Differential Regulation of Strand-Specific Transcripts from Arabidopsis Centromeric Satellite Repeats

Bruce P. May, Zachary B. Lippman, Yuda Fang, David L. Spector, Robert A. Martienssen

Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, United States of America

Introduction

The centromeric regions of animal and plant chromosomes are large assemblies of short (approximately 151 to 340 base pairs [bp]) satellite repeats in head-to-tail orientation with interspersed retroelements. In Arabidopsis thaliana, these comprise 177- to 179-bp satellite repeats (cen180, also known as pAL1 and AtCon; Figure 1A), Athila LTR-retroelements, and 106B repeats, which are 398-bp internal portions of Athila2 LTRs [1–3]. The only noted common feature among eukaryotic alpha satellites is a binding site for CENP-B, an evolutionary relative of pogo transposase that is necessary for de novo centromere formation [4,5] but not for the centromeres of the human Y chromosome and some marker chromosomes that lack CENP-B [6,7]. In G2 of the cell cycle, a modified histone H3, CENP-A in humans and CenH3 (HTR12) in Arabidopsis, is incorporated into centromeric nucleosomes independently of DNA replication [8–10]. A complex of proteins, some of which are directly recruited by CENP-A, then assembles to form the kinetochore that moves the mitotic chromatid or meiotic univalent poleward along depolymerizing microtubules during anaphase.

Although a specific binding site for CENP-A has been hypothesized, constitutive expression of CENP-A results in ectopic deposition at sites throughout the genome [11] and CENP-A has been observed to spread to formerly euchromatic regions lacking repeats during neocentromere creation [12,13]. Transgenic arrays of human centromeric repeats did not attract CENP-A until the cells were treated with an inhibitor of histone deacetylase [14] and yeast Cse4 can functionally substitute for CENP-A in human cells [15]. Thus, deposition of CENP-A may be guided by epigenetic features other than DNA sequence [16,17]. Centromeres appear cytologically as central constrictions flanked by conspicuous pericentromeric heterochromatin. Pericentromeric heterochromatin appears transcriptionally silent: it is depleted of histone H3 methylated at lysine 9 (H3K9me2) as well as histone acetylation, enriched for histone H3 methylated at lysine 9 (H3K9me2), and in mammals, filamentous fungi, and plants, it is enriched for 5-methylcytosine incorporated into the DNA [18–21]. In Schizosaccharomyces pombe, which lacks..
DNA methylation, the central region mediates attachment to the kinetochore while the heterochromatic outer regions attract H3K9me2, Swi6, and cohesin [22]. Unexpectedly, the heterochromatic repeats are transcribed and subject to RNA interference (RNAi) [23]. RNAi guides histone modification via an interaction between complexes containing small interfering RNAs (siRNAs) and nascent transcripts at the locus [24–26]. H3K9me2 in turn binds a repressor, Swi6 in yeast and heterochromatic protein 1 (HP1) in plants and animals, that propagates a more densely coiled, transcriptionally silenced chromatin structure [22]. In fission yeast, Swi6 interacts with the cohesin complex [27] so that the outer regions remain associated longer during mitosis than the central region or the chromosome arms [28]. Lagging chromosomes are observed at high frequency (20%) in mutants defective in heterochromatin formation such as clr4 (H3K9me2 methyltransferase), swi6 (HP1 homolog), rad21/cohesin, and RNAi [24,29].

This connection between RNAi and centromeric silencing has since been extended to mammalian, insect, and avian cells, but it is not known if the interaction with cohesin is common to other organisms [30–32].

DNA methylation is found in many eukaryotes (although not extensively, if at all, in S. pombe, Drosophila, or Caenorhabditis elegans) and interacts with histone deacetylation and H3K9me2, but the order of steps is unclear: In Arabidopsis, loss of H3K9me2 is accompanied by loss of DNA methylation [36–39] and loss of DNA methylation is accompanied by loss of H3K9me2 [39–41], although the latter effect may not be direct [20]. The DNA methyltransferase CHROMOMETHYLASE3 (CMT3), the histone H3 lysine-9 methyltransferase KRYPTONITE (KYP), and the RNAi component AGO4 maintain hypermethylation induced by transcription of long inverted repeats [37,42–45]. In contrast, DNA METHYLTRANSFERASE1 (MET1) and HISTONE DEACETYLASE6 (HDAC6) are required to maintain silencing of promoters induced by transcription of short inverted repeats [46]. Despite the presence of siRNA, other components of the siRNA pathway have not been recovered in these screens, and AGO4 is not required for silencing mediated in trans by short inverted repeats [45]. But AGO4 is required, along with DICER-LIKE3 (DCL3) and RNA-DEPENDENT RNA POLYMERASE2 (RDR2), for DRM1- and DRM2-dependent DNA methylation and silencing of transgenes when they are introduced via agrobacterium-mediated transformation [47].
Similarly, two pathways have been identified for transposon silencing in *Arabidopsis*. One involves CMT3, KYP, and AGO1, and the other involves MET1, DDM1, and HDA6 [39]. In this second pathway, resilienting of some but not other transposons in backcrosses suggests a role for siRNA in the establishment of silencing [39]. For most transposons, a role for siRNA in the maintenance of silencing is not apparent in *dcl3* mutants, although other dicers can compensate for the loss of *DCL3* [48] so that such a role cannot be ruled out.

We have found that centromeric satellite repeats in *Arabidopsis* are transcribed strand-specifically and that the transcripts remain associated in the nucleus. Only a subset of 180-bp repeats is transcribed, but the repeats are partially silenced by an RNAi-based system including *DCL1, AGO1, CMT3, KYP,* and *HDA6* [39]. Loss of silencing is inherited epigenetically in backcrosses. Other 180-bp repeats are silent in wild-type (WT) but transcribed in *met1, ddm1,* and *hda6*. At the same time, *athila2* LTR retroelements direct the transcription of these satellite repeats, and silencing is restored in *met1* and *hda6* backcrosses. Consistent with different patterns of retroelement insertion, different repeats are transcribed in the Landsberg and Columbia strains of *Arabidopsis,* providing a possible mechanism for the rapid evolution of centromeric satellite repeats.

Results

Centromeric Repeats Are Transcribed and Remain in the Nucleus

*Arabidopsis* cen180 repeat families are 78% to 96% identical in sequence. Variant repeats are located on different chromosomes [49] and may, like human alpha satellites, form subdomains [50], but each repeat contains common conserved regions [2]. We designed cen180 Fc and Rc primers (Figure 1A) to recognize two conserved regions, and these primers amplified a large number of repeats from genomic DNA. Transcripts from centromeric repeats were detected by RT followed by PCR. “F-strand transcripts” are transcribed from the forward (Watson) DNA strand and reverse transcribed into cDNA using an F primer; “R-strand transcripts” are transcribed from the reverse (Crick) DNA strand and reverse transcribed into cDNA using an R primer. Using the highly conserved Fc and Rc primers, transcripts from both strands could be detected (Figure 1B). However, transcripts from each strand could be derived from different subfamilies of repeats. In order to test this possibility, specific primers were used (F primer and R primer, Figure 1A) to amplify only a subset of cen180 repeats [20]. In this case, we found that F-strand transcripts were much more abundant than R-strand transcripts in vegetative and floral tissues, suggesting strand-specific transcription of individual subfamilies of repeats (Figure 1B). Transcription from only one strand was consistent with transcripts initiating in tandem orientation from the repeats, while the amplification of 180-bp multimers indicated that transcripts did not terminate within each repeat. Transcription levels were highest in young seedlings and developing inflorescences, indicating expression in dividing cells.

The same subfamily of cen180 repeats were amplified from genomic DNA and used as strand-specific probes for in situ hybridization (see Materials and Methods). In longitudinal sections of WT inflorescences, F-strand transcripts were more abundant than R-strand transcripts (Figure 2), consistent with RT-PCR (Figure 1B). Young, proliferating tissues (Figure 2A and 2B) had more signal than mature tissue (Figure 2C and 2D). Strand and tissue specificity indicated that the signal came from cellular RNA and was not due to background hybridization with fortuitously denatured DNA. Under higher magnification, small punctate nuclear signals were seen with the F-strand probe, consistent with nascent transcripts remaining at the centromere (Figure 2E). In contrast, what little signal could be detected with the R-strand probe was nuclear but not punctate (Figure 2B and 2D).

Satellite Repeat Transcripts Are Regulated by DNA Methylation, Histone Modification, and RNAi

In *S. pombe*, *H3K9me2* depends on RNA silencing of pericentromeric repeats [23], and vice versa [51]. Similarly, in *Arabidopsis*, *H3K9me2*, RNAi, and DNA methylations are mutually interdependent [39,51]. We examined the level of pericentromeric transcripts in *Arabidopsis* mutants defective in DNA and histone methylation as well as in RNAi (Figure 3A). Elevated transcript levels were detected in *del1*–9 (a weak allele of *DICER-LIKE1*), *ago1*–9 (a strong allele of *ARGONAUTE1*), *rdr2*–1 (a T-DNA insertion in *RNA POLYMERASE2*), *del5*–1 (a T-DNA insertion in *DICER-LIKE3*), *kyp2* (a splicing mutation in histone H3 K9 methyltransferase), *hda6/sil1* (an allele of *HISTONE DEACETYLASE6*), *cm3-m5662* (a null allele of the CNG DNA methyltransferase), *met1–1* (a strong allele of the CG DNA methyltransferase), and *ddm1–2* (a hypomorphic allele of the swi5/m Chromatin Remodeling ATPase). Strand-specific RT-PCR (Figure 3A) indicated that F-strand transcripts were elevated in abundance in *cm3, kyp, hda6,* and the RNAi mutants, whereas both F- and R-strand transcripts accumulated in *met1* and *ddm1*. In backcrosses to WT, elevated transcript levels were inherited epigenetically from *cm3, ddm1,* and *kyp*. Reestablishment of silencing was observed for both strands in *met1* backcrosses and on one strand in *hda6* plus backcrosses. *del1*–9 and *ago1*–9 were not tested in this way due to sterility [39]. Northern analysis of *met1* and *del1* mutants revealed heterogeneous transcripts ranging from approximately 0.1 kb to more than 5 kb, indicating multiple, irregular transcription initiation and/or termination sites (Figure S1).
Satellite Repeat Subfamilies Are Silenced by Different Mechanisms

One explanation for this differential regulation was that further subfamilies of F + R repeats were differentially regulated on each strand. We therefore examined which particular repeats were expressed in met1 compared with dcl1, ago1, and WT using restriction and sequence analysis of cDNA. Digestion of PCR products from cDNA and genomic DNA with TaqI revealed that transcribed repeats, which were sensitive to TaqI digestion, were a minority of those amplified from genomic DNA, most of which were not sensitive (Figure 3B). Further, RT-PCR products from met1 were digested by Sau3AI, indicating they had Sau3AI sites, but cDNAs from dcl1 were not digested and must therefore differ in sequence. Transcribed repeats were sequenced and grouped by cDNA sequence similarity using CLUSTALW. Whereas those from WT, dcl1, and ago1 formed mixed clusters, transcripts from met1 were from a distinct subfamily of repeats (Figure S2).

Diagnostic 20-mers were designed from the most highly diverged satellite cDNAs (Figure 3C) and hybridized to strand-specific RT-PCR products from the various mutants (Figure 3D). Hybridization was exclusive and complementary. F-strand transcripts from the subfamily of repeats expressed in WT were elevated in cmt3, kyp, hda6, dcl1, and ago1, but transcripts from the R strand were not seen. Elevated transcript levels were inherited epigenetically when kyp, hda6, and cmt3 mutants were backcrossed to WT. The subfamily of repeats expressed in met1 was transcribed from both strands in met1 and hda6 but only from the F strand in ddm1. These transcripts could not be detected in WT and were epigenetically inherited in hda6/+. In this respect, they resemble ATGP1 gypsy-class LTR retrotransposons [39,41,52]. Additional R-strand transcripts, detected by RT-PCR in ddm1, did not hybridize with either probe, indicating a third subfamily of differentially regulated repeats.

The differentially regulated subfamilies of cen180 repeats are very similar, so sequence differences are unlikely to account for differences in regulation. A more likely explanation was that some cen180 transcripts might be driven by read-through from adjacent retroelements. The 106B LTR-like repeats and Athila-class LTR retroelements are integrated throughout pericentromeric regions [53] and are interspersed in random orientation with respect to cen180s (data not shown). F-strand 106B transcripts were detected in WT but were only weakly up-regulated in cmt3, kyp, and hda6 (Figure 3A). Transcripts from both strands were up-regulated...
in *ddm1* and *met1*, and R-strand transcripts were up-regulated in *hda6*. In this respect, 106B repeats resembled both classes of cen180 repeats, and we examined the possibility they were co-transcribed using cDNA that was reverse transcribed with 106B or cen180 primers (Figure 3E). F-strand co-transcripts originating in cen180 repeat arrays could be detected in WT and *ddl1*. In contrast, reverse strand co-transcripts appeared to originate within 106B and could only be detected in *met1*. Subsequent RACE PCR showed that these co-transcripts originated in the LTR of *Athila2* itself (not shown). Other co-transcripts were not seen. This indicated that differential genetic regulation of each class of cen180 repeat transcript could be explained by differential origin of the transcripts in 106B/*Athila2* and cen180 repeats, respectively.

**Polymorphic Regulation of Repeats**

Most of the mutants described above were isolated in the Landsberg genetic background, but *vd* and *ddl3* were isolated in Columbia, and so this WT strain was also assayed for transcripts. Surprisingly, Columbia and Landsberg transcripts amplified with conserved primers *Fc* and *Rc* differed in their regulation (Figures 3A and 4). We therefore sequenced cen180 RT-PCR products from *ddl3* and *met1* mutants of Columbia and designed diagnostic primers for amplification of genomic DNA and cDNA (Figure 4). Primers corresponding to *ddl3*-and *met1*-specific transcripts amplified cDNA from each corresponding mutant in Columbia, as expected. However, no cDNA could be amplified from *met1* or * ddl3* mutants in Landsberg despite the presence of such repeats in the Landsberg genome (indicated by amplification of genomic DNA). Conversely, specific transcripts detected in *ddl3* mutants in Landsberg were found to accumulate in *met1* mutants of Columbia, indicating that the regulation of this class of repeats had changed during the divergence of these two ecotypes. Repeats transcribed specifically in *met1* mutants of Landsberg do not appear to be present in the Columbia genome. The origin of this natural variation is discussed below.

**Histone Modifications Associated with Satellite Repeats**

We attempted to examine histone methylation associated with centromeric repeats using semiquantitative chromatin immunoprecipitation, although this proved to be very difficult as previously reported [20,40] because of the very high copy numbers involved. Nonetheless, while quantitative changes were not reliable, qualitative changes could be detected (Figure S3). H3K9me2 was substantially reduced in *kyp* and *ago1* over both classes of repeat, resembling transposons such as ATCOPIA4 [39] as well as pericentromeric repeats and retrotransposons in *S. pombe* and *Drosophila* [23,31]. In *met1* and *ddl1*, H3K9me2 was largely lost from 106B but only slightly reduced at cen180. H3K4me2 underwent a modest but heritable increase in *ddl1* and was found associated with each class of repeat in WT, perhaps reflecting the transcription of the repeats. The association of H3K4me2 with heterochromatic pericentromeric repeats has also been reported in *S. pombe* [23], while in *Drosophila* [33] it is interspersed with Cid in more central domains.

DNA methylation of centromeric repeats has been extensively investigated by gel blot analysis, and we have obtained similar results (Figure S4). CNG methylation is partial in WT and reduced in *ddl1*, *cmt3*, and *kyp* [21,37,42,43], but it is unaffected in *hda6* [54]. CG methylation is more extensive than CNG methylation at the centromere in WT and is lost only in *met1* and *ddl1*. However, *met1* affects transcripts initiating in 106B/*Athila2* LTRs and not those initiating within the class of cen180 repeats regulated by RNAi (Figure 3E). *hda6* has a similar effect on 106B transcripts but has no effect on CG methylation, although H3K9me2 is reduced ([54] and

---

**Figure 4. Polymorphic Regulation of cen180 Transcripts**

Primers were designed to recognize sample cen180 cDNAs from *ddl3* and *met1* mutants of Landsberg erecta (see Figure 3C) and *ddl3* and *met1* mutants of Columbia. PCR amplification using genomic DNA templates (DNA) tested for the presence of the particular repeats in the genome of each ecotype. RT-PCRs using RNA detected transcripts in WT (wt) and mutants of each ecotype as indicated. The cen180 Fc + Rc primers served as positive controls in the presence of reverse transcriptase and negative controls in its absence. RNA was prepared from aerial tissues from 28-day-old plants.

DOI: 10.1371/journal.pgen.0010079.g004
Lagging Chromosomes Were Not Observed in Mutants Defective in Silencing

In the fission yeast *S. pombe*, mutants defective in centromeric silencing recruit, but fail to retain, cohesin in anaphase pericentromeric heterochromatin, because of the loss of H3K9me2 and associated Swi6 [61]. We therefore examined anaphase in centromeric silencing mutants in *Arabidopsis*. Developing floral or root tissue was fixed in paraformaldehyde and stained with DAPI to detect the presence of lagging or otherwise abnormal chromosomes during anaphase. In WT anaphase cells (N = 20), rdr2 mutant cells (N = 49), and *ddm1* mutant cells (N = 23), no abnormality was seen (not shown). Further, these mutants were indistinguishable from WT in growth and fertility, unlike mutants in *rad21/cohesin*, which are semisterile due to defects in meiosis [34]. One explanation might be that H3K9me2 was largely retained in *rad2* and *ddm1*, so that any interaction with Swi6 homologs and cohesin was retained. However, mutants in the H3K9 methyltransferase *kyp* are also fully fertile, and lagging chromosomes were not observed in *kyp* anaphase cells either (N = 24). Differences between the yeast and plant systems are discussed below.

Discussion

Silencing of Centromeric Transcripts in Arabidopsis and Fission Yeast

*Arabidopsis* centromeric repeats are transcribed and, for at least one subfamily, transcripts from one strand accumulate predominantly and are regulated posttranscriptionally by RNAi, similar to the situation in *S. pombe*. This is consistent with the tandem orientation of these large blocks of repeats: putative promoter sequences found in each repeat would be expected to align in the same orientation as each other (Figure 6). Recently, each strand of the pericentromeric repeats in *Arabidopsis* was found to differ in DNA methylation [62]. While certainly an intriguing result, these differentially methylated sequences were located outside the satellite repeats and mostly composed of retrotransposons integrated in both orientations. Therefore, the significance of this methylation for strand-specific transcription of the satellite repeats is not yet clear.

Even though they are transcribed from only one strand, tandem repeats can theoretically generate siRNA by reiterated rounds of RdRP replication and Dicer degradation [63]. The requirement for both *DCL1/AGO1* and *RDR2/DCL3* indicates both 21- and 24-nucleotide siRNAs may be involved. Although we could barely detect 21-nucleotide siRNA on blots, rare centromeric siRNA of this size has been detected by sequencing [64]. For other repeat subfamilies, both strands were silent, but transcription could be detected in mutants defective in the DNA methyltransferase *MET1* and the type I histone deacetylase HDA6. These transcripts arose by read-out from LTR retrotransposons inserted among the repeats (see Figure 3E).

In fission yeast, RNAi is required to maintain normal levels of H3K9me2 at the outer pericentromeric repeats [23] and mutants have lagging chromosomes at anaphase [24,29]. A similar repeat is responsible for aspects of silencing at the mating type locus, but additional *cis*-acting elements are required for silencing in the absence of RNAi, requiring histone deacetylation instead [65]. In *Arabidopsis*, *dcl3* and *rdr2*...
mutants had little or no 24-nucleotide centromeric satellite siRNAs, yet had no detectable defects in H3K9me2 accumulation or in centromere function. This is also evidence of redundant mechanisms for maintaining H3K9me2, one RNAi based and the other depending on CG methylation and histone deacetylation of retrotransposons. However, kyp mutants lost most H3K9me2 without anaphase defects, although other H3K9 methyltransferases may be redundant with KYP [19], just as other Dicers are redundant with DCL3 [48]. Nevertheless, we cannot exclude the possibility that heterochromatic association of cohesin via H3K9me2 is not required for mitotic chromosome disjunction in plants [34], which lack centrosomes and may differ from animals in this respect.

If RNAi at the centromere is dispensable but transcripts are still found, this raises the possibility that the transcripts themselves may have some function. Centromeric transcripts from CRM retrotransposons and CentC satellite repeats are associated with immunoprecipitated kinetochores in maize, although neither full-length transcripts nor siRNAs were detected on Northern blots, so the origin and fate of these transcripts are unknown [66]. Similarly, major and minor satellite repeats are transcribed from mouse centromeres, but in this case RNA interference is thought to play a role [30,67]. Our results indicate that centromeric transcripts can be long, persist in the nucleus, and show some correlation with mitotic activity and may associate with chromosomes.

Whether the transcripts serve to recruit factors in addition to the RNAi apparatus remains to be seen.

The role of DDM1 in siRNA production is unclear, but it has been shown that DDM1 is required for mitotic chromosome disjunction in plants [34], although neither full-length transcripts nor siRNAs were detected. DDM1 is also involved in histone modification and DNA replication, and DCL3 is required for RNAi at the centromere [39]. HDA6 may act downstream of DDM1 as CG methylation is unaffected in hda6 mutants [39]. RNAi from cen180 repeats was also retained in hda6, but it was reduced in ddm1 and F-strand silencing was not restored in ddm1. These results are consistent with a silencing complex, composed of DDM1, MET1, and HDA6, while de novo histone modification is guided by siRNA [39]. HDA6 may act downstream of MET1 as CG methylation is unaffected in hda6 mutants [39]. RNAi from cen180 repeats was also retained in hda6, but it was reduced in ddm1 and F-strand silencing was not restored in ddm1. These results are consistent with a silencing complex, composed of DDM1, MET1, and HDA6, while de novo histone modification is guided by siRNA [39]. HDA6 may act downstream of MET1 as CG methylation is unaffected in hda6 mutants [39]. RNAi from cen180 repeats was also retained in hda6, but it was reduced in ddm1 and F-strand silencing was not restored in ddm1. These results are consistent with a silencing complex, composed of DDM1, MET1, and HDA6, while de novo histone modification is guided by siRNA [39]. HDA6 may act downstream of MET1 as CG methylation is unaffected in hda6 mutants [39].

This could account for the difference between 106B and cen180 siRNA accumulation in ddm1. If siRNA processing was reduced but 106B was transcriptionally up-regulated, this could lead to an overall increase in siRNA in ddm1. In contrast, if cen180 repeats were not transcriptionally up-regulated, siRNA would be decreased in ddm1. Combined losses of siRNA and H3K9me2 are most severe in ddm1, and chromocenters are severely disrupted [69], indicating they may play a role in heterochromatin association even though lagging chromosomes were not observed.

**Maintenance and Reestablishment of Silencing**

Silencing of centromeric repeats was restored in met1+/− following backcrosses to WT, as was methylation [70]. However, transgene silencing and methylation were not restored in similar backcrosses, although allelic differences could be responsible [52]. Transposon methylation could also be restored in met1+/− but only for those transposons that retained H3K9me2 and siRNA in met1 [39]. RNAi from cen180 repeats was also retained in met1 and hda6, but it was reduced in ddm1 and F-strand silencing was not restored in ddm1. However, transgene silencing and methylation were not restored in similar backcrosses, although allelic differences could be responsible [52].

**Evolution of Centromeres**

Centromeres are dynamic components of genome evolution. In sequenced arrays of human 171-bp satellites, the central repeats appear to be the youngest, with progressively older repeats on the outside. Formation and homogenization of new repeats are hypothesized to occur via unequal recombination, so that one repetitive sequence comes to dominate the centromere [50]. CENP-A has been proposed to co-evolve with such repeats, although binding to new, homogenous repeats has not been directly tested [72]. Transcribed repeats could also be replicated via RT and retrotransposition. New copies, being identical in sequence, would still be recognized by siRNA. As repeats age, retrotransposon insertions would recruit MET1 and DDM1, silencing the repeats transcriptionally and allowing them to diverge (Figure 6). This model predicts that younger RNAi-regulated repeats should be toward the center of the centromere and older MET1-regulated repeats should be on the flanks. In support, BLAST analysis showed that MET1-regulated repeats do indeed predominate in the sequenced portion of the genome, i.e., the exterior regions of the centromere (Figure S5). The regulation of subfamilies of centromeric repeats has diverged in the Columbia and Landsberg ecotypes, which are closely related [73]. Although
the centromeric heterochromatin has not been completely sequenced in either strain, the pattern of retrotransposon insertions elsewhere in the genome is quite different [74], and this could account for the differential silencing of individual subfamilies of centromeric repeats in each strain (Figure 6). The transcription of centromeric repeats provides an attractive model for speciation. In wide crosses, paternal chromosomes might be destabilized if the pericentromeric repeats no longer match the sequence of maternal siRNA. This mechanism could contribute to hybrid incompatibility in polyploids as well as the loss of paternal chromosomes in wide crosses [73,76].

Materials and Methods

**Plant material.** cmi3-m5662, met1–1 (E. Richards), ddm1–2 (E. Richards), klp2 (S. Jacobsen), hda6/sil1 (I. Furner), del1–9 (S. Jacobsen), agol–9–9 mutations were introgressed into the Landsberg erecta ecotype. del13–1 (J. Carrington) and rde2–1 (J. Carrington) are in the Columbia ecotype and were obtained from the Syngenta Arabidopsis Insertion Library (SAIL) collection of T-DNA insertions (Syngenta Biotechnology Inc. [SBI], Research Triangle Park, North Carolina, United States). Plants were grown 28 days (16 h light, 8 h dark) for RNA and chromatin preparation. Sources for seed stocks are indicated in parentheses.

**Transcript analysis.** RNA was prepared, reverse transcribed into DNA, and amplified as described [39]. RT-PCRs were performed with 100 ng of RNA per reaction using the OneStep kit from Qiagen (Valencia, California, United States) according to the manufacturer’s protocol. Negative controls to detect contaminating DNA were performed on the RNA preparations using cen180 F + R or cen180 Fc + Rc primers but no RT. The following reaction conditions were used in each cycle: 94 °C for 20 s, 60 °C for 30 s, and 72 °C for 1 min. Southern hybridization was performed using standard methods [77].

**Primer sequences** were selected from conserved and variable regions of cen180 repeats, also known as pAIL and AtCon (see Supplementary Information). The cen180 F (forward) primer is conserved in Arabidopsis and related species, while the R (reverse) primer is specific to subsets of repeats [20]. The cen180 Fc and Rc primers match conserved regions within the repeat. For CLUSTALW analysis, sequences were determined for 25 to 30 cDNA clones from each mutant. A cen180 F primer with a T3 promoter and a cen180 R primer with a T7 promoter were used to amplify cen180 repeats, and these products were transcribed in vitro to prepare strand-specific probes for in situ hybridization, performed according to the protocol of Jeffrey Long (Salk Institute for Biological Studies, San Diego, California, United States) for protocol, see http://www.its.caltech.edu/~plantlab/protocols/oligoprep.pdf).

**Histone H3 methylation.** Chromatin immunoprecipitation (ChIP) was performed as previously described [39,40] using conserved F primers and R primers specific to each class of repeat. Semiquantification of the cen180 repeat ChIP data was performed by comparing PCR results from three different cycle numbers (13, 15, and 17 cycles), which were then analyzed by Southern blotting. In this way, the PCR of such highly repetitive sequences was maintained in the linear phase to avoid PCR saturation. DNA samples from each genotype were then normalized to each other by amplifying dilutions of total input DNA. Semiquantitative data were then obtained by comparing amplification with each set of primer pairs within the same ChIP extraction, which served as internal controls. In this way, control primers such as actin, whose association with lysine-9 is unclear, could be avoided. In all cases, mock precipitation with no control primers such as actin, whose association with lysine-9 is unclear, could be avoided. In all cases, mock precipitation with no antibody yielded little or no product. In Figure S3, PCR conditions were 17 cycles for F and R primers and 28 cycles for del1 and met-specific primers. PCR products were blotted and probed with radiolabeled cen180 repeats.

**siRNA hybridization.** Total RNA was purified, large RNAs were removed by precipitation in 5% PEG-8000/0.5 M NaCl, and 50 μg of the resulting enriched small RNA was loaded per lane in a 10% polyacrylamide/urea gel and separated by electrophoresis [39]. The gel was elec troblotted and probed with in vitro transcribed, radiolabeled RNA. Imaging was performed on a PhosphorImager (Fuji, Tokyo, Japan).

**Supporting Information**

Figure S1. Northern Analysis of cen180 Transcripts
Total RNA (50 μg) from WT (lane 1), met1 mutants (lane2), and del1 mutants (lane 3) were separated on an agarose gel, blotted, and probed with radiolabeled transcripts corresponding to each strand (F and R) of cen180 repeats. Heterogeneous high-molecular-weight transcripts hybridize to both strands in met1 mutants. Prominent bands cross-hybridize with DNA (left). Found at DOI: 10.1371/journal.pgen.0010079.g001 (171 KB PDF).

Figure S2. CLUSTALW Analysis of Sequences of Transcribed Repeats
An unrooted tree of cen180 cDNA sequences from different mutants in the Landsberg ecotype. Sequence identifiers beginning with w are from WT; d, from del1; a, from agol; and m, from met1. The met1 branches are red. Sequences found in WT, agol, and del1 were not found in met1.

Found at DOI: 10.1371/journal.pgen.0010079.g002 (183 KB PDF).

Figure S3. Methylation of Lysine-9 (K9) and Lysine-4 (K4) of Histone H3 at Centromeric Repeats
Chromatin was immunoprecipitated with anti–dimethyl K9 antibody (K9), anti–dimethyl K4 antibody (K4), or no antibody (na). Precipitated samples and total input chromatin (tot) were then PCR-amplified with cen180 F + R primers, F + dcl-specific primers, F + met-specific primers, and 106b F + R primers.

Found at DOI: 10.1371/journal.pgen.0010079.g003 (1.9 MB PDF).

Figure S4. Centromeric Repeat Methylation
Genomic DNAs were prepared from 3-week-old plants of the indicated genotypes. Methylation changes at (A) cen180 repeats and (B) 106b repeats were assayed by digestion with HpaII (left) andMspII (right) for each repeat followed by blotting and Southern hybridization.

Found at DOI: 10.1371/journal.pgen.0010079.g004 (775 KB PDF).

Figure S5. Comparison of cen180 cDNAs from WT (wt), del3 Mutants, and met1 Mutants with the Sequenced Region of the Arabidopsis Genome
Sequences of 32 cDNAs from WT and each mutant in the Columbia ecotype were BLASTed against the Columbia genome (TIGR version 5, http://www.tigr.org) using a cutoff of E < 0.0001. The matches for each repeat were collected and sorted into bins according to score. Each bin is 10 score points.

Found at DOI: 10.1371/journal.pgen.0010079.g005 (1.2 MB PDF).

**Text S1. Primer Sequences**
The same oligonucleotide primers were used for RT-PCR, ChIP, and probe amplification.

Found at DOI: 10.1371/journal.pgen.0010079.s001 (24 KB DOC).

**Accession Numbers**
The dBEST accession numbers in GenBank ([http://www.ncbi.nlm.nih.gov/Genbank]) for the centromeric satellite cDNA sequences are Dv671393 to Dv671628.

**Acknowledgments**
We thank I. Furner, J. Carrington, S. Jacobsen, and E. Richards for Arabidopsis strains and E. Lam, K. Dawe, and G. Karpen for helpful discussions. ZBL was an Arnold and Mabel Beckman fellow of the Watson School of Biological Sciences. This work was supported by National Science Foundation Plant Genome Program grant DBI-077617 to DLS and RAM and grant DBI-077774 to RAM.

**Competing interests.** The authors have declared that no competing interests exist.

**Author contributions.** BPM, ZBL, and RAM conceived and designed the experiments. BPM, ZBL, and YF performed the experiments. BPM, ZBL, YF, DLS, and RAM analyzed the data. BPM and ZBL contributed reagents/materials/analysis tools. BPM, ZBL, and RAM wrote the paper.

---

**Figure S6.**...
Dicer is essential for formation of the heterochromatin structure in vegetative cells. Nat Cell Biol 6: 78–79.

33. Sullivan BA, Karpen GH (2004) Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. Nat Struct Mol Biol 11: 1076–1085.

34. Blackburn EM, Lister C, Page T, Franz J, Findlay K, et al. (1999) The DFI1 gene of Arabidopsis is required for meiotic chromosome segregation and belongs to the REC8/RAD21 cohesin gene family. Plant J 19: 463–472.

35. Gaudin Y, Libault M, Pontouc E, Jusl T, Zhao G, et al. (2001) Mutations in LIKE HETEROCRICHOMATIN PROTEIN 1 affect flowering time and plant architecture in Arabidopsis. Development 128: 4847–4858.

36. Tamara H, Selker EU (2001) A histone H3 methyletransferase controls DNA methylation in Neurospora crassa. Nature 413: 227–233.

37. Jackson JP, Lindroth AM, Cao X, Jacobsen SE (2002) Control of CpG methylation DNA methylation by the KRYPTONITE histone H3 methyltransferase. Nature 416: 556–560.

38. Lehnerz B, Ueda Y, Derijck AA, Brainsmeweg U, Perez-Burgos L, et al. (2003) Sun59-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. Curr Biol 13: 1192–1200.

39. Lippmann Z, May B, Yordan C, Singer T, Martienssen R (2003) Distinct mechanisms determine transposon inheritance and methylation via small interfering RNA and histone modification. PLoS Biol 1: e67. DOI: 10.1371/journal.pbio.0000687.

40. Gendrel AV, Lippmann Z, Yordan C, Colot V, Martienssen RA (2002) Dependence of heterochromatic histone H3 methylations patterns on the Arabidopsis gene DDM1. Science 297: 1873–1873.

41. Soppe WJ, Haenicka Z, Houben A, Kakutani T, Meister A, et al. (2002) DNA methylation controls histone H3 lysine 9 methylation and heterochromatin assembly in Arabidopsis. EMBO J 21: 6277–6290.

42. Lindroth AM, Cao X, Jackson JP, Zilberman D, McCallum CM, et al. (2001) Requirement of CHROMOMETHYLASE3 for maintenance of CpG methylation. Science 292: 977–980.

43. Bartel L, Malagouzis G, Bender J (2001) Arabidopsis centm3 chromomethylase mutations block non-CG methylation and silencing of an endogenous gene. Genetics 157: 1753–1758.

44. Malagouzis G, Bartel L, Bender J (2002) An Arabidopsis SET domain protein required for maintenance but not establishment of DNA methylation. EMBO J 21: 6842–6852.

45. Zilberman D, Cao X, Jacobsen SE (2003) ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. Science 299: 716–719.

46. Lu F, Mette MF, Matze A, Matze A (2004) The role of MET1 in RNA-directed de novo and maintenance methylation of CG dinucleotides. Plant Mol Biol 54: 793–804.

47. Chan SW, Zilberman D, Xie Z, Johansen LK, Carrington JC, et al. (2004) RNA silencing genes control de novo DNA methylation. Science 303: 1336.

48. Gascoillii V, Mallory AC, Bartel DP, Vacherot H (2005) Partially redundant functions of Arabidopsis Dicer-like enzymes and a role for DCL4 in producing trans-acting siRNAs. Curr Biol 15: 1494–1500.

49. Heslop-Harrison JS, Murata M, Ogura Y, Schwarzacher T, Motoyoshi I (1999) Polymorphisms and genomic organization of repetitive DNA from heteromorphic regions of Arabidopsis thaliana. Plant Mol Biol 34: 65–69.

50. Schramke V, Allshire R (2003) Hairpin RNAs and retrotransposon LTRs effect RNAi and chromatin-based gene silencing. Science 301: 1069–1074.

51. Sahe H, Scheid OM, Pavkovic J (2005) Maintenance of Cpg methylation is essential for epigenetic inheritance during plant gametogenesis. Nature 434: 65–69.

52. Prasun P, Aristomancy C, Alonso-Blanco C, Fischer TC, Torres-Ruiz RA, et al. (1998) Cytogetics for the model system Arabidopsis thaliana. Plant J 13: 867–876.

53. Auffaetz W, Mette MF, van der Winder J, Matze M, Matze A (2002) A putative histone decytoase needed to enhance DNA methylation induced by double-stranded RNA. EMBO J 21: 6852–6861.

54. Reinhart BJ, Bartel DP (2002) Small RNAs that function in RNA interference and histone modification. PLoS Biol 1: e67. DOI: 10.1371/journal.pbio.0000687.

55. Li F, Goto DB, Zaratiegui M, Tang X, Martienssen R, et al. (2005) Two novel RNA-dependent DNA methylation genes in Arabidopsis thaliana. Mol Gen Genet 273: 1309–1319.

56. Verdel A, Derijck AA, Braunschweig U, Perez-Burgos L, et al. (2003) Rapid creation of BAC-based human artificial chromosome vectors by transposition with synthetic alpha-satellite arrays. Nucleic Acids Res 33: 563–570.

57. Auffaetz W, Mette MF, Matze A, Matze A (2004) The role of MET1 in RNA-directed de novo and maintenance methylation of CG dinucleotides. Plant Mol Biol 54: 793–804.

58. Chan SW, Zilberman D, Xie Z, Johansen LK, Carrington JC, et al. (2004) RNA silencing genes control de novo DNA methylation. Science 303: 1336.

59. Gascillii V, Mallory AC, Bartel DP, Vacherot H (2005) Partially redundant functions of Arabidopsis Dicer-like enzymes and a role for DCL4 in producing trans-acting siRNAs. Curr Biol 15: 1494–1500.

60. Heslop-Harrison JS, Murata M, Ogura Y, Schwarzacher T, Motoyoshi I (1999) Polymorphisms and genomic organization of repetitive DNA from heteromorphic regions of Arabidopsis thaliana. Plant Mol Biol 34: 65–69.

61. Schramke V, Allshire R (2003) Hairpin RNAs and retrotransposon LTRs effect RNAi and chromatin-based gene silencing. Science 301: 1069–1074.

62. Sahe H, Scheid OM, Pavkovic J (2005) Maintenance of Cpg methylation is essential for epigenetic inheritance during plant gametogenesis. Nature 434: 65–69.

63. Prasun P, Aristomancy C, Alonso-Blanco C, Fischer TC, Torres-Ruiz RA, et al. (1998) Cytogetics for the model system Arabidopsis thaliana. Plant J 13: 867–876.

64. Auffaetz W, Mette MF, van der Winder J, Matze M, Matze A (2002) A putative histone decytoase needed to enhance DNA methylation induced by double-stranded RNA. EMBO J 21: 6852–6861.

65. Reinhart BJ, Bartel DP (2002) Small RNAs that function in RNA interference and histone modification. PLoS Biol 1: e67. DOI: 10.1371/journal.pbio.0000687.

66. Li F, Goto DB, Zaratiegui M, Tang X, Martienssen R, et al. (2005) Two novel RNA-dependent DNA methylation genes in Arabidopsis thaliana. Mol Gen Genet 273: 1309–1319.

67. Auffaetz W, Mette MF, Matze A, Matze A (2004) The role of MET1 in RNA-directed de novo and maintenance methylation of CG dinucleotides. Plant Mol Biol 54: 793–804.

68. Chan SW, Zilberman D, Xie Z, Johansen LK, Carrington JC, et al. (2004) RNA silencing genes control de novo DNA methylation. Science 303: 1336.

69. Gascillii V, Mallory AC, Bartel DP, Vacherot H (2005) Partially redundant functions of Arabidopsis Dicer-like enzymes and a role for DCL4 in producing trans-acting siRNAs. Curr Biol 15: 1494–1500.

70. Heslop-Harrison JS, Murata M, Ogura Y, Schwarzacher T, Motoyoshi I (1999) Polymorphisms and genomic organization of repetitive DNA from heteromorphic regions of Arabidopsis thaliana. Plant Mol Biol 34: 65–69.

71. Schramke V, Allshire R (2003) Hairpin RNAs and retrotransposon LTRs effect RNAi and chromatin-based gene silencing. Science 301: 1069–1074.

72. Sahe H, Scheid OM, Pavkovic J (2005) Maintenance of Cpg methylation is essential for epigenetic inheritance during plant gametogenesis. Nature 434: 65–69.
centromeric regions in Arabidopsis. Proc Natl Acad Sci U S A 100: 11133–11138.

63. Martienssen RA (2003) Maintenance of heterochromatin by RNA interference of tandem repeats. Nat Genet 33: 213–214.

64. Gustafson AM, Allen E, Givan S, Smith D, Carrington JC, et al. (2005) ASRP: The Arabidopsis Small RNA Project Database. Nucleic Acids Res 33: D637–D640.

65. Martienssen RA, Zaratiegui M, Goto DB (2005) RNA interference and heterochromatin in the fission yeast Schizosaccharomyces pombe. Trends Genet 21: 450–456.

66. Topp CN, Zhong CX, Dawe RK (2004) Centromere-encoded RNAs are integral components of the maize kinetochore. Proc Natl Acad Sci U S A 101: 15986–15991.

67. Rudert F, Bronner S, Garnier JM, Dolle P (1995) Transcripts from opposite strands of gamma satellite DNA are differentially expressed during mouse development. Mamm Genome 6: 76–83.

68. Schnitzler G, Sif S, Kingston RE (1998) Human SWI/SNF interconverts a nucleosome between its base state and a stable remodeled state. Cell 94: 17–27.

69. Fransz P, De Jong JH, Lysak M, Castiglione MR, Schubert I (2002) Interphase chromosomes in Arabidopsis are organized as well defined chromocenters from which euchromatin loops emanate. Proc Natl Acad Sci U S A 99: 11484–11489.

70. Kankel MW, Ramsey DE, Stokes TL, Flowers SK, Haag JR, et al. (2003) Arabidopsis MET1 cytosine methyltransferase mutants. Genetics 163: 1169–1172.

71. Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, et al. (1998) Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature 393: 386–389.

72. Malik HS, Henikoff S (2002) Conflict begets complexity: The evolution of centromeres. Curr Opin Genet Dev 12: 711–718.

73. Sharbel TF, Hauhold B, Mitchell-Olds T (2000) Genetic isolation by distance in Arabidopsis thaliana: Biogeography and postglacial colonization of Europe. Mol Ecol 9: 2109–2118.

74. Borevitz JO, Liang D, Plouffe D, Chang HS, Zhu T, et al. (2003) Large-scale identification of single-feature polymorphisms in complex genomes. Genome Res 13: 513–523.

75. Osborn TC, Pires JC, Birchler JA, Auger DL, Chen ZJ, et al. (2003) Understanding mechanisms of novel gene expression in polyploids. Trends Genet 19: 141–147.

76. Ananiev EV, Riera-Lizarazu O, Rines HW, Phillips RL (1997) Oat-maize chromosome addition lines: A new system for mapping the maize genome. Proc Natl Acad Sci U S A 94: 3524–3529.

77. Sambrook J, Russell DW (2001) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor (New York): Cold Spring Harbor Laboratory Press. 999 pp.

Centromeric Transcripts in Arabidopsis