the plant journal

SEB Society for Experimental Biology

The Plant Journal (2018) doi: 10.1111/tpj.13910

Maternal components of RNA-directed DNA methylation are required for seed development in *Brassica rapa*

Jeffrey W. Grover[1,†], Timmy Kendall[2,†], Abdul Baten[3], Diane Burgess[4], Michael Freeling[4], Graham J. King[3] and Rebecca A. Mosher[1,2,*]

[1]*Department of Molecular & Cellular Biology, The University of Arizona, Tucson, AZ 85721, USA,*
[2]*The School of Plant Sciences, The University of Arizona, Tucson, AZ 85721, USA,*
[3]*Southern Cross Plant Science, Southern Cross University, Lismore, NSW 2480, Australia, and*
[4]*Department of Plant and Microbial Biology, The University of California, Berkeley, CA 94720, USA*

Received 17 August 2017; revised 10 March 2018; accepted 13 March 2018.
*For correspondence (e-mail rmosher@email.arizona.edu).
†These authors contributed equally.

SUMMARY

Small RNAs trigger repressive DNA methylation at thousands of transposable elements in a process called RNA-directed DNA methylation (RdDM). The molecular mechanism of RdDM is well characterized in Arabidopsis, yet the biological function remains unclear, as loss of RdDM in Arabidopsis causes no overt defects, even after generations of inbreeding. It is known that 24 nucleotide Pol IV-dependent siRNAs, the hallmark of RdDM, are abundant in flowers and developing seeds, indicating that RdDM might be important during reproduction. Here we show that, unlike Arabidopsis, mutations in the Pol IV-dependent small RNA pathway cause severe and specific reproductive defects in *Brassica rapa*. High rates of abortion occur when seeds have RdDM mutant mothers, but not when they have mutant fathers. Although abortion occurs after fertilization, RdDM function is required in maternal somatic tissue, not in the female gametophyte or the developing zygote, suggesting that siRNAs from the maternal soma might function in filial tissues. We propose that recently outbreeding species such as *B. rapa* are key to understanding the role of RdDM during plant reproduction.

Keywords: *Brassica rapa*, polymerase IV, polymerase V, RNA-directed DNA methylation, seed development, siRNA.

INTRODUCTION

Small RNAs are a common mechanism to protect a genome, with most eukaryotes deploying at least one variety of small RNA to transcriptionally or post-transcriptionally silence transposons (Malone and Hannon, 2009; Holoch and Moazed, 2015). Suppression of transposons is particularly important in the germ line to prevent inheritance of reactivated elements or transposon-based mutations (Iwasaki *et al.*, 2015). In *Drosophila*, Piwi-interacting RNAs (piRNAs) are produced in pre-meiotic nurse cells that surround and support the egg cell and are loaded into the egg before fertilization (Brennecke *et al.*, 2008; Klattenhoff and Theurkauf, 2008). These piRNAs trigger additional piRNA production after fertilization to suppress transposons in the zygotic germ line. The maternal initiation of piRNA-mediated silencing means that transposons inherited solely from the paternal genome can evade silencing and retrotranspose in the zygote.

Plants do not encode piRNAs, and instead use small interfering RNAs (siRNAs) to silence transposons (Matzke and Mosher, 2014; Fultz *et al.*, 2015). In the Arabidopsis male gametophyte, 21 nucleotide (nt) epigenetically activated RNAs (easiRNAs) are produced from transposon sequences in the pollen vegetative nucleus, a haploid cell that supports the sperm cells but does not contribute to the zygote (Slotkin *et al.*, 2009). These easiRNAs are transferred to the sperm cells, where they can silence homologous transcripts post-transcriptionally (Martínez *et al.*, 2016).

In vegetative tissues, 24 nt Pol IV-dependent siRNAs (p4-siRNAs) are produced from transposons by the sequential action of RNA polymerase IV, RNA-dependent RNA polymerase 2 (RDR2), and DCL3. P4-siRNAs then associate with ARGONAUTE4 and RNA polymerase V and transcriptionally silence transposons in a process called

RNA-directed DNA methylation (RdDM) (Matzke and Mosher, 2014). P4-siRNAs are expressed highly from the maternal genome during Arabidopsis seed development (Mosher *et al.*, 2009; Pignatta *et al.*, 2014) and have been implicated in controlling the effective parental dosage during seed development (Autran *et al.*, 2011; Lu *et al.*, 2012), and in establishing boundaries between euchromatin and heterochromatin (Zemach *et al.*, 2013; Li *et al.*, 2015). However, it is not clear whether p4-siRNAs are produced from the female gametophyte, the filial tissues (embryo and endosperm), or the somatic tissues surrounding these reproductive structures (Mosher *et al.*, 2009).

Neither paternal easiRNAs nor maternal p4-siRNAs are required for normal seed development in Arabidopsis (Mosher *et al.*, 2009; Martínez *et al.*, 2016), although paternal 21 nt siRNAs have been implicated in failure of triploid endosperm during unbalanced crosses (Borges *et al.*, 2018; Martínez *et al.*, 2018). However, Arabidopsis is an inbreeding species with few active transposons and therefore might be a poor system to study the function of small RNAs during reproduction. RdDM mutants in tomato and maize display pleiotropic phenotypes, including reproductive defects (Hollick, 2010; Gouil and Baulcombe, 2016), suggesting that p4-siRNAs and RdDM might be important during sexual reproduction in genomes with higher transposon content. Here we show that the loss of p4-siRNAs in *Brassica rapa* causes a specific defect during seed production. P4-siRNA production is required in maternal somatic tissue for the development of embryo and endosperm, suggesting that there is communication between diploid generations in the seed.

RESULTS AND DISCUSSION

To better understand the role of p4-siRNAs during reproduction, we isolated mutations impacting the production and activity of p4-siRNAs in *B. rapa*, a recently outbreeding relative of Arabidopsis whose genome is approximately 40% transposons (Wang *et al.*, 2011). In Arabidopsis, NRPD1 (the largest subunit of Pol IV) and RDR2 are required for p4-siRNA production, while NRPE1 (the largest subunit of Pol V) produces nascent transcripts that are required for p4-siRNA activity (Matzke and Mosher, 2014). Each of these genes has reverted to a single copy in the mesopolyploid *B. rapa* genome (Huang *et al.*, 2013) and we identified putative loss-of-function mutations in each gene (Table S1).

After repeated backcrossing to remove other mutations, we sequenced small RNA transcriptomes from the mutants (Table S2). As in other plants, 24 nt siRNAs are the most abundant size class in *B. rapa* reproductive tissues (Figure 1a). These siRNAs are strongly reduced in *braA.nrpd1.a-2* and *braA.rdr2.a-2* (hereafter *braA.nrpd1* and *braA.rdr2*), indicating they are p4-siRNAs. Compared with the almost complete loss of 24 nt siRNAs in *braA.rdr2*, *braA.nrpd1* retains a low level of 24 nt siRNAs, suggesting that this allele might not be a null mutation. As in Arabidopsis (Mosher *et al.*, 2008; Wang and Axtell, 2017), *braA.nrpe1.a-1* (*braA.nrpe1*) does not eliminate *B. rapa* 24 nt siRNAs (Figure 1a,b).

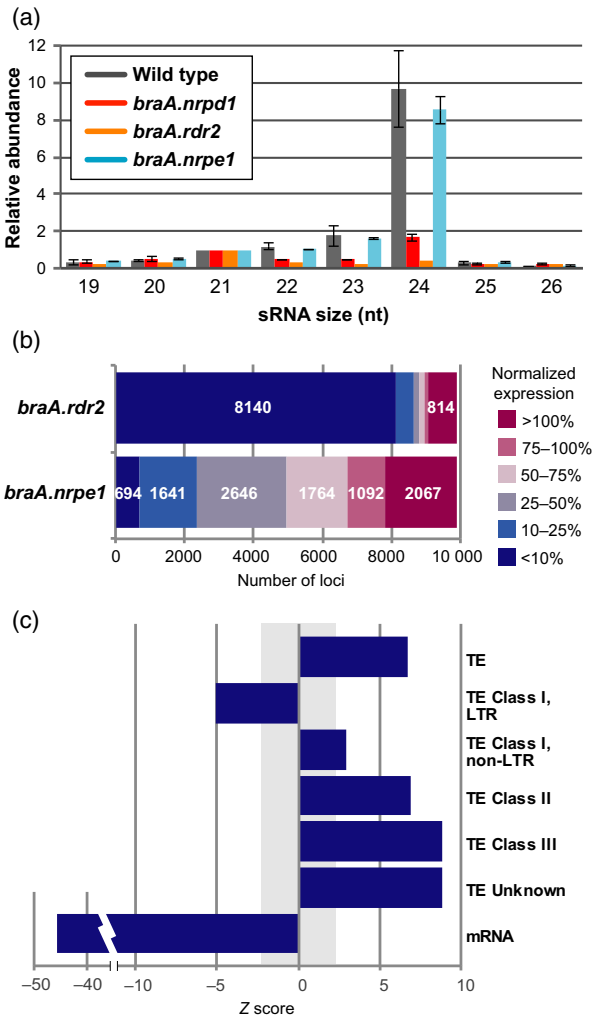

Figure 1. Small RNA analysis of *B. rapa* RdDM mutants.
(a) Size profile of 21 nt normalized genome-matching reads demonstrates that 24 nt small RNAs are the most abundant class of small RNAs in unfertilized ovules. These sequences are lost or strongly reduced in the *braA.rdr2* and *braA.nrpd1* mutants, indicating they are p4-siRNAs. Error bars depict standard deviation of the mean.
(b) Most small RNA-generating loci have sharply reduced sRNA expression in *braA.rdr2*, while there is a range of reduced expression in the *braA.nrpe1* mutant.
(c) P4-siRNA loci are depleted for mRNA annotations and enriched for transposons (TE), particularly DNA elements (Class II) and Helitrons (Class III). The gray shaded area indicates a $P > 0.01$.

We defined 9904 loci that produce small RNAs in the *B. rapa* genome (Figure S1). As expected, over 80% of small RNA-generating loci require *BraA.RDR2* (Figure 1b), indicating that p4-siRNA loci are the predominant class of small RNA-generating loci in *B. rapa*. We defined p4-siRNA

loci as those with at least a 90% reduced expression in *braA.rdr2* compared with wild-type. These loci are strongly depleted for protein-coding genes and enriched for transposons, including non-LTR retrotransposons, DNA elements, and Helitrons (Figure 1c). Surprisingly, LTR retrotransposons are depleted among the p4-siRNA loci, which might reflect Pol IV's preference for transposons in or near euchromatin (Zemach *et al.*, 2013; Li *et al.*, 2015). About half of all small RNA loci are downregulated at least 50% in *braA.nrpe1* (Figure 1b), demonstrating that although this mutation has a small effect on total p4-siRNA accumulation, it has a strong effect on siRNA production at specific loci.

Transcriptome analyses identified 1616 and 1803 genes that are differentially expressed by at least two-fold ($P < 0.05$ after FDR) in *braA.nrpd1* ovules and in seeds 10 days after fertilization, respectively (Tables S3 and S4). Upregulated genes were more likely to overlap p4-siRNA loci both before and after fertilization, while downregulated genes showed significant overlap with p4-siRNA loci only after fertilization (Figure 2a). Additionally, some differentially expressed genes accumulated p4-siRNA in putative promoters (Figure S2). *BraA.TT8* is a transcription factor responsible for the black seed coat that is characteristic of *Brassica* species. In *B. rapa* R-o-18 a Helitron transposon is inserted in the second intron of *BraA.TT8*, causing loss of expression and yellow seeds (Li *et al.*, 2012). We detected small RNAs from this Helitron in wild-type, but not in *braA.nprd1* or *braA.rdr2*, suggesting that p4-siRNAs could contribute to transcriptional repression of *BraA.TT8* (Figure 2b). Consistent with this hypothesis, *BraA.TT8* was significantly upregulated during *braA.nrpd1* seed development (Figure 2c) and the resulting seeds were slightly darker in color (Figure 2d), indicating that loss of p4-siRNA accumulation at least partially relieved transcriptional repression due to transposon insertion (Figure 2c,d).

We observed no vegetative growth defects associated with loss of RdDM in *B. rapa*, although *braA.nrpe1* plants flowered slightly later. All three mutants developed normal flowers and their siliques (seed pods) enlarged following self-fertilization. Mutant siliques, however, remained smaller than wild-type siliques and produced many fewer viable seeds (Figure 3a). The few seeds that were produced from homozygous mutant plants were smaller, lighter, and less uniformly round when compared to wild-type seeds (Figures 2d and 3b,c). However, these seeds were viable and produced plants that were indistinguishable from first-generation homozygous mutants. Hence, mutant seeds that progressed through embryogenesis went on to have normal vegetative development. These second-generation homozygous mutants produced viable seed at a similar frequency to their first-generation parents, suggesting that there is neither a progressive defect associated with loss of

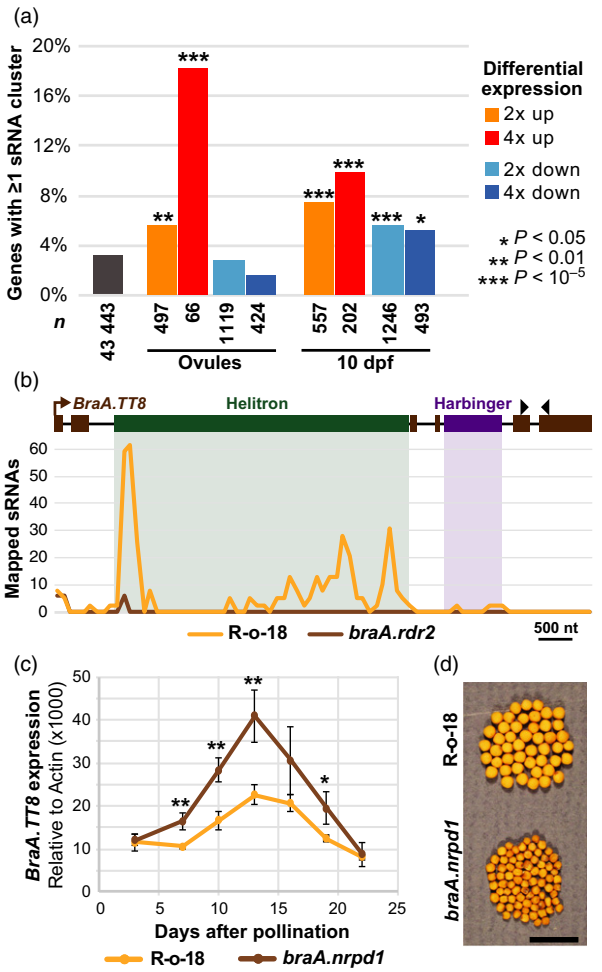

Figure 2. P4-siRNA accumulation is associated with changes in gene expression.
(a) Differentially expressed genes (at least two-fold change, adjusted $P < 0.05$) are significantly more likely to overlap at least 1 p4-siRNA cluster compared to all genes in the *B. rapa* genome.
(b) R-o-18 small RNAs map to transposon sequences in *BraA.TT8* introns but are absent in *braA.rdr2*. Mapped sRNAs are normalized to million mapped 21-mers and were counted in 100 bp bins.
(c, d) (c) *BraA.TT8* expression is significantly upregulated in *braA.nrpd1* seeds during seed coat development, resulting in slightly darker seeds (d).

p4-siRNAs nor a compensatory change in seeds produced from p4-siRNA mutants.

Careful assessment of seed weight in Arabidopsis *nrpd1* mutants uncovered a small but significant reduction in seed weight in each of three different genetic backgrounds (Figure 3d). This pattern indicated that p4-siRNAs might have a conserved influence on seed production that varies only in severity between *B. rapa* and Arabidopsis.

To better understand what caused the reduced seed production in *B. rapa* p4-siRNA mutants, we observed early events during seed development. Upon self-fertilization or manual pollination, mutant siliques elongated rapidly and remained indistinguishable from wild-type for

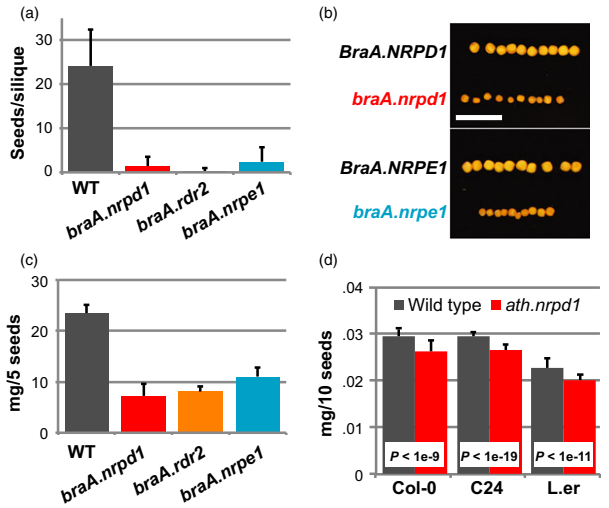

Figure 3. RdDM mutants are defective for seed production.
(a) The number of seeds per silique is strongly reduced in all three *B. rapa* RdDM mutants ($n > 200$ siliques from at least two individuals).
(b) Seed from *B. rapa* RdDM mutants are smaller and less uniform in shape than seed from wild-type siblings (scale bar represents 10 mm).
(c) *B. rapa* mutant seeds are also significantly lighter than wild-type ($n > 24$ measurements).
(d) Arabidopsis *nrpd1* mutant seeds are significantly lighter than wild-type in three different genetic backgrounds ($n > 35$ measurements). All error bars depict standard deviation.

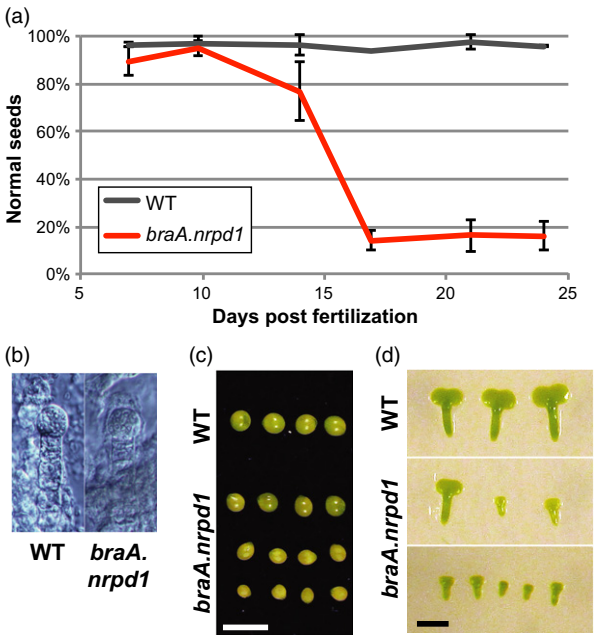

Figure 4. RdDM mutants display asynchronous seed abortion.
(a) Wild-type and mutant seeds contain identical globular embryos at 7 days post fertilization (dpf).
(b) Siliques were dissected at 7–24 dpf and the development of seeds was assessed. Seeds were considered abnormal if they were substantially smaller than wild-type or beginning to shrivel. In total, 50–300 seeds were counted for each of 2 or 3 replicates for each data point. Error bars depict standard deviation of biological replicates.
(c) Wild-type seeds are uniformly green, round, and plump, while *braA.nrpd1* seeds are variable in size, color, and turgidity at 17 dpf. Even the largest *braA.nrpd1* seeds (top row) are smaller than wild-type seeds and contain smaller and developmentally retarded embryos. The lower two rows of *braA.nrpd1* seeds demonstrate the variation in seed size. Scale bar represents 5 mm.
(d) Embryos dissected from 17 dpf wild-type or wild-type-like *braA.nrpd1* seeds. Wild-type embryos are uniform in size and developmental stage (late torpedo) while *braA.nrpd1* embryos are variable. Scale bar represents 1 mm.

at least 2 weeks. Within these siliques, mutant ovules increased in size only after pollination, indicating that fertilization and early seed development began normally. Light microscopy of cleared seeds further confirmed that fertilization occurred and that the embryogenesis program began correctly (Figure 4a). By day 14 (torpedo stage of embryogenesis), many of the developing mutant seeds were smaller than wild-type seeds, and smaller and/or shriveled seeds became more abundant as development progressed (Figure 4b). By 17 days after fertilization, diverse seed phenotypes were observed in mutant siliques, ranging from clearly aborted to nearly normal (Figure 4c). However, even the roundest and plumpest mutant seeds were smaller than age-matched wild-type seeds and contained embryos of diverse sizes and developmental states (Figure 4d). Endosperm from mutant seeds was also variable in size and developmental stage and was considerably smaller than wild-type endosperm. This variability remained evident in mature mutant siliques, which contained a range of aborted seed – from small flecks of completely degenerated seeds to large, but clearly shriveled and unviable seeds. This diversity suggests that in *B. rapa*, loss of p4-siRNAs or their activity caused asynchronous seed abortion, rather than a defect at one fixed point of development.

In Arabidopsis, as in *Drosophila*, the maternal genotype determines whether small RNAs are produced after fertilization; developing seeds with *nrpd1* mothers have strongly reduced p4-siRNA expression even when fertilized by wild-type pollen (Mosher *et al.*, 2009; Lu *et al.*, 2012). To determine whether the reproductive defect in *B. rapa* is controlled by maternal or zygotic genotype, we reciprocally crossed homozygous mutants to wild-type and scored seed production. The genotype of pollen donors had no impact on the seed set, while seed production in a mutant × wild-type cross (female genotype listed first in all crosses) was indistinguishable from a mutant × mutant cross (Figure 5a). Furthermore, the seeds produced from mutant × wild-type crosses were smaller and less round than seeds produced from the reciprocal wild-type × mutant cross, indicating that the maternal genotype is responsible for both aspects of the reproductive defect – seed number and seed quality.

Although in Arabidopsis dissected embryo and endosperm show few protein-coding genes with parental bias (Nodine and Bartel, 2012), developing seeds are biased toward maternal small RNA (Mosher *et al.*, 2009; Pignatta

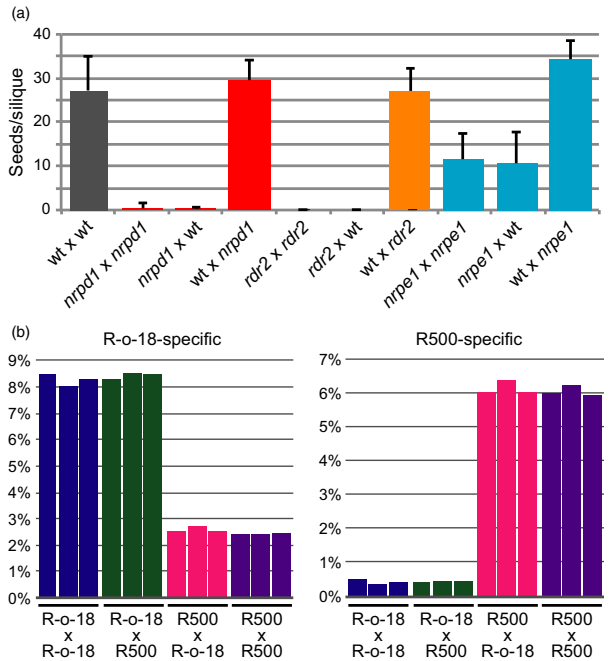

Figure 5. Seed production and siRNA populations are maternally controlled in *B. rapa*.

(a) Reciprocal crosses between RdDM mutants and wild-type demonstrates that maternal genotype is the determining factor in seed production. Viable seeds were counted for <10 siliques per cross (maternal parent listed first). Error bars indicate standard deviation.

(b) Seeds from reciprocal R-o-18× R-500 crosses were dissected 14 days after fertilization and subjected to small RNA sequencing. The resulting libraries (three biological replicates per genotype) were mapped onto both parental genomes. Reads matching only to the R-o-18 genome (left) and those matching only the R-500 (right) demonstrate that paternal genotype (listed at bottom) does not substantially influence the small RNA population.

et al., 2014). To assess parental contribution of small RNAs in *B. rapa*, we sequenced small RNAs from reciprocal intraspecific (R-o-18 × R-500) crosses at 14 days after fertilization and mapped small RNAs to both genomes independently. Although genome-specific siRNAs do not provide a strict measure of genome-of-origin because of missing sequence in each assembly, comparison of the relative level of genome-specific reads between crosses gives an indication of the contribution from the maternal and paternal genomes. As in Arabidopsis (Mosher *et al.*, 2008), there is no difference in the level of genome-specific reads with different paternal genotypes, indicating that the paternally contributed chromosomes do not contribute significantly to the small RNA transcriptome of the seed as a whole (Figure 5b). This observation indicated that most p4-siRNA production is either from maternally inherited chromosomes in the filial tissues, or from the integuments, which are diploid maternal tissue that develops into the seed coat after fertilization.

We previously hypothesized that p4-siRNAs are maternal-specific after fertilization due to epigenetic patterns established by p4-siRNAs in the female gametophyte before fertilization (Mosher *et al.*, 2009; Mosher, 2010). However, perturbation of various epigenetic marks did not influence maternal-specific p4-siRNA production (Mosher *et al.*, 2009, 2011), raising the possibility that p4-siRNAs in the developing seed are instead produced from maternal sporophytic tissue. The clear and quantitative phenotype associated with loss of p4-siRNAs in *B. rapa* provides a useful system to determine whether p4-siRNAs are required in the diploid maternal sporophyte, the haploid female gametophyte, or both. To address this question, we genotyped progeny from crosses with heterozygous mutants, which have a wild-type allele in the sporophyte and in 50% of the resulting gametophytes. A gametophytic defect would result in partial loss of seed production and distortion of the expected allelic ratio in the progeny. Instead, we observed that seed size, shape, and number were indistinguishable from wild-type, and detected the expected 1:1 ratio of alleles in all crosses, indicating that both male and female gametophytes function normally even when carrying a p4-siRNA mutation (Table 1). Combined with the observation that homozygous mutant mothers have defective seed production, we concluded that the seed production defect is due to loss of p4-siRNA production or activity in the maternal sporophytic tissue, even though the defect manifests after successful gametogenesis and fertilization.

The requirement for functional p4-siRNA in maternal sporophytic tissues does not rule out the possibility that p4-siRNAs are also required in zygotic tissues. To measure any defect in seed production due to the zygotic genotype we measured seed production and genotypes in crosses between heterozygous mothers and homozygous mutant fathers. Seed production was normal (both in phenotype and number) and resulted in equal development of heterozygous and homozygous mutant progeny, confirming that seed abortion is controlled solely by maternal sporophytic genotype (Table 1).

Our analysis of *B. rapa* mutants demonstrated that p4-siRNA production or activity is required in *B. rapa* maternal sporophytic tissue for successful seed development. This process is reminiscent of the maternal sporophytic requirement for *NRPD2* (the second subunit of Pol IV and V) for suppression of maternally transmitted *EVADE* retrotransposons before meiosis in Arabidopsis (Garcia *et al.*, 2005). While it is possible that seed abortion in *B. rapa* RdDM mutants is due to transposition of reactivated transposons before meiosis or during gametogenesis, we think this hypothesis is unlikely to be correct because mutant ovules are indistinguishable from wild-type ovules before fertilization and during early embryogenesis (Figure 4b). It is also possible that RdDM is required specifically during integument development as mutations that reduce integument growth can result in

Table 1 Seed abortion is controlled by maternal sporophytic genotype

	Progeny genotype[a]			
	+/+	+/−	−/−	*P*-value[b]
Het. × wild-type (WT)				
nrpd1-2/+ × WT	32	34		n.s.
nrpe1-1/+ × WT	25	31		n.s.
rdr2-2/+ × WT	30	20		n.s.
wild-type (WT) × Het.				
WT × *nrpd1-2*/+	29	34		n.s.
WT × *nrpe1-1*/+	35	31		n.s.
WT × *rdr2-2*/+	25	30		n.s.
−Het. × homo.				
nrpd1-2/+ × *nrpd1-2*		68	57	n.s.
nrpe1-1/+ × *nrpe1-1*		65	60	n.s.
rdr2-2/+ × *rdr2-2*		65	57	n.s.

[a]+/+, WT; +/−, heterozygous; −/−, homozygous mutant.
[b]*P*-values from chi-squared test. n.s., not significant.

smaller endosperm (Garcia *et al.*, 2005) and our analysis of *BraA.TT8* demonstrated that p4-siRNAs can influence gene expression in the integument (Figure 2). However it is not clear why RdDM would be required specifically in integuments when other somatic development is unaffected.

Instead, we propose that p4-siRNAs produced from maternal sporophytic tissue might move to the developing endosperm and/or embryo to control transposons during seed development. These p4-siRNAs could be produced in the integuments, or alternatively, move in the phloem from vegetative tissues. Loss of RdDM could negatively affect seed development in several possible ways. Massive deregulation of transposons would explain the asynchronous and nearly complete seed abortion we observed, and is consistent with reports of transposon deregulation linked to seed abortion during the triploid block (Erdmann *et al.*, 2017; Borges *et al.*, 2018; Martínez *et al.*, 2018). It is also possible that mobile maternal p4-siRNAs are required for correct expression of specific transposon-associated genes, such as AGAMOUS-like transcription factors, which are upregulated in Arabidopsis seeds lacking maternal *NRPD1* (Lu *et al.*, 2012). Finally, mobile maternal p4-siRNAs might function across the genome to maintain the effective parental dosage in endosperm and their loss could unbalance parental contributions and lead to endosperm failure (Gehring, 2013; Yang *et al.*, 2017).

Our data suggest that control of developing progeny by parental somatic small RNA production is a more common mechanism than was previously appreciated. In addition to maternal loading of piRNAs in flies, ciliates also use small RNAs from the parental macronuclei to program the macronuclei of daughter cells (Mochizuki and Gorovsky, 2004). Our results offer another example of parental tissues influencing developing progeny through small RNAs, and suggest that this process is a common biological mechanism.

EXPERIMENTAL PROCEDURES

Plant growth, genotyping, and phenotyping

All plants were grown in a greenhouse at 18°C with supplemental lighting to 16 h daylight. The TILLING population was created in R-o-18, an inbred line derived from the self-compatible, yellow-seeded subspecies *trilocularis* (Stephenson *et al.*, 2010). *B. rapa* is generally an outbreeding species, although self-compatible varieties such as *trilocularis* have arisen in the last 3000–5000 years (Downey and Rimmer, 1993; Qi *et al.*, 2017). Because the TILLING population is heavily mutagenized, putative loss-of-function alleles were backcrossed six times to the parental line before generating homozygous mutants. For all crosses, flowers were emasculated prior to anthesis and were pollinated 1 day later. The resulting plants were genotyped with molecular markers using standard approaches. *braA.nrpd1.a-2* (at BraA09006512) was genotyped with AAACGGGAGACAGCTTCTTACG and AGGAACGT ACCCGTGAAGACAGACT and digested with *Hin*fI; *braA.rdr2.a-2* (at BraA09004909) was amplified with TTCGAAATGTGCTGCTAGG ATGAGC and CTGAAGGAACTGCGGTCAACG and digested with *Alu*I; *braA.nrpe1.a-1* (at BraA03002228) was amplified with CTGGCCCGGCTTGGACATTTTTCC and ACAAGTCTTCCAGACTT AAACTAAATCCTT and digested with *Afl*II.

To assess embryogenesis, seeds were manually dissected at 17 days after fertilization, or were cleared as described in (Wang *et al.*, 2017). *B. rapa* and Arabidopsis seed weight was measured from seed that had dried for at least 4 weeks. *B. rapa* seeds were measured in pools of 5, while Arabidopsis seeds were measured in pools of 10.

Small RNA sequencing and analysis

Unfertilized ovules from *braA.nrpd1.a-2*, *braA.rdr2.a-2*, and *braA.nrpe1.a-1* were collected less than 24 h prior to anthesis and total RNA was prepared in a method adapted from White and Kaper (1989). Briefly, frozen tissue was ground to a fine powder and mixed with at least five volumes of extraction buffer (100 mM glycine, 10 mM EDTA, 100 mM NaCl, 2% SDS, pH 9.5), prior to organic extraction and precipitation. Small RNA was enriched with a mirVana miRNA isolation kit (Thermo Fisher Scientific) prior to NEBNext small RNA library preparation (New England Biolabs). Three independent biological replicates were prepared for each genotype. Barcoded libraries were pooled and sequenced on an Illumina HiSeq 2500 at The University of Missouri DNA Core Facility.

Trimmed reads from unfertilized ovules were obtained from the sequencing facility and aligned to the *B. rapa* R-o-18 genome (version 1.2, A Baten and G.J. King, unpublished) using Bowtie v1.2.2 (Langmead *et al.*, 2009). Genome-mapping reads were further filtered by aligning to the Rfam database v12.1 (Nawrocki *et al.*, 2015) and removing small RNAs matching structural and noncoding RNAs (excluding miRNAs and miRNA precursors) annotated in *B. rapa* or *A. thaliana*. Reads mapping to the *A. thaliana* TAIR10 assembly (Berardini *et al.*, 2015) chloroplast and mitochondrial genomes, and reads <19 or >26 nt were also removed with a custom Python script (Table S2). In all cases, Bowtie options were -v 0, -m 50, –best, -a, and –nomaqround. –norc was used for the Rfam filtering step only. No mismatches were allowed for any step (bowtie option -v 0). Size profiles of independent replicates

were similar, and therefore replicates were pooled for further analyses.

For small RNA locus annotation, filtered reads from pooled wild-type libraries were mapped to the R-o-18 genome using ShortStack v3.8.1 (Axtell, 2013; Johnson *et al.*, 2016) allowing no mismatches. ShortStack options were –mismatches 0, –mmap u, and –mincov 2 rpm. Reads within 75 bp (default value for –pad) were merged into loci and were considered for further analysis if they met a minimum read depth of two reads per million (approximately 12 reads in wild-type). Pooled libraries for each genotype were then realigned with ShortStack to compare expression at wild-type-defined loci. To address over-sampling, raw read counts were normalized by the number of mapped 21 nt reads in each pooled library after confirming that putative miRNA machinery (*AGO1*, *DCL1*, *HYL1*, *SE*, and *DRB2*) was not differentially expressed in *braA.nrpd1* ovules. Principle component analysis was performed on rlog-transformed siRNA counts within loci using DESeq2 (Love *et al.*, 2014) and plotted with the ggplot2 R package.

siRNA loci associated with genomic features were determined with bedtools intersect v2.26.0 (Quinlan and Hall, 2010) using default options. Genomic coordinates of siRNA loci were randomized 100× using bedtools shuffle, with option –noOverlapping, and z-scores were calculated by the formula:

$$Z = \frac{\text{overlaps}_{\text{observed}} - \text{mean overlaps}_{\text{randomized}}}{\text{sd}_{\text{randomized}}}$$

Reads from intraspecific crosses were filtered and mapped as above except both the R-o-18 and R-500 genomes (K. Greenham and C.R. McClung, unpublished) were used for mapping (Table S5). To maximize mapping of siRNAs, a larger but unordered and unannotated R-o-18 assembly was used (A Baten and G.J. King, unpublished). Genome-specific reads were defined as reads with a perfect match in only one of the two parental genomes.

Small RNA sequence data were deposited at the NCBI Short Read Archive (SRA) under accession numbers SRP114437 (RdDM mutants) and SRP114469 (intraspecific crosses).

mRNA sequencing and analysis

Total RNA from R-o-18 and *braA.nrpd1* was prepared as described above and sent to the University of Missouri DNA Core for TruSeq mRNA library preparation and sequencing. Ovules were collected less than 1 day before anthesis, and whole seeds were collected 10 days after manual pollination. Trimmed RNA-seq reads were aligned to the R-o-18 genome with STAR v2.5.4b (Dobin *et al.*, 2013) using default options. Reads overlapping annotated genes were determined with the Bioconductor GenomicAlignments R package (Lawrence *et al.*, 2013) using the summarizeOverlaps function and options mode = 'union', singleEnd = TRUE, ignore. strand = TRUE. Principle component analysis was performed on rlog-transformed counts per gene with DESeq2 and visualized with ggplot2 (Figure S3). Differentially expressed genes were identified using DESeq2 (Love *et al.*, 2014) and overlapping siRNA loci were determined with bedtools intersect (Quinlan and Hall, 2010).

RNA sequence data were deposited at the NCBI Short Read Archive (SRA) under accession number SRP132223.

Transposable element annotation

Transposable elements were annotated in the *B. rapa* R-o-18 genome with RepeatMasker (v. 4-0.5) using RMBlast and a combined library of exemplar repetitive DNA elements from the following sources: *Brassica* and Arabidopsis repetitive DNA elements from RepBase (http://www.girinst.org/repbase/) updated through vol17, issue 4; *Brassica rapa* RepeatModeler-generated dataset (Cheng *et al.*, 2016); *Arabidopsis thaliana* transposons from TAIR10 (ftp://ftp.arabidopsis.org/home/tair/Genes/TAIR10_genome_release/, downloaded 3/27/2017); *Brassica* MITE, TRIM and SINEs (Murukarthick *et al.*, 2014); *Brassica rapa* MITEs (Chen *et al.*, 2014); and *Brassica* hATs (Nouroz *et al.*, 2015). As the RepeatModeler-generated dataset of repetitive elements included 1037 unclassified elements, 176 of the unclassified elements were manually annotated and used to classify an additional 226 elements based on blastn hits with e-values <1E-10. Unclassified elements were also blasted against *Arabidopsis thaliana* CDS sequences (blastn e-value <1E-05); 72 of the elements with blastn hits were manually annotated and classified either as gene families and removed, or as Helitron or Mutator-MuDR elements with captured sequences (23 elements); the remaining 66 non-analyzed repetitive elements were removed from the library.

ACKNOWLEDGEMENTS

We are grateful to Dr Robertson McClung and Dr Kathleen Greenham from Dartmouth University for sharing the *B. rapa* R-500 genome prior to publication. Thanks also to Fran Robson at RevGenUK for TILLING advice, and Dr Nathan Bivens and the staff at the University of Missouri DNA Core.

FUNDING

This research was funded by the NSF Plant Genome Research Program (IOS-1546825 to R.A.M. and M.F.).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. Principle component analysis of *B. rapa* sRNA clusters.

Figure S2. Small RNA accumulation is associated with differentially expressed genes.

Figure S3. Principle component analysis of *B. rapa* RNA-seq.

Table S1. *B. rapa* RdDM mutant alleles.

Table S2. Small RNA sequencing of RdDM mutants.

Table S5. Small RNA sequencing of reciprocal crosses.

Table S3. Differentially expressed genes in unfertilized ovules.

Table S4. Differentially expressed genes in 10 dpf seeds.

REFERENCES

Autran, D., Baroux, C., Raissig, M.T. *et al.* (2011) Maternal epigenetic pathways control parental contributions to Arabidopsis early embryogenesis. *Cell*, **145**, 707–719.

Axtell, M.J. (2013) ShortStack: comprehensive annotation and quantification of small RNA genes. *RNA*, **19**, 740–751.

Berardini, T.Z., Reiser, L., Li, D., Mezheritsky, Y., Muller, R., Strait, E. and Huala, E. (2015) The Arabidopsis information resource: making and mining the "gold standard" annotated reference plant genome. *Genesis*, **53**, 474–485.

Borges, F., Parent, J.-S., Van Ex, F., Wolff, P., Martínez, G., Köhler, C. and Martienssen, R.A. (2018) Transposon-derived small RNAs triggered by miR845 mediate genome dosage response in Arabidopsis. *Nat. Genet.* **109**, 14746.

Brennecke, J., Malone, C.D., Aravin, A.A., Sachidanandam, R., Stark, A. and Hannon, G.J. (2008) An epigenetic role for maternally inherited piRNAs in transposon silencing. *Science*, **322**, 1387–1392.

Chen, J., Hu, Q., Zhang, Y., Lu, C. and Kuang, H. (2014) P-MITE: a database for plant miniature inverted-repeat transposable elements. *Nucleic Acids Res.* **42**, D1176–D1181.

Cheng, F., Sun, C., Wu, J., Schnable, J., Woodhouse, M.R., Liang, J., Cai, C., Freeling, M. and Wang, X. (2016) Epigenetic regulation of subgenome dominance following whole genome triplication in *Brassica rapa*. *New Phytol.* **211**, 288–299.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.

Downey, R.K. and Rimmer, S.R. (1993) Agronomic improvement in oilseed. *Brassicas Advances in Agronomy* **50**, 1–66.

Erdmann, R.M., Satyaki, P.R.V., Klosinska, M. and Gehring, M. (2017) A small RNA pathway mediates allelic dosage in endosperm. *Cell Rep.* **21**, 3364–3372. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=29262317&retmode=ref&cmd=prlinks.

Fultz, D., Choudury, S.G. and Slotkin, R.K. (2015) Silencing of active transposable elements in plants. *Curr. Opin. Plant Biol.* **27**, 67–76.

Garcia, D., Fitz Gerald, J.N. and Berger, F. (2005) Maternal control of integument cell elongation and zygotic control of endosperm growth are coordinated to determine seed size in Arabidopsis. *Plant Cell*, **17**, 52–60.

Gehring, M. (2013) Genomic imprinting: insights from plants. *Annu. Rev. Genet.* **47**, 187–208.

Gouil, Q. and Baulcombe, D.C. (2016) DNA methylation signatures of the plant chromomethyltransferases. *PLoS Genet.* **12**, e1006526.

Hollick, J.B. (2010) Paramutation and development. *Annu. Rev. Cell Dev. Biol.* **26**, 557–579.

Holoch, D. and Moazed, D. (2015) RNA-mediated epigenetic regulation of gene expression. *Nat. Rev. Genet.* **16**, 71–84.

Huang, Y., Kendall, T. and Mosher, R. (2013) Pol IV-dependent siRNA production is reduced in *Brassica rapa*. *Biology*, **2**, 1210–1223.

Iwasaki, Y.W., Siomi, M.C. and Siomi, H. (2015) PIWI-interacting RNA: its biogenesis and functions. *Annu. Rev. Biochem.* **84**, 405–433.

Johnson, N.R., Yeoh, J.M., Coruh, C. and Axtell, M.J. (2016) Improved placement of multi-mapping small RNAs. *G3 (Bethesda)*, **6**, 2103–2111.

Klattenhoff, C. and Theurkauf, W. (2008) Biogenesis and germline functions of piRNAs. *Development*, **135**, 3–9.

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

Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M.T. and Carey, V.J. (2013) Software for computing and annotating genomic ranges. *PLoS Comput. Biol.* **9**, e1003118.

Li, X., Chen, L., Hong, M. *et al.* (2012) A large insertion in bHLH transcription factor BrTT8 resulting in yellow seed coat in *Brassica rapa*. *PLoS ONE*, **7**, e44145.

Li, Q., Gent, J.I., Zynda, G. *et al.* (2015) RNA-directed DNA methylation enforces boundaries between heterochromatin and euchromatin in the maize genome. *Proc. Natl Acad. Sci. USA*, **112**, 14728–14733.

Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550.

Lu, J., Zhang, C., Baulcombe, D.C. and Chen, J. (2012) Maternal siRNAs as regulators of parental genome imbalance and gene expression in endosperm of Arabidopsis seeds. *Proc. Natl Acad. Sci. USA*, **109**, 5529–5534.

Malone, C.D. and Hannon, G.J. (2009) Small RNAs as guardians of the genome. *Cell*, **136**, 656–668.

Martínez, G., Panda, K., Köhler, C. and Slotkin, R.K. (2016) Silencing in sperm cells is directed by RNA movement from the surrounding nurse cell. *Nat. Plants*, **2**, 16030.

Martínez, G., Wolff, P., Wang, Z., Moreno-Romero, J., Santos-González, J., Conze, L.L., DeFraia, C., Slotkin, R.K. and Köhler, C. (2018) Paternal easiRNAs regulate parental genome dosage in Arabidopsis. *Nat. Genet.* **57**, 5. Available at: http://www.nature.com/articles/s41588-017-0033-4.

Matzke, M.A. and Mosher, R.A. (2014) RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nat. Rev. Genet.* **15**, 394–408.

Mochizuki, K. and Gorovsky, M.A. (2004) Conjugation-specific small RNAs in Tetrahymena have predicted properties of scan (scn) RNAs involved in genome rearrangement. *Genes Dev.* **18**, 2068–2073.

Mosher, R.A. (2010) Maternal control of Pol IV-dependent siRNAs in Arabidopsis endosperm. *New Phytol.* **186**, 358–364.

Mosher, R.A., Schwach, F., Studholme, D. and Baulcombe, D.C. (2008) PolIVb influences RNA-directed DNA methylation independently of its role in siRNA biogenesis. *Proc. Natl Acad. Sci. USA*, **105**, 3145–3150.

Mosher, R.A., Melnyk, C.W., Kelly, K.A., Dunn, R.M., Studholme, D.J. and Baulcombe, D.C. (2009) Uniparental expression of PolIV-dependent siRNAs in developing endosperm of Arabidopsis. *Nature*, **460**, 283–286.

Mosher, R.A., Tan, E.H., Shin, J., Fischer, R.L., Pikaard, C.S. and Baulcombe, D.C. (2011) An atypical epigenetic mechanism affects uniparental expression of Pol IV-dependent siRNAs. *PLoS ONE*, **6**, e25756.

Murukarthick, J., Sampath, P., Lee, S.-C., Choi, B.-S., Senthil, N., Liu, S. and Yang, T.-J. (2014) BrassicaTED – a public database for utilization of miniature transposable elements in Brassica species. *BMC Res. Notes*, **7**, 379.

Nawrocki, E.P., Burge, S.W., Bateman, A. *et al.* (2015) Rfam 12.0: updates to the RNA families database. *Nucleic Acids Res.* **43**, D130–D137.

Nodine, M.D. and Bartel, D.P. (2012) Maternal and paternal genomes contribute equally to the transcriptome of early plant embryos. *Nature*, **482**, 94–97.

Nouroz, F., Noreen, S. and Heslop-Harrison, J.S. (2015) Identification, characterization and diversification of non-autonomous hAT transposons and unknown insertions in Brassica. *Genes and Genomics*, **37**, 945–958.

Pignatta, D., Erdmann, R.M., Scheer, E., Picard, C.L., Bell, G.W. and Gehring, M. (2014) Natural epigenetic polymorphisms lead to intraspecific variation in Arabidopsis gene imprinting. *Elife*, **3**, e03198.

Qi, X., An, H., Ragsdale, A.P., Hall, T.E., Gutenkunst, R.N., Chris Pires, J. and Barker, M.S. (2017) Genomic inferences of domestication events are corroborated by written records in *Brassica rapa*. *Mol. Ecol.* **26**, 3373–3388.

Quinlan, A.R. and Hall, I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.

Slotkin, R.K., Vaughn, M., Borges, F., Tanurdzic, M., Becker, J.D., Feijo, J.A. and Martienssen, R.A. (2009) Epigenetic reprogramming and small RNA silencing of transposable elements in Pollen. *Cell*, **136**, 461–472.

Stephenson, P., Baker, D., Girin, T., Perez, A., Amoah, S., King, G.J. and Østergaard, L. (2010) A rich TILLING resource for studying gene function in *Brassica rapa*. *BMC Plant Biol.* **10**, 62.

Wang, F. and Axtell, M.J. (2017) AGO4 is specifically required for heterochromatic siRNA accumulation at Pol V-dependent loci in *Arabidopsis thaliana*. *Plant J.* **90**, 37–47.

Wang, X., Wang, H., Wang, J. *et al.* (2011) The genome of the mesopolyploid crop species *Brassica rapa*. *Nat. Genet.* **43**, 1035–1039.

Wang, Y., Tsukamoto, T., Noble, J.A., Liu, X., Mosher, R.A. and Palanivelu, R. (2017) Arabidopsis LORELEI, a maternally expressed imprinted gene, promotes early seed development. *Plant Physiol.* **175**, 758–773.

White, J.L. and Kaper, J.M. (1989) A simple method for detection of viral satellite RNAs in small plant tissue samples. *J. Virol. Methods*, **23**, 83–93.

Yang, R., Zheng, Z., Chen, Q. *et al.* (2017) The developmental regulator PKL is required to maintain correct DNA methylation patterns at RNA-directed DNA methylation loci. *Genome Biol.* **18**, 103.

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