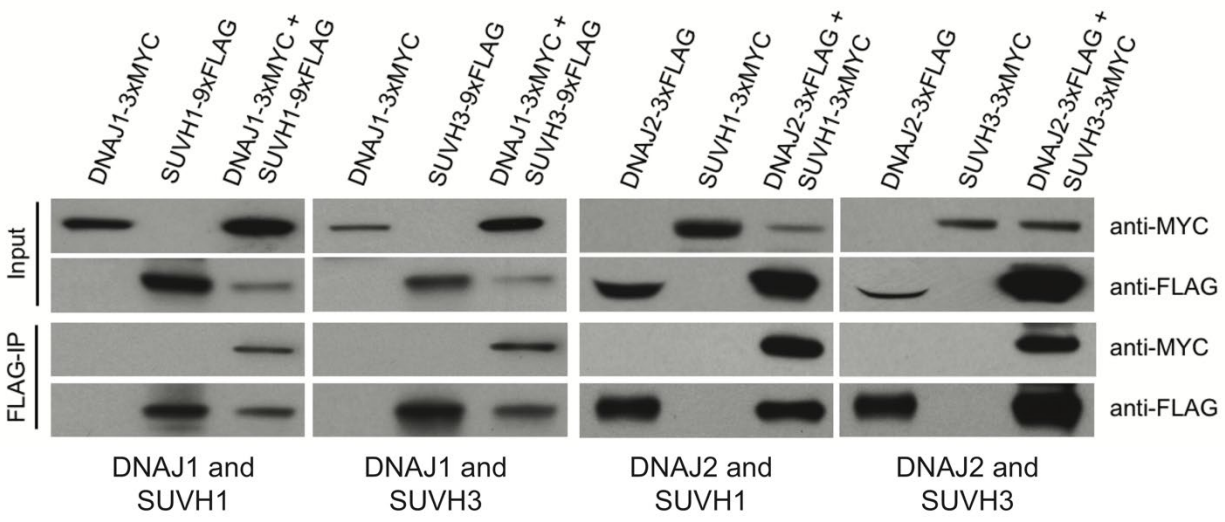

Fig. S14. Co-immunoprecipitation of SUVH1/3 with DNAJ1/2. Pairwise co-immunoprecipitation of SUVH1/3 with DNAJ1/2 after transient expression in *N. benthamiana*.

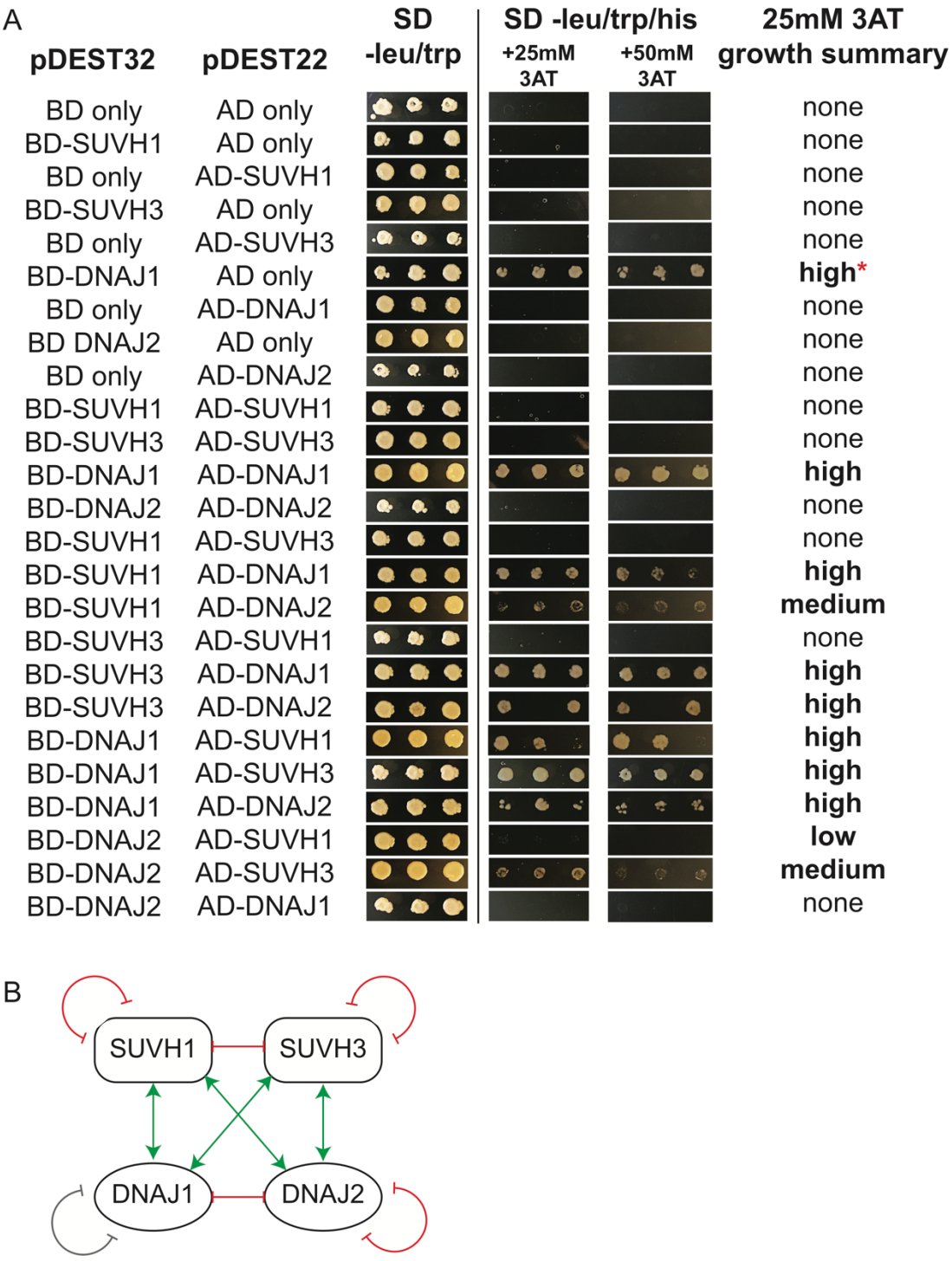

Fig. S15. Yeast two-hybrid identifies direct interaction between SUVH1/3 and DNAJ1/2. (A) Co-transformation setup is shown on left. Growth of transformed yeast (in biological triplicates) is shown on permissive SD -leu/-trp media, and on SD -leu/-trp/-his + competitive 3AT inhibitor

(at 25mM or 50mM). **(B)** Model to summarize direct interaction results. Green arrows indicate interaction observed, red lines indicate no evidence for interaction observed. Grey line for DNAJ1 indicates that we are unable to determine if homotypic interaction exists as DNAJ1 activates expression of the reporter when co-expressed with an empty pDEST22(AD) vector (see *).

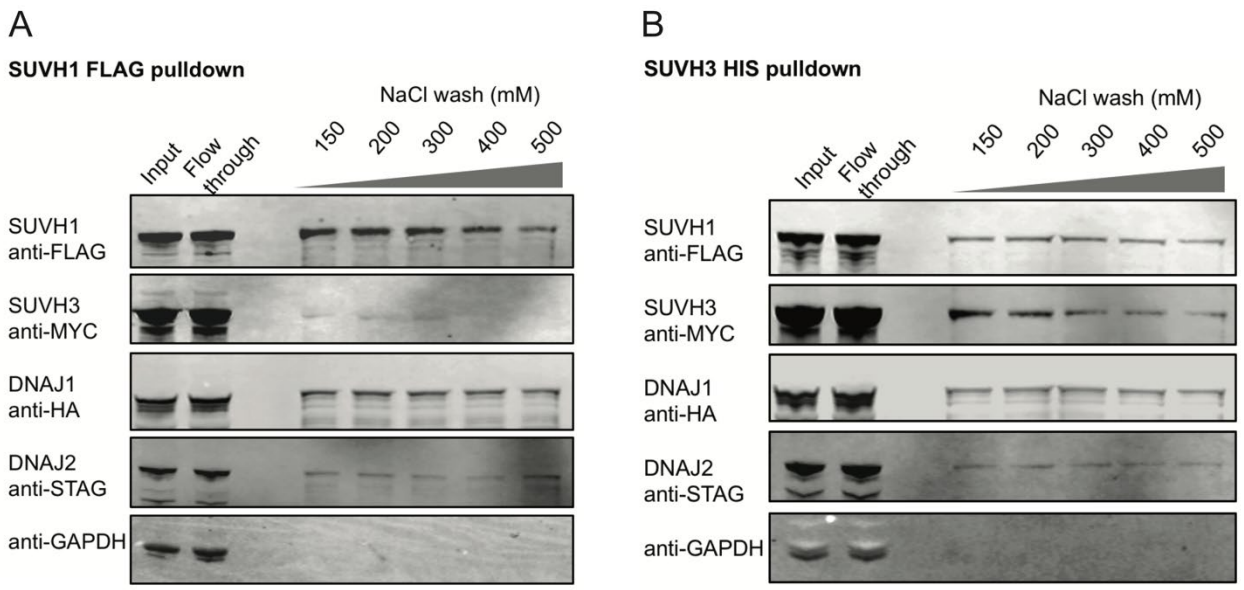

Fig. S16. Bacterial co-immunoprecipitation with salt stringency washes: (A) Co-immunoprecipitation of MBP-FLAG-SUVH1, MBP-MYC-SUVH3-6xHIS, MBP-HA-DNAJ1-V5 and MBP-STag-DNAJ2-V5 after co-expression and anti-FLAG immunoprecipitation or **(B)** Ni-NTA HIS resin purification. Proteins were eluted with either FLAG peptide (A) or Imidazole (B) after performing the 3x salt stringency washes as indicated. GAPDH is used as a negative control.

Proteins	**wild type genotype** uniq. peptide avg. (four tech reps)		**suvh1/3 genotype** uniq. peptide avg. (four tech reps)		**dnaj1/2 genotype** uniq. peptide avg. (four tech reps)	
	non-meth	meth	non-meth	meth	non-meth	meth
DNAJ1 (AT5G64360)	0	4	0	0	0	0
DNAJ2 (AT2G01710)	0	10	0	0	0	0
SUVH1 (AT5G04940)	0.25	15.5	0	0	0	6.75
SUVH3 (AT1G73100)	0.5	7	0	0	0	1.75
SUVH4 (AT5G13960)	3.25	14	1.25	7.25	1.75	9.75
SUVH6 (AT2G22740)	0.5	14.5	0	8	0	4

Fig. S17. *SUPERMAN* (*SUP*) CHH context comparative mass spectrometry pulldown in WT, *suvh1/3* and *dnaj1/2* genotypes. Table shows number of unique peptides pulled down with either methylated or unmethylated oligonucleotides using the genotypes indicated.

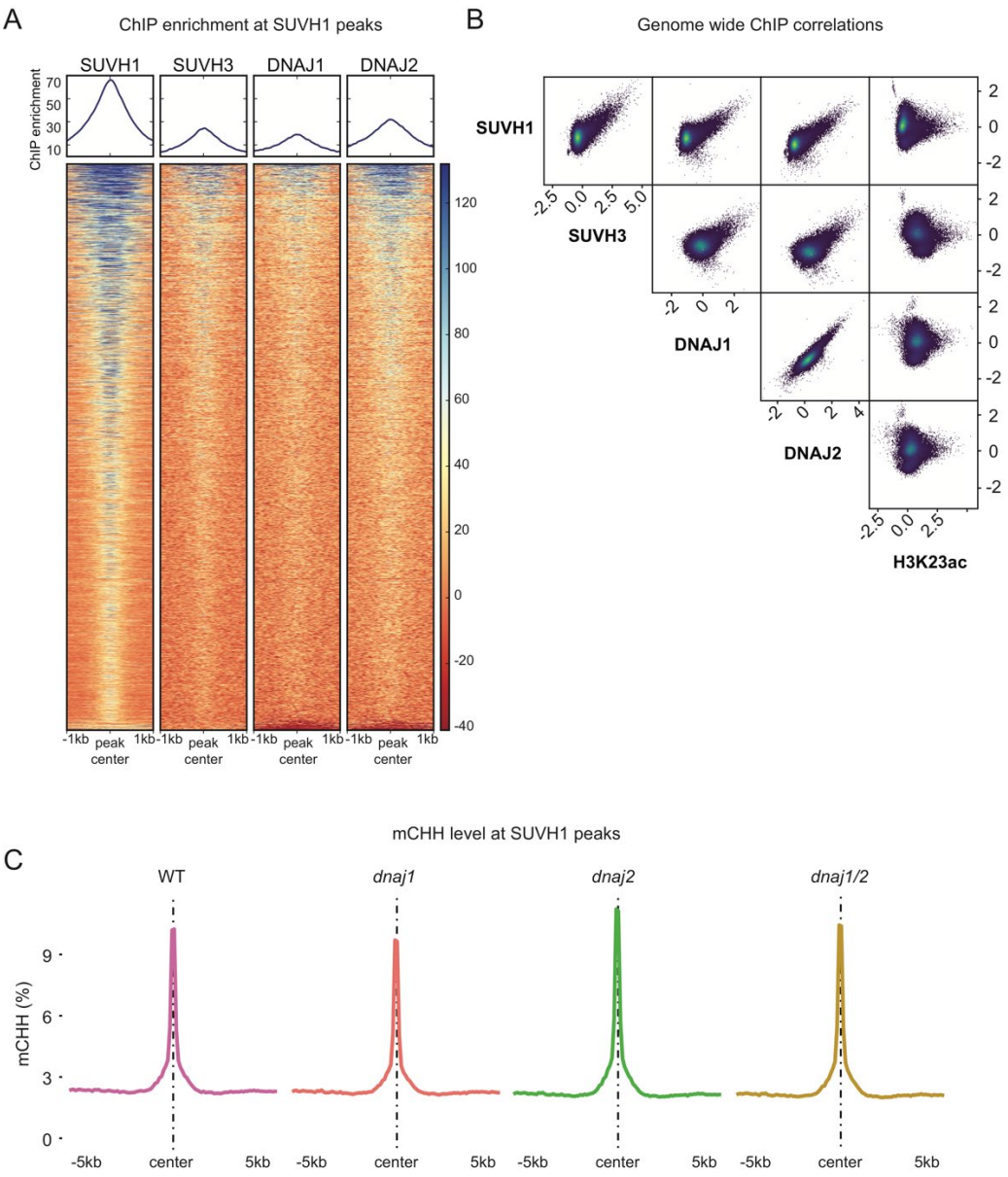

Fig. S18. DNAJ1/2 co-localize with SUVH1/3 and are not required for methylation maintenance. (A) ChIP-seq enrichment of SUVH1, SUVH3, DNAJ1 and DNAJ2 tagged lines (normalized reads of FLAG-tagged versions minus wild type) over SUVH1 peaks. **(B)** Pairwise scatterplots for ChIP-seqs at 1kb bins across the genome (see Fig. 3C). **(C)** CHH methylation levels in genotypes indicated over SUVH1 defined ChIP-seq peaks.

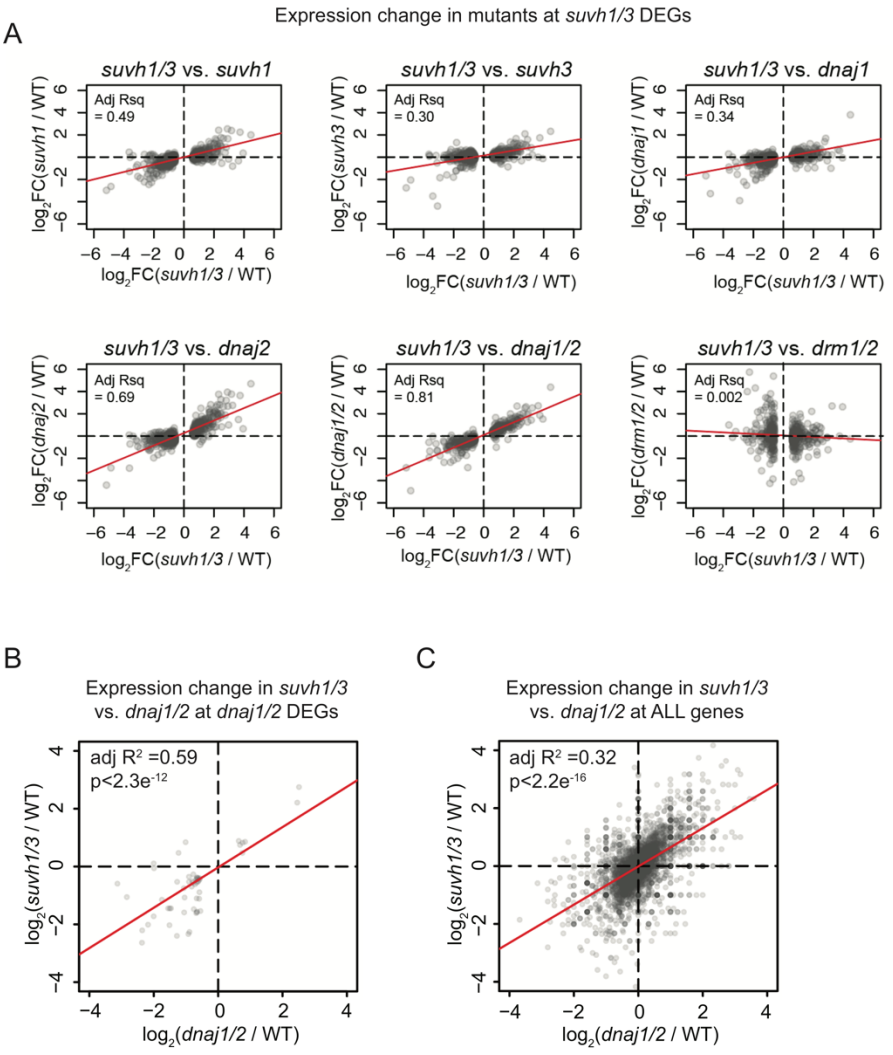

Fig. S19. Correlation between *suvh1/3* differentially expressed genes and other mutant genotypes. (A) Scatterplot of FPKM $\log_2$ fold change over wild type (WT) of mutants indicated (y-axis) vs. *suvh1/3* (x-axis) at *suvh1/3* differentially expressed genes. Red lines indicate line of best fit and adjusted R^2 value is indicated. *drm1/2* (bottom right) is used as a negative control with data from (*20*). **(B)** Scatterplot of FPKM $\log_2$ fold change over wild type (WT) of *dnaj1/2* (x-axis) vs. *suvh1/3* (y-axis) at *dnaj1/2* differentially expressed genes. **(C)** Scatterplot of FPKM $\log_2$ fold change over wild type (WT) of *dnaj1/2* (x-axis) vs. *suvh1/3* (y-axis) over all genes.

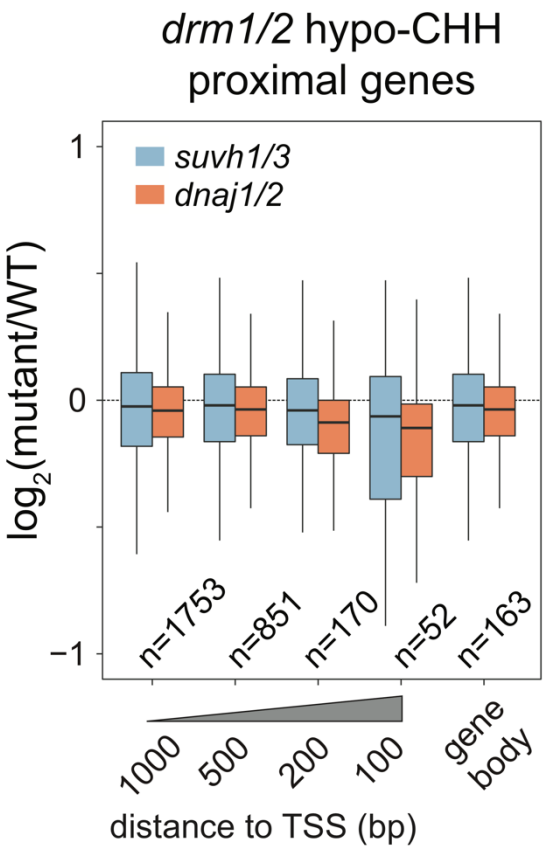

Fig. S20. *SUVH1/3* and *DNAJ1/2* promote the expression of RdDM proximal genes. Boxplot of expression change of genes proximal to *drm1/2* hypo CHH DMRs. n = number of genes.

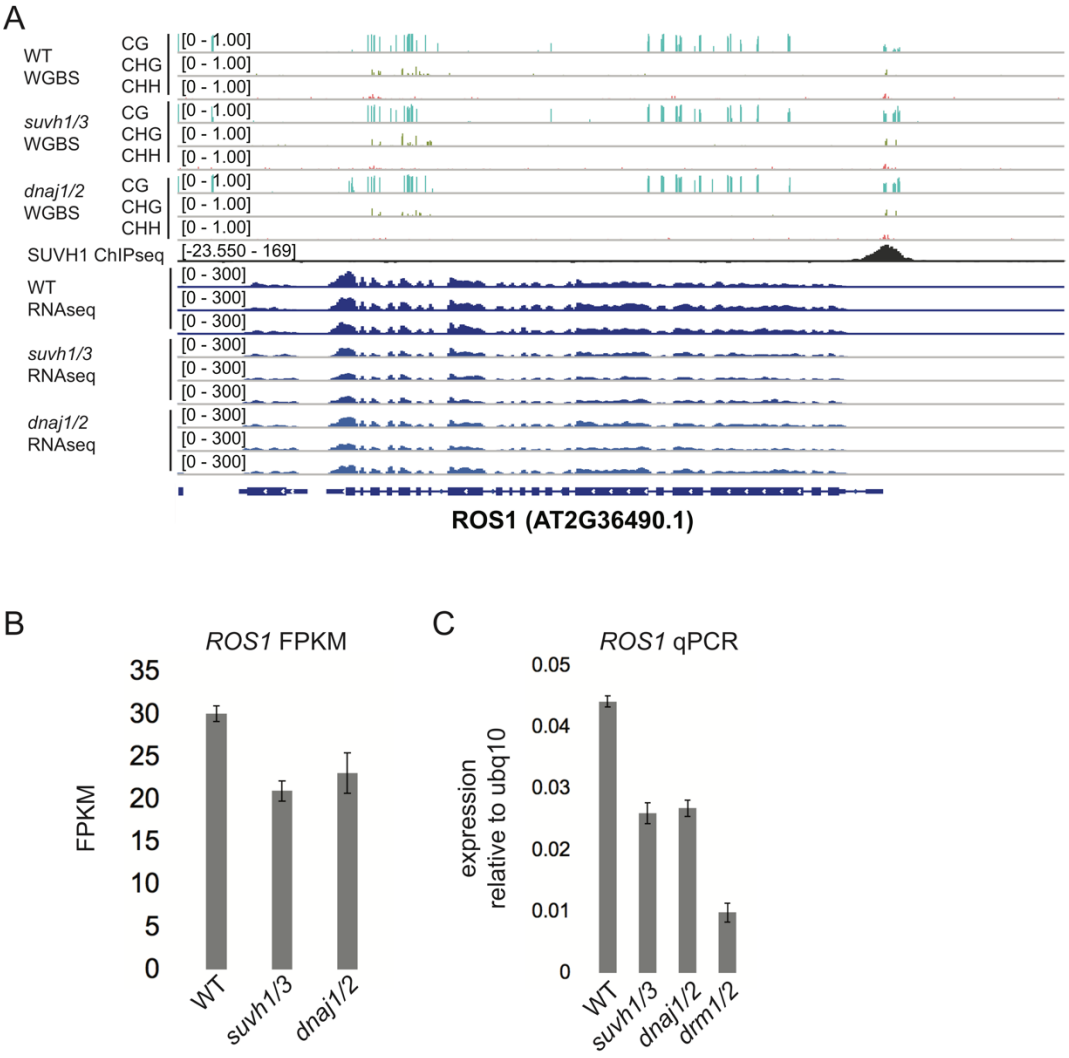

Fig. S21. *SUVH1/3* and *DNAJ1/2* are required for full expression of ROS1. (A) Browser track at the *ROS1* locus showing RNA-seq and WGBS from WT, *suvh1/3* and *dnaj1/2* genotypes along with ChIP-seq of SUVH1, showing reduced expression of *ROS1* in *suvh1/3* and *dnaj1/2* while methylation levels are retained. **(B)** RNA-seq expression of *ROS1* in the genotypes indicated. Error bars represent s.e.m. from three biological replicates. **(C)** qRT-PCR on *ROS1* in genotypes indicated, from 7 day old seedlings, as described by(*11*). *drm1/2* is used as a control genotype, known to be required for the expression of *ROS1*. Error bars represent s.e.m. from three biological replicates.

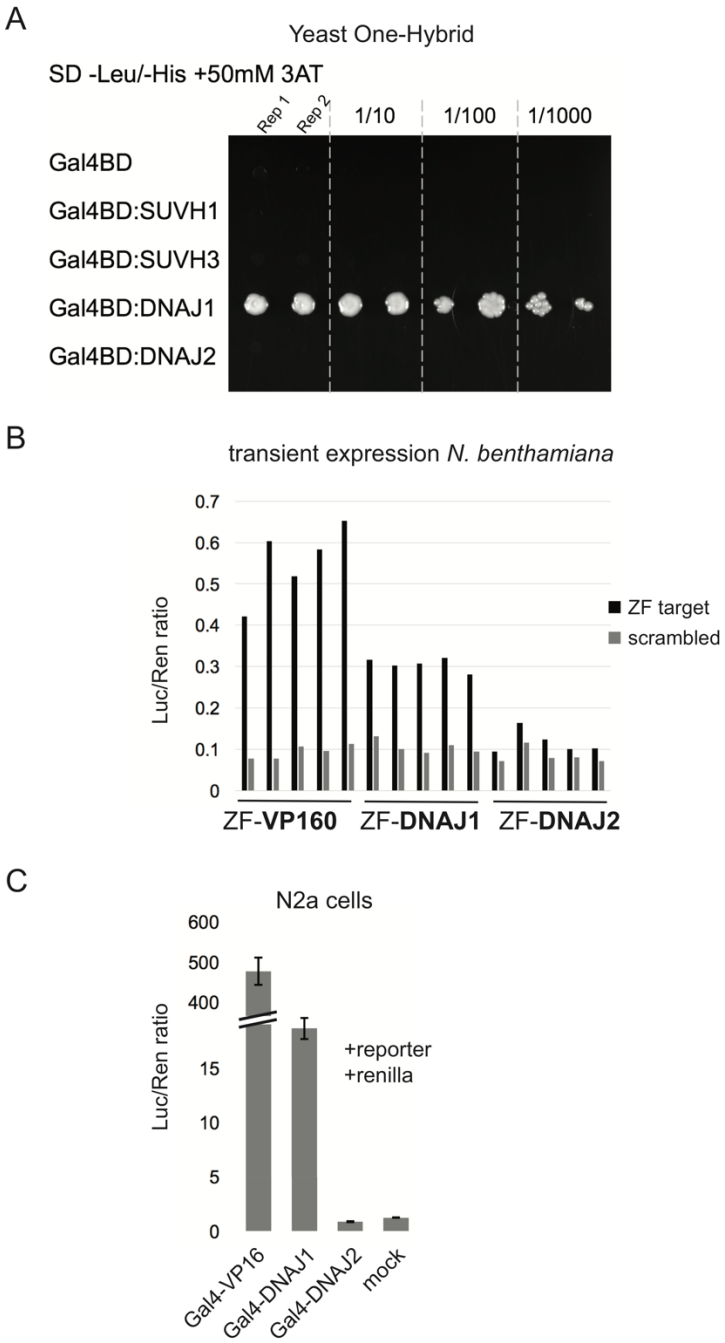

Fig. S22. Recruitment of DNAJ1 activates reporter gene expression in yeast, *N. benthamiana*, and N2a cell lines (A) Yeast-one-hybrid of Gal4BD fused constructs. Two biological replicates each, at 0, 1/10, 1/100 and 1/1000 cellular dilutions. **(B)** Luc/Ren ratio from *N. benthamiana* transient expression with ZF108 fused proteins co-expressed with a reporter construct containing either the ZF108 target site or scrambled sequence in the LUC promoter. VP160, containing ten

copies of the VP16 transcriptional activator, is used as a positive control. Five independent biological replicates from different plant leaves were used per combination. **(**C**)** Luc/Ren ratio from N2a FlipIN cell line transfections with plasmids indicated. VP16 is used as a positive control. Mock = reporter and renilla plasmids only. Error bars represent 95% confidence intervals from three biological replicates.

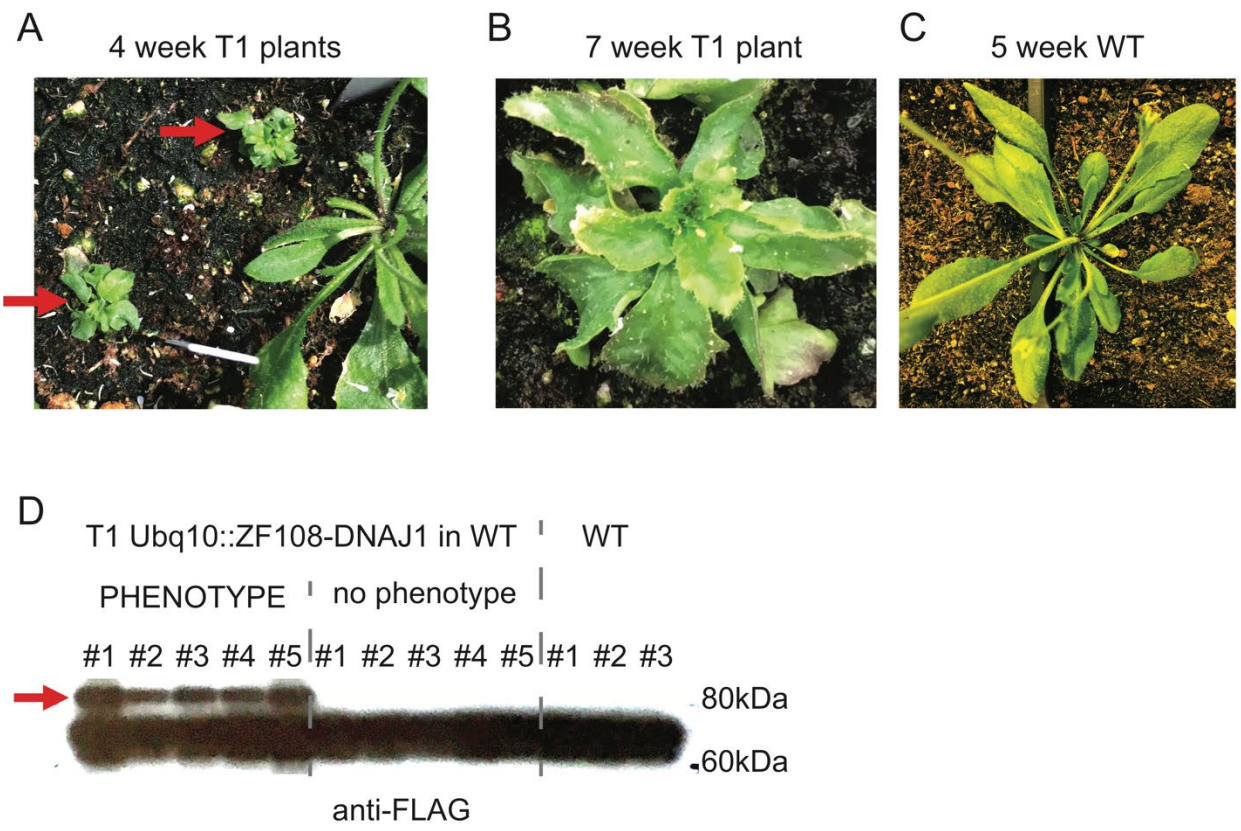

Fig. S23. *UBQ10::*ZF108-DNAJ1 transgenic plants display severe morphological defects. (A) Four-week-old *UBQ10::*ZF108-DNAJ1 in wild type (WT) T1 transgenic plants (red arrows) growing alongside a phenotypically normal Col-0 *A. thaliana* plant. **(B)** Seven-week-old *UBQ10::*ZF108-DNAJ1 T1 transgenic plant, with severe morphological defects that does not flower. **(C)** Five-week-old WT non-transgenic phenotypically normal flowering Col-0 *A. thaliana* plant for comparison. **(D)** Western blot of 5 randomly chosen *UBQ10::*ZF108-DNAJ1 T1 transgenic plants that displayed morphological phenotypes vs. 5 randomly chosen *UBQ10::*ZF108-DNAJ1 T1 transgenic plants that did not display morphological phenotypes and three non-transgenic control plants. Only the T1 transgenic plants with morphological phenotypes express the transgene (upper band, see red arrow). Lower cross reactive band serves as a loading control.

UBQ10::ZF108-DNAJ1			
	Up DEGs	**Down DEGs**	**Non DEGs**
Total (gene number)	1689	1442	30469
Overlap with UBQ10::ZF108-DNAJ1 peaks	474 (28.1%)	184 (12.8%)	3509 (11.5%)
observed/expected ratio	2.26	1.03	0.93
Hypergeometric test	7.70E-71 (over-represented)	0.348 (n.s.)	1.72E-46 (under-represented)
UBQ10::DNAJ1			
	Up DEGs	**Down DEGs**	**Non DEGs**
Total (gene number)	10	1	33591
Overlap with UBQ10::ZF108-DNAJ1 peaks	1 (10%)	0 (0%)	4166 (12.4%)
observed/expected ratio	0.81	0	1.00
Hypergeometric test	0.643 (n.s.)	NA	0.596 (n.s.)
UBQ10::ZF108-YPET			
	Up DEGs	**Down DEGs**	**Non DEGs**
Total (gene number)	380	55	33167
Overlap with UBQ10::ZF108-DNAJ1 peaks	46 (12.1%)	7 (12.7%)	4114 (12.4%)
observed/expected ratio	0.98	1.03	1.00
Hypergeometric test	0.469 (n.s.)	0.532 (n.s.)	0.481 (n.s.)

Background reference:
Total number of genes = 33602
Total number of UBQ10::ZF108-DNAJ1 peaks = 4951
Genes that overlap with UBQ10::ZF108-DNAJ1 peaks = 4167 (12.4%)

Fig. S24. Table showing association of ZF108-DNAJ1 binding with differentially expressed genes. Table shows the number of differentially expressed genes (FDR<0.01, FC>1.5), and tests for the significance of overlap (hypergeometric tests) between these gene sets and the ZF108-DNAJ1 binding sites, in the transgenic plant lines indicated.

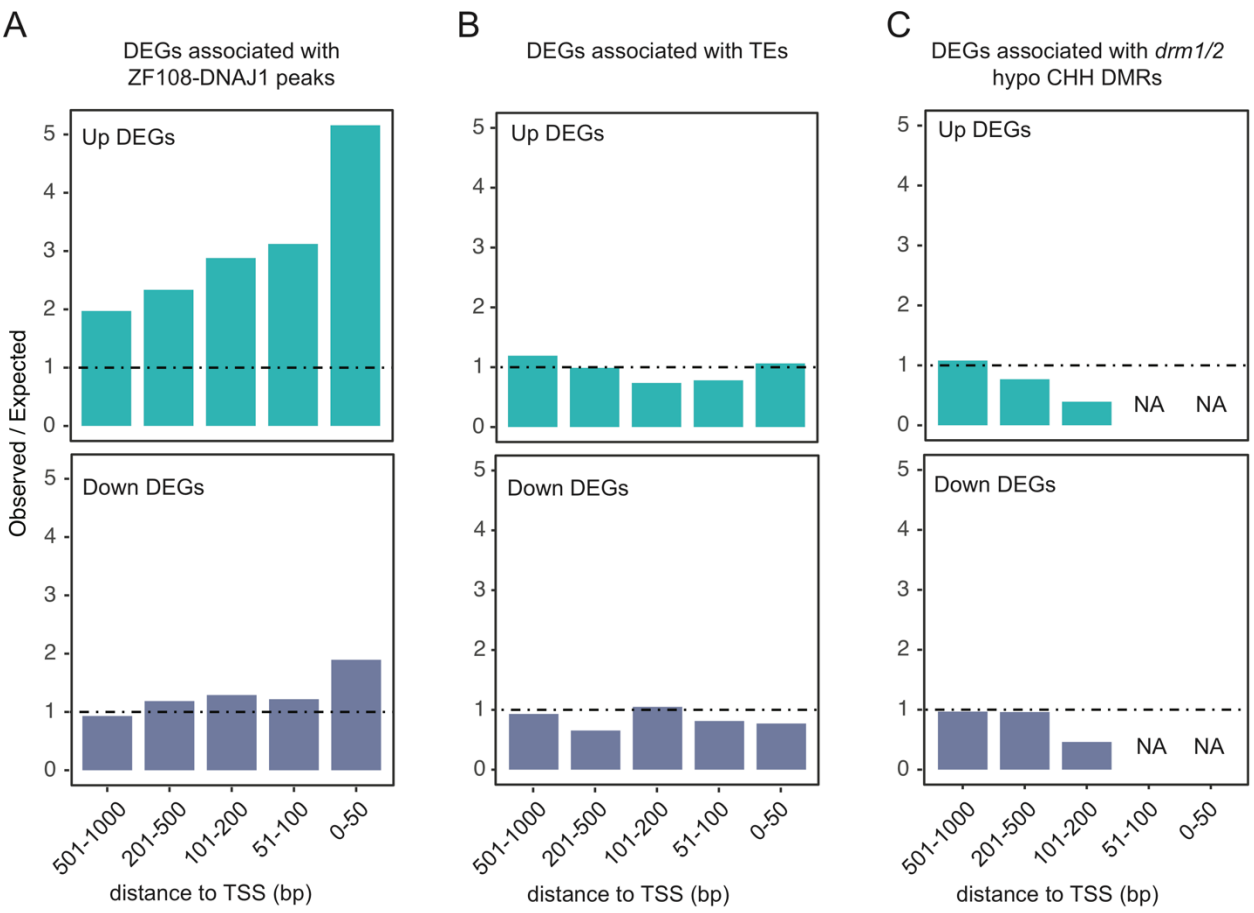

Fig. S25. Association of ZF108-DNAJ1 DEGs with ChIP binding sites, TEs and *drm1/2* hypo CHH DMRs. Observed / expected ratio for promoter overlap of ZF108-DNAJ1 differentially expressed genes proximal to **(A)** ZF108-DNAJ1 peaks, **(B)** transposable elements (TEs), or **(C)** *drm1/2* hypo CHH DMRs. Note that (A) is the same data as presented Fig. 4D, shown again here for comparison.

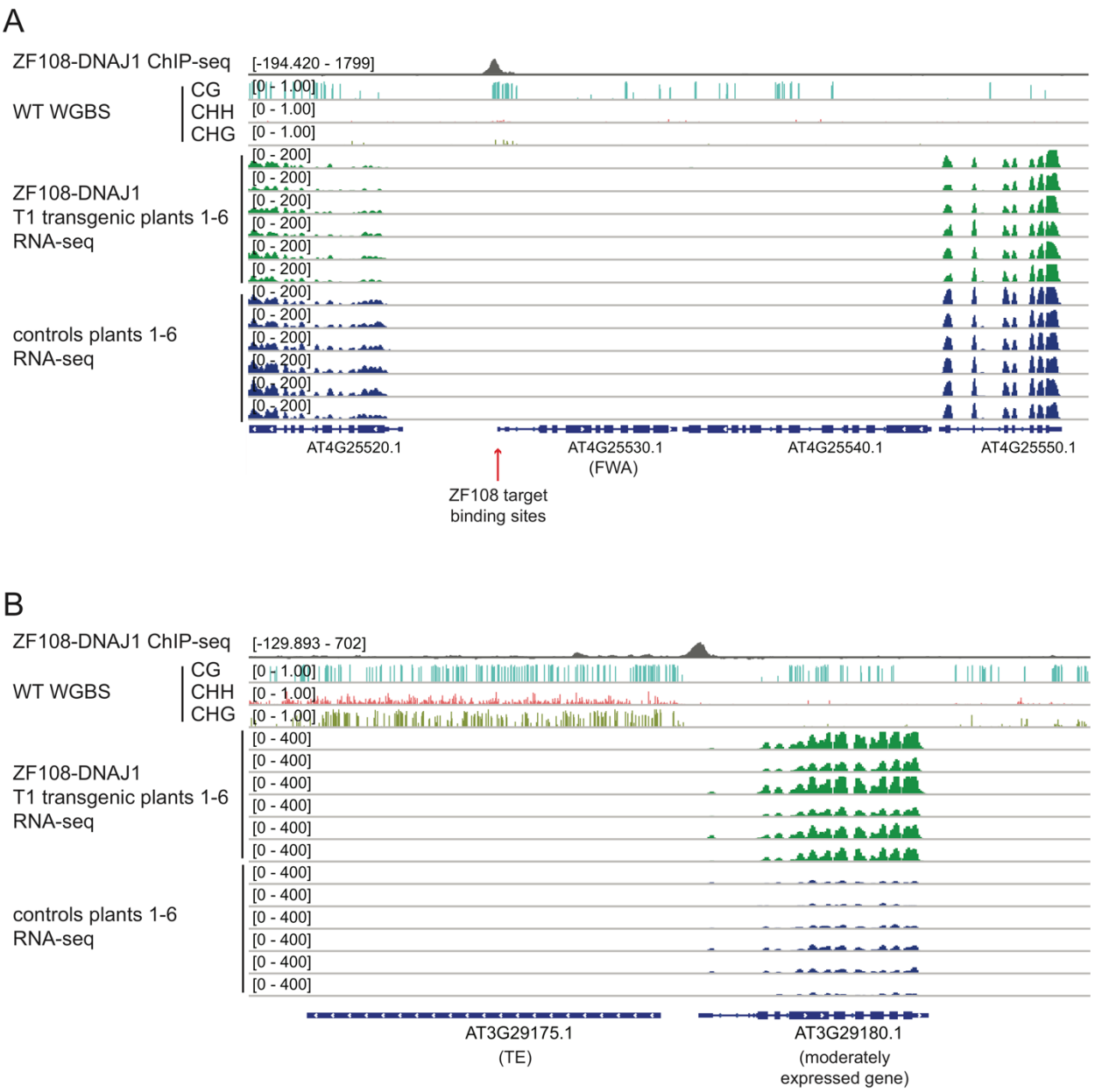

Fig. S26. ZF108-DNAJ1 promotes the expression of mildly expressed genes. (A) Browser track showing the *UBQ10::*ZF108-DNAJ1 ChIP-seq profile at FWA (also see Fig 4A) along with RNA-seq from 6 independent T1 *UBQ10::*ZF108-DNAJ1 transgenic plants (green) and 6 control plants (blue). Note lack of induced expression of *FWA* in ZF108-DNAJ1 transgenic lines. **(B)** Browser track at a ZF108-DNAJ1 associated upregulated gene, AT3G29180. Note that AT3G29180 is

expressed in WT, and shows increased expression in ZF108-DNAJ1, while the adjacent TE (methylated and silent in WT) remains silent.

References and Notes

1. J. Du, L. M. Johnson, S. E. Jacobsen, D. J. Patel, DNA methylation pathways and their crosstalk with histone methylation. *Nat. Rev. Mol. Cell Biol.* **16**, 519–532 (2015). doi:10.1038/nrm4043 Medline

2. S. Feng, S. J. Cokus, X. Zhang, P.-Y. Chen, M. Bostick, M. G. Goll, J. Hetzel, J. Jain, S. H. Strauss, M. E. Halpern, C. Ukomadu, K. C. Sadler, S. Pradhan, M. Pellegrini, S. E. Jacobsen, Conservation and divergence of methylation patterning in plants and animals. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 8689–8694 (2010). doi:10.1073/pnas.1002720107 Medline

3. A. Zemach, I. E. McDaniel, P. Silva, D. Zilberman, Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* **328**, 916–919 (2010). doi:10.1126/science.1186366 Medline

4. M. A. Matzke, R. A. Mosher, RNA-directed DNA methylation: An epigenetic pathway of increasing complexity. *Nat. Rev. Genet.* **15**, 394–408 (2014). doi:10.1038/nrg3683 Medline

5. A. E. Van't Hof, P. Campagne, D. J. Rigden, C. J. Yung, J. Lingley, M. A. Quail, N. Hall, A. C. Darby, I. J. Saccheri, The industrial melanism mutation in British peppered moths is a transposable element. *Nature* **534**, 102–105 (2016). doi:10.1038/nature17951 Medline

6. T. Stuart, S. R. Eichten, J. Cahn, Y. V. Karpievitch, J. O. Borevitz, R. Lister, Population scale mapping of transposable element diversity reveals links to gene regulation and epigenomic variation. *eLife* **5**, e20777 (2016). doi:10.7554/eLife.20777 Medline

7. L. Quadrana, A. Bortolini Silveira, G. F. Mayhew, C. LeBlanc, R. A. Martienssen, J. A. Jeddeloh, V. Colot, The Arabidopsis thaliana mobilome and its impact at the species level. *eLife* **5**, e15716 (2016). doi:10.7554/eLife.15716 Medline

8. J. D. Hollister, B. S. Gaut, Epigenetic silencing of transposable elements: A trade-off between reduced transposition and deleterious effects on neighboring gene expression. *Genome Res.* **19**, 1419–1428 (2009). doi:10.1101/gr.091678.109 Medline

9. J. A. Law, S. E. Jacobsen, Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat. Rev. Genet.* **11**, 204–220 (2010). doi:10.1038/nrg2719 Medline

10. K. Bahar Halpern, T. Vana, M. D. Walker, Paradoxical role of DNA methylation in activation of FoxA2 gene expression during endoderm development. *J. Biol. Chem.* **289**, 23882–23892 (2014). doi:10.1074/jbc.M114.573469 Medline

11. B. P. Williams, D. Pignatta, S. Henikoff, M. Gehring, Methylation-sensitive expression of a DNA demethylase gene serves as an epigenetic rheostat. *PLOS Genet.* **11**, e1005142 (2015). doi:10.1371/journal.pgen.1005142 Medline

12. M. Lei, H. Zhang, R. Julian, K. Tang, S. Xie, J.-K. Zhu, Regulatory link between DNA methylation and active demethylation in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 3553–3557 (2015). doi:10.1073/pnas.1502279112 Medline

13. K. Shibuya, S. Fukushima, H. Takatsuji, RNA-directed DNA methylation induces transcriptional activation in plants. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 1660–1665 (2009). doi:10.1073/pnas.0809294106 Medline

14. J. Moreno-Romero, H. Jiang, J. Santos-González, C. Köhler, Parental epigenetic asymmetry of PRC2-mediated histone modifications in the *Arabidopsis* endosperm. *EMBO J.* **35**, 1298–1311 (2016). doi:10.15252/embj.201593534 Medline

15. D. Kappei, M. Scheibe, M. Paszkowski-Rogacz, A. Bluhm, T. I. Gossmann, S. Dietz, M. Dejung, H. Herlyn, F. Buchholz, M. Mann, F. Butter, Phylointeractomics reconstructs functional evolution of protein binding. *Nat. Commun.* **8**, 14334 (2017). doi:10.1038/ncomms14334 Medline

16. L. O. Baumbusch, T. Thorstensen, V. Krauss, A. Fischer, K. Naumann, R. Assalkhou, I. Schulz, G. Reuter, R. B. Aalen, The Arabidopsis thaliana genome contains at least 29 active genes encoding SET domain proteins that can be assigned to four evolutionarily conserved classes. *Nucleic Acids Res.* **29**, 4319–4333 (2001). doi:10.1093/nar/29.21.4319 Medline

17. S. Li *et al*., SUVH1, a Su(var)3-9 family member, promotes the expression of genes targeted by DNA methylation. *Nucleic Acids Res.* **44**, 608–620 (2016). Medline

18. J. Du, L. M. Johnson, M. Groth, S. Feng, C. J. Hale, S. Li, A. A. Vashisht, J. Gallego-Bartolome, J. A. Wohlschlegel, D. J. Patel, S. E. Jacobsen, Mechanism of DNA methylation-directed histone methylation by KRYPTONITE. *Mol. Cell* **55**, 495–504 (2014). doi:10.1016/j.molcel.2014.06.009 Medline

19. W. Liu, S. H. Duttke, J. Hetzel, M. Groth, S. Feng, J. Gallego-Bartolome, Z. Zhong, H. Y. Kuo, Z. Wang, J. Zhai, J. Chory, S. E. Jacobsen, RNA-directed DNA methylation involves co-transcriptional small-RNA-guided slicing of polymerase V transcripts in Arabidopsis. *Nat. Plants* **4**, 181–188 (2018). doi:10.1038/s41477-017-0100-y Medline

20. H. Stroud, T. Do, J. Du, X. Zhong, S. Feng, L. Johnson, D. J. Patel, S. E. Jacobsen, Non-CG methylation patterns shape the epigenetic landscape in Arabidopsis. *Nat. Struct. Mol. Biol.* **21**, 64–72 (2014). doi:10.1038/nsmb.2735 Medline

21. A. Zemach, M. Y. Kim, P.-H. Hsieh, D. Coleman-Derr, L. Eshed-Williams, K. Thao, S. L. Harmer, D. Zilberman, The Arabidopsis nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. *Cell* **153**, 193–205 (2013). doi:10.1016/j.cell.2013.02.033 Medline

22. X. Zhong, C. J. Hale, J. A. Law, L. M. Johnson, S. Feng, A. Tu, S. E. Jacobsen, DDR complex facilitates global association of RNA polymerase V to promoters and evolutionarily young transposons. *Nat. Struct. Mol. Biol.* **19**, 870–875 (2012). doi:10.1038/nsmb.2354 Medline

23. J. P. Jackson, A. M. Lindroth, X. Cao, S. E. Jacobsen, Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* **416**, 556–560 (2002). doi:10.1038/nature731 Medline

24. L. M. Johnson, J. Du, C. J. Hale, S. Bischof, S. Feng, R. K. Chodavarapu, X. Zhong, G. Marson, M. Pellegrini, D. J. Segal, D. J. Patel, S. E. Jacobsen, SRA- and SET-domain-

containing proteins link RNA polymerase V occupancy to DNA methylation. *Nature* **507**, 124–128 (2014). doi:10.1038/nature12931 Medline

25. T. Y. Erkina, A. M. Erkine, Nucleosome distortion as a possible mechanism of transcription activation domain function. *Epigenetics Chromatin* **9**, 40 (2016). doi:10.1186/s13072-016-0092-2 Medline

26. L. G. Lowder, D. Zhang, N. J. Baltes, J. W. Paul 3rd, X. Tang, X. Zheng, D. F. Voytas, T.-F. Hsieh, Y. Zhang, Y. Qi, A CRISPR/Cas9 Toolbox for Multiplexed Plant Genome Editing and Transcriptional Regulation. *Plant Physiol.* **169**, 971–985 (2015). doi:10.1104/pp.15.00636 Medline

27. X. Wang, D. Wiegel, L. M. Smith, Transposon variants and their effects on gene expression in Arabidopsis. *PLOS Genet*. **9**, e1003255 (2013). doi:10.1371/journal.pgen.1003255.

28. D. Meng, M. Dubin, P. Zhang, E. J. Osborne, O. Stegle, R. M. Clark, M. Nordborg, Limited Contribution of DNA Methylation Variation to Expression Regulation in Arabidopsis thaliana. *PLOS Genet.* **12**, e1006141 (2016). doi:10.1371/journal.pgen.1006141 Medline

29. K. Naito, F. Zhang, T. Tsukiyama, H. Saito, C. N. Hancock, A. O. Richardson, Y. Okumoto, T. Tanisaka, S. R. Wessler, Unexpected consequences of a sudden and massive transposon amplification on rice gene expression. *Nature* **461**, 1130–1134 (2009). doi:10.1038/nature08479 Medline

30. V. Sundaram, T. Wang, Transposable Element Mediated Innovation in Gene Regulatory Landscapes of Cells: Re-Visiting the "Gene Battery" Model. *BioEssays* **40**, 170155 (2018). doi:10.1002/bies.201700155 Medline

31. H. Ito, H. Gaubert, E. Bucher, M. Mirouze, I. Vaillant, J. Paszkowski, An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. *Nature* **472**, 115–119 (2011). doi:10.1038/nature09861 Medline

32. H. Y. Park, S.-Y. Lee, H.-Y. Seok, S.-H. Kim, Z. R. Sung, Y.-H. Moon, EMF1 interacts with EIP1, EIP6 or EIP9 involved in the regulation of flowering time in Arabidopsis. *Plant Cell Physiol.* **52**, 1376–1388 (2011). doi:10.1093/pcp/pcr084 Medline

33. S. J. Clough, A. F. Bent, Floral dip: A simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J.* **16**, 735–743 (1998). doi:10.1046/j.1365-313x.1998.00343.x Medline

34. S. E. Jacobsen, E. M. Meyerowitz, Hypermethylated SUPERMAN epigenetic alleles in arabidopsis. *Science* **277**, 1100–1103 (1997). doi:10.1126/science.277.5329.1100 Medline

35. I. R. Henderson, S. E. Jacobsen, Tandem repeats upstream of the Arabidopsis endogene SDC recruit non-CG DNA methylation and initiate siRNA spreading. *Genes Dev.* **22**, 1597–1606 (2008). doi:10.1101/gad.1667808 Medline

36. W. J. Soppe, S. E. Jacobsen, C. Alonso-Blanco, J. P. Jackson, T. Kakutani, M. Koornneef, A. J. M. Peeters, The late flowering phenotype of fwa mutants is caused by gain-of-function epigenetic alleles of a homeodomain gene. *Mol. Cell* **6**, 791–802 (2000). doi:10.1016/S1097-2765(05)00090-0 Medline

37. X. Cao, S. E. Jacobsen, Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. *Proc. Natl. Acad. Sci. U.S.A.* **99** (Suppl 4), 16491–16498 (2002). doi:10.1073/pnas.162371599 Medline

38. J. Cox, M. Mann, MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008). doi:10.1038/nbt.1511 Medline

39. J. S. Harrison, E. M. Cornett, D. Goldfarb, P. A. DaRosa, Z. M. Li, F. Yan, B. M. Dickson, A. H. Guo, D. V. Cantu, L. Kaustov, P. J. Brown, C. H. Arrowsmith, D. A. Erie, M. B. Major, R. E. Klevit, K. Krajewski, B. Kuhlman, B. D. Strahl, S. B. Rothbart, Hemi-methylated DNA regulates DNA methylation inheritance through allosteric activation of H3 ubiquitylation by UHRF1. *eLife* **5**, e17101 (2016). doi:10.7554/eLife.17101 Medline

40. E. Rajakumara, J. A. Law, D. K. Simanshu, P. Voigt, L. M. Johnson, D. Reinberg, D. J. Patel, S. E. Jacobsen, A dual flip-out mechanism for 5mC recognition by the Arabidopsis SUVH5 SRA domain and its impact on DNA methylation and H3K9 dimethylation in vivo. *Genes Dev.* **25**, 137–152 (2011). doi:10.1101/gad.1980311 Medline

41. X. Li, C. J. Harris, Z. Zhong, W. Chen, R. Liu, B. Jia, Z. Wang, S. Li, S. E. Jacobsen, J. Du, Mechanistic insights into plant SUVH family H3K9 methyltransferases and their binding to context-biased non-CG DNA methylation. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E8793–E8802 (2018). Medline

42. J. Gallego-Bartolomé, J. Gardiner, W. Liu, A. Papikian, B. Ghoshal, H. Y. Kuo, J. M.-C. Zhao, D. J. Segal, S. E. Jacobsen, Targeted DNA demethylation of the *Arabidopsis* genome using the human TET1 catalytic domain. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E2125–E2134 (2018). doi:10.1073/pnas.1716945115 Medline

43. K. H. Emami, M. Carey, A synergistic increase in potency of a multimerized VP16 transcriptional activation domain. *EMBO J.* **11**, 5005–5012 (1992). doi:10.1002/j.1460-2075.1992.tb05607.x Medline

44. J. D. Buenrostro, B. Wu, H. Y. Chang, W. J. Greenleaf, ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr. Protoc. Mol. Biol.* **109**, 1–9 (2015). Medline

45. Z. Lu, B. T. Hofmeister, C. Vollmers, R. M. Dubois, J. Schmitz, Combining ATAC-seq with nuclei sorting for discovery of cis-regulatory regions in plant genomes. *Nucleic Acids Res.* **45**, e41 (2017). Medline

46. C. J. Harris, D. Husmann, W. Liu, F. E. Kasmi, H. Wang, A. Papikian, W. A. Pastor, G. Moissiard, A. A. Vashisht, J. L. Dangl, J. A. Wohlschlegel, S. E. Jacobsen, Arabidopsis AtMORC4 and AtMORC7 Form Nuclear Bodies and Repress a Large Number of Protein-Coding Genes. *PLOS Genet.* **12**, e1005998 (2016). doi:10.1371/journal.pgen.1005998 Medline

47. D. Kim, G. Pertea, C. Trapnell, H. Pimentel, R. Kelley, S. L. Salzberg, TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013). doi:10.1186/gb-2013-14-4-r36 Medline

48. Y. Xi, W. Li, BSMAP: Whole genome bisulfite sequence MAPping program. *BMC Bioinformatics* **10**, 232 (2009). doi:10.1186/1471-2105-10-232 Medline

49. Z. Gu, R. Eils, M. Schlesner, Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016). doi:10.1093/bioinformatics/btw313 Medline

50. Y. Zhang, C. J. Harris, Q. Liu, W. Liu, I. Ausin, Y. Long, L. Xiao, L. Feng, X. Chen, Y. Xie, X. Chen, L. Zhan, S. Feng, J. J. Li, H. Wang, J. Zhai, S. E. Jacobsen, Large-scale comparative epigenomics reveals hierarchical regulation of non-CG methylation in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E1069–E1074 (2018). Medline

51. B. Langmead, C. Trapnell, M. Pop, S. L. Salzberg, Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009). doi:10.1186/gb-2009-10-3-r25 Medline

52. Y. Zhang, T. Liu, C. A. Meyer, J. Eeckhoute, D. S. Johnson, B. E. Bernstein, C. Nusbaum, R. M. Myers, M. Brown, W. Li, X. S. Liu, Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008). doi:10.1186/gb-2008-9-9-r137 Medline

53. L. Shen, N. Shao, X. Liu, E. Nestler, ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. *BMC Genomics* **15**, 284 (2014). doi:10.1186/1471-2164-15-284 Medline

54. F. Ramírez, D. P. Ryan, B. Grüning, V. Bhardwaj, F. Kilpert, A. S. Richter, S. Heyne, F. Dündar, T. Manke, deepTools2: A next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016). doi:10.1093/nar/gkw257 Medline

55. W. A. Pastor, W. Liu, D. Chen, J. Ho, R. Kim, T. J. Hunt, A. Lukianchikov, X. Liu, J. M. Polo, S. E. Jacobsen, A. T. Clark, TFAP2C regulates transcription in human naive pluripotency by opening enhancers. *Nat. Cell Biol.* **20**, 553–564 (2018). doi:10.1038/s41556-018-0089-0 Medline

56. S. Anders, W. Huber, Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010). doi:10.1186/gb-2010-11-10-r106 Medline

57. H. T. Rube, W. Lee, M. Hejna, H. Chen, D. H. Yasui, J. F. Hess, J. M. LaSalle, J. S. Song, Q. Gong, Sequence features accurately predict genome-wide MeCP2 binding in vivo. *Nat. Commun.* **7**, 11025 (2016). doi:10.1038/ncomms11025 Medline

58. G. Moissiard, S. Bischof, D. Husmann, W. A. Pastor, C. J. Hale, L. Yen, H. Stroud, A. Papikian, A. A. Vashisht, J. A. Wohlschlegel, S. E. Jacobsen, Transcriptional gene silencing by Arabidopsis microrchidia homologues involves the formation of heteromers. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 7474–7479 (2014). doi:10.1073/pnas.1406611111 Medline

59. X. Zhong, J. Du, C. J. Hale, J. Gallego-Bartolome, S. Feng, A. A. Vashisht, J. Chory, J. A. Wohlschlegel, D. J. Patel, S. E. Jacobsen, Molecular mechanism of action of plant DRM de novo DNA methyltransferases. *Cell* **157**, 1050–1060 (2014). doi:10.1016/j.cell.2014.03.056 Medline

60. I. Ausin, M. V. C. Greenberg, D. K. Simanshu, C. J. Hale, A. A. Vashisht, S. A. Simon, T. F. Lee, S. Feng, S. D. Española, B. C. Meyers, J. A. Wohlschlegel, D. J. Patel, S. E. Jacobsen, INVOLVED IN DE NOVO 2-containing complex involved in RNA-directed DNA methylation in Arabidopsis. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 8374–8381 (2012). doi:10.1073/pnas.1206638109 Medline