Title: Nurse cell-derived small RNAs define paternal epigenetic inheritance in *Arabidopsis*

Authors: Jincheng Long[1]*, James Walker[1]*, Wenjing She[1]*, Billy Aldridge[1], Hongbo Gao[1,2], Samuel Deans[1], Martin Vickers[1], Xiaoqi Feng[1]†

Affiliations:

[1]Department of Cell and Developmental Biology, John Innes Centre, Norwich NR4 7UH, UK.

[2]Current affiliation: Joint Center for Single Cell Biology, School of Agriculture and Biology, Shanghai Jiao Tong University, Shanghai, 200240, China.

*These authors contributed equally to this work.

†Corresponding author. Email: xiaoqi.feng@jic.ac.uk

Abstract

The plant male germline undergoes DNA methylation reprogramming, which methylates genes *de novo* and thereby alters gene expression and regulates meiosis. Here we reveal the molecular mechanism underlying this reprogramming. We demonstrate that genic methylation in the male germline, from meiocytes to sperm, is established by 24 nucleotide siRNAs transcribed from transposons with imperfect sequence homology. These siRNAs are synthesized by meiocyte nurse cells (tapetum) via activity of CLSY3, a chromatin remodeler absent in other anther cells. Tapetal siRNAs govern germline methylation throughout the genome, including the inherited methylation patterns in sperm. Tapetum-derived siRNAs also silence germline transposons, thereby safeguarding genome integrity. Our results reveal that tapetal siRNAs are sufficient to reconstitute germline methylation patterns and drive functional methylation reprogramming throughout the male germline.

Keywords

Germline, DNA methylation, reprogramming, RdDM, siRNA, gene regulation, transposon silencing, meiocyte, tapetum, *Arabidopsis*.

Main Text

Methylation of the 5th carbon of cytosine carries essential regulatory functions in eukaryotic genomes, including transcriptional regulation of genes and transposable elements (TEs) (*1-3*). DNA methylation patterns are faithfully replicated during cell division, thus allowing methylation to have homeostatic functions during development (*1, 2*). However, animal and plant germlines undergo DNA methylation reprogramming (*4-6*). In mammals, impaired reprogramming is associated with misregulation of genes and TEs, and arrests male meiosis (*7*). In plants, DNA methylation reprogramming occurs during gametogenesis, when global methylation levels change (*8-11*) and thousands of loci are actively demethylated in gamete companion cells (*9, 10*). DNA methylation reprogramming in the *Arabidopsis thaliana* male germline also involves *de novo* methylation at hundreds of genes, which regulates gene expression and meiosis (*6*). This germline-specific methylation is catalyzed by the small RNA-directed DNA methylation pathway (RdDM) (*6*), which generally methylates transposable elements (TEs) (*12*). How RdDM targets genes specifically in the male germline, and how particular genes are selected for methylation, are unknown.

24-nt siRNA pattern in meiocytes

In RdDM, 24-nt small interfering RNAs (siRNAs), which are produced from transcripts synthesized by RNA polymerase IV (Pol IV, a plant-specific derivative of Pol II) and RNA-dependent RNA polymerase 2 (RDR2), guide methyltransferases DRM1 and 2 to target loci via association with a homologous transcript generated by Pol V (another plant-specific derivative of Pol II) (*12*). To understand how specific loci are methylated by RdDM exclusively in the male germline, we sequenced siRNAs from isolated *Arabidopsis* male meiocytes (prophase I, mostly

pachytene stage; Fig. 1A). In somatic tissues such as leaves, roots and seedlings, 24-nt siRNAs associate with methylated RdDM loci (~ 10,000 loci, predominantly TEs (*6*); Fig. 1, B and C, and fig. S1A). In relation to somatic tissues, nearly all (98%) of these canonical RdDM loci have substantially fewer 24-nt siRNAs in meiocytes (Fig. 1, B and C, and fig. S1, A and B), and CHH methylation at these loci is reduced (Fig. 1, B and D). However, when normalized to 21-nt microRNAs, the levels of 24-nt siRNAs at canonical RdDM loci are similar in meiocytes and soma (fig. S1C). The vast relative difference in 24-nt siRNA abundance occurs because male meiocyte 24-nt siRNAs are concentrated ($\log_2$(RPKM+1) > 7) in 854 clusters, most of which (80%) do not overlap canonical RdDM loci. These clusters have higher levels of siRNAs and DNA methylation in meiocytes compared to somatic tissues (Fig. 1, B to D, and fig. S1, A and D), and overlap the 724 loci we previously identified as hypermethylated in the male germline (*6*) (fig. S1E).

Consistent with RdDM targeting, the majority (93%, 797) of meiocyte siRNA clusters have significantly reduced CHH methylation in *drm1 drm2* double mutant (simplified as *drm*) or *rdr2* mutant plants (Fig. 1, B and D, and table S1), and like canonical RdDM loci (*13*), overlap mostly non-coding regions (fig. S1F). These 797 siRNA clusters span 588 Kb (0.4% of the genome) but comprise 94% of the clustered 24-nt siRNAs, mostly (70%) overlapping TEs (TEs make up 18% of the genome; fig. S1G and table S2). Altogether, our results demonstrate that male meiocytes have a distinctive RdDM-associated siRNA profile, with the majority of 24-nt siRNAs targeting a small number of hypermethylated loci (we will refer to these 797 loci as HyperTEs; table S2).

TE-derived siRNAs methylate genes with mismatches

We found few 24-nt siRNAs in meiocytes associated with genes hypermethylated in the male germline (MetGenes; Fig. 1, B and C, and fig. S1, A and G), even though MetGenes are targeted

by RdDM (*6*). A comparison revealed that MetGenes and HyperTEs share similar sequences, leading us to hypothesize that MetGenes may be targeted by siRNAs produced from HyperTEs. Supporting this hypothesis, we found that 24-nt siRNAs generated from HyperTEs can be aligned to MetGenes if up to three mismatches are allowed (Fig. 1, B and E, and table S3). The association of siRNAs with MetGenes is not caused by random mapping of mismatched siRNAs, because neither all genes nor random control loci associate with mismatched siRNAs (Fig. 1E). Furthermore, we find that for 89% of HyperTE-associated siRNAs that map to MetGenes, MetGenes are the best hits in the genome aside from the source HyperTEs. These analyses suggest that siRNAs produced from HyperTEs cause methylation at MetGenes with similar sequences.

To test the causal relationship between HyperTE-associated siRNAs and the methylation at MetGenes, we created two independent CRISPR/Cas9 lines with deletions of a HyperTE (HyperTE224, *AT2TE15980*; Fig. 1F and fig. S2A). We observed abolishment of methylation at the predicted target MetGene (MetGene186, *AT3G04230*; Fig. 1F and table S3) in meiocytes isolated from both deletion lines (Fig. 1G), whereas overall methylation did not change (fig. S2B). We applied this experimental strategy to another HyperTE (HyperTE315, *AT2TE72995*), siRNAs from which are predicted to target two MetGenes (MetGene24, *AT1G15520;* and MetGene86, *AT1G56410*) (fig. S2C and table S3). Consistently, we observed complete loss of methylation at both target MetGenes in meiocytes of the HyperTE315-deletion lines (Fig. 1H and fig. S2, B and D). These results demonstrate that methylation at MetGenes is induced *in trans* by siRNAs derived from HyperTEs.

Meiocytes and tapetum share siRNA profiles

In RdDM, Pol IV and Pol V preferentially associate with methylated DNA, making RdDM a self-reinforcing pathway, in which methylation promotes the generation of methylation-inducing siRNA (*14, 15*). Therefore, regardless of the initial guiding siRNA, Pol V-mediated DNA methylation attracts the Pol IV pathway *in situ* to generate perfectly matching siRNAs (*12, 14*). The observation that few perfectly matching siRNAs are associated with MetGenes (Fig. 1, B, C and E) suggests that the Pol IV pathway is quiescent in meiocytes, with MetGene-inducing siRNAs generated from HyperTEs in other cells.

Male meiocytes are enclosed by a layer of tapetal nurse cells (*16*) (Fig. 1A). The meiocytes and tapetal cells are connected via cytoplasmic channels called plasmodesmata (*17, 18*), which are regarded as the most likely route for intercellular siRNA movement in plants (*19-21*). Furthermore, 24-nt siRNAs have been proposed to accumulate in maize tapetum during early meiosis (*22*), exactly the stage when plasmodesmata connect meiocytes and tapetal cells. Therefore, siRNAs that induce methylation at MetGenes may be generated by HyperTEs in the tapetum and transported into meiocytes. To test this hypothesis, we developed a method to isolate tapetal cells. We generated *Arabidopsis* transgenic plants carrying GFP driven by a tapetum-specific promoter (*pA9*, simplified as *pTP*; Fig. 2A) (*23, 24*) and isolated GFP-positive tapetal cells via fluorescence-activated cell sorting (FACS; fig. S3A). The isolated tapetal cells are of high purity, evaluated through fluorescence microscopy (>95% purity) and RNA sequencing, which shows enrichment of tapetum-specific genes and depletion of somatic and meiotic genes (table S4).

DNA methylation analysis of tapetal cells revealed that HyperTEs are hypermethylated in the tapetum in comparison to somatic tissues, at levels even higher than in meiocytes (Figs. 1B and 2B). As in meiocytes (Fig. 1, B and D), HyperTE methylation in the tapetum requires RDR2 (Figs.

1B and 2B). Tapetal cells also show enrichment of 24-nt siRNAs at HyperTEs (Fig. 2, C and D, and fig. S3B), similar to meiocytes (Fig. 1C). As in meiocytes (Fig. 1C), siRNA levels at canonical RdDM loci are relatively low but similar in absolute terms to somatic cells (Fig. 2C and fig. S1C) and DNA methylation at these loci is low (Figs. 1B and 2B). Genome-wide, 24-nt siRNAs in identified siRNA clusters strongly correlate between tapetal cells and meiocytes (Fig. 2E). As in meiocytes (Fig. 1C), few siRNAs match MetGenes perfectly in tapetal cells (Fig. 2C). The resemblance between siRNA profiles suggests *en masse* siRNA movement from tapetum to meiocytes, which is consistent with knowledge of intercellular siRNA transport through plasmodesmata (*25*). In contrast to meiocytes, MetGenes are not methylated in tapetal cells (Figs. 1B and 2B). This is consistent with tapetum having an active Pol IV pathway, because RdDM-mediated methylation should not exist in the absence of perfectly matching siRNAs in such cells. Our results also imply that, unlike in meiocytes, the tapetal Pol V pathway is unable to induce DNA methylation with mismatched siRNAs.

Tapetal siRNAs drive meiocyte methylation reprogramming

To test whether tapetum-derived siRNAs work *in trans* to induce methylation in meiocytes, we created a genetic mosaic system with 24-nt siRNA biogenesis confined to the tapetum. We expressed RDR2 (fused with a Flag tag) in the *rdr2* null mutant background using the tapetum-specific *pA9* promoter (simplified as *pTP::RDR2 rdr2*; Fig. 3A). Immunolocalization with anti-Flag antibodies in anther cross-sections confirmed the specific induction of RDR2 in the tapetum in two independent mosaic lines (Fig. 3A). In the meiocytes isolated from both lines, methylation at MetGenes and HyperTEs was increased to levels comparable with wild type (Fig. 3, B and C, and fig. S4, A and B). As a control, we generated a *pTP::POL-V pol-v* mosaic line. As Pol V

mediates DNA methylation rather than siRNA biogenesis, it must work *in cis*. Consistently, meiocytes from two independent lines show no methylation at MetGenes (Fig. 3, B and C, and fig. S4C shows Pol V protein is correctly expressed). These results demonstrate that activity of the Pol IV pathway in the tapetum is sufficient to methylate MetGenes and HyperTEs in meiocytes.

RDR2 produces double-stranded RNA (dsRNA) precursors, which are converted to 24-nt siRNAs by DCL3 (*12*). Because the dsRNA precursors have been shown to move between cells (*19, 20*), we sought to understand whether these molecules or siRNAs are the primary signals that move from the tapetum to meiocytes. We therefore generated two independent mosaic lines with DCL3 confined to the tapetum (*pTP::DCL3 dcl3*; fig. S4C shows DCL3 is correctly expressed). Meiocytes from both lines show wild-type methylation at MetGenes and HyperTEs, whereas *dcl3* mutants have reduced methylation at both sets of loci in meiocytes (Fig. 3, B and C). This indicates that tapetal 24-nt siRNAs are sufficient to trigger methylation at MetGenes and HyperTEs in meiocytes. Furthermore, *DCL3* expression is 9.4 times greater in tapetal cells than in meiocytes (fig. S4D), suggesting most 24-nt siRNA dicing occurs in the tapetum. These results suggest that 24-nt siRNAs are the predominant non-cell-autonomous signal that originates in the tapetum and drives methylation in meiocytes.

HyperTEs produce siRNAs via CLSY3

We next investigated how the distinctive siRNA profile in the tapetum is generated (Fig. 2, C and D). Four putative chromatin remodelers, CLASSY1-4 (CLSYs), recruit Pol IV to produce siRNAs at largely discrete sets of loci (*26*). Therefore, we evaluated the relationship of the *clsy1*/*2*/*3*/*4*-dependent siRNA clusters identified in floral buds (*26*) with HyperTEs. We found a proportion of HyperTEs (66%; 524 out of 797) overlap *clsy3*-dependent siRNA clusters (804 loci), whereas

only 1% to 8% of HyperTEs overlap *clsy1*-, *2*- or *4*- dependent clusters (fig. S5A). Consistently, in both tapetal cells and meiocytes, relatively few 24-nt siRNAs are associated with *clsy1,2*- or *clsy4*-dependent clusters, whereas *clsy3*-dependent clusters are enriched in siRNAs (Fig. 4A). In contrast in leaves, most siRNAs are associated with *clsy1,2*- and *clsy4*-dependent clusters (Fig. 4A).

We next examined the transcription of *CLSY*s in the tapetum, meiocytes and somatic tissues. We found that *CLSY1* and *2* are the predominant homologs in somatic tissues, whereas *CLSY3* is the most highly expressed *CLSY* in the tapetum (Fig. 4B). To validate this result and examine the expression pattern of CLSY3, we generated a *CLSY3-Venus* fusion line (*pCLSY3::Venus-CLSY3*). Confocal microscopy shows CLSY3 is undetectable in seedlings and leaves but enriched in carpels and anthers (Fig. 4C and fig. S5B). In the anther, CLSY3 is specifically expressed in tapetal cells and absent from all other cell types, including meiocytes (Fig. 4C). Based on these results, CLSY3 is likely responsible for the distinctive siRNA profile in tapetal and meiocyte cells. CLSY3 is also enriched in ovules (fig. S5B), where accumulation of 24-nt siRNAs from ~200 loci was recently observed in *Brassica* (*27*), suggesting a similar methylation reprogramming phenomenon on the female side. CLSY3 activity also likely contributes to the high siRNA abundance in the floral bud (Fig. 4A), a mixture of cells and tissues that express different CLSY proteins.

To test our hypothesis that tapetal CLSY3 drives the production of HyperTE-associated siRNAs, we isolated meiocytes from the *clsy3* mutant and performed siRNA sequencing. We also examined *clsy4* meiocytes, because *CLSY3* and *4* were reported to be partially redundant (*26*) and *CLSY4* transcripts are detected in the tapetum (Fig. 4B). We found that 24-nt siRNAs at HyperTEs

are lost in *clsy3* mutant meiocytes, similarly to *pol iv* mutant meiocytes, whereas *clsy4* meiocytes are unaffected (Fig. 4D). Consistently, we found loss of methylation at HyperTEs and MetGenes in *clsy3* meiocytes, whereas *clsy1*, *2* or *4* mutant meiocytes and *clsy1;2* double mutant meiocytes have wild-type levels of methylation (Fig. 4E). These results establish that CLSY3 is required for the production of HyperTE-associated siRNAs in the tapetum and the consequent methylation of HyperTEs and MetGenes. Given that in anthers CLSY3 is absent from meiocytes and only expressed in the tapetum (Fig. 4C), these results further demonstrate that siRNA biogenesis is suppressed in meiocytes and meiocyte siRNAs are derived from the tapetum.

Tapetal siRNAs shape sperm DNA methylation

As the methylation at MetGenes persists throughout male germline development to gametogenesis (*6*) (Fig. 1B), we wondered whether tapetum-derived siRNAs can drive methylation throughout the germline. To test this, we analyzed DNA methylation in sperm isolated from the *clsy3* mutant. DNA methylation at MetGenes and HyperTEs is reduced in *clsy3* mutant sperm to levels resembling *drm* mutant sperm (Fig. 5, A and B, and fig. S6). In contrast, methylation at canonical RdDM loci is not affected by the *clsy3* mutation (Fig. 5, A and B, and fig. S6).

To confirm the competence of tapetal siRNAs to shape sperm methylation, we examined sperm from the *pTP::RDR2 rdr2* mosaic line. DNA methylation at MetGenes and HyperTEs is restored to wild-type levels in *pTP::RDR2 rdr2* sperm (including CG, CHG and CHH methylation; Fig. 5, A and B, and fig. S6). Furthermore, methylation at canonical RdDM loci is also restored (Fig. 5, A and B, and fig. S6), in line with soma-like absolute levels of 24-nt siRNA at canonical RdDM loci in the tapetum (fig. S1C). These results demonstrate that tapetal siRNAs drive DNA methylation reprogramming in the male germline all the way to the gametes.

Tapetal siRNAs silence germline transposons

MetGene methylation was shown to regulate gene expression and male meiosis (*6*). Besides MetGenes, our results demonstrate that tapetal siRNAs also induce methylation at TEs in the male germline (Figs. 3B, 4D and 5A). Therefore, we investigated whether siRNAs serve to silence germline TEs in addition to their gene regulatory function. To address this question, we performed RNA sequencing of tapetal cells and pollen from wild type and the *drm* mutant. Analysis of these data in combination with analogous data from meiocytes and leaves (*6*) identified retrotransposons from a Gypsy family (*GP1*) that are specifically activated in RdDM-defective germline cells (Fig. 6A). *GP1* transcription is not activated in *drm* mutant leaves, but strongly activated in the tapetum and pollen (and slightly in meiocytes; Fig. 6A). We observed loss of methylation in RdDM (*drm* and *rdr2*) mutants at the LTRs in the tapetum and meiocytes (fig. S7A), indicating that *GP1* is directly suppressed by RdDM.

To further examine *GP1* activity during germline development, we generated a YFP reporter line driven by *GP1*'s LTR sequence ($GP1_{LTR}$::*YFP*) and crossed it to *rdr2* and *drm* mutants. Consistent with the RNA-seq data, YFP signal was undetectable in wild type and specifically observed in the anthers of *rdr2* and *drm* mutants (Fig. 6, B to D). In both mutants, *GP1* activity can be first observed in the tapetum at a stage prior to meiosis (Fig. 6B). Subsequently, YFP signal is strongest in the tapetum at the onset of meiosis, and in microspores (and to a lesser extent pollen) after meiosis (Fig. 6, B to D). These observations confirm that *GP1* is specifically activated in the germline and tapetum when RdDM is compromised. Although DNA methylation of the *GP1* LTRs requires RdDM in the soma, tapetum and germline (fig. S7A), loss of RdDM does not activate

GP1 in somatic tissues (Fig. 6A). This suggests that *GP1* targets expression in reproductive cells, possibly by exploiting transcription factors specific to these cell types.

To confirm that tapetum is the origin of *GP1* silencing siRNAs, we first took advantage of the fact that microspores are haploid meiotic products in which genetic segregation has occurred (Figs. 1A and 6C). In a mutant carrying a recessive heterozygous *rdr2* mutation, the diploid tapetal cells (of *rdr2*/+ genotype) are able to produce siRNAs, whereas half of the haploid microspores (the ones of *rdr2* genotype) are not (Fig. 6C). We did not detect *GP1* activity in any anther tissue, including the tapetum and microspores, of $GP1_{LTR}$*::YFP rdr2/+* plants (Fig. 6D), confirming that *GP1*-silencing siRNAs are produced in diploid somatic cells. In contrast, half (50.7%, n = 1391) of the microspores from $GP1_{LTR}$*::YFP drm/+* plants exhibit YFP signal (Fig. 6D), consistent with the *cis* action of DRM methyltransferases.

To validate the silencing of *GP1* by tapetal siRNAs, we crossed the *GP1* YFP reporter into the *pTP::RDR2 rdr2* genetic background that limits siRNA production to the tapetum. No *GP1* activity was observed in $GP1_{LTR}$*::YFP pTP::RDR2 rdr2* anthers (Fig. 6D), demonstrating the competence of tapetal siRNAs to silence *GP1* in the male germline. *GP1* is a canonical RdDM locus (fig. S7, A and B), further demonstrating that tapetal siRNA production and germline activity are not confined to HyperTEs (fig. S8).

Discussion

Our study demonstrates that DNA methylation reprogramming in the male germline is driven by tapetum-derived siRNAs. While quiescent in 24-nt siRNA biogenesis, male meiocytes can use imperfectly matching siRNAs to target DNA methylation, suggesting the Pol V branch of RdDM

is hypersensitive in meiocytes (fig. S8). Thus, meiosis appears to be a key stage for genome surveillance, during which RdDM targets genes and TEs with sequences sufficiently similar to those of TEs that produce siRNAs in the tapetum (*e.g.*, *GP1*) (fig. S8). Methylation of genes allows transcriptional regulation by RdDM in the germline (*6*), which has the potential to improve fitness and hence be selected during evolution at specific loci.

Tapetal siRNAs can induce DNA methylation not only in meiocytes, but throughout the male germline (Fig. 5A). Furthermore, tapetal siRNAs are competent to restore not only sperm MetGene methylation, but also the full spectrum of sperm RdDM (Fig. 5A, and figs. S6 and S8). The central role of tapetal siRNAs in determining the paternally inherited DNA methylome illustrates the reprogramming potential of nurse cells. This theme has also emerged in studies of animal germlines (*28*), indicating convergent functional evolution between plant and animal germline epigenetic reprogramming systems.

Materials and methods summary

Arabidopsis male meiocytes, tapetal cells and sperm nuclei were isolated by microdissection or FACS to produce siRNA, RNA, and bisulfite sequencing libraries. Sequencing data was processed using the following software: Bowtie, Bismark, Kallisto, Shortstack, Cutadapt, TrimGalore and MarkDuplicates. HyperTEs were identified by requiring higher levels of 24-nt siRNA compared to 25-nt siRNA in wildtype meiocytes as well as requiring a loss of meiotic CHH methylation in *drm2* or *rdr2* mutant meiocytes. Canonical RdDM loci and MetGenes were previously reported (*6*). HyperTE CRISPR deletion lines were created using guide RNAs listed in table S5. Immunolocalization of *pA9::RDR2-Flag rdr2* anthers was examined via a Zeiss DM6000

microscope. Expression of *pA9::Pol-V-3×Flag* and *pA9::DCL3-3×Flag* was confirmed by inflorescence immunoblot assay. *CLSY3-Venus* and *GP1$_{LTR}$::YFP* lines were examined using a Leica SP5 confocal microscope.

REFERENCES AND NOTES

1. Z. D. Smith, A. Meissner, DNA methylation: roles in mammalian development. *Nat Rev Genet* **14**, 204-220 (2013).
2. H. Zhang, Z. Lang, J. K. Zhu, Dynamics and function of DNA methylation in plants. *Nat Rev Mol Cell Biol* **19**, 489-506 (2018).
3. A. Zemach, D. Zilberman, Evolution of eukaryotic DNA methylation and the pursuit of safer sex. *Curr Biol* **20**, R780-785 (2010).
4. G. F. Wang, C. Kohler, Epigenetic processes in flowering plant reproduction. *J Exp Bot* **68**, 797-807 (2017).
5. S. Seisenberger *et al.*, Reprogramming DNA methylation in the mammalian life cycle: building and breaking epigenetic barriers. *Philos T R Soc B* **368**, (2013).
6. J. Walker *et al.*, Sexual-lineage-specific DNA methylation regulates meiosis in Arabidopsis. *Nature Genetics* **50**, 130-137 (2018).
7. M. V. C. Greenberg, D. Bourc'his, The diverse roles of DNA methylation in mammalian development and disease. *Nature Reviews Molecular Cell Biology* **20**, 590-607 (2019).
8. M. W. Schmid *et al.*, Extensive epigenetic reprogramming during the life cycle of Marchantia polymorpha. *Genome Biol* **19**, 9 (2018).
9. C. A. Ibarra *et al.*, Active DNA demethylation in plant companion cells reinforces transposon methylation in gametes. *Science* **337**, 1360-1364 (2012).
10. J. P. Calarco *et al.*, Reprogramming of DNA methylation in pollen guides epigenetic inheritance via small RNA. *Cell* **151**, 194-205 (2012).
11. P. H. Hsieh *et al.*, Arabidopsis male sexual lineage exhibits more robust maintenance of CG methylation than somatic tissues. *Proc Natl Acad Sci U S A* **113**, 15132-15137 (2016).
12. M. A. Matzke, R. A. Mosher, RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nat Rev Genet* **15**, 394-408 (2014).
13. T. K. To *et al.*, RNA interference-independent reprogramming of DNA methylation in Arabidopsis. *Nat Plants* **6**, 1455-1467 (2020).

14. J. M. Wendte, C. S. Pikaard, The RNAs of RNA-directed DNA methylation. *Biochim Biophys Acta Gene Regul Mech* **1860**, 140-148 (2017).

15. H. Y. Kuo, E. L. Jacobsen, Y. Long, X. Chen, J. Zhai, Characteristics and processing of Pol IV-dependent transcripts in Arabidopsis. *J Genet Genomics* **44**, 3-6 (2017).

16. X. Feng, D. Zilberman, H. Dickinson, A conversation across generations: soma-germ cell crosstalk in plants. *Dev Cell* **24**, 215-225 (2013).

17. E. A. Mamun, L. C. Cantrill, R. L. Overall, B. G. Sutton, Cellular organisation and differentiation of organelles in pre-meiotic rice anthers. *Cell Biol Int* **29**, 792-802 (2005).

18. R. Sager, J. Y. Lee, Plasmodesmata in integrated cell signalling: insights from development and environmental signals and stresses. *J Exp Bot* **65**, 6337-6358 (2014).

19. C. W. Melnyk, A. Molnar, D. C. Baulcombe, Intercellular and systemic movement of RNA silencing signals. *Embo J* **30**, 3553-3563 (2011).

20. D. E. Pyott, A. Molnar, Going mobile: non-cell-autonomous small RNAs shape the genetic landscape of plants. *Plant Biotechnol J* **13**, 306-318 (2015).

21. L. Liu, X. Chen, Intercellular and systemic trafficking of RNAs in plants. *Nat Plants* **4**, 869-878 (2018).

22. J. Zhai *et al.*, Spatiotemporally dynamic, cell-type-dependent premeiotic and meiotic phasiRNAs in maize anthers. *Proc Natl Acad Sci U S A* **112**, 3146-3151 (2015).

23. X. Feng, H. G. Dickinson, Tapetal cell fate, lineage and proliferation in the Arabidopsis anther. *Development* **137**, 2409-2416 (2010).

24. W. Paul, R. Hodge, S. Smartt, J. Draper, R. Scott, The isolation and characterization of the tapetum-specific Arabidopsis thaliana A9 gene. *Plant Mol Biol* **19**, 611-622 (1992).

25. C. Faulkner, Plasmodesmata and the symplast. *Curr Biol* **28**, R1374-R1378 (2018).

26. M. Zhou, A. M. S. Palanca, J. A. Law, Locus-specific control of the de novo DNA methylation pathway in Arabidopsis by the CLASSY family. *Nature Genetics* **50**, 865-873 (2018).

27. J. W. Grover *et al.*, Abundant expression of maternal siRNAs is a conserved feature of seed development. *Proc Natl Acad Sci U S A* **117**, 15305-15315 (2020).

28. K. F. Toth, D. Pezic, E. Stuwe, A. Webster, The piRNA pathway guards the germline genome against transposable elements. *Adv Exp Med Biol* **886**, 51-77 (2016).

ACKNOWLEDGMENTS

We thank the John Innes Centre Bioimaging Facility (Sergio Lopez, Eva Wegel and Kim Findlay) for their assistance with microscopy, and the Norwich BioScience Institute Partnership Computing infrastructure for Science Group for High Performance Computing resources. **Funding:** This work was funded by a European Research Council Starting Grant ('SexMeth' 804981; J.L., J.W., and X.F.), a Sainsbury Charitable Foundation studentship (J.W.), two Biotechnology and Biological Sciences Research Council (BBSRC) grants (BBS0096201 and BBP0135111; W.S., M.V., and X.F.), two John Innes Foundation studentships (B.A. and S.D.), and a BBSRC David Phillips Fellowship (BBL0250431; H.G. and X.F.). **Author contributions:** J.L., J.W., and X.F. designed the study and wrote the manuscript, J.L., W.S., B.A., H.G., and S.D. performed the experiments, and J.L., J.W., B.A., H.G., S.D., M.V., and X.F. analyzed the data. **Competing interests:** The authors declare no competing interests. **Data and material availability:** All sequencing data have been deposited in the Gene Expression Omnibus (GEO) under accession number GSE161625. Accession numbers of published datasets used in this study are listed in table S6. Published software used in this study include Bowtie v1.2.2 (https://doi.org/10.1002/0471250953.bi1107s32), Bismark v0.22.2 (https://doi.org/10.1093/bioinformatics/btr167), Kallisto v0.43.0 (https://doi.org/10.1038/nbt0816-888d), Shortstack v3.8.5 (https://doi.org/10.1534/g3.116.030452), and Cutadapt v1.15 (https://doi.org/10.1089/cmb.2017.0096). TrimGalore v0.4.1 and MarkDuplicates v1.141 are available from https://github.com/FelixKrueger/TrimGalore and https://github.com/broadinstitute/picard. All remaining data are in the main paper or the supplement.

SUPPLEMENTARY MATERIALS

Materials and Methods

Figs. S1-S8

Tables S1-S6

References (*29-56*)

MDAR Reproducibility Checklist

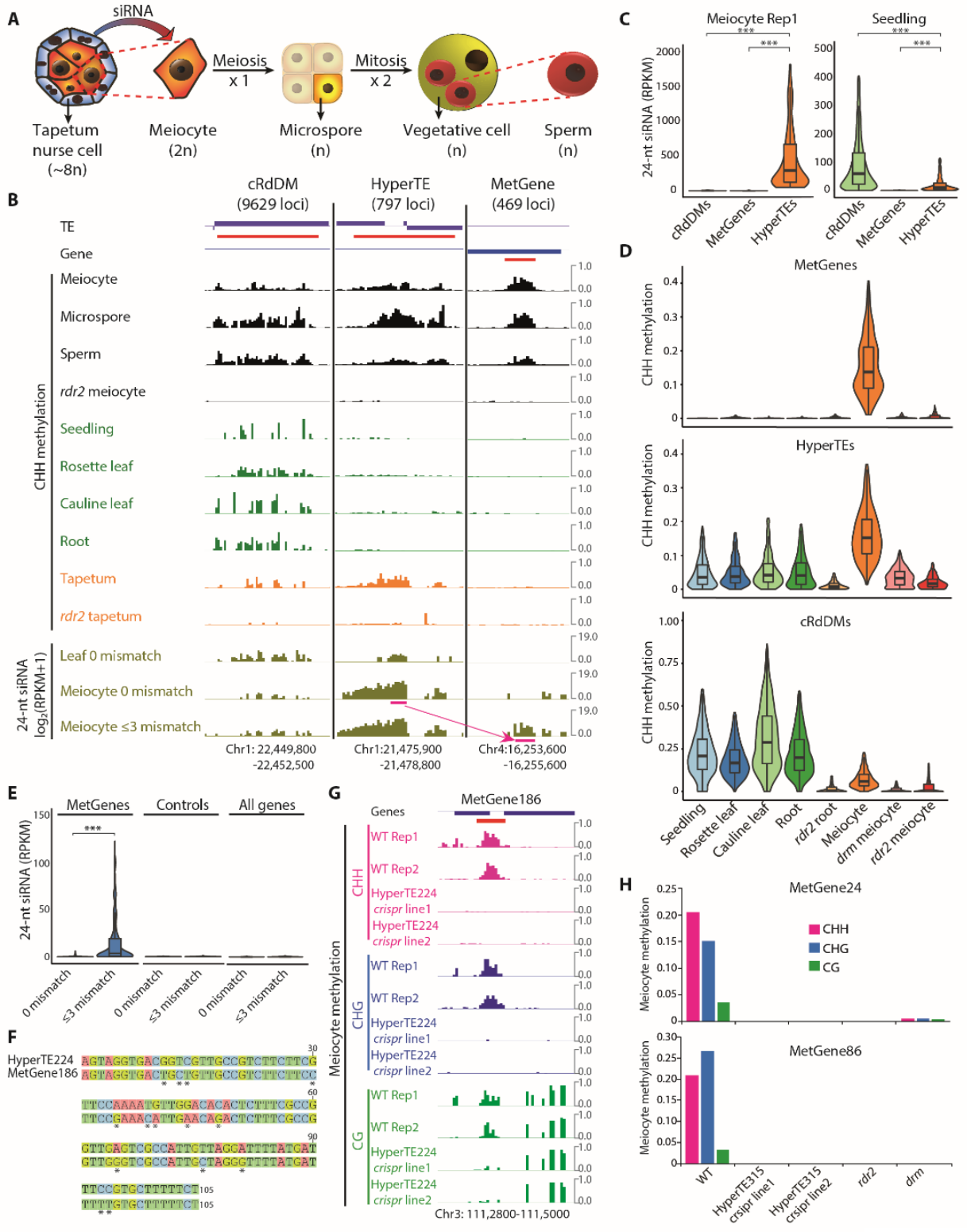

Fig. 1. HyperTE-derived siRNAs drive MetGene methylation in male meiocytes

(**A**) A schematic of male germline development in *Arabidopsis*. (**B**) CHH methylation and 24-nt siRNA abundance (mapped with 0 or ≤3 mismatches) at a canonical RdDM locus (cRdDM), a HyperTE and

a MetGene across tissues and cells as indicated. Each locus is underlined in red. siRNAs that target the MetGene with mismatches originate from the HyperTE, with both source and sink regions underlined in magenta. (**C** and **D**) Levels of 24-nt siRNAs (C) and CHH methylation (D) at cRdDMs, HyperTEs and MetGenes in male meiocytes and indicated somatic tissues. (**E**) Abundance of meiocyte 24-nt siRNAs mapped with 0 or ≤3 mismatches to MetGenes, a control set of random loci of similar sizes, and all genes. (**F**) DNA sequence alignment between HyperTE224 and MetGene186, with mismatches marked by asterisks. (**G**) DNA methylation at MetGene186 (underlined in red) in meiocytes from wild-type (WT) and two independent HyperTE224 deletion lines. (**H**) DNA methylation levels at MetGene24 and 86 in meiocytes from indicated genotypes. *** $P < 2.2e-16$, Kolmogorov-Smirnov test; Rep, biological replicate; RPKM, reads per kilobase per million (C, E).

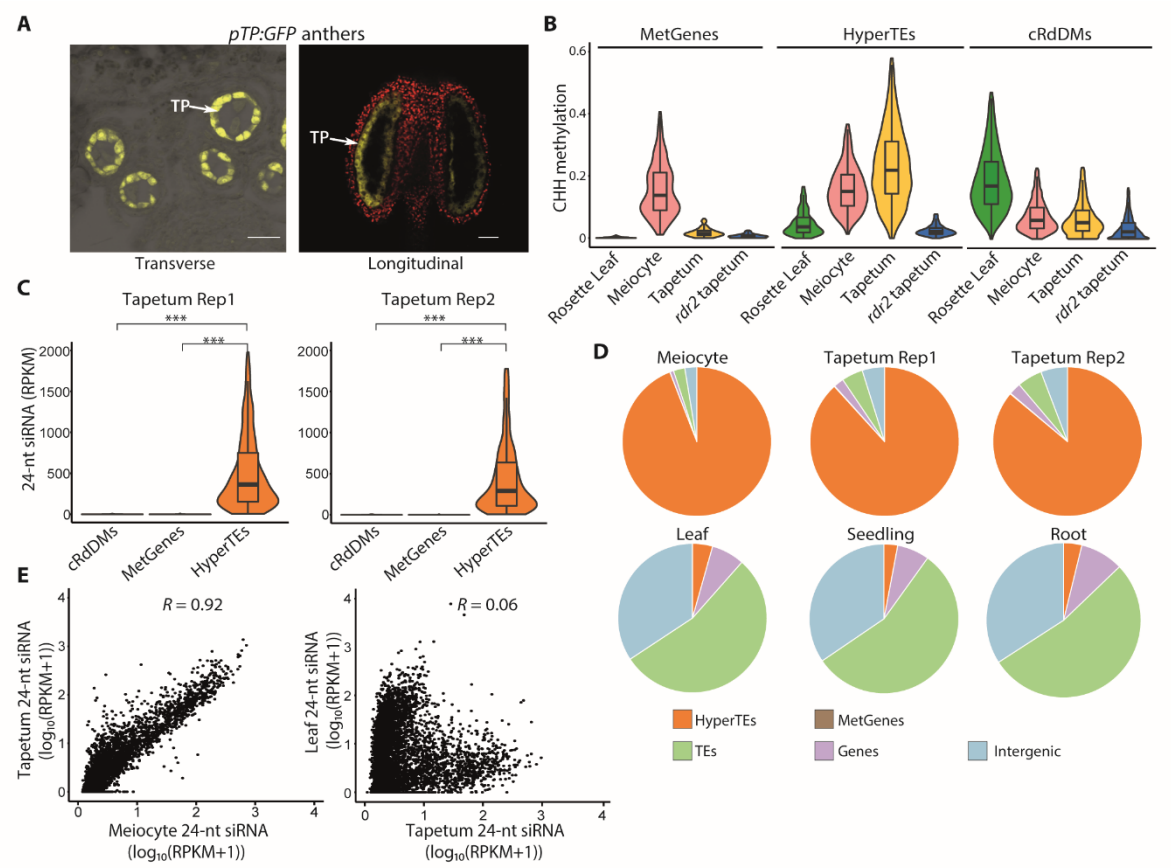

Fig. 2. Tapetal nurse cells have a similar 24-nt siRNA profile to meiocytes

(**A**) Images of *pTP::GFP* anthers. TP, tapetum; yellow, GFP fluorescence; red, auto-fluorescence. Scale bars, 25 μm. (**B**) CHH methylation at MetGenes, HyperTEs and cRdDMs in the leaf, meiocyte and tapetum. (**C**) 24-nt siRNA levels at cRdDMs, MetGenes and HyperTEs in the tapetum (Rep, biological replicate). *** $P < 2.2e\text{-}16$, Kolmogorov-Smirnov test. (**D**) Proportions of 24-nt siRNAs associated with different features in the meiocyte, tapetum, and somatic tissues. (**E**) Correlations (Pearson's *R*) between 24-nt siRNA abundance in the indicated cells/tissues at meiocyte (left; n = 3,712) or tapetum (right; n = 5,841) siRNA clusters.

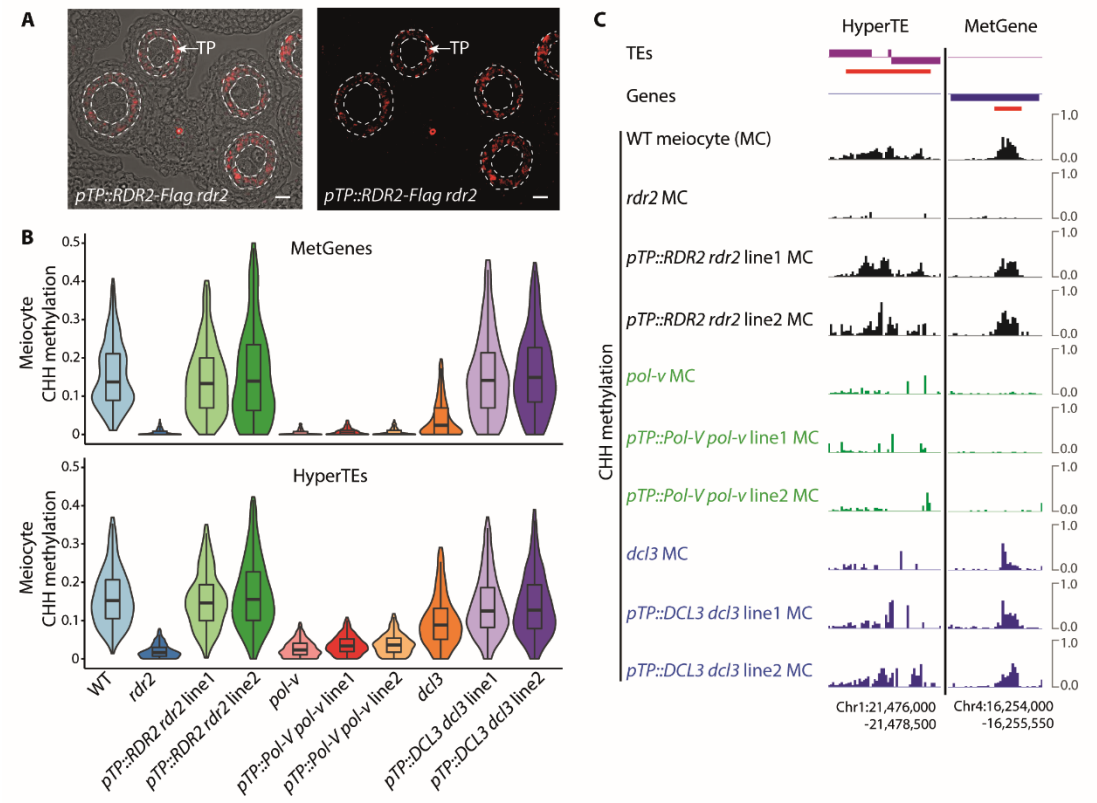

Fig. 3. siRNAs that drive methylation in meiocytes are produced by tapetal cells

(**A**) Immunostaining of a *pTP::RDR2-Flag rdr2* anther. Anti-Flag signal (red) is observed specifically in the tapetum (TP, outlined in white). Scale bars, 10 μm. (**B** and **C**) CHH methylation at MetGenes and HyperTEs (B; examples shown in C, underlined in red) in the meiocytes from WT, RdDM mutants and their corresponding mosaic lines (two independent transgenic lines for each).

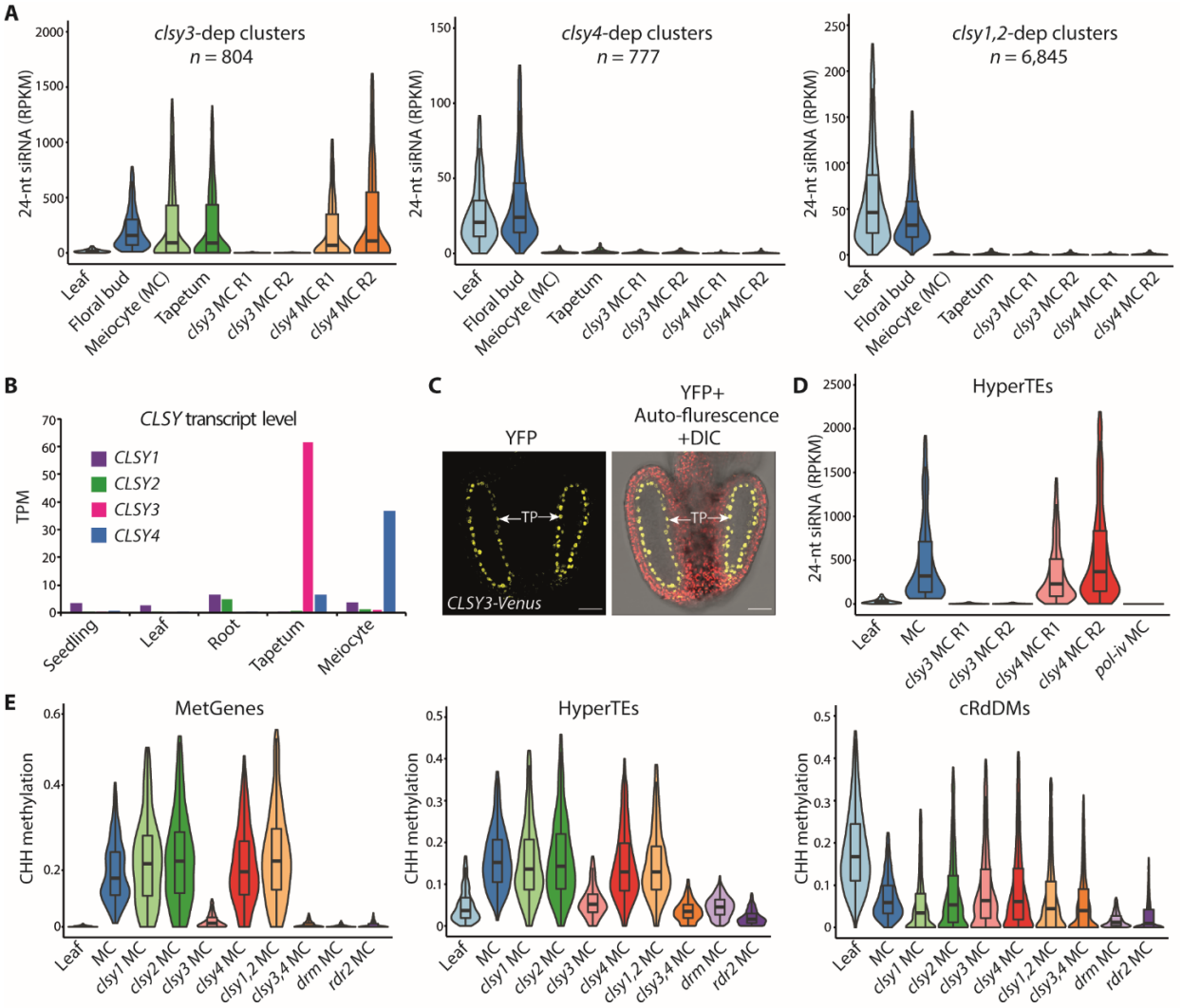

Fig. 4. CLSY3 is required for HyperTE siRNAs

(**A**) 24-nt siRNA abundance at *clsy3*-, *clsy4*-, *clsy1*,*2*- dependent (dep) clusters in indicated tissues and cells. (**B**) Abundance of *CLSY* transcripts across tissues and cells. TPM, transcripts per million. (**C**) CLSY3-Venus localization in the tapetum (TP). Scale bars, 25 µm. (**D**) 24-nt siRNA levels at HyperTEs in the leaf and meiocytes (MC) from indicated genotypes. (**E**) CHH methylation at MetGenes, HyperTEs and cRdDMs in the leaf and MC across genotypes. MC, meiocyte; R, biological replicate (A, D, E).

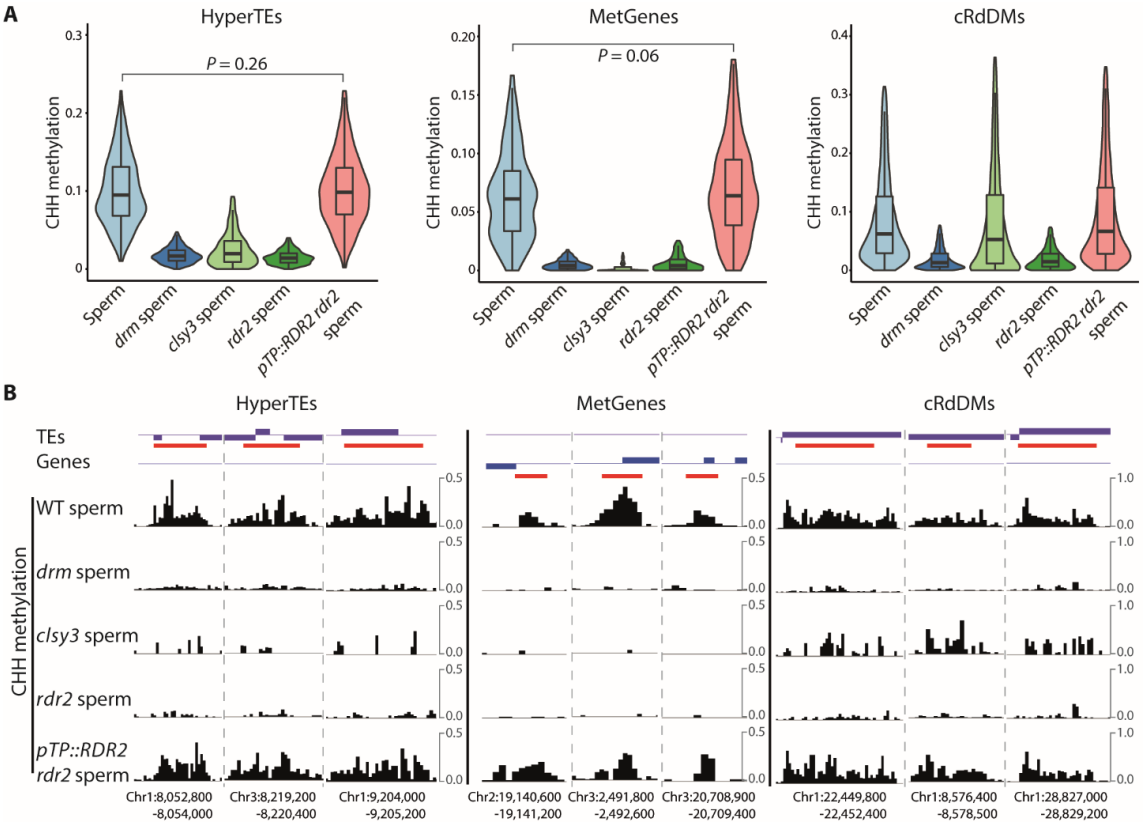

Fig. 5. Tapetal siRNAs determine the sperm DNA methylome

(**A** and **B**) CHH methylation at HyperTEs, MetGenes, and cRdDMs (A; examples displayed in B, underlined in red) in sperm cells from wild-type and indicated genotypes. *P* values, Kolmogorov-Smirnov test.

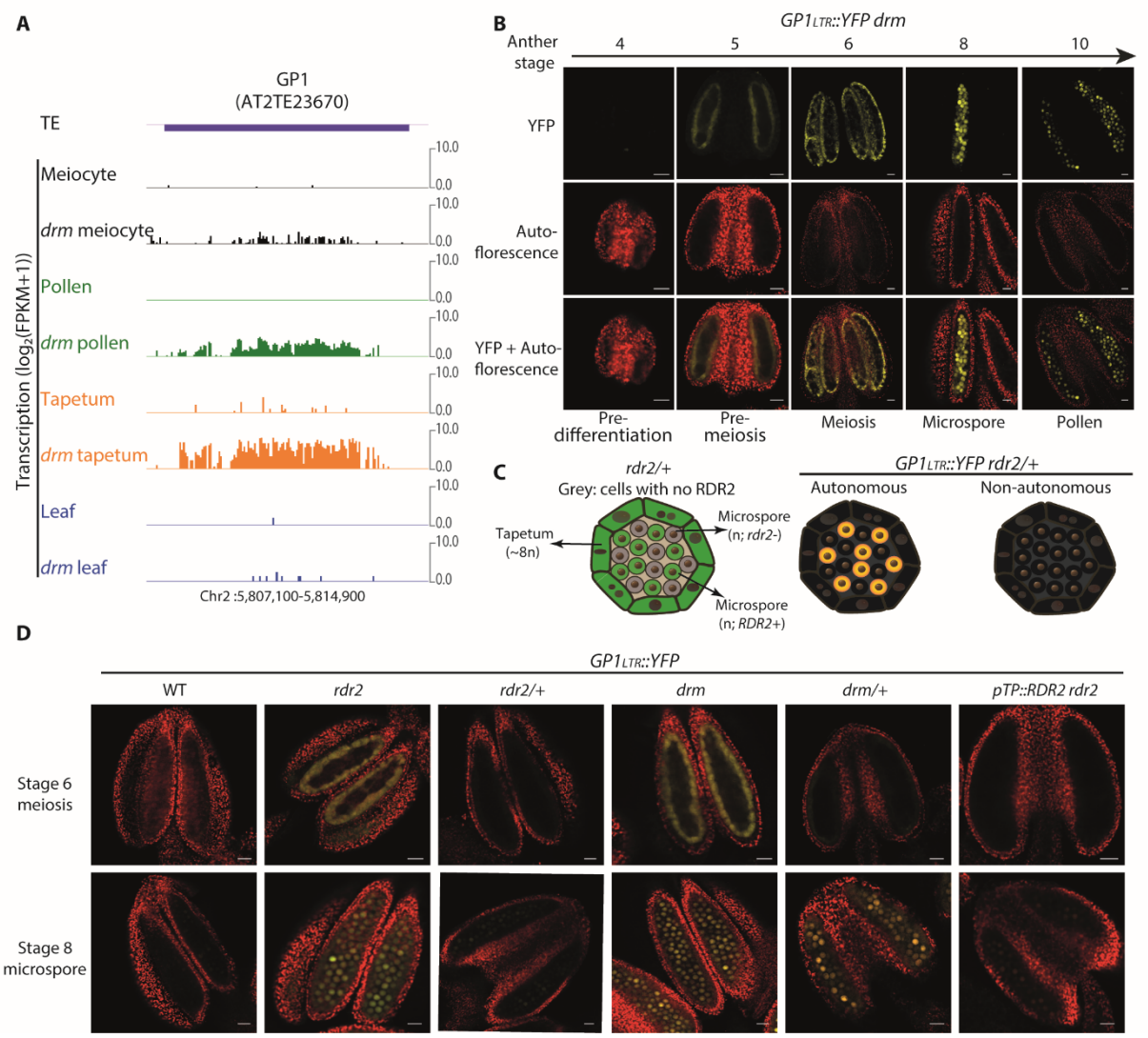

Fig. 6. Tapetal siRNAs silence the *GP1* transposon in the germline

(**A**) *GP1* transcription in indicated cells and tissues. FPKM, fragments per kilobase per million. (**B**) Confocal images of *GP1$_{LTR}$::YFP drm* anthers during development. (**C**) Schematics illustrating RDR2 functionality in *rdr2*/+ heterozygous mutant anther (left) and the expected outcome if *GP1*-silencing siRNA is cell autonomous (middle) or non-autonomous (right). (**D**) *GP1$_{LTR}$::YFP* anthers in indicated genetic backgrounds at stage 6 (upper panels) and 8 (lower panels). Yellow, YFP; red, auto-fluorescence (B, D). Scale bars, 20 μm (B) and 25 μm (D).