

Supplementary Materials for

MORC family ATPases required for heterochromatin condensation and gene silencing

Guillaume Moissiard¹, Shawn J. Cokus¹, Joshua Cary¹, Suhua Feng¹, Allison C. Billi², Hume Stroud¹, Dylan Husmann¹, Ye Zhan³, Bryan R. Lajoie³, Rachel P. McCord³, Christopher J. Hale¹, Wei Feng⁴, Scott D. Michaels⁴, Alison R. Frand⁵, Matteo Pellegrini^{1,6}, Job Dekker³, John K. Kim², and Steve Jacobsen^{1,5,6,7}*

¹Department of Molecular, Cell, and Developmental Biology, University of California at Los Angeles, Terasaki Life Sciences Building, 610 Charles Young Drive East, Los Angeles, California 90095-723905, USA.

²Life Sciences Institute and Department of Human Genetics, University of Michigan, Ann Arbor, Michigan, USA.

³Lazare Research Building 570N, Program in Systems Biology and Gene Function and Expression, University of Massachusetts Medical School, 364 Plantation Street, Worcester, MA, 01605, USA

⁴Department of Biology, Indiana University, Bloomington, Indiana, USA.

⁵Department of Biological Chemistry, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, California, USA.

⁶Eli & Edythe Broad Center of Regenerative Medicine & Stem Cell Research, University of California Los Angeles, Los Angeles, California, USA.

⁷Howard Hughes Medical Institute, University of California Los Angeles, Los Angeles, California, USA.

*Correspondence to: jacobsen@ucla.edu

Material and Methods:

Plant material and growing conditions.

All mutants are in the Columbia (Col) ecotype. *atmorc6-3* (GK_599B06) and *atmorc1-4* (SAIL_1239_C08) T-DNA lines were obtained from GABI-Kat (24) at University of Bielefeld, Germany and ABRC at Ohio State University, respectively. *drm2-2* (SALK_150863.37.35) and *cmt3-11* (SALK_148381) were previously described (4). T-DNA insertions were confirmed by PCR-based genotyping. Primer sequences are described in [Table S4](#). *Arabidopsis* plants were grown under continuous light.

Cloning of *SDC::GFP*.

NLS-GFP-35S terminator was PCR amplified and cloned into pCambia3300. The *SDC* promoter corresponding to a region of ~2.4 kb upstream of *SDC* transcriptional start site was PCR amplified from wild type genomic DNA and cloned into pCR2.1 TOPO vector (Invitrogen). Quick change site-directed mutagenesis (Stratagene) was performed to create a polymorphism (NlaIII -> BamHI) within the *SDC* promoter, which was subsequently mobilized into pCambia3300 upstream of NLS-GFP sequence. *drm2 cmt3* double mutant plants were transformed with the *SDC::GFP* construct using the *Agrobacterium*-mediated floral dip method (25). Transgenic plants showing strong GFP fluorescence were backcrossed with a wild type plant to ensure proper silencing of *SDC::GFP* in the F₁ generation. F₁ plants were self-crossed and their progenies (F₂) were screened for GFP fluorescence and PCR-based genotyped to obtain the following genetic backgrounds: *SDC::GFP* wt, *SDC::GFP drm2*, *SDC::GFP cmt3* and *SDC::GFP drm2 cmt3*. Primer sequences used for *SDC::GFP* cloning are described in [Table S4](#).

Cloning of *pAtMORC1::AtMORC1-Myc* and *pAtMORC6::AtMORC6-Myc*.

AtMORC1 and *AtMORC6* genomic regions were PCR amplified and the Myc epitope was added to the C-terminus of each protein as previously described (26). In both cases, the amplified region includes a ~1Kb promoter sequence upstream of the respective transcriptional start site. Primer sequences are described in [Table S4](#).

EMS mutagenesis, GFP screening and mapping analyses.

Two thousand seeds from *SDC::GFP* wt and *SDC::GFP cmt3* lines were mutagenized in 0.3% EMS solution for 13 hours with rotation. Seeds were subsequently washed with water and planted onto soil. For each background, approximately one thousand M2 populations were collected and subsequently screened for GFP fluorescence under UV light using a Leica MZ16F Fluorescence Stereomicroscope coupled with the GFP Plus fluorescence filter. Pictures were taken using the DFC300 FX digital camera kit. Mapping and identification of the three EMS mutations responsible for the phenotypes were performed by bulk segregant analysis coupled with deep genome re-sequencing as previously described (27), using single nucleotide polymorphisms (SNPs) between the Landsberg (*Ler*) and Col ecotypes derived *de novo* from data from a large number of mapping crosses.

Western Blotting.

Western blots against GFP were performed using the GFP-specific antibody (Invitrogen, AA1122). Western blots against Myc were performed as previously described (26).

RNA analyses.

Total RNAs were extracted from two-week-old seedlings using Trizol (Ambion RNA technology). Two μg of total RNAs were subsequently used to generate libraries for High Throughput RNA sequencing (TruSeq RNA, Illumina) per manufacturer instruction. For RNA-seq analyses, sequencing reads were mapped with Bowtie (28) allowing up to 2 mismatches. Gene and transposon expression was measured by calculating reads per kilobase per million mapped reads (RPKM) (29). *p*-values were calculated using Fisher's exact test and Benjamini corrected for multiple testing (30). Differentially expressed elements in wild type and mutants were defined by applying $\log_2(\text{mutant}/\text{wild type}) > 2$ and $P < 0.05$ cutoffs. For quantitative PCR analysis, total RNAs were converted into cDNA using SuperScript III Reverse Transcriptase (Invitrogen) per manufacturer instructions. Quantitative PCR was carried out using SyBr Green PCR mastermix (Roche) and gene- or transposon-specific primers (see [Table S4](#)) on a Stratagene Mx real-time thermocycler.

DNA methylation analyses.

Whole genome BS-seq libraries were performed as previously described (9), except directly with pre-methylated final adapters. BS-seq data was mapped with BS seeker as previously described (31). For traditional bisulfite sequencing, genomic DNA extracted from two-week-old seedlings was bisulfite converted using MethylEasy (Human Genetic Signature) and processed as previously described (5). Primer sequences used for bisulfite sequencing are described in [Table S4](#). Southern blot was performed as previously described (2).

H3K9me2 ChIP-seq analyses.

Two grams of 3-week-old seedlings were crosslinked with formaldehyde and chromatin-IP experiments were performed as previously described (32). A mouse monoclonal antibody was used for H3K9me2 immunoprecipitation (Abcam ab1220). ChIP-seq library was generated per manufacturer instructions (Illumina). Demultiplexed (by exact match to canonical 6-mers) single end 50-mer HiSeq PF-passing reads were aligned to the TAIR8 reference genome using Bowtie 1, keeping all hits with at most two or fewer mismatches in the first 28 cycles and with total sum of Phred quality scores at mismatches up to 100, further filtered to only keep reads with a unique hit of fewest total mismatches, retaining only that unique hit. Reads were extended downstream to total length 220 nucleotides to reflect nominal library fragment lengths and single-stranded per-base pair coverage tallied, normalized by total nuclear chromosome coverage to account for variation in sequencing depth.

siRNA analyses.

Total RNAs were extracted from flowers using Trizol (Ambion RNA technology) and siRNAs were purified as previously described (33) with the following modifications. Polyacrylamide gel-excised siRNAs were eluted in 0.3M NaCl overnight at 4°C. Gel debris were filtered using 5µm Filter tubes (IST Engineering Inc) and, ethanol-precipitated siRNAs were resuspended in 5ul of nuclease-free H₂O to subsequently generate libraries for High Throughput small RNA sequencing (TruSeq small RNA, Illumina) per manufacturer instruction. Small RNA libraries were Illumina sequenced to 50bp length, and resulting reads were trimmed for adapter sequences and then aligned to the TAIR8 genome using Bowtie (28).

Hi-C analyses.

Two grams of 3-week-old seedling leaves were crosslinked with formaldehyde as previously described (32). Hi-C experiments were performed as previously described (34), with the exception that plant nuclei were prepared following a previously published *Arabidopsis* ChIP protocol (32). Hi-C libraries were sequenced on a HiSeq 2000 sequencer (Illumina) obtaining paired end 50+50 nucleotide reads. Sequencing reads were mapped to the TAIR8 *A. thaliana* reference genome using Bowtie 1 to obtain all zero-mismatch hits of ends independently and keeping only paired end read pairs with each end having exactly one hit, obtaining 21,379,391 wild type and 14,815,038 *atmorc6-1* pairs. Paired reads with ends aligning to the same HindIII fragment were discarded. Hi-C interaction counts were summed within disjoint symmetric 2-D bins 200 kilobase pairs tiling the genome. HindIII fragments which overlapped regions of poor reference genome quality were excluded from the analysis. Genomic 1-D bins (rows and columns) in which > 50% of the sequence length was excluded by these filters were treated as missing data, excluded from further analyses, and appear in figures as empty regions. The “raw” coverage of each genomic 2-D bin was taken as the number of paired end reads lying in that bin. Whole genomic 1-D bins with a total coverage more than 3 standard deviations greater than or less than the mean were excluded and then the matrix of interactions was corrected for 1-D bin coverage variation as previously described (34) using 50 iterations of that procedure. The comparison between the wild type and *atmorc6-1* mutant was expressed as the difference divided by the mean within each bin with smoothing plus or minus one bin.

Immunofluorescence.

Immunofluorescence experiments examining chromocenter condensation were performed as previously described (35) with the following modifications. Leaves from three-week-old plants were fixed in 4% paraformaldehyde in TRIS buffer (10mM TRIS pH 7.5, 10mM EDTA, 100mM NaCl) for 20 minutes and washed twice in TRIS buffer. Leaves were chopped in 400 microliters lysis buffer (15mM TRIS pH 7.5, 2mM EDTA, 0.5mM spermine, 80mM KCl, 20mM NaCl, 0.1% Triton X-100) and filtered through a 35 micron cell strainer. Five microliters of nuclei suspension was added to sorting buffer (100mM TRIS pH 7.5, 50mM KCl, 2mM MgCl₂, .05% Tween-20, 20.5% sucrose) and air dried on microscope slides for two hours and then

post-fixed in 4% paraformaldehyde in PBS for 20 minutes. Slides were washed three times in PBS and incubated in blocking buffer (3% BSA, 10% horse serum in PBS) for 30 minutes at 37°C. Nuclei were incubated at 4°C overnight in mouse monoclonal antibody against H3K9me2 (Abcam ab1220; 1:200). Slides were washed in PBS and incubated with goat anti-mouse FITC antibody (Abcam ab7064; 1:200) for 90 minutes at room temperature. Following PBS washes, nuclei were counterstained and mounted in Vectashield mounting media with DAPI (Vector H-1200). Nuclei were analyzed with a Zeiss Axio Imager Z1 microscope at 100X magnification and images were captured with a Hamamatsu ORCA-ER Camera. For detection of Myc epitope tagged proteins, nuclei were isolated as above. Following preparation of nuclei suspension, nuclei were spun down for 2 minutes at 2,000rpm and resuspended in PBS. Blocking and antibody incubations were performed in suspension, followed by pellet washing with PBS. Myc epitope was detected with mouse monoclonal antibody (Abcam 9E10; 1:200) and goat anti-mouse FITC (Abcam ab7064; 1:200). Two microliters of prepared nuclei were mounted in Vectashield media. Nuclei were analyzed with the Applied Precision DeltaVision DV Live Cell Imaging System using Olympus IX-71 Customized Inverted Microscope and Photometrics CoolSNAP HQ2 CCD Camera (Figure 3). Additional images of nuclei were taken using a Zeiss LSM 510 META confocal microscope with Axiocam camera (Figure S15).

RNA interference in *C. elegans*.

RNAi experiments were carried out as reported previously (36) using the *eri-1(mg366); [wIs54(scm::gfp)]* strain, which shows increased sensitivity to RNAi. Briefly, bacterial strains carrying plasmids expressing double-stranded RNA targeting *morc-1* or *rde-4* were obtained from the Ahringer RNAi library (37). Hatched L1 *eri-1(mg366); [wIs54(scm::gfp)]* larvae were cultured on empty vector (L4440), *morc-1*, or *rde-4* RNAi bacteria for two generations at 22.5°C. Images of F₁ L4 larvae were captured on an Olympus BX61 epifluorescence compound microscope with a Hamamatsu ORCA ER camera using Slidebook 4.0.1 digital microscopy software (Intelligent Imaging Innovations) and processed using ImageJ.

Supplementary Figures:

Fig. S1. (A) Schematic representation of the *SDC::GFP* construct. The *SDC* promoter carries seven tandem repeats (black arrows) targeted by DNA methylation. The red bar corresponds to the *Simian Virus 40 (SV40)* Nuclear Localization Signal fused to GFP. **(B)** Schematic of AtMORC1 and AtMORC6 proteins showing the GHKL and S5 ATPase domains together with putative Coiled-Coil (CC) domains. The location of nonsense mutations within the protein sequences is shown for each EMS allele. aa; amino acid. Based on contextual associations of prokaryotic and eukaryotic MORC family members with associated domains, as well as with genes in operons, eukaryotic MORCs were predicted to have originated from bacterial restriction modification systems, and in eukaryotes have been proposed to remodel chromatin superstructure in response to epigenetic signals such as histone and DNA methylation (12). For instance, topoisomerases and MutL (a factor involved in methylated DNA directed mismatch repair) proteins use ATP hydrolysis to mediate large movements and looping of DNA. Similar to other GHKL ATPases (38), MORC3 in mouse was shown to act as a molecular clamp, interacting with itself constitutively through its coiled-coil domain, and also interacting via its ATPase domain in an ATP-dependent manner (39). In this way, MORC protein architectures are also reminiscent of the structural maintenance of chromosomes (SMC) family of proteins, which control chromosome condensation and cohesion (40). In addition, a very distant class of ATPase homologs contain the GHKL domain fused with the hinge and coiled coil domains of SMC-like ATPases (12) one of which, SmcHD1, is required for maintenance of X chromosome inactivation in mouse (41). Interestingly, in *Arabidopsis*, the protein DMS3/IDN1, which contains an SMC-like Hinge domain,

has also been involved in gene silencing (42, 43) and GMI1, which is a GHKL protein carrying a Hinge domain has recently been involved in DNA repair (44).

Fig. S2. Mapping of *cmt3* #7, wt #67 and *cmt3* #49 mutations by bulk segregant analysis coupled with whole genome re-sequencing (27). (A to C) Depletion in percentage of Landsberg (*Ler*) single nucleotide polymorphisms (SNPs), defining the linkage interval for the population *cmt3* #7 (chromosome 4 at ~17-18 megabases) (A), and defining the linkage intervals for the populations wt #67 (B) and *cmt3* #49 (C) (chromosome 1 at ~6-7 megabases). Red arrows mark the linkage intervals.

Fig. S3. Characterization of *atmorc1* and *atmorc6* EMS and T-DNA mutant alleles. (A) Location of EMS point mutations in *atmorc1-3*, *atmorc6-1*, and *atmorc6-2* alleles confirmed by traditional DNA sequencing. *AtMORC1* and *AtMORC6* reference sequences are shown for comparison. bp, base pair. (B) Gene structures of *AtMORC1* and *AtMORC6* showing exons (E, dark gray boxes) and introns (gray lanes). Light gray boxes correspond to 5' and 3' untranslated regions. The locations of EMS point mutations and T-DNA insertions are described above the gene structures. (C and D) Relative *AtMORC1* and *AtMORC6* RNA levels (C), and *SDC* RNA level (D) in *atmorc1-4* and *atmorc6-3* over wild type assayed by RT-qPCR and normalized to *ACTIN7*. Errors bars indicate standard deviation based on three independent biological replicates. Red arrows shown in (B) correspond to the primer locations within *AtMORC1* and *AtMORC6* mRNAs used in (C). (E) Genetic complementation tests and backcrosses showing that wt #67, *cmt3* #49, and *atmorc6-3* are three recessive allelic mutations in *AtMORC6* (top panels), while *cmt3* #7 and *atmorc1-4*

are two recessive allelic mutations in *AtMORC1* (two bottom left panels). The cross between wt #67 and *cmt3* #7 confirms that these two mutations are non-allelic (bottom right panel). Pictures represent leaves from F₁ plants observed under UV light for GFP fluorescence.

Fig. S4. Similar sets of TEs and protein-coding genes are upregulated in *atmorc1* and *atmorc6*. (A) Relative fold increase of *SDC* and four TE transcripts in *atmorc6-1*, *cmt3-11* and *cmt3-11 atmorc1-3* over wild type assayed by RT-qPCR and normalized to *ACTIN7*. Errors bars indicate standard deviation based on three independent biological replicates. (B-D) Venn diagrams showing overlap of upregulated TEs (B), upregulated protein-coding genes (C), and upregulated protein-coding genes associated with DNA methylation (D) in *atmorc1* and *atmorc6*. These analyses include the genes that were 4-fold up-regulated in RNA-seq experiments in both the EMS and T-DNA alleles. (E) Gene ontology (GO) analyses of genes upregulated in *atmorc1* and *atmorc6* showing no significant over-representation of any GO category in both *MORC* mutants. GO *p*-values are shown in parenthesis.

Fig. S5. DNA methylation at all protein-coding genes and transposons is not altered in *atmorc1* and *atmorc6*. (A and B) Metaplot analyses showing DNA methylation percent within protein-coding genes (A) and transposons (B) in *atmorc1-4*, *atmorc6-3* and wild type plants. The gray vertical lines mark the boundaries between 1 kilobase (Kb) upstream regions and gene bodies or transposons (left), and between gene bodies or transposons and 1Kb downstream regions (right).

Fig. S6. H3K9me2 is not altered in *atmorc1* and *atmorc6*. (A) \log_2 ratios of H3K9me2/H3 in *atmorc1-4* and *atmorc6-3* over wild type do not show H3K9me2 changes in *atmorc6* or *atmorc1* mutants at the *SDC* promoter region. The H3K9me2/H3 wild type ratio is shown over the *SDC* promoter region in which H3K9me2 is enriched. RNA-seq reads in the different genetic backgrounds are shown to define the *SDC* transcribed region. (B) Metaplot analyses showing \log_2 ratios of H3K9me2/H3 in *atmorc1-4* and *atmorc6-3* over wild type at the set of transposable elements and DNA-methylated genes that are upregulated in *atmorc1-4* and *atmorc6-3*. The gray vertical lines mark the boundaries between 1 kilobase upstream regions and transposons/PCGs (left) and between transposons/PCGs and 1 kilobase downstream regions (right). PCGs, Protein-coding genes.

Fig. S7. Small RNA accumulation is unaltered in *atmorc1* and *atmorc6*. Metaplot analyses showing no difference in the level of siRNAs in *atmorc1-4* and *atmorc6-3* compared to wild type at the set of transposable elements that are upregulated in *atmorc1-4* and *atmorc6-3*. The gray vertical lines mark the boundaries between 2 kilobases upstream regions and transposons (left) and between transposons and 2 kilobases downstream regions (right). Data is illustrated for different size classes of small RNAs.

Fig. S8. AtMORC1 and AtMORC6 are required for chromocenter condensation. Percentage of nuclei showing decondensed, partially decondensed (intermediate) and wild type chromocenters in *atmorc6-1*, *cmt3-11 atmorc1-3* and in *atmorc1-3 atmorc6-1* double mutants in comparison to control backgrounds after immunostaining of

nuclei using an antibody against H3K9me2. Pictures on top panels show examples of the three different patterns of chromocenters.

Fig. S9. H3K9me2 and DAPI staining show similar chromocenter condensation patterns in nuclei defined in Fig. S8 as nuclei showing wild type, intermediate or decondensed chromocenters in *cmt3-11 atmorc1-3*.

Fig. S10. H3K9me2 and DAPI staining show similar chromocenter condensation patterns in nuclei defined in Fig. S8 as nuclei showing wild type, intermediate or decondensed chromocenters in *atmorc6-1*.

Fig. S11. H3K9me2 and DAPI staining show similar chromocenter condensation patterns in nuclei defined in Fig. S8 as nuclei showing wild type, intermediate or decondensed chromocenters in *atmorc1-3 atmorc6-1*.

Fig. S12. Loci upregulated in *atmorc1* and *atmorc6* mostly localize to pericentromeric regions. Chromosomal views showing the log₂ ratios (*atmorc6-3/wild type* and *atmorc1-4/wild type*) of RNA sequencing reads in 100 Kb bins. The two red bars on each chromosome delimit the pericentromeric region with white circles representing the centromeres.

Fig. S13. Interaction matrix of the *atmorc6-1* genome from Hi-C analysis. Positions along the 5 chromosomes are shown from left to right and top to bottom, and each pixel represents interactions from uniquely mapping paired end reads in 200 kilobase bins. Black bars and circles mark the positions of the pericentromeric and telomeric

regions, respectively. Light grey regions represent areas masked out due to problematic mapping. Black bars show separation between chromosomes.

Fig. S14. Myc tagged AtMORC1 and AtMORC6 complement the EMS mutant lines.

(A) Western blot using an antibody against Myc confirms that pAtMORC1::AtMORC1-Myc and pAtMORC6::AtMORC6-Myc are expressed in transformed plants but not in untransformed control. (B) Similar analyses using an antibody against GFP show that lines expressing AtMORC1-Myc and AtMORC6-Myc do not express the *SDC::GFP* transgene, confirming that pAtMORC1::AtMORC1-Myc and pAtMORC6::AtMORC6-Myc proteins are functional. *cmt3-11 atmorc1-3* and *atmorc6-1* mutants are used as GFP positive controls for AtMORC1-Myc and AtMORC6-Myc complementation, respectively. Coomassie staining of the large Rubisco subunit (*rbcL*) is used as loading control.

Fig. S15. AtMORC6 bodies are adjacent to chromocenters. Additional images of nuclei expressing AtMORC6-Myc were obtained using a Zeiss LSM 510 META confocal microscope. Eight images are displayed, taken at depth intervals of 1 micron. Multiple chromocenters are bordered by AtMORC6 bodies.

Fig. S1

A



B

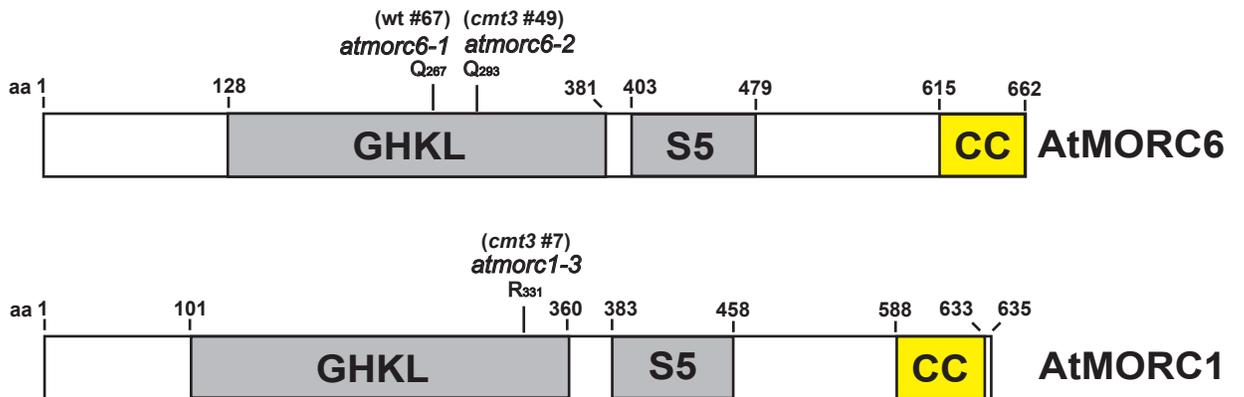
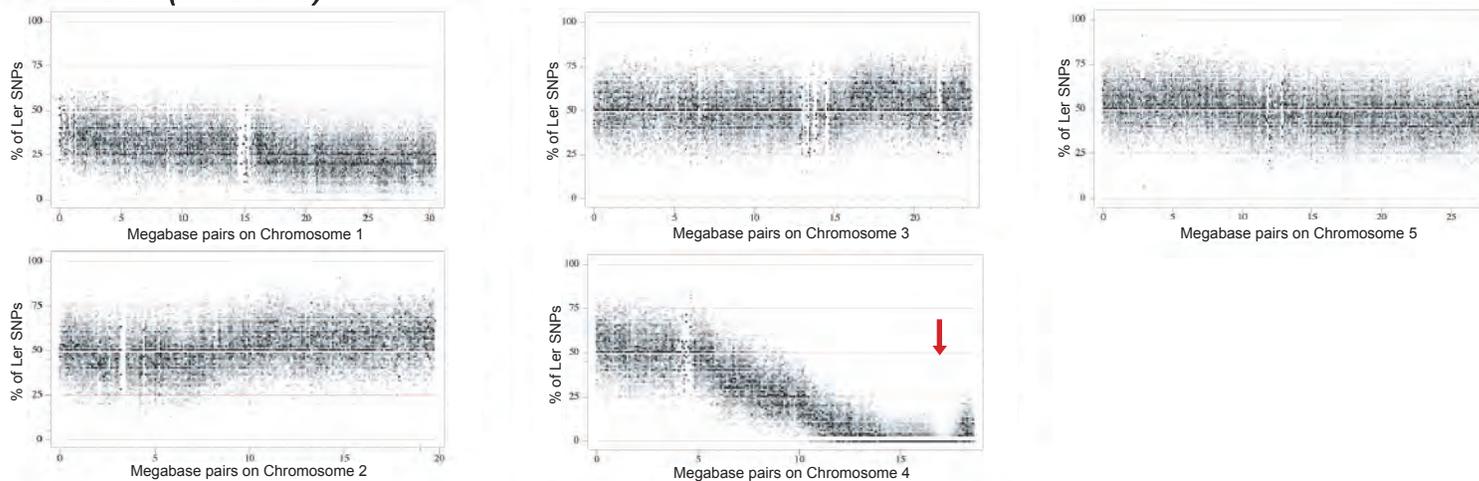


Fig. S2

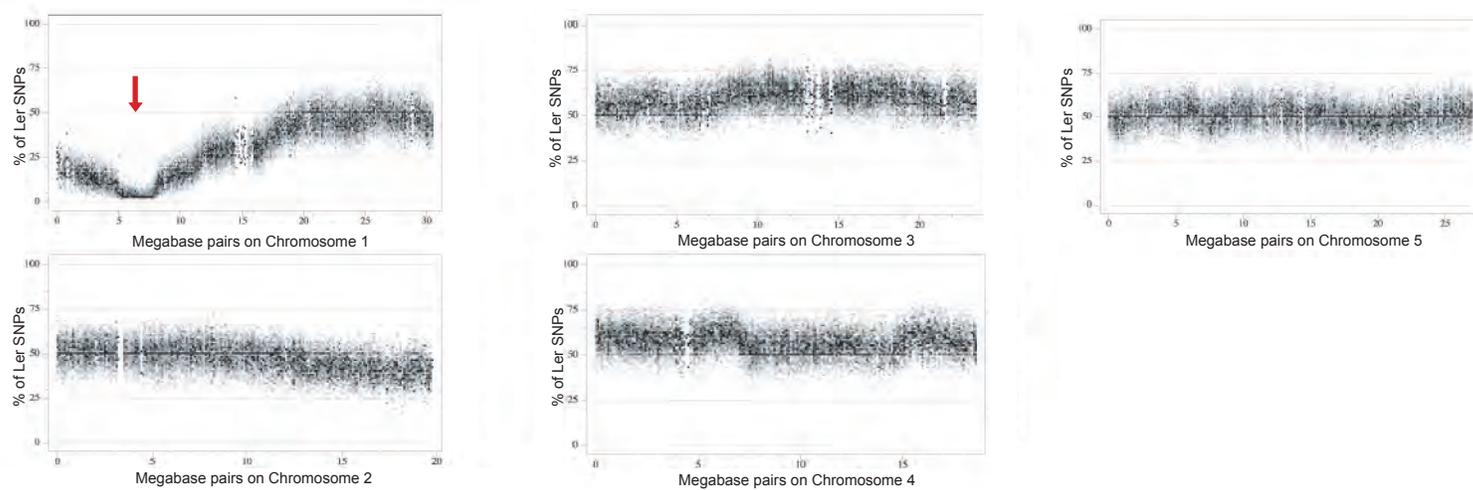
A

cmt3 #7 (atmorc1-3)



B

wt #67 (atmorc6-1)



C

cmt3 #49 (atmorc6-2)

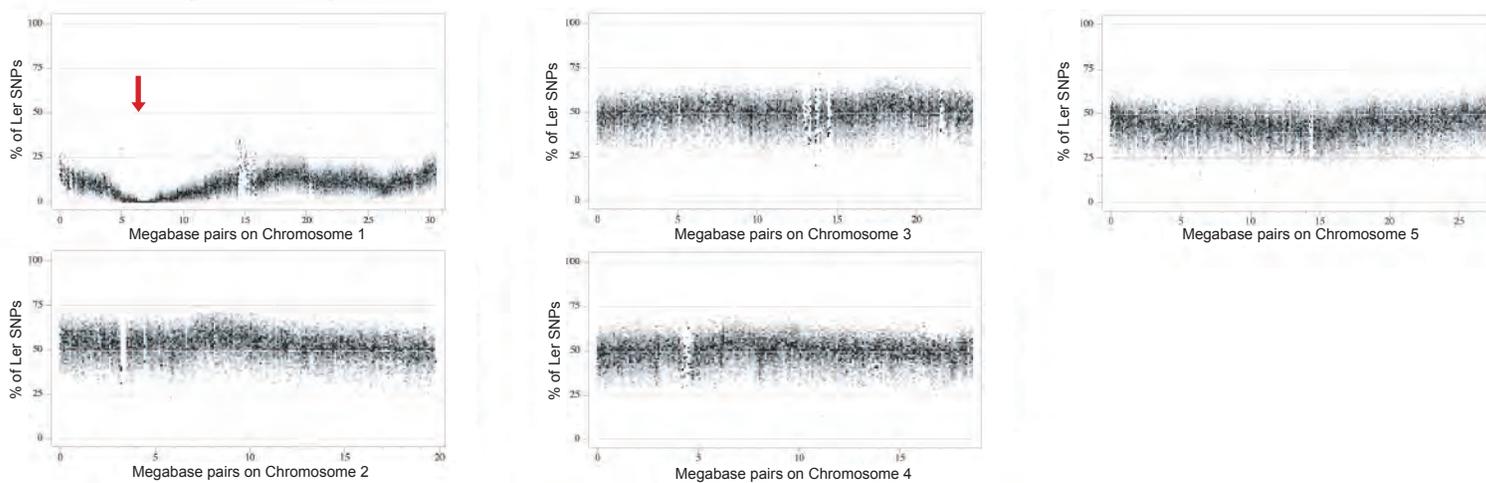
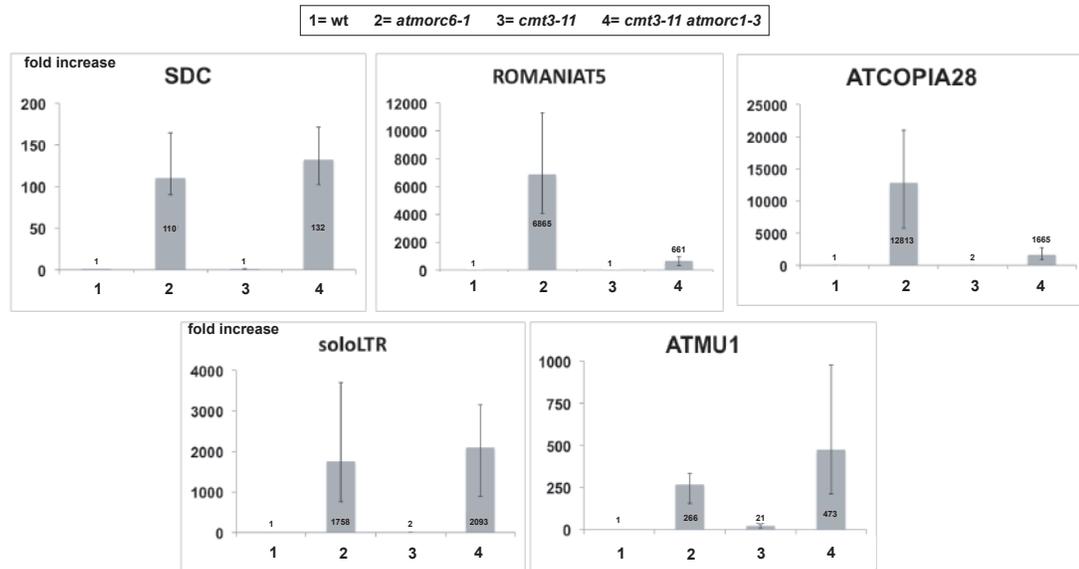


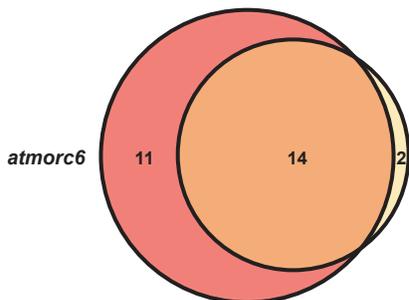
Fig. S4

A



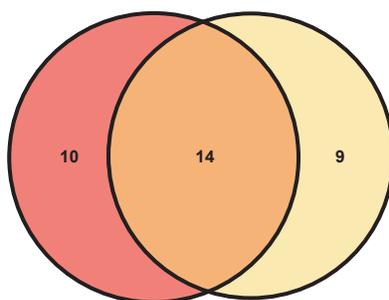
B

Transposable Elements



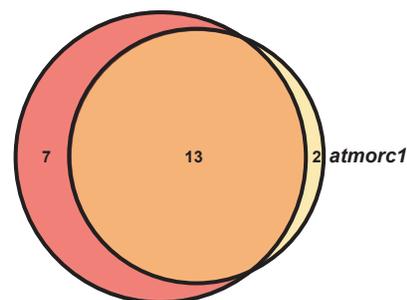
C

Protein coding Genes



D

DNA-Methylated Protein coding Genes



E

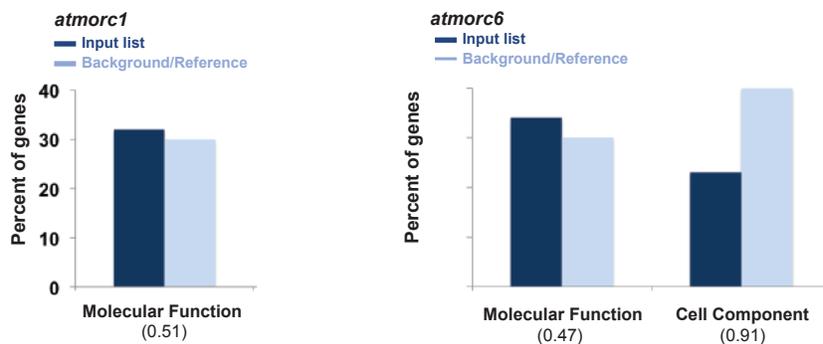
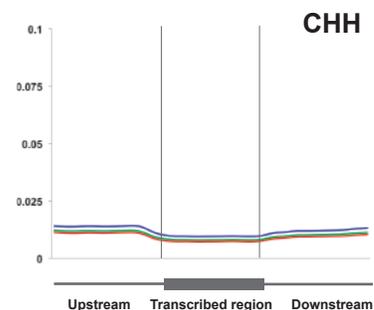
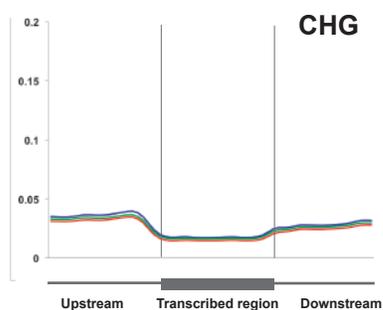
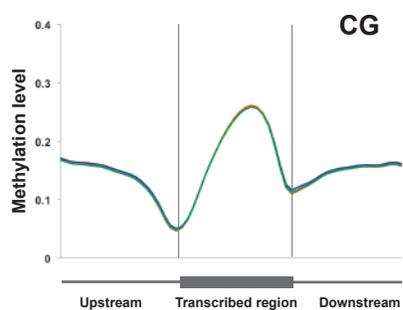
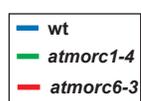


Fig. S5

A



B

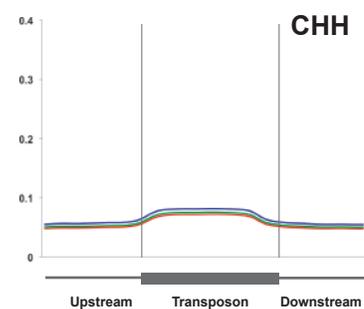
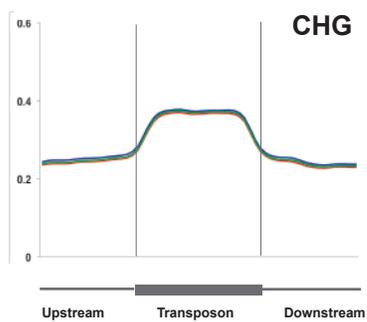
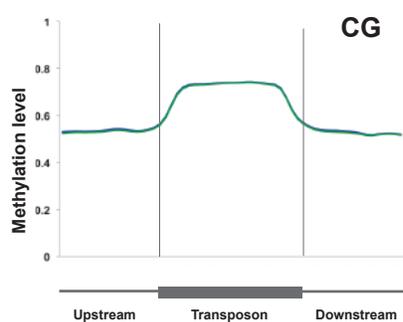


Fig. S7

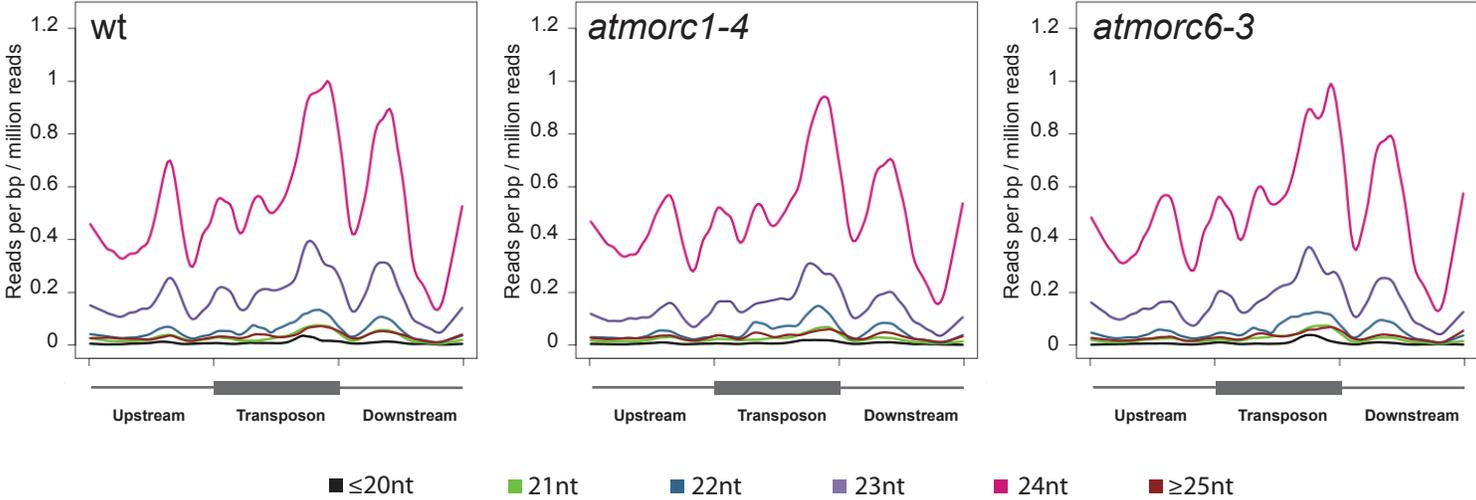


Fig. S8

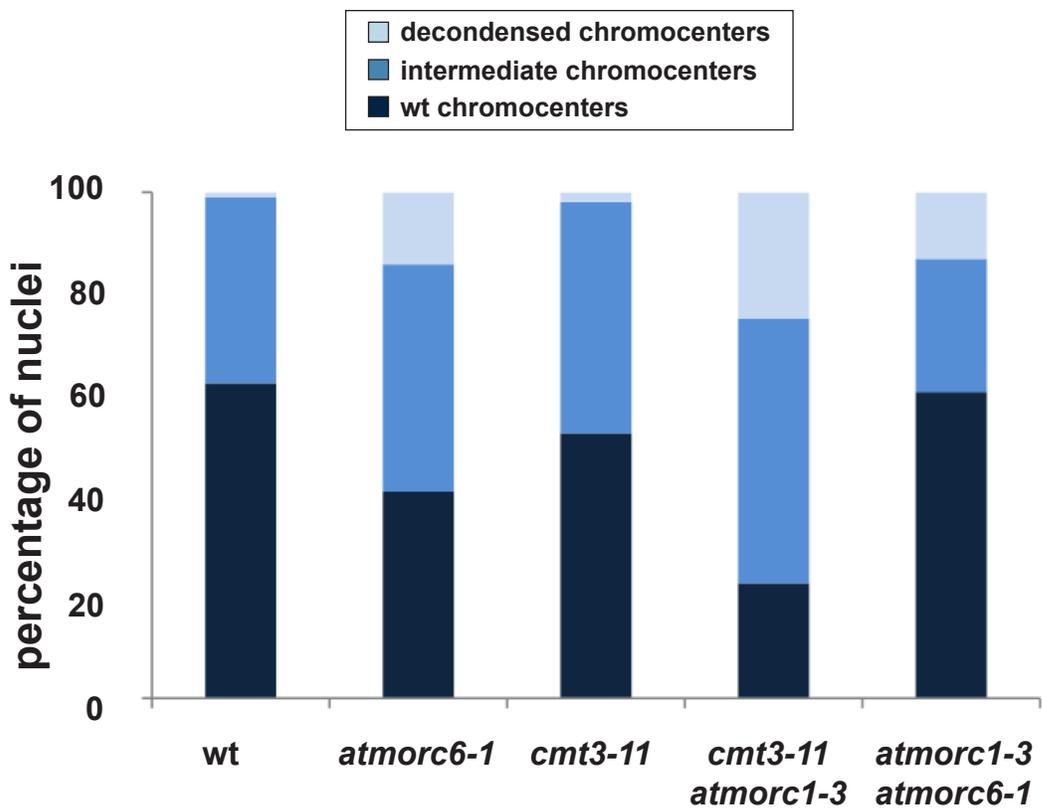
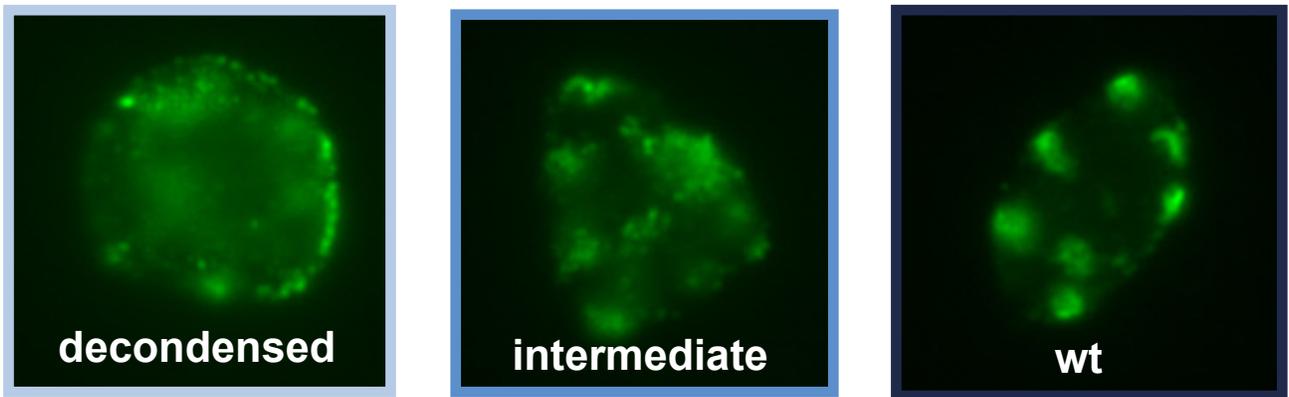


Fig. S9

cmt3-11 atmorc1-3

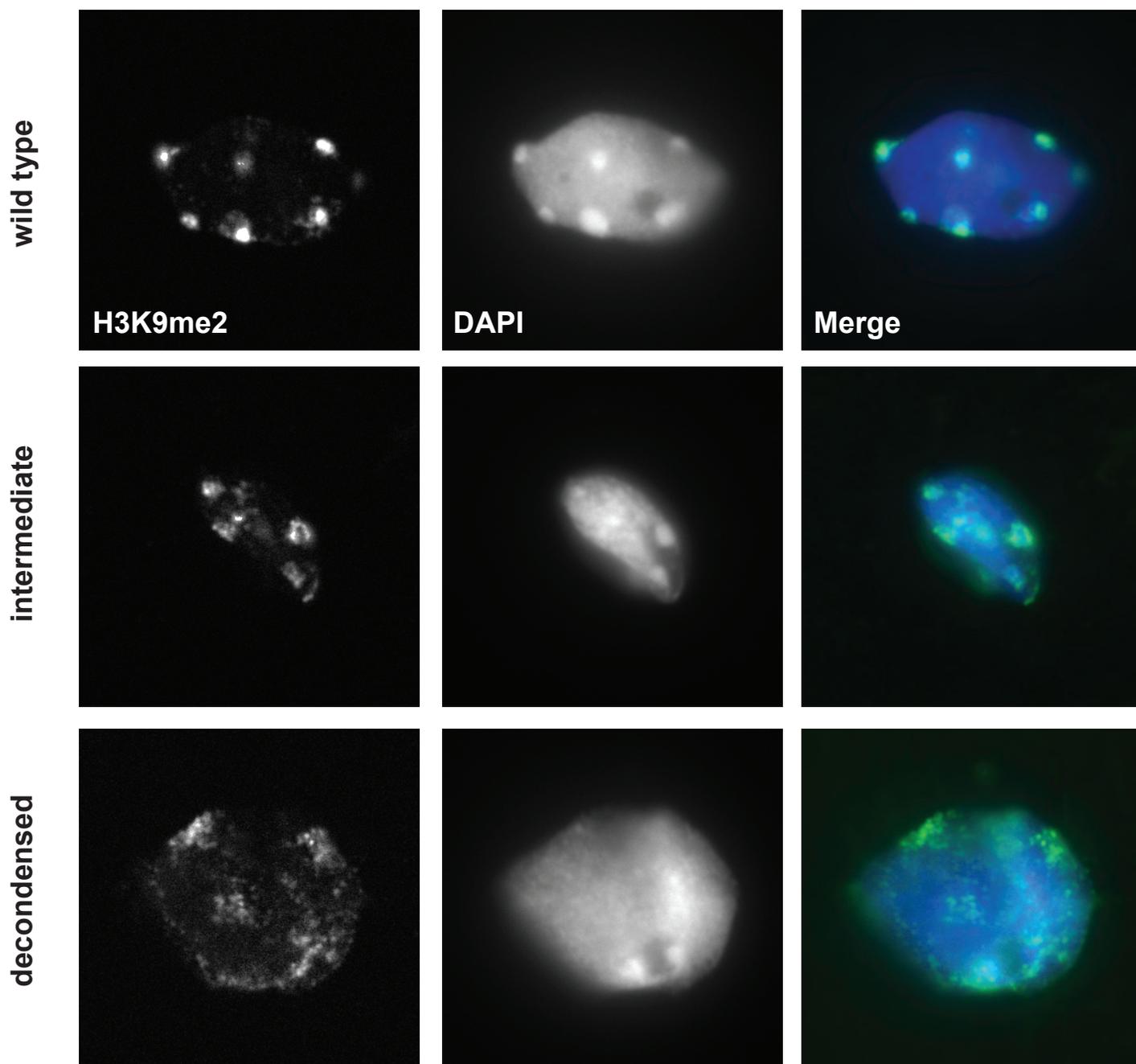


Fig. S10

atmorc6-1

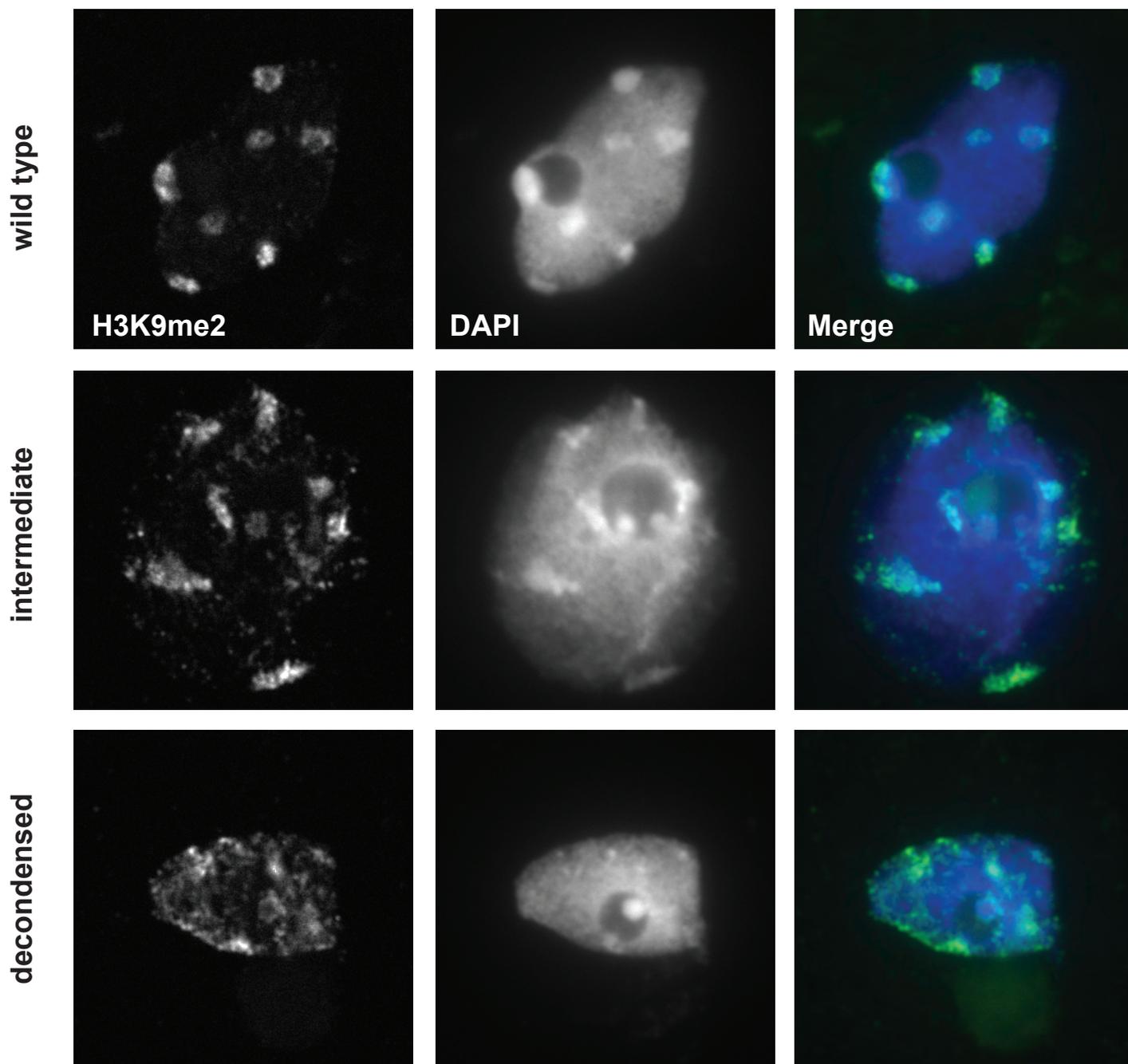


Fig. S11

atmorc1-3 atmorc6-1

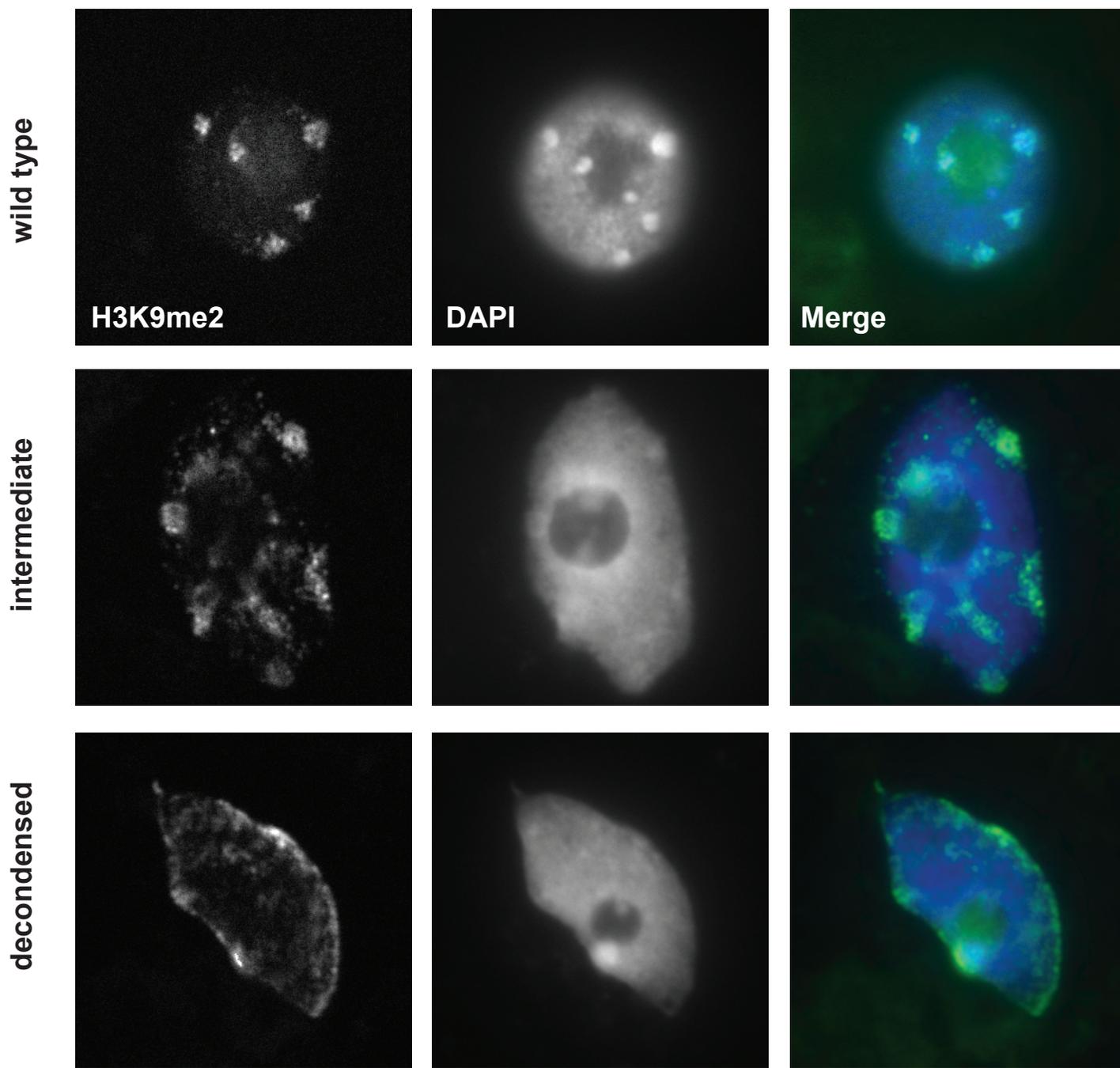


Fig. S12

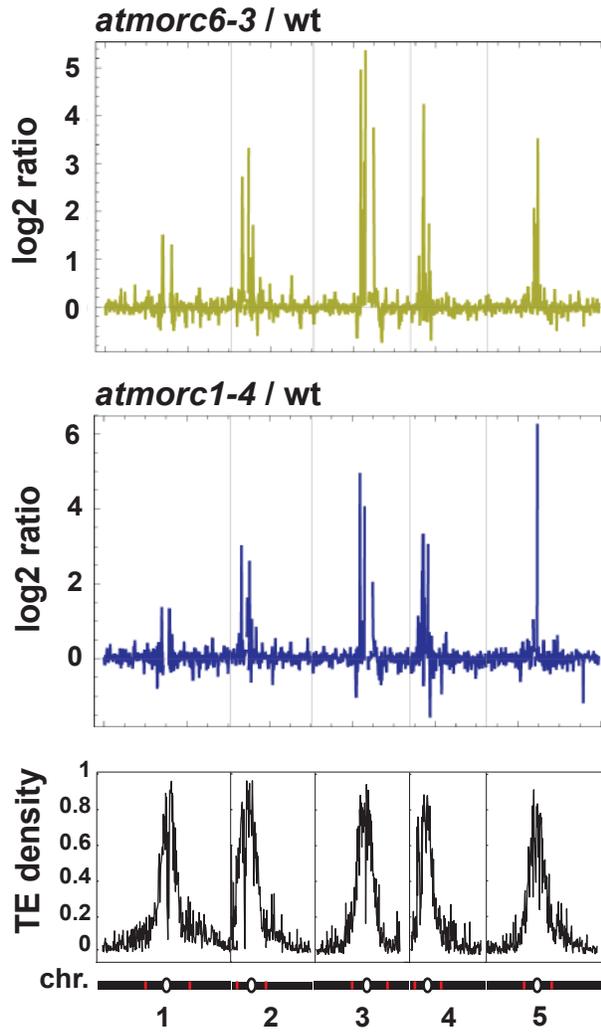


Fig. S13

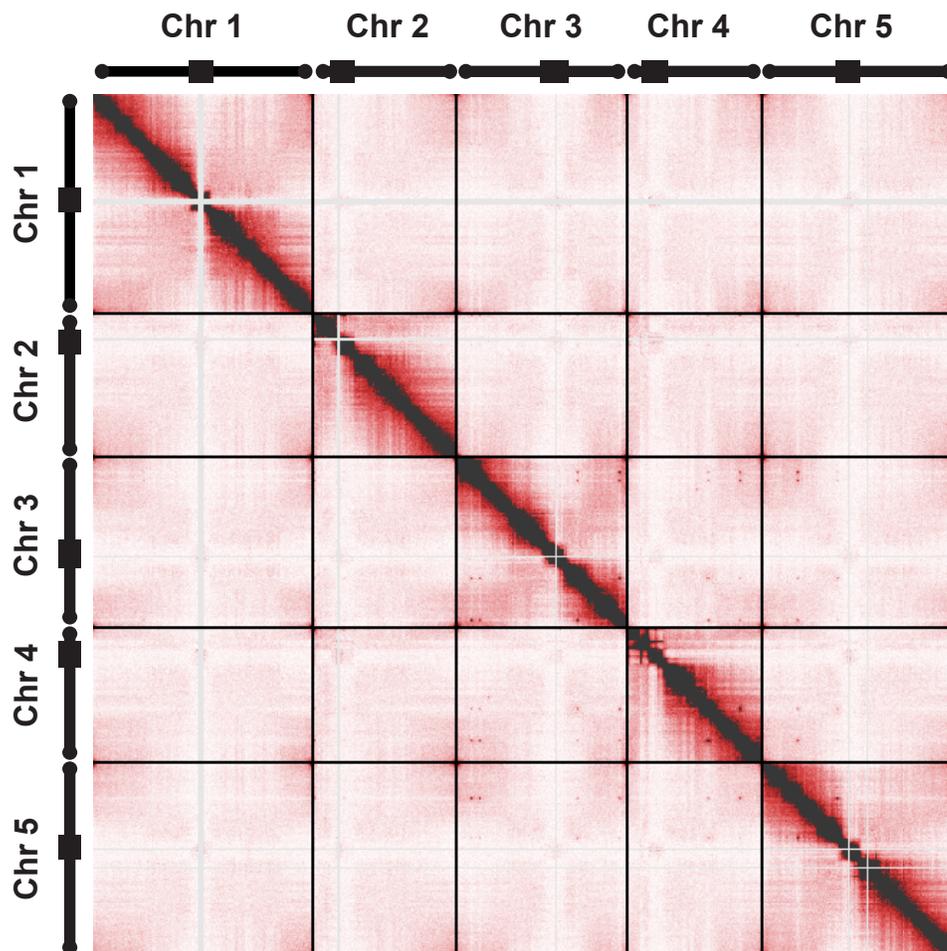


Fig. S14

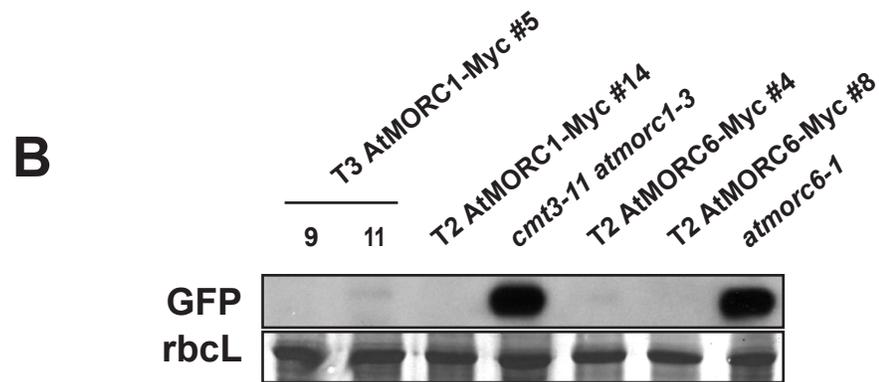
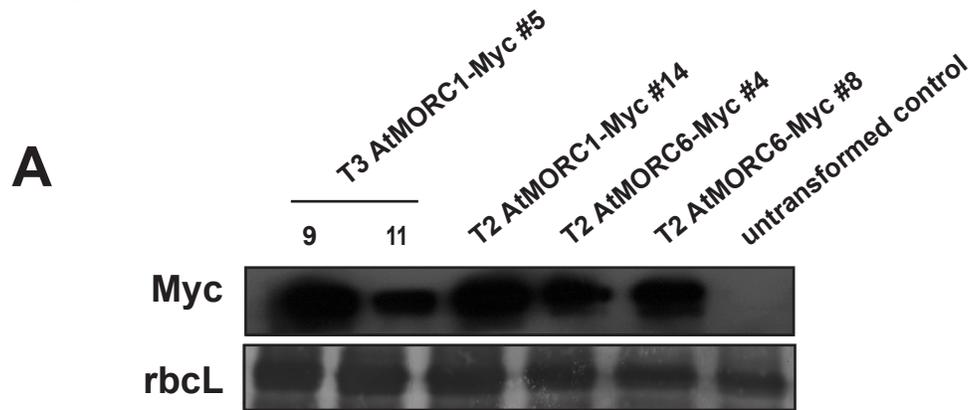


Fig. S15

pAtMORC6::AtMORC6-Myc

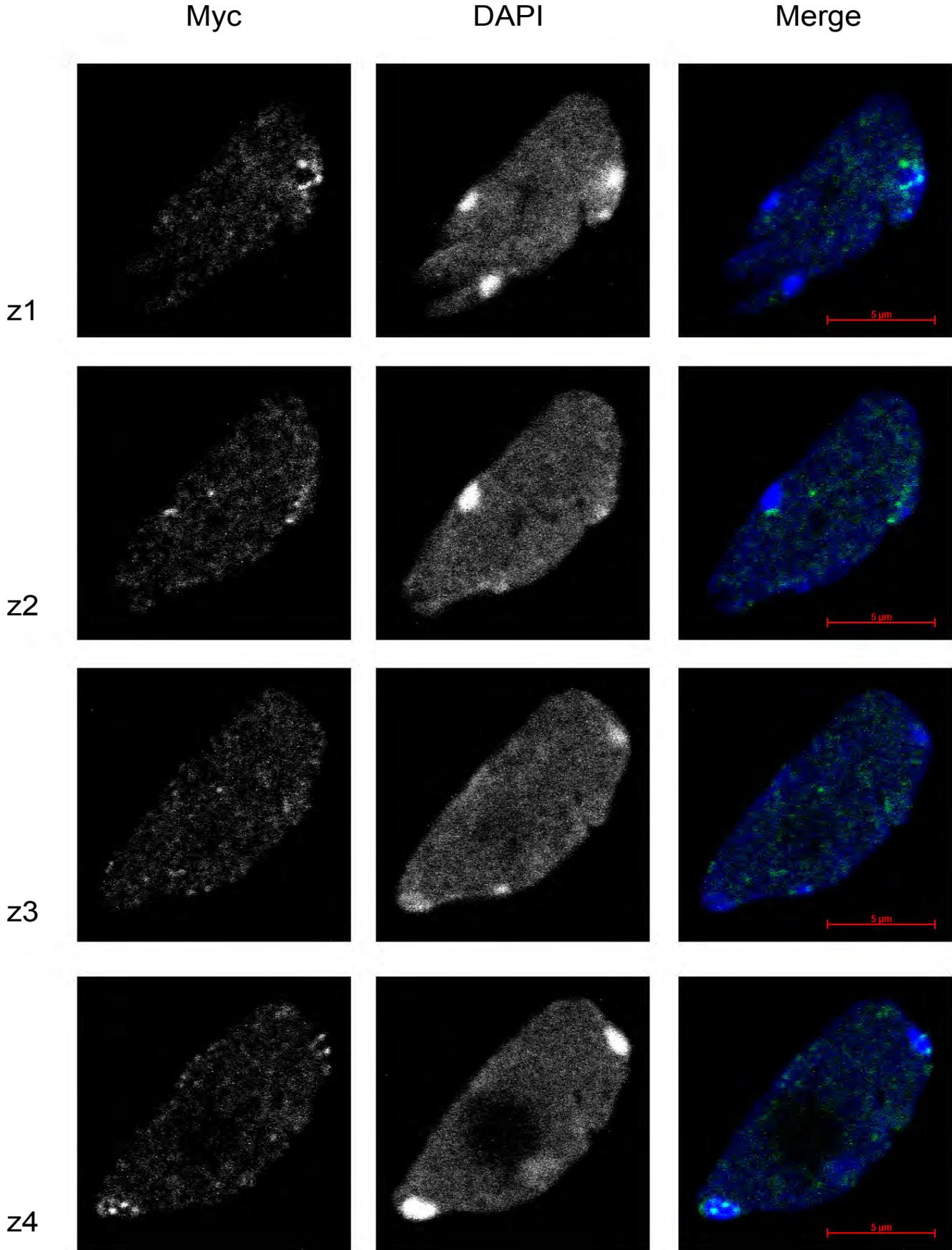


Fig. S15 (continued)

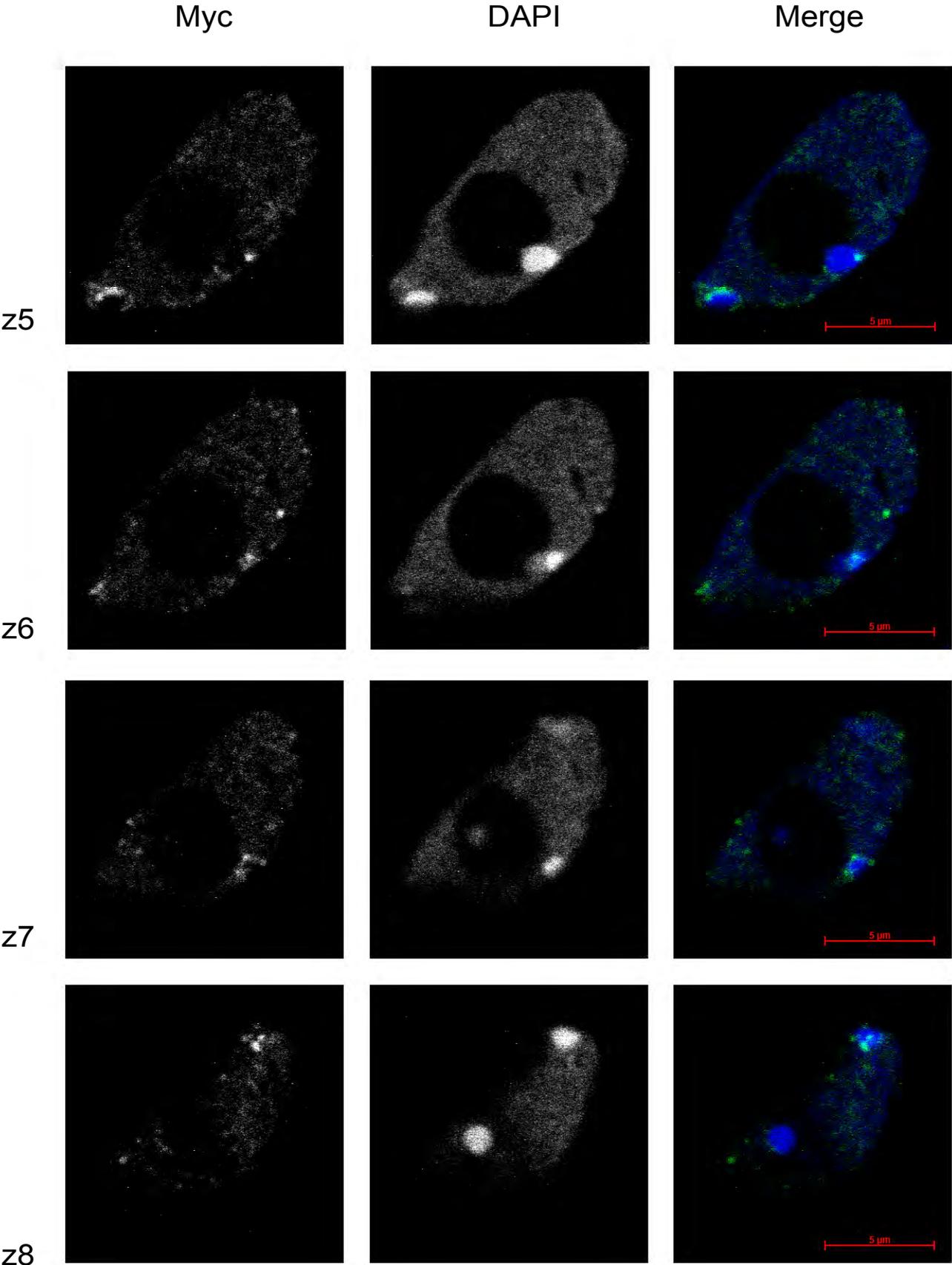


Table S1: Lists of Transposable elements (TEs) 4-fold upregulated in both EMS and T-DNA *atmorc6* and *atmorc1* mutant alleles. The TE soloLTR was identified by RT-qPCR. Therefore, it does not appear in Venn diagram shown in Fig. S4b.

TEs upregulated in <i>atmorc6</i>			
NAME	FAMILY NAME	SUPERFAMILY NAME	CHROMOSOMAL LOCATION
AT1TE43225	ROMANIAT5	LTR/Copia	chr1:13,232,205-13,236,935
AT1TE45510	ATENSPM6	DNA/En-Spm	chr1:13,872,370-13,872,595
AT1TE51360	ROMANIAT5	LTR/Copia	chr1:15,613,068-15,617,844
AT2TE07145	ATCOPIA95	LTR/Copia	chr2:1,547,658-1,552,030
AT2TE13060	ATCOPIA32	LTR/Copia	chr2:2,995,748-2,996,367
AT2TE16220	HELITRON2	RC/Helitron	chr2:3,749,179-3,756,447
AT2TE18240	ATIS112A	DNA/Harbinger	chr2:4,343,480-4,344,832
AT2TE28020	ATMU1	DNA/MuDR	chr2:6,888,343-6,891,692
AT2TE41170	ATIS112A	DNA/Harbinger	chr2:9,660,885-9,661,799
AT3TE51895	ROMANIAT5	LTR/Copia	chr3:12,605,745-12,609,307
AT3TE51900	ATCOPIA28	LTR/Copia	chr3:12,609,702-12,610,774
AT3TE51910	VANDAL18NA	DNA/MuDR	chr3:12,612,304-12,613,201
AT3TE51930	ATGP1	LTR/Gypsy	chr3:12,615,787-12,623,418
AT3TE60425	ATHILA3	LTR/Gypsy	chr3:14,799,085-14,800,505
AT3TE60460	ATHILA6A	LTR/Gypsy	chr3:14,806,306-14,815,940
AT3TE63540	ATHILA2	LTR/Gypsy	chr3:15,722,463-15,733,680
AT4TE04415	ATIS112A	DNA/Harbinger	chr4:860,921-864,078
AT4TE09845	ATIS112A	DNA/Harbinger	chr4:2,076,980-2,077,351
AT4TE15005	ATHILA3	LTR/Gypsy	chr4:3,266,370-3,288,501
AT4TE15025	ATHILA0_I	LTR/Gypsy	chr4:3,273,946-3,285,539
AT4TE15030	ATHILA6A	LTR/Gypsy	chr4:3,274,431-3,285,232
soloLTR	soloLTR	retroelement soloLTR	chr5:9,871,837-9,872,408
AT5TE39630	ATIS112A	DNA/Harbinger	chr5:10,901,855-10,903,021
AT5TE47200	ATHILA2	LTR/Gypsy	chr5:13,334,474-13,345,770
AT5TE48715	ATREP15	RC/Helitron	chr5:13,705,744-13,706,653
AT5TE48720	HELITRONY1E	RC/Helitron	chr5:13,706,654-13,706,934
TEs upregulated in <i>atmorc1</i>			
NAME	FAMILY NAME	SUPERFAMILY NAME	CHROMOSOMAL LOCATION
AT1TE43225	ROMANIAT5	LTR/Copia	chr1:13,232,205-13,236,935
AT1TE45510	ATENSPM6	DNA/En-Spm	chr1:13,872,370-13,872,595
AT1TE51360	ROMANIAT5	LTR/Copia	chr1:15,613,068-15,617,844
AT2TE13060	ATCOPIA32	LTR/Copia	chr2:2,995,748-2,996,367
AT2TE18240	ATIS112A	DNA/Harbinger	chr2:4,343,480-4,344,832
AT2TE28020	ATMU1	DNA/MuDR	chr2:6,888,343-6,891,692
AT2TE28025	ATGP10	LTR/Gypsy	chr2:6,891,693-6,892,236
AT3TE51895	ROMANIAT5	LTR/Copia	chr3:12,605,745-12,609,307
AT3TE51900	ATCOPIA28	LTR/Copia	chr3:12,609,702-12,610,774
AT3TE60460	ATHILA6A	LTR/Gypsy	chr3:14,806,306-14,815,940
AT4TE04415	ATIS112A	DNA/Harbinger	chr4:860,921-864,078
AT4TE09845	ATIS112A	DNA/Harbinger	chr4:2,076,980-2,077,351
soloLTR	soloLTR	retroelement soloLTR	chr5:9,871,837-9,872,408
AT5TE39630	ATIS112A	DNA/Harbinger	chr5:10,901,855-10,903,021
AT5TE48715	ATREP15	RC/Helitron	chr5:13,705,744-13,706,653
AT5TE48720	HELITRONY1E	RC/Helitron	chr5:13,706,654-13,706,934

AT5TE48740	ATLINE1_4	LINE/L1	chr5:13,712,007-13,715,840
------------	-----------	---------	----------------------------

Table S2: Lists of Protein coding genes 4-fold upregulated in both EMS and T-DNA *atmorc6* and *atmorc1* mutant alleles. With the exception of genes in italics, all genes in these lists are located in pericentromeric regions and show DNA methylation or H3K9me2 silencing marks.

Protein coding genes upregulated in <i>atmorc6</i>	
NAME	GENE ANNOTATION
AT1G33840	unknown protein
AT1G35730	APUM9 (ARABIDOPSIS PUMILIO 9); RNA binding
AT1G36675	glycine-rich protein
AT1G53480	unknown protein
AT1G59930	unknown protein
AT1G60110	jacalin lectin family protein
AT1G67105	non coding RNA
AT2G07215	unknown protein
AT2G07240	Ulp1 protease family protein
AT2G10975	unknown protein
AT2G13770	unknown protein
AT2G17690	SDC, F-box family protein
<i>AT3G20340</i>	<i>unknown protein</i>
AT3G29639	unknown protein
AT3G30842	ATPDR10/PDR10 (PLEIOTROPIC DRUG RESISTANCE 10)
AT3G33528	unknown protein
AT3G42850	galactokinase, putative
<i>AT4G12490</i>	<i>protease inhibitor/seed storage/lipid transfer protein (LTP) family protein</i>
AT5G07550	GRP19 (Glycine rich protein 19)
AT5G07560	GRP20 (Glycine rich protein 20); nutrient reservoir
<i>AT5G15500</i>	<i>ankyrin repeat family protein</i>
AT5G35480	unknown protein
AT5G35490	unknown protein
<i>AT5G36910</i>	<i>THI2.2 (THIONIN 2.2); toxin receptor binding (THI2.2)</i>
Protein coding genes upregulated in <i>atmorc1</i>	
NAME	GENE ANNOTATION
<i>AT1G19610</i>	<i>LCR78/PDF1.4 (Low-molecular-weight cysteine-rich 78)</i>
<i>AT1G26390</i>	<i>FAD-binding domain-containing protein</i>
<i>AT1G26410</i>	<i>FAD-binding domain-containing protein</i>
AT1G33840	unknown protein
AT1G35730	APUM9 (ARABIDOPSIS PUMILIO 9)
AT1G36675	glycine-rich protein (AT1G36675)
<i>AT1G47890</i>	<i>disease resistance family protein</i>
AT2G07215	unknown protein
AT2G10975	unknown protein
AT2G13770	unknown protein
AT2G17690	SDC, F-box family protein
<i>AT2G17740</i>	<i>DC1 domain-containing protein</i>
<i>AT2G45220</i>	<i>pectinesterase family protein</i>
AT3G29639	unknown protein
AT3G30842	ATPDR10/PDR10 (PLEIOTROPIC DRUG RESISTANCE 10)
AT3G33528	unknown protein
AT3G55970	oxidoreductase, 2OG-Fe(II) oxygenase family protein

AT4G12490	<i>protease inhibitor/seed storage/lipid transfer protein (LTP) family protein</i>
AT4G18170	<i>WRKY28 (WRKY DNA-binding protein 28); transcription factor</i>
AT5G07550	GRP19 (Glycine rich protein 19)
AT5G35480	unknown protein
AT5G35490	unknown protein
AT5G35510	unknown protein

Table S3: Lists of Transposable elements (TEs) and protein-coding genes (PCGs) that were 4-fold upregulated in the *atmorc1-3 atmorc6-1* double mutant. * defines loci found to be 4-fold upregulated only in the *atmorc1-3 atmorc6-1* double and not in the *atmorc1* or *atmorc6* single mutants. All the other loci have been found to be upregulated in either EMS or T-DNA *atmorc1* and/or *atmorc6* alleles at least in one RNA sequencing dataset.

TEs upregulated in the <i>atmorc1-3 atmorc6-1</i> double mutant			
NAME	FAMILY NAME	SUPERFAMILY NAME	CHROMOSOMAL LOCATION
AT1TE30845	HELITRON1	RC/Helitron	chr1:9,574,349-9,575,358
AT1TE43225	ROMANIAT5	LTR/Copia	chr1:13,232,205-13,236,935
AT1TE45135	ATHILA2	LTR/Gypsy	chr1:13,872,370-13,872,595
AT1TE45510	ATENSPM6	DNA/En-Spm	chr1:13,872,370-13,872,595
AT1TE53070	ATCOPIA87	LTR/Copia	chr1:16,118,258-16,118,383
AT1TE59745	ATCOPIA49	LTR/Copia	chr1:18,005,845-18,011,038
AT2TE07145	ATCOPIA95	LTR/Copia	chr2:1,547,658-1,552,030
AT2TE08840	TA11	LINE/L1	chr2:1,921,517-1,925,279
AT2TE13060	ATCOPIA32	LTR/Copia	chr2:2,995,748-2,996,367
AT2TE15415	ATGP10	LTR/Gypsy	chr2:3,533,343-3,539,331
AT2TE15880	ATHILA6A	LTR/Gypsy	chr2:3,643,453-3,652,409
AT2TE16220	HELITRON2	RC/Helitron	chr2:3,749,179-3,756,447
AT2TE18240	ATIS112A	DNA/Harbinger	chr2:4,343,480-4,344,832
AT2TE19615	VANDAL21	DNA/MuDR	chr2:4,736,486-4,744,627
AT2TE19625	ATHILA4C	LTR/Gypsy	chr2:4,738,315-4,739,542
AT2TE29450	ATCOPIA70	LTR/Copia	chr2:7,231,166-7,236,206
AT2TE38900	ATCOPIA76	LTR/Copia	chr2:9,194,572-9,198,735
AT2TE38905	ATGP2N	LTR/Gypsy	chr2:9,198,736-9,200,805
AT2TE66360	ATCOPIA50	LTR/Copia	chr2:14,938,850-14,939,092
AT3TE50570	VANDAL3	DNA/MuDR	chr3:12,177,936-12,189,847
AT3TE50595	ATHILA2	LTR/Gypsy	chr3:12,189,865-12,193,865
AT3TE51895	ROMANIAT5	LTR/Copia	chr3:12,605,745-12,609,307
AT3TE51900	ATCOPIA28	LTR/Copia	chr3:12,609,702-12,610,774
AT3TE51910	VANDAL18NA	DNA/MuDR	chr3:12,612,304-12,613,201
AT3TE51930	ATGP1	LTR/Gypsy	chr3:12,615,787-12,623,418
AT3TE60425	ATHILA3	LTR/Gypsy	chr3:14,799,08- 14,800,505
AT3TE60460	ATHILA6A	LTR/Gypsy	chr3:14,806,306-14,815,940
AT3TE63540	ATHILA2	LTR/Gypsy	chr3:15,722,463-15,733,680
AT4TE09845	ATIS112A	DNA/Harbinger	chr4:2,076,980-2,077,351
AT4TE10335	ATCOPIA58	LTR/Copia	chr4:2,198,163-2,203,991
AT4TE15005	ATHILA3	LTR/Gypsy	chr4:3,266,370-3,288,501
AT4TE15025	ATHILA0_I	LTR/Gypsy	chr4:3,273,946-3,285,539
AT4TE15030	ATHILA6A	LTR/Gypsy	chr4:3,274,431-3,285,232
AT4TE16900	ATHILA	LTR/Gypsy	chr4:3,838,635-3,843,992
AT4TE17090	ATCOPIA41	LTR/Copia	chr4:3,909,249-3,910,990
AT4TE17115	ATHILA6A	LTR/Gypsy	chr4:3,915,207-3,916,819
AT4TE20120	ATENSPM5	DNA/En-Spm	chr4:4,830,277-4,836,762
AT4TE25590	ATCOPIA49	LTR/Copia	chr4:6,067,394-6,072,632
AT5TE35950	HELITRONY1D	RC/Helitron	chr5:9,871,833-9,872,483
AT5TE39170	ATGP1	LTR/Gypsy	chr5:10,764,247-10,772,507
AT5TE39630	ATIS112A	DNA/Harbinger	chr5:10,901,855-10,903,021

AT5TE43315	ATHILA	LTR/Gypsy	chr5:12,182,325-12,187,588
AT5TE46210	ATHILA8B	LTR/Gypsy	chr5:13,008,325-13,013,667
AT5TE47200	ATHILA2	LTR/Gypsy	chr5:13,334,474-13,345,770
PCGs upregulated in <i>in the atmorc1-3 atmorc6-1 double mutant</i>			
NAME	GENE ANNOTATION		
AT3G20710	F-box protein-related		
AT4G05370	unknown protein		
AT3G13220	ABC transporter family protein		
AT3G15440	unknown protein		
AT3G22860	TIF3C2 (eukaryotic translation initiation factor 3 subunit C2)		
AT5G29560	Ca ²⁺ -binding EF hand family protein		
AT1G45063	copper ion binding / electron carrier		
AT3G44460	DPBF2 (BASIC LEUCINE ZIPPER TRANSCRIPTION FACTOR 67)		
AT3G30842	ATPDR10/PDR10 (PLEIOTROPIC DRUG RESISTANCE 10)		
AT1G27570	phosphatidylinositol 3- and 4-kinase family protein		
AT2G17690	F-box family protein		
AT3G29639	unknown protein		
AT2G04050	MATE efflux family protein		
AT5G60260	unknown protein		
AT4G05380	AAA-type ATPase family protein		
AT1G33840	unknown protein		
AT2G10975	unknown protein		
AT5G18840	sugar transporter, putative		
AT1G36675	glycine-rich protein		
AT1G35730	APUM9 (ARABIDOPSIS PUMILIO 9); RNA binding (APUM9)		
AT3G01345	unknown protein		
AT1G15150	MATE efflux family protein		
AT4G36850	PQ-loop repeat family protein / transmembrane family protein		
AT1G27565	unknown protein		
AT2G07213	other_rna		
AT5G23020	MAM-L (METHYLTHIOALKYMALATE SYNTHASE-LIKE); 2-isopropylmalate synthase		
AT5G45095	unknown protein		
AT4G20820*	FAD-binding domain-containing protein		
AT2G07215	unknown protein		
AT3G14380	integral membrane family protein		
AT3G47340	ASN1 (DARK INDUCIBLE 6)		
AT5G07700*	MYB76 (myb domain protein 76); DNA binding / transcription factor		
AT5G41080	glycerophosphoryl diester phosphodiesterase family protein		
AT3G33528	unknown protein		
AT3G09450	unknown protein		
AT5G38386	F-box family protein		
AT2G18193	AAA-type ATPase family protein		
AT5G36910	THI2.2 (THIONIN 2.2); toxin receptor binding (THI2.2)		
AT3G44990*	XTR8 (xyloglucan:xyloglucosyl transferase 8); hydrolase, acting on glycosyl bonds (XTR8)		
AT3G01600	ANAC044 (Arabidopsis NAC domain containing protein 44); transcription factor		
AT1G07450*	tropinone reductase, putative / tropine dehydrogenase, putative		
AT4G15680	glutaredoxin family protein		

Table S4: Sequences of primers used in this study		
T-DNA genotyping		
JP9339	atmorc6-3 LP	GGAAAGCTGGAAGCTATAATGATG
JP9340	atmorc6-3 RP	GATGACATCTGCCCAAGTCTC
JP7707	GABI-KAT LB O8409	ATATTGACCATCATACTCATTGC
JP8938	atmorc1-4 LP	CGTATCTCAGCCGCTAACTTG
JP8939	atmorc1-4 RP	AAGCAGCTGCAGTGGATTATG
JP2207	LB3 SAIL T-DNA	TAGCATCTGAATTTTATAACCAATCTCGATACAC
JP8509	drm2-2 LP	AGATCGCTTCCAGAGTTAGCC
JP8510	drm2-2 RP	TTGTCGCAAAAAGCAAAAGAG
JP2922	cmt3-11 LP	TAACGGAAGGATGCCAGATT
JP2923	cmt3-11 RP	CAAGAAATGGGCTGTTGACAT
JP2410	LBA1 SALK T-DNA	TGGTTCACGTAGTGGGCCATCG
RT-qPCR		
JP2452	ACTIN 7 LP	TCGTGGTGGTGAGTTTGTAC
JP2453	ACTIN 7 RP	CAGCATCATACAAGCATCC
JP3395	SDC LP	AATGTAAGTTGTAAACATTTGAACGTGACC
JP3396	SDC RP	CAGGCATCCGTAGAACTCATGAGC
JP9642	ROMANIAT5 LP	GTATCCTTTGGCCCGGTATT
JP9643	ROMANIAT5 RP	GCCTCTTCGAAATGCCATAA
JP9055	ATCOPIA28 LP	AGTCCTTTTGGTTGCTGAACA
JP9056	ATCOPIA28 RP	CCGGATGTAGCAACATTCAT
JP9640	soloLTR LP	AACTAACGTCATTACATACATCTTG
JP9641	soloLTR RP	AATTAGGATCTTGTTTGCCAGCTA
JP9057	ATMU1 LP	TAATTTGGCTGACGGAATCAC
JP9058	ATMU1 RP	ATTTGGGGGAAAACAAATGAG
JP9646	atmorc6-3 LP	CATGTGCACCCTATGTTCCCTT
JP9647	atmorc6-3 RP	ATCCCTTGGATTTGTGGTTTTT
JP9680	atmorc1-4 LP	ATCAAGGAGGCCCTAAACTT
JP9681	atmorc1-4 RP	TGTGACAGTGATTTTGCCAGT
BS-sequencing		
JP6349	SDC BSPCR LP	GAAAAAGTTGGAATGGGTTTGGAGAGTTTAA
JP6350	SDC BSPCR RP	CAACAAACCCTAATATATTTTATATTTAAAC
JP9775	ROMANIAT5 BSPCR LP	GTAAGTGGATTAGTTATTTAAAGAGAGTT
JP9776	ROMANIAT5 BSPCR RP	ATAAATAAACATCATCTACATCTTATAA
JP9120	ATCOPIA28 BSPCR LP	TATTTATTTYGTTTCATTTGGATTAGTTTT
JP9121	ATCOPIA28 BSPCR RP	ACRATATCAAATAATTATCATCATCTTAA
JP9377	soloLTR BSPCR LP	GATATAAAGGAATGGTTAGATAATATGYGATT
JP9378	soloLTR BSPCR RP	CRATATAACTCAAATTTATATTACTCTTAA
JP9177	ATMU1 BSPCR LP	TTATGAATTAGTTAGTTATAGTTTGTATT
JP9178	ATMU1 BSPCR RP	ATTCCTCRTCTTCTACAACATCATTTAA
Cloning		
JP5275	PstI SDCpro fwd	GTAAGTGCAGTGATGCTCTAACAATTCTTCCACAAGACC
JP5276	PstI SDCpro rev	GAAAGTGCAGTTCTCTCCCTGTTTTGCTACTATTG
JP5235	HindIII-NLS-eGFP FWD	TGGCAAGCTTATGGCTCCAAAGAAGAAGAAAGGTCATG
JP5236	HindIII 35Sterminator REV	TAGCAAGCTTCTCTCAACACATGAGCGAAACCCTATAAG
JP8644	ATMORC1 Ctertag LP	CACCGTTGATTTGGTTTTGTCTGGTC
JP8645	ATMORC1 Ctertag RP	AACTTGTTCATCTCCTTCTTC
JP8646	ATMORC6 Ctertag LP	CACCAGTATGATGTGAGGTTAGTGAG

JP8647	ATMORC6_Ctertag_RP	CGTATTACATTTCTTCTGTGC
--------	--------------------	-----------------------

References:

1. J. A. Law, S. E. Jacobsen, *Nat Rev Genet* **11**, 204 (2010).
2. J. P. Jackson, A. M. Lindroth, X. Cao, S. E. Jacobsen, *Nature* **416**, 556 (2002).
3. F. Malagnac, L. Bartee, J. Bender, *EMBO J* **21**, 6842 (2002).
4. X. Zhang *et al.*, *Cell* **126**, 1189 (2006).
5. I. R. Henderson, S. E. Jacobsen, *Genes Dev* **22**, 1597 (2008).
6. H. G. Kang, J. C. Kuhl, P. Kachroo, D. F. Klessig, *Cell Host Microbe* **3**, 48 (2008).
7. H. G. Kang *et al.*, *The Plant cell* **22**, 918 (2010).
8. Z. Wang, M. Gerstein, M. Snyder, *Nat Rev Genet* **10**, 57 (2009).
9. S. J. Cokus *et al.*, *Nature* **452**, 215 (2008).
10. M. L. Watson *et al.*, *Proc Natl Acad Sci U S A* **95**, 14361 (1998).
11. N. Inoue *et al.*, *Hum Mol Genet* **8**, 1201 (1999).
12. L. M. Iyer, S. Abhiman, L. Aravind, *Biology direct* **3**, 8 (2008).
13. P. Fransz, J. H. De Jong, M. Lysak, M. R. Castiglione, I. Schubert, *Proc Natl Acad Sci U S A* **99**, 14584 (2002).
14. E. Lieberman-Aiden *et al.*, *Science* **326**, 289 (2009).
15. P. Amedeo, Y. Habu, K. Afsar, O. Mittelsten Scheid, J. Paszkowski, *Nature* **405**, 203 (2000).
16. A. V. Probst, P. F. Fransz, J. Paszkowski, O. Mittelsten Scheid, *Plant J* **33**, 743 (2003).
17. V. J. Simpson, T. E. Johnson, R. F. Hammen, *Nucleic Acids Res* **14**, 6711 (1986).
18. S. Kennedy, D. Wang, G. Ruvkun, *Nature* **427**, 645 (2004).
19. H. Tabara *et al.*, *Cell* **99**, 123 (1999).
20. M. A. Carmell *et al.*, *Dev Cell* **12**, 503 (2007).
21. D. Bourc'his, T. H. Bestor, *Nature* **431**, 96 (2004).
22. A. A. Aravin *et al.*, *Mol Cell* **31**, 785 (2008).
23. M. W. Kankel *et al.*, *Genetics* **163**, 1109 (2003).
24. N. Kleinboelting, G. Huep, A. Kloetgen, P. Viehoveer, B. Weisshaar, *Nucleic Acids Res* **40**, D1211 (2012).
25. S. J. Clough, A. F. Bent, *Plant J* **16**, 735 (1998).
26. J. A. Law, A. A. Vashisht, J. A. Wohlschlegel, S. E. Jacobsen, *PLoS Genet* **7**, e1002195 (2011).
27. M. V. Greenberg *et al.*, *Epigenetics* **6**, 344 (2011).
28. B. Langmead, C. Trapnell, M. Pop, S. L. Salzberg, *Genome Biol* **10**, R25 (2009).
29. A. Mortazavi, B. A. Williams, K. McCue, L. Schaeffer, B. Wold, *Nat Methods* **5**, 621 (2008).
30. Y. Benjamini, Hochberg, Y., *Journal of the Royal Statistical Society Series B* **57**: **12.**, (1995).
31. P. Y. Chen, S. J. Cokus, M. Pellegrini, *BMC Bioinformatics* **11**, 203 (2010).
32. J. Lee *et al.*, *Plant Cell* **19**, 731 (2007).
33. C. Lu, B. C. Meyers, P. J. Green, *Methods* **43**, 110 (2007).
34. Y. Zhang *et al.*, *Cell* **148**, 908 (2012).
35. W. J. Soppe *et al.*, *EMBO J* **21**, 6549 (2002).
36. J. K. Kim *et al.*, *Science* **308**, 1164 (2005).
37. R. S. Kamath *et al.*, *Nature* **421**, 231 (2003).
38. Y. Y. Polosina, C. G. Cupples, *Mutation research* **705**, 228 (2010).

39. Y. Mimura, K. Takahashi, K. Kawata, T. Akazawa, N. Inoue, *J Cell Sci* **123**, 2014 (2010).
40. T. Hirano, *Genes Dev* **16**, 399 (2002).
41. M. E. Blewitt *et al.*, *Nat Genet* **40**, 663 (2008).
42. I. Ausin, T. C. Mockler, J. Chory, S. E. Jacobsen, *Nat Struct Mol Biol* **16**, 1325 (2009).
43. T. Kanno *et al.*, *Nat Genet* **40**, 670 (2008).
44. G. Bohmdorfer *et al.*, *Plant J* **67**, 420 (2011).