

Epigenetic reprogramming rewires transcription during the alternation of generations in *Arabidopsis*

Michael Borg¹, Ranjith K Papareddy¹, Rodolphe Dombey¹, Elin Axelsson¹, Michael D Nodine¹, David Twell^{1,2}, Frédéric Berger^{1*}

¹Gregor Mendel Institute (GMI), Austrian Academy of Sciences, Vienna, Austria;

²Department of Genetics, University of Leicester, Leicester, United Kingdom

Abstract Alternation between morphologically distinct haploid and diploid life forms is a defining feature of most plant and algal life cycles, yet the underlying molecular mechanisms that govern these transitions remain unclear. Here, we explore the dynamic relationship between chromatin accessibility and epigenetic modifications during life form transitions in *Arabidopsis*. The diploid-to-haploid life form transition is governed by the loss of H3K9me2 and DNA demethylation of transposon-associated *cis*-regulatory elements. This event is associated with dramatic changes in chromatin accessibility and transcriptional reprogramming. In contrast, the global loss of H3K27me3 in the haploid form shapes a chromatin accessibility landscape that is poised to re-initiate the transition back to diploid life after fertilisation. Hence, distinct epigenetic reprogramming events rewire transcription through major reorganisation of the regulatory epigenome to guide the alternation of generations in flowering plants.

Introduction

Epigenetic reprogramming refers to the global erasure and remodelling of epigenetic marks during development. DNA and histone methylation are reprogrammed during germline differentiation and after fertilisation to reconfigure transcription in the mammalian embryo (*Hajkova, 2011; Reik et al., 2001*). In plants and algae, meiosis marks the transition from the diploid sporophyte to haploid gamete-producing gametophytes, while the union of gametes at fertilisation re-initiates the transition back to the diploid phase (*Hackenberg and Twell, 2019; Haig, 2008; Hisanaga et al., 2019*). The marked differences in development, morphology, and gene expression between the gametophyte and sporophyte (*Borg et al., 2009; Couceiro et al., 2015; Dickinson and Grant-Downton, 2009; Heesch et al., 2019*) suggest that epigenetic patterns are reprogrammed between these two life forms. Nevertheless, the nature of epigenetic reprogramming and its influence on transcriptional activity during the alternation of generations remain poorly understood.

Ancestral land plants represented by extant bryophytes typically have a dominant haploid gametophytic phase and require Polycomb Repressive Complex 2 (PRC2) activity to repress precocious sporophyte development (*Okano et al., 2009*). DNA methylation also undergoes extensive reprogramming between the haploid and diploid phase in bryophytes, although it remains unclear whether this is required for the transitions between the two life forms (*Ikeda et al., 2018; Schmid et al., 2018; Yaari et al., 2015*). In contrast, flowering plants have a dominant diploid sporophytic phase that culminates with the production of meiotic spores from somatic tissue within the flower. Haploid spores develop into multicellular male or female gametophytes depending on whether they originate in the anther or ovary, respectively (*Dickinson and Grant-Downton, 2009*). The female gametophyte in flowering plants is comprised of an embryo sac containing the egg and accessory cells (*Erbasol Serbes et al., 2019*). Pollen represents the male gametophyte. It develops from

***For correspondence:**

Frederic.berger@gmi.oeaw.ac.at

Competing interests: The authors declare that no competing interests exist.

Funding: See page 18

Received: 07 August 2020

Accepted: 25 January 2021

Published: 25 January 2021

Reviewing editor: Richard Amasino, University of Wisconsin Madison, United States

© Copyright Borg et al. This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

eLife digest Each pollen grain from a flowering plant houses sperm, which contain half of the genes needed to make a new plant, and a companion or vegetative cell (VC) that serves to deliver sperm to the egg. The genes in the vegetative cell and those in the sperm are identical to the genes of the plant they come from, so how can this set of identical genetic information produce such different cells?

Both DNA and histones, the proteins that pack and order DNA, can be chemically modified locally through a process called methylation. The location of these modifications can affect how genetic information in the DNA is read to make different types of cells. The use of processes like methylation to regulate whether genes are switched on or off is called epigenetics. So what role does epigenetics play in plant pollen?

To answer this question, Borg et al. examined the epigenetics of pollen in *Arabidopsis thaliana*, a widely studied plant and common weed. In vegetative cells, DNA methylation is lost together with a different methylation mark (H3K9me2), which unlocks several genes needed for pollen to transport sperm. By contrast, sperm loses an entirely different methylation mark, called H3K27me3, which unlocks a different set of genes that help to prepare development of a new plant once sperm fertilizes the egg. Through these different set of epigenetic changes, activity increases at different groups of genes that are important for shaping the function of each pollen cell type.

These results reveal how the loss of DNA and histone methylation are important for plants to reproduce sexually via pollen. This offers insights into the evolution of plants and other related life forms. Learning about plant reproduction may also help to increase food production by improving crop yields.

asymmetric division of the haploid microspore, which produces a germ cell that becomes engulfed in the cytoplasm of its sister vegetative cell (VC) (**Figure 1A; Borg et al., 2009**). The germ cell represents the male germline initial cell and undergoes a distinct developmental programme to divide and differentiate into two sperms (**Borg et al., 2009**). Sperm differentiation is intricately linked with epigenetic reprogramming of sperm chromatin in *Arabidopsis*. Targeted removal of repressive H3K27me3 marks from histone-based sperm chromatin facilitates the transcription of genes required for sperm differentiation, which are normally silenced by H3K27me3 during sporophytic life (**Borg et al., 2020**).

In contrast to sperm, the events that govern differentiation of the VC remain largely unknown. VC differentiation involves acquiring the competency to grow a pollen tube for the delivery of sperm to the embryo sac, where fertilisation of the egg cell re-initiates diploid sporophyte development (**Hamamura et al., 2012**). Asymmetric microspore division is essential for pollen patterning since chemically induced symmetry results in two daughter cells with VC fate (**Eady et al., 1995**), suggesting that VC specification is the default developmental fate in the male gametophyte. VC nuclear (VN) chromatin organisation is highly diffuse compared to sperm nuclei (SN) and somatic cell nuclei (**Borg and Berger, 2015**). This is caused by the decondensation of pericentromeric heterochromatin, which involves the depletion of linker histone H1 and H3K9me2 (**He et al., 2019; Mérai et al., 2014; Schoft et al., 2009; Slotkin et al., 2009**). Heterochromatin decondensation licenses the active demethylation of pericentromeric DNA sequences by the gametophyte-specific DNA glycosylase DEMETER (DME) (**Andreuzza et al., 2010; He et al., 2019; Ibarra et al., 2012; Schoft et al., 2011**), which stimulates transcription of transposable elements (TEs) and epigenetically activated small RNA (easiRNA) (**Creasey et al., 2014; Slotkin et al., 2009**). Despite the large-scale nature of these epigenetic reconfigurations, it remains unclear if and how these might relate to differentiation of the VC, which ultimately embodies the paternal sporophyte-to-gametophyte transition in flowering plants.

Here, we assess chromatin and transcriptional reprogramming during the sporophyte-to-gametophyte transition in developing *Arabidopsis* pollen. We reveal widespread reprogramming of accessible chromatin across this transition and map hundreds of loci that gain accessibility in each pollen cell type. We demonstrate how the rewiring of *cis*-regulatory activity is intricately linked with the differential reprogramming of repressive epigenetic marks after microspore division. In the VC, accessible chromatin is shaped by the loss of constitutive heterochromatin and active cytosine

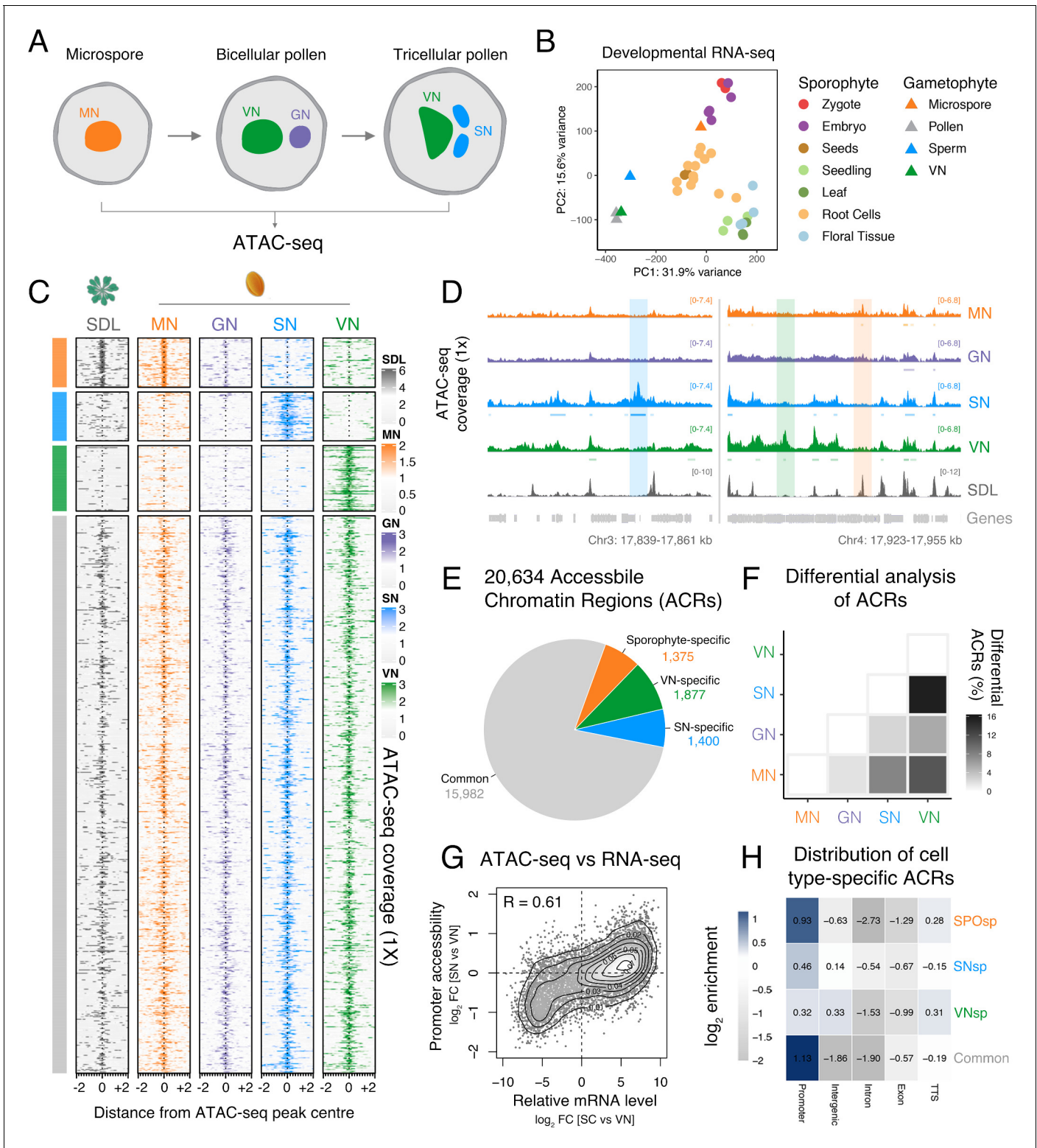


Figure 1. Chromatin accessibility is extensively reprogrammed in pollen. (A) Schematic of Arabidopsis pollen development. Microspores divide asymmetrically to produce a vegetative cell nucleus (VN, green) and germ cell nucleus (GN, purple). While the vegetative cell terminally differentiates, the germ cell divides to produce two sperm nuclei (SN, blue). Details of cellular membranes have been excluded for simplicity. ATAC-seq was performed on FACS-enriched microspores and on fluorescence-activated nuclear sorting (FANS)-isolated GN, SN, and VN. (B) Principal component analysis comparing Arabidopsis RNA-seq transcriptomes from sporophytic and male gametophytic cell types and tissues. Published RNA-seq data sets were reanalysed or described previously (Borg et al., 2020; Hofmann et al., 2019; Nodine and Bartel, 2012; Wang et al., 2020; Zhao et al., 2019). Figure 1 continued on next page

Figure 1 continued

(C) Heatmap centred on pollen accessible chromatin regions (ACRs) comparing seedlings (SDL), microspore nuclei (MN), GN, SN, and VN. Clustering is based on ACR specificity to MN (orange), SN (blue), or VN (green). ACRs in common with seedlings are shown in grey. (D) Genome browser comparison of ATAC-seq tracks in MN, GN, SN, VN, and SDL. Examples of SN-specific (blue), VN-specific (green), and sporophyte and MN-specific (orange) ACRs are indicated with colour shading. ATAC-seq signal is normalised to $1\times$ Arabidopsis genome coverage. (E) Pie chart summary of the specificity of pollen ACRs determined in this study. (F) Heatmap summarising the percentage of differential ACRs detected in pairwise comparisons between the MN, GN, SN, and VN. Differential accessibility was assessed within the 20,634 pollen ACRs shown in panel E. (G) Pairwise correlation of relative chromatin accessibility at promoters and relative mRNA levels of the associated transcribed genes between SN and VN. R represents Pearson's correlation coefficient. (H) Genomic distribution of ATAC-seq peaks in each different group of ACRs. The \log_2 enrichment or depletion is calculated relative to the frequency of genomic features in the Arabidopsis genome.

The online version of this article includes the following figure supplement(s) for figure 1:

Figure supplement 1. Isolation of key cell types across Arabidopsis pollen development.

Figure supplement 2. Transcriptomic profiling performed in this study.

Figure supplement 3. ATAC-seq profiling performed in this study.

Figure supplement 4. ATAC-seq profiling across Arabidopsis pollen development.

demethylation of *cis*-regulatory regions that are predicted to be bound by pollen-expressed transcription factors (TFs) involved in VC differentiation. In contrast, chromatin accessibility in sperm is shaped by the loss of H3K27me₃, resulting in a chromatin state that forecasts the re-initiation of sporophyte development in the next generation. Our findings thus reveal how epigenetic reprogramming orchestrates the rewiring of transcription during haploid–diploid life form transitions in flowering plants.

Results

Pollen cell types have highly distinct transcriptional output

As sperm cells are embedded within the VC cytoplasm, existing Arabidopsis whole pollen transcriptomes represent a composite profile derived from both cell types (Becker *et al.*, 2003; Honys and Twell, 2004). To circumvent this issue, we generated an RNA-seq profile from highly pure fractions of the VN (Figure 1—figure supplement 1A and B) and compared this with the transcriptome of isolated sperm (Figure 1—figure supplement 2A–D; Borg *et al.*, 2020; Borges *et al.*, 2008). Consistent with most pollen RNA deriving from the VC, the VN transcriptome strongly correlated with that of whole pollen but was poorly correlated with that of sperm (Figure 1—figure supplement 2C and D). This supported the VN transcriptome as a suitable proxy for VC-specific gene expression, as shown for nuclear transcriptomes from other cell types (Deal and Henikoff, 2011; Slane *et al.*, 2015). A principal component analysis illustrates how the transcriptomes of two cells constituting mature pollen are highly distinct from each other and from transcriptomes of sporophytic cell types and tissues (Figure 1B). Gene expression in haploid microspores was more related to diploid tissues of the sporophyte than to the two differentiated pollen cell types. This indicates that major transcriptional reprogramming establishes distinct regulatory controls after unequal division of the microspore to specify the sperm and the VC.

Mapping accessible chromatin during pollen development

To assess chromatin reprogramming in the male gametophyte, we profiled chromatin accessibility in the distinct cell types of the developing Arabidopsis pollen lineage using Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq) (Buenrostro *et al.*, 2013; Figure 1A). We isolated highly pure fractions of SN and VN using a combination of cell-type specific markers and fluorescence-activated nuclear sorting (FANS) (Figure 1—figure supplement 1A–D), while microspores were isolated based on their size and auto-fluorescent properties (Figure 1—figure supplement 1E–G; Borges *et al.*, 2012). To isolate germ cell nuclei (GN), we developed a tailored marker line that is active solely in the germ cell and performed FANS with nuclei isolated from developing spores (Figure 1—figure supplement 1H and I). For each cell type, we generated at least two ATAC-seq biological replicates and validated our profiles in several ways (Figure 1—figure supplement 3A). First, biological replicates for each stage were highly correlated, while modest cross-

correlation between tissues indicated distinct chromatin landscapes (**Figure 1—figure supplement 3B and C**). Second, ATAC-seq peaks were highly enriched at promoters (**Figure 1—figure supplement 4A**). Third, highly transcribed genes showed higher accessibility around transcription start sites (TSSs) than genes with low expression (**Figure 1—figure supplement 4B**). Fourth, ATAC-seq profiles in SN and VN had characteristic periodicity arising from polynucleosome fragments (**Figure 1—figure supplement 3D**). Although no periodicity was evident in profiles from microspore nuclei (MN) and GN (**Figure 1—figure supplement 3D**), we observed substantial overlap with profiles from other cell types as well as strong correlation of chromatin accessibility with active genes in the microspore and sperm, respectively (**Figure 1C and D**; **Figure 1—figure supplement 4B**). Our ATAC-seq profiles thus map distinct chromatin accessibility landscapes that span the inception and development of the male gametophyte.

Chromatin accessibility is extensively reprogrammed in pollen

We first examined differences in chromatin accessibility between the microspore, sperm, and vegetative cell. We identified a total of 20,634 ATAC-seq peaks covering over 15 Mb of the Arabidopsis genome (**Figure 1C**; **Supplementary file 1**). We defined peak regions recovered by ATAC-seq as accessible chromatin regions (ACRs). Chromatin accessibility in microspores, differentiated pollen cell types, and seedlings strongly overlapped (15,982 peaks, 77.5%) (**Figure 1C–E**). These common peaks were associated with genes involved in core housekeeping functions that maintain expression across the entire life cycle, such as mRNA metabolism, protein transport, cell cycle control, and osmotic stress response (**Figure 1—figure supplement 4C**). A set of sporophyte-specific ACRs (6.7%; 1375 peaks) remained accessible between seedlings and MN, had reduced accessibility in the GN, but later lost accessibility in the SN and VN (**Figure 1C**). These occurred at genes involved in photosynthesis and energy metabolism (**Figure 1—figure supplement 4C**), which are dispensable in the short-lived haploid gametophyte since it derives resources from the host parental sporophyte. Our analysis did not reveal *de novo* gains in chromatin accessibility in GN compared with MN but rather a gradual loss in accessibility among sporophytic-specific ACRs (**Figure 1C**).

Pairwise comparisons between the pollen cell types revealed a gradual increase in differential accessibility across pollen development (**Figure 1F**; **Figure 1—figure supplement 3E**; **Supplementary file 1**). These differences were most pronounced at the mature pollen stage, with 6.8% (1400 peaks) and 9.1% (1877 peaks) cell type-specific ACRs unique to SN or VN, respectively (**Figure 1C–E**; **Supplementary file 1**). Heatmaps centred on ATAC-seq peaks illustrate the cell type-specificity of SN-specific and VN-specific ACRs (**Figure 1C**). VN-specific ACRs were strongly associated with functions unique to pollen biology such as pollination, cell tip growth, and morphogenesis (**Figure 1—figure supplement 4C**). In sperm, we observed high promoter accessibility at genes directly regulated by DUO POLLEN 1 (DUO1), an essential R2R3-type MYB TF that controls sperm differentiation (**Figure 1—figure supplement 4D**; **Borg et al., 2011**; **Brownfield et al., 2009**; **Rotman et al., 2005**). Importantly, the relative accessibility of promoter regions was positively correlated ($R = 0.61$) with differential expression of the associated genes between sperm and the VN (**Figure 1G**), consistent with the distinct transcriptional state of each cell type. SN-specific and VN-specific ACRs were less prevalent at promoters but had an increased incidence within distal intergenic regions compared with common ACRs (**Figure 1H**). Transcribed genes most proximal to these intergenic ACRs were differentially expressed in a cell type-specific manner (**Figure 1—figure supplement 4E**), suggesting distal regulatory influence. VN-specific ACRs were also highly enriched at transcriptional termination sites (TTSs) compared to other ACRs (**Figure 1H**; **Figure 1—figure supplement 4B**). In summary, ATAC-seq has revealed hundreds of unique ACRs specific to the sperm or VC, which likely encompass putative *cis*-regulatory elements that modulate transcription and differentiation of each pollen cell type. These *cis*-regulatory regions are distributed between proximal promoters and distal regions, suggesting the involvement of long-range chromatin interactions in the regulation of pollen gene expression.

Pericentromeric sequences gain chromatin accessibility in the male gametophyte

In Arabidopsis, condensed heterochromatin is organised into chromocenters and is comprised of pericentromeric sequences enriched in H3K9me2 and DNA methylation (**Feng and Michaels, 2015**).

These repressive epigenetic marks silence TEs that accumulate within pericentromeric regions (Feng and Michaels, 2015). Because pericentromeric heterochromatin is lost in the VN (Schoft et al., 2009), we explored how this might impact the landscape of accessible chromatin. ATAC-seq peak density was reduced within pericentromeric regions compared to the arms of each of the five Arabidopsis chromosomes (Figure 2A). We also noted a slight but consistent increase in ATAC-seq peak density within pericentromeric regions of the VN compared to SN (Figure 2A; Figure 2—figure supplement 1A). Consistent with this, and unlike common or SN-specific ACRs, VN-specific ACRs were relatively enriched for H3K9me2 in somatic tissues but depleted for active H3K4me3 and H3K27ac histone marks, with just over one-third (36.5%; 686 peaks) marked by H3K9me2 in leaves (Figure 2B and C; Figure 2—figure supplement 1B). Similarly, almost one-third of VN-specific ACRs significantly overlapped with a TE insertion (Figure 2D and E), with DNA/MuDR and LINE/L1 elements notably over-represented (Figure 2F). Accordingly, TEs specifically activated in pollen showed increased promoter accessibility specifically in the VN (Figure 2G; He et al., 2019). VN-specific ACRs were also significantly associated with regions that produce pollen siRNAs, including 21–22-nt class easiRNAs known to preferentially accumulate in pollen (Figure 2C and D; Borges et al., 2018; Martinez et al., 2018; Slotkin et al., 2009). These observations highlight how the VN gains chromatin accessibility within genomic regions normally silenced in sporophytic tissues, including TE-rich pericentromeric regions that lose constitutive heterochromatin, undergo transcription, and produce small non-coding RNAs.

DME demethylates predicted TF binding sites in ACRs of the VN

The loss of constitutive heterochromatin in the VN is accompanied by selective DNA demethylation by DME, a DNA glycosylase that demethylates cytosines at a subset of genes flanked by transposons and repeats (Choi et al., 2002; Schoft et al., 2011). We thus explored how DME-dependent demethylation might impact chromatin accessibility in the VN. We first performed differential methylation analysis using published DNA methylomes to delimit hypermethylated cytosines in the VN of *dme-2/+* mutants compared to wild type (WT) (Figure 3—figure supplement 1A–E; Supplementary file 2; Ibarra et al., 2012). Strikingly, VN-specific ACRs were significantly enriched for regions demethylated by DME in the VN (Figure 2C and D), with just over half of all VN-specific ACRs (982 peaks; 52.3%) containing DME-dependent hypermethylated cytosines (Figure 3A; Supplementary file 3). A small fraction of common ACRs were also targeted for demethylation (2052 peaks; 12.8%) (Figure 3B; Supplementary file 3). Metaplots centred over DME-targeted ACRs illustrate the demethylation mediated by DME in all cytosine contexts of the VN (Figure 3D), which is mirrored by higher levels of chromatin accessibility in the VN than in SN (Figure 3C). Consistent with the VN-specific expression and activity of DME (Park et al., 2017), the levels of DNA methylation remained unchanged over these regions in WT and *dme-2/+* sperm. DME activity is thus unlikely to stimulate a gain in chromatin accessibility alone, at least at common ACRs, but rather modulates differential DNA methylation levels within these shared open chromatin sites. This suggests that chromatin accessibility likely precedes DME activity, consistent with the requirement of histone H1 depletion for DNA demethylation in the VN (He et al., 2019).

To further probe the role of DME, we performed motif enrichment analysis to determine which TFs are likely to be bound within DME-targeted ACRs. We used a set of experimentally derived DNA binding motifs generated by DAP-seq, an in vitro binding assay of the sequence specificity of TF-genomic DNA interactions (O'Malley et al., 2016). Fifty TFs expressed in the VN were present in the DAP-seq database, of which two-thirds (36 of 50; 72.0%) had significantly enriched motifs within DME-targeted ACRs (Figure 3E). These included motifs bound by MYB101 (Figure 3E), a pollen-specific TF that acts redundantly with MYB97 and MYB120 to regulate pollen tube growth (Leydon et al., 2013; Liang et al., 2013). The binding affinity of most TFs in Arabidopsis is strongly inhibited by DNA methylation, which was determined by performing the DAP-seq assay with methylated (DAP) or non-methylated (ampDAP) DNA (O'Malley et al., 2016). Of the 36 TFs predicted to bind within DME-targeted ACRs, 25 had been assessed in this manner (O'Malley et al., 2016), half of which (14 of 25; 56.0%) were inhibited by DNA methylation (Figure 3F). This suggests that DNA demethylation by DME facilitates the binding of several methylation-sensitive TFs within DME-targeted ACRs. Consistently, the predicted binding sites for these methylation-sensitive TFs were substantially hypo-methylated in VN from WT pollen (Figure 3G–J; Figure 3—figure supplement 2A–J). In *dme-2/+* mutant pollen, these predicted binding sites were hyper-methylated to around half of

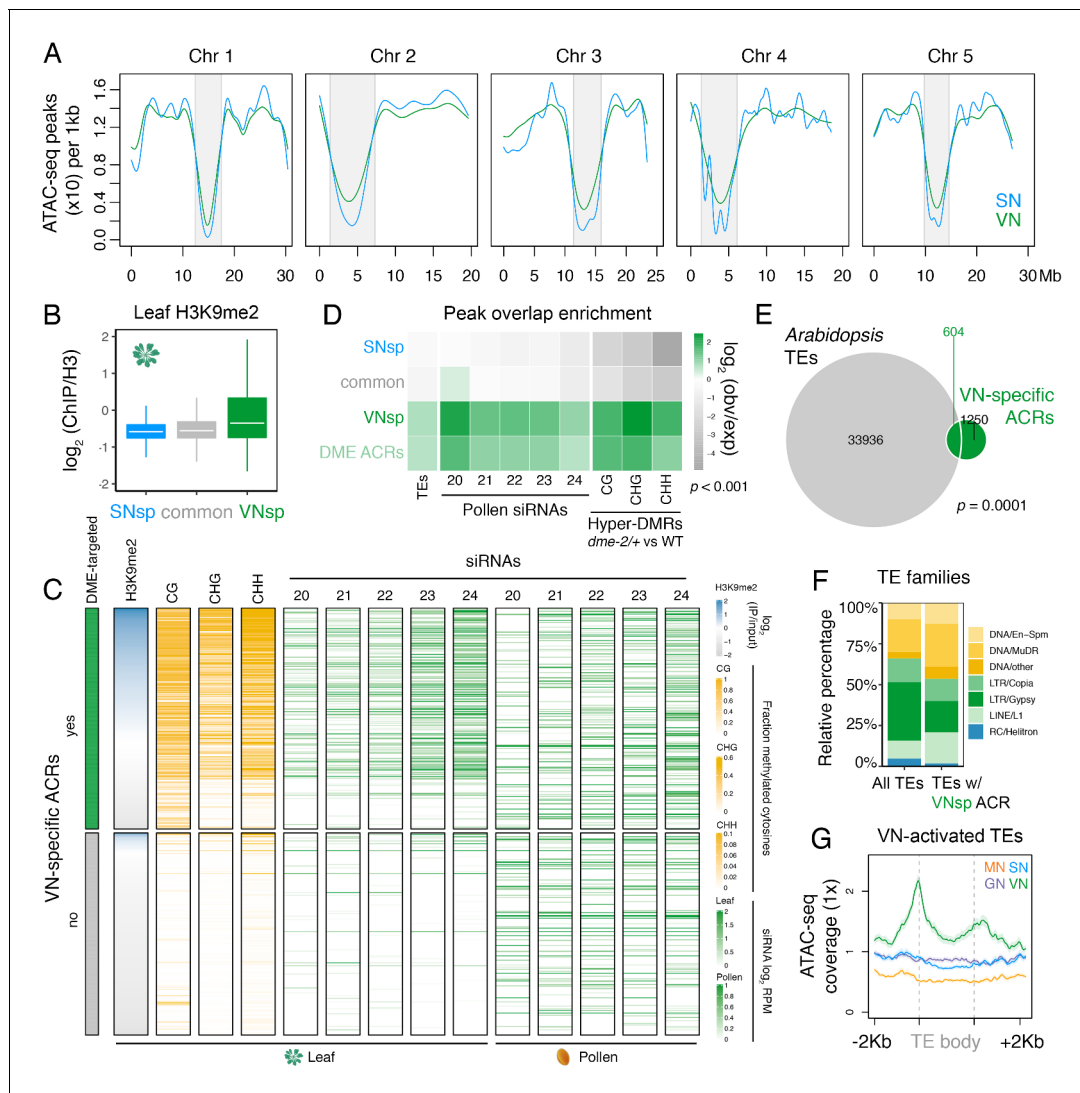


Figure 2. Pericentromeric sequences gain chromatin accessibility in the VN. **(A)** Distribution of accessible chromatin region (ACR) density over the five Arabidopsis chromosomes calculated in 10 kb bins. Pericentromeric regions are indicated with grey shading. **(B)** Levels of H3K9me2 marks in leaf tissue (log₂ ChIP-seq enrichment relative to H3) (*Baerenfaller et al., 2016*) at SN-specific, VN-specific, and common ACRs. Each boxplot indicates minimum and maximum values as well as 25th, 50th, and 75th quartiles. **(C)** Heatmap summarising the epigenetic state of VN-specific ACRs. Shown are the log₂ enrichment of H3K9me2 (relative to H3), the proportion of CG, CHH, and CHG methylation (*Stroud et al., 2013*) and different size classes of short-interfering RNA (siRNAs) in leaves (*Papareddy et al., 2020*) alongside pollen siRNAs (*Borges et al., 2018*). VN-specific ACRs are grouped by the presence or absence of cytosines demethylated by DEMETER (DME) in the VN. **(D)** Heatmap summarising the overlap enrichment of SN-specific, VN-specific, DME-targeted, and common ACRs with TEs, pollen siRNAs (*Slotkin et al., 2009*), and *dme-2/+* hyper-DMRs. Fold changes were determined using hypergeometric tests compared with random Arabidopsis genomic regions ($n = 10,000,000$ permutations). **(E)** Pairwise overlap between VN-specific ACRs and Arabidopsis transposable elements (TEs). Significance of the enriched overlap (p-value) was determined in the hypergeometric test shown in panel C. **(F)** Distribution of TE gene classes associated with VN-specific ACRs alongside the relative frequency in the Arabidopsis genome. **(G)** Averaged ATAC-seq enrichment over TE genes specifically re-activated in the VN. Plotted is the ATAC-seq signal normalised to 1 × Arabidopsis genome coverage.

The online version of this article includes the following figure supplement(s) for figure 2:

Figure supplement 1. DEMETER (DME) demethylates accessible chromatin uniformly across the VN.

the levels in WT sperm, consistent with the heterozygous nature of the *dme-2* mutant (*Figure 3G–J*; *Figure 3—figure supplement 2A–J*). These observations further delimit DME-mediated demethylation to motifs targeted by TFs expressed in the VN, which include several that are unable to bind efficiently in the presence of DNA methylation. In summary, we propose that a substantial portion of genomic regions that become preferentially or specifically accessible in the VN undergo DME-

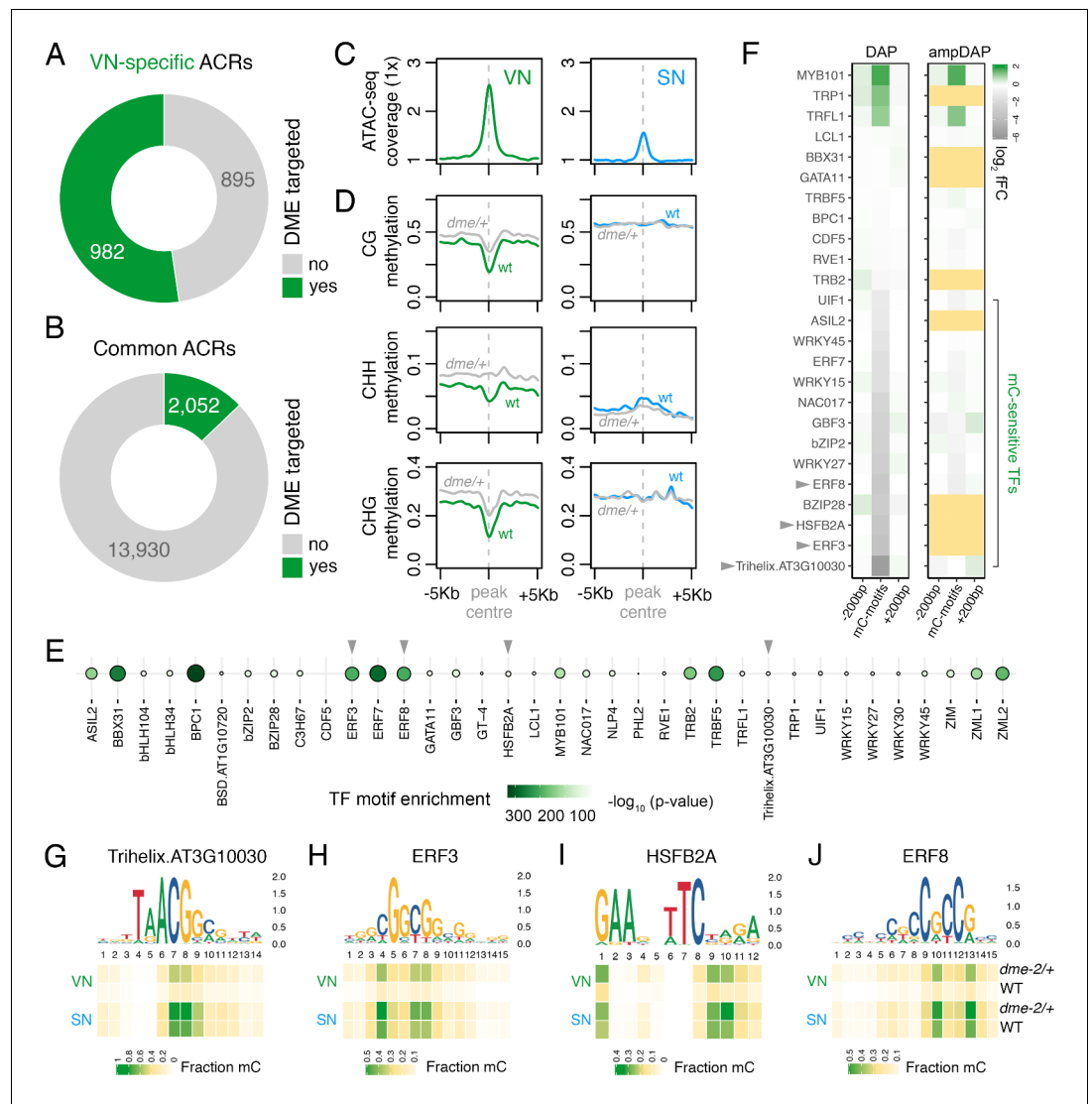


Figure 3. DEMETER (DME) demethylates predicted TF binding sites in accessible chromatin. **(A and B)** Relative proportion of VN-specific accessible chromatin regions (ACRs) **(A)** and common ACRs **(B)** that overlap with cytosines demethylated by DME in the VN. **(C and D)** Averaged metaplots comparing ATAC-seq enrichment **(C)** and the proportion of CG, CHH, and CHG methylation **(D)** in sperm (blue) and the VN (green) over DME-targeted ACRs. Regions with at least five differentially methylated cytosines were plotted. Coloured lines represent wild-type (WT) levels while grey lines indicate the level in nuclei isolated from *dme-2/+* mutant pollen. **(E)** Motifs of VN-expressed TFs enriched within DME-targeted ACRs. Significance (p-value) was assessed using a two-sided Fisher's exact test. TFs shown in panels F–I are marked with grey triangles. **(F)** Binding preference for the VN-expressed TF motifs enriched within DME-targeted ACRs. Left: DAP-seq binding fold-change at methylated motifs (mC-motifs) relative to neighbouring motifs (± 200 bp). Right: ampDAP-seq binding fold-change at the same mC-motifs without DNA methylation. Yellow boxes indicate TFs without a score due to too few motifs. Data is from previously published DAP-seq experiments (O'Malley et al., 2016). TFs shown in panels F–I are marked with grey triangles. **(G–J)** DNA methylation level at predicted binding sites within DME-targeted ACRs for the four highest methylation-sensitive TFs shown in panel E – Trihelix.AT3G10030 **(G)**, ERF3 **(H)**, HSF2A **(I)**, and ERF8 **(J)**. Each heatmap summarises the proportion of methylated cytosines over predicted binding sites in VN and SN of WT and *dme-2/+* pollen. Single nucleotides are indicated by the corresponding position frequency matrix logo plotted above each heatmap.

The online version of this article includes the following figure supplement(s) for figure 3:

Figure supplement 1. Differential methylated region (DMR) analysis between VN from wild-type (WT) and *dme-2/+* mutant pollen.

Figure supplement 2. DEMETER (DME) demethylates predicted TF binding sites specifically in the VN.

mediated DNA demethylation, which likely licenses the binding of several VN-expressed TFs involved in regulating the sporophyte-to-gametophyte transition.

DNA demethylation by DME regulates the expression of pollen-tube related genes

DME-targeted ACRs clustered into two groups distinguished by a relative enrichment or depletion of H3K9me₂, H3K27me₁, and histone H1 in sporophytic tissue (**Figure 2—figure supplement 1C**). DME activity in open chromatin thus targets the genome uniformly in both pericentromeric regions and chromosome arms (**Figure 2—figure supplement 1D**). The 3034 DME-targeted ACRs were strongly enriched for TEs and for regions that produce siRNAs (**Figure 2C and D**), with around one-fifth (661 peaks, 21.8%) lying distally within intergenic regions (**Figure 4A; Supplementary file 3**). However, the majority of DME-targeted ACRs (2373 peaks, 78.2.8%) lied in the vicinity of protein-coding genes (**Figure 4A; Supplementary file 3**), which were significantly over-represented amongst VN-expressed genes (**Figure 4B**). Importantly, we observed a small but highly significant overlap between genes with a DME-targeted ACR and genes differentially expressed in *dme-2/+* mutant pollen (**Figure 4C and D; Figure 2—figure supplement 1E**), confirming that DME directly modulates gene expression during pollen development. These 27 genes included protein kinases, cysteine-rich peptides, metabolic proteins, as well as transcription factors, gene functions that are all known to be involved in pollen tube growth (**Figure 4D; Supplementary file 4; Supplementary file 5; Higashiyama and Takeuchi, 2015; Qu et al., 2015**).

The most significantly downregulated gene in *dme-2/+* pollen was a pollen-specific cysteine-rich RECEPTOR-LIKE SERINE/THREONINE KINASE (*RFK2*) (**Figure 4C**), which is embedded within H3K9me₂-marked heterochromatin in the sporophyte (**Figure 4E; Figure 2—figure supplement 1C**). The upstream promoter region of *RFK2* was heavily hypomethylated in VN from WT pollen and contained several DME-demethylated cytosines including at predicted binding sites for bZIP2, ERF3, and ERF8 (**Figure 4E**), three DNA methylation-sensitive TFs (**Figure 3F**). Similarly, the downregulated AT2G16030 locus, which encodes a pollen-specific SAM-dependent methyltransferase also marked with sporophytic H3K9me₂ (**Figure 2—figure supplement 1C**), contained several DME-demethylated cytosines at the predicted binding sites of ERF3, ERF7, ERF8 and a BSD domain TF (**Figure 4F**). Interestingly, DME-targeted ACRs were significantly associated with genes upregulated in growing pollen tubes, which occurs upon interaction of semi in vivo grown pollen tubes (SIVPT) with female pistil tissue (**Figure 4B**, SIVPT-induced) (**Leydon et al., 2017; Qin et al., 2009**). These 352 genes encode several examples of well-characterised pollen-specific receptor like-kinases, calcium-dependent protein kinases, cation/H⁺ exchanger proteins, and RALF-like signalling peptides (**Figure 2—figure supplement 1C**, grey labels) (**Myers et al., 2009; Sze et al., 2004; Takeuchi and Higashiyama, 2016**), around a quarter of which had little to no transcripts detectable in dry pollen (**Figure 2—figure supplement 1F**). Thus, DME demethylation acts to poise *cis*-regulatory regions for the onset of events taking place after pollination to facilitate the *de novo* wave of gene expression that occurs in the growing pollen tube (**Leydon et al., 2017; Qin et al., 2009**). In conclusion, chromatin accessibility in the VN strongly reflects the loss of constitutive heterochromatin, which facilitates DME-mediated demethylation of *cis*-regulatory elements predicted to be bound by methylation-sensitive TFs. VC differentiation is associated with these *cis*-regulatory regions, many of which are embedded within TE-rich regions of constitutive heterochromatin and inaccessible to transcription during sporophytic life.

Accessible chromatin in sperm forecasts the sporophytic transition

In Arabidopsis, sperm chromatin is epigenetically reprogrammed through the global loss of H3K27me₃, which involves the sperm-specific deposition of histone H3.10 and active demethylation by Jumonji-C family H3K27 demethylases *EARLY FLOWERING 6* (*ELF6*), *RELATIVE OF ELF6* (*REF6*), and *JUMONJI 13* (*JMJ13*) (**Borg et al., 2020**). We thus explored how H3K27me₃ reprogramming might impact chromatin accessibility in sperm. We previously showed that the global loss of H3K27me₃ in sperm is accompanied by selective priming with or without H3K4me₃ (**Borg et al., 2020; Figure 5A**). Consistently, H3K4me₃-primed genes were marked by increased chromatin accessibility in their flanking promoter region and were enriched for SN-specific ACRs (**Figure 5A and B**, primed cluster). SN-specific ACRs were depleted for active marks in sporophytic

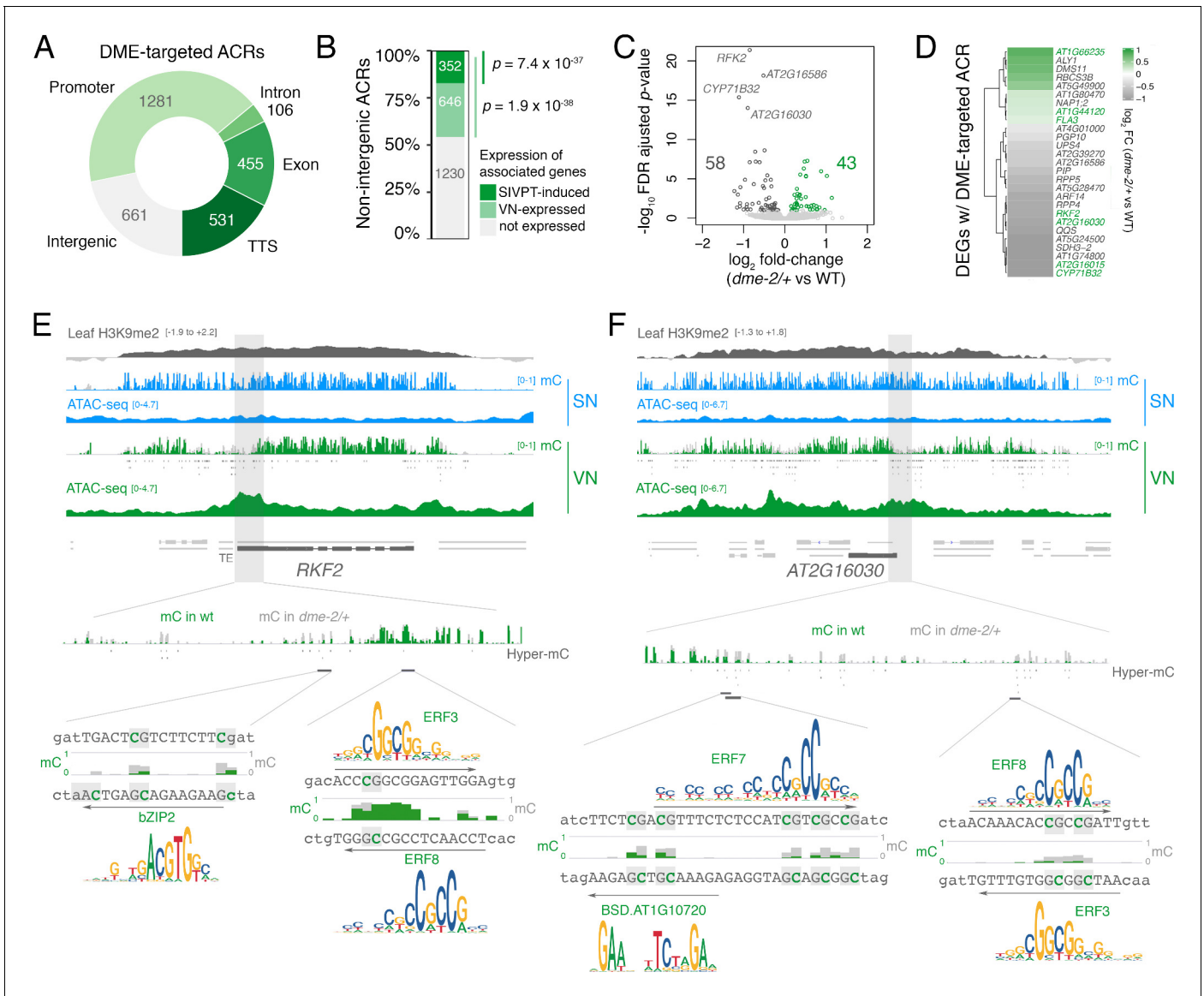


Figure 4. DEMETER (DME) activity regulates the expression of pollen tube-related genes. **(A)** Distribution of genomic features associated with DME-targeted accessible chromatin regions (ACRs). **(B)** Expression status of genes lying in the vicinity (<900 bp) of a DME-targeted ACR. Genes were classified as being expressed in the VN (TPM >1, light green) or by being upregulated during semi in vivo pollen tube (SIVPT) growth (dark green) (Leydon et al., 2017). Statistical enrichment of the overlap was determined using pairwise two-sided Fisher’s exact tests. **(C)** Volcano plot of differentially expressed genes (DEGs) in *dme-2/+* pollen compared to wild type (WT). Upregulated genes are shown in green while downregulated genes are shown in dark grey. DEGs were defined as having an FDR-adjusted p-value <0.1. The top four most significantly downregulated genes are labelled. **(D)** Heatmap illustrating the magnitude of expression changes for DME-dependent DEGs directly associated with a DME-targeted ACR. Plotted is the log₂ fold-change in *dme-2/+* pollen compared to WT. VN-specific genes are highlighted in green. **(E and F)** Browser view of two VN-specific direct DME target genes – *RECEPTOR-LIKE SERINE/THREONINE KINASE (RFK2)* (E) and the *AT2G16030* locus encoding a SAM-dependent methyltransferase (F). Both genes are marked with H3K9me2 in the sporophyte (log₂ ChIP-seq enrichment relative to H3). Tracks illustrate ATAC-seq signals and the proportion of methylated cytosines in all contexts (mC) in SN (blue) and VN (green), with DNA methylation levels in *dme-2/+* overlaid in grey. DME-dependent hyper-methylated cytosines (HMCs) are marked below. Shading indicates the flanking promoter region shown in the close-up view. Sequences surrounding predicted TF binding sites are shown, with the level of cytosine methylation at HMCs highlighted in green and grey for WT and *dme-2/+*, respectively. Arrows indicate the orientation and position of TF motifs with a corresponding position frequency matrix logo plotted above.

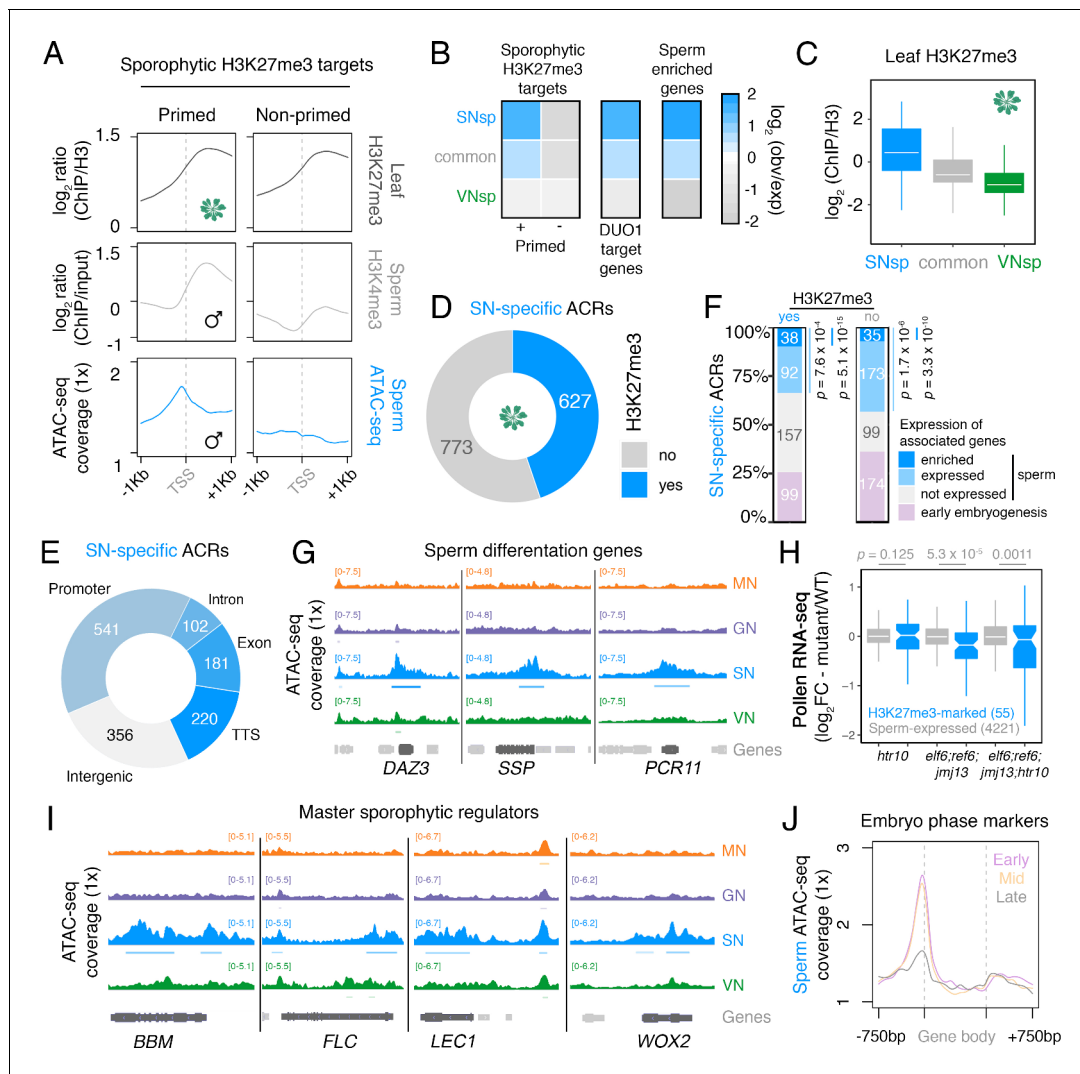


Figure 5. Chromatin accessibility in sperm forecasts sporophytic development. (A) Sperm ATAC-seq profiles reflect the reprogrammed state of sporophytic H3K27me3 target genes. Plotted are averaged levels of leaf H3K27me3 and sperm H3K4me3 (\log_2 ChIP-seq enrichment relative to control) together with sperm ATAC-seq signals. ChIP-seq data and clustering for the presence and absence of H3K4me3 enrichment in sperm were described previously (Borg et al., 2020). (B) Heatmap summarising the overlap enrichment of SN-specific, VN-specific, and common accessible chromatin regions (ACRs) with somatic Polycomb target genes, DUO1 target genes, and sperm-enriched genes. Fold-enrichment was determined using hypergeometric tests compared with random Arabidopsis genomic regions ($n = 10,000,000$ permutations). (C) Levels of H3K27me3 marks in leaf tissue (\log_2 ChIP-seq enrichment relative to H3) at SN-specific, VN-specific, and common ACRs. Each boxplot indicates minimum and maximum values as well as 25th, 50th, and 75th quartiles. (D) Relative proportion of SN-specific ACRs that overlap with a somatic H3K27me3 domain. (E) Distribution of genomic features associated with SN-specific ACRs. (F) Expression status of genes in the vicinity (<900 bp) of an SN-specific ACR with or without H3K27me3 in the sporophyte. Genes were classified as being expressed in sperm (TPM >1, light blue), by having enriched expression in sperm (dark blue), or by expression in the early zygote or embryos but not sperm (pink). Non-expressed genes in sperm are shown in grey. TPM >1 or <1 was used as a threshold for expressed and non-expressed genes, respectively. Statistical enrichment was determined using pairwise two-sided Fisher's exact tests. (G) Chromatin accessibility at three sperm-specific genes associated with sperm differentiation – *DUO1-ACTIVATED ZINC FINGER 1 (DAZ1)*, *SHORT SUSPENSOR 1 (SSP1)* and *PLANT CADMIUM RESISTANCE 11 (PCR11)*. Tracks represent the ATAC-seq signal normalised to $1 \times$ Arabidopsis genome coverage for nuclei of the microspore (MN), germ cell (GN), sperm (SN), and vegetative cell (VN). (H) Differential expression between *htr10*, *elf6;ref6*; *jmj13*, and *elf6;ref6*; *jmj13*; *htr10* pollen relative to wild type (WT). Each boxplot indicates the minimum and maximum values as well as 25th, 50th, and 75th quartiles. Sample size (n) of genes associated with an H3K27me3-repressed ACR (blue) and all sperm-expressed genes (grey) is shown. Plot is restricted to genes with >10 counts in at least one RNA-seq sample and/or replicate. Statistical analysis was performed using two-sided Kolmogorov-Smirnov tests. (I) Chromatin accessibility at four major developmental regulators transcribed during early sporophyte development – *BABY BOOM (BBM)*, *FLOWERING LOCUS C (FLC)*, *LEAFY COTYLEDON 1 (LEC1)*, and *WUSCHEL-RELATED HOMEBOX 2 (WOX2)*. Tracks represent the ATAC-seq signal normalised to $1 \times$ Arabidopsis genome coverage for nuclei of the microspore (MN), germ cell (GN), sperm (SN), and vegetative cell (VN). (J) Averaged ATAC-seq enrichment over genes with enriched expression during early (pink), mid (yellow), and late (grey) embryogenesis, which represent

Figure 5 continued on next page

Figure 5 continued

the pre-cotyledon phase, the transition phase, and the mature green phase, respectively (Hofmann et al., 2019). Plotted is the sperm ATAC-seq signal normalised to 1 × Arabidopsis genome coverage.

leaf tissue but highly enriched for H3K27me3 (Figure 5C; Figure 2—figure supplement 1B). Almost one half of the SN-specific ACRs (627; 44.8%) overlapped a sporophytic H3K27me3 domain (Figure 5D; Supplementary file 6). The remaining SN-specific ACRs (773; 55.2%) lacked H3K27me3 in the sporophyte (Figure 5D; Supplementary file 6), suggesting sporophytic repression by an unknown pathway.

The majority of SN-specific ACRs (1044 peaks; 74.6%) were associated with a protein-coding gene (Figure 5E), around one-third of which were significantly over-represented for sperm-expressed genes (Figure 5F). These included several direct DUO1 target genes, including *DUO1-ACTIVATED ZINC FINGER (DAZ3)* and *PLANT CADMIUM RESISTANCE 11 (PCR11)*, as well as the paternally transcribed embryonic regulator *SHORT SUSPENSOR (SSP)* (Figure 5G; Bayer et al., 2009; Borg et al., 2011). The genes neighbouring H3K27me3-marked ACRs were significantly downregulated upon retention of H3K27me3, which occurs in mutant sperm compromised for the pathways that remove H3K27me3 (Figure 5H; Borg et al., 2020). Incidentally, H3K27me3 in GN remains comparable to levels in microspores but is specifically lost in divided sperm (Borg et al., 2020), further supporting our observation that major reorganisation of chromatin accessibility in sperm coincides with H3K27me3 reprogramming. These results reveal how H3K27me3 reprogramming exposes *cis*-regulatory elements to the transcriptional machinery during sperm specification.

In addition to genes transcribed in sperm, one quarter of the H3K27me3-repressed ACRs (110 peaks; 27.0%) were also associated with genes expressed specifically during the initiation of sporophytic development in the early embryo (Figure 5F). Strikingly, several master embryonic regulators all had high levels of chromatin accessibility in sperm, including *BABY BOOM (BBM)*, *LEAFY COTYLEDON 1 (LEC1)*, and *WUSCHEL-RELATED HOMEBOX 2 (WOX2)*, as well as the floral repressor *FLOWERING LOCUS C (FLC)* (Figure 5I; Bosca et al., 2011; Horstman et al., 2017; Tao et al., 2017). This trend was broadly evident at genes with enriched expression during the early and middle phases of embryogenesis, which had higher levels of promoter accessibility than genes with enriched expression later during embryo maturation (Figure 5J; Hofmann et al., 2019). Thus, sperm chromatin is highly derived from the gametophytic state observed in the VN, since it is distinguished by accessible chromatin that forecasts the sporophytic transition in the next generation.

Discussion

Here, we assess the landscape of accessible chromatin in the developing Arabidopsis male gametophyte and reveal major chromatin reprogramming at the transition between haploid and diploid life. Importantly, we show how the reprogramming of accessible chromatin and haploid gene expression is intricately linked with the differential erasure of distinct repressive epigenetic marks – H3K27me3 and H3K9me2-associated DNA methylation. Once microspore division is complete, the programmed removal of H3K9 and DNA methylation in the VC facilitates the expression of a fraction of the genes required for somatic haploid development. In parallel, H3K27me3 removal in sperm exposes *cis*-regulatory regions at developmental genes required to initiate transition back to diploid life after fertilisation. While substantial differences in ACRs have been observed across sporophytic Arabidopsis tissues (Sullivan et al., 2014), more closely related cell types appear to show few specific open chromatin sites (Dorrity et al., 2020; Maher et al., 2018). Despite being separated by a single cell division, thousands of differential ACRs distinguish the SN and VN, which is the result of the epigenetic reprogramming that occur across the male sporophyte-to-gametophyte transition. Future cell type-specific analysis of the sporophyte and female gametophyte promises to reveal whether the extent of transcriptional rewiring in the male haploid phase is shared during other major developmental transitions in the plant life cycle.

The epigenetic reprogramming events underlying the male haploid transition likely trace back to ancient, conserved mechanisms that accompany life form transitions in early land plants (Ikeda et al., 2018; Okano et al., 2009; Schmid et al., 2018; Yaari et al., 2015). Alternating haploid–diploid life forms in eukaryotes are likely derived from the dominant haploid phase of ancestral

unicellular organisms (Niklas *et al.*, 2014). Multicellularity is proposed to have evolved in these haploid organisms first and was then followed by meiosis to mitigate spontaneous whole-genome duplications (Lenormand *et al.*, 2016; Niklas *et al.*, 2014; Wilkins and Holliday, 2009). Such events are likely to have initiated the first haploid–diploid life cycles during eukaryotic evolution (Niklas *et al.*, 2014). Subsequent evolution of sexual reproduction through the innovation of gametes and fertilisation would have differentiated diploid development from that in the haploid phase (Niklas *et al.*, 2014). In light of this, our findings provide insights into the molecular mechanisms that were likely adopted to control life form transitions during eukaryotic evolution. H3K27me3 likely evolved in unicellular eukaryotes prior to the repressive feedback loop that couples DNA with H3K9 methylation in land plants (Bewick *et al.*, 2017; Schuettengruber *et al.*, 2017). We propose that H3K27me3 reprogramming at the haploid-to-diploid transition might be a remnant of an ancestral role in the first haplo-diplo phasic life cycles. We also postulate that the DNA-H3K9 methylation feedback loop, which is normally associated with transposon silencing in flowering plants (Feng and Michaels, 2015), might have been co-opted to control diploid-to-haploid life form transitions during early land plant evolution. Importantly, our findings implicate H3K9 methylation in the transcriptional repression of not only transposons but also lineage specification in flowering plants, which has developmental relevance during early mouse organogenesis (Nicetto *et al.*, 2019) but likely also in the specification of the female gametophyte (Jiang *et al.*, 2017).

The vegetative cell represents the somatic cell of the male gametophyte and undergoes loss of constitutive heterochromatin, which facilitates active DNA demethylation by DME (He *et al.*, 2019; Mérai *et al.*, 2014; Schoft *et al.*, 2009; Schoft *et al.*, 2011). Chromatin accessibility in the VN supports the expression of TEs and easiRNAs caused by this epigenetic reconfiguration (Ibarra *et al.*, 2012; Slotkin *et al.*, 2009). Our findings suggest that the reprogramming and demethylation of pericentromeric heterochromatin are primarily required for VC differentiation. We identify several putative *cis*-regulatory elements actively demethylated by DME in the VN, many of which are silenced by constitutive heterochromatin in the sporophyte and are predicted to be bound by pollen-expressed TFs strongly repelled by DNA methylation (O'Malley *et al.*, 2016). DME is essential for pollen viability and pollen tube germination in certain *Arabidopsis* ecotypes (Schoft *et al.*, 2011), which is consistent with our findings that DME regulates the expression of several pollen tube-related genes. Intriguingly, the 3' end of highly transcribed genes show prominent chromatin accessibility specifically in the VN. Physical proximity of the promoter and terminator through intragenic looping often facilitates rapid re-initiation of transcription (Grzechnik *et al.*, 2014). Alternatively, these regions might be occupied by insulating proteins that prevent transcriptional read-through into neighbouring genes (West *et al.*, 2002). Either mechanism would contribute to the high levels of transcription that are likely required to fuel rapid pollen tube growth (Qin *et al.*, 2009; Wang *et al.*, 2008).

DME-dependent demethylation preferentially targets short AT-rich TEs (Ibarra *et al.*, 2012), suggesting that gametophytic development is at least partly specified through de-repression of TE-associated *cis*-regulatory sequences. Rewiring of gene expression by transposons and endogenous retroviruses is proposed to have contributed to the evolution and diversification of mammals (Johnson, 2019; Lynch *et al.*, 2011). During early embryogenesis in mice and humans, accessible chromatin is widely shaped by fixed TE insertions that strongly associate with *cis*-regulatory elements (Wu *et al.*, 2016; Wu *et al.*, 2018). Novel TE and endovirus integrations can create novel promoters or TF binding sites that alter the regulation of nearby or even distant genes (Rebollo *et al.*, 2012). We speculate whether TE integrations have imposed repression on the gametophytic programme during diploid life, which in turn could have facilitated the evolution of haploid–diploid life cycles in land plants. This hypothesis supports the general proposal that the fixation and selection of transposon-derived *cis*-regulatory elements marks specific periods of evolution within a lineage (Jacques *et al.*, 2013; Jordan *et al.*, 2003). In conclusion, our findings reveal the intricate relationship between epigenetic reprogramming and the regulatory epigenome during the alternation of generations in flowering plants.

Materials and methods

Plant materials and growth conditions

Arabidopsis thaliana seeds for ProHTR10:HTR10-Clover (Kawashima *et al.*, 2014) and *dme-2/+* (Choi *et al.*, 2002) were described previously. Plants were grown in long day (16 hr light/8 hr dark) conditions at 22°C.

Plasmid construction and plant transformation

ProVCK:NTF and ProDUO1:MDB-NTF were generated using MultiSite Gateway Technology (Thermo Scientific). For the ProVCK:NTF construct, the NTF reporter with a stop codon (Deal and Henikoff, 2011) was recombined downstream of the vegetative cell-specific VCK promoter into the multisite gateway binary vector pB7m24GW,3 (VIB-UGent Gateway Vectors). The pDONRP4P1R-ProVCK entry clone was described previously (Grant-Downton *et al.*, 2013). For the ProDUO1:MDB-NTF construct, the CYCB1;one mitotic destruction box (MDB) was fused in frame with the NTF reporter and driven with the DUO1 promoter in a strategy similar to that described previously (Borg *et al.*, 2014). The pDONRP4P1R-ProDUO1 entry clone was described previously (Brownfield *et al.*, 2009). The NTF cDNA was amplified from a ProGL2:NTF construct (Deal and Henikoff, 2011) using the primers NTF-attB1-F and NTF-attB2-R and cloned into pDONR221 or the primers NTF-attB2-F and NTF-attB3-R for cloning into pDONR2RP3. The MDB cDNA without a stop codon was amplified with the primers MDB-attB1-F and MDB-attB2-R and cloned into pDONR221. The resulting binary vectors were transformed into a homozygous ProUBQ14:BirA marker line in Col-0 background.

The primer sequences were:

```
NTF-attB1-F GGGACAAGTTTGTACAAAAAAGCAGGCTCTATGGATCA
TTCAGCGAAAACCAC
NTF-attB2-R GGGGACCACTTTGTACAAGAAAGCTGGGTCTCAAGATCCACCAGTATCCTC
NTF-attB2-F GGGGACCACTTTGTACAAGAAAGCTGGGTCTCAAGATCCACCAGTATCCTC
NTF-attB3-R GGGGCAACTTTGTATAATAAAGTTGTTCAAGATCCACCAGTATCCTC
MDB-attB1-F GGGACAAGTTTGTACAAAAAAGCAGGCTTCATGATGACTTCTCGTTTCGATTG
TTC
MDB-attB2-R GGGGACCACTTTGTACAAGAAAGCTGGGTCTTCTCTCGAGCAGCAAC
TAAACC
```

Fluorescence-activated cell sorting

Dry pollen from the ProVCK:NTF and ProHTR10:HTR10-Clover marker lines was harvested using a method published previously (Johnson-Brousseau and McCormick, 2004). SN and VN were isolated by FANS using a method described previously (Borges *et al.*, 2012). Dry pollen was resuspended in ice-cold Galbraith Buffer (with 1% Triton X-100) (Galbraith *et al.*, 1983) supplemented with 10 mM β -mercaptoethanol and 1 \times complete protease inhibitor cocktail (Roche). After gentle rotation at room temperature for ~15 min, 500 μ m glass beads were added to the pollen suspension and vortexed vigorously to disrupt pollen grains. Debris was removed by filtering the suspension through a 10 μ m nylon mesh. The crude suspension was immediately subjected to FANS on a FACS Aria III cell sorter. The sorter was run in standard configuration, using a 70 μ m ceramic nozzle with 1 \times PBS running at a constant pressure of 20 psi. A 488 nm blue laser was used to excite both GFP (i.e. NTF) and Clover and detected with a 530/30 nm bandpass filter.

Intact microspores were isolated by fluorescence-activated cell sorting (FACS) based on a method described previously (Borges *et al.*, 2012). Unopened buds were picked from Col-0 wild-type plants and ground gently in ice-cold pollen extraction buffer (PEB: 10 mM CaCl_2 , 2 mM MES, 1 mM KCl, 1% H_3BO_3 , 10% Sucrose, pH 7.5). The crude spore preparation was then filtered through two layers of miracloth and the spore suspension centrifuged at 800 g for 5 min at 4°C to pellet the spores. The resulting pellet was resuspended in PEB, filtered through a 20 μ m mesh, and immediately subjected to FACS on a FACS Aria III cell sorter. The sorter was run in standard configuration, using a 100 μ m ceramic nozzle with 1 \times PBS running at a constant pressure of 20 psi. Microspores were gated for high angle scatter (SSC) and autofluorescence in the GFP channel using a 488 nm blue laser and 530/30 nm bandpass filter. Microspore purity and integrity were assessed by DAPI staining and microscopy immediately after FACS-based purification.

GN were isolated by FANS from a crude spore population isolated from the ProDUO1:MDB-NTF marker line. Unopened buds were picked from this line and ground gently in ice-cold 0.3 M mannitol solution. The crude spore preparation was filtered through a 100 μ m nylon mesh and Triton X-100 added to a final concentration of 1%. The spore suspension was centrifuged at 800 g for 5 min at 4°C to pellet the spores. The resulting pellet was resuspended in ice-cold Galbraith Buffer (with 1% Triton X-100) (Galbraith et al., 1983) supplemented with 10 mM β -mercaptoethanol and 1 \times complete protease inhibitor cocktail (Roche). 500 μ m glass beads were added to the spore suspension and vortexed vigorously to disrupt the spores. Debris was removed by filtering the suspension through a 10 μ m nylon mesh. The crude suspension was immediately subjected to FANS on a FACS Aria III cell sorter. The sorter was run in standard configuration, using a 70 μ m ceramic nozzle with 1 \times PBS running at a constant pressure of 20 psi. A 488 nm blue laser was used to excite GFP (i.e. NTF) and detected with a 530/30 nm bandpass filter.

RNA-seq analysis

For analysis of *dme-2/+* pollen, dry pollen was harvested from Col-*gl* and BASTA-sprayed *dme-2/+* plants using a method published previously (Johnson-Brousseau and McCormick, 2004). Pollen grains were disrupted in a Precellys tissue homogeniser (Bertin Instruments) and total RNA isolated using an RNeasy Micro Kit (Qiagen). For vegetative cell nuclei, FACS-isolated nuclei from the ProVCK:NTF line were sorted directly into Trizol reagent and precipitated with isopropanol. RNA-seq libraries were generated from the resulting RNA with Smart-seq2 (Picelli et al., 2014) using independent biological replicates. The libraries were sequenced on an Illumina HiSeq 2500 using 50 bp paired-end and single-end reads for VN and pollen samples, respectively.

Adapter trimming was performed using TrimGalore version 0.4.1 RRID:SCR_011847 (<https://github.com/FelixKrueger/TrimGalore>). Reads were aligned to the Arabidopsis genome (TAIR10) using the STAR aligner version 2.5.2a (Dobin et al., 2013). Transcripts per million (TPM) values were generated using Kallisto version 0.43.1 (Bray et al., 2016) with an index built on TAIR10 cDNA sequences (*Arabidopsis_thaliana.TAIR10.cdna.all.fa.gz*). Published RNA-seq data sets were analysed and described previously (Borg et al., 2020) with the addition of microspore RNA-seq data (Wang et al., 2020). Principal component analysis between different Arabidopsis tissues and cell types was based on the mean TPM value of corresponding biological replicates. Differential gene expression analysis between Col-*gl* and *dme-2/+* pollen was performed using DESeq2 version 1.22.2 (Love et al., 2014) for transcripts that had 10 counts or more in at least one sample. Differentially expressed genes (DEGs) were classified as having an FDR adjusted p-value <0.1. Overlap enrichment of DEGs with gene lists of interest was determined with the R package GeneOverlap 1.18.0 function *newGOM* (<https://github.com/shenlab-sinai/GeneOverlap>; RRID:SCR_018419).

ATAC-seq analysis

For GN, SN, and VN ATAC-seq, ~50,000 FACS-purified nuclei were immediately resuspended in 25 μ l transposase reaction mix (12.5 μ l 2 \times TD buffer, 2.0 μ l Nextera Tn5 [Illumina Cat #FC-121–1030], and nuclease-free water). For microspore ATAC-seq, FACS-purified spores were centrifuged at 800 g for 5 min at 4°C and the spore pellet resuspended in 500 μ l lysis buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl₂, 1% IGEPAL CA-630). The microspores were pelleted once more at 500 g for 10 min at 4°C and ~50,000,000 added to a 25 μ l transposase reaction mix (12.5 μ l 2 \times TD buffer, 2.0 μ l Nextera Tn5 [Illumina Cat #FC-121–1030], and nuclease-free water). The Tn5 reaction mixes were incubated at 37°C for 1 hr and tagmented DNA fragments recovered with a MinElute PCR Purification Kit (Qiagen). Sequencing-ready libraries were generated as described previously (Buenrostro et al., 2015). The libraries were sequenced on an Illumina HiSeq 2500 using 50 bp paired-end reads. All ATAC-seq profiles were generated with at least two independent biological replicates.

Adapter trimming was performed with cutadapt version 1.9 using *-minimum-length 5 -overlap 1 -a 'CTGTCTTATACATCTCCGAGCCACGAGAC' -A 'CTGTCTTATACATCTGACGCTGCCGACGA'* (<https://github.com/marcelm/cutadapt/tree/v1.9.1>; RRID:SCR_011841). Reads were mapped to the Arabidopsis genome (TAIR10) with Bowtie2 version 2.1.0 using *-end-to-end -X 2000 -x* (Langmead and Salzberg, 2012) and subsequently filtered for duplicate reads using Picard tools *MarkDuplicates* version 1.141 (<https://github.com/broadinstitute/picard>), RRID:SCR_006525. Only

reads with a mapQ >10 were used in our analysis. Biological replicates were subsequently merged for downstream analysis after confirming high correlation among replicates. Genome bigwig coverage files were generated with the deepTools version 2.2.4 utility *bamCoverage* using `-normalizeTo1 × 119146348 -binSize=10 -smoothLength=100` (Ramírez et al., 2014). A cross-correlation matrix based on Pearson's correlation coefficient was generated by comparing bigwig coverage files of each replicate using deepTools version 2.5.0.1 utility *multiBigwigSummary* (Ramírez et al., 2014). Bigwig coverage files were visualised along the TAIR10 genome using IGV version 2.7.2 (Robinson et al., 2011). ATAC-seq peaks were called using the MACS2 version 2.1.0 *callpeak* function using `-B -q 1e-5 -f BAMPE` on merged replicate BAM files (Zhang et al., 2008). Annotation of ATAC-seq peaks to genes was performed using the R package ChIPseeker version 1.6.7 using Arabidopsis TxDb transcript metadata (*TxDb.Athaliana.BioMart.plantsmart28*). The region range to TSS was set to 900 bp (Yu et al., 2015), such that peaks >900 bp from the closest gene were considered as intergenic. Relative enrichment of the ATAC-seq peak groups to genome features was performed using the *annotatePeaks.pl* script in HOMER version 4.9 (Benner et al., 2017).

Differential analysis of ATAC-seq peaks

For differential ATAC-seq peak analysis, all SN and VN peaks were combined and peaks lying within 50 bp of each other merged into one. Sporophyte-specific peaks were determined as microspore peaks that did not overlap with this combined set of peaks. This resulted in a master peak set with a total of 20,634 peaks. Reads mapping to the master peak set for each replicate were counted using the *featureCounts* function in Subread version 2.0.1 and used to generate FRiP scores. The resulting counts were used for pairwise differential analysis between the cell types using DESeq2 version 1.22.2 with default settings (Love et al., 2014). VN-specific and SN-specific peaks were classified as having an FDR adjusted p-value <0.05 and a log₂ fold-change >1 or < -1, respectively. Common peaks were considered to have a log₂ fold-change between 1 and -1.

Analysis of ATAC-seq peak features

Chromosomal plots of ATAC-seq peak density based on the average number of peaks were calculated in 10 kb bins using bedtools *map* version 2.26 (<https://bedtools.readthedocs.io/en/latest/>). ATAC-seq heatmaps were generated using the R package EnrichedHeatmap (Gu et al., 2018). Averaged ATAC-seq profiles were generated using the EnrichedHeatmap function *normalizeToMatrix* and plotted using a custom script in R. Gene ontology (GO) enrichment was performed using the gProfileR package in R (Reimand et al., 2011) using gSCS correction, strong hierarchical filtering, and limited to biological process GO terms. Levels of histone marks associated with the ATAC-seq peaks were calculated with bedtools *map* version 2.26 (<https://bedtools.readthedocs.io/en/latest/>) using ChIP-seq coverage files analysed and described previously (Borg et al., 2020). Histone H1 ChIP-seq data was published previously (Wollmann et al., 2017). Peak overlap enrichment of ATAC-seq peak groups with TEs, siRNAs, and DMRs was performed with GAT using 10,000 permutations with random Arabidopsis genomic regions (<https://gat.readthedocs.io/en/latest/>).

Small RNA re-analysis

Publicly available sRNA-seq reads obtained from Slotkin et al., 2009 and Borges et al., 2018 were analysed as described previously (Papareddy et al., 2020). First, sRNA reads were adaptor trimmed using Cutadapt v1.9.1 (<https://github.com/marcelm/cutadapt/tree/v1.9.1>). Trimmed sequences of 18–30 bp were aligned to the Arabidopsis genome (TAIR10) using STAR aligner version 2.7 (Dobin et al., 2013) requiring zero mismatches and ≤100 multiple end-to-end alignments. The resulting SAM files were used as input for *readmapIO.py* software (Schon et al., 2018) to reassign multimappers with a 'rich-get-richer' algorithm. The output bedFiles were sorted, condensed, and normalised for total genome matching reads. The level of siRNAs for VN-specific ACRs presented in the heatmap represent the mean reads per million (RPM) over the peaks including ±10 bp upstream and downstream.

DNA methylation re-analysis

DNA methylation data sets were obtained from *Ibarra et al., 2012* and analysed as described previously (*Papareddy et al., 2020*). Briefly, sequenced reads were quality filtered and adaptor trimmed using TrimGalore version 0.6.2 (<https://github.com/FelixKrueger/TrimGalore>). Bisulphite-converted reads were aligned against the TAIR10 genome using Bismark version 0.22.2 with *bismark -non_directional -q -score-min L,0,-0.4* (*Krueger and Andrews, 2011*). BAM files containing clonal deduplicated and uniquely mapped reads were then used as substrate for Methylpy software version 1.2.9 (<https://bitbucket.org/schultzmatt/methylpy>) to extract weighted methylation rates at each cytosine as described previously (*Schultz et al., 2015*). Bisulphite conversion rates were calculated using the unmethylated chloroplast genome. Differentially methylated regions (DMRs) were defined using *Methylpy*. Differentially methylated cytosines (DMCs) with coverage ≥ 4 overlapping reads were identified by root mean square tests and a false discovery rate ≤ 0.05 . Differentially methylated sites within 250 bp were collapsed into DMRs. DMRs were further filtered by discarding regions with < 5 DMCs. Statistical analyses and associated figures were generated with a custom script in R. Sequence logo for DMCs was generated using the R package ggseqlogo (*Wagih, 2017*).

Analysis of DME-targeted accessible chromatin

DME-targeted ACRs were defined as ATAC-seq peaks in the VN that overlapped with a hyper-DMC in *dme-2/+* VN compared to WT (see *DNA methylation analysis*). Overlap enrichment of genes associated with non-intergenic DME-targeted ACRs against gene lists of interest was determined with the R package GeneOverlap version 1.18.0 (<https://github.com/shenlab-sinai/GeneOverlap>). Motif enrichment was performed using motifs in the DAP-seq data set (*O'Malley et al., 2016*) for transcription factors with TPM > 1 in the VN using *AME* in the MEME suite version 5.0.4 using all ATAC-seq peaks as a background model (*Bailey et al., 2009*). Binding \log_2 fold-change for these TFs in DAP-seq and ampDAP-seq experiments was described previously (*O'Malley et al., 2016*). Predicted TF binding sites within DME-targeted ACRs were determined using *FIMO* in the MEME suite version 5.0.4 with a FIMO threshold p-value $< 1e-5$ and a background Markov model generated from all ATAC-seq peaks (*Bailey et al., 2009*). Pollen tube-induced genes were determined from previous semi in vivo pollen tube (SIVPT) data sets, with all genes in the intersection between SNP-based and microarray-based experiments considered (*Leydon et al., 2017; Qin et al., 2009*). ChIP-seq data for leaf H3K9me2 was described previously (*Baerenfaller et al., 2016*).

Analysis of H3K27me3-repressed accessible chromatin

H3K27me3-repressed ACRs were defined as ATAC-seq peaks in SN that overlapped with somatic H3K27me3 domains described previously (*Borg et al., 2020*) using the R package ChIPpeakAnno function *findOverlapsOfPeaks* (*Zhu et al., 2010*). The clustering of somatic H3K27me3 target genes was based on H3K4me3 enrichment in sperm, with clusters 1 and 2 (primed) and cluster 3 (unprimed) described previously (*Borg et al., 2020*). Genome bigwig coverage files for somatic H3K27me3 and sperm H3K4me3 were also described previously (*Borg et al., 2020*). Overlap enrichment of genes associated with non-intergenic H3K27me3-repressed ACRs against gene lists of interest was determined with the R package GeneOverlap version 1.18.0 (<https://github.com/shenlab-sinai/GeneOverlap>). Peak overlap enrichment of ATAC-seq peak groups with different features was performed using GAT with 10,000 permutations of random Arabidopsis genomic regions (<https://gat.readthedocs.io/en/latest/>). DUO1 target genes were defined as all genes significantly induced by DUO1 overexpression in seedlings as described previously (*Borg et al., 2011*). Sperm-enriched genes were defined previously (*Borg et al., 2020*). Early embryogenesis genes were defined as genes with a TPM > 1 in 14 hr zygotes or 2 cell embryos and a TPM < 1 in sperm, using RNA-seq data described previously (*Borg et al., 2020*). Early and late embryo-enriched genes were defined previously and corresponded to pre-cotyledon and mature green phase markers, respectively (*Hofmann et al., 2019*).

Acknowledgements

We thank Z Lorković and J M Watson for critical reading of the manuscript, S Akimcheva for guidance and technical support, and Yeonhee Choi for *dme-2* seeds. We thank the Next Generation

Sequencing and Plant Sciences facilities at the Vienna BioCenter Core Facilities GmbH, and the CLIP-CBE HPC team. This work was supported through core funding from the Gregor Mendel Institute, external grants from the FWF (P26887, I4258) and ERA-CAPS (EVO-REPRO I2163-B16). MB was supported through the FWF Lise Meitner fellowship M1818. RPK and MDN were supported by the European Research Council under the European Union's Horizon 2020 Research and Innovation Program (grant 637888 to MDN). DT was supported by Biotechnology and Biological Research Council funding (BB/I011269/1 and BB/N005090).

Additional information

Funding

Funder	Grant reference number	Author
Austrian Science Fund	P26887	Frédéric Berger
Austrian Science Fund	I 4258	Frédéric Berger
Austrian Science Fund	I2163-B16	Frédéric Berger
Austrian Science Fund	M1818	Michael Borg
European Commission	ERC 637888	Michael D Nodine
Biotechnology and Biological Sciences Research Council	BB/I011269/1	David Twell
Biotechnology and Biological Sciences Research Council	BB/N005090	David Twell

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Michael Borg, Conceptualization, Data curation, Software, Formal analysis, Funding acquisition, Validation, Investigation, Visualization, Methodology, Writing - original draft, Writing - review and editing; Ranjith K Papareddy, Rodolphe Dombey, Formal analysis; Elin Axelsson, Data curation, Software, Methodology; Michael D Nodine, David Twell, Funding acquisition; Frédéric Berger, Conceptualization, Formal analysis, Supervision, Funding acquisition, Validation, Investigation, Methodology, Project administration, Writing - review and editing

Author ORCIDs

Michael Borg  <http://orcid.org/0000-0002-3982-3843>

Rodolphe Dombey  <http://orcid.org/0000-0002-3670-4128>

Elin Axelsson  <http://orcid.org/0000-0003-4382-1880>

Frédéric Berger  <https://orcid.org/0000-0002-3609-8260>

Decision letter and Author response

Decision letter <https://doi.org/10.7554/eLife.61894.sa1>

Author response <https://doi.org/10.7554/eLife.61894.sa2>

Additional files

Supplementary files

- Supplementary file 1. Pollen ATAC-seq accessible chromatin regions (ACRs) detailed in this study.
- Supplementary file 2. Differentially hyper-methylated cytosines (DMCs) in vegetative cell nucleus (VN) from *dme-2/+* pollen compared to wild type (WT) described in this study.
- Supplementary file 3. DEMETER (DME)-targeted accessible chromatin regions (ACRs) described in this study.
- Supplementary file 4. Differential gene expression analysis of wild-type (WT) and *dme-2/+* pollen.

- Supplementary file 5. Pollen-tube related genes directly regulated by DEMETER (DME)-dependent demethylation.
- Supplementary file 6. H3K27me3-repressed accessible chromatin regions (ACRs) described in this study.
- Transparent reporting form

Data availability

Deep-sequencing data that support the findings of this study have been deposited in the Gene Expression Omnibus (GEO) under accession code GSE155369. Re-analysis of previously published DNA methylomes from *dme-2/+* pollen (Ibarra et al., 2012), and siRNAs from leaves (Papareddy et al., 2020) and pollen (Borges et al., 2018; Slotkin et al., 2009) were deposited in the GEO under accession code GSE155369.

The following dataset was generated:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Borg M, Berger F	2020	Epigenetic reprogramming rewires transcription during the alternation of generations in <i>Arabidopsis</i>	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE155369	NCBI Gene Expression Omnibus, GSE155369

The following previously published datasets were used:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Ibarra CA, Feng X, Schoft VK, Hsieh TF, Uzawa R, Rodrigues JA, Zemach A, Chumak N, Machlicova A, Nishimura T, Rojas D, Fischer RL, Tamaru H, Zilberman D	2012	Active DNA demethylation in plant companion cells reinforces transposon methylation in gametes	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE38935	NCBI Gene Expression Omnibus, GSE38935
Borg M, Jacob Y, Susaki D, LeBlanc C	2020	Targeted reprogramming of H3K27me3 resets epigenetic memory in plant paternal chromatin	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE120669	NCBI Gene Expression Omnibus, GSE120669
Borges F, Parent JS, van Ex F, Wolff P	2018	Transposon-derived small RNAs triggered by miR845 mediate genome dosage response in <i>Arabidopsis</i>	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE106117	NCBI Gene Expression Omnibus, GSE106117
Slotkin RK, Vaughn M, Borges F, Tanurdzić M	2009	Epigenetic reprogramming and small RNA silencing of transposable elements in pollen	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE61028	NCBI Gene Expression Omnibus, GSE61028
Papareddy RK, Páldi K, Paulraj S, Kao P	2020	Chromatin regulates expression of small RNAs to help maintain transposon methylome homeostasis in <i>Arabidopsis</i>	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE152971	NCBI Gene Expression Omnibus, GSE152971

References

- Andreuzza S, Li J, Guitton AE, Faure JE, Casanova S, Park JS, Choi Y, Chen Z, Berger F. 2010. DNA LIGASE I exerts a maternal effect on seed development in *Arabidopsis thaliana*. *Development* **137**:73–81. DOI: <https://doi.org/10.1242/dev.041020>, PMID: 20023162
- Baerenfaller K, Shu H, Hirsch-Hoffmann M, Fütterer J, Opitz L, Rehrauer H, Hennig L, Gruissem W. 2016. Diurnal changes in the histone H3 signature H3K9ac|H3K27ac|H3S28p are associated with diurnal gene expression in *Arabidopsis*. *Plant, Cell & Environment* **39**:2557–2569. DOI: <https://doi.org/10.1111/pce.12811>, PMID: 27487196

- Bailey TL**, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. 2009. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Research* **37**:W202–W208. DOI: <https://doi.org/10.1093/nar/gkp335>, PMID: 19458158
- Bayar M**, Nawy T, Giglione C, Galli M, Meinel T, Lukowitz W. 2009. Paternal control of embryonic patterning in *Arabidopsis thaliana*. *Science* **323**:1485–1488. DOI: <https://doi.org/10.1126/science.1167784>, PMID: 19286558
- Becker JD**, Boavida LC, Carneiro J, Haury M, Feijó JA. 2003. Transcriptional profiling of *Arabidopsis* tissues reveals the unique characteristics of the pollen transcriptome. *Plant Physiology* **133**:713–725. DOI: <https://doi.org/10.1104/pp.103.028241>, PMID: 14500793
- Benner C**, Heinz S, Glass CK. 2017. HOMER - Software for Motif Discovery and Next Generation Sequencing Analysis. v4.11. <http://homer.ucsd.edu/homer/>
- Bewick AJ**, Niederhuth CE, Ji L, Rohr NA, Griffin PT, Leebens-Mack J, Schmitz RJ. 2017. The evolution of CHROMOMETHYLASES and gene body DNA methylation in plants. *Genome Biology* **18**:65. DOI: <https://doi.org/10.1186/s13059-017-1195-1>, PMID: 28457232
- Borg M**, Brownfield L, Twell D. 2009. Male gametophyte development: a molecular perspective. *Journal of Experimental Botany* **60**:1465–1478. DOI: <https://doi.org/10.1093/jxb/ern355>, PMID: 19213812
- Borg M**, Brownfield L, Khatab H, Sidorova A, Lingaya M, Twell D. 2011. The R2R3 MYB transcription factor DUO1 activates a male germline-specific regulon essential for sperm cell differentiation in *Arabidopsis*. *The Plant Cell* **23**:534–549. DOI: <https://doi.org/10.1105/tpc.110.081059>, PMID: 21285328
- Borg M**, Rutley N, Kagale S, Hamamura Y, Gherghinoiu M, Kumar S, Sari U, Esparza-Franco MA, Sakamoto W, Rozwadowski K, Higashiyama T, Twell D. 2014. An EAR-Dependent regulatory module promotes male germ cell division and sperm fertility in *Arabidopsis*. *The Plant Cell* **26**:2098–2113. DOI: <https://doi.org/10.1105/tpc.114.124743>, PMID: 24876252
- Borg M**, Jacob Y, Susaki D, LeBlanc C, Buendía D, Axelsson E, Kawashima T, Voigt P, Boavida L, Becker J, Higashiyama T, Martienssen R, Berger F. 2020. Targeted reprogramming of H3K27me3 resets epigenetic memory in plant paternal chromatin. *Nature Cell Biology* **22**:621–629. DOI: <https://doi.org/10.1038/s41556-020-0515-y>, PMID: 32393884
- Borg M**, Berger F. 2015. Chromatin remodelling during male gametophyte development. *The Plant Journal* **83**:177–188. DOI: <https://doi.org/10.1111/tpj.12856>, PMID: 25892182
- Borges F**, Gomes G, Gardner R, Moreno N, McCormick S, Feijó JA, Becker JD. 2008. Comparative transcriptomics of *Arabidopsis* sperm cells. *Plant Physiology* **148**:1168–1181. DOI: <https://doi.org/10.1104/pp.108.125229>, PMID: 18667720
- Borges F**, Gardner R, Lopes T, Calarco JP, Boavida LC, Slotkin RK, Martienssen RA, Becker JD. 2012. FACS-based purification of *Arabidopsis* microspores, sperm cells and vegetative nuclei. *Plant Methods* **8**:44. DOI: <https://doi.org/10.1186/1746-4811-8-44>, PMID: 23075219
- Borges F**, Parent JS, van Ex F, Wolff P, Martínez G, Köhler C, Martienssen RA. 2018. Transposon-derived small RNAs triggered by miR845 mediate genome dosage response in *Arabidopsis*. *Nature Genetics* **50**:186–192. DOI: <https://doi.org/10.1038/s41588-017-0032-5>, PMID: 29335544
- Boscá S**, Knauer S, Laux T. 2011. Embryonic development in *Arabidopsis thaliana*: from the zygote division to the shoot meristem. *Frontiers in Plant Science* **2**:93. DOI: <https://doi.org/10.3389/fpls.2011.00093>, PMID: 22639618
- Bray NL**, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology* **34**:525–527. DOI: <https://doi.org/10.1038/nbt.3519>, PMID: 27043002
- Brownfield L**, Hafidh S, Borg M, Sidorova A, Mori T, Twell D. 2009. A plant germline-specific integrator of sperm specification and cell cycle progression. *PLOS Genetics* **5**:e1000430. DOI: <https://doi.org/10.1371/journal.pgen.1000430>, PMID: 19300502
- Buenrostro JD**, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. 2013. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nature Methods* **10**:1213–1218. DOI: <https://doi.org/10.1038/nmeth.2688>
- Buenrostro JD**, Wu B, Chang HY, Greenleaf WJ. 2015. Atac-seq: a method for assaying chromatin accessibility genome-wide. *Current Protocols in Molecular Biology* **109**:21.29.1–21.29.21. DOI: <https://doi.org/10.1002/0471142727.mb2129s109>
- Choi Y**, Gehring M, Johnson L, Hannon M, Harada JJ, Goldberg RB, Jacobsen SE, Fischer RL. 2002. DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in *Arabidopsis*. *Cell* **110**:33–42. DOI: [https://doi.org/10.1016/S0092-8674\(02\)00807-3](https://doi.org/10.1016/S0092-8674(02)00807-3), PMID: 12150995
- Couceiro L**, Le Gac M, Hunsperger HM, Mauger S, Destombe C, Cock JM, Ahmed S, Coelho SM, Valero M, Peters AF. 2015. Evolution and maintenance of haploid-diploid life cycles in natural populations: the case of the marine Brown alga *Ectocarpus*. *Evolution* **69**:1808–1822. DOI: <https://doi.org/10.1111/evo.12702>
- Creasey KM**, Zhai J, Borges F, Van Ex F, Regulski M, Meyers BC, Martienssen RA. 2014. miRNAs trigger widespread epigenetically activated siRNAs from transposons in *Arabidopsis*. *Nature* **508**:411–415. DOI: <https://doi.org/10.1038/nature13069>, PMID: 24670663
- Deal RB**, Henikoff S. 2011. The INTACT method for cell type-specific gene expression and chromatin profiling in *Arabidopsis thaliana*. *Nature Protocols* **6**:56–68. DOI: <https://doi.org/10.1038/nprot.2010.175>, PMID: 21212783
- Dickinson HG**, Grant-Downton R. 2009. Bridging the generation gap: flowering plant gametophytes and animal germlines reveal unexpected similarities. *Biological Reviews* **84**:589–615. DOI: <https://doi.org/10.1111/j.1469-185X.2009.00088.x>, PMID: 19725820

- Dobin A**, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**:15–21. DOI: <https://doi.org/10.1093/bioinformatics/bts635>, PMID: 23104886
- Dorrrity MW**, Alexandre C, Hamm M, Vigil A-L, Fields S, Queitsch C, Cuperus J. 2020. The regulatory landscape of *Arabidopsis thaliana* roots at single-cell resolution. *bioRxiv*. DOI: <https://doi.org/10.1101/2020.07.17.204792>
- Eady C**, Lindsey K, Twell D. 1995. The significance of microspore division and division symmetry for vegetative Cell-Specific transcription and generative cell differentiation. *The Plant Cell* **7**:65–74. DOI: <https://doi.org/10.2307/3869838>, PMID: 12242352
- Erbasol Serbes I**, Palovaara J, Groß-Hardt R. 2019. Development and function of the flowering plant female gametophyte. *Current Topics in Developmental Biology* **131**:401–434. DOI: <https://doi.org/10.1016/bs.ctdb.2018.11.016>, PMID: 30612625
- Feng W**, Michaels SD. 2015. Accessing the inaccessible: the organization, transcription, replication, and repair of heterochromatin in plants. *Annual Review of Genetics* **49**:439–459. DOI: <https://doi.org/10.1146/annurev-genet-112414-055048>, PMID: 26631514
- Galbraith DW**, Harkins KR, Maddox JM, Ayres NM, Sharma DP, Firoozabady E. 1983. Rapid flow cytometric analysis of the cell cycle in intact plant tissues. *Science* **220**:1049–1051. DOI: <https://doi.org/10.1126/science.220.4601.1049>, PMID: 17754551
- Grant-Downton R**, Kourmpetli S, Hafidh S, Khatib H, Le Trionnaire G, Dickinson H, Twell D. 2013. Artificial microRNAs reveal cell-specific differences in small RNA activity in pollen. *Current Biology* **23**:R599–R601. DOI: <https://doi.org/10.1016/j.cub.2013.05.055>, PMID: 23885870
- Grzechnik P**, Tan-Wong SM, Proudfoot NJ. 2014. Terminate and make a loop: regulation of transcriptional directionality. *Trends in Biochemical Sciences* **39**:319–327. DOI: <https://doi.org/10.1016/j.tibs.2014.05.001>, PMID: 24928762
- Gu Z**, Eils R, Schlesner M, Ishaque N. 2018. EnrichedHeatmap: an R/Bioconductor package for comprehensive visualization of genomic signal associations. *BMC Genomics* **19**:234. DOI: <https://doi.org/10.1186/s12864-018-4625-x>, PMID: 29618320
- Hackenberg D**, Twell D. 2019. In: The evolution and patterning of male gametophyte development. *Current Topics in Developmental Biology* **131**:257–298. DOI: <https://doi.org/10.1016/bs.ctdb.2018.10.008>, PMID: 30612620
- Haig D**. 2008. Homologous versus antithetic alternation of generations and the origin of sporophytes. *The Botanical Review* **74**:395–418. DOI: <https://doi.org/10.1007/s12229-008-9012-x>
- Hajkova P**. 2011. Epigenetic reprogramming in the germline: towards the ground state of the epigenome. *Philosophical Transactions of the Royal Society B: Biological Sciences* **366**:2266–2273. DOI: <https://doi.org/10.1098/rstb.2011.0042>
- Hamamura Y**, Nagahara S, Higashiyama T. 2012. Double fertilization on the move. *Current Opinion in Plant Biology* **15**:70–77. DOI: <https://doi.org/10.1016/j.pbi.2011.11.001>, PMID: 22153653
- He S**, Vickers M, Zhang J, Feng X. 2019. Natural depletion of histone H1 in sex cells causes DNA demethylation, heterochromatin decondensation and transposon activation. *eLife* **8**:e42530. DOI: <https://doi.org/10.7554/eLife.42530>, PMID: 31135340
- Heesch S**, Serrano-Serrano M, Luthringer R, Peters AF, Destombe C, Cock JM, Valero M, Roze D, Salamin N, Coelho S. 2019. Evolution of life cycles and reproductive traits: insights from the Brown algae. *bioRxiv*. DOI: <https://doi.org/10.1101/530477>
- Higashiyama T**, Takeuchi H. 2015. The mechanism and key molecules involved in pollen tube guidance. *Annual Review of Plant Biology* **66**:393–413. DOI: <https://doi.org/10.1146/annurev-arplant-043014-115635>, PMID: 25621518
- Hisanaga T**, Yamaoka S, Kawashima T, Higo A, Nakajima K, Araki T, Kohchi T, Berger F. 2019. Building new insights in plant gametogenesis from an evolutionary perspective. *Nature Plants* **5**:663–669. DOI: <https://doi.org/10.1038/s41477-019-0466-0>
- Hofmann F**, Schon MA, Nodine MD. 2019. The embryonic transcriptome of *Arabidopsis thaliana*. *Plant Reproduction* **32**:77–91. DOI: <https://doi.org/10.1007/s00497-018-00357-2>, PMID: 30610360
- Honys D**, Twell D. 2004. Transcriptome analysis of haploid male gametophyte development in *Arabidopsis*. *Genome Biology* **5**:R85. DOI: <https://doi.org/10.1186/gb-2004-5-11-r85>, PMID: 15535861
- Horstman A**, Li M, Heidmann I, Weemen M, Chen B, Muino JM, Angenent GC, Boutilier K. 2017. The BABY BOOM transcription factor activates the LEC1-ABI3-FUS3-LEC2 network to induce somatic embryogenesis. *Plant Physiology* **175**:848–857. DOI: <https://doi.org/10.1104/pp.17.00232>, PMID: 28830937
- Ibarra CA**, Feng X, Schoft VK, Hsieh TF, Uzawa R, Rodrigues JA, Zemach A, Chumak N, Machlicova A, Nishimura T, Rojas D, Fischer RL, Tamaru H, Zilberman D. 2012. Active DNA demethylation in plant companion cells reinforces transposon methylation in gametes. *Science* **337**:1360–1364. DOI: <https://doi.org/10.1126/science.1224839>, PMID: 22984074
- Ikeda Y**, Nishihama R, Yamaoka S, Arteaga-Vazquez MA, Aguilar-Cruz A, Grimanelli D, Pogorelnik R, Martienssen RA, Yamato KT, Kohchi T, Hirayama T, Mathieu O. 2018. Loss of CG methylation in *Marchantia polymorpha* causes disorganization of cell division and reveals unique DNA methylation regulatory mechanisms of Non-CG methylation. *Plant and Cell Physiology* **59**:2421–2431. DOI: <https://doi.org/10.1093/pcp/pcy161>, PMID: 30102384
- Jacques PÉ**, Jeyakani J, Bourque G. 2013. The majority of Primate-Specific regulatory sequences are derived from transposable elements. *PLOS Genetics* **9**:e1003504. DOI: <https://doi.org/10.1371/journal.pgen.1003504>, PMID: 23675311

- Jiang H**, Moreno-Romero J, Santos-González J, De Jaeger G, Gevaert K, Van De Slijke E, Köhler C. 2017. Ectopic application of the repressive histone modification H3K9me2 establishes post-zygotic reproductive isolation in *Arabidopsis thaliana*. *Genes & Development* **31**:1272–1287. DOI: <https://doi.org/10.1101/gad.299347.117>
- Johnson WE**. 2019. Origins and evolutionary consequences of ancient endogenous retroviruses. *Nature Reviews Microbiology* **17**:355–370. DOI: <https://doi.org/10.1038/s41579-019-0189-2>, PMID: 30962577
- Johnson-Brousseau SA**, McCormick S. 2004. A compendium of methods useful for characterizing *Arabidopsis* pollen mutants and gametophytically-expressed genes. *The Plant Journal* **39**:761–775. DOI: <https://doi.org/10.1111/j.1365-313X.2004.02147.x>, PMID: 15315637
- Jordan IK**, Rogozin IB, Glazko GV, Koonin EV. 2003. Origin of a substantial fraction of human regulatory sequences from transposable elements. *Trends in Genetics* **19**:68–72. DOI: [https://doi.org/10.1016/S0168-9525\(02\)00006-9](https://doi.org/10.1016/S0168-9525(02)00006-9), PMID: 12547512
- Kawashima T**, Maruyama D, Shagirov M, Li J, Hamamura Y, Yelagandula R, Toyama Y, Berger F. 2014. Dynamic F-actin movement is essential for fertilization in *Arabidopsis thaliana*. *eLife* **3**:e04501. DOI: <https://doi.org/10.7554/eLife.04501>
- Krueger F**, Andrews SR. 2011. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**:1571–1572. DOI: <https://doi.org/10.1093/bioinformatics/btr167>, PMID: 21493656
- Langmead B**, Salzberg SL. 2012. Fast gapped-read alignment with bowtie 2. *Nature Methods* **9**:357–359. DOI: <https://doi.org/10.1038/nmeth.1923>, PMID: 22388286
- Lenormand T**, Engelstädter J, Johnston SE, Wijnker E, Haag CR. 2016. Evolutionary mysteries in meiosis. *Philosophical Transactions of the Royal Society B: Biological Sciences* **371**:20160001. DOI: <https://doi.org/10.1098/rstb.2016.0001>
- Leydon AR**, Beale KM, Woroniecka K, Castner E, Chen J, Horgan C, Palanivelu R, Johnson MA. 2013. Three MYB transcription factors control pollen tube differentiation required for sperm release. *Current Biology* **23**:1209–1214. DOI: <https://doi.org/10.1016/j.cub.2013.05.021>, PMID: 23791732
- Leydon AR**, Weinreb C, Venable E, Reinders A, Ward JM, Johnson MA. 2017. The molecular dialog between flowering plant reproductive partners defined by SNP-Informed RNA-Sequencing. *The Plant Cell* **29**:984–1006. DOI: <https://doi.org/10.1105/tpc.16.00816>, PMID: 28400492
- Liang Y**, Tan ZM, Zhu L, Niu QK, Zhou JJ, Li M, Chen LQ, Zhang XQ, Ye D. 2013. MYB97, MYB101 and MYB120 function as male factors that control pollen tube-synergid interaction in *Arabidopsis thaliana* fertilization. *PLoS Genetics* **9**:e1003933. DOI: <https://doi.org/10.1371/journal.pgen.1003933>, PMID: 24278028
- Love MI**, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**:550. DOI: <https://doi.org/10.1186/s13059-014-0550-8>, PMID: 25516281
- Lu Z**, Hofmeister BT, Vollmers C, DuBois RM, Schmitz RJ. 2017. Combining ATAC-seq with nuclei sorting for discovery of cis-regulatory regions in plant genomes. *Nucleic Acids Research* **45**:e41. DOI: <https://doi.org/10.1093/nar/gkw1179>, PMID: 27903897
- Lynch VJ**, Leclerc RD, May G, Wagner GP. 2011. Transposon-mediated rewiring of gene regulatory networks contributed to the evolution of pregnancy in mammals. *Nature Genetics* **43**:1154–1159. DOI: <https://doi.org/10.1038/ng.917>, PMID: 21946353
- Maher KA**, Bajic M, Kajala K, Reynoso M, Pauluzzi G, West DA, Zumstein K, Woodhouse M, Bubb K, Dorrity MW, Queitsch C, Bailey-Serres J, Sinha N, Brady SM, Deal RB. 2018. Profiling of accessible chromatin regions across multiple plant species and cell types reveals common gene regulatory principles and new control modules. *The Plant Cell* **30**:15–36. DOI: <https://doi.org/10.1105/tpc.17.00581>, PMID: 29229750
- Martinez G**, Wolff P, Wang Z, Moreno-Romero J, Santos-González J, Conze LL, DeFraia C, Slotkin RK, Köhler C. 2018. Paternal easiRNAs regulate parental genome dosage in *Arabidopsis*. *Nature Genetics* **50**:193–198. DOI: <https://doi.org/10.1038/s41588-017-0033-4>
- Mérai Z**, Chumak N, García-Aguilar M, Hsieh TF, Nishimura T, Schoft VK, Bindics J, Slusarz L, Arnoux S, Opravil S, Mechtler K, Zilberman D, Fischer RL, Tamaru H. 2014. The AAA-ATPase molecular chaperone Cdc48/p97 disassembles sumoylated centromeres, decondenses heterochromatin, and activates ribosomal RNA genes. *PNAS* **111**:16166–16171. DOI: <https://doi.org/10.1073/pnas.1418564111>, PMID: 25344531
- Myers C**, Romanowsky SM, Barron YD, Garg S, Azuse CL, Curran A, Davis RM, Hatton J, Harmon AC, Harper JF. 2009. Calcium-dependent protein kinases regulate polarized tip growth in pollen tubes. *The Plant Journal* **59**:528–539. DOI: <https://doi.org/10.1111/j.1365-313X.2009.03894.x>, PMID: 19392698
- Nicetto D**, Donahue G, Jain T, Peng T, Sidoli S, Sheng L, Montavon T, Becker JS, Grindheim JM, Blahnik K, Garcia BA, Tan K, Bonasio R, Jenuwein T, Zaret KS. 2019. H3K9me3-heterochromatin loss at protein-coding genes enables developmental lineage specification. *Science* **363**:294–297. DOI: <https://doi.org/10.1126/science.aau0583>, PMID: 30606806
- Niklas KJ**, Cobb ED, Kutschera U. 2014. Did meiosis evolve before sex and the evolution of eukaryotic life cycles? *BioEssays* **36**:1091–1101. DOI: <https://doi.org/10.1002/bies.201400045>, PMID: 25143284
- Nodine MD**, Bartel DP. 2012. Maternal and paternal genomes contribute equally to the transcriptome of early plant embryos. *Nature* **482**:94–97. DOI: <https://doi.org/10.1038/nature10756>, PMID: 22266940
- O'Malley RC**, Huang SC, Song L, Lewsey MG, Bartlett A, Nery JR, Galli M, Gallavotti A, Ecker JR. 2016. Cistrome and epicistrome features shape the regulatory DNA landscape. *Cell* **166**:1598. DOI: <https://doi.org/10.1016/j.cell.2016.08.063>, PMID: 27610578
- Okano Y**, Aono N, Hiwatashi Y, Murata T, Nishiyama T, Ishikawa T, Kubo M, Hasebe M. 2009. A polycomb repressive complex 2 gene regulates apogamy and gives evolutionary insights into early land plant evolution. *PNAS* **106**:16321–16326. DOI: <https://doi.org/10.1073/pnas.0906997106>, PMID: 19805300

- Papareddy RK**, Páldi K, Paulraj S, Kao P, Lutzmayer S, Nodine MD. 2020. Chromatin regulates expression of small RNAs to help maintain transposon methylome homeostasis in *Arabidopsis*. *Genome Biology* **21**:251. DOI: <https://doi.org/10.1186/s13059-020-02163-4>, PMID: 32943088
- Park JS**, Frost JM, Park K, Ohr H, Park GT, Kim S, Eom H, Lee I, Brooks JS, Fischer RL, Choi Y. 2017. Control of DEMETER DNA demethylase gene transcription in male and female gamete companion cells in *Arabidopsis thaliana*. *PNAS* **114**:2078–2083. DOI: <https://doi.org/10.1073/pnas.1620592114>, PMID: 28130550
- Picelli S**, Faridani OR, Björklund AK, Winberg G, Sagasser S, Sandberg R. 2014. Full-length RNA-seq from single cells using Smart-seq2. *Nature Protocols* **9**:171–181. DOI: <https://doi.org/10.1038/nprot.2014.006>, PMID: 24385147
- Qin Y**, Leydon AR, Manziello A, Pandey R, Mount D, Denic S, Vasic B, Johnson MA, Palanivelu R. 2009. Penetration of the stigma and style elicits a novel transcriptome in pollen tubes, pointing to genes critical for growth in a pistil. *PLOS Genetics* **5**:e1000621. DOI: <https://doi.org/10.1371/journal.pgen.1000621>, PMID: 19714218
- Qu LJ**, Li L, Lan Z, Dresselhaus T. 2015. Peptide signalling during the pollen tube journey and double fertilization. *Journal of Experimental Botany* **66**:5139–5150. DOI: <https://doi.org/10.1093/jxb/erv275>, PMID: 26068467
- Ramírez F**, Dündar F, Diehl S, Grüning BA, Manke T. 2014. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Research* **42**:W187–W191. DOI: <https://doi.org/10.1093/nar/gku365>, PMID: 24799436
- Rebollo R**, Romanish MT, Mager DL. 2012. Transposable elements: an abundant and natural source of regulatory sequences for host genes. *Annual Review of Genetics* **46**:21–42. DOI: <https://doi.org/10.1146/annurev-genet-110711-155621>, PMID: 22905872
- Reik W**, Dean W, Walter J. 2001. Epigenetic reprogramming in mammalian development. *Science* **293**:1089–1093. DOI: <https://doi.org/10.1126/science.1063443>, PMID: 11498579
- Reimand J**, Arak T, Vilo J. 2011. G:profiler—a web server for functional interpretation of gene lists (2011 update). *Nucleic Acids Research* **39**:W307–W315. DOI: <https://doi.org/10.1093/nar/gkr378>, PMID: 21646343
- Robinson JT**, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. *Nature Biotechnology* **29**:24–26. DOI: <https://doi.org/10.1038/nbt.1754>, PMID: 21221095
- Rotman N**, Durbarry A, Wardle A, Yang WC, Chaboud A, Faure JE, Berger F, Twell D. 2005. A novel class of MYB factors controls sperm-cell formation in plants. *Current Biology* **15**:244–248. DOI: <https://doi.org/10.1016/j.cub.2005.01.013>, PMID: 15694308
- Schmid MW**, Giraldo-Fonseca A, Rövekamp M, Smetanin D, Bowman JL, Grossniklaus U. 2018. Extensive epigenetic reprogramming during the life cycle of *Marchantia polymorpha*. *Genome Biology* **19**:9. DOI: <https://doi.org/10.1186/s13059-017-1383-z>, PMID: 29368664
- Schoft VK**, Chumak N, Mosiolek M, Slusarz L, Komnenovic V, Brownfield L, Twell D, Kakutani T, Tamaru H. 2009. Induction of RNA-directed DNA methylation upon decondensation of constitutive heterochromatin. *EMBO Reports* **10**:1015–1021. DOI: <https://doi.org/10.1038/embor.2009.152>, PMID: 19680290
- Schoft VK**, Chumak N, Choi Y, Hannon M, Garcia-Aguilar M, Machlicova A, Slusarz L, Mosiolek M, Park JS, Park GT, Fischer RL, Tamaru H. 2011. Function of the DEMETER DNA glycosylase in the *Arabidopsis thaliana* male gametophyte. *PNAS* **108**:8042–8047. DOI: <https://doi.org/10.1073/pnas.1105117108>, PMID: 21518889
- Schon MA**, Kellner MJ, Plotnikova A, Hofmann F, Nodine MD. 2018. NanoPARE: parallel analysis of RNA 5' ends from low-input RNA. *Genome Research* **28**:1931–1942. DOI: <https://doi.org/10.1101/gr.239202.118>, PMID: 30355603
- Schuettengruber B**, Bourbon HM, Di Croce L, Cavalli G. 2017. Genome regulation by polycomb and trithorax: 70 years and counting. *Cell* **171**:34–57. DOI: <https://doi.org/10.1016/j.cell.2017.08.002>, PMID: 28938122
- Schultz MD**, He Y, Whitaker JW, Hariharan M, Mukamel EA, Leung D, Rajagopal N, Nery JR, Ulrich MA, Chen H, Lin S, Lin Y, Jung I, Schmitt AD, Selvaraj S, Ren B, Sejnowski TJ, Wang W, Ecker JR. 2015. Human body epigenome maps reveal noncanonical DNA methylation variation. *Nature* **523**:212–216. DOI: <https://doi.org/10.1038/nature14465>, PMID: 26030523
- Slane D**, Kong J, Schmid M, Jürgens G, Bayer M. 2015. Profiling of embryonic nuclear vs. cellular RNA in *Arabidopsis thaliana*. *Genomics Data* **4**:96–98. DOI: <https://doi.org/10.1016/j.gdata.2015.03.015>, PMID: 26484189
- Slotkin RK**, Vaughn M, Borges F, Tanurđić M, Becker JD, Feijó JA, Martienssen RA. 2009. Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. *Cell* **136**:461–472. DOI: <https://doi.org/10.1016/j.cell.2008.12.038>, PMID: 19203581
- Stroud H**, Greenberg MV, Feng S, Bernatavichute YV, Jacobsen SE. 2013. Comprehensive analysis of silencing mutants reveals complex regulation of the *Arabidopsis* methylome. *Cell* **152**:352–364. DOI: <https://doi.org/10.1016/j.cell.2012.10.054>, PMID: 23313553
- Sullivan AM**, Arsovski AA, Lempe J, Bubb KL, Weirauch MT, Sabo PJ, Sandstrom R, Thurman RE, Neph S, Reynolds AP, Stergachis AB, Vernot B, Johnson AK, Haugen E, Sullivan ST, Thompson A, Neri FV, Weaver M, Diegel M, Mnaimneh S, et al. 2014. Mapping and dynamics of regulatory DNA and transcription factor networks in *A. thaliana*. *Cell Reports* **8**:2015–2030. DOI: <https://doi.org/10.1016/j.celrep.2014.08.019>, PMID: 25220462
- Sze H**, Padmanaban S, Cellier F, Honys D, Cheng NH, Bock KW, Conéjéro G, Li X, Twell D, Ward JM, Hirschi KD. 2004. Expression patterns of a novel *AtCHX* gene family highlight potential roles in osmotic adjustment and K⁺ homeostasis in pollen development. *Plant Physiology* **136**:2532–2547. DOI: <https://doi.org/10.1104/pp.104.046003>, PMID: 15347787

- Takeuchi H**, Higashiyama T. 2016. Tip-localized receptors control pollen tube growth and LURE sensing in *Arabidopsis*. *Nature* **531**:245–248. DOI: <https://doi.org/10.1038/nature17413>, PMID: 26961657
- Tao Z**, Shen L, Gu X, Wang Y, Yu H, He Y. 2017. Embryonic epigenetic reprogramming by a pioneer transcription factor in plants. *Nature* **551**:124–128. DOI: <https://doi.org/10.1038/nature24300>, PMID: 29072296
- Wagih O**. 2017. Ggseqlogo: a versatile R package for drawing sequence logos. *Bioinformatics* **33**:3645–3647. DOI: <https://doi.org/10.1093/bioinformatics/btx469>, PMID: 29036507
- Wang Y**, Zhang WZ, Song LF, Zou JJ, Su Z, Wu WH. 2008. Transcriptome analyses show changes in gene expression to accompany pollen germination and tube growth in *Arabidopsis*. *Plant Physiology* **148**:1201–1211. DOI: <https://doi.org/10.1104/pp.108.126375>, PMID: 18775970
- Wang Z**, Butel N, Santos-González J, Borges F, Yi J, Martienssen RA, Martínez G, Köhler C. 2020. Polymerase IV Plays a Crucial Role in Pollen Development in *Capsella*. *The Plant Cell* **32**:950–966. DOI: <https://doi.org/10.1105/tpc.19.00938>
- West AG**, Gaszner M, Felsenfeld G. 2002. Insulators: many functions, many mechanisms. *Genes & Development* **16**:271–288. DOI: <https://doi.org/10.1101/gad.954702>, PMID: 11825869
- Wilkins AS**, Holliday R. 2009. The evolution of meiosis from mitosis. *Genetics* **181**:3–12. DOI: <https://doi.org/10.1534/genetics.108.099762>, PMID: 19139151
- Wollmann H**, Stroud H, Yelagandula R, Tarutani Y, Jiang D, Jing L, Jamge B, Takeuchi H, Holec S, Nie X, Kakutani T, Jacobsen SE, Berger F. 2017. The histone H3 variant H3.3 regulates gene body DNA methylation in *Arabidopsis thaliana*. *Genome Biology* **18**:94. DOI: <https://doi.org/10.1186/s13059-017-1221-3>, PMID: 28521766
- Wu J**, Huang B, Chen H, Yin Q, Liu Y, Xiang Y, Zhang B, Liu B, Wang Q, Xia W, Li W, Li Y, Ma J, Peng X, Zheng H, Ming J, Zhang W, Zhang J, Tian G, Xu F, et al. 2016. The landscape of accessible chromatin in mammalian preimplantation embryos. *Nature* **534**:652–657. DOI: <https://doi.org/10.1038/nature18606>, PMID: 27309802
- Wu J**, Xu J, Liu B, Yao G, Wang P, Lin Z, Huang B, Wang X, Li T, Shi S, Zhang N, Duan F, Ming J, Zhang X, Niu W, Song W, Jin H, Guo Y, Dai S, Hu L, et al. 2018. Chromatin analysis in human early development reveals epigenetic transition during ZGA. *Nature* **557**:256–260. DOI: <https://doi.org/10.1038/s41586-018-0080-8>, PMID: 29720659
- Yaari R**, Noy-Malka C, Wiedemann G, Auerbach Gershovitz N, Reski R, Katz A, Ohad N. 2015. DNA METHYLTRANSFERASE 1 is involved in (m)CG and (m)CCG DNA methylation and is essential for sporophyte development in *Physcomitrella patens*. *Plant Molecular Biology* **88**:387–400. DOI: <https://doi.org/10.1007/s11103-015-0328-8>, PMID: 25944663
- Yu G**, Wang LG, He QY. 2015. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**:2382–2383. DOI: <https://doi.org/10.1093/bioinformatics/btv145>, PMID: 25765347
- Zhang Y**, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biology* **9**:R137. DOI: <https://doi.org/10.1186/gb-2008-9-9-r137>, PMID: 18798982
- Zhao P**, Zhou X, Shen K, Liu Z, Cheng T, Liu D, Cheng Y, Peng X, Sun MX. 2019. Two-Step Maternal-to-Zygotic transition with Two-Phase parental genome contributions. *Developmental Cell* **49**:882–893. DOI: <https://doi.org/10.1016/j.devcel.2019.04.016>, PMID: 31080059
- Zhu LJ**, Gazin C, Lawson ND, Pagès H, Lin SM, Lapointe DS, Green MR. 2010. ChIPpeakAnno: a bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC Bioinformatics* **11**:237. DOI: <https://doi.org/10.1186/1471-2105-11-237>, PMID: 20459804