Title
MORC Family ATPases Required for Heterochromatin Condensation and Gene Silencing

Author
Cary, Joshua Reynolds

Publication Date
2012

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

MORC Family ATPases Required for Heterochromatin Condensation and Gene Silencing

A thesis submitted in partial satisfaction
of the requirements for the degree Master of Arts
in Molecular, Cell, and Developmental Biology

by

Joshua Reynolds Cary

2012

ABSTRACT OF THE THESIS

MORC Family ATPases Required for
Heterochromatin Condensation and Gene Silencing

by

Joshua Reynolds Cary

Master of Arts in Molecular, Cell, and Developmental Biology
University of California, Los Angeles, 2012
Professor Steven E. Jacobsen, Chair

Transposons and genes are silenced in eukaryotes through DNA methylation. In this work, mutations in the *Arabidopsis* genes *CRT1* and *CRH6*, two ATPases in the conserved *Microrchidia (MORC)* family, are found to release silencing of DNA methylated transposons and genes without an accompanying loss of DNA methylation. The loci upregulated in *crt1* and *crh6* mutants are located primarily in the regions of pericentromeric heterochromatin. CRT1 and CRH6 are located in small nuclear bodies adjacent to the regions of condensed pericentromeric heterochromatin called chromocenters. In the absence of MORC function, chromocenters decondense and interactions between pericentromeric and euchromatic regions increase. The single *Caenorhabditis elegans* MORC homolog is also shown to be required for gene silencing. These studies demonstrate that MORC ATPases are required for proper heterochromatin condensation and have a conserved role in gene silencing in eukaryotes.

The thesis of Joshua Reynolds Cary is approved.

Jau-Nian Chen

Amander T. Clark

Steven E. Jacobsen, Committee Chair

University of California, Los Angeles

2012

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

This thesis is based on work previously published in an article of the same title written by Moissiard, G., Cokus, S. J., Cary, J., Feng, S., Billi, A. C., Stroud, H., Husmann, D., Zhan, Y., Lajoie, B. R., McCord, R. P., Hale, C. J., Feng, W., Michaels, S. D., Frand, A. R., Pellegrini, M., Dekker, J., Kim, J. K. and Jacobsen, S. E, in *Science* 336: 1448-1451. The methods section of this thesis is taken from the submitted version of the manuscript. Figures 5E, 7B, 7C, 10 and 11 present my data and analyses that are not previously published. The remaining figures are from the submitted version of the manuscript. Figures 8, 9, 13C, and 13D present results from experiments I performed and published in the paper. Figures 3 and 4 present data which I assisted in analyzing. I thank Wei Fang and Scott Michaels for their expertise in microscopy and for the images presented in Figure 13C; Ye Zhan, Bryan Lajoie, Rachel Patton McCord, and Job Dekker for their performance of Hi-C experiments and data presented in Figure 12; Allison Billi, Alison Frand, and John Kim for their contributions to the story of MORC function in *C. elegans* and the data presented in Figure 14; and Shawn Cokus and Matteo Pellegrini for their many contributions to the analysis of genomic data in this project. I am grateful to members of the Jacobsen lab, including Suhua Feng, Hume Stroud, Dylan Husmann, Christopher Hale, Julie Law, Elena Caro Bernat, and especially Guillaume Moissiard, for their training and assistance in many important experiments and collaboration on this project. Many thanks to Dr. Steve Jacobsen for his mentorship and support during my graduate work at UCLA. I thank the UCLA Cellular and Molecular Biology Training Program and the Ruth L. Kirschstein National Research Service Award GM007185 for financial support of my graduate education.

INTRODUCTION

Transcriptional gene silencing of transposable elements (TEs), DNA repeats, and genes is correlated with methylation of DNA at cytosines in eukaryotes (Goll and Bestor 2005). In *Arabidopsis*, methylation occurs in three cytosine contexts: CG, CHG, and CHH, where H is an A, T, or C. DNA METHYLTRANSFERASE1 (MET1) maintains methylation at CG sites, CHROMOMETHYLASE3 (CMT3) methylates CHG sites, and DOMAINS REARRANGED METHYLTRANSFERASE2 (DRM2) is largely responsible for persistent *de novo* methylation of asymmetric CHH sites (Law and Jacobsen 2010). Repressed DNA elements are also correlated with histone H3 lysine 9 dimethylation (H3K9me2) (Jackson et al. 2002; Malagnac et al. 2002). These epigenetic signals are essential for proper gene expression patterns, as loss of DNA methylation, such as that observed in *met1* mutants, has drastic impacts on gene overexpression and transposon reactivation (Zhang et al. 2006).

These effects are not isolated to plants but are found in diverse organisms exhibiting DNA methylation, including mammals, in which methylation occurs primarily at CG sites (Law and Jacobsen 2010). DNA methyltransferase 1 (Dnmt1), the mammalian homolog of *Arabidopsis* MET1, preferentially catalyzes the methylation of hemimethylated sites in order to maintain genomic methylation patterns (Goll and Bestor 2005). Dnmt1 function is essential in mammalian development, as mice homozygous for null mutations in *Dnmt1* lose the vast majority of DNA methylation in the genome and fail to develop to the 8-somite stage of embryonic development (Lei et al. 1996). Dnmt3a and Dnmt3b, homologs of *Arabidopsis* DRM2, operate mainly in embryogenesis and catalyze *de novo* methylation to establish patterns of DNA methylation (Law and Jacobsen 2010). Mice harboring mutations in both *Dnmt3a* and *Dnmt3b* are unable to methylate endogenous and newly introduced retroviral DNA and show embryonic developmental defects similar to those observed in *Dnmt1* null embryos (Okano et al. 1999). DNA methyltransferase 3-like (Dnmt3L), a protein which is related to the other Dnmt3

methyltransferases but lacks the catalytic domains, is required for fertility but not viability in mice (Goll and Bestor 2005). For example, loss of Dnmt3L prevents *de novo* methylation of retrotransposons, and the resulting transposon reactivation causes pronounced defects in meiosis of spermatocytes and leads to sterility (Bourc'his and Bestor 2004). Thus, epigenetic silencing achieved by DNA methylation is a requirement for proper development.

To search for additional genes involved in DNA methylation and silencing pathways in *Arabidopsis*, an ethyl methanesulfonate (EMS) mutagenesis screen was carried out to identify mutants displaying reactivation of a methylated and silenced transgene (Moissiard et al. 2012). The promoter of the gene *suppressor of drm2 cmt3 (SDC)* carries seven tandem repeats which are methylated and silenced (Henderson and Jacobsen 2008). Silencing of this gene is achieved by the redundant activities of DRM2 and CMT3, as *SDC* silencing is lost in *drm2 cmt3* double mutants but is maintained in a *drm2* or *cmt3* single mutant (Henderson and Jacobsen 2008). A sensor construct with the *SDC* promoter driving expression of GFP was used in EMS screens in wild-type and *cmt3* mutant backgrounds to identify mutants showing upregulation of the *SDC::GFP* transgene (Moissiard et al. 2012). These screens identified two homologous genes, *At4g36290* and *At1g19100*, which, when mutated, each showed a loss of *SDC* transgene silencing.

Mutations in *At4g36290*, previously named *COMPROMISED RECOGNITION OF TCV-1* (*CRT1*), have been shown to reduce resistance to the turnip crinkle virus (Kang et al. 2008; Kang et al. 2010). Mutations in *At1g19100,* referred to as *CRT1 HOMOLOG 6* (*CRH6*), had not been previously reported. The *Arabidopsis* genome includes seven homologs in the *At4g36290* family (Kang et al. 2010). The EMS mutagenesis screen results suggested that, in addition to any roles in viral resistance, genes in this family also play important functions in epigenetic silencing. *CRT1* and *CRH6* are homologs of mouse *Microrchidia 1 (MORC1),* an ATPase whose function is essential for completion of meiosis in spermatogenesis (Watson et al. 1998; Inoue et al. 1999). MORC family proteins with this ATPase motif belong to the GHKL superfamily, which includes

gyrase, Hsp90, histidine kinase, and MutL, and shares an unconventional Bergerat ATP-binding fold (Dutta and Inouye 2000). These enzymes carry out diverse functions in the cell which include alterations in DNA topology. Phylogenetic analyses have found that MORC proteins are present in a wide variety of eukaryotes, suggesting an early evolutionary origin (Iyer et al. 2008). Interestingly, eukaryotic MORC proteins have been hypothesized, based on their fusions to various DNA- and peptide-binding domains and other contextual information, to remodel chromatin by catalyzing DNA manipulations in response to epigenetic signals (Iyer et al. 2008). This context suggested a mechanism by which *CRT1* and *CRH6* might regulate expression of *SDC* and other methylated DNA sequences.

To characterize the role of the two *Arabidopsis MORCs* in gene silencing, first, independent mutations in the two genes were obtained to verify causation of *SDC* upregulation. *SDC::GFP* reactivation in the knock-out transferred DNA (T-DNA) insertion lines *crt1-4* and *crh6-3* occurred as it did in the EMS mutant alleles and confirmed that these two genes were responsible for the *SDC* transgene overexpression.

Next, genome-wide transcription analyses with RNA sequencing (RNA-seq) were used to identify the TEs and protein-coding genes with altered expression in the *crt1* and *crh6* mutants. Specific loci were also analyzed by real-time quantitative polymerase chain reaction (RT-qPCR) to assess changes in expression compared to wild-type plants.

Upon finding that the transcription defects in *crt1* and *crh6* extended beyond the *SDC* transgene to many transposons and genes, whole-genome bisulfite sequencing (BS-seq) was used to determine the effects of mutation in these genes on DNA methylation levels. Additionally, specific loci were probed for DNA methylation levels by traditional bisulfite sequencing. Patterns of the H3K9me2 mark were also assessed using chromatin immunoprecipitation sequencing (ChIP-seq). These approaches were used to investigate changes in epigenetic silencing marks to correlate them with the changes in gene expression. Surprisingly, DNA methylation levels and H3K9me2 modification did not appear to be altered in

crt1 and *crh6* relative to wild-type plants, suggesting that these proteins are involved in silencing downstream of DNA methylation or through a separate, novel pathway.

Due to the indications that MORC proteins might act as chromatin remodeling enzymes, chromatin organization in the mutants was analyzed by immunofluorescence. Heterochromatin in *Arabidopsis* is mainly located in the pericentromeric regions of the five chromosomes, and these regions condense in wild-type cells into compact chromocenters at the nuclear periphery (Fransz et al. 2002). These chromocenters consist mostly of repetitive DNA sequences that are marked by DNA methylation as well as H3K9me2 modification (Fransz et al. 2002; Probst et al. 2003), while euchromatin is marked by histone acetylation modifications, including histone H4 lysine 12 acetylation (H4K12Ac) (Soppe et al. 2002). Patterns of chromocenter organization in wild-type *Arabidopsis* nuclei were analyzed by DAPI staining, H3K9me2 immunostaining, and H4K12Ac immunostaining and compared to the patterns observed in *crt1* and *crh6*. Chromocenters were observed to decondense in the *MORC* mutants. This decondensation effect was similar to that seen in nuclei from plants with a mutation in the chromatin remodeling factor *DDM1* (Probst et al. 2003). For further analysis of DNA organization, Hi-C analyses (Lieberman-Aiden et al. 2009) were performed for wild-type and *crh6* to investigate chromatin interactions across the genome and confirm increased interaction between pericentromeric regions and euchromatic arms.

To correlate heterochromatin organization with the action of MORC proteins at chromocenters, nuclear localization of the MORC proteins was assessed with CRT1 and CRH6 myc-tagged transgenes. Immunofluorescence of the tagged proteins was used to identify their location relative to chromocenters.

In order to more broadly investigate the role of MORC family proteins in gene regulation, silencing was examined in the worm *Caenorhabditis elegans* (Moissiard et al 2012). This species lacks DNA methylation (Simpson et al. 1986), thus permitting the investigation of MORC function independent of DNA methylation pathways. GFP transgene silencing was

studied in the *eri-1* sensitized background (Kennedy et al. 2004) and RNA interference (RNAi) was used to knockdown *C. elegans morc-1*. Involvement of MORCs in gene silencing in plants and animals provides confirmation of a general and fundamental role for MORC family ATPases in eukaryotic gene silencing.

RESULTS

EMS mutagenesis screens using the *SDC::GFP* sensor (Fig. 1A) in wild-type and *cmt3* mutant backgrounds revealed the wild-type #67, *cmt3* #7 and *cmt3* #49 mutants (Moissiard et al. 2012). These three mutations showed overexpression of GFP as seen in the *drm2 cmt3* double mutant (Fig. 1B, C). Mapping experiments by bulk segregant analysis and deep genome resequencing revealed the EMS mutant allele in *cmt3* #7 to be a mutation in *At4g36290* (*CRT1*) and the mutant alleles in wild-type #67 and *cmt3* #49 to be mutations in *At1g19100* (*CRH6*), a homolog of *CRT1* (Moissiard et al. 2012). These three mutations each represented nonsense mutations located within the GHKL domain (Fig. 2A). To confirm that *crt1* and *crh6* mutations were responsible for the *SDC::GFP* reactivation, independent knockout T-DNA insertion mutations in these genes were obtained (Fig 2B, C). These mutants, *crt1-4* and *crh6-3*, also showed dramatic increases in *SDC* expression (Fig. 2D). With this confirmation of the EMS mutations in *CRT1* and *CRH6*, the mutants *cmt3* #7, wild-type #67, and *cmt3* #49 were renamed *crt1-3*, *crh6-1*, and *crh6-2*, respectively.

To characterize the role of *CRT1* and *CRH6* in silencing other methylated loci like *SDC*, RNA-seq experiments (Wang et al. 2009) were used to determine the changes in genome-wide transcription patterns in the mutants. Comparisons of overexpressed loci in the mutants relative to wild-type, using a threshold of four-fold of greater increase in reads in RNA-seq experiments with both the T-DNA and EMS mutants, revealed many TEs to be upregulated in *crt1* and in *crh6* (Fig. 3A). Using RT-qPCR to validate the genomic data, specific transposons were demonstrated to undergo massive reactivation in both the T-DNA and EMS mutants (Fig. 3B, C). Some protein-coding genes, mostly ones with DNA methylation, were also upregulated, and the TEs and genes overexpressed in *crt1* overlapped to a large extent with those overexpressed in *crh6* (Fig. 4).

Given the disruption of silencing at many loci, *crt1* and *crh6* were examined for alterations in epigenetic silencing marks. Bisulfite sequencing of the *SDC* promoter in the EMS mutagenized lines indicated no change in DNA methylation levels (Fig. 5E). Although CHG methylation levels decreased significantly in the *cmt3* mutant, as expected, the *crh6-1* and *crt1-3* mutants did not show any significant alterations in DNA methylation levels relative to controls. This result prompted the examination of DNA methylation levels across the genome. Results from BS-seq experiments with the T-DNA insertion mutants revealed no significant changes in DNA methylation in transposons and flanking regions for the set of TEs upregulated in *crt1-4* or in *crh6-3* (Fig. 5A, B). Additionally, DNA methylation levels for all TEs as well as for all protein-coding genes were similar in *crt1-4* and *crh6-3* compared to wild-type (Fig. 5C, D). Traditional bisulfite sequencing of the *SDC* promoter as well four highly upregulated TEs confirmed the lack of significant differences in DNA methylation levels in the mutants (Fig. 5E). Investigation into the patterns of H3K9me2 modification, another mark associated with silencing, using ChIP-seq experiments showed similar levels of the mark at the *SDC* promoter in *crt1-4* and *crh6-3* compared to wild-type (Fig. 6A). Comparison of H3K9me2 in mutant and wild-type at all upregulated TEs and genes and flanking regions similarly yielded no differences (Fig. 6B). Together, these results indicated loss of silencing of TEs and genes in the *crt1* and *crh6* mutants despite the lack of any noticeable effects on epigenetic silencing marks.

Since upregulated loci did not display changes in epigenetic modifications, other common features among the upregulated loci were examined, including chromosomal position. Plots showing the location of upregulated loci along the five *Arabidopsis* chromosomes were created and indicated that most of the overexpressed loci were located near the centromeres (Fig. 7A). Additional plots showed that the majority of upregulated TEs and genes, especially the DNA methylated genes, were located within the boundaries of the pericentromeric heterochromatin on each chromosome (Fig. 7B, C). These analyses suggested a potential mode of action specifically in this area of the genome.

With the data indicating expression defects primarily in the regions of pericentromeric heterochromatin, as well as indications that MORC proteins might function as chromatin remodeling enzymes, additional experiments were conducted to examine heterochromatin organization in the *crt1* and *crh6* mutants. Chromocenters, consisting of condensed pericentromeric heterochromatin marked by H3K9me2, were examined in immunofluorescence experiments. H3K9me2 immunostaining revealed bright spots at the chromocenters, and DAPI staining showed similar patterns of compacted DNA at the chromocenters (Fig. 8A, B, C). Nuclei from *crt1-3* and *crh6-1* mutant plants indicated a degree of chromocenter decondensation, with some showing the wild-type pattern of five to eight bright and compact chromocenter spots, some showing an intermediate level of expansion in the size of chromocenters, and others showing a more severe decondensation and obscuring of the chromocenters (Fig. 8A, B). Nuclei with similar decondensation phenotypes were observed in the *crh6-1 crt1-3* double mutant (Fig. 8C). Quantification of the distribution of chromocenter patterns revealed an increase in the frequency of decondensed chromocenters in the mutants (Fig 9). While intermediate levels of chromocenter decondensation were observed frequently among wild-type and *cmt3* nuclei, those with severely decondensed chromocenters accounted for less than two percent of the total. In contrast, nuclei from *crh6-1, crt1-3*, and *crh6-1 crt1-3* plants showed severe decondensation, in which the individual chromocenters could not be counted, with a frequency of fifteen percent or more. This analysis also demonstrated that the *crh6-1 crt1-3* double mutant did not show a greater degree of decondensation that either single mutant.

One interesting observation was that even in nuclei exhibiting severe decondensation of chromocenters, one or two chromocenters often remained intact (Fig. 8). In some nuclei the difference was stark, wherein the chromocenter near the nucleolus was seen as a bright, compact spot and all others were faint and decondensed (Fig. 10). These chromocenters represent the nucleolar organizing regions (NOR), which contain the 45S rDNA and are found on chromosomes 2 and 4 in *Arabidopsis* (Fransz et al. 2002). Unlike the majority of

heterochromatin in the *Arabidopsis* genome, the NORs are not located near the centromeres but are instead found near the telomeres on chromosomes 2 and 4. This region of the genome usually remained condensed, as viewed via DAPI staining and H3K9me2 immunostaining, in each of the wild-type and mutant backgrounds examined.

In addition to DAPI staining and H3K9me2 immunostaining, histone acetylation patterns were detected by immunofluorescence. Whereas DNA methylation and H3K9me2 mark heterochromatin in Arabidopsis, various histone acetylation marks, including histone H4K12Ac, are correlated with euchromatin (Soppe et al. 2002). H4K12Ac immunostaining in wild-type and mutant nuclei revealed a broad distribution across the nucleus (Fig. 11). Staining was excluded from the nucleolus and was somewhat brighter in the interior of the nucleus, where the euchromatin is located, than at the edges. Wild-type nuclei usually displayed chromocenters marked by H3K9me2 at the periphery of the nucleus, outside the area of bright H4K12Ac staining (Fig. 11), although this was difficult to observe simultaneously for all chromocenters in a nucleus at a single focal point. Mutant nuclei often showed subtle differences, with greater colocalization of H3K9me2 and H4K12Ac marks, especially at the nuclear periphery. This result provides additional evidence of altered chromatin organization in *MORC* mutants.

The previous cytological studies revealed altered chromatin architecture in the *MORC* mutants, in which the normally compact and peripheral chromocenters decondense and expand into a broader and more interior region of the nucleus where the euchromatic arms of the chromosomes are found. This data suggested that, in the mutants, the region of pericentromeric heterochromatin, found in the chromocenters, would interact less exclusively with itself and interact more frequently with regions of euchromatin. Hi-C experiments were performed to directly assess chromatin interactions across the entire genome. Mapping of interactions in wild-type showed the strongest interactions among neighboring chromatin regions, as seen along the diagonal axis (Fig. 12A). Euchromatic regions in each chromosome were observed to interact more frequently with the euchromatic regions of the other four chromosomes than with

the pericentromeric heterochromatin, as expected. Mapping of interactions in the *crh6* mutant indicated a similar overall pattern (Fig. 12B). However, calculation of the differences in interactions in *crh6* compared to wild-type revealed important variations (Fig. 12C). Frequency of interaction between the pericentromeric region of each chromosome and the euchromatic regions of each of the five chromosomes increased, as shown in red pixels, while the frequency of interaction between euchromatic regions, especially within each individual chromosome, decreased as shown in blue pixels. Because the sum of relative differences in this plot is set to zero, this analysis cannot distinguish whether the sole change in the mutant was increased interactions of pericentromeric regions with euchromatin, or whether the mutant in fact affected architecture of the euchromatic arms as well. In either case, this data is consistent with the cytological analysis of chromocenter condensation and confirms heterochromatin decondensation in the *crh6* mutant.

In order to verify the action of CRT1 and CRH6 at pericentromeric heterochromatin, complementing Myc-tagged transgenic proteins were expressed in the EMS-mutagenized lines for immunolocalization studies (Fig. 13A, B). Anti-Myc immunostaining showed that CRT1 and CRH6 proteins formed small nuclear bodies near chromocenters (Fig. 13C). These bodies usually numbered one to three per chromocenter and were found adjacent to but not inside the chromocenter. Additional confocal images of CRH6-Myc localization at several depth intervals per nucleus showed CRH6 bodies at the periphery of multiple chromocenters (Fig. 13D). These experiments indicated that the *Arabidopsis* MORC proteins are situated at the border of pericentromeric heterochromatin.

Additional studies examined the function of *morc-1*, the single MORC homolog in *C. elegans*, to determine if the gene has a role in silencing in an animal model (Moissiard et al. 2012). RNAi knockdown of *morc-1* resulted in reactivation of a silenced GFP transgene in seam cells (Fig. 14). Loss of GFP silencing following *morc-1* knockdown occurred to a similar extent as

in knockdown of *rde-4*, a key component of *C. elegans* gene silencing (Tabara et al. 1999). Thus, MORC proteins are required for proper gene silencing in organisms beyond *Arabidopsis*.

DISCUSSION

These studies represent the first observations of the involvement of Arabidopsis *MORC* homologs in gene silencing. Mutations in *CRT1* or *CRH6* result in loss of silencing of *SDC*, a gene with DNA methylation in the tandem repeats of the promoter. *SDC* silencing is lost in the *drm2 cmt3* double mutant, in which CHG and CHH methylation are strongly reduced (Cao and Jacobsen 2002). However, in the *MORC* mutants, *SDC* is reactivated without a reduction of DNA methylation. T-DNA insertion mutations in *CRT1* and *CRH6* displayed similar upregulation of *SDC* as in the EMS mutant alleles discovered in the mutagenesis screen. Examination of genome-wide transcription levels showed many transposons and genes were upregulated in both the T-DNA and EMS mutants. Yet BS-seq experiments indicated that these loci did not show changes in DNA methylation levels in the mutants, and H3K9me2 ChIP-seq experiments corroborated the lack of change in epigenetic silencing marks at the upregulated loci. These results suggested that the MORC proteins may act downstream of epigenetic silencing marks or may reinforce silencing through another pathway independent of DNA methylation.

Loci overexpressed in the *MORC* mutants were found to be located primarily in the regions of pericentromeric heterochromatin. These regions, which condense into compact chromocenters at the nuclear periphery in wild-type nuclei, were found to decondense in the mutants. Decondensation was independently evaluated with Hi-C experiments, which indicated more frequent interactions of pericentromeric heterochromatin with euchromatic regions in the *crh6* mutant. In addition, *MORC* proteins were found in small nuclear bodies adjacent to but not inside the chromocenters. Together, these results suggest that the MORCs act to silence targets through manipulation of chromatin architecture, as has been hypothesized previously (Iyer et al. 2008). In nuclei lacking CRT1 or CRH6 function, the regions of pericentromeric heterochromatin fail to maintain proper condensation in chromocenters and expand out across

the nucleus. Loci in these decondensed regions are transcribed at significantly higher levels without changes in associated epigenetic silencing marks. Future examination of the biochemical activity of MORCs will reveal whether they have topoisomerase activity or other catalytic activity which alters DNA structure.

One notable observation among the studies of *Arabidopsis MORC* mutants is the similarity of the *crt1* and the *crh6* single mutant phenotypes. The genes and transposons upregulated in each single mutant showed a great degree of overlap (Fig. 4) and the chromocenter condensation patterns of each mutant were similar (Figs. 8, 9). Additionally, both CRT1 and CRH6 proteins colocalize with chromocenter peripheries in small nuclear bodies of similar appearance (Fig. 13C). This data suggests that CRT1 and CRH6 may act together in a complex, with the loss of either protein showing similar effects. Also, the decondensation phenotype of the *crt1 crh6* double mutant was no more severe than either single mutant (Figs. 8, 9) indicating that the loss of both MORCs does not have an additive effect. Future studies of the double mutant will determine whether the silencing defects and impact on epigenetic modifications, if any, differ in comparison the single mutants. Protein structure studies may reveal how MORCs interact in complexes to achieve their function.

Another noteworthy result from the immunofluorescence experiments was the lack of decondensation of specific chromocenters in nuclei with otherwise decondensed heterochromatin. In nuclei exhibiting decondensation, the chromocenters adjacent to nucleoli usually remained intact (Fig. 8, 10). These chromocenters represent the nucleolar organizing regions (NOR) on chromosomes 2 and 4 rather than pericentromeric heterochromatin found in the other chromocenters (Fransz et al. 2002). Previous work has determined that mutations which reduce DNA methylation levels across the genome, such as *met1* and *ddm1*, reduce the size of chromocenters and cause dispersion of pericentromeric heterochromatin from chromocenters; however, the 45S rDNA repeats, found in the NORs, remain colocalized with chromocenters in these mutants (Soppe et al. 2002). This result suggests additional

mechanisms exist which maintain heterochromatin condensation of NORs even when reduction of DNA methylation leads to dispersion of pericentromeric heterochromatin. Consistent with this scenario, mutations in *CRT1* and *CRH6* are found to disrupt pericentromeric condensation without affecting organization of the NORs. This suggests that the mechanisms which maintain the chromatin architecture of NORs can operate independently of DNA methylation and MORC activity.

The analyses of H4K12Ac immunostaining indicate increased colocalization of H4K12Ac modification with H3K9me2 in *MORC* mutants, as expected for nuclei with defects in heterochromatin assembly. This effect is especially noticeable at the nuclear periphery, a region where H3K9me2-marked heterochromatin is located and H4K12Ac-marked euchromatin is excluded (Soppe et al. 2002). In mutants with decondensed chromocenters, H4K12Ac was more visible at the nuclear periphery (Fig. 11), suggesting either increased localization of euchromatin at the periphery, increased acetylation of pericentromeric heterochromatin, or both. Conclusive evidence supporting these hypotheses awaits future experimental determination of the distribution of H4K12Ac and other acetylation marks in *MORC* mutants.

Other *Arabidopsis* genes encoding chromatin remodeling enzymes are known to affect gene silencing. Mutations in theSWI2/SNF2-like chromatin remodeling factor *DDM1* result in reduced DNA methylation levels and in decondensation of chromocenters (Probst et al. 2003). Mutations in another member of the SWI/SNF2 family, the plant-specific *MOM1*, cause a loss of silencing of methylated loci without reduction in the DNA methylation levels (Amedeo et al. 2000). However, *mom1* mutants do not display chromocenter decondensation (Probst et al. 2003). *MORC* mutants have some phenotypes in common with mutants in *DDM1* and *MOM1*, yet *MORC* is unique in affecting heterochromatin condensation without affecting levels of DNA methylation. These results indicate the cooperation of multiple pathways to maintain proper gene silencing. DNA methylation represents one level of regulation, and DDM1 participates in this level as it is required for proper maintenance of DNA methylation (Jeddeloh et al. 1999).

Heterochromatin compaction represents an additional, possibly independent level, as loss of *CRT1* or *CRH6* results in TE reactivation and chromocenter decondensation. *MOM1*, whose mode of action remains unknown (Habu 2010), likely functions separately from the MORC pathway since it does not affect the organization of chromocenters.

The Arabidopsis MORC gene *CRT1* was first described and named for its role as a gene responsible for viral resistance (Kang et al. 2008). Thus, it appears that this gene may play roles in multiple cellular processes including disease resistance and gene silencing. One possibility accounting for these disparate functions may be the pleiotropic effects of mutations in genes required for silencing on disease resistance. Previous studies have shown that mutations in *AGO4*, a component of the RNA-directed DNA methylation (RdDM) pathway, affect susceptibility to virulent bacteria (Agorio and Vera 2007) and mutations in *Pol V*, another essential player in RdDM, also reduce plant immunity (Lopez et al. 2011). Likewise, loss of *CRT1* potentially weakens plant disease resistance as a secondary effect of the loss of gene silencing.

Reactivation of a silenced transgene in *C. elegans* following *morc-1* RNAi knockdown demonstrated that MORC is required for transcriptional repression in worms as it is in *Arabidopsis*. Intriguingly, *C. elegans* lacks DNA methylation, and therefore silencing by MORC occurs via a DNA methylation independent mechanism in this species. This observation lends support for the idea that *Arabidopsis* MORCs silence their target loci through a mechanism that is independent of, rather than downstream of, DNA methylation. Examination of genome-wide transcription defects as well as chromatin organization in *C. elegans* nuclei following *morc-1* knockdown can assess whether phenotypes of *MORC* mutants in *Arabidopsis* are also seen in worms and provide clues about the conserved function of MORCs across eukaryotes.

In addition to worms, *MORC* homologs are found in mammals. The mouse *Morc-1* gene, named for microrchidia, or small testes, was originally identified as a mutation that causes arrest of spermatogenesis at the zygotene stage of meiosis I (Watson et al. 1999). As previously mentioned, loss of transposon silencing, as occurs in Dnmt3L-deficient mouse cells, leads to

arrest of spermatocyte development at the same stage (Bourc'his and Bestor 2004). Mutations in other components of silencing in the germ line, such as MIWI2, similarly induce demethylation and loss of silencing of transposons as well as failure in spermatocytes to progress to the pachytene stage (Carmell et al. 2007). The resemblance between the developmental arrests in mice lacking *Morc-1* and those lacking other silencing components may indicate that the arrest in *Morc*-1 is also caused by inability to repress transpsons. Future experiments exploring the transcriptome of *Morc*-1 knockout mice may support this idea. Such evidence of the function of MORC in mice will further establish the conserved role of MORCs in eukaryotic gene silencing.

In summary, these studies have identified a novel role of MORC family ATPases in gene silencing. In Arabidopsis, the MORC genes *CRT1* and *CRH6* are each required for proper silencing of many TEs and genes, especially those located in the regions of pericentromeric heterochromatin. Mutation in these genes release silencing without disrupting DNA methylation or H3K9me2 modification. The MORC proteins form bodies near chromocenters and are responsible for maintaining the condensation of pericentromeric heterochromatin, though they appear to act through a separate pathway than several other known chromatin remodeling enzymes. MORC is also required for gene silencing in *C. elegans*, and may be responsible for TE repression in mammals as well. Future studies of the biochemical activity and protein structure of *Arabidopsis* MORCs will elaborate on their mechanism of silencing, and future work in animal models will further reveal the function of this conserved gene family across many eukaryotic lineages.

FIGURE 1A

Figure 1A. Schematic of the *SDC::GFP* transgene. Black arrows represent the seven tandem repeats in the *SDC* promoter which are targeted by DNA methylation. The red bar represents the SV40 nuclear localization signal fused to GFP.

FIGURE 1B

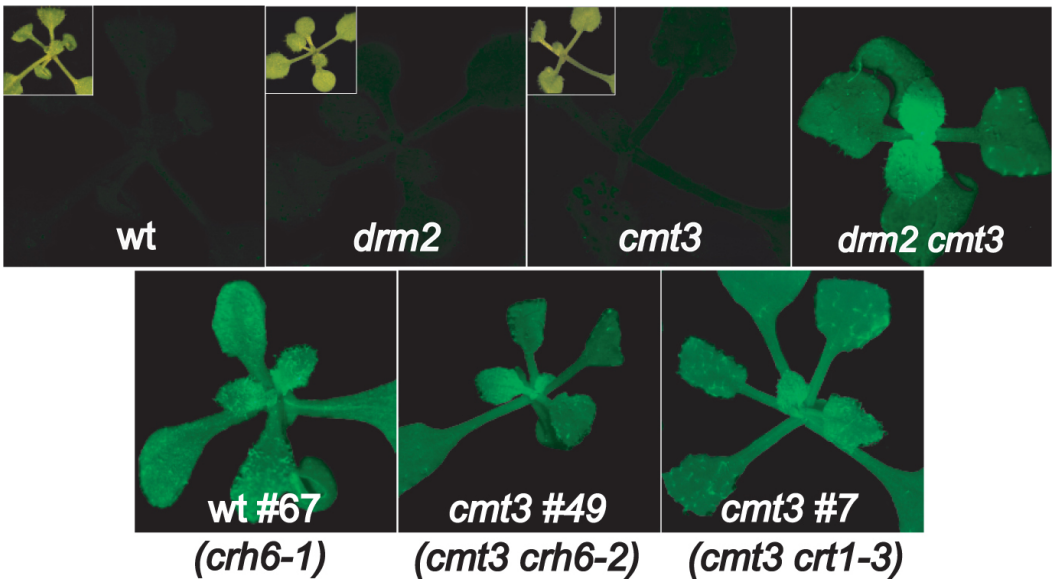


Figure 1B. GFP expression in wild-type, *drm2* single mutant, *cmt3* single mutant, *drm2 cmt3* double mutant, and EMS-mutagenized lines. Wild-type and single mutant plants with the *SDC::GFP* transgene did not display GFP fluorescence under UV light, while the double mutant and the three EMS-mutagenized lines shown did display GFP fluorescence.

FIGURE 1C

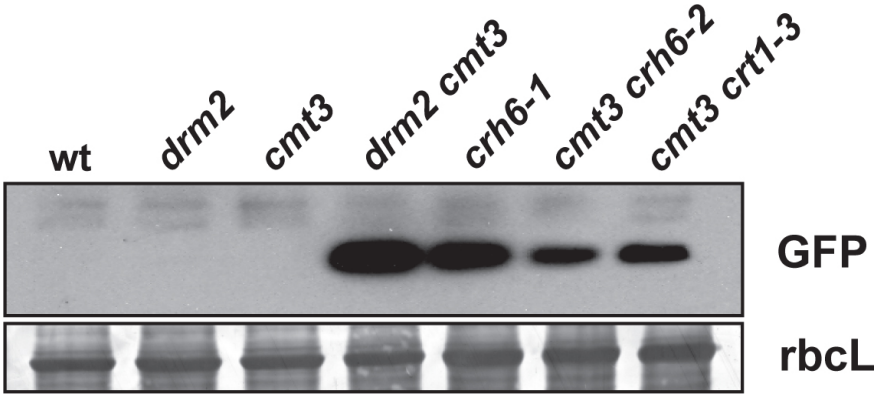


Figure 1C. GFP expression in wild-type, *drm2* single mutant, *cmt3* single mutant, *drm2 cmt3* double mutant, and EMS-mutagenized lines. Western blot with anti-GFP antibody confirms upregulation of the *SDC::GFP* transgene in the double mutant and in the EMS-mutagenized lines shown. Coomassie staining of rbcL (large Rubisco subunit) is a loading control.

FIGURE 2A

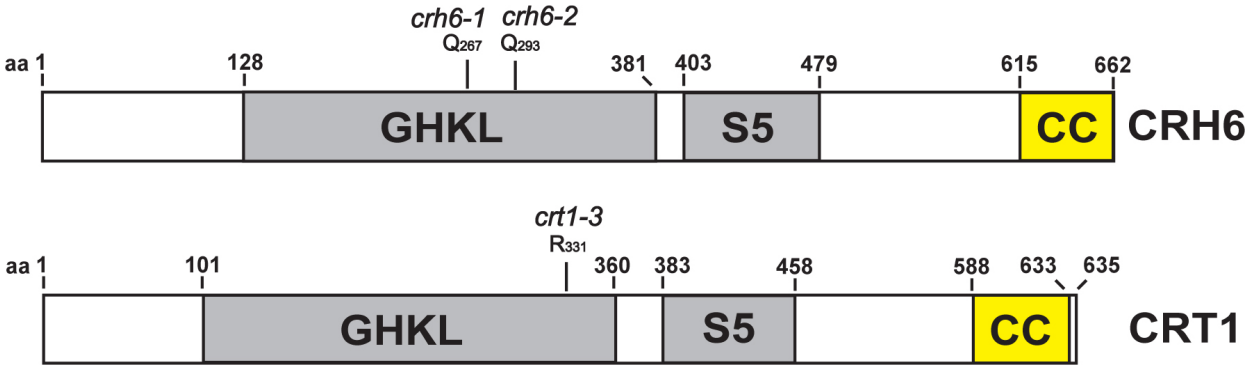


Figure 2A. Schematic of CRH6 and CRT1 proteins. The locations of the GHKL and S5 ATPase and the putative coiled-coil (CC) domains are shown. The locations of the mutation in each EMS-mutagenized line are shown. Each of the three mutations represents a premature stop codon.

FIGURE 2B

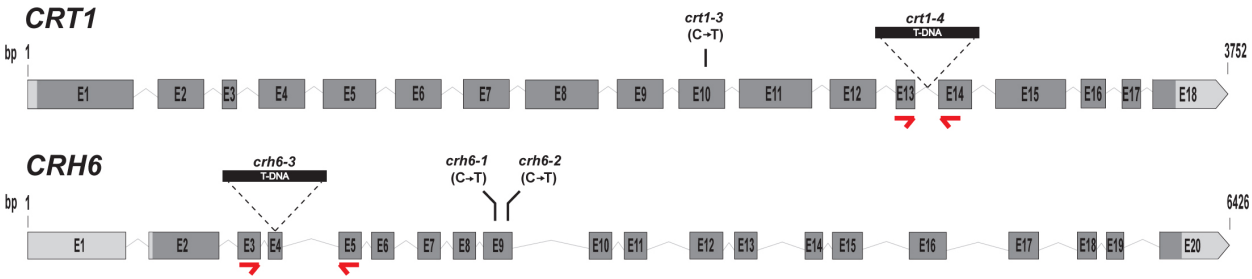


Figure 2B. Structure of the CRT1 and CRH6 genes. Dark gray boxes represent exons, gray lines represent introns, and gray boxes represent untranslated regions. Red arrows represent the primer locations within each gene used for RT-qPCR analysis of RNA levels. The location and nature of the EMS mutations and T-DNA insertions are shown.

FIGURE 2C

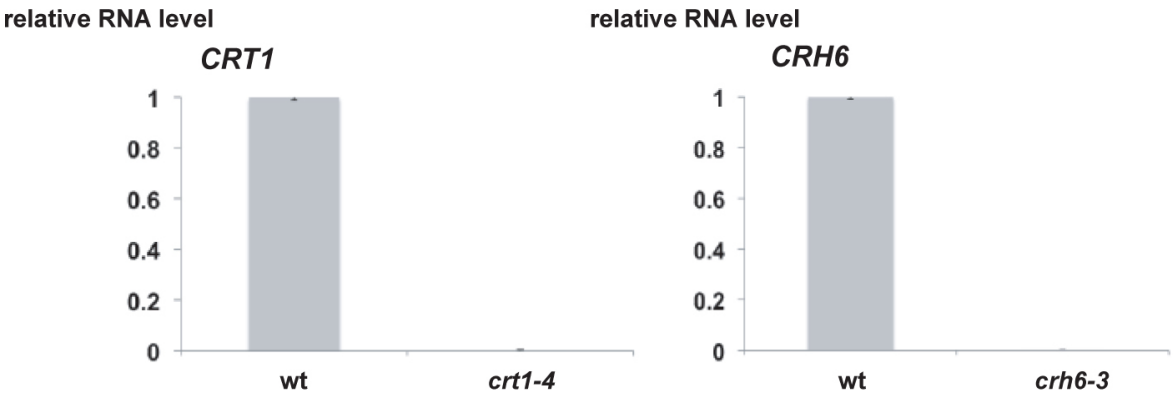


Figure 2C. Relative expression of *CRT1* and *CRH6* in T-DNA insertion lines. RNA levels were analyzed by RT-qPCR using primers with locations as indicated in Figure 2B. Levels were normalized to *ACTIN7* and compared to expression levels in wild-type plants.

FIGURE 2D

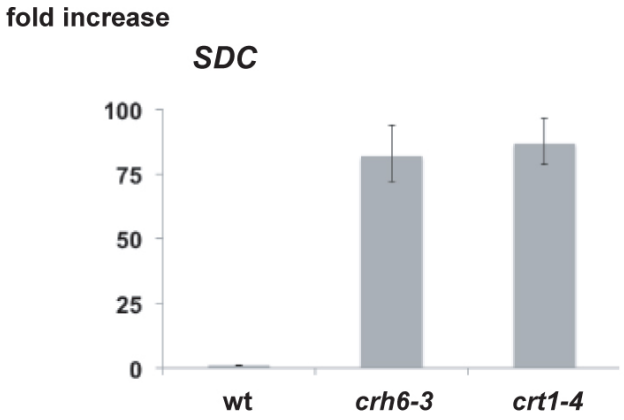


Figure 2D. Relative expression of *SDC* in T-DNA insertion lines. RNA levels were analyzed by RT-qPCR. Levels were normalized to *ACTIN7* and compared to expression levels in wild-type plants.

FIGURE 3A

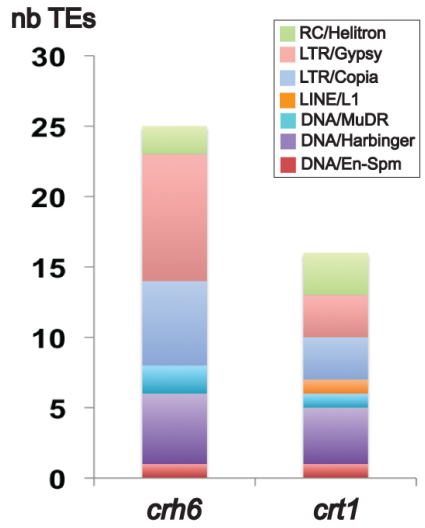


Figure 3A. Overexpression of TEs in *crt1* and *crh6* mutants. Expression levels were quantified with RNA-seq. Only TEs overexpressed in both EMS and T-DNA mutants with a four-fold or greater increase relative to wild-type and with p≤0.05 are shown. Number of overexpressed TEs, grouped by superfamily, are indicated.

FIGURE 3B

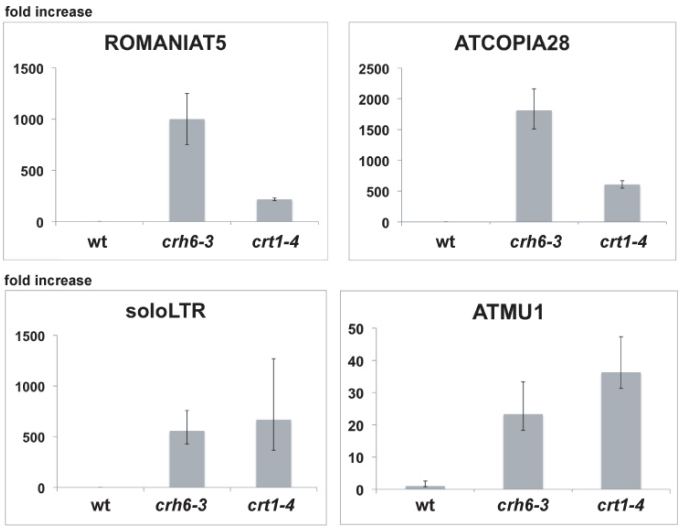


Figure 3B. Overexpression of TEs in *crt1* and *crh6* T-DNA insertion mutants. RNA levels for four TEs were quantified using RT-qPCR, normalized to *ACTIN7*, and compared to wild-type. Error bars display standard deviation of three independent biological replicates.

FIGURE 3C

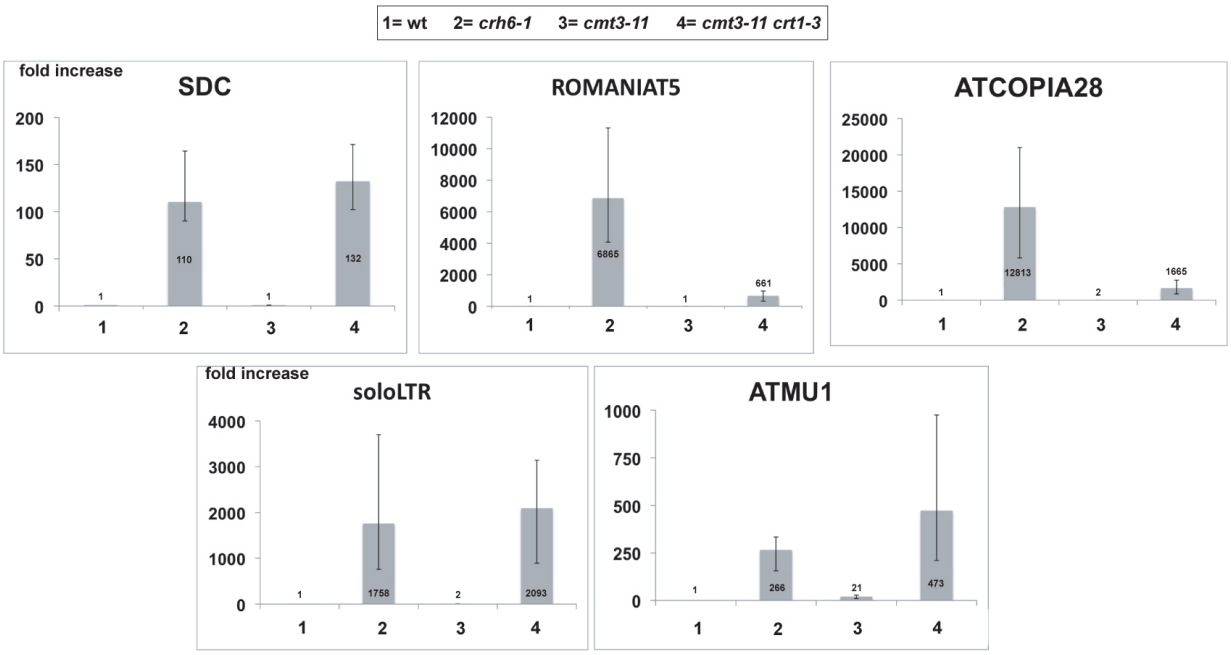


Figure 3C. Overexpression of *SDC* and TEs in *crt1* and *crh6* EMS mutants. RNA levels were quantified using RT-qPCR, normalized to *ACTIN7*, and compared to wild-type. Error bars display standard deviation of three independent biological replicates.

FIGURE 4

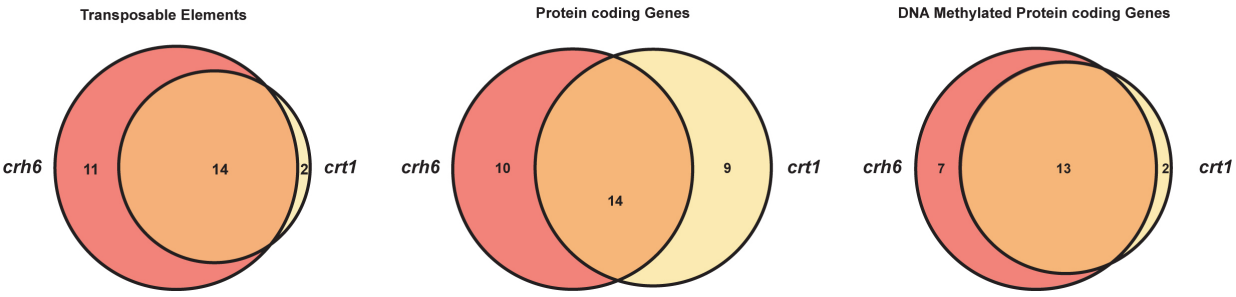


Figure 4. Comparison of overexpressed loci in *crt1* and *crh6* mutants. Expression levels were quantified with RNA-seq. Only those loci overexpressed in both EMS and T-DNA mutants, in each case with a four-fold or greater increase relative to wild-type, are shown. Venn diagrams display the overlap in numbers of overexpressed loci in *crt1* compared to *crh6* for TEs, protein-coding genes, and protein-coding genes that are DNA methylated.

FIGURE 5A

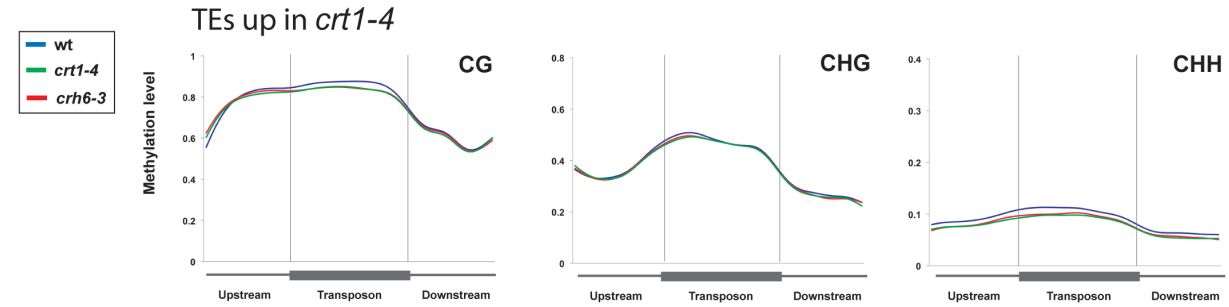


Figure 5A. DNA methylation levels of TEs upregulated in *crt1*. Whole-genome bisulfite sequencing (BS-seq) was used to analyze DNA methylation patterns in the TEs that were overexpressed in *crt1* mutant relative to wild-type. DNA methylation in the CG, CHG, and CHH contexts is compared between wild-type, *crt1* T-DNA mutant, and *crh6* T-DNA mutant plants. Metaplots shown DNA methylation levels within the transposon boundaries marked in gray vertical lines as well as 1 kilobase upstream and downstream of the TE.

FIGURE 5B

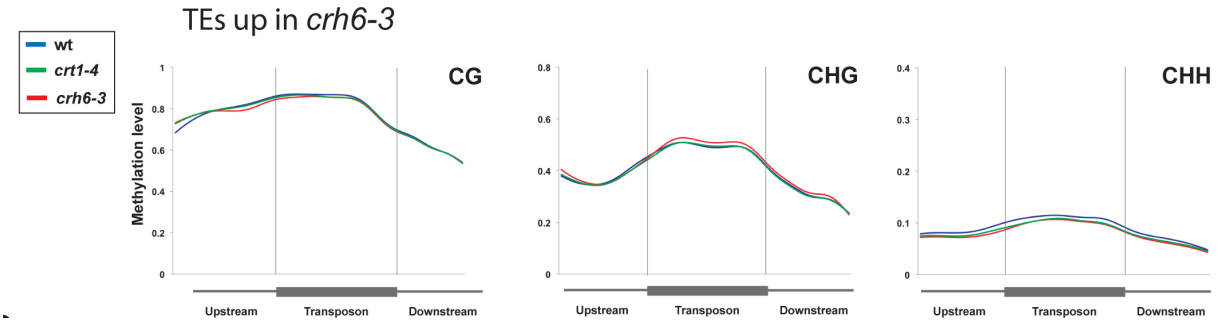


Figure 5B. DNA methylation levels of TEs upregulated in *crh6*. Whole-genome bisulfite sequencing (BS-seq) was used to analyze DNA methylation patterns in the TEs that were overexpressed in *crh6* mutant relative to wild-type. Metaplots are constructed as in Figure 5A.

FIGURE 5C

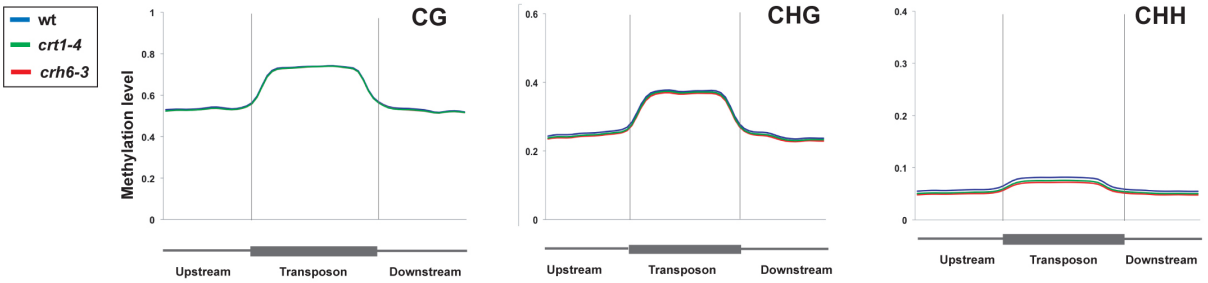


Figure 5C. DNA methylation levels of all TEs. Whole-genome bisulfite sequencing (BS-seq) was used to analyze DNA methylation patterns in all transposons. Metaplots are constructed as in Figure 5A.

FIGURE 5D

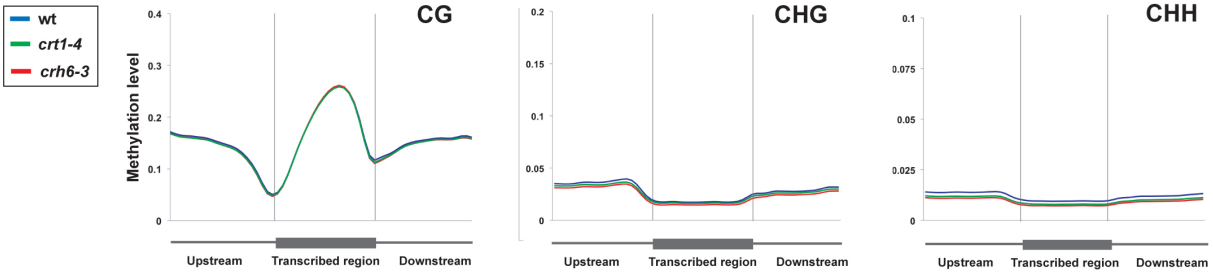


Figure 5D. DNA methylation levels of all protein-coding genes. Whole-genome bisulfite sequencing (BS-seq) was used to analyze DNA methylation patterns in all genes. Metaplots are constructed as in Figure 5A and show DNA methylation levels within the gene body with boundaries marked in gray vertical lines as well as 1 kilobase upstream and downstream of the gene.

FIGURE 5E

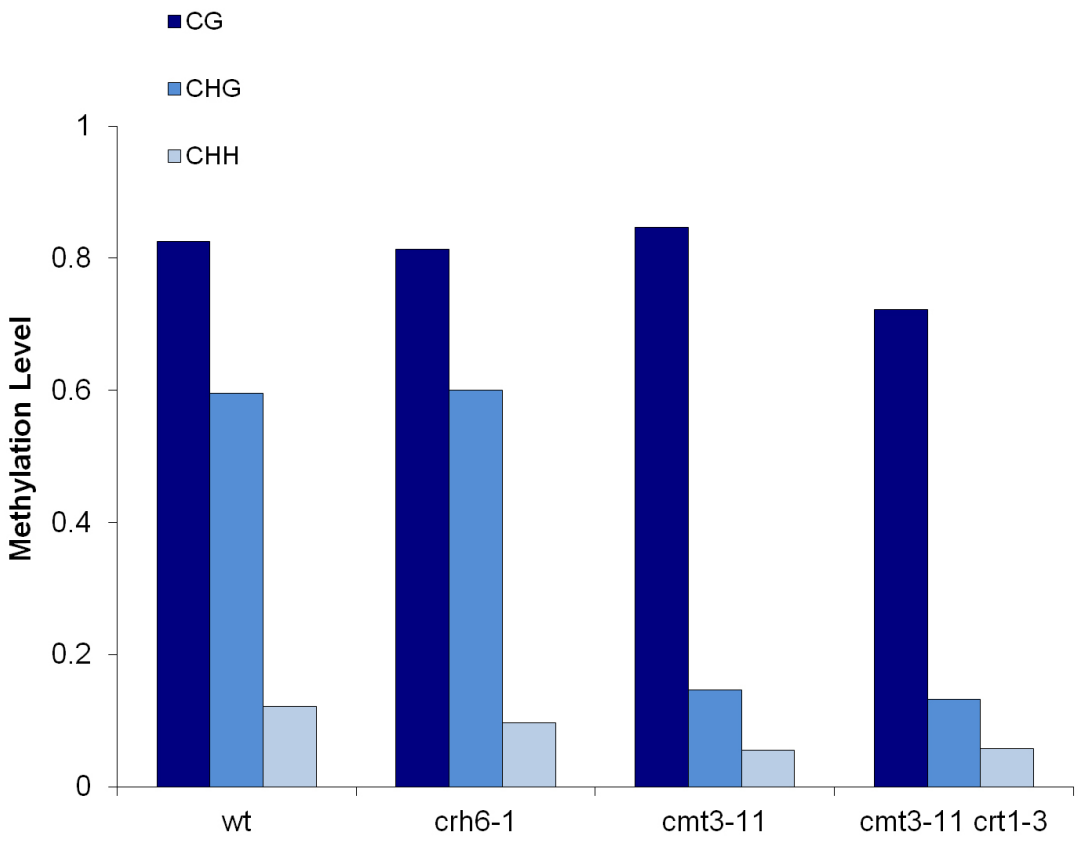


Figure 5E. DNA methylation levels of *SDC* promoter in EMS mutagenized lines analyzed by traditional bisulfite sequencing. Percent DNA methylation in CG, CHG, and CHH contexts is shown for *SDC* promoter in wild-type, *crh6-1* mutant, *cmt3-11* and *cmt3-11 crt1-3* mutant plants. Wild-type is the control for *crh6* EMS mutant (found in the mutagenesis screen in the wild-type background), and *cmt3-11* is the control for the *crt1* EMS mutant (found in the screen in the *cmt3* background). Twenty-four clones were analyzed for each locus.

FIGURE 5F

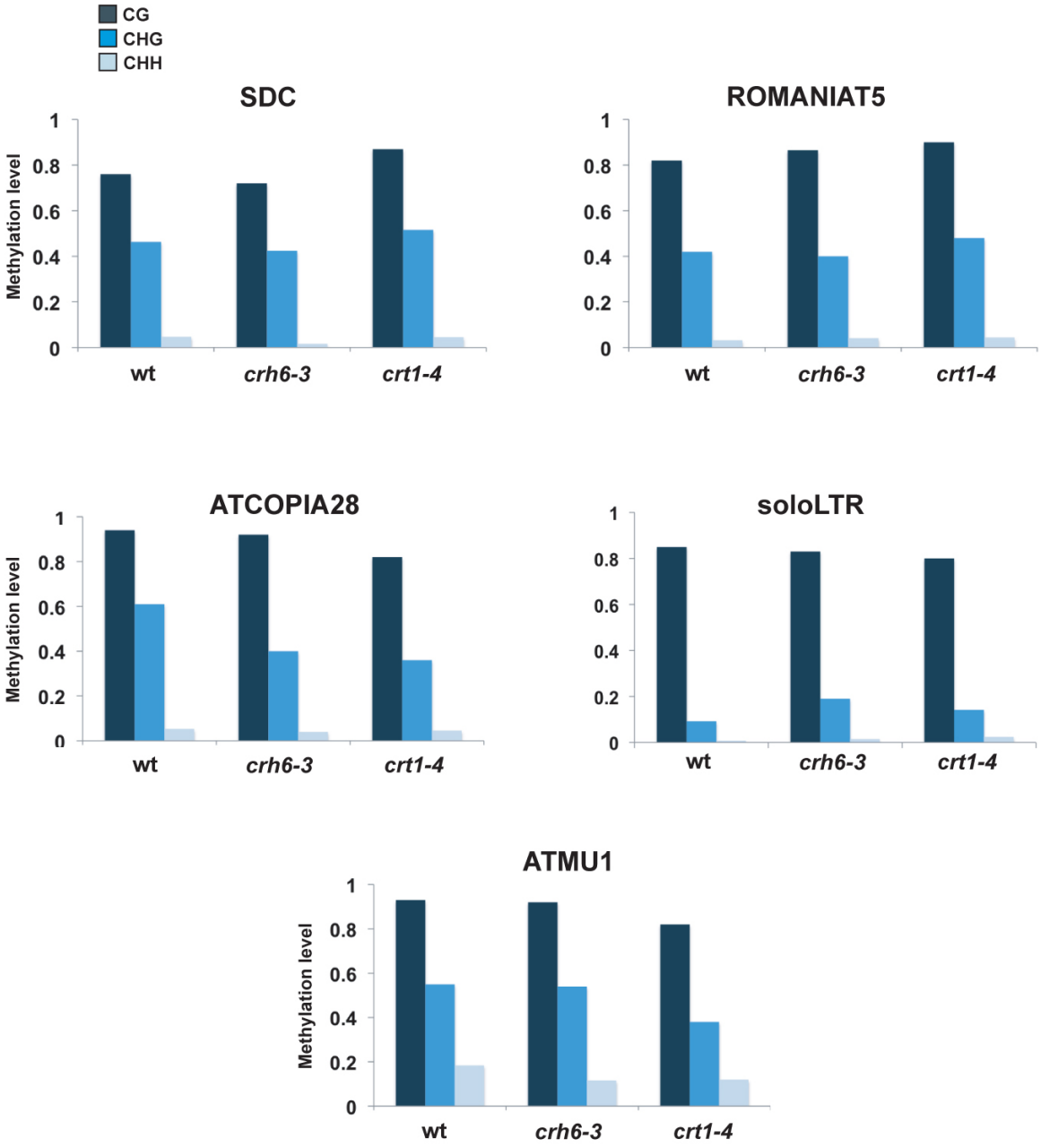


Figure 5F. DNA methylation levels of upregulated loci in T-DNA insertion mutants analyzed by traditional bisulfite sequencing. Percent DNA methylation in CG, CHG, and CHH contexts is shown for *SDC* and four upregulated TEs in wild-type, *crh6-3* mutant and *crt1-4* mutant plants. Twenty-four clones were analyzed for each locus.

FIGURE 6A

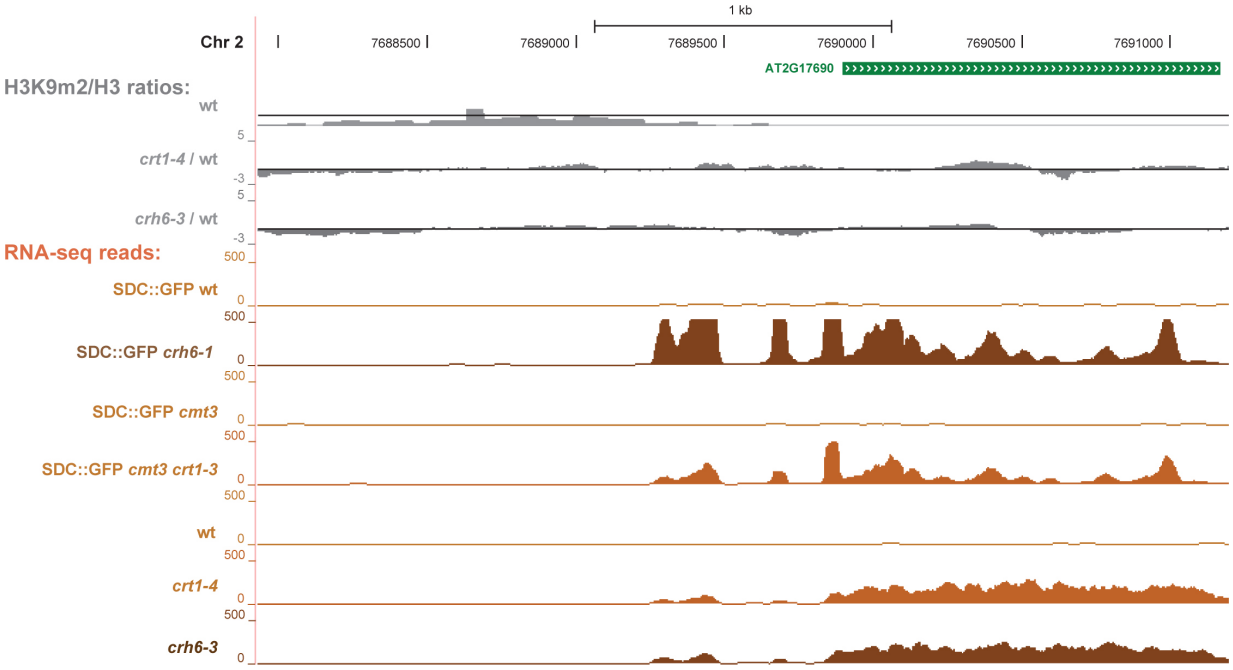


Figure 6A. Analysis of H3K9me2 levels at the *SDC* promoter. ChIP-seq reads for H3K9me2 were normalized to reads for H3 for wild-type, *crt1* mutant and *crh6* mutant plants. The $\log_2$ ratio of H3K9me3/H3 was compared between the mutants and wild-type. RNA-seq reads in the indicated genotypes are displayed to define the *SDC* transcribed region.

FIGURE 6B

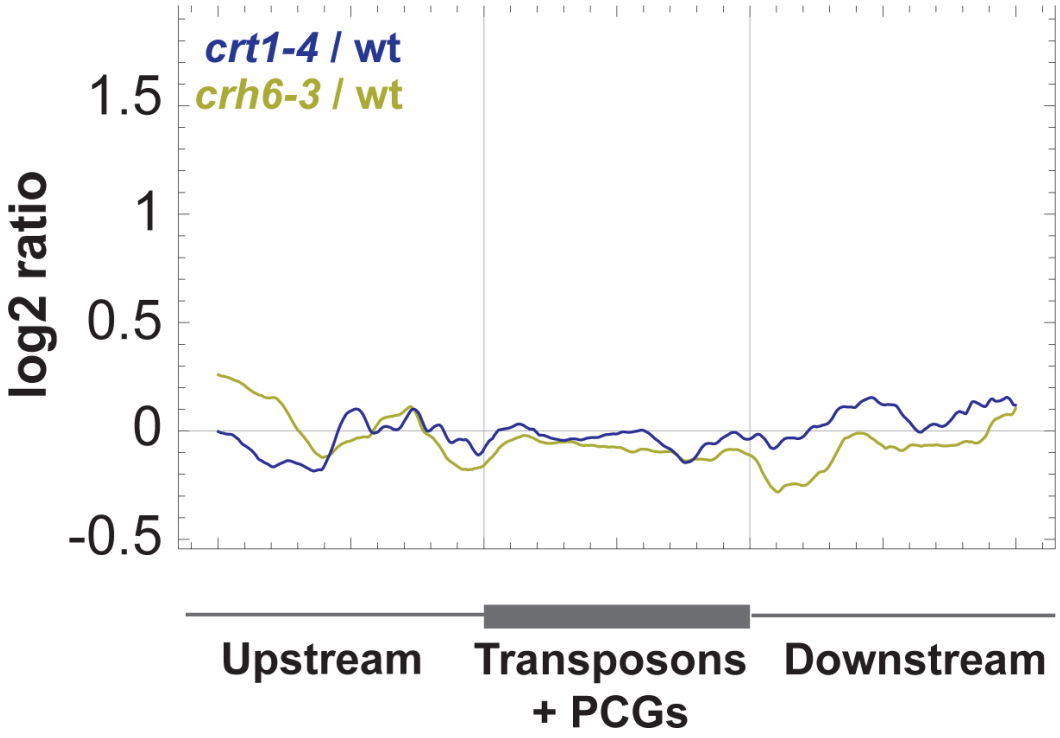


Figure 6B. Comparison of H3K9me2 levels in *crt1* and *crh6* mutants for all upregulated TEs and DNA methylated genes. Metaplot shows the $\log_2$ ratio of H3K9me3/H3 for *crt1* mutant and *crh6* mutant divided by the same ratio in wild-type plants. H3K9me2 levels are shown for the set of all transposons and DNA methylated genes upregulated in *crt1* and *crh6* relative to wild-type. Gray vertical lines show the boundaries between the transposon/gene and the regions spanning 1 kilobase upstream and downstream.

FIGURE 7A

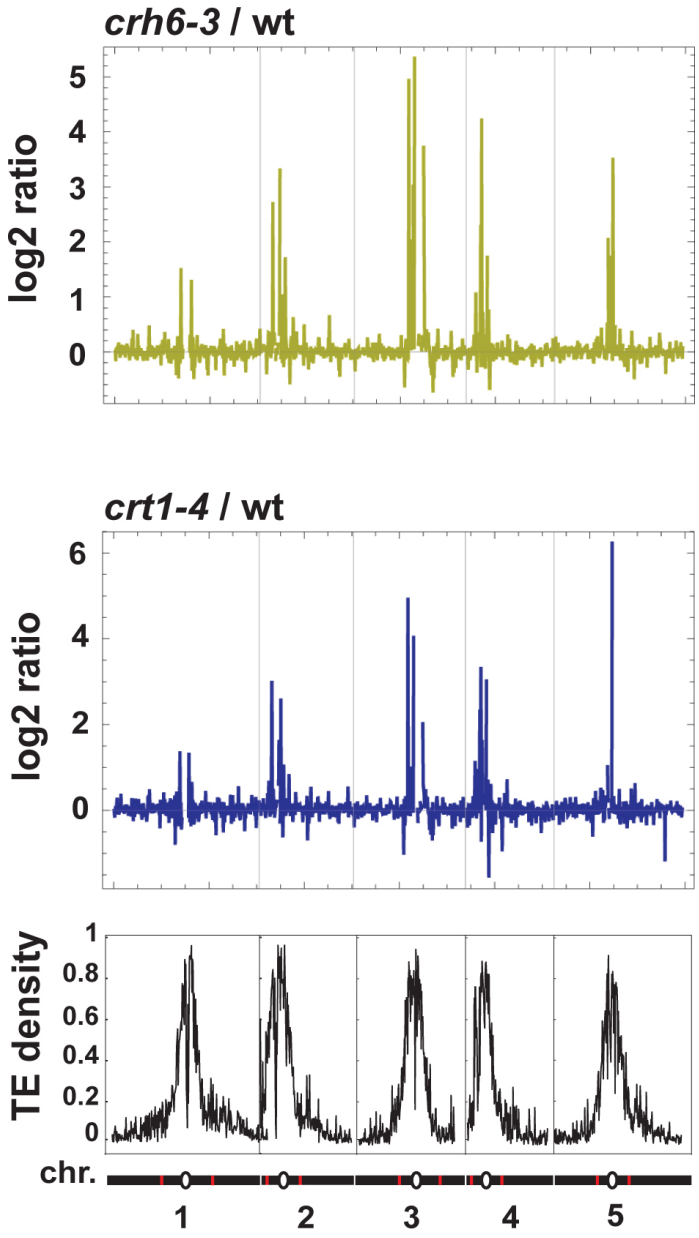


Figure 7A. Chromosomal location of upregulated loci in *crt1* and *crh6* mutants. The $\log_2$ ratio of number of RNA-seq reads in each mutant divided by wild-type is shown in 100 kilobase bins. White circles represent the centromeres and red lines represent the borders of the pericentromeric region of each chromosome.

FIGURE 7B

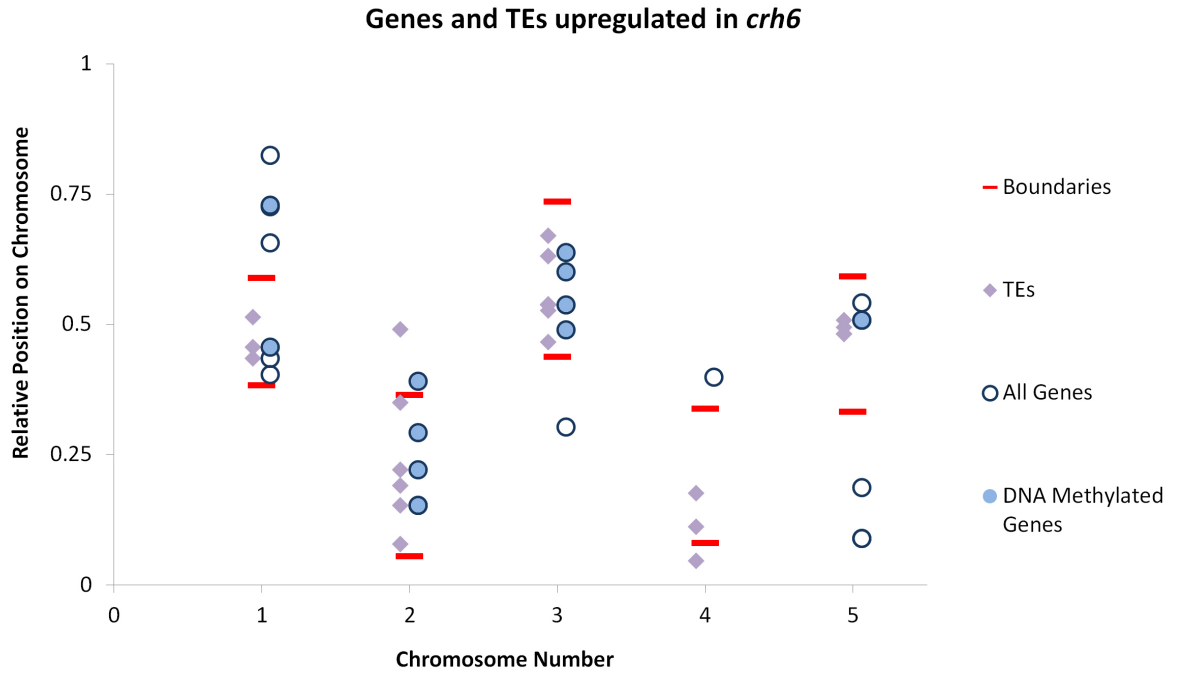


Figure 7B. Plot of chromosomal locations of loci upregulated in *crh6* mutant. The selection criteria for the set of upregulated loci in RNA-seq experiments is described in Figure 4. The boundaries of pericentromeric heterochromatin, as previously described (Bernatavichute et al. 2008), are shown as red bars along the relative position on each chromosome. While 26% of the *Arabidopsis* genome consists of pericentromeric heterochromatin, 58% of upregulated genes, 85% of upregulated DNA methylated genes, and 92% of upregulated TEs in the *crh6* mutant are found within the boundaries.

FIGURE 7C

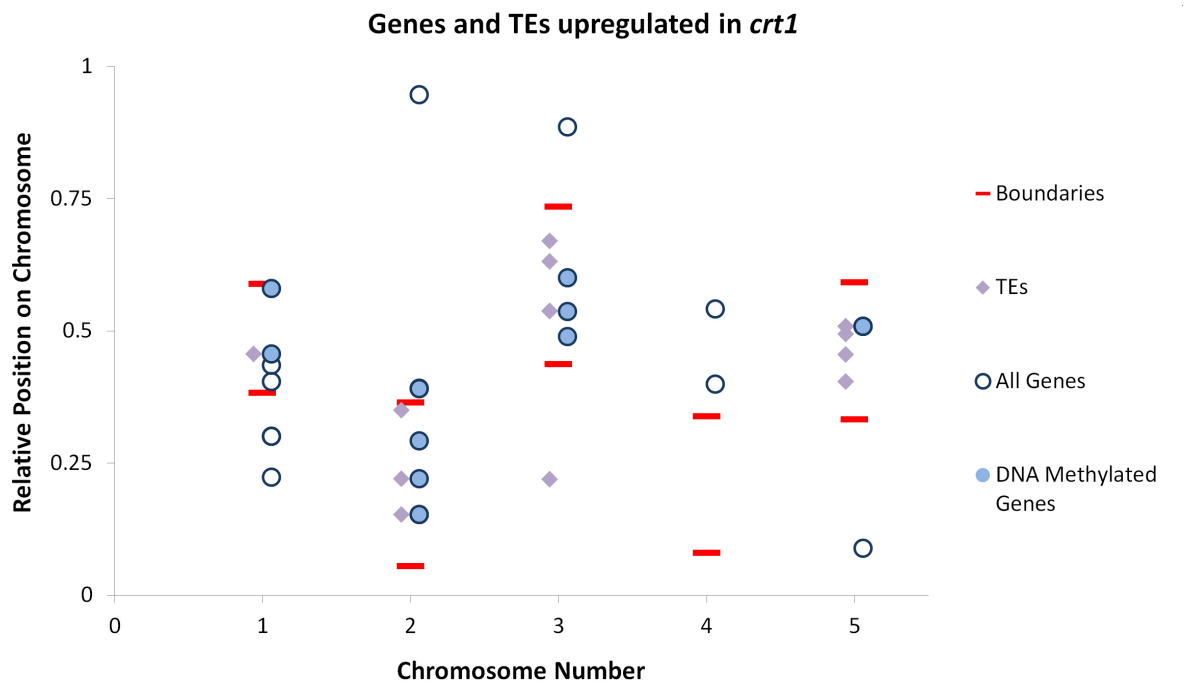


Figure 7C. Plot of chromosomal locations of loci upregulated in *crt1* mutant, constructed as in Figure 7B. For the *crt1* mutant, 58% of upregulated genes, 91% of upregulated DNA methylated genes, and 93% of upregulated TEs are located within the boundaries of pericentromeric heterochromatin.

FIGURE 8A

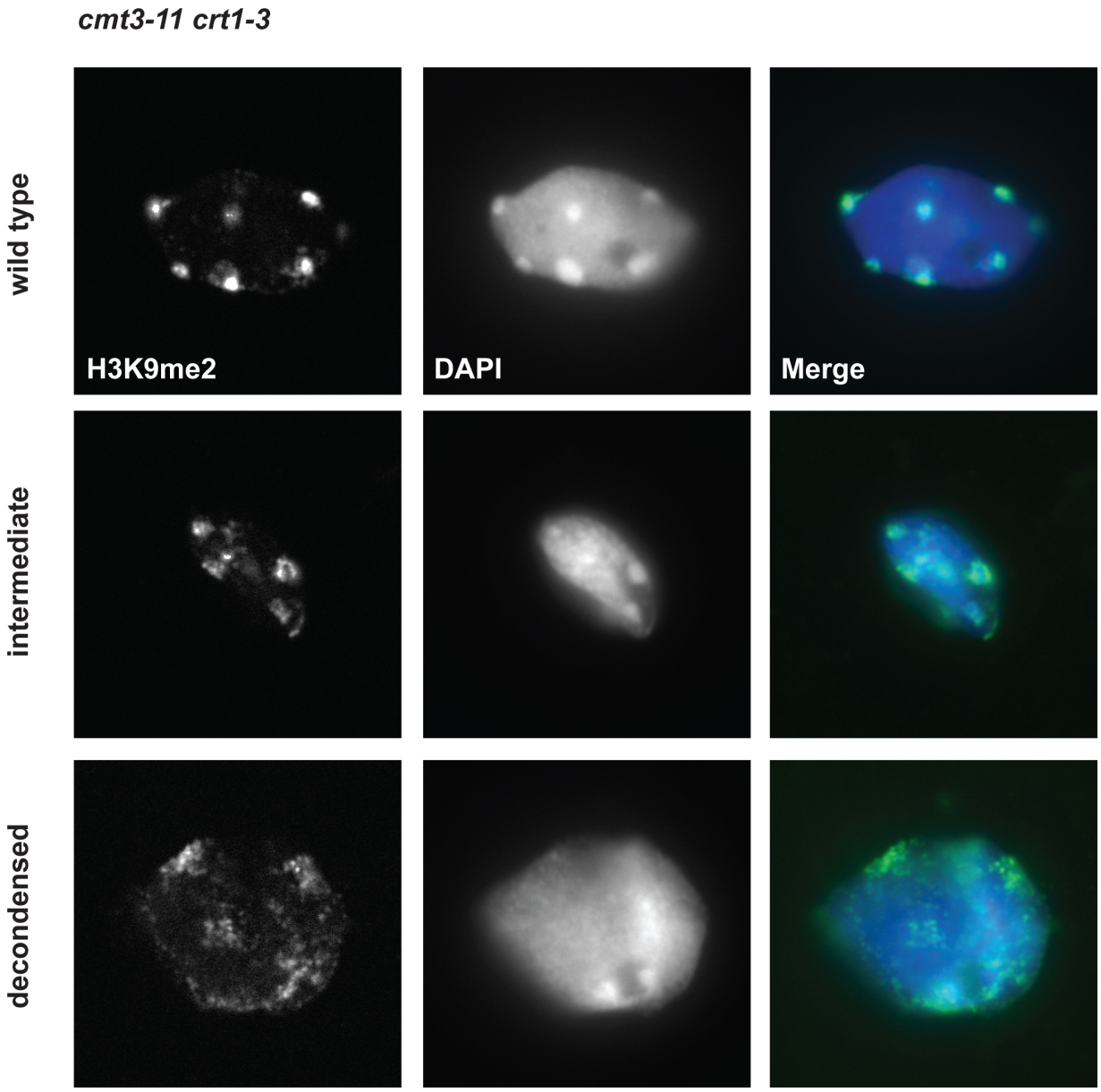


Figure 8A. Chromocenter condensation in *crt1* EMS mutant nuclei. DAPI staining and H3K9me2 immunostaining show chromocenters in nuclei representative of wild-type condensation patterns, partially decondensed (intermediate) patterns, and decondensed patterns.

FIGURE 8B

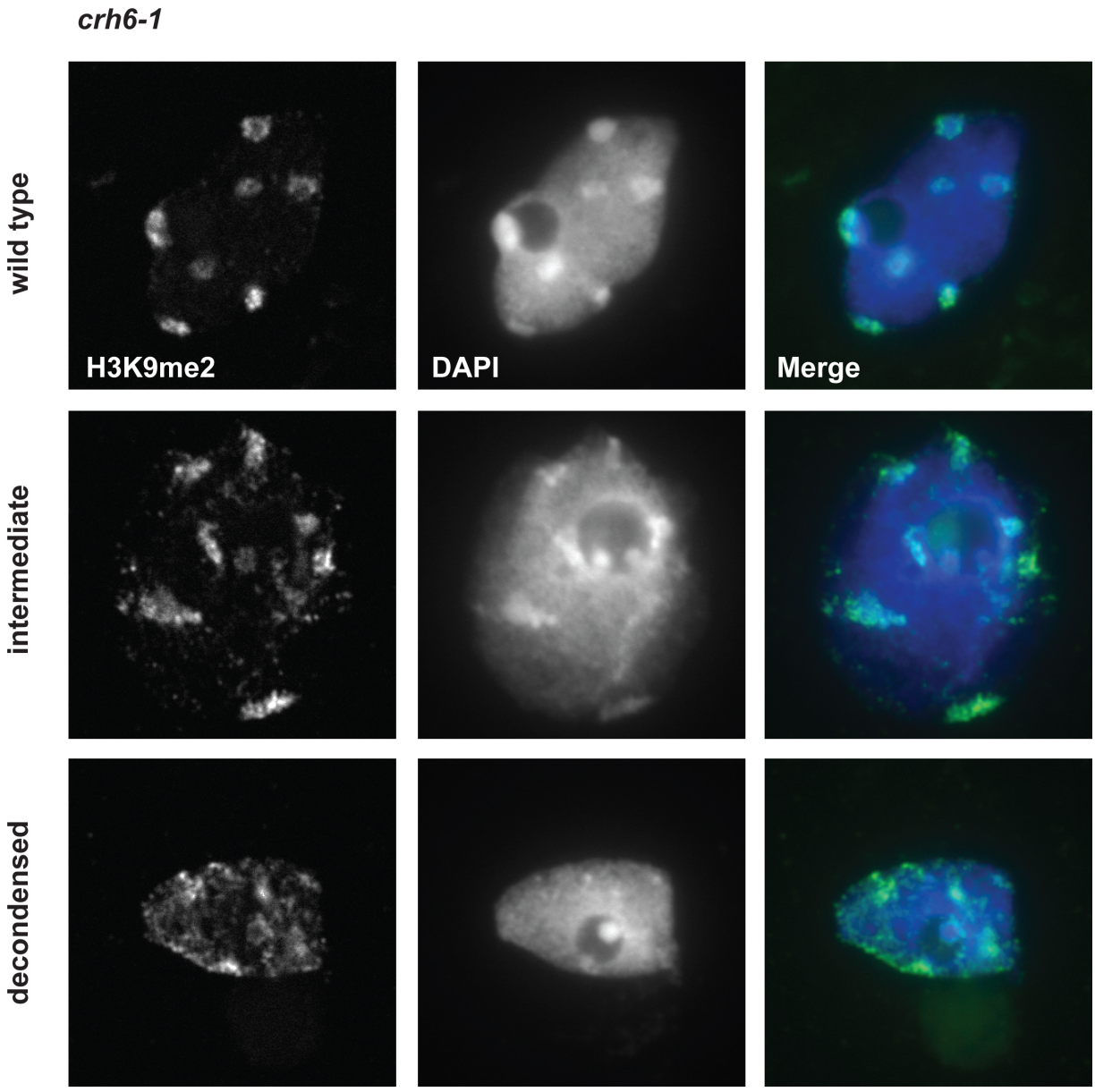


Figure 8B. Chromocenter condensation in *crh6* EMS mutant nuclei shown as in Figure 8A.

FIGURE 8C

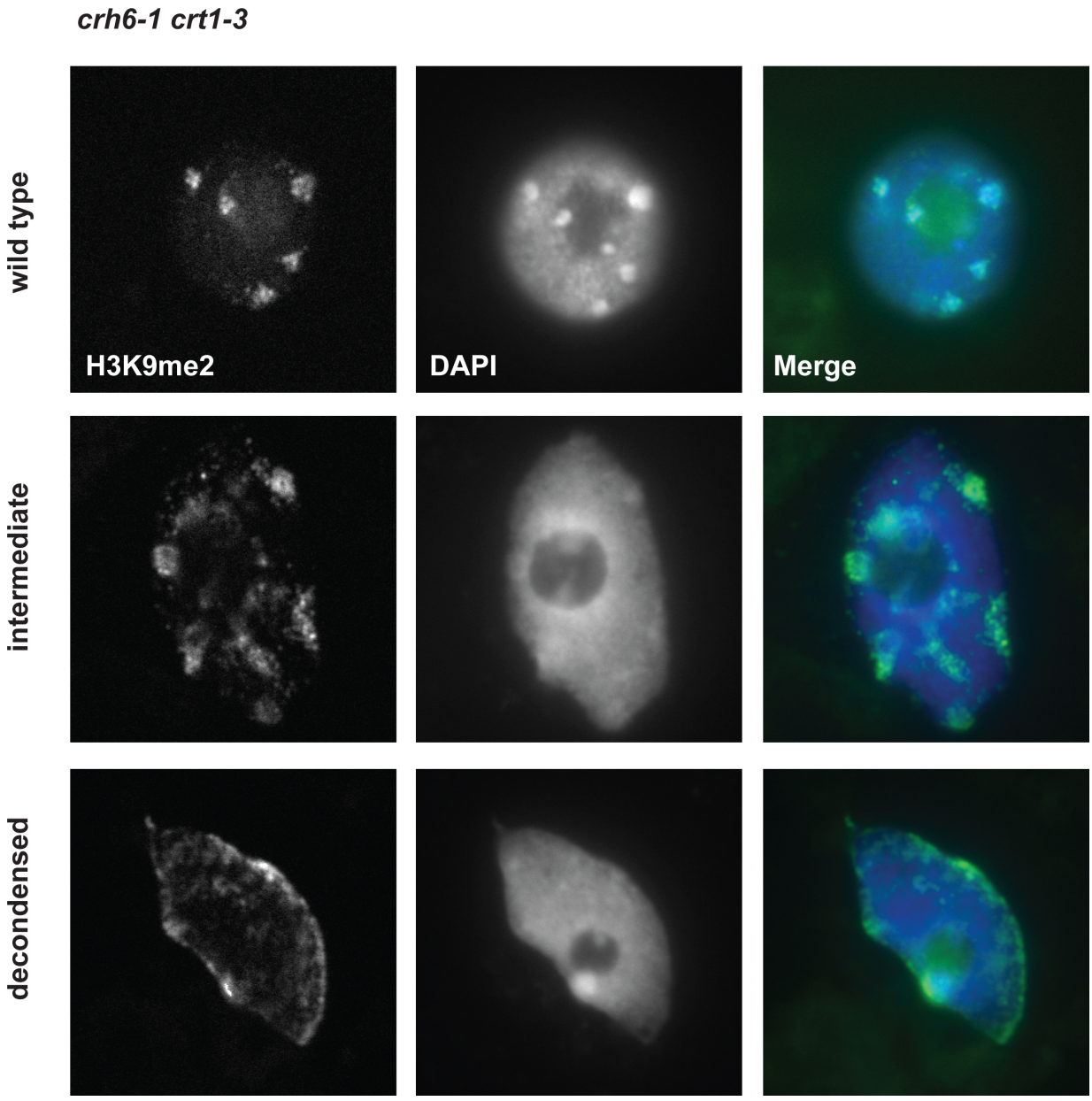


Figure 8C. Chromocenter condensation in *crt1 crh6* double mutant nuclei shown as in Figure 8A.

FIGURE 9

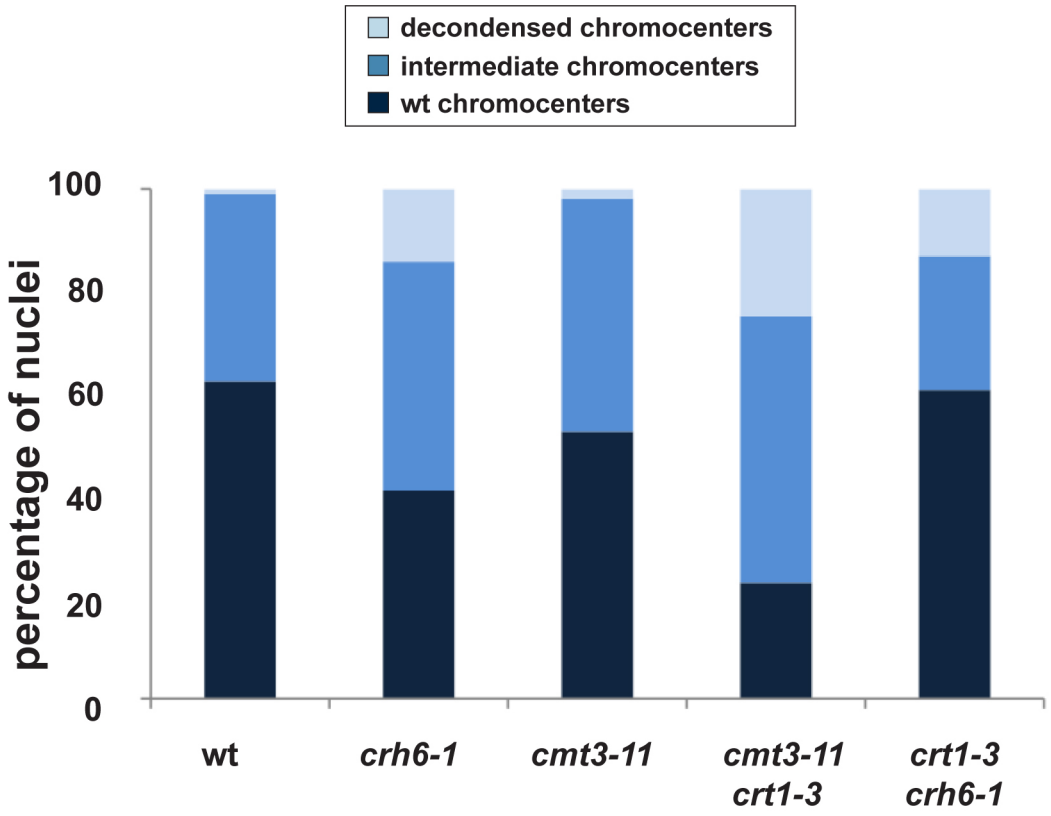


Figure 9. Patterns of chromocenter decondensation in mutant and wild-type nuclei. Percentages of nuclei showing chromocenter patterns as defined in Figure 8 were analyzed by H3K9me2 immunostaining. One hundred nuclei in each genetic background were analyzed. Wild-type nuclei serve as the control for *crh6-1* mutants, and *cmt3-11* nuclei serve as the control for *cmt3-11 crt1-3* mutants.

Figure 10

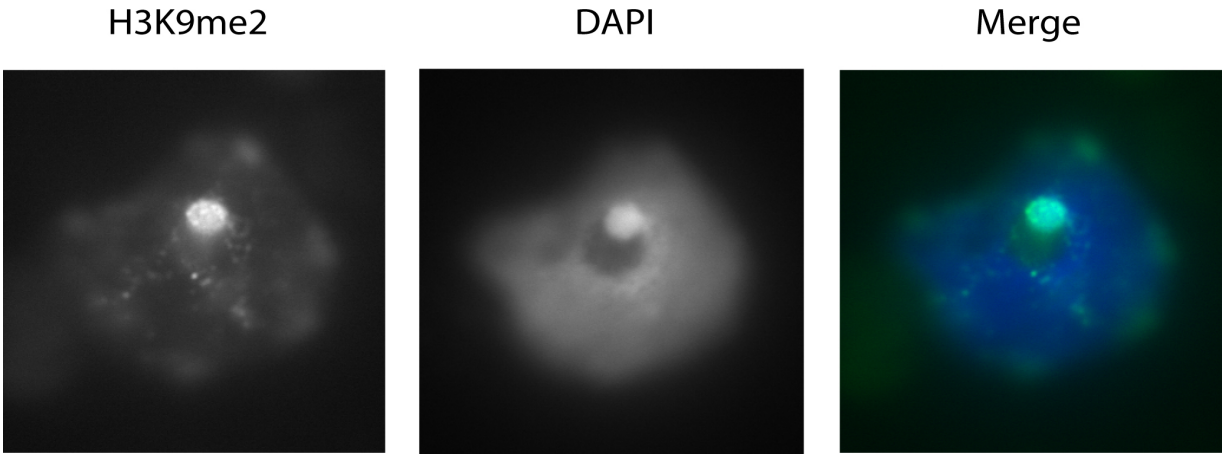


Figure 10. Image of chromocenter condensation in *crh6-1* mutant, taken as in Figure 8A. The brightest chromocenter, corresponding to the nucleolar organizing region (NOR), is not decondensed.

Figure 11

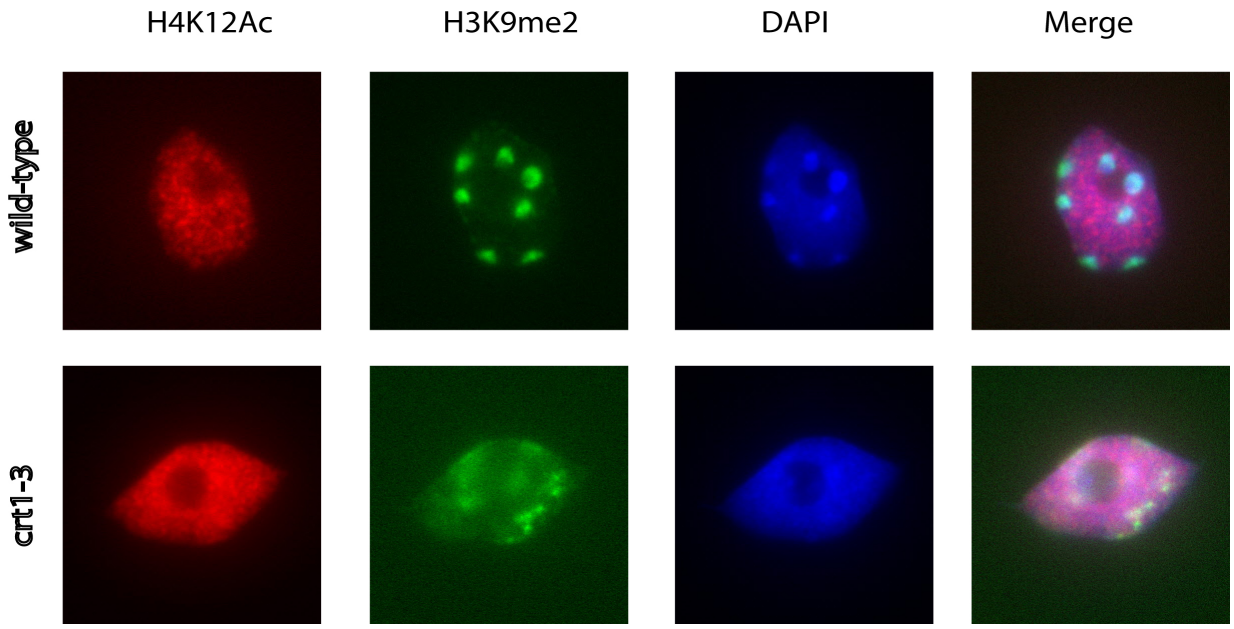


Figure 11. Images of wild-type and crt1-3 mutant nuclei, taken as in Figure 8A, with additional immunostaining using anti-H4K12Ac antibody.

FIGURE 12A

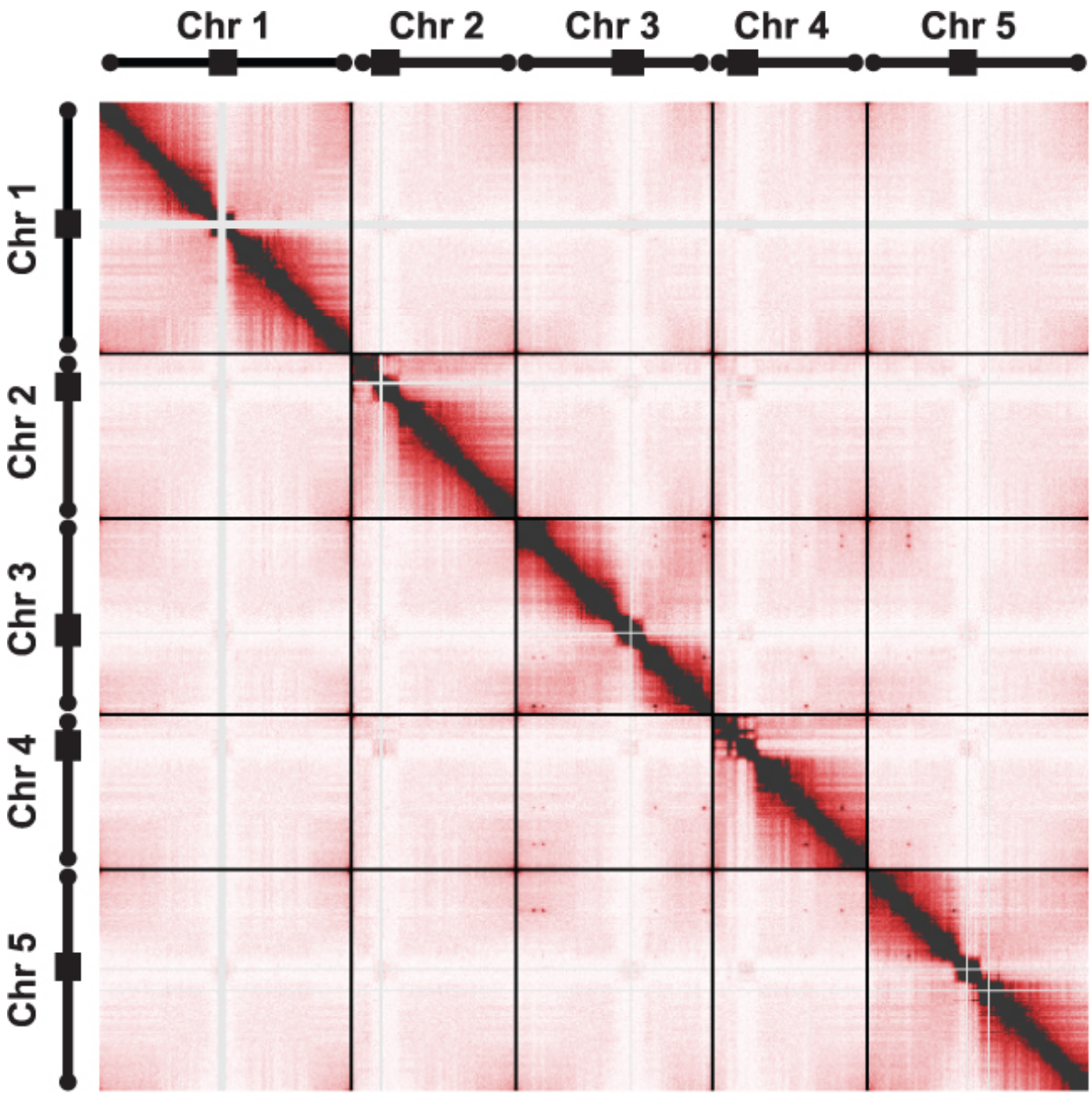


Figure 12A. Interaction matrix of the *Arabidopsis* genome in wild-type plants. Hi-C analysis of uniquely mapping paired end reads across the five chromosomes are shown with each pixel showing interactions within a 200 kilobase bin. Black bars represent the pericentromeric regions and black circles represent the telomeres. Black lines separate each chromosome and gray lines show areas with problematic mapping.

FIGURE 12B

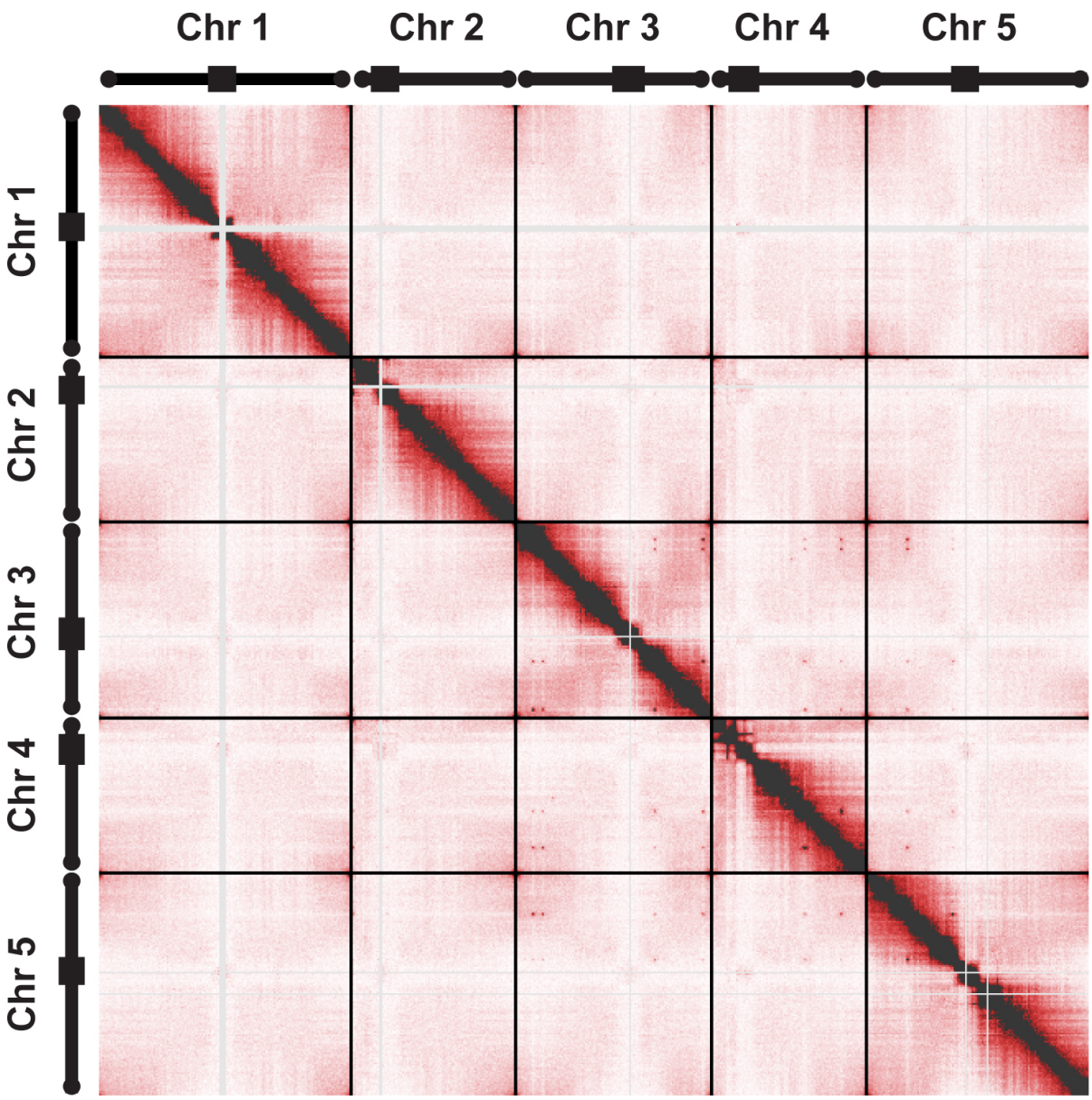


Figure 12B. Interaction matrix of the *Arabidopsis* genome in *crh6-1* mutant plants shown as in Figure 12A.

FIGURE 12C

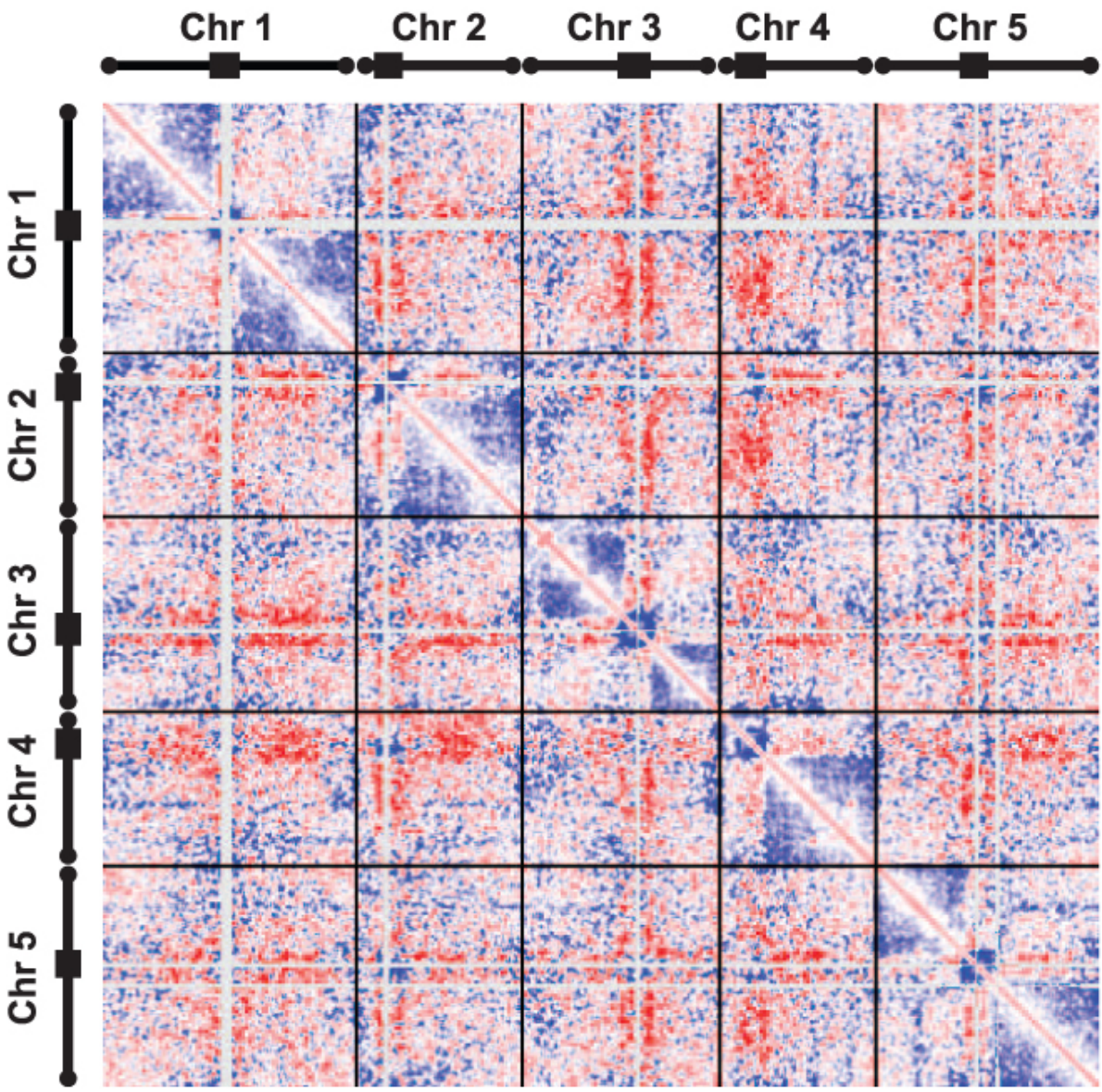


Figure 12C. Comparison of genomic interactions in *crh6* mutant and wild-type plants. Differences between the two interactions matrices are shown with red indicating enrichment and blue indicating depletion of interactions in *crh6* mutant relative to wild-type.

FIGURE 13A

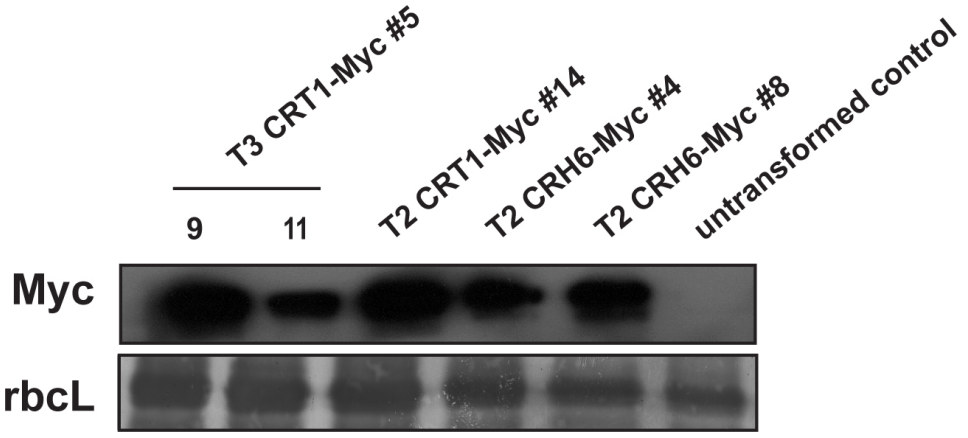


Figure 13A. Expression of pCRT1::CRT1-Myc and pCRH6::CRH6-Myc. Western blot with anti-Myc antibody confirms expression of myc-tagged transgene in transformed lines but not untransformed control.

FIGURE 13B

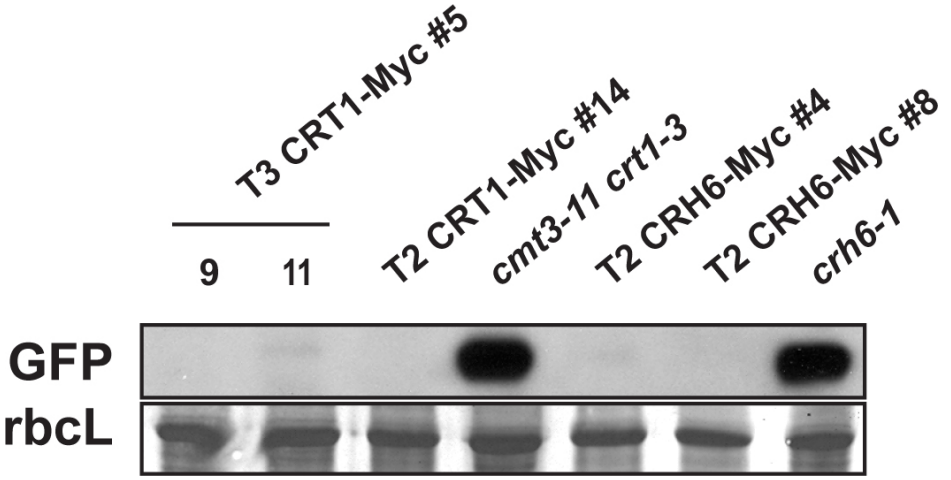


Figure 13B. Complementation test with myc-tag transgene and EMS mutant allele. Western blot with anti-GFP antibody confirms that myc-tagged proteins are functional and compelement the loss-of-function EMS alleles, thereby silencing the *SDC::GFP* transgene. *cmt3-11 crt1-3* and *crh6*-1 are GFP positive controls, and Coomassie staining of rbcL (large Rubisco subunit) is a loading control.

FIGURE 13C

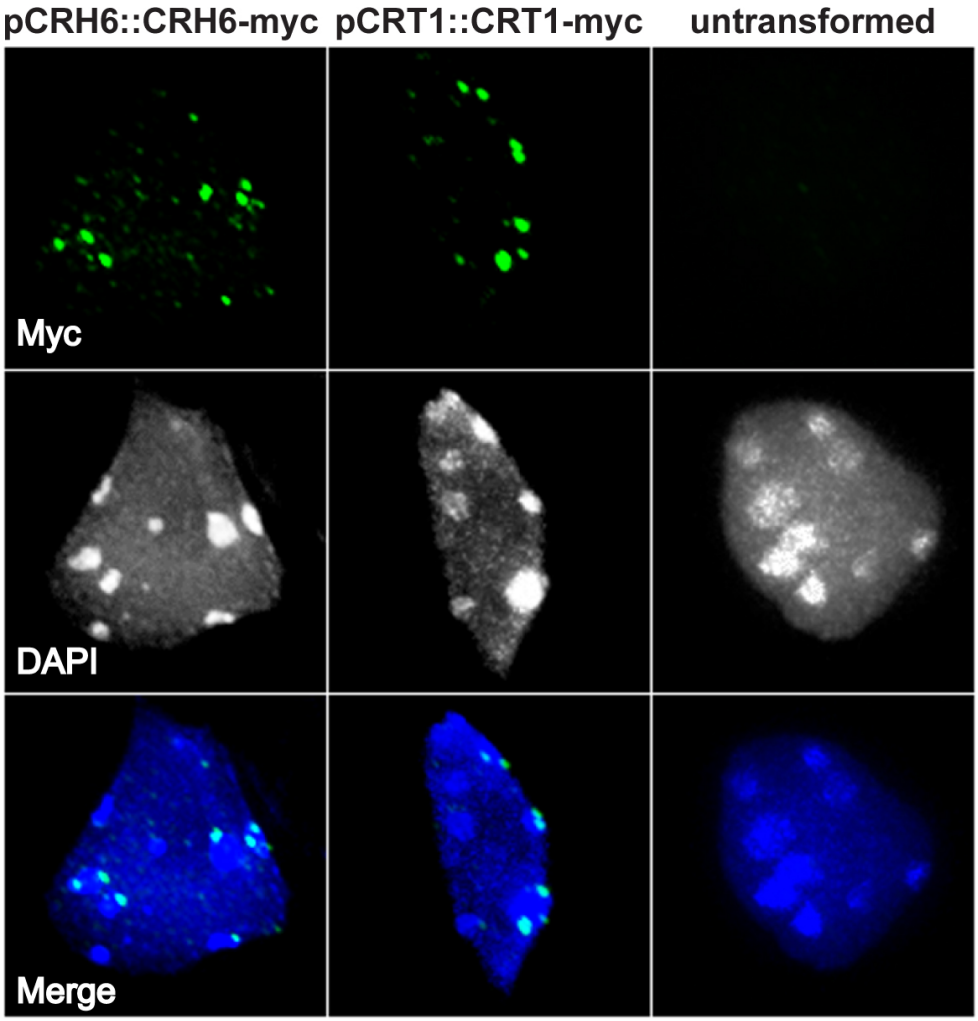
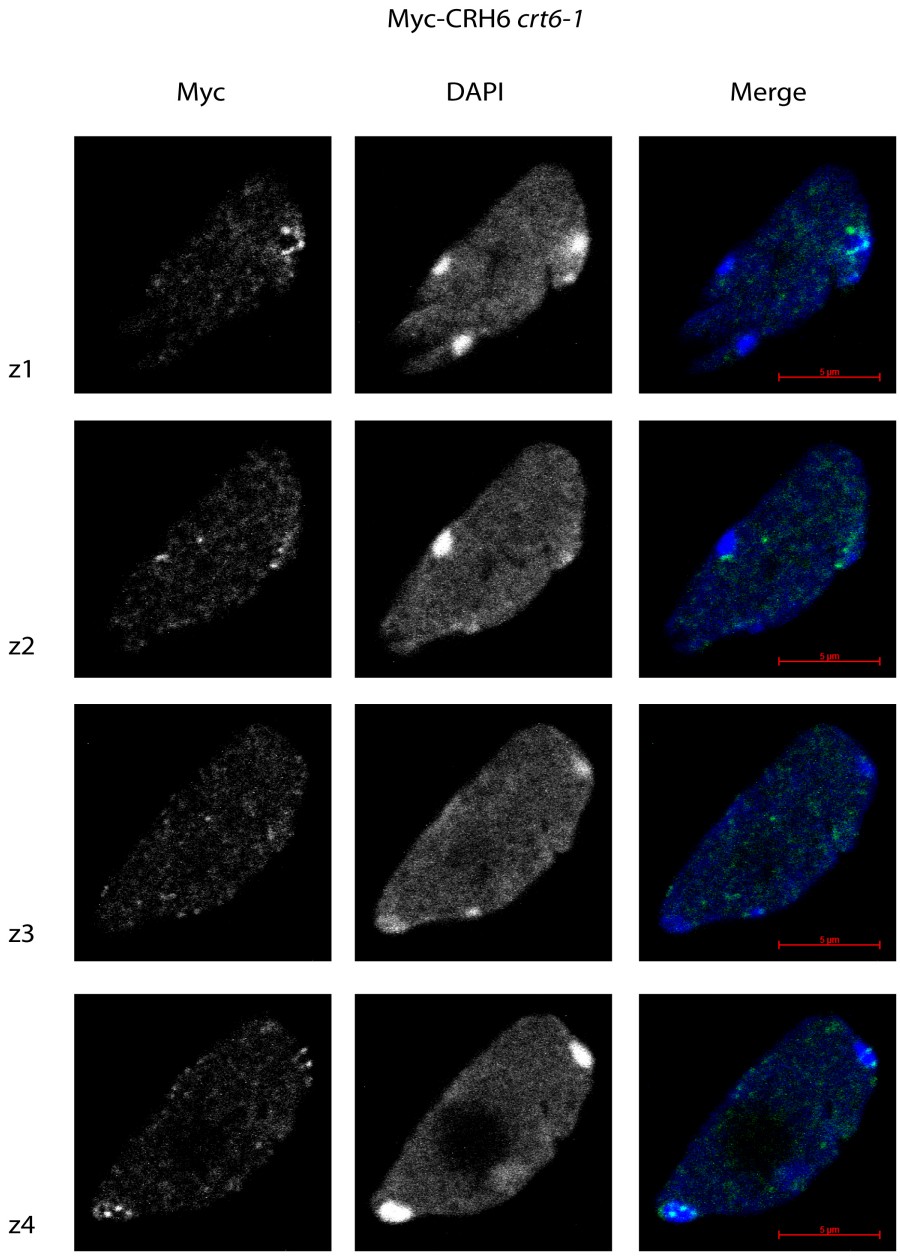


Figure 13C. Nuclear localization of myc-tagged CRT1 and CRH6. Anti-Myc immunostaining showed location of CRT1 and CRH6 proteins relative to chromocenters, shown as concentrated areas of DNA by DAPI staining.

FIGURE 13D



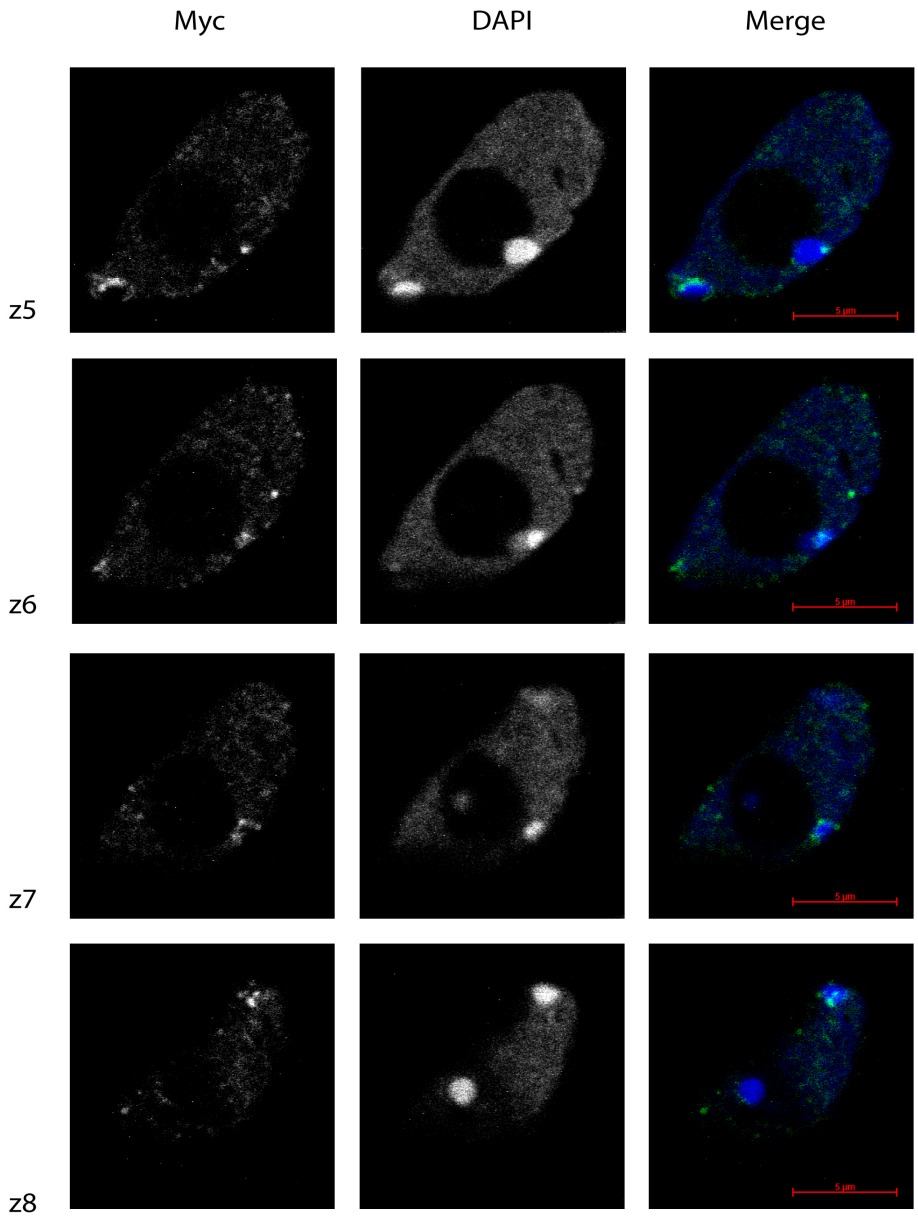


Figure 13D. Nuclear localization of myc-tagged CRH6. Nuclei were prepared as in Figure 11C and images captured with a confocal microscope. This nucleus, shown in eight images at depth intervals of 1 micron, is representative of the localization of CRH6 to the periphery of multiple chromocenters in each nucleus.

FIGURE 14

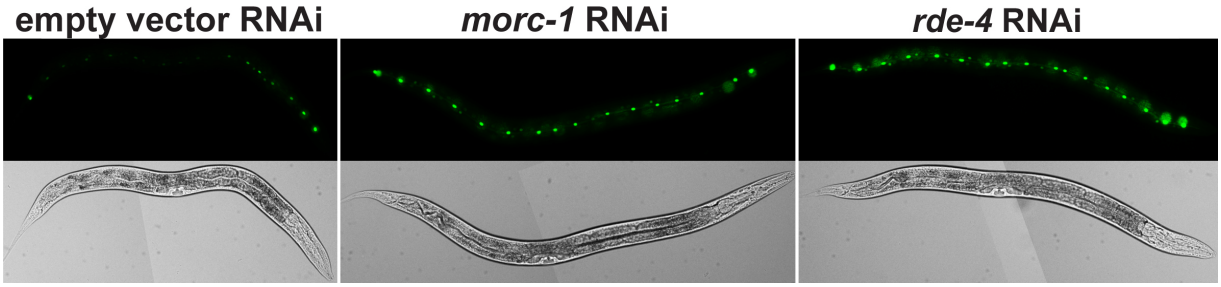


Figure 14. Loss of silencing in *morc-1* depleted *C. elegans* seam cells. Worms were fed bacteria expressing double-stranded RNA and silencing of a seam cell-specific GFP transgene was observed. Empty vector RNAi served as a control for GFP silencing and *rde-4* served as a control for GFP transgene overexpression.

MATERIALS AND METHODS

Plant material and growing conditions

All mutants are in the Columbia (Col) ecotype. Crh6-3 (GK_599B06) and crt1-4 (SAIL_1239_C08) T-DNA lines were obtained from GABI-Kat (Kleinboelting et al. 2012) 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 (Zhang et al. 2006). T-DNA insertions were confirmed by PCR-based genotyping. Primer sequences are described in Table S4 in Moissiard et al., 2012. 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 (Clough et al. 1998). 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 obtained 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 in Moissiard et al., 2012.

Cloning of pCRT1::CRT1-Myc and pCRH6::CRH6-Myc

CRT1 and CRH6 genomic regions were PCR amplified and the Myc epitope was added to the C-terminus of each protein as previously described (Law et al. 2011). 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 in Moissiard et al., 2012.

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 GPF 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 (Greenberg et al. 2011), 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 (Law et al. 2011).

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 (Langmead et al. 2009) allowing up to 2 mismatches. Gene and transposon expression was measured by calculating reads per kilobase per million mapped reads (RPKM) (Mortazavi et al. 2008). p-values were calculated using Fisher's exact test and Benjamini corrected for multiple testing (Benjamini and Hochberg 1995). Differentially expressed elements in wild type and mutants were defined by applying log2(mutant /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 in Moissiard et al., 2012) on a Stratagene Mx real-time thermocycler.

DNA methylation analyses

Whole genome BS-seq libraries were performed as previously described (Cokus et al. 2008), except directly with pre-methylated final adapters. BS-seq data was mapped with BS seeker as previously described (Chen et al. 2010). 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 (Henderson and Jacobsen 2008). Primer sequences used for bisulfite sequencing are described in Table S4 in Moissiard et al., 2012. Southern blot was performed as previously described (Jackson et al. 2002).

H3K9me2 ChIP-seq analyses

Two grams of 3-week-old seedlings were crosslinked with formaldehyde and chromatin-IP experiments were performed as previously described (Lee et al. 2007). 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 Pfpassing 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.

Hi-C analyses

Two grams of 3-week-old seedling leaves were crosslinked with formaldehyde as previously described (Lee et al. 2007). Hi-C experiments were performed as previously described (Zhang et al. 2012), with the exception that plant nuclei were prepared following a previously published Arabidopsis ChIP protocol (Lee et al. 2007). 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 crh6-1 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 (Zhang et al. 2012) using 50 iterations of that procedure. The comparison between the wild type and crh6-1

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 (Soppe et al. 2002) 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 13C). Additional images of nuclei were taken using a Zeiss LSM 510 META confocal microscope with Axiocam camera (Figure 13D).

RNA interference in C. elegans

RNAi experiments were carried out as reported previously (Kim et al. 2005) 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 (Kamath et al. 2003). 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.

REFERENCES

Agorio, A. & Vera, P. 2007. ARGONAUTE4 is Required for Resistance to *Pseudomonas syringae* in *Arabidopsis*. *Plant Cell* 19: 3778-3790.

Amedeo, P., Habu, Y., Afsar, K., Mittelsten Scheid, O. & Paszkowski, J. 2000. Disruption of the plant gene *MOM* releases transcriptional silencing of methylated genes. *Nature* 405: 203-206.

Benjamini, Y. & Hochberg, Y. 1995. Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. B* 57: 289-300.

Bernatavichute, Y. V., Zhang, X., Cokus, S., Pellegrini, M. & Jacobsen, S. E. 2008. Genome-Wide Association of Histone H3 Lysine Nine Methylation with CHG DNA Methylation in *Arabidopsis thaliana*. *PLoS ONE* 3: e3156.

Bourc'his, D. & Bestor, T. H. 2004. Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature* 431: 96-99.

Carmell, M. A., Girard, A., Van De Kant, H. J. G., Bourc'his, D., Bestor, T. H., De Rooji, D. G. & Hannon, G. J. 2007. MIWI2 Is Essential for Spermatogenesis and Repression of Transposons in the Mouse Male Germline. *Dev. Cell* 12: 503-514.

Cao, X. & Jacobsen S. E. 2002. Locus-specific control of asymmetric and CpNpG methylation by the *DRM* and *CMT3* methyltransferase genes. *Proc. Natl. Acad. Sci.* 99: 16491-16498.

Chen, P. Y., Cokus, S. J. & Pellegrini, M. 2010. BS Seeker: precise mapping for bisulfite sequencing. *BMC Bioinformatics* 11: 203.

Clough, S. J. & Bent, A. F. 1998. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J* 16: 735-743.

Cokus, S. J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C. D., Pradhan, S., Nelson, S. F., Pellegrini, M., & Jacobsen, S. E. 2008. Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature* 452: 215-219.

Dutta, R. & Inouye, M. 2000. GHKL, an emergent ATPase/kinase superfamily. *Trends Biochem. Sci. 25*: 24-28.

Fransz, P., De Jong, J. H., Lysak, M., Castiglione, M. R. & Schubert, I. 2002. Interphase chromosomes in *Arabidopsis* are organized as well defined chromocenters from which euchromatin loops emanate. *Proc. Natl. Acad. Sci.* 99: 14584-14589.

Goll, M. G. & Bestor, T. H. 2005. Eukaryotic Cytosine Methyltransferases. *Annu. Rev. Biochem.* 74: 481-514.

Greenberg, M. V. C., Ausin, I., Chan, S. W. L., Cokus, S. J., Cuperus, J. T., Feng, S., Law, J. A., Chu, C., Pellegrini, M., Carrington, J. C. & Jacobsen, S. E. 2011. Identification of genes required for de novo DNA methylation in Arabidopsis. *Epigenetics* 3: 344-354.

Habu, Y. 2010. Epigenetic silencing of endogenous repetitive sequences by MORPHEUS' MOLECULE1 in *Arabidopsis thaliana*. *Epigenetics* 5: 562-565.

Henderson, I. R. & Jacobsen, S. E. 2008. Tandem repeats upstream of the *Arabidopsis* endogene SDC recruit non-CG DNA methylation and initiate siRNA spreading. *Genes Dev*. 22: 1597-1606.

Inoue, N., Hess, K. D., Moreadith, R. W., Richardson, L. L., Handel, M. A., Watson, M. L., & Zinn, A. R. 1999. New gene family defined by MORC, a nuclear protein required for mouse spermatogenesis. *Human Mol. Genet*. 8: 1201-1207.

Iyer, L. M., Abhiman, S. & Aravind, L. 2008. MutL homologs in restriction-modification systems and the origin of eukaryotic MORC ATPases. *Biol. Direct* 3: 8.

Jackson, J. P., Lindroth, A. M., Cao, X., & Jacobsen, S. E. 2002. Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* 416: 556-560.

Jeddeloh, J. A., Stokes, T. L. & Richards, E. J. 1999. Maintenance of genomic methylation requires a SWI2/SNF2-like protein. *Nat. Genet*. 22: 94-97.

Kamath, R. S. *et al*. 2003. Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. *Nature* 421: 231-237.

Kang, H. G., Kuhl, J. C., Kachroo, P., & Klessig, D. F. 2008. CRT1, an *Arabidopsis* ATPase that Interacts with Diverse Resistance Proteins and Modulates Disease Resistance to Turnip Crinkle Virus. *Cell Host Microbe* 3: 48-57.

Kang, H. G., Oh, C. S., Sato, M., Katagiri, F., Glazebrook, J., Takahashi, H., Kachroo, P., Martin, G. B., & Klessig, D. F. 2010. Endosome-Associate CRT1 Functions Early in *Resistance* Gene-Mediated Defense Signaling in *Arabidopsis* and Tobacco. *Plant Cell* 22: 918-936.

Kennedy, S., Wang, D. & Ruvkun, G. 2004. A conserved siRNA-degrading RNase negatively regulates RNA interference in *C. elegans*. *Nature* 427: 645-649.

Kim, J. K., Gabel, H. W., Kamath, R. S., Tewari, M., Pasquinelli, A., Rual, J., Kennedy, S., Dybbs, M., Bertin, N., Kaplan, J. M., Vidal, M. & Ruvkun, G. 2005. Functional Genomic Analysis of RNA Interference in *C. elegans*. *Science* 308: 1164-1167.

Kleinboelting, N., Huep, G., Kloetgen, A., Viehoever, P., & Weisshaar, B. 2011. GABI-Kat SimpleSearch: new features of the *Arabidopsis thaliana* T-DNA mutant database. *Nucleic Acids Res*. 40: 1211-1215.

Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 10: R25.

Law, J. A. & Jacobsen, S. E. 2010. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat. Rev. Genet*. 11: 204-220.

Law, J. A., Vashisht, A. A., Wohlschlegel, J. A. & Jacobsen, S. E. 2011. SHH1, a Homeodomain Protein Required for DNA Methylation, As Well As RDR2, RDM4, and Chromatin Remodeling Factors, Associate with RNA Polymerase IV. *PLoS Genet*. 7: e1002195.

Lee, J., He, K., Stolc, V., Lee, H., Figueroa, P., Gao, Y., Tongprasit, W., Zhao, H., Lee, I. & Deng, X. W. 2007. Analysis of Transcription Factor HY5 Genomic Binding Sites Revealed Its Hierarchical Role in Light Regulation of Development. *Plant Cell* 19: 731-749.

Lei, H., Oh, S. P., Okano, M., Juttermann, R., Goss, K. A., Jaenisch, R. & Li, E. 1996. De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development* 122: 3195-3205.

Lopez, A., Ramirez, V., Garcia-Andrade, J., Flors, V. & Vera, P. 2011. The RNA Silencing Enzyme RNA Polymerase V is Required for Plant Immunity. *PLoS Genet*. 7: e1002434.

Lieberman-Aiden, E. *et al.* 2009. Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. *Science* 326: 289-293.

Malagnac, F., Bartee, L., & Bender, J. 2002. An *Arabidopsis* SET domain protein required for maintenance but not establishment of DNA methyation. *EMBO J*. 21: 6842-6852.

Moissiard, G., Cokus, S. J., Cary, J., Feng, S., Billi, A. C., Stroud, H., Husmann, D., Zhan, Y., Lajoie, B. R., McCord, R. P., Hale, C. J., Feng, W., Michaels, S. D., Frand, A. R., Pellegrini, M., Dekker, J., Kim, J. K. & Jacobsen, S. E. 2012. MORC Family ATPases Required for Heterochromatin Condensation and Gene Silencing. *Science* 336: 1448-1451.

Mortazavi, A., Williams, B. A., McCue, K., Shaeffer, L. & Wold, B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* 5: 621-628.

Okano, M., Bell, D. W., Haber, D. A. & Li, E. 1999. DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development. *Cell* 99: 247-257.

Probst, A. V., Fransz, P. F., Paszkowski, J. & Mittelsten Scheid, O. 2003. Two means of transcriptional reactivation within heterochromatin. *Plant J*. 33: 743-749.

Simpson, V. J., Johnson, T. E. & Hammen, R. F. 1986. *Caenorhabditis elegans* DNA does not contain 5-methylcytosine at any time during development or aging. *Nucleic Acids Res*. 14: 6711-6719.

Soppe, W. J. J., Jasencakova, Z., Houben. A., Kakutani, T., Meister, A., Huang, M. S., Jacobsen, S. E., Schubert, I. & Fransz, P. F. 2002. DNA methylation controls histone H3 lysine 9 methylation and heterochromatin assembly in *Arabidopsis*. *EMBO J*. 21: 6549-6559.

Tabara, H., Sarkissian, M., Kelly, W. G., Fleenor, J., Grishok, A., Timmons, L., Fire, A. & Mello, C. C. 1999. The *rde-1* Gene, RNA Interference, and Transposon Silencing in *C elegans*. *Cell* 99: 123-132.

Wang, Z., Gerstein, M., & Snyder, M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet*. 10: 57-63.

Watson, M. L., Zinn, A. R., Inoue, N., Hess, K. D., Cobb, J., Handel, M. A., Halaban, R., Duchene, C. C., Albright, G. M., & Moreadith, R. W. 1998. Identification of *morc* (*microrchidia*), a mutation that results in arrest of spermatogenesis at an early meiotic stage in the mouse. *Proc. Natl. Acad. Sci.* 95: 14361-14366.

Zhang, X., Yazaki, J., Sundaresan, A., Cokus, S., Chan, S. W., Chen, H., Henderson, I. R., Shinn, P., Pellegrini, M., Jacobsen, S. E., & Ecker, J. R. 2006. Genome-wide High-Resolution Mapping and Functional Analysis of DNA Methylation in *Arabidopsis*. *Cell* 126: 1189-1201.

Zhang, Y., McCord, R. P., Ho, Y. J., Lajoie, B. R., Hildebrand, D. G., Simon, A. C., Becker, M. S., Alt, F. W. & Dekker, J. 2012. Spatial Organization of the Mouse Genome and Its Role in Recurrent Chromosomal Translocations. *Cell* 148: 908-921.