nature communications

Article

https://doi.org/10.1038/s41467-024-51513-4

ARID1 is required to regulate and reinforce H3K9me2 in sperm cells in Arabidopsis

Received: 17 January 2024

Accepted: 8 August 2024

Published online: 16 August 2024

Check for updates

Lei Li[1], Huaihao Yang[1], Yi Zhao[1], Qianqian Hu[1], Xiaotuo Zhang[1], Ting Jiang[1], Hua Jiang[2] & Binglian Zheng[1]

Heterochromatin de-condensation in companion gametic cells is conserved in both plants and animals. In plants, microspore undergoes asymmetric pollen mitosis (PMI) to produce a vegetative cell (VC) and a generative cell (GC). Subsequently, the GC undergoes pollen mitosis (PMII) to produce two sperm cells (SC). Consistent with heterochromatin de-condensation in the VC, H3K9me2, a heterochromatin mark, is barely detected in VC. However, how H3K9me2 is differentially regulated during pollen mitosis remains unclear. Here, we show that H3K9me2 is gradually evicted from the VC since PMI but remain unchanged in the GC and SC. ARID1, a pollen-specific transcription factor that facilitates PMII, promotes H3K9me2 maintenance in the GC/SC but slows down its eviction in the VC. The genomic targets of ARID1 mostly overlaps with H3K9me2 loci, and ARID1 recruits H3K9 methyltransferase SUVH6. Our results uncover that differential pattern of H3K9me2 between two cell types is regulated by ARID1 during pollen mitosis.

Epigenetic reprogramming is a highly conserved phenomenon during germ cell maturation. Epigenetic marks, including DNA methylation and histone modifications, are reprogrammed in the gametes to reset the genomic transcriptome of the next generation[1]. In mammals, paternal chromatin is extensively reprogrammed through the global erasure of DNA methylation and the exchange of histones with protamines[2,3]. In plants, sperm cell formation is established by mitotic divisions after meiosis[4]. The male meiotic product, termed as microspore, undergoes asymmetric pollen mitosis (PMI) to form a bicellular pollen comprised of a larger vegetative cell (VC) and a smaller generative cell (GC). In contrast to arrest at G1 phase of the VC, the GC undergoes the second pollen mitosis (PMII) to produce two sperm cells (SC). Therefore, a mature pollen grain contains two SC completely engulfed in the cytoplasm of the VC. Unlike sperm in mammals, pollen maintains relatively constant DNA methylation, and histones are retained in sperm chromatin[5]. Recent findings in Arabidopsis show that H3K27me3, a typical repressive histone mark for gene expression, was abolished in the SC while regained in the VC after PMII[6–8], suggesting that H3K27me3 undergoes reprogramming during pollen mitosis. Moreover, dynamics in chromatin accessibility and transcriptional reprogramming have been detected during pollen mitosis[9]. However, although it is better understood why constitutive heterochromatin is decondensed in the VC, less is known about its fate during sperm cell formation.

Chromatin condensation is usually associated with heterochromatin formation at constitutive pericentromeric regions[10,11]. In general, repressive histone modifications and DNA methylation contribute to heterochromatin formation[12]. Among these epigenetic marks, Histone 3 lysine 9 dimethylation (H3K9me2) is generally thought as a typical heterochromatin mark in plants[12]. In Arabidopsis, H3K9me2 accumulates in pericentromeric regions of chromosomes which are generally transcriptionally inactive. DAPI staining and immunofluorescence assay using an anti-H3K9me2 body usually exhibit 5–10 bright and condensed dots in the nucleus in Arabidopsis. Consistent with relaxation of chromatin in companion cells, transposable elements (TE) are reactivated in the VC facilitated by the absence of histone 1 (H1) and DME-mediated DNA demethylation[13–15]. Subsequently, TE transcripts are processed into siRNAs that translocate to reinforce TE silencing in the sperm cells[13,16–18].

We previously reported that a pollen transcription factor ARID1 (AT-Rich Interacting Domain-containing protein 1) promotes sperm cell formation by facilitating expression of *DUO1*, a gate-controlled

[1]State Key Laboratory of Genetic Engineering, School of Life Sciences, Fudan University, Shanghai, China. [2]The Leibniz Institute for Plant Genetics and Crop Plant Research, Stadt Seeland, Germany. e-mail: zhengbl@fudan.edu.cn

transcription factor of PMII[19–23]. ARID1 exhibits nuclear body-like localization and it is constitutively expressed in the VC, but accumulates in early GC rather than in the SC[24,25]. Moreover, the accumulation of ARID1 in the GC is partially dependent on small RNA effector protein AGO9 and both ARID1 and AGO9 are involved in small RNA movement to reinforce heterochromatin silencing during pollen mitosis[26,27]. In this study, we show that ARID1 is essential for maintenance of H3K9me2 in the GC and SC and gradual eviction in the VC during pollen mitosis via recruiting SUVH6, a pollen-enriched H3K9 methyltransferase. Different from that H3K27me3 is regained in the VC but abolished in the SC after PMII[6,7], H3K9me2 is gradually evicted in the VC after PMI and completely abolished before PMII. In contrast, H3K9me2 is steadily maintained in the GC and SC. Genomic target site analyzes show that ARID1 preferentially binds to those regions with H3K9me2, potentially enhances heterochromatin maintenance in sperm cells. Therefore, the maintenance of H3K9me2 during sperm cell formation is regulated by ARID1 during pollen mitosis.

Results

Differential pattern of H3K9me2 between VC and GC/SC initiates after PMI

Immunofluorescence analyzes show that H3K9me2 is highly detected in the SC but absent in the VC of mature pollen[7,28], consistent with that heterochromatin is de-condensed in the VC[7,27,28]. To investigate when and how the differential pattern of H3K9me2 between VC and SC occurs during pollen mitosis, we performed immunofluorescence assay analysis of H3K9me2 using an anti-H3K9me2 antibody throughout pollen mitosis. We show that H3K9me2 signal is clearly detected in microspores that are released after meiosis (Fig. 1a), indicating that H3K9me2 is normally deposited or maintained after meiosis. However, after PMI, more than 90% of early bicellular pollen (EBP) still exhibit H3K9me2 signal in the VC (Fig. 1a, b), although the intensity of H3K9me2 is reduced compared to that of microspore (Fig. 1c). Even though at middle stage of bicellular pollen (MBP), H3K9me2 signal is detected in the VC of more than 50% of bicellular pollen (Fig. 1b). Of note, the intensity of H3K9me2 is successively reduced in the VC of MBP (Fig. 1c). As pollen development goes on, H3K9me2 signal is gradually evicted from the VC until completely abolishes at late stage of bicellular pollen (LBP) (Fig. 1a). After PMII, H3K9me2 signal is only detected in the SC (Fig. 1a), which is consistent with previous findings[7,27,28]. In contrast, H3K9me2 signal is easily detected in the GC after PMI (Fig. 1a). Importantly, the signal intensity of H3K9me2 in the GC is much higher than that of the VC and is comparable to that in microspore (Fig. 1c), indicating that H3K9me2 is actively inherited rather than passively diluted into two daughter cells. Most importantly, the level of H3K9me2 reaches the maximum in early GC, and can be maintained till the SC (Fig. 1c). These observations show that the eviction of H3K9me2 in the VC is a progressive process while the maintenance of H3K9me2 in the GC and SC is a constant process. Notably, a recent study using ChIP-seq analysis shows that H3K9me2 is detected in both VN and SN, arguing whether H3K9me2 is really lost in the VN[8]. Based on strong signals of H3K27me3 in the VN using immunofluorescence[6,7], it is impossible that the absence of H3K9me2 in the VN is caused by an artifact of chromatin de-condensation. Instead, the loss of H3K9me2 in the VN is consistent with the disappearance of pericentromeric foci[7,26,28].

The fact that H3K9me2 is differentially regulated between VC and GC/SC, indicates that certain regulators of H3K9me2 might be involved in the regulation of H3K9me2 dynamics. Based on our previous finding that ARID1 is highly dynamic between the GC and the VC at EBP and MBP[24], we hypothesize that ARID1 is involved in the regulation of differential pattern formation of H3K9me2 during pollen mitosis. Consistent with the finding that *DUO1* is immediately expressed in the GC after PMI[19], which is synchronous with the timing of ARID1 accumulation in the GC (Supplementary Fig. 1a). As pollen mitosis goes on, ARID1 is gradually decrease in the GC of EBP and MBP and finally eliminates in the GC of LBP and SC (Supplementary Fig. 1a). Statistical analyzes confirm that ARID1 and H3K9me2 spatiotemporally overlaps during pollen mitosis (Fig. 1b, d).

Given that ARID1 exhibited nuclear body localization (Supplementary Fig. 1)[24], alike to H3K9me2 foci (Fig. 1a), we wanted to examine whether ARID1-nuclear body is co-localized with H3K9me2 foci during pollen mitosis. As shown in Fig. 1e, ARID1 nuclear bodies are completely located within H3K9me2 foci from microspore to EBP and MBP. However, ARID1-nuclear body and H3K9me2 foci separate with each other from LBP till mature pollen (Fig. 1e), since both are reciprocally detected in two types of cells at LBP and MP (Fig. 1e). Statistical analyzes show that more than 80% of microspores and EBP/MBP exhibit overlapped localization with both ARID1 nuclear body and H3K9me2 foci (Fig. 1f). Once the differential pattern of H3K9me2 between two cell types has been formed, less than 10% of ARID1 nuclear body locates within H3K9me2 foci (Fig. 1f). Of note, the ARID1 body partially co-localizes with the nucleolus marker Fibrillarin (FIB2) during pollen mitosis, particularly in the VC of mature pollen (Supplementary Fig. 1b). Collectively, the spatiotemporal overlap among ARID1 and H3K9me2 in EBP/MBP indicates that ARID1 has a potential to regulate H3K9me2 during pollen mitosis.

ARID1 promotes the maintenance of H3K9me2 in the GC/SC but slows down its eviction in the VC during pollen mitosis

To understand whether the differential pattern formation of H3K9me2 between the VC and the GC/SC is regulated by ARID1, we first performed western blotting assay to access the total levels of H3K9me2 in mature pollen from Col-0 and *arid1* (hereafter named *arid1*), a previously isolated mutant[24]. The result shows that the level of H3K9me2 is significantly reduced in *aird1* mutant pollen (Supplementary Fig. 2a). Then, we examined the dynamics of H3K9me2 throughout pollen mitosis using anti-H3K9me2 immunofluorescence assays. In microspore (UM), the signal intensity of H3K9me2 is comparable between Col-0 and *arid1* (Fig. 2a, the top panels), indicating that ARID1 is not involved in H3K9me2 maintenance in the meiotic products. However, the signal intensities of H3K9me2 in the VC of EBP and MBP in *arid1* are much lower than those in Col-0 (Fig. 2a, b). In LBP and mature pollen, H3K9me2 signals are barely detected in both Col-0 and *arid1* (Fig. 2a, b), indicating that the eviction extent of H3K9me2 is aggravated when the function of ARID1 is impaired. In contrast to that the comparable intensities of H3K9me2 among UM, the GC, and the SC in Col-0, H3K9me2 intensity becomes immediately decreased in early GC of *arid1* after PMI (Fig. 2a, c). Moreover, the decrease trend of the H3K9me2 level lasts in the GC of whole bicellular stages and becomes even worse in the SC of *arid1* (Fig. 2a, c). Importantly, the intensities of H3K9me2 in the GC are much lower than that in microspore in *arid1* (Fig. 2a, c). Of note, H3K9me2 intensities are comparable between the GN and SCN, indicating H3K9me2 is maintained in the GC/SC after PMI. H3K9me2 intensities in the SC of *arid1* are indeed much lower than that in the SC of Col-0 (Fig. 2a, c), indicating that ARID1 is necessary for maintaining a proper level of H3K9me2 in the SC, although ARID1 protein only accumulates in the VC after PMII. The lower level of H3K9me2 in the GC of *arid1* is unable to be restored by pollen mitosis progression, further supporting an irreplaceable role of ARID1 in maintenance of H3K9me2 in the GC/SC after PMI.

Since the loss of ARID1 causes reduced levels of H3K9me2 in the GC and SC, we hypothesize that overexpression of ARID1 in the sperm cells may increase the levels of H3K9me2. Therefore, we generated transgenic lines expressing *ARID1-YFP* driven by the GC/SC-specific promoter *HTR10* (*proHTR10::ARID1-YFP*, (Supplementary Fig. 2b) and examined the levels of H3K9me2 in mature pollen. Immunofluorescence analyzes show that the levels of H3K9me2 in the sperm cells of *proHTR10::ARID1-YFP* pollen are significantly higher than those in Col-0 (Fig. 2d; and Supplementary Fig. 2c), further supporting the

https://doi.org/10.1038/s41467-024-51513-4

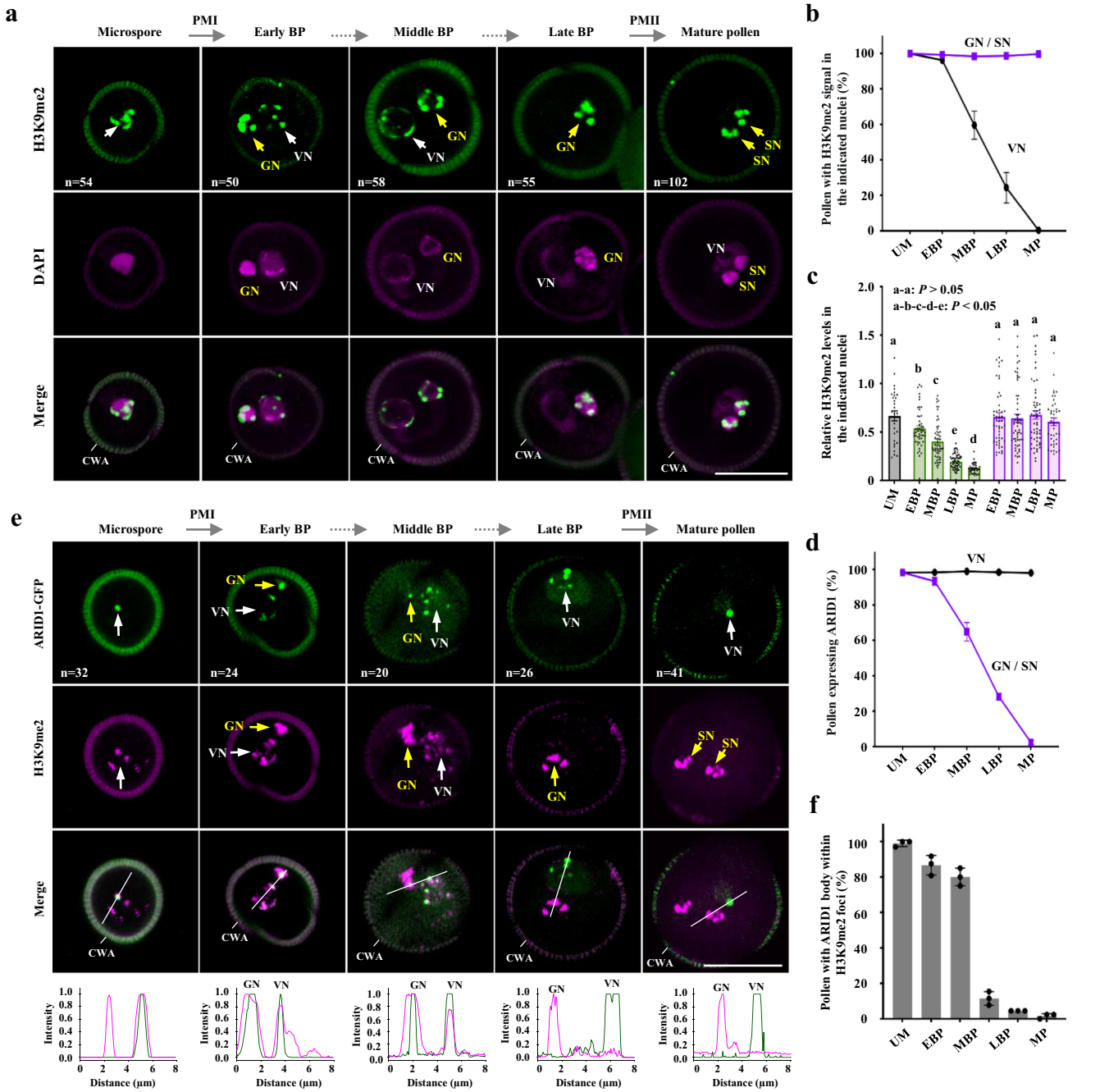

Fig. 1 | H3K9me2 is differentially regulated between VC and GC/SC after PMI. **a** Immunofluorescence analysis showing H3K9me2 signal during pollen development in Col-0. DAPI staining indicates location of nucleus of indicated cells. **b** Ratio of pollen with H3K9me2 signal in the indicated nuclei of (**a**). Data are presented as mean ± SE, n = 3. More than 50 pollen were analyzed for each. **c** Quantitative analysis of H3K9me2 intensities in (**a**). The H3K9me2 intensities were calculated relative to DAPI. Each dot represents one pollen. *P* values calculated using one-sided ANOVA followed by individual two-sample Tukey tests. Boxplots marked as a, b, c, d and e are significantly different between groups ($P < 0.05$) but not within the group ($P > 0.05$). More than 30 pollen were analyzed for each. **d** Statistical analysis showing dynamic localization of ARID1-GFP during pollen development in Supplementary Fig. 1a. Data are presented as mean ± SE, n = 3. More than 20 pollen were analyzed for each. **e** ARID1 nuclear body is colocalized with H3K9me2 foci before PMII. The transgenic plants expressing *proARID1::ARID1-GFP* was used for anti-H3K9me2 immunofluorescence assay. The intensity profiles of ARID1-GFP and H3K9me2 signals are displayed along the white lines (shown in the merged panels). **f** Statistical analysis showing the ratio of pollen with ARID1 nuclear bodies co-localizing with H3K9me2 foci in (**e**). Data are presented as mean ± SE, n = 3. VN, vegetative nucleus; GN, generative nucleus; SN, sperm nuclei. PMI and PMII, pollen mitosis I and II. UM, unicellular microspore; EBP, MBP, and LBP indicate early, middle, and late bicellular pollen, respectively; MP, mature pollen. CWA, cell wall autofluorescence. Scale bar, 10 μm. n indicates the number of analyzed pollen.

role of ARID1 in the regulation of H3K9me2. Changes in heterochromatin levels may lead to alterations in nucleus size, prompting us to further investigate the biological consequence of H3K9me2 maintenance in pollen. We compared the changes in nucleus size of the vegetative nucleus and sperm nucleus in the *arid1* mutant, ARID1 overexpression lines (the transgenic plants carrying *proHTR10::ARID1-YFP*), and the *suvh456* mutant (which involves three H3K9 methyltransferase SUVH4, SUVH5, SUVH6 in Arabidopsis). DAPI staining analyzes show that the nucleus size of the VN in both *arid1* and *suvh456* mutants is significantly larger than those in Col-0 (Fig. 2e). Instead,

https://doi.org/10.1038/s41467-024-51513-4

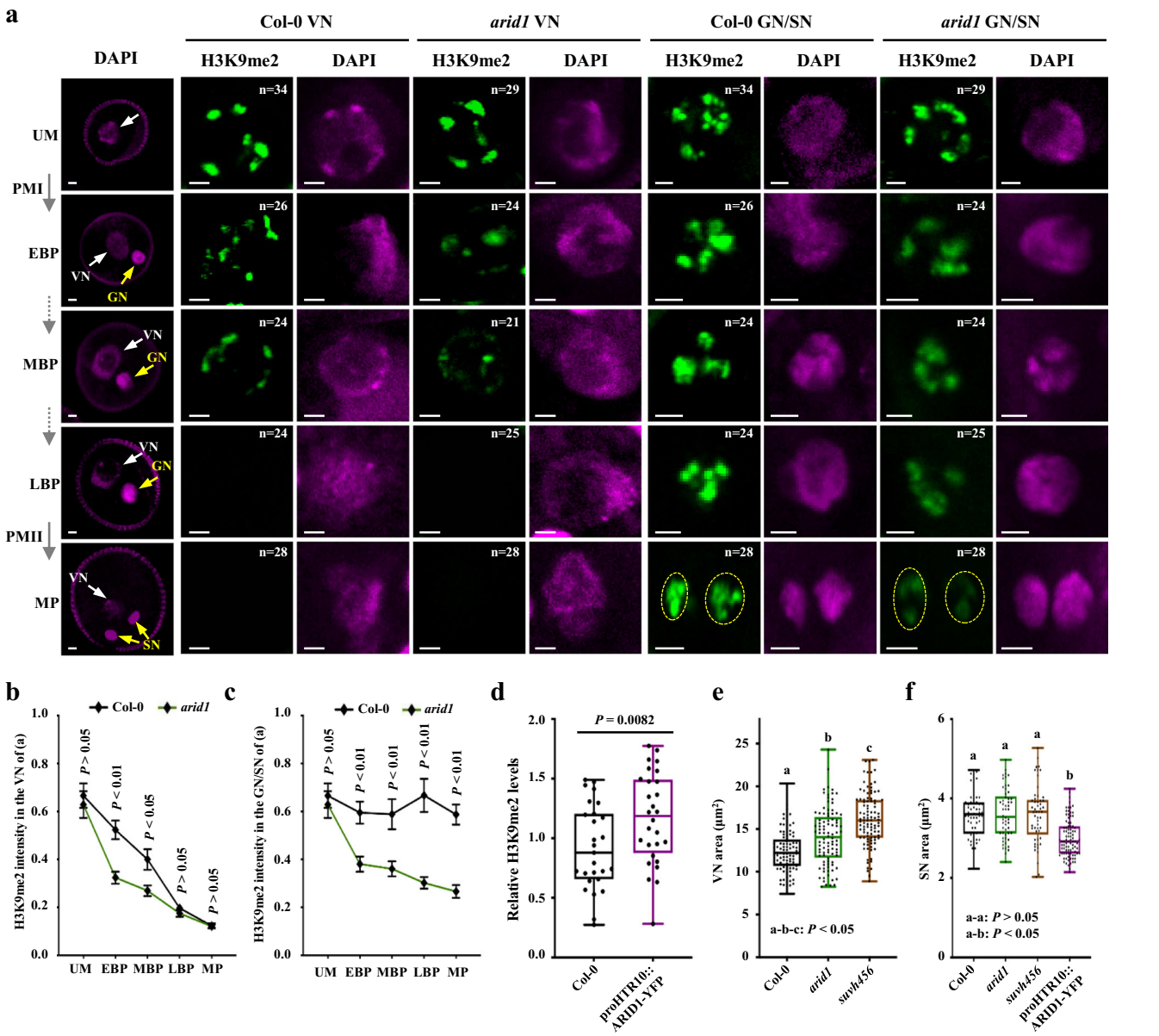

Fig. 2 | ARID1 regulates the maintenance of H3K9me2 during pollen mitosis. **a** Immunofluorescence analysis showing H3K9me2 dynamics between Col-0 and *arid1* during pollen development. VN, vegetative nucleus; GN, generative nucleus; SN, sperm nuclei. Yellow dashed circle indicate two SN. PMI, pollen mitosis I; PMII, pollen mitosis II; UM, unicellular microspore; EBP, early bicellular pollen; MBP, middle bicellular pollen; LBP, late bicellular pollen; MP, mature pollen. n indicates the number of analyzed pollen. Scale bar, 2 μm. **b**–**c** Quantitative analysis of (**a**). H3K9me2 intensities in VN (**b**) and GN/SN (**c**) are calculated relative to the DAPI. Data are presented as mean ± SE, n = 3. Statistical significance was determined by unpaired two-sided Student's *t*-test. More than 20 pollen were analyzed for each. **d** Quantitative analysis of H3K9me2 intensities in the SN of the transgenic plants expressing *proHTR10::ARID1-YFP*. Relative intensities were calculated relative to DAPI. Each dot represents one pollen. Boxplots indicate minimum and maximum values, as well as 25th, 50th and 75th quartiles. *P* value was determined by unpaired two-sided Student's *t*-test. More than 25 pollen were analyzed for each. **e**, **f** Quantitative analysis of VN (**e**) and SN (**f**) area in indicated genotypes. Each dot represents a VN or SN. Boxplots indicate minimum and maximum values, as well as 25th, 50th and 75th quartiles. *P* values were calculated using one-sided ANOVA followed by individual two-sample Tukey tests. Boxplots marked as **a**, **b** and **c** are significantly different between groups ($P < 0.05$) but not within the group ($P > 0.05$). More than 70 VN or SN were analyzed for each.

overexpression of ARID1 results in a smaller nucleus size in sperm cells (Fig. 2f). Therefore, we conclude that the regulation of H3K9me2 is essential for maintaining a proper nucleus size of the two cell types in pollen. Notably, the nucleus size of SN is comparable between Col-0 and *arid1* or *suvh456* (Fig. 2f), suggesting that other factors may compensate for the roles of ARID1 and SUVH4, SUVH5, SUVH6 in the regulation of nucleus size in sperm cells.

Through protein blast searching, we identified another gene, *At4G11400*, which bears similarity to ARID1, containing both an ARID domain and an ELM2 domain. For convenience, here we name it ARID6. In comparison to the pattern of specific enrichment observed for ARID1 in pollen (Supplementary Fig. 2d), ARID6 exhibits much lower expression levels across various tissues (Supplementary Fig. 2d). Then, we examined whether ARID6 is redundant with ARID1 to regulate H3K9me2 maintenance in pollen. immunofluorescence assay shows that no difference is detected in *arid6* compared to that in Col-0 (Supplementary Figs. 2e, f). Moreover, H3K9me2 intensity is comparable between *arid1* and *arid1 arid6*, supporting that ARID6 is not required for regulation of H3K9me2 in pollen (Supplementary Figs. 2e, f). Taken together, our findings suggest that ARID1 plays a

crucial role in regulating the differential pattern formation of H3K9me2. Furthermore, ARID1-dependent H3K9me2 regulation controls the nucleus size of the two cell types in pollen.

ARID1 is necessary for H3K9me2 maintenance in pollen at a genome-wide scale

To substantiate the influence of ARID1 on H3K9me2 after pollen mitosis, and assess which specific sites are dependent on ARID1 for H3K9me2 at a genome-wide scale, we collected mature pollen from Col-0 and *arid1* to perform ChIP-seq experiments using anti-H3K9me2 antibody. Of note, due to the difficulty of the GC and the VC purified from bicellular pollen for ChIP-seq analysis, it is accepted to judge the influence of ARID1 on the final output of H3K9me2 using mature pollen. ChIP-seq sequencing data are highly reproducible (Supplementary Figs. 3a, b), and the detailed alignment statistics for each library are available in Supplementary Table 1. In total, we identified 3692 peaks of H3K9me2 in mature pollen (Supplementary Data 1). Based on the extent of H3K9me2 dependence on ARID1 in mature pollen, we define the H3K9me2 targets into two clusters. Cluster 1, consisting of 790 peaks, exhibits a higher level of H3K9me2 in Col-0, which is overall downregulated to 16.16% in *arid1* (Fig. 3a, b). In contrast, Cluster 2, consisting of 2902 peaks, exhibits a lower level of H3K9me2 in Col-0 but is downregulated to 49.72% in *arid1* (Fig. 3a–c). Accordingly, Cluster 1 peaks are strictly located at constitutive pericentromeric regions and Cluster 2 are slightly expanded outside of pericentromeric regions (Supplementary Fig. 3c). IGV analysis confirm that six representative TE loci from Cluster 2 are dependent on ARID1, respectively (Fig. 3d). These results indicate that ARID1 is widely involved in the regulation of H3K9me2 in mature pollen.

To further understand the relationship between ARID1 and heterochromatin maintenance, we analyzed the correlation between ARID1-dependent H3K9me2 regions and other heterochromatin epigenetic marks, including H2A.W[29], H3K27me1[6], and DNA methylation[14]. The results show that the dependency of H3K9me2 on ARID1 is clearly anti-correlated with the levels of H3K27me1, H2A.W and DNA methylation (Fig. 3e, f). Unlike Cluster 1 peaks, Cluster 2 peaks, which have a stronger dependency on ARID1, harbor fewer heterochromatic epigenetic marks (Fig. 3e, f). To investigate whether H3K9me2 target loci are different between somatic and reproductive cells, we compared the H3K9me2 peaks between mature pollen and leaves of three-week-old plants[30]. We show that in mature pollen, 2,075 out of 3,692 H3K9me2 loci are de novo peaks, with more than 95% of these de novo H3K9me2 loci being ARID1-dependent (Supplementary Fig. S3d). Heatmap analyzes confirm that theses de novo H3K9me2 peaks are indeed enriched in mature pollen compared to leaf tissues (Fig. 3g). Among these de novo H3K9me2 peaks in pollen, 78.51% (1629) are located within TE regions (Fig. 3h), which contain fewer LTR/Gypsy elements and more RC/Helitron elements compared to the genome (Supplementary Fig. S3e). While most of the de novo H3K9me2 peaks are from TE regions, 17.20% (357) are gene peaks and 2.99% (62) are noncoding RNA peaks (Fig. 3h). However, these de novo H3K9me2 loci in pollen barely overlap with pollen tube genes previously found to be silenced by DNA methylation and H3K9me2 but reactivated after pollen germination[9,31] (Supplementary Fig. S3f). Consistent with the lack of relevance between pollen tube genes and ARID1-dependent de novo H3K9me2 loci in pollen, ARID1 shows minimal binding to these pollen tube genes (Supplementary Fig. S3f), indicating that the role of de novo ARID1-dependent H3K9me2 is limited in the expression of these pollen tube genes. Collectively, our results conclusively demonstrate that ARID1 is widely involved in the regulation of H3K9me2 in pollen.

ARID1 is recruited to H3K9me2 target genomic regions in pollen

As ARID1 co-localizes within H3K9me2 foci (Fig. 1) and ARID1 regulates the differential pattern formation of H3K9me2 during pollen mitosis (Figs. 2, 3), we hypothesized that ARID1 might be recruited to H3K9me2 targets. To test whether it is the case, we characterized the occupancy of ARID1 at a genome-wide scale, and compared the overlap between ARID1- and H3K9me2-bound peaks in mature pollen. We performed Cut & Tag experiments using anti-GFP antibody in mature pollen of the transgenic plants expressing *proARID1::ARID1-GFP* followed by high-throughput sequencing. DAPI staining indicates that ARID1-GFP can rescue the single sperm-like cell defect phenotype of the *arid1* mature pollen, demonstrating the in vivo function of ARID1-GFP (Supplementary Fig. 4a). High throughput sequencing data are highly reproducible (Supplementary Fig. 4b). We employed the MACS2 (model-based analysis for ChIP-seq) peaking calling method to identify ARID1-bound peaks. In total, we identified 2431 ARID1 peaks (Supplementary Data 2). Among the 2431 ARID1 peaks, 25.46% of (619/2431) and 71.25% (1732/2431) of peaks are mapped to genic regions and transposable elements (TE) regions, respectively (Fig. 4a). A motif discovery analysis using MEME-ChIP identified two enriched motifs from the sequences of ARID1-bound peaks: ATACTCRAYCAY (R represents A or G; Y represents C or T; E-value = 3.0e-081) and AGCTTTGWTRGT (W represents A or T; E-value = 8.9e-078) (Supplementary Fig. 4c). Moreover, regions with higher densities of these two motifs are preferentially bound by ARID1 (Supplementary Fig. 4d), indicating that ARID1 is recruited to chromatin with specific sequence preferences.

Since ARID1 can be recruited to TE chromatin and ARID1 is required for maintenance of H3K9me2 in pollen (Figs. 2, 3), we then examined to what extent ARID1 and H3K9me2 co-target in the genome. Among the 2431 ARID1 peaks, 79.4% (1931/2431) overlap with H3K9me2 peaks (Fig. 4b). Specifically, 98.7% (1710/1732) of ARID1-bound TE peaks overlap with H3K9me2 peaks (Fig. 4c), indicating that ARID1 preferentially binds to TE regions marked by H3K9me2. Metaplots and heatmap analyzes show that ARID1-bound TE peaks exhibit a higher level of H3K9me2, while ARID1-bound gene peaks barely overlap with H3K9me2 peaks (Fig. 4d, e). The distribution pattern along the chromosomes of both ARID1-bound TE peaks is similar to that of H3K9me2-bound peaks (Fig. 4f), indicating that ARID1 prefers to bind pericentromeric regions. To investigate the correlation between ARID1 binding intensity and H3K9me2 targeting, we categorized the ARID1-bound peaks into two types based on binding intensity. Type 1 exhibits high affinity for ARID1 at 897 peaks, while Type 2 shows low affinity for ARID1 at 1534 peaks (Fig. 4g). Moreover, the higher the ARID1 binding intensity, the higher the level of H3K9me2 targeting (Fig. 4h), indicating the existence of a positive regulatory loop between ARID1 and H3K9me2 in these genomic regions. More importantly, by comparing the H3K9me2 levels of the two types of ARID1 peaks in Col-0 and *arid1* mutant pollen, we found that Type 2 regions exhibit 34.15% less H3K9me2, while Type 1 regions exhibit 5.69% less H3K9me2 (Fig. 4h). Similar to ARID1-dependent H3K9me2 peaks, the binding sites of ARID1 tend to enrich all heterochromatin epigenetic marks, including H3K9me2, H2A.W, and H3K27me1 but not H3K27me3 (Fig. 4i). Taken together, we conclude that ARID1 is recruited to H3K9me2-targeted genomic regions and both promote each other reciprocally.

ARID1 interacts with SUVH6 in SRA domain-dependent manner in pollen

Since ARID1 is required for the maintenance of H3K9me2 during pollen mitosis (Figs. 2, 3), and ARID1 tends to bind to H3K9me2-targeted regions (Fig. 4), is there a possibility that ARID1 interacts with the three H3K9 methyltransferases SUVH4, SUVH5, and SUVH6 to regulate the differential regulation of H3K9me2 during pollen mitosis? To answer this question, we first examined expression of *SUVH4, 5, 6*. The public RNA-seq database[32] show that, compared to a much lower expression of *SUVH4* and *SUVH5*, *SUVH6* is the most abundant one in mature pollen (Supplementary Fig. 5a). Moreover, a recent RNA-seq analysis using VN and SN[33] shows that *SUVH4* is barely detected, *SUVH6* is

https://doi.org/10.1038/s41467-024-51513-4

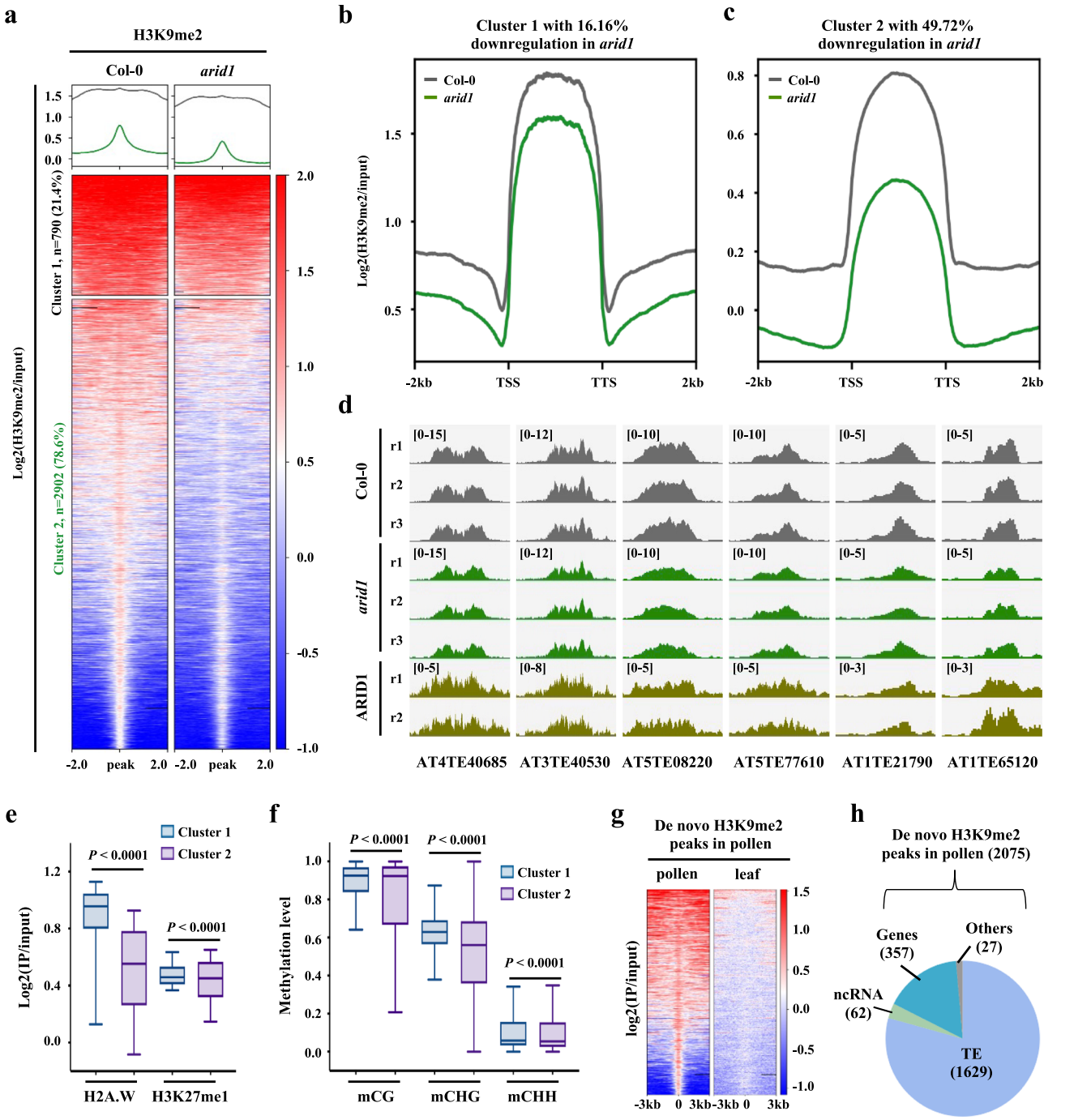

Fig. 3 | ARID1 is necessary for loci characterized by low levels of H3K9me2 in pollen. a Heatmap analysis of H3K9me2 peaks in mature pollen of Col-0 and *arid1*, respectively. Cluster 1 and 2 peaks are shown based on the extent decreased in *arid1*. The abundance of H3K9me2 is shown as log2(H3K9me2/input). **b**–**c** Metaplots showing the overall occupancy abundance of H3K9me2 at two clusters of peaks in Col-0 and *arid1*. The abundance of H3K9me2 is shown as log2(H3K9me2/input). **d** Genome browser showing the abundance of H3K9me2 and ARID1 at TE loci. Numbers in the brackets indicate the value of the normalized expression levels (CPM). rep1, rep2 and rep3 indicate three sets of replicates. **e** Box plots show the abundance of H2A.W and H3K27me1 at Cluster 1 and Cluster 2 peaks, respectively. ChIP signals are shown as log2(IP/input). **f** Box plots show the levels of CG, CHG and CHH methylation at Cluster 1 and Cluster 2 peaks, respectively. In (**e**) and (**f**), Boxplots indicate minimum and maximum values, as well as 25th, 50th and 75th quartiles. *P* value was determined by unpaired two-sided Student's *t*-test. **g** Heatmap analysis showing abundance of de novo H3K9me2 peaks in pollen compared to somatic cells. The abundance is represented as log2(IP/input). **h** Pi chart illustrating genomic annotation for de novo H3K9me2 peaks in pollen. 2075 de novo peaks were mapped to the TAIR10 genome. TE, transposable elements.

mainly detected in the VN while *SUVH5* accumulated in the SN (Supplementary Fig. 5b). To further characterize the spatiotemporal patterns of SUVH4, SUVH5, and SUVH6 proteins during pollen development, we transformed three plasmids into the *suvh456* mutant which expressed *YFP-SUVH4/5/6* driven by their endogenous promoters, respectively. The introduction of either *YFP-SUVH4/5/6* construct into the *suvh456* mutant can fully restore the H3K9me2 level in the SC (Supplementary Fig. 5c), indicating that YFP-SUVH4/5/6 are biologically functional, and SUVH4, SUVH5, SUVH6 play a redundant role in H3K9me2 maintenance in mature pollen. Then we observed

https://doi.org/10.1038/s41467-024-51513-4

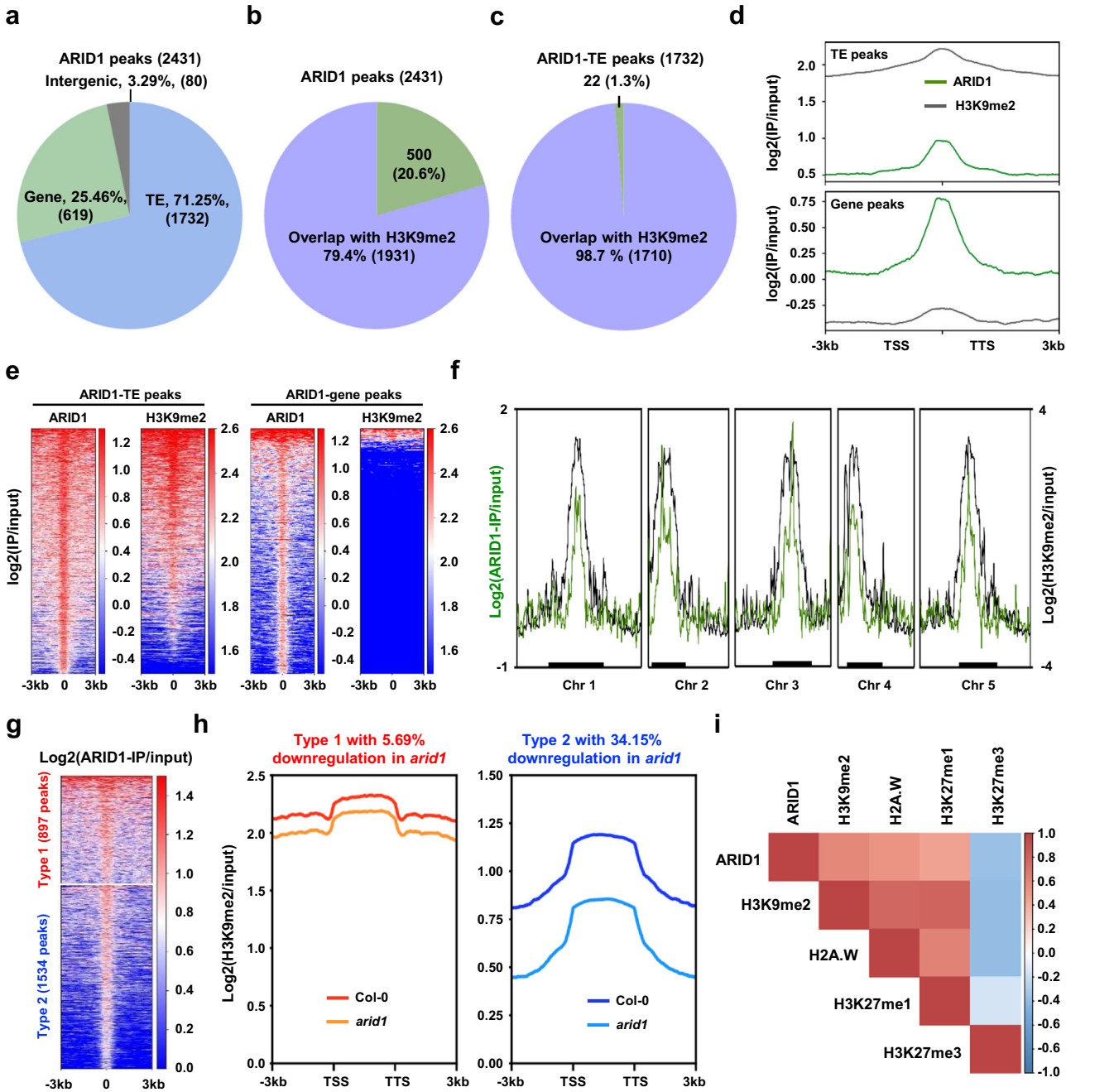

Fig. 4 | Genomic binding of ARID1 overlaps with H3K9me2 in pollen. a Pi chart showing genomic annotation for ARID1 binding peaks. Total 2431 peaks identified by Cut & Tag analysis (IP/input) were mapped to the TAIR10 genome. TE, transposable elements. **b**–**c** Pi chart showing percentage of ARID1 binding peaks with H3K9me2 (**b**) and ARID1 binding TE peaks with H3K9me2 (**c**). **d** Metaplots showing overall occupancy of ARID1 and H3K9me2 at ARID1 binding TE and gene peaks, respectively. The abundance is represented as log2(IP/input). **e** Heatmap analysis showing overall abundance of ARID1 and H3K9me2 at ARID1-bound TE peaks (left) and gene peaks (right). The abundance is represented as log2(IP/input). **f** Distribution of ARID1 and H3K9me2 occupancy along the five chromosomes. The black lines indicate the pericentromeric region of each chromosome. The abundance is represented as log2(IP/input). **g** Heatmap analysis showing abundance of ARID1 at all binding peaks. ARID1 peaks were grouped into two categories based on binding intensity of ARID1. The abundance is represented as log2(IP/input). **h** Metaplots showing abundance of H3K9me2 at ARID1-bound two type peaks, respectively. The abundance is represented as log2(IP/input). **i** Pearson correlation matrix among ARID1 occupancy, H3K9me2, H2A.W, H3K27me1, and H3K27me3.

subcellular localization of YFP-SUVH4/5/6 throughout pollen mitosis by fluorescence microscope analysis. As shown in Fig. 5a and Supplementary Fig. 5d, YFP-SUVH4, SUVH5, and SUVH6 are all detected in microspores. However, more than 90% EBP and 50–70% of MBP exhibit YFP signals in both VC and GC, respectively (Fig. 5a, b, Supplementary Fig. 5d). In contrast, only ~10% of LBP exhibit YFP signals in the GC, although all three proteins accumulate normally in the VC of LBP (Fig. 5a; Supplementary Fig. 5d), indicating SUVH4, SUVH5, and SUVH6, alike to ARID1, are evicted gradually from the GC as PMII is approaching. After PMII, YFP-SUVH5 and SUVH6 are only detected in the VC but not in the SC (Fig. 5a; Supplementary Fig. 5d), and YFP-SUVH4 is invisible in mature pollen (Supplementary Fig. 5d). Of note, although *SUVH5* mRNA is detected in the SN (Supplementary Fig. 5b), YFP-SUVH5 protein is only observed in the VN (Supplementary Fig. 5d).

https://doi.org/10.1038/s41467-024-51513-4

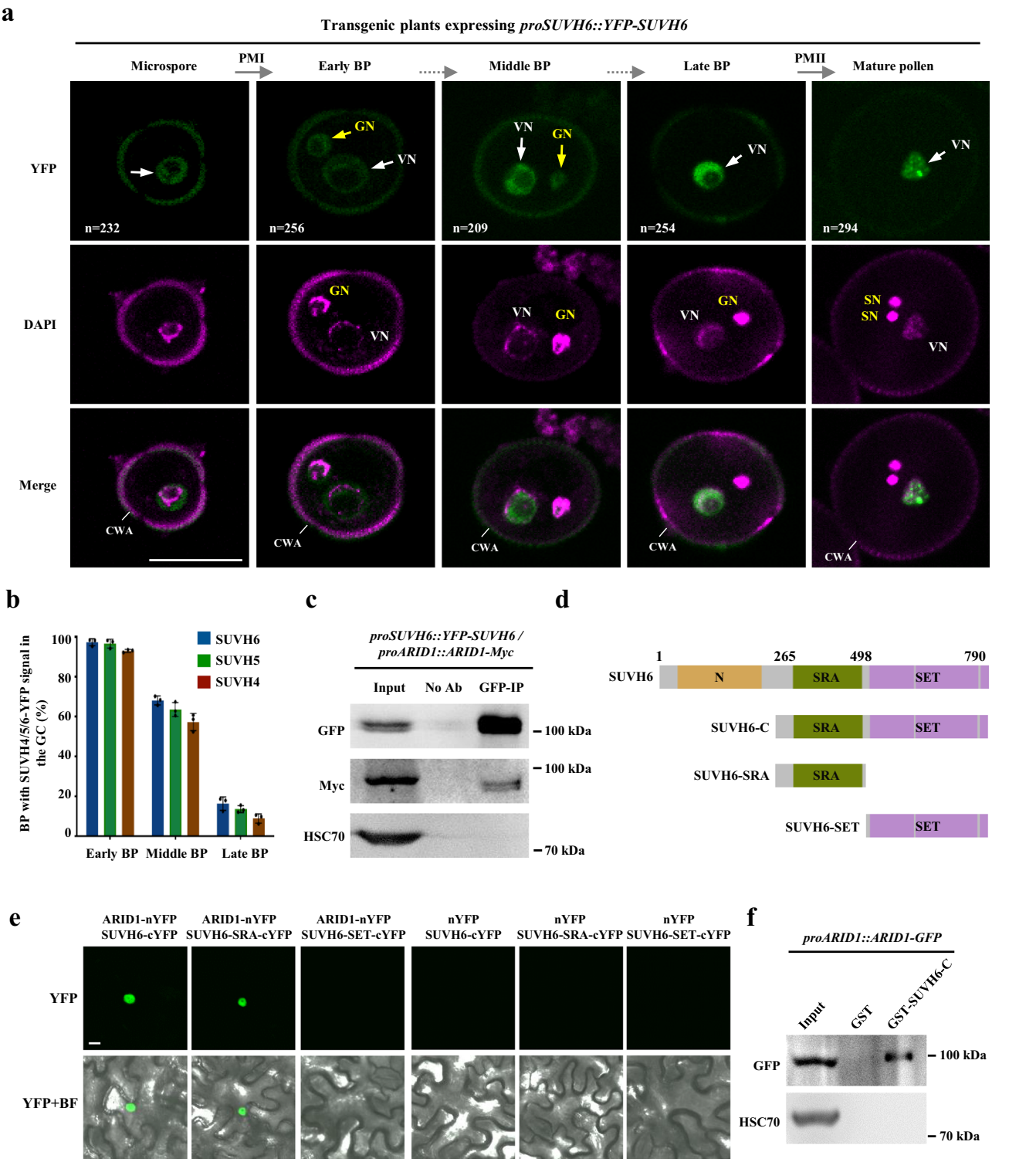

Fig. 5 | ARID1 interacts with SUVH6 that exhibits ARID1-like spatiotemporal dynamics during pollen mitosis. a Subcellular localization of SUVH6 during pollen development. BP, bicellular pollen. VN, vegetative nucleus; GN, generative nucleus; SN, sperm nuclei. PMI, pollen mitosis I; PMII, pollen mitosis II; White and yellow arrows indicate YFP-SUVH6 signal in VN and GN/SN, respectively. n, total analyzed number of pollen. CWA, cell wall autofluorescence. Scale bar, 10 μm. **b** Statistical analysis of (**a**) showing incidence of different stages of bicellular pollen expressing YFP-SUVH4, -SUVH5/, -SUVH6 in generative cell (GC). Data are presented as mean ± SE, n = 3. **c** Co-IP experiments showing the ARID1-SUVH6 interaction in pollen. Mature pollens of the transgenic plants were used. HSC70 was used as the negative control. **d** Schematic showing the protein domain structures of SUVH6. Numbers indicate the position of amino acids. **e** BiFC analysis of the ARID1-SUVH6 interaction. ARID1 and full length or truncated SUVH6 were analyzed. Scale bar, 10 μm. **f** Semi in vivo pull down assay showing interaction between ARID1 and SUVH6. GST-SUVH6-C was incubated with mature pollen lysates of the *proARID1::ARID1-GFP* transgenic plants. GST and HSC70 were used as the negative control. The indicator of molecular weight is shown in (**c**) and (**f**). Source data are provided as a Source Data file.

Statistical analyzes confirm that the incidence of three proteins accumulation in the GC is gradually decreased as pollen mitosis goes on (Fig. 5b). The phenomenon of SUVH4, SUVH5, and SUVH6 accumulation in early GC is coincide with the finding that H3K9me2 is differentially regulated after PMI.

Given that ARID1 and SUVH4, SUVH5, and SUVH6 spatiotemporally overlap during pollen mitosis (Figs. 1 and 5), we hypothesized that ARID1 interacts with SUVH4, SUVH5, and SUVH6 in pollen. Because SUVH6 is the most abundant one, the following experiments were performed by focusing on SUVH6. Co-IP experiment was performed using mature pollen of the transgenic plants co-expressing *proSUVH6::YFP-SUVH6* and *proARID1::ARID1-Myc*. The results show that ARID1 interacts with SUVH6 in mature pollen (Fig. 5c). Given that SUVH6 harbors N, SRA, and SET domains (Fig. 5d), we then performed BiFC experiments to test which domain mediates the ARID1-SUVH6 interaction. By introducing the paired plasmids expressing the split YFP fused with ARID1 and SUVH6, SUVH6-SRA, SUVH6-SET, respectively, into tobacco leaves, we show that the SRA domain rather than the SET domain of SUVH6 is necessary for the interaction with ARID1 (Fig. 5e). Lastly, we performed semi in vivo pull down experiments using recombinant GST-SUVH6-C protein (including SRA and SET domains) to incubate with pollen lysates of the transgenic plants expressing *proARID1::ARID1-GFP*. In contrast to the negative control GST, GST-SUVH6-C pulled down ARID1-GFP (Fig. 5f), further supporting that the SRA domain is sufficient to mediate the interaction with ARID1. Therefore, we conclude that ARID1 and SUVH4, SUVH5, and SUVH6 not only overlap spatiotemporally but also interact with each other, implying that ARID1 promotes the maintenance of H3K9me2 via regulating histone methyltransferases.

ARID1 promotes SUVH6 accumulation and recruits SUVH6 to H3K9me2 loci

Since ARID1 regulates maintenance of H3K9me2 during pollen mitosis (Fig. 2), and ARID1 interacts with SUVH6 (Fig. 5), we hypothesize that ARID1 may facilitate spatiotemporal pattern formation of SUVH6 and help SUVH6 to access to the substrate. To test this hypothesis, we introduced the *suvh456* mutant plants expressing *proSUVH6::YFP-SUVH6* into *arid1* by genetic crosses. We show that the signal intensity of YFP-SUVH6 is comparable in microspores between *suvh456* and *suvh456 arid1* (Fig. 6a, b). However, compared to that in *suvh456*, the intensities of YFP-SUVH6 in *suvh456 arid1* are slightly decreased in the VC of EBP/MBP (Fig. 6a, b). Moreover, the proportion of bicellular pollen with YFP-SUVH6 signal in the GC is significantly decreased in *suvh456 arid1* (Fig. 6a, c), indicating that ARID1 is required for maintenance of SUVH6 spatiotemporal pattern after PMI, especially accumulation in the GC. Surprisingly, SUVH6 clearly accumulates in the VC with a condensed big dot in *suvh456* (Fig. 6a), and the condensed big dot of YFP-SUVH6 in the VC is completely diffused in mature pollen of *suvh456 arid1* (Fig. 6a). In contrast, *suvh4/5/6* mutation has no effects on expression and localization of ARID1 in pollen (Supplementary Fig. 6a). We thus conclude that ARID1 is essential for subcellular localization of SUVH6 during pollen mitosis.

To conclusively demonstrate the dependency of subcellular localization of SUVH6 on ARID1, we examined subcellular localization of SUVH6 and ARID1 in the VC of mature pollen from the doubly transgenic plants expressing *proSUVH6::YFP-SUVH6* and *proARID1::ARID1-RFP*. As expected, the big dot of SUVH6 and ARID1 nuclear body are completely co-localized in the VC of mature pollen (Fig. 6d). To understand the biological significance of SUVH6 regulation by ARID1, we performed ChIP-seq experiments on SUVH6 using an anti-GFP antibody in mature pollen of transgenic plants expressing *proSUVH6::YFP-SUVH6* in Col-0 and *arid1*, followed by high-throughput sequencing. The total protein levels of YFP-SUVH6 in mature pollen are comparable between Col-0 and *arid1* (Supplementary Fig. 6b), excluding the possibility that the reduced occupancy of SUVH6 at TE loci is caused by the reduced level of SUVH6 in mature pollen. 154 SUVH6-bound peaks were identified (Supplementary Data 3), and among these, 86.4% (133/154) and 84.4% (130/154) of peaks overlap with H3K9me2 and ARID1 peaks, respectively (Supplementary Fig. 4c). We then examined whether the *arid1* mutation affects the recruitment of SUVH6 at these sites. Metaplot analysis shows that the recruitment of SUVH6 is slightly reduced to 9.21% in the *arid1* mutant (Fig. 6e), indicating that ARID1 is involved in the recruitment of SUVH6 at genomic sites.

Since ARID1 interacts with not only histone deacetylases[24] but also SUVH6 (Fig. 5), ARID1 is recruited to H3K9me2 targets in vivo (Fig. 4), and ARID1 is involved in the regulation of H3K9me2 maintenance during pollen mitosis (Figs. 2 and 3), we hypothesized that ARID1 binds to H3K9 directly. By expressing the recombinant GST-ARID1 protein (Supplementary Fig. 6d), we first performed GST pull down assay by incubating equal amounts of GST-ARID1-His with equal amounts of biotinylated histone peptides, followed by western blotting analysis using anti-GST antibody. The results show that ARID1 binds all four histone H3 peptides, but exhibit the strongest binding to the N-terminus of H3 harboring 1^{st}–21^{th} amino acids (Fig. 6f, g). In contrast, ARID1 has the weakest binding to the C-terminus of H3 harboring 21^{th}–44^{th} amino acids, and moderate binding to H3K9me1 and H3K9me2 (Fig. 6f, g). Statistical analyzes from three biological replicates further indicate that ARID1 prefers to bind to unmethylated N terminal H3 containing lysine 9 (H3K9). Next, we measured the affinity of ARID1 with different H3 peptides using fluorescence polarization (FP) analysis. The results further confirm a preferential binding of ARID1 to unmethylated H3 (1–21) (Fig. 6h). The results show that the dissociation constant (Kd) value of ARID1 to H3 (1-21) was measured to be ~4.73 μM (Fig. 6h). In contrast, the Kd values of ARID1 are increased to ~ 21.18 and ~40.46 μM for H3K9me1 and H3K9me2, respectively (Fig. 6h). Notably, the Kd value of ARID1 for H3 (21-44) is ~240.0 μM (Fig. 6h), indicating that ARID1 has preference to K9 rather than K27, albeit both residues can be modified. Taken together, we conclude that ARID1 is involved in regulating H3K9me2 maintenance via not only regulating spatiotemporal expression of SUVH6 but also binding unmethylated H3K9.

To exclude the possibility that VC-depleted H3K9me2 is due to dominant function of IBM1, an H3K9 demethylase in Arabidopsis[34], we examined IBM1 localization during pollen development by analyzing the transgenic plants expressing IBM1-GFP driven by the native *IBM1* promoter[35]. We show that *IBM1* is ubiquitously but lowly expressed throughout pollen mitosis, with a slightly increase in the VC of mature pollen (Supplementary Fig. 7a). However, the level of H3K9me2 in the SC is comparable to between Col-0 and *ibm1* (Supplementary Figs. 7b, c), consistent with the finding that IBM1 is only responsible for H3K9me2 demethylation in euchromatin[34]. Due to the transgenerational progressive effects on the *ibm1* mutant, we performed anti-H3K9me2 immunofluorescence assay using the *ibm1* mutant isolated from the *ibm1*/+ heterozygote plants. Even though *IBM1* was specifically expressed in the SC (Supplementary Fig. 7d), the level of HK9me2 is only slightly increased in the SC (Supplementary Figs. 7e, f), further supporting that IBM1 plays a minor role in the regulation of H3K9me2 in pollen.

Discussion

Epigenetic reprogramming widely occurs during germ cell development from plants to animals, in which heterochromatin silencing is usually attenuated in the companion/nurse cells to reinforce TE silencing in the germ cells[14,28,36]. Typical signs of heterochromatin relaxation in plants are decrease of H3K9me2 and DNA methylation. In contrast to recent progress that H3K27me3 is deposited only in the GC after PMI but completely depleted in the SC and regained in the VC after PMII[7,9], whether and how H3K9me2 is regulated is largely unexplored. Our findings show that H3K9me2 is maintained in the GC after

https://doi.org/10.1038/s41467-024-51513-4

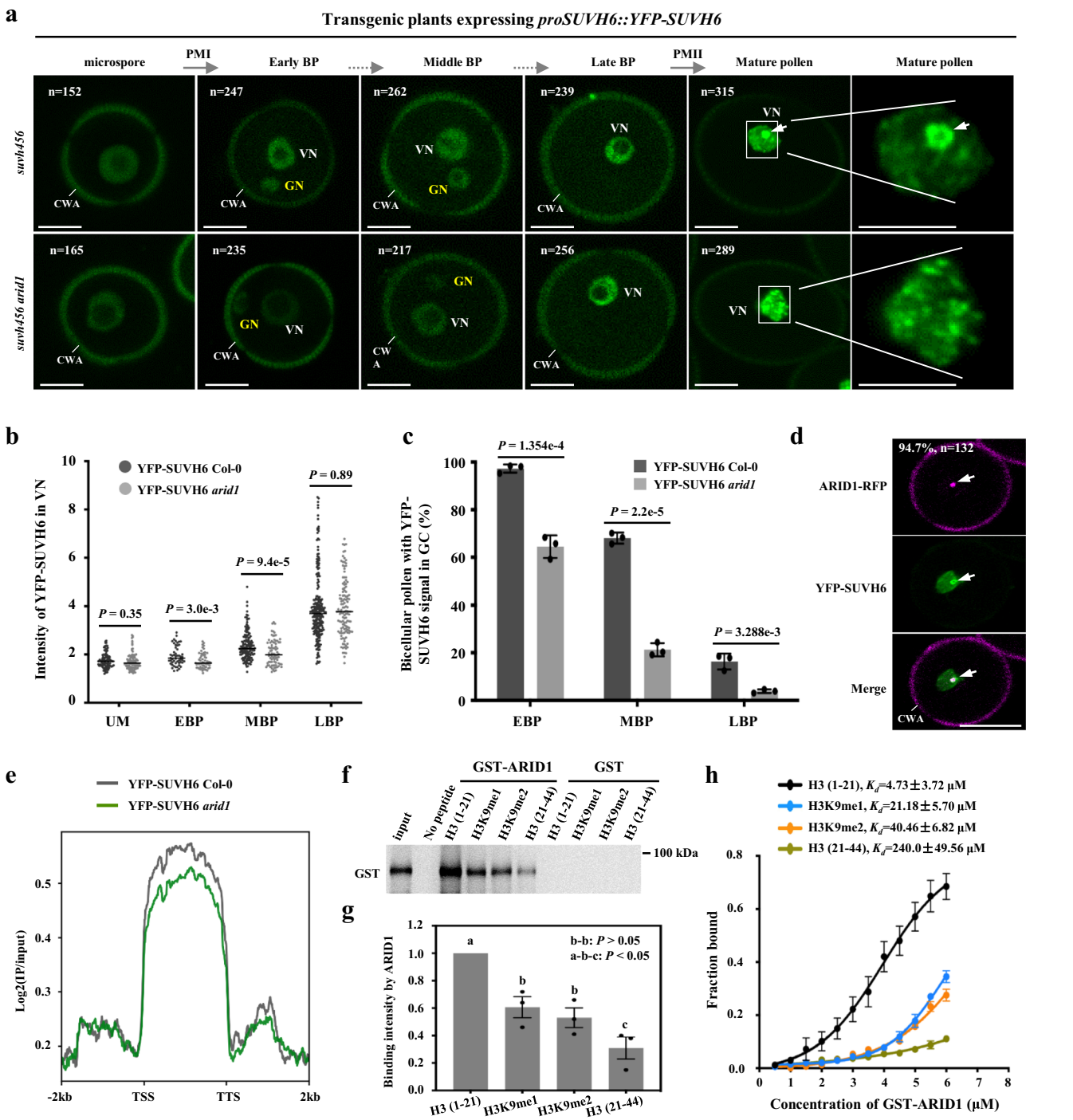

Fig. 6 | ARID1 promote SUVH6 accumulation and recruits SUVH6. a Localization of YFP-SUVH6 in *suvh456* and *suvh456 arid1* during pollen development. The arrow indicates the large foci signal of YFP-SUVH6. VN, vegetative nucleus; GN, generative nucleus; SN, sperm nuclei; PMI, pollen mitosis I; PMII, pollen mitosis II. BP, bicellular pollen. n indicates the total number of analyzed pollen. CWA, cell wall auto-fluorescence. Scale bar, 5 μm. **b** Quantitative analysis of YFP-SUVH6 signal in the VN before PMII shown in (**a**). Intensity of YFP-SUVH6 signal was normalized to the autofluorescence. Each dot represents one pollen. The black bar represents the median. *P* value was determined by unpaired two-sided Student's *t*-test. More than 50 pollen were analyzed for each. **c** Percentage of bicellular pollen with YFP-SUVH6 signal in the GC shown in (**a**). Data are presented as mean ± SE, n = 3. Statistical significance was determined by unpaired two-sided Student's *t*-test. **d** ARID1 co-localizes with SUVH6 in the foci in mature pollen. White arrows indicate the foci signal. Scale bar, 10 μm. **e** Metaplots showing overall occupancy abundance of SUVH6 in Col-0 and *arid1*. The abundance is represented as log2(IP/input). **f** Pull down assays showing the binding of ARID1 with H3 peptides. GST-ARID1-His was incubated with biotin-labeled peptides. The precipitates were separated by SDS-PAGE followed by western blotting analysis using anti-GST antibody. GST-His protein was used as the control. Three biological replicates were performed. The indicator of molecular weight is shown. **g** Quantitative analysis of (**f**). The intensities of ARID1 binding to H3K9me1 (1-21), H3K9me2 (1-21), and H3 (21-44) were relative to that of H3 (1–21). Data are presented as mean ± SE, n = 3. *P* values were calculated using one-sided ANOVA followed by individual two-sample Tukey tests. Boxplots marked as **a**, **b** and **c** are significantly different between groups ($P < 0.05$) but not within the group ($P > 0.05$). **h** Fluorescence polarization (FP) assay showing preferential binding of ARID1 with unmodified H3 (1-21). Data points and Kd values were calculated from three independent experiments, error bars are 0.95 confidence intervals (CIs). Data are presented as mean ± SE, n = 3. Source data are provided as a Source Data file.

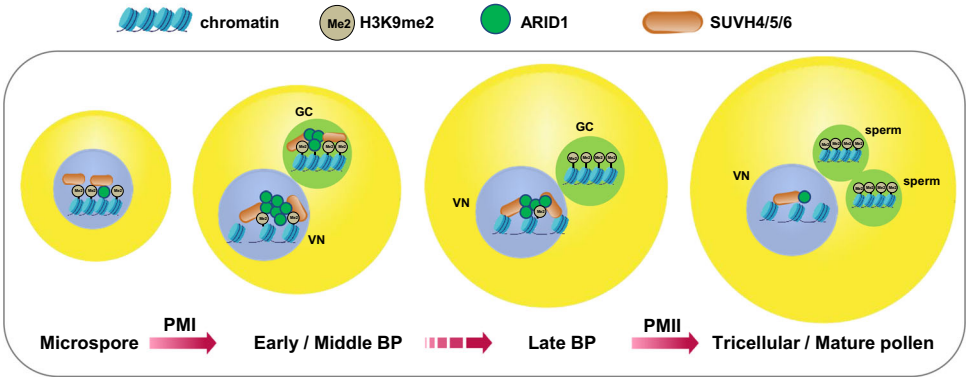

Fig. 7 | A model for ARID1 promoting H3K9me2 maintenance in pollen. After PMI, the chromatin becomes de-condensed in the VC but remain compacted in the GC. H3K9me2 is progressively decreased in the VN while remains unchanged in the GN and SN. During the regulation of this differential pattern of H3K9me2 between VC and GC/SC, ARID1 slows down decrease of H3K9me2 in the VN but promotes maintenance of H3K9me2 in the GN and SN. Mechanistically, ARID1 facilitates SUVH4/5/6 to localize in early GC, and recruits SUVH4/5/6 to chromatin. PMI, pollen mitosis I; PMII, pollen mitosis II; VN, vegetative nucleus; GC, generative cell; SC, sperm cell; BP, bicellular pollen. The icons of "chromatin" and "SUVH4/5/6" created with BioRender.com released under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International license.

PMI but gradually evicted in the VC. This differential pattern of H3K9me2 between two types of cells is regulated by a pollen-specific transcription factor ARID1.

However, it is paradoxical why H3K9me2 is barely detected in the VC but enriched in the SC, despite SUVH5/6 accumulating in the VC rather than the SC. We propose a model to describe how ARID1 is involved in regulating maintenance of H3K9me2 during pollen mitosis (Fig. 7). After meiosis, H3K9me2 is normally detected in microspores and its maintenance is independent of ARID1. However, after PMI, the constitutive heterochromatin becomes de-condensed in the VC but compacted in the GC. H3K9me2 is progressively lost in the VC while clearly maintained in the GC until the SC. ARID1, with highly dynamic localization between the VC and the GC after PMI, promotes maintenance of H3K9me2 in the GC/SC, while slows down eviction of H3K9me2 in the VC. The loss of ARID1 leads to a reduction in accumulation of SUVH4/5/6 proteins, decreased accessibility of SUVH4/5/6 to unmethylated H3K9, and reduced occupancy of SUVH4/5/6 at target genomic regions. Consequently, this results in impairment of H3K9me2 maintenance in pollen. Although the levels of histone modifications are determined by the coordinated actions of histone methyltransferases and histone demethylases, how H3K9me2 is established in plants remains unclear. In pollen, it is obvious that SUVH4/5/6 specificity is not equivalent to that of H3K9me2. Instead, ARID1 as an H3K9 binding protein may play a necessary role in H3K9me2 specificity.

It is intriguing why, despite the accumulation of both ARID1 and SUVH6, there is insufficient induction of H3K9me2 in vegetative nuclei. In parallel, we provide evidence that ectopic expression of ARID1 can result in increased levels of H3K9me2 in sperm cells (Fig. 2d; S2b-c). We hypothesize that the absence of other heterochromatic factors, such as DDM1[13,29], H2A.W[37], and H1[15], in the VN compared to SN may explain this phenomenon. These additional factors might contribute to the regulation of H3K9me2 levels differently between the VN and the SN. In the SN, there appears to be a robust regulatory loop involving H3K9me2 and an environment reinforced by these heterochromatin epigenetic factors. Ectopic expression of ARID1 in the SN enhances the coupling efficiency, thereby promoting H3K9me2. This notion is supported by our observation that ARID1 binding sites are more closely associated with heterochromatic epigenetic marks (Fig. 3e, f, Fig. 4i). Further investigation into the interplay between ARID1, SUVH6, and other heterochromatic factors in different cell types will provide deeper insights into the mechanisms underlying the differential regulation of H3K9me2.

Because the male germ cell identity and the following cell division need to be rapidly completed, the transient accumulation of ARID1 in early GC have two aims. One is to promote expression of *DUO1* for pollen development[24], the other is to recruit SUVH4/5/6 for maintenance of differential pattern of H3K9me2 between companion vegetative cell and male germ cells during pollen mitosis. Compared to that in the GC, ARID1 is more abundant in the VC in early and middle bicellular pollen, which may play a buffering role to antagonize the shock caused by heterochromatin relaxation. During this timing window, ARID1 attempts to slow down the speed of H3K9me2 eviction in the VC, possibly hindering over-reactivation of some transposable elements. This point can be supported by the finding that increased derepression of transposable elements was detected in *arid1* mutant pollen[24]. When the function of ARID1 is impaired, not only DUO1-dependent developmental processes are disrupted, but also maintenance of H3K9me2 is significantly impaired, finally affecting nuclear size.

Methods

Plant materials and reagents

Columbia-0 (Col-0) was used in this study. Seeds of *arid1-1* (SALK_047099), *suvh456* (*suvh4/SALK_044606*; *suvh5/SALK_074957*; *suvh6/SAIL_864_E08*), *proARID1::ARID1-GFP*, and *proARID1::ARID1-Myc* were described. Transgenic plants of *proSUVH4/5/6::YFP-SUVH4/5/6* and *proHTR10::ARID1-YFP* were generated in this study. Plants were grown in soil under a 16-h-light/8-h-dark photoperiod at 22°C. Protease inhibitor cocktail (Roche, Cat.11873580001), DAPI (Roche, Cat.236276), GFP-Trap beads (Chromotek, Cat.gta-20), Protein A agaroseab1220 (Upstate, Cat. #16-157), anti-H3K9me2 (Abcam, Cat.ab1220), anti-GFP (Roche, Cat.11814460001), secondary antibody Alexa Fluor 488 (Jackson, Cat.711-545-152) and secondary antibody Alexa Fluor 555 (Thermo, Cat. A21424) were purchased from indicated companies.

Plasmid construction

For *proSUVH4/5/6::YFP-SUVH4/5/6*, genomic DNA was amplified from Col-0 and cloned into pCambia2302. For *GST-SUVH6*, the 265th −790th fragment of CDS was amplified from Col-0 and cloned into *pGEX-4T-1*. For *proIBM1::IBM1-GFP*, genomic DNA was amplified from Col-0 and cloned into *pMDC107*. For *ARID1-nYFP*, CDS was amplified from Col-0 and cloned into *pXY106*. For *SUVH6-cYFP*, CDS was amplified from Col-0 and cloned into *pXY104*. For *GST-ARID1-His*, ARID1 CDS was amplified from Col-0 and cloned into *pGEX-4T-1*. For *proHTR10::ARID1-YFP*,

ARID1 CDS was amplified from Col-0 and cloned into *pHTR10::GW-YFP*. Primer information is shown in Supplementary Data 4.

Chromatin immunoprecipitation

Approximately 150 μL of mature pollen were resuspended in ice-cold pollen nuclei buffer (1.3 mM H_3BO_3, 3.6 mM $CaCl_2{\cdot}2H_2O$, 0.74 mM KH_2PO_4, 438 mM Sucrose, 7 mM MOPS, 0.83 mM $MgSO_4{\cdot}7H_2O$, pH 6.0), and then transferred to a 2 mL tube containing 2 mm BashingBead. The mixture was vortexed at the 49 Hz for 3 min with a homogenizer. Debris was removed by filtering the suspension through a 10 μm nylon mesh. Nuclei were precipitated by centrifugation and resuspended in M1 buffer (10 mM sodium phosphate pH 7.0, 100 mM NaCl, 1 M hexylene glycol, 10 mM 2-mercaptoethanol, and protease inhibitor) containing 1% formaldehyde and fixed for 20 min. The cross-linking reaction was stopped by the addition of 125 mM glycine. Nuclei were precipitated by centrifugation and washed three times with M2 buffer (10 mM sodium phosphate pH 7.0, 100 mM NaCl, 1 M hexylene glycol, 10 mM $MgCl_2$, 0.5% Triton X-100, 10 mM 2-mercaptoethanol, and protease inhibitor), and then further washed once with M3 buffer (10 mM sodium phosphate pH 7.0, 100 mM NaCl, 10 mM 2-mercaptoethanol, and protease inhibitor). Nuclei pellets were rinsed and resuspended in sonication buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 0.1% SDS, and protease inhibitor), and then sonicated with a sonicator (QSonica, Q800R3) for 30 min (20 s on / 40 s off). After shearing, debris was removed by centrifugation, and the solution containing chromatin fragments was diluted with three times its volume of ChIP dilution buffer (16.7 mM Tris-HCl pH 8.0, 167 mM NaCl, 1.2 mM EDTA, 1.1% Triton X-100, 0.01% SDS, and protease inhibitor). Then, 1% of the solution was aliquoted as the 'input,' and the remaining portion was used to immunoprecipitate the protein-DNA complex.

For H3K9me2 ChIP, after dilution, protein G magnetic beads were added to the sheared chromatin and incubated at 4°C for 1 hour. Pre-cleared samples were collected and incubated with anti-H3K9me2 antibodies at 4°C overnight with rotation. After incubation, samples were mixed with 30 μL of protein G magnetic beads, incubated at 4°C for 3 hours with rotation. For SUVH6 ChIP, after dilution, the pre-cleared samples were collected and incubated with GFP-trap beads at 4°C overnight with rotation. The beads were then washed twice with low salt buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM EDTA, 1% Triton X-100, and 0.1% SDS), twice with high salt buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 2 mM EDTA, 1% Triton X-100, and 0.1% SDS), once with LiCl buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 0.25 M LiCl, 1% NP-40, and 0.1% deoxycholic acid), and twice with TE buffer (10 mM Tris-HCl pH 8.0 and 1 mM EDTA). Immunoprecipitated DNA was incubated at 65 °C using 150 μL of elution buffer (1% SDS, 10 mM EDTA, 0.1 M NaHCO3) twice with shaking rigorously. After elution, the DNA-protein complex and input samples were reverse-crosslinked using 12 μL 5 M NaCl at 65 °C for 6 hours. 1 μL of Proteinase K, 10 μL of 0.5 M EDTA, 20 μL of 1 M Tris pH 6.5, 10 μg of RNase A were added to digest proteins and RNA. DNA was then purified using phenol:chloroform and precipitated with sodium acetate, glycogen and ethanol overnight at −20 °C. The precipitated DNA was used for the library construction. ChIP-Seq libraries were prepared using the VAHTS Universal DNA Library Prep Kit for Illumina V4 (Vazyme, ND610) following the manufacturer's instructions.

For ARID1-GFP CUT & Tag libraries, pollen nuclei were used for library construction using the Hyperactive Universal CUT & Tag Assay Kit for Illumina (Vazyme, TD903) following the manufacturer's instructions. The nuclear extracts from the transgenic plants expressing *proARID1::ARID1-GFP* were sonicated, and genomic DNA was used as the input control for library construction. The libraries were sequenced with an Illumina NovaSeq 6000 or Illumina Hiseq 3000 to generate pair-end 150 bp reads. Samples were prepared from three or two biological replicates.

Sequencing data analysis

Following a FastQC quality control, adapter trimming was performed using TrimGalore. Reads were mapped to the Arabidopsis genome (TAIR10) using Bowtie2[38], and subsequently filtered for duplicate reads using Sambamba. A cross-correlation matrix based on Pearson's correlation coefficient was generated by comparing the processed BAM files using deepTools utility bamCorrelate[39]. Biological replicates were merged for subsequent analysis. The peaks of ARID1 were identified by MACS2[40] with the parameters of $p < 0.00001$. The peaks of H3K9me2 and SUVH6 were identified by MACS2[40] with the parameters of $q < 0.05$. Peaks located within 2 kb upstream of the transcription start site (TSS) or 2 kb downstream of the transcription termination site (TTS) were considered gene-associated or TE-associated peaks, respectively. The distribution of peaks across genic or TE regions was made by deeptools program. The binding intensity log2(IP/input) was z-transformed across genotypes, and the resulting z scores were clustered using k-means. Visualizations of the clustering results were generated using deeptools (v2.5.0.1) or ggplot2 (v.3.1.0). Bigwig coverage files were visualized along the TAIR10 genome using IGV version 2.4.10. ARID1-enriched motif was identified by the MEME-ChIP[41] with the the default parameters. FIMO[42] was used to count individual motif occurrences.

Immunofluorescence assay

Pollen grains at different stages were collected into a 1.5 mL tube containing 4% paraformaldehyde and fixed for 30 minutes at room temperature. After rinsing for 3 times with PBS buffer, pollen grains were digested in the enzymatic solution (10 mM citrate (pH 4.5), 1% (w/v) cellulase, 0.5% (w/v) macerozyme at 37 °C for 60 min, rinsed with PBS for 3 times, and permeabilized for 30 min with 0.1% Triton X-100, and rinsed with PBS for 3 times. After 2 h blocking with 1% BSA at 37 °C, pollen grains were incubated overnight at 4 °C with anti-H3K9me2 at a 1:200 dilution in 1% BSA and then rinsed with PBS for 3 times. After 2 h of incubation with the secondary anti-Alexa Fluor 488 or Fluor 555 in 1% BSA at 37 °C, pollen was rinsed with PBS for 3 times and stained with DAPI. Pollen was observed with an Olympus FV3000 laser confocal microscope and the images were processed using Adobe Photoshop and Image J. The signal of H3K9me2 was compared to that of DAPI or the background.

Co-immunoprecipitation (Co-IP)

~100 μL mature pollen were collected and mixed with ~100 μL glass beads (425–600 μm), 4-5 zircons (2 mm, about 100 μL) and 600 μL lysis buffer (50 mM Tris_HCl (pH 7.5), 150 mM NaCl, 5 mM $MgCl_2$, 0.2% Nonidet P-40, 2 mM DTT, 10% glycerol, protease inhibitor). The mixture was vortexed at the 60 HZ for 3 min with homogenizer, and centrifuged for 15 min at 16,000 g twice at 4°C. The supernatant was pre-cleared with protein A agarose, and the lysate was incubated with GFP-Trap agarose beads or Flag® M2 beads at 4°C for 3 h with gentle rotation. The immune complexes were washed for five times with 1 mL of lysis buffer. Proteins retained on the beads were resolved on SDS-PAGE followed by western blots. The image of western blots was measured with Tanon-5200 Gel Image System software.

Protein expression

A codon plus (DE3) *E. coli* strain was used for expression of recombinant proteins. The bacteria were cultured in LB, and 1 mM IPTG was used to induce protein expression at 16°C for 16–18 h. For purification of the GST-SUVH6 and GST proteins, the cell pellets were resuspended in 1 × PBS containing 1 mM PMSF before sonication and centrifugation. The supernatant was passed through a Glutathione Sepharose 4B column, and bound proteins were washed with 1 × PBS containing 2 mM L-Glutathione, then the proteins were used for the GST-pull down assay. For purification of GST-ARID1-His and GST-His, the cell pellets

were resuspended in lysis buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 25 mM imidazole, pH 8.0) containing 1 mM PMSF. After sonication and centrifugation, the supernatant was passed through a Ni-NTA agarose column followed by elution with the buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 500 mM imidazole, pH 8.0). The eluted proteins were used for in vitro histone binding assay and fluorescence polarization (FP) assay.

In vitro histone binding assays

The biotinylated histone peptides H3 (1-21), H3K9me1 (1-21), H3K9me2 (1-21), and H3 (21-44) (SciLight Biotechnology, Beijing) were used for in vitro histone binding assays. 1 μg of peptides was incubated with 30 μL Streptavidin beads (Thermo Scientific) at 4 °C for 2 h with rotation in 300 μL binding/wash buffer (20 mM Tris-Cl pH 8.0, 150 mM NaCl, 0.1% Triton X-100). After washing for three times, 2 μg of GST-ARID1-His were added and incubated at 4 °C for 3 h with rotation in 300 μL binding/wash buffer. The beads were washed three times and resuspended in 30 μL 2 × SDS sample buffer, the bound proteins were subjected to SDS-PAGE. GST-His protein was used as the negative control.

Fluorescence polarization (FP) assay

GST-ARID1-His proteins were serially diluted into the buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl) from 0.5 to 6 μM in 12 reaction tubes with a final volume of 30 μL. An equal volume of 5′FAM-labeled peptides (Sangon Biotech, Shanghai) was added to reach a final concentration of 10 nM. The reaction mixture was incubated at room temperature for 1 h in the dark. The fluorescence anisotropy was measured using Synergy2 (BioTek, USA). The K_d values and the fitting errors were calculated with GraphPad Prism 9.5 software by using the sigmoidal dose-response model.

Semi in vivo pull down assay

GST or GST-SUVH6 were purified by glutathione Sepharose 4B affinity chromatography. 1 mL of total lysates from mature pollen of the transgenic plants expressing *proARID1::ARID1-GFP* were applied to the beads-recombinant protein mixture (~ 2 μg protein) and incubated at 4°C for 3 h with gentle rotation. The retained mixture was washed five times with washing buffer (50 mM Tris_HCl (pH 7.5), 150 mM NaCl, 0.2% Triton X-100). The pellets were analyzed by western blots.

BiFC analysis

The plasmids were transformed into *Agrobacterium* strain GV3101 and co-infiltrated in young leaves of *N. benthamiana* and then grown in the dark for 1 d. After two days of long-day conditions (16 h light/8 h dark) growth, the infiltrated part of leaves was analyzed using confocal microscopy

Quantification and statistical analysis

Statistical procedures are shown in the figure captions. Image J software was used for quantification of H3K9me2 levels in pollen nuclei. Statistical methods for assessing peak overlap included one-sided permutation overlap tests or two-sided Fisher's exact tests, respectively. All boxplots, bar plots, Venn overlaps and line charts were generated by R or GraphPad Prism 9.5. Statistical significance was determined by one-way ANOVA or Student's *t*-test.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Data supporting the findings of this work are available within the paper and its Supplementary Information files. A reporting summary for this Article is available as a Supplementary Information file. The datasets and plant materials generated and analyzed during the current study are available from the corresponding author upon request. High throughput sequencing data has been deposited in the Gene Expression Omnibus (GEO) database and can be accessed with the accession number GSE268716. Source data are provided with this paper.

References

1. Hajkova, P. Epigenetic reprogramming in the germline: towards the ground state of the epigenome. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **366**, 2266–2273 (2011).
2. Reik, W., Dean, W. & Walter, J. Epigenetic reprogramming in mammalian development. *Science* **293**, 1089–1093 (2001).
3. Sasaki, H. & Matsui, Y. Epigenetic events in mammalian germ-cell development: reprogramming and beyond. *Nat. Rev. Genet* **9**, 129–140 (2008).
4. McCormick, S. Control of male gametophyte development. *Plant Cell* **16**, S142–S153 (2004).
5. Borg, M. & Berger, F. Chromatin remodelling during male gametophyte development. *Plant J.* **83**, 177–188 (2015).
6. Borg, M. et al. Targeted reprogramming of H3K27me3 resets epigenetic memory in plant paternal chromatin. *Nat. Cell Biol.* **22**, 621–629 (2020).
7. Huang, X. & Sun, M. X. H3K27 methylation regulates the fate of two cell lineages in male gametophytes. *Plant Cell* **34**, 2989–3005 (2022).
8. Zhu, D. et al. Distinct chromatin signatures in the Arabidopsis male gametophyte. *Nat. Genet* **55**, 706–720 (2023).
9. Borg, M. et al. Epigenetic reprogramming rewires transcription during the alternation of generations in Arabidopsis. *Elife* **10**, e61894 (2021).
10. Lippman, Z. & Martienssen, R. The role of RNA interference in heterochromatic silencing. *Nature* **431**, 364–370 (2004).
11. Matzke, M. A. & Birchler, J. A. RNAi-mediated pathways in the nucleus. *Nat. Rev. Genet* **6**, 24–35 (2005).
12. Feng, W. & Michaels, S. D. Accessing the inaccessible: the organization, transcription, replication, and repair of heterochromatin in plants. *Annu Rev. Genet* **49**, 439–459 (2015).
13. Slotkin, R. K. et al. Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. *Cell* **136**, 461–472 (2009).
14. Ibarra, C. A. et al. Active DNA demethylation in plant companion cells reinforces transposon methylation in gametes. *Science* **337**, 1360–1364 (2012).
15. He, S., Vickers, M., Zhang, J. & Feng, X. Natural depletion of histone H1 in sex cells causes DNA demethylation, heterochromatin decondensation and transposon activation. *Elife* **8**, e42530 (2019).
16. Martinez, G., Panda, K., Kohler, C. & Slotkin, R. K. Silencing in sperm cells is directed by RNA movement from the surrounding nurse cell. *Nat. Plants* **2**, 16030 (2016).
17. Borges, F. et al. Transposon-derived small RNAs triggered by miR845 mediate genome dosage response in Arabidopsis. *Nat. Genet* **50**, 186–192 (2018).
18. Martinez, G. et al. Paternal easiRNAs regulate parental genome dosage in Arabidopsis. *Nat. Genet* **50**, 193–198 (2018).
19. Rotman, N. et al. A novel class of MYB factors controls sperm-cell formation in plants. *Curr. Biol.* **15**, 244–248 (2005).
20. Brownfield, L. et al. A plant germline-specific integrator of sperm specification and cell cycle progression. *PLoS Genet* **5**, e1000430 (2009).
21. Borg, M. et al. The R2R3 MYB transcription factor DUO1 activates a male germline-specific regulon essential for sperm cell differentiation in Arabidopsis. *Plant Cell* **23**, 534–549 (2011).
22. Higo, A. et al. Transcription factor DUO1 generated by neofunctionalization is associated with evolution of sperm differentiation in plants. *Nat. Commun.* **9**, 5283 (2018).

23. Rauf, A., Khatab, H., Borg, M. & Twell, D. Genetic control of generative cell shape by DUO1 in Arabidopsis. *Plant Reprod.* **36**, 243–254 (2023).
24. Zheng, B., He, H., Zheng, Y., Wu, W. & McCormick, S. An ARID domain-containing protein within nuclear bodies is required for sperm cell formation in *Arabidopsis thaliana*. *PLoS Genet* **10**, e1004421 (2014).
25. Li, L., Wu, W., Zhao, Y. & Zheng, B. A reciprocal inhibition between ARID1 and MET1 in male and female gametes in Arabidopsis. *J. Integr. Plant Biol.* **59**, 657–668 (2017).
26. Wu, W. & Zheng, B. Intercellular delivery of small RNAs in plant gametes. *N. Phytol.* **224**, 86–90 (2019).
27. Wu, W. et al. Heterochromatic silencing is reinforced by ARID1-mediated small RNA movement in Arabidopsis pollen. *N. Phytol.* **229**, 3269–3280 (2021).
28. Schoft, V. K. et al. Induction of RNA-directed DNA methylation upon decondensation of constitutive heterochromatin. *EMBO Rep.* **10**, 1015–1021 (2009).
29. Osakabe, A. et al. The chromatin remodeler DDM1 prevents transposon mobility through deposition of histone variant H2A.W. *Nat. Cell Biol.* **23**, 391–400 (2021).
30. Stroud, H. et al. Non-CG methylation patterns shape the epigenetic landscape in Arabidopsis. *Nat. Struct. Mol. Biol.* **21**, 64–72 (2014).
31. Khouider, S. et al. Male fertility in Arabidopsis requires active DNA demethylation of genes that control pollen tube function. *Nat. Commun.* **12**, 410 (2021).
32. Zhang, H. et al. A comprehensive online database for exploring approximately 20,000 public Arabidopsis RNA-Seq libraries. *Mol. Plant* **13**, 1231–1233 (2020).
33. Misra, C. S., Sousa, A. G. G., Barros, P. M., Kermanov, A. & Becker, J. D. Cell-type-specific alternative splicing in the Arabidopsis germline. *Plant Physiol.* **192**, 85–101 (2023).
34. Saze, H., Shiraishi, A., Miura, A. & Kakutani, T. Control of genic DNA methylation by a jmjC domain-containing protein in *Arabidopsis thaliana*. *Science* **319**, 462–465 (2008).
35. Cheng, J. et al. H3K9 demethylases IBM1 and JMJ27 are required for male meiosis in *Arabidopsis thaliana*. *N. Phytol.* **235**, 2252–2269 (2022).
36. Wang, L., Dou, K., Moon, S., Tan, F. J. & Zhang, Z. Z. Hijacking oogenesis enables massive propagation of LINE and retroviral transposons. *Cell* **174**, 1082–1094.e1012 (2018).
37. Borges, F. et al. Loss of small-RNA-directed DNA methylation in the plant cell cycle promotes germline reprogramming and somaclonal variation. *Curr. Biol.* **31**, 591–600.e594 (2021).
38. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
39. Ramirez, F., Dundar, F., Diehl, S., Gruning, B. A. & Manke, T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* **42**, W187–W191 (2014).
40. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
41. Machanick, P. & Bailey, T. L. MEME-ChIP: motif analysis of large DNA datasets. *Bioinformatics* **27**, 1696–1697 (2011).
42. Grant, C. E., Bailey, T. L. & Noble, W. S. FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**, 1017–1018 (2011).

Acknowledgements

This work was supported by the National Natural Science Foundation of China (32025005, 31830045, M-0398 to B.Z.; 32100277 to L.L.).

Author contributions

L.L. and B.Z. conceptualized the study. L.L. performed most experiments. H.Y., Yi.Z., Q.H., X.Z., T.J., and H.J. helped collect pollen, construct plasmids, analyzed data, and revise the manuscript. L.L. and B.Z. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-51513-4.

Correspondence and requests for materials should be addressed to Binglian Zheng.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.