

ARTICLE

Open Access

Four putative SWI2/SNF2 chromatin remodelers have dual roles in regulating DNA methylation in Arabidopsis

Dong-Lei Yang^{1,2}, Guiping Zhang¹, Lili Wang², Jingwen Li¹, Dachao Xu², Cuiru Di², Kai Tang^{1,3}, Lan Yang¹, Liang Zeng¹, Daisuke Miki¹, Cheng-Guo Duan¹, Huiming Zhang¹ and Jian-Kang Zhu^{1,3}

Abstract

DNA methylation is a conserved epigenetic mark that is critical for many biological processes in plants and mammals. In Arabidopsis, the antagonistic activities of RNA-directed DNA methylation (RdDM) and ROS1-dependent active DNA demethylation are key for the dynamic regulation of locus-specific DNA methylation. However, the molecular factors that coordinate RdDM and active demethylation are largely unknown. Here we report that CLSY4 and its three paralogous SWI2/SNF2-type chromatin-remodeling proteins function in both RdDM and DNA demethylation in Arabidopsis. We initially identified CLSY4 in a genetic screen for DNA demethylation factors and subsequently demonstrated that it also is important in RdDM. Comprehensive genetic analyses using single and high order mutants of CLSY family proteins revealed their roles as double agents in the balance between methylation and demethylation reactions. The four CLSY proteins collectively are necessary for the canonical RdDM pathway; at the same time, each CLSY likely mediates DNA demethylation at specific loci where DNA methylation depends on RdDM. These results indicate that the four chromatin-remodeling proteins have dual functions in regulating genomic DNA methylation, and thus provide new insights into the dynamic regulation of DNA methylation in a model multicellular eukaryotic organism.

Introduction

DNA methylation is an important epigenetic mark that confers transcriptional regulation of genes and transposable elements^{1–4}. Aberrant DNA methylation can disrupt normal developmental processes and cause disease symptoms in plants and mammals^{5,6}. A specific DNA methylation pattern is an outcome of dynamic regulation by methylation establishment and maintenance, together with methylation removal activities.

In plants, de novo DNA methylation can be established by the RNA-directed DNA methylation (RdDM) pathway, in which complementary pairing between 24-nt siRNAs and nascent scaffold RNAs, together with protein–protein interactions, recruits the DNA methyltransferase DRM2 for DNA methylation^{4,7–10}. Production of nearly all 24-nt siRNAs depends on the plant-specific RNA polymerase Pol IV, which generates single-stranded non-coding RNAs that serve as templates for downstream processing by RDR2 (RNA-dependent RNA polymerase 2) and DCL3 (Dicer-like 3) to produce 24-nt siRNAs^{11–17}. Pol IV is recruited to RdDM targets in part by a SAWADEE domain-containing protein, SHH1/DTF1, which binds via its Tudor domain to histone H3 with methylated H3K9^{18,19}. Both Pol IV and DTF1/SHH1 associate with CLSY1^{18,19}, which is a chromatin-remodeling protein involved in Pol IV-dependent siRNA production²⁰.

Correspondence: Dong-Lei Yang (dlyang@njau.edu.cn) or Jian-Kang Zhu (jkzhu@sibs.ac.cn)

¹Shanghai Center for Plant Stress Biology and Center of Excellence in Molecular Plant Sciences, Chinese Academy of Science, Shanghai 200032, China

²State Key Laboratory of Crop Genetics and Germplasm Enhancement, Nanjing Agricultural University, Nanjing 210095, China

Full list of author information is available at the end of the article.

© The Author(s). 2018



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

However, it appears that CLSY1 only affects a part of Pol IV-dependent RdDM²¹, and it is unclear whether other chromatin-remodeling proteins are involved in Pol IV-dependent siRNA production and DNA methylation.

Active DNA demethylation counteracts the establishment and maintenance of DNA methylation to prevent DNA hypermethylation at specific loci²². In plants, a family of bifunctional 5-mC DNA glycosylases initiates active DNA demethylation through a base excision repair pathway^{23–26}. ROS1 is a major DNA demethylase in Arabidopsis and is particularly important for counteracting DNA methylation established by the RdDM pathway, although it also demethylates RdDM-independent methylation at some genomic regions^{22,23,27,28}. The *ROS1* gene promoter contains a short DNA sequence termed MEMS (DNA METHylation Monitoring Sequence). DNA methylation is positively correlated with *ROS1* gene expression and is co-regulated by the RdDM pathway, by the CG methylation maintenance methyltransferase MET1, and by ROS1 itself;^{29–31} thus MEMS serves as an indicator of general methylation and demethylation activities and enables the coordination of DNA methylation and active DNA demethylation through transcriptional regulation of *ROS1*. Although it is evident that DNA methylation is dynamically regulated by methylation and demethylation processes, the underlying mechanisms of this fine dynamic regulation, besides the transcriptional regulation of *ROS1*, remain to be discovered.

In this study, our genetic screen for Arabidopsis DNA demethylation factors led to the isolation of a *clsy4* mutant. Investigations on single, triple, and quadruple mutants of CLSY1, CLSY2, CLSY3, and CLSY4 revealed that these four chromatin-remodeling proteins not only redundantly function in Pol IV-dependent RdDM, but also individually mediate active DNA demethylation at specific genomic regions including RdDM targets. Our results show that the four CLSY proteins collectively are necessary for Pol IV-dependent siRNA production and DNA methylation in the canonical RdDM pathway. Importantly, each CLSY also mediates DNA demethylation at loci where DNA methylation depends on one or more of the other three CLSY proteins, thereby providing a novel mechanism underlying the dynamic regulation of DNA methylation in Arabidopsis.

Results

Identification of CLSY4/CHR40 as a DNA demethylation factor

In searching for molecular factors involved in active DNA demethylation, we conducted a genetic screen in a T-DNA mutant collection (Supplementary Table S1) using Chop-PCR to assess the DNA methylation status of two genomic loci, Pm36 and Pm27 (Supplementary Fig. S1a). Pm36 and Pm27 are located in the 3' regions of

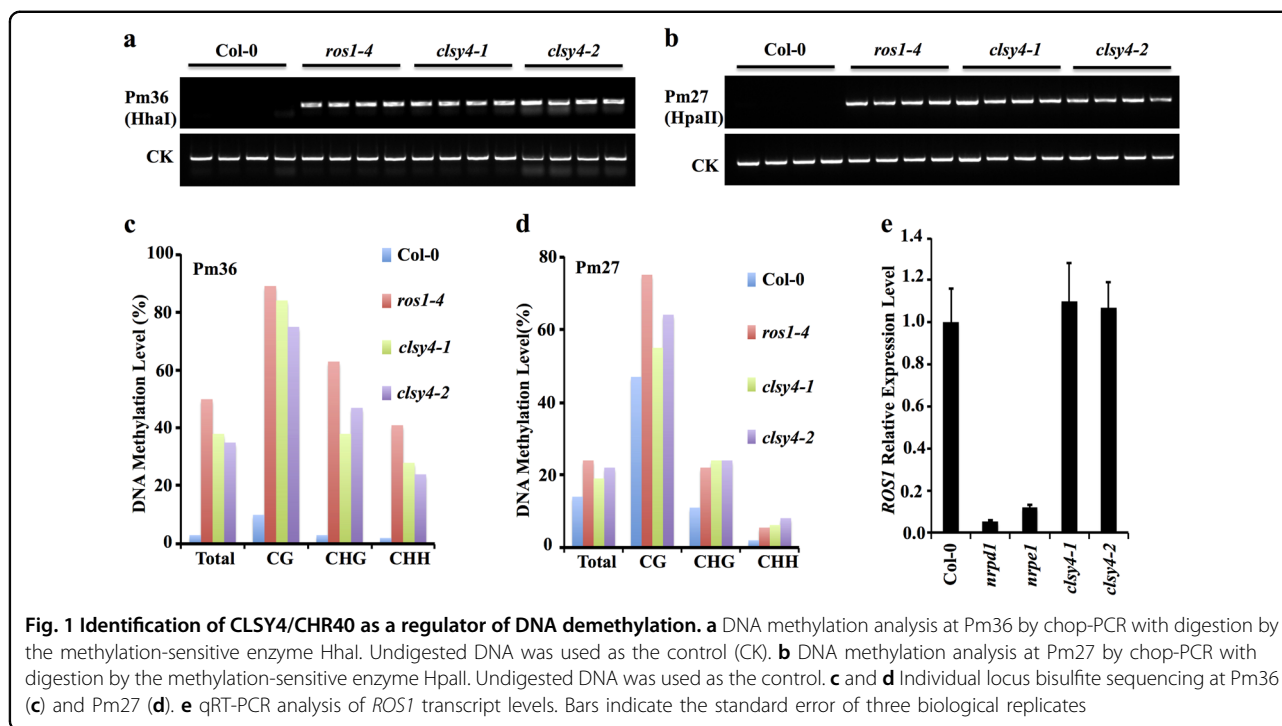
At1g26400 and At1g26390, respectively, and both show DNA hypermethylation in the *ros1* mutant and were used successfully as markers to isolate *idm1* (*increased DNA methylation 1*) and *idm2* (*increased DNA methylation 2*) mutants that are defective in the regulation of active DNA demethylation (Supplementary Fig. S1a)^{32,33}. In a T-DNA insertion mutant (SALK_102252) of At3g24340, both Pm36 and Pm27 could be amplified after methylation-sensitive enzyme digestion, suggesting DNA hypermethylation in this mutant (Fig. 1a, b; Supplementary Fig. S1b, c). Moreover, another T-DNA insertion mutant allele for this gene (SALK_003876) also exhibited DNA hypermethylation (Fig. 1a, b; Supplementary Fig. S1b, c). At3g24340 encodes Chromatin Remodeling 40 (CHR40), a member of the ATP-dependent chromatin remodeler protein family that is also known as CLSY4³⁴. To confirm the DNA hypermethylation phenotype, we performed bisulfite sequencing at the two loci and found that DNA methylation levels in all cytosine contexts (CG, CHG, and CHH, where H represents A, T, and C) were increased in both *clsy4/chr40* alleles as is also the case in the *ros1* mutant (Fig. 1c, d).

Previous reports showed that the reduced *ROS1* expression in various RdDM mutants causes increased DNA methylation at certain loci in the genome^{22,30,35}. Given that CHR40 was previously found to be associated with the Pol IV protein complex³⁴, we suspected that DNA hypermethylation in *clsy4/chr40* mutants might be caused indirectly by a decrease in *ROS1* expression. However, real-time PCR assays showed that the transcript level of *ROS1* was not changed in either of the *clsy4/chr40* mutant alleles compared to wild-type plants (Fig. 1e), suggesting that CLSY4/CHR40 is more likely to be directly involved in the ROS1-mediated active DNA demethylation pathway.

In addition, transgenic expression of *CLSY4* driven by its native promoter rescued the CHH and CHG hypermethylation phenotype in the *clsy4-1/chr40-1* mutant (Supplementary Fig. S1d, e), supporting the inference that the mutations in *CLSY4/CHR40* are responsible for DNA hypermethylation at the tested loci.

CLSY4 affects DNA methylation at many genomic regions

CLSY4/CHR40 belongs to the SWI2/SNF2 family of chromatin-remodeling proteins, and is in the clade of six plant-specific members (Supplementary Fig. S2a) that include CLSY1 and DRD1; DRD1 is a component of the DDR complex required for Pol V function in RdDM³⁶. A previous study reported that CLSY4 could be co-purified with the largest subunit of Pol IV, suggesting its possible role in RdDM³⁴. However, our results above suggested that CLSY4 functions in DNA demethylation. To investigate this inconsistency, we performed high coverage (~60x) whole-genome bisulfite sequencing of the *clsy4-1/*

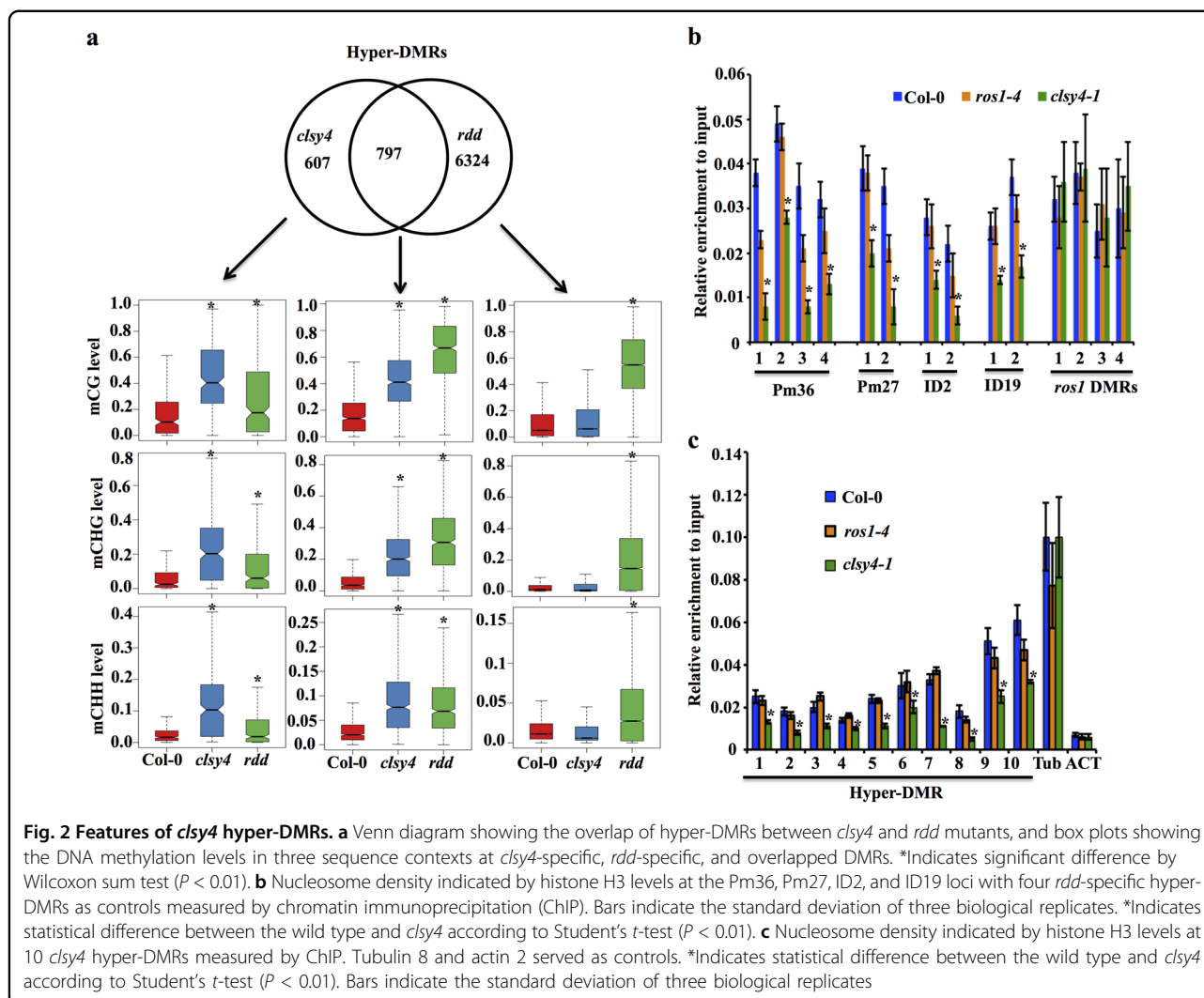


chr40-1 mutant in order to examine the DNA methylome of *clsy4* mutant plants (Supplementary Table S2). We identified 1404 hypermethylated differentially methylated regions (hyper-DMRs) and 314 hypomethylated DMRs (hypo-DMRs) in *clsy4-1* mutant plants compared to wild-type plants (Supplementary Fig. S2b). For the hyper-DMRs, DNA methylation in three sequence contexts, including CG, CHG, and CHH, was increased in the *clsy4-1/chr40-1* mutant (Fig. 2a). This pattern of increased methylation in all three sequence contexts was also observed in *ros1* single and *ros1dml2dml3* (*rdd*) triple mutants³². According to our analysis, about 60% of the hyper-DMRs in *clsy4-1/chr40-1* also showed hypermethylation in the *rdd* mutant (Fig. 2a). Even for the *clsy4*-specific hyper-DMRs, DNA methylation in CG, CHG, and CHH contexts was also significantly increased in *rdd* mutant plants, although the increase was less than that in *clsy4* (Fig. 2a), possibly due to the remaining 5-mC DNA glycosylase DME in *rdd*.

Increased DNA methylation levels at two hyper-DMRs, ID2 (increased DNA methylation 2) and ID19, were confirmed by individual locus bisulfite sequencing (Supplementary Fig. S2c, d). The siRNA abundance at Pm36, ID2, and ID19 was not increased in *clsy4* mutants, as shown by northern blotting (Supplementary Fig. S2e), indicating that the DNA hypermethylation was not due to increased accumulation of siRNAs. Given that *CLSY4* encodes a putative ATP-dependent chromatin-remodeling protein that might regulate nucleosome positioning, we assessed the nucleosome occupancy at the hyper-

DMRs³⁷ using anti-histone H3 for ChIP-qPCR assays after MNase digestion. At four genomic regions (Supplementary Fig. S3a) that were confirmed to be hypermethylated by individual locus bisulfite sequencing in both *clsy4* and *ros1* (Fig. 1c, d; Supplementary Fig. S2c, d), the ChIP signals were significantly lower in *clsy4* mutant plants than in wild-type plants, indicating that the nucleosome occupancy at those regions was reduced in *clsy4* (Fig. 2b). In contrast, the signals in the wild type and the *ros1* mutant were comparable (Fig. 2b), indicating that *ROS1* does not affect nucleosome positioning. In addition, at four *rdd* mutant-specific hyper-DMRs (Supplementary Fig. S3b), the nucleosome occupancy was not decreased in the *clsy4* mutant (Fig. 2b), suggesting a direct role of *CLSY4* in the regulation of active demethylation. At 10 other randomly selected hyper-DMRs of *clsy4* (Supplementary Fig. S4), the *clsy4* mutant also showed lower nucleosome occupancy than the wild type and *ros1* (Fig. 2c). These results suggest that *CLSY4* likely facilitates active DNA demethylation through nucleosome remodeling.

To determine whether the increased DNA methylation in *clsy4* affects gene expression, we selected 12 genes near the hyper-DMRs and measured their expression levels (Supplementary Figs. S5, S6). Eight of the 12 selected genes showed reduced transcript levels in the *clsy4* mutant (Supplementary Figs. S5, S6). One gene, At3g44910, showed an increased expression in *clsy4* and *ros1* (Supplementary Fig. S6a), providing a case in which DNA methylation is positively correlated with the



transcript level. Taken together, the results show that CLSY4-dependent DNA demethylation contributes to gene regulation.

Among the 314 hypo-DMRs identified in the *clsy4-1/chr40-1* mutant, only 16 overlapped with *nrdp1* hypo-DMRs, suggesting that a large subset of the hypo-DMRs in *clsy4* mutant are not RdDM target loci (Fig. 3a). Four common hypo-DMRs were selected for validation by individual locus bisulfite sequencing, which showed that DNA methylation was indeed reduced in both the *clsy4* and *nrdp1* mutants relative to the Col-0 wild type and *ros1* mutant (Fig. 3b–e). siRNAs were abolished at the tested loci in the *nrdp1* mutant and were not reduced in the *clsy4* mutant (Fig. 3f). siRNA accumulation at several known RdDM target loci such as Solo-LTR, AtSN1, and GP1, was unaffected in *clsy4* (Supplementary Fig. S7a); DNA methylation at these loci was also unaffected in *clsy4* mutant plants (Supplementary Fig. S7b–d). Altogether, our genome-wide analyses of DMRs in *clsy4* revealed that

CLSY4 is likely involved in DNA demethylation and, to a lesser degree, in Pol IV-dependent RdDM.

The relationship between RdDM and DNA demethylation mediated by CLSY4

To investigate the relationship between CLSY4-mediated DNA demethylation and RdDM, we carried out genetic analysis in the *clsy4-1nrpd1-3* and *clsy4-1nrpe1-11* double mutants at three *clsy4* hyper-DMRs (Pm36, ID2, and ID19). These tested loci are also ROS1 target loci (Fig. 4a–c). At these loci, the DNA methylation in CHG and CHH contexts was substantially lower in the double mutants *clsy4nrpd1*, *clsy4nrpe1*, *ros1nrpd1*, and *ros1nrpe1* than that in the *clsy4* and *ros1* mutants (Fig. 4a–c), indicating that these loci are RdDM targets. The variation in CG methylation was substantial and included full maintenance at PM36 in *clsy4nrpd1*, *clsy4nrpe1*, *ros1nrpd1*, and *ros1nrpe1*; partial maintenance at ID2 in *ros1nrpd1* and *ros1nrpe1*; and complete

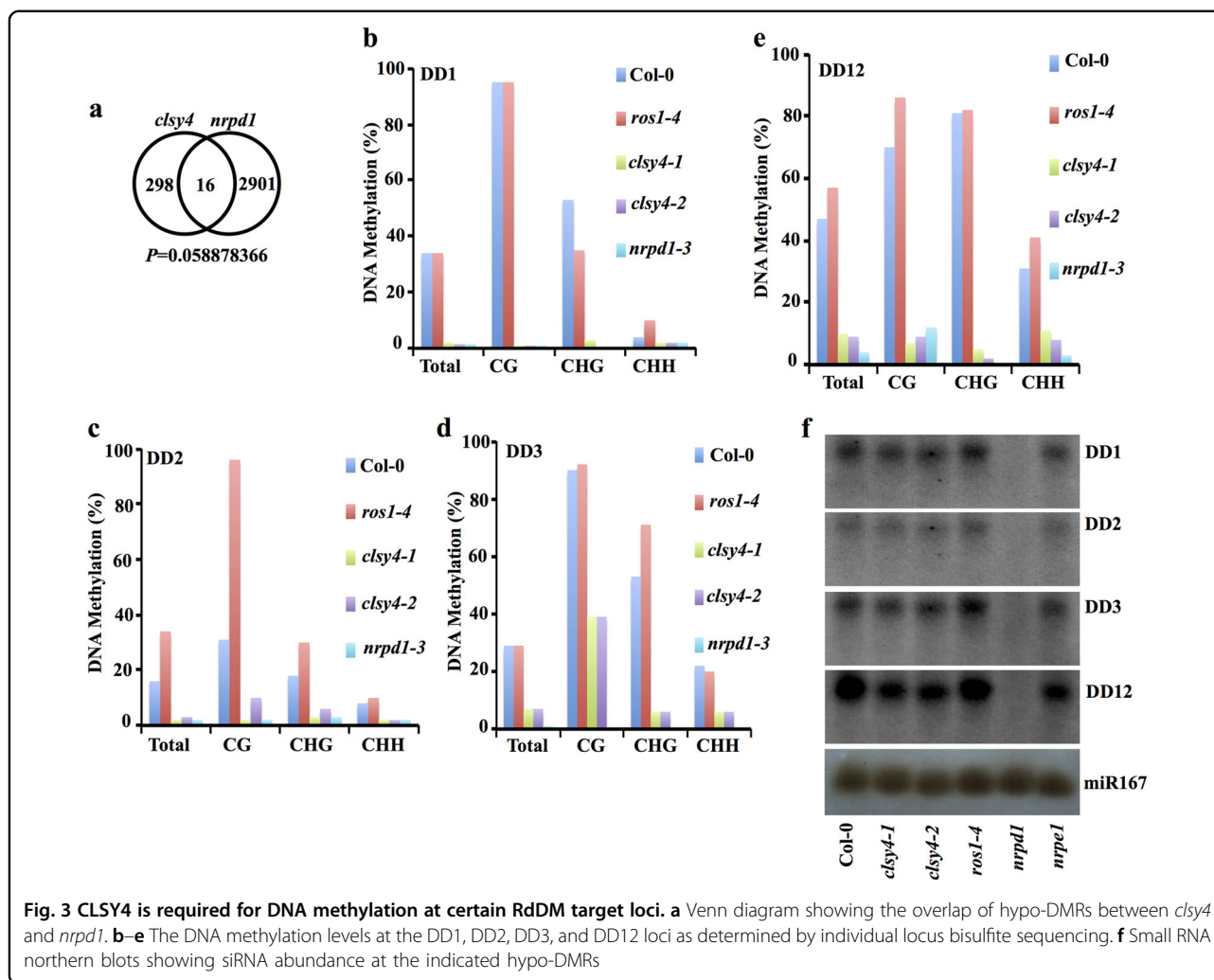


Fig. 3 CLSY4 is required for DNA methylation at certain RdDM target loci. **a** Venn diagram showing the overlap of hypo-DMRs between *cly4* and *nrpd1*. **b–e** The DNA methylation levels at the DD1, DD2, DD3, and DD12 loci as determined by individual locus bisulfite sequencing. **f** Small RNA northern blots showing siRNA abundance at the indicated hypo-DMRs

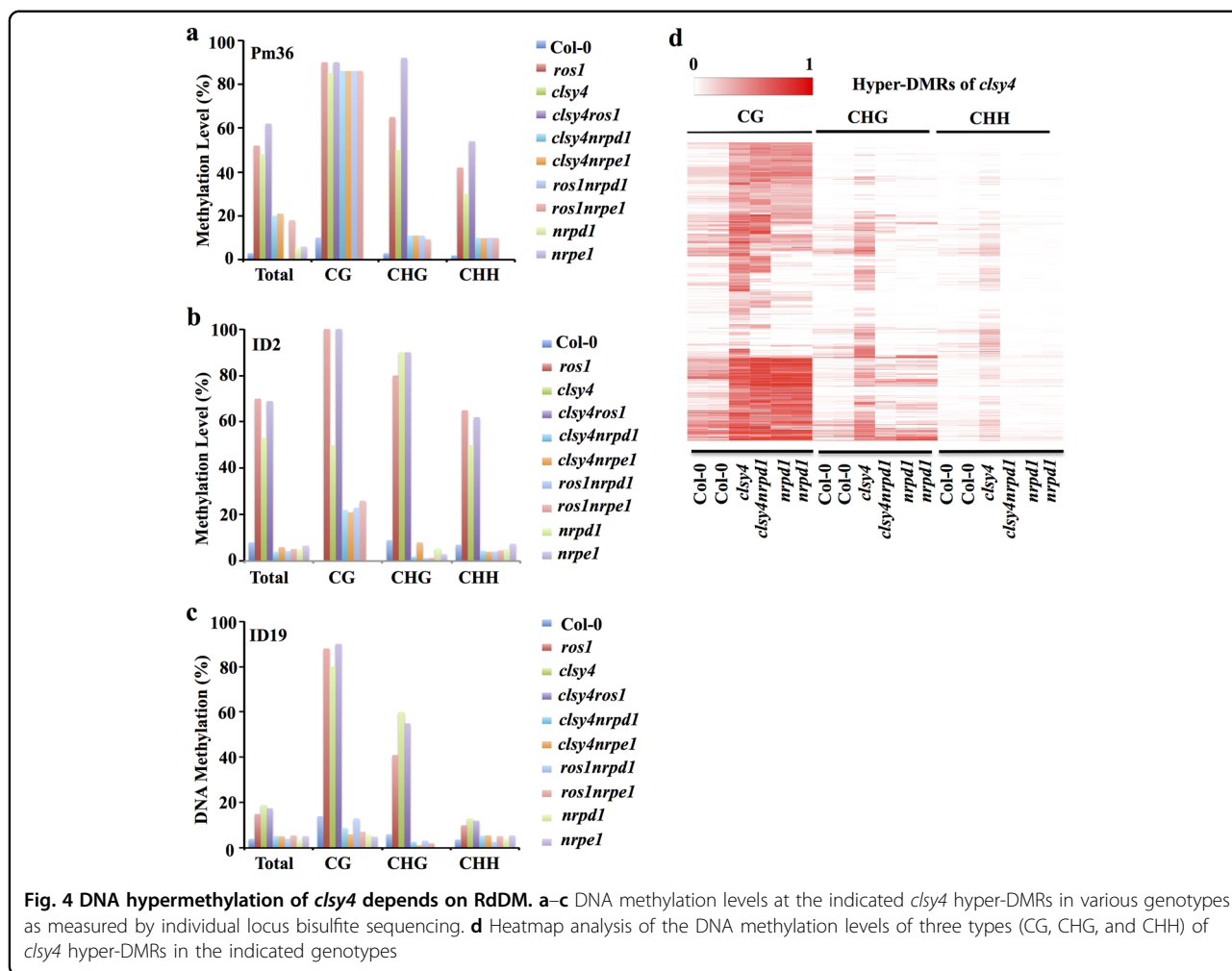
reduction at ID19 in *cly4nrpd1*, *cly4nrpd1*, *ros1nrpd1*, and *ros1nrpd1* (Fig. 4a–c). These results suggest that the DNA hypermethylation in *cly4* depends on the RdDM pathway. We did not observe any additive effect between *cly4* and *ros1* by comparing the DNA methylation levels at these loci in *cly4* and *ros1* single mutants to those in the *ros1cly4* double mutant (Fig. 4a–c). These genetic analysis results, together with the data showing that the DNA methylation levels at *cly4* hyper-DMRs were increased in *rdd* (Fig. 2a), strongly suggest that CLSY4/CHR40 and the ROS1 family DNA demethylases function in the same genetic pathway for DNA demethylation.

To further examine the dependence of the *cly4* hyper-DMRs on RdDM, we conducted high coverage (~58x) methylome analysis of the *cly4nrpd1* double mutant using whole-genome bisulfite sequencing (Supplementary Table S2). The cytosine methylation level at *cly4* hyper-DMRs was clearly reduced in the *cly4nrpd1* double mutant (Supplementary Fig. 8). In the CHG and CHH sequence contexts, DNA methylation in the double

mutant was significantly lower than that in the Col-0 wild type, but was comparable to that in the *nrpd1* mutant, whereas DNA methylation in the CG context was only slightly reduced in the double mutant (Supplementary Fig. S8a–d). Heatmap analysis also showed that the CHH methylation at most *cly4* hyper-DMRs was decreased in the *cly4nrpd1* double mutant (Fig. 4d), while CHG methylation was reduced at the majority of the hyper-DMRs, and CG methylation was reduced at a subset of the hyper-DMRs in the double mutant (Fig. 4d). These results suggest that the hypermethylation of non-CG sequences throughout the genome in *cly4* depends on RdDM.

Four CLSYs redundantly regulate DNA methylation at RdDM target loci

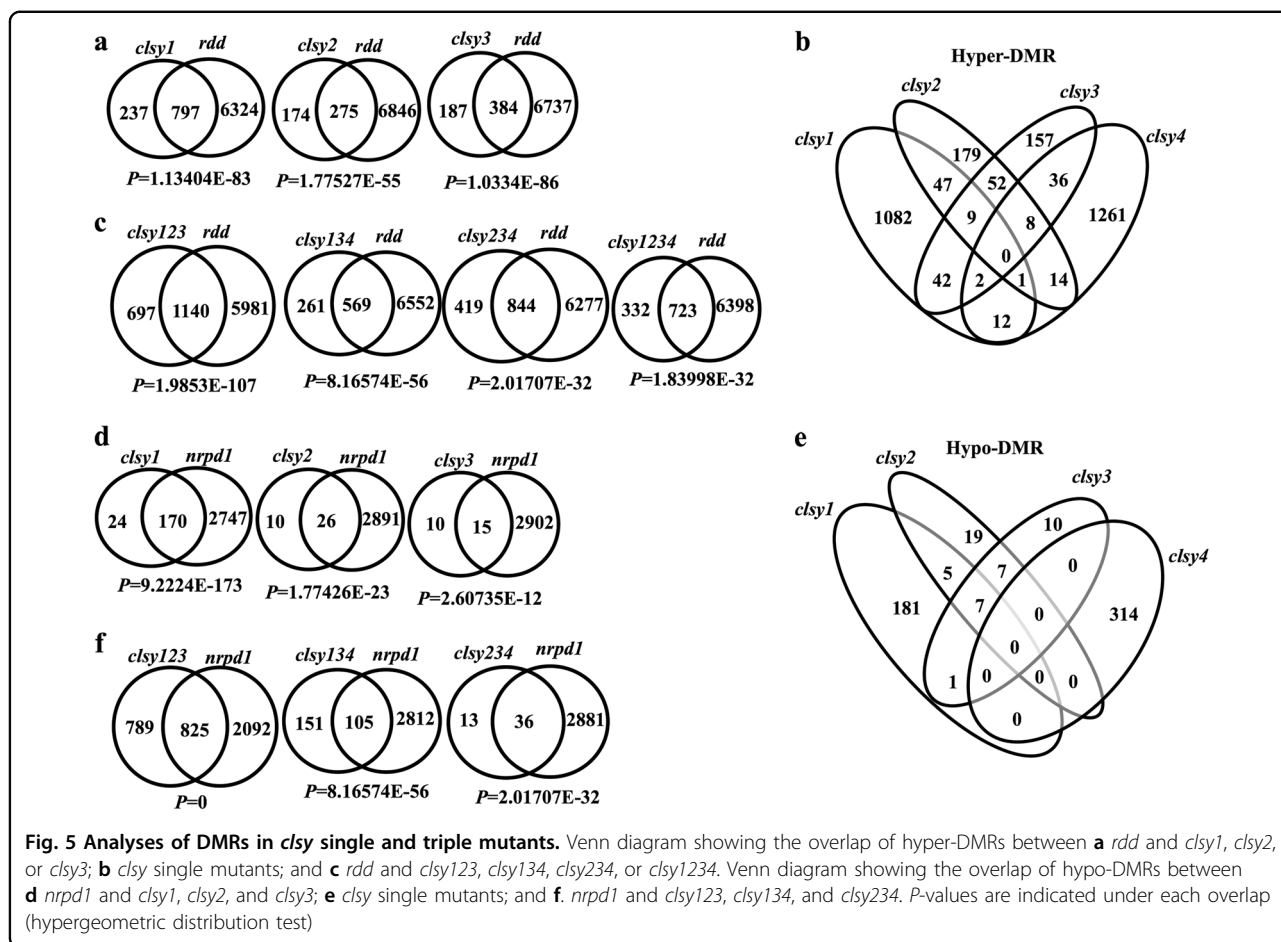
The Arabidopsis CLSY1, CLSY2, CLSY3, and CLSY4 are four closely related chromatin-remodeling proteins and thus may have redundant functions in regulating DNA methylation. To understand their genetic relationships, we crossed their respective mutants and



constructed high order mutants including three triple mutants (*clsy123* refers to *clsy1clsy2clsy3*, *clsy134* refers to *clsy1clsy3clsy4*, and *clsy234* refers to *clsy2clsy3clsy4*) and a quadruple mutant (*clsy1234* refers to *clsy1clsy2clsy3clsy4*) (Supplementary Table S3). Whole-genome bisulfite sequencing was conducted in the high order mutants together with the single mutants with high coverage (Supplementary Table S2). Hundreds of hyper-DMRs were identified in *clsy1*, *clsy2*, and *clsy3* single mutants, with the majority being shared by the *rdd* mutant, i.e., 77% in *clsy1*, 61% in *clsy2*, and 67% in *clsy3* overlapped with the hyper-DMRs in *rdd* (Fig. 5a). Consistent with the hypermethylation in *clsy* single mutants, the CHG and CHH methylation level of transposons was higher in *clsy1*, *clsy2*, and *clsy3* than that in the wild type (Supplementary Fig. S9a). The overlapped hyper-DMRs among the four *clsy* single mutants were very limited (Fig. 5b), suggesting that each CLSY controls DNA demethylation in a locus-specific manner. The majority of the hyper-DMRs in the triple mutants and the quadruple mutant were also shared by *rdd*, with 62% overlap for *clsy123*, 69% overlap for

clsy134, 67% overlap for *clsy234*, and 69% overlap for *clsy1234* (Fig. 5c).

In contrast to the hypo-DMRs in *clsy4* (Fig. 3a), the majority of hypo-DMRs in the other three *clsy* single mutants were shared by *nrdp1* (Fig. 5d), indicating important roles of CLSY1, 2, and 3 in RdDM. Again, each CLSY affected DNA methylation in largely distinct genomic regions (Fig. 5e). Even though certain hypo-DMRs were identified in the four *clsy* single mutants, DNA methylation levels at gene body and transposon regions in the single mutants were not as low as those in the *nrdp1* mutant (Supplementary Fig. S9a), suggesting that the four CLSYs are largely redundant in mediating RdDM. The triple mutant *clsy123* had the highest number of hypo-DMRs ($n = 1614$), while *clsy134* had an intermediate number ($n = 256$), and *clsy234* had the lowest number ($n = 49$) (Fig. 5f). The majority of the hypo-DMRs in the three triple mutants were RdDM targets (Fig. 5f). In addition, among the three triple mutants, *clsy123* had the lowest and *clsy234* had the highest CHH methylation levels at both gene and transposon regions



(Supplementary Fig. S9b), suggesting a predominant role of CLSY1 and a minor role of CLSY4 in controlling of DNA methylation at RdDM target regions.

The CHG and CHH methylation levels at gene body regions in the *clsy1234* mutant were quite similar to those in *nrpd1* (Supplementary Fig. S9b). At transposons, which are the major RdDM targets, the CHG and CHH methylation levels were reduced more in *clsy1234* than in *nrpd1* (Supplementary Fig. S9b). Thousands of hypo-DMRs were identified in *clsy1234*, the majority of which were shared by *nrpd1* (Fig. 6a). Importantly, the methylation level was also reduced in *nrpd1* at *clsy1234*-specific hypo-DMRs, and vice versa (Fig. 6a). Heatmap analysis showed highly similar patterns of CHG and CHH hypomethylation in *clsy1234* and *nrpd1* at *nrpd1* hypo-DMRs (Fig. 6b). DNA hypomethylation at RdDM target loci was observed in the single mutants *clsy1* and *clsy4*, whereas *clsy2* and *clsy3* showed little difference compared to the wild type (Fig. 6b). Non-CG methylation was moderately reduced in *clsy123* and was less reduced in *clsy134*, whereas *clsy234* showed little reduction in non-CG methylation (Fig. 6b). The methylation analyses of

RdDM regions of the single and triple mutants also suggested that CLSY1 predominates among the four CLSY proteins in mediating RdDM, which might explain why CLSY1 but not the other three CLSY proteins has been identified from genetic screens for RdDM factors²⁰. These results demonstrate that the four CLSY members redundantly control DNA methylation at RdDM target loci. DNA methylation at several known RdDM target loci was also examined using chop-PCR. The chop-PCR results showed that IGN5, AtSN1, and Solo-LTR were hypomethylated in *clsy1234* as well as in the RdDM mutants *nrpd1* and *nrpe1* (Fig. 6c), supporting the inference that RdDM was blocked in the quadruple mutant.

***clsy* mutations impair siRNA production and transposon silencing**

Because DRD1 is a putative chromatin-remodeling protein that is closely related to the CLSYs and that functions to facilitate Pol V transcription³⁶, we tested whether the CLSY proteins might also be involved in Pol V transcription to produce the scaffold RNAs essential for AGO4 targeting³⁸. Quantitative RT-PCR assays showed

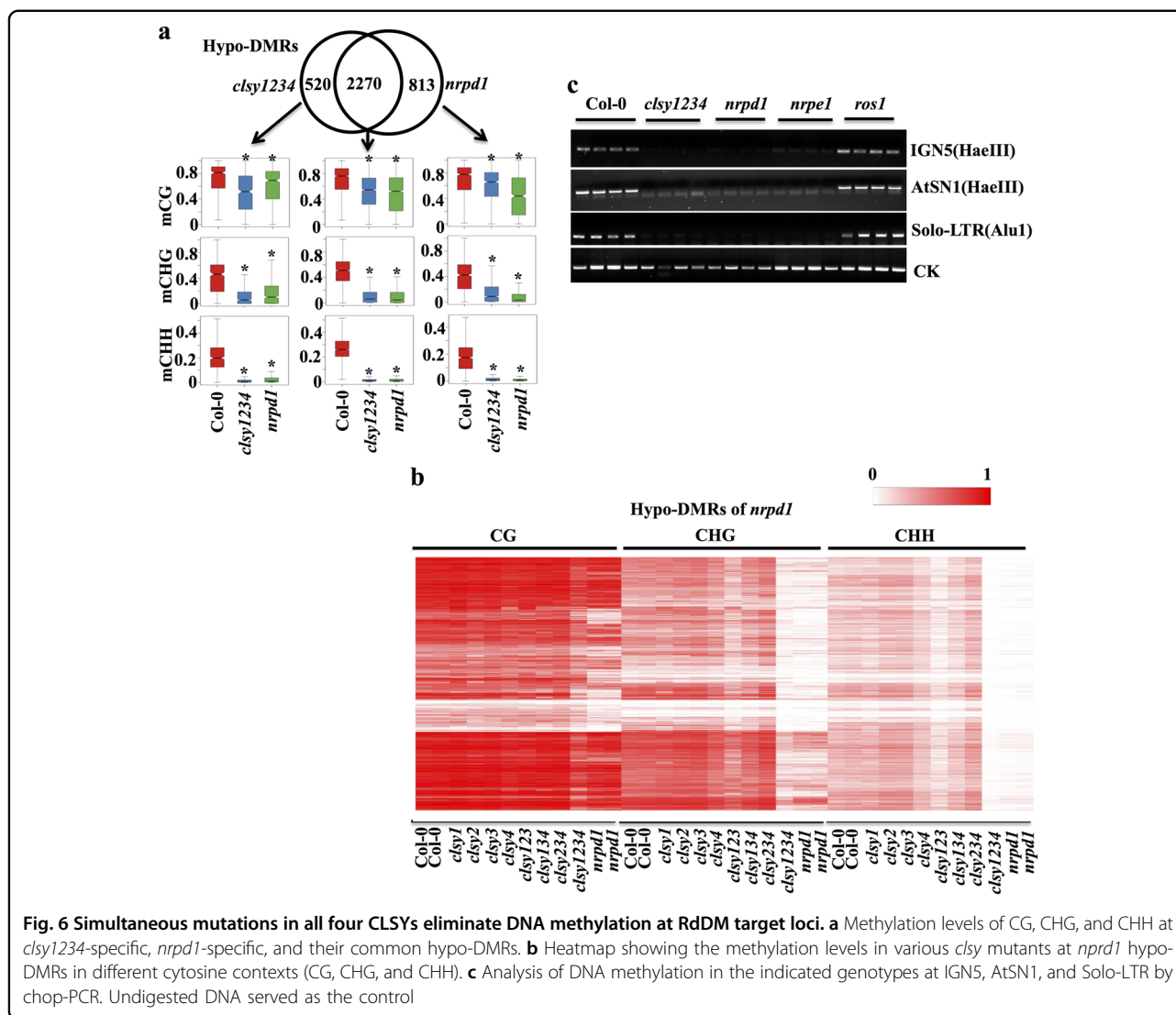
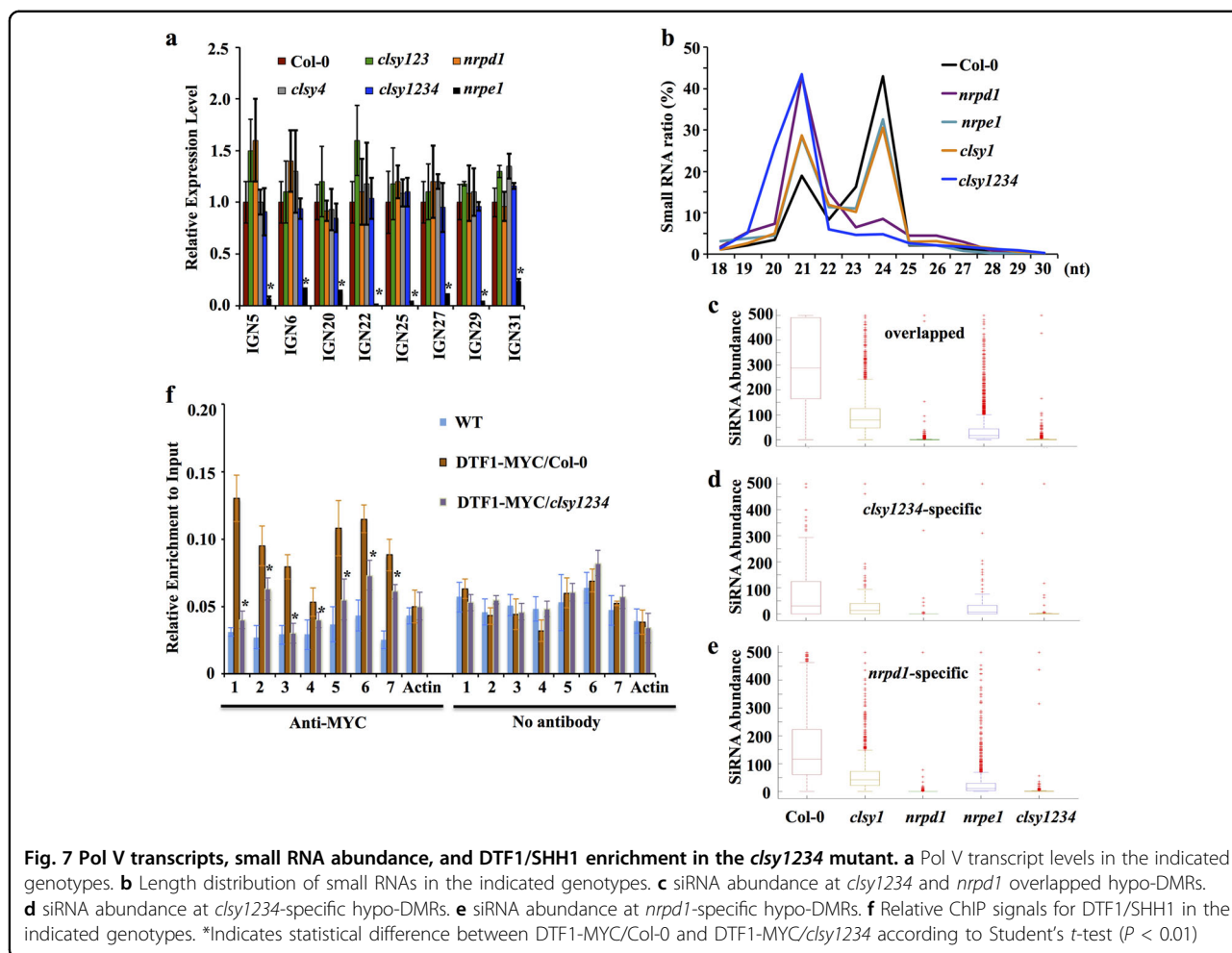


Fig. 6 Simultaneous mutations in all four CLSYs eliminate DNA methylation at RdDM target loci. a Methylation levels of CG, CHG, and CHH at *clsy1234*-specific, *nrpd1*-specific, and their common hypo-DMRs. **b** Heatmap showing the methylation levels in various *clsy* mutants at *nrpd1* hypo-DMRs in different cytosine contexts (CG, CHG, and CHH). **c** Analysis of DNA methylation in the indicated genotypes at IGN5, AtSN1, and Solo-LTR by chop-PCR. Undigested DNA served as the control

that the transcript levels of eight Pol V scaffold RNAs in *clsy4*, *clsy123*, and *clsy1234* were comparable to those in the wild type (Fig. 7a), suggesting that the CLSYs are not involved in Pol V transcription. The four CLSYs proteins are physically associated with Pol IV, which indicates their potential roles in siRNA biosynthesis³⁴. To test this possibility, we conducted small RNA sequencing using *clsy1234* mutant plants. As was the case in *nrpd1*, 24-nt siRNAs were almost eliminated in the quadruple mutant (Fig. 7b). In the *clsy1* single mutant, 24-nt siRNAs were slightly reduced, resembling the *nrpe1* mutant (Fig. 7b). These patterns suggest that the four CLSYs are collectively required for 24-nt siRNA production. At the *clsy1234*-specific, *nrpd1*-specific, and their overlapped hypo-DMRs (Fig. 6a), 24-nt siRNAs were nearly eliminated in *clsy1234* and *nrpd1* mutants but were moderately reduced in *clsy1* and *nrpe1* mutants (Fig. 7c–e),

supporting the inference that the four CLSYs redundantly regulate siRNA biogenesis at all RdDM target loci. Because DTF1/SHH1 is required for Pol IV function and 24-nt siRNA biogenesis^{18,19}, we examined DTF1/SHH1 enrichment at several RdDM target loci (Supplementary Fig. S10) and found that the chromatin occupancy of DTF1/SHH1 was less in *clsy1234* plants than that in wild-type plants (Fig. 7f). This suggested that the four CLSYs are required for proper DTF1/SHH1 binding at the RdDM target loci.

ROS1 expression level is an excellent indicator of RdDM activities, because *ROS1* expression positively correlates with RdDM-dependent DNA methylation in its promoter region^{30,39}. Like *nrpd1* and *nrpe1*, *clsy1234* showed a drastic reduction in *ROS1* transcript levels (Fig. 8a). Among the four single *clsy* mutants, *clsy1* had the lowest *ROS1* expression, while *clsy123* had the lowest



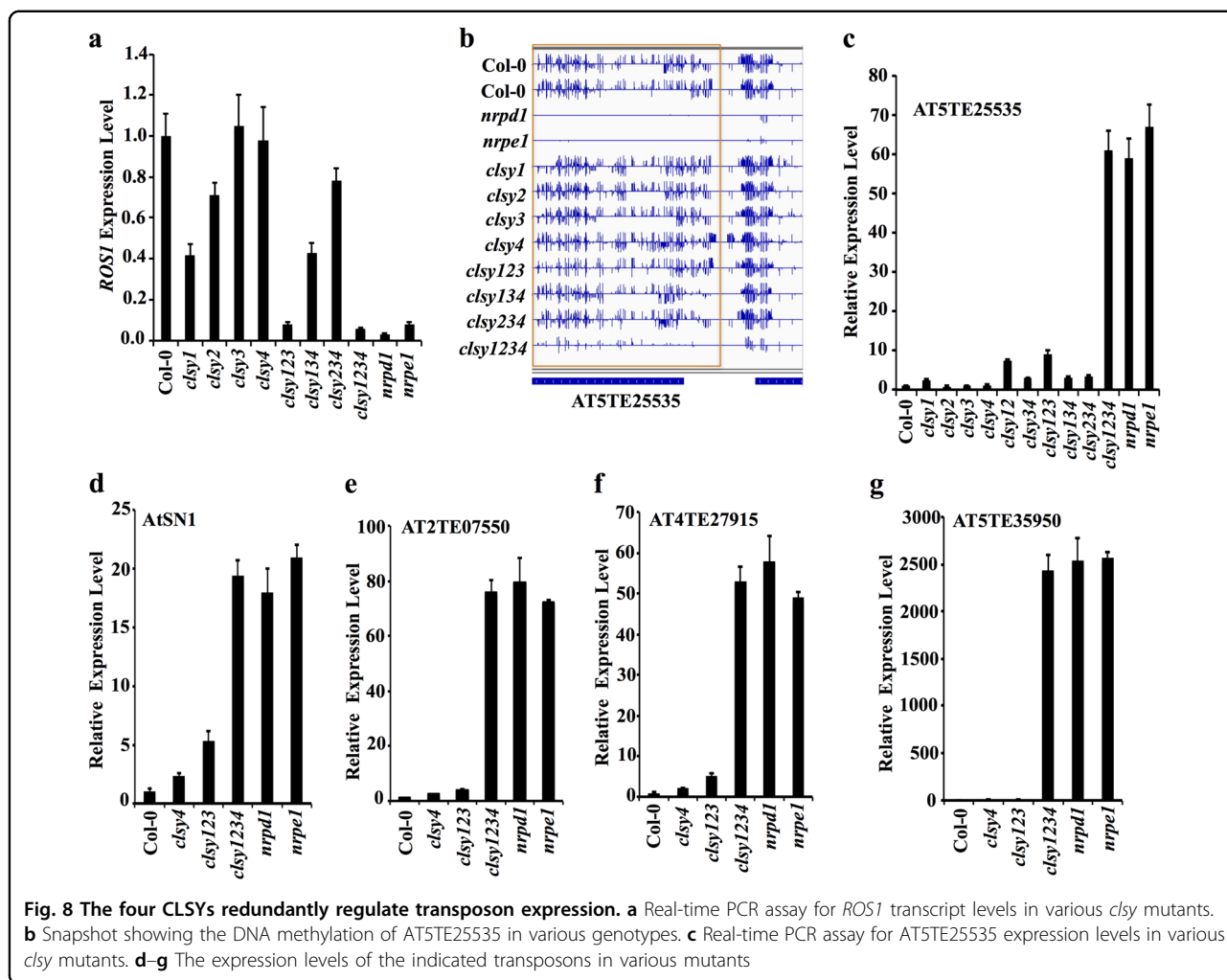
ROS1 expression among the three triple mutants (Fig. 8a). These patterns of *ROS1* gene repression are consistent with the DNA hypomethylation phenotypes in the mutants (Fig. 6b).

Because RdDM can transcriptionally silence transposons, we selected several TEs that are known to be released in the *drd1* mutant⁴⁰, and examined their DNA methylation levels and transcript levels in various *cly* mutants and *nrpd1* and *nrpe1* mutants. As shown in Fig. 8b, DNA methylation at AT5TE25535 was nearly eliminated in *cly1234* but not in the other *cly* mutants (Fig. 8b). Consistent with the DNA methylation patterns, the expression level of AT5TE25535 in *cly1234* was increased by about 60 fold, which was comparable to that in *nrpd1* and *nrpe1* mutants; in the other *cly* mutants, in contrast, AT5TE25535 expression levels were either not changed or were only slightly increased (Fig. 8c). As was the case at AT5TE25535, DNA methylation at the other four tested transposons was eliminated in *cly1234*, *nrpd1*, and *nrpe1* mutants, but remained largely unchanged in

the other *cly* mutants (Supplementary Fig. S11a–d). Their transcript levels were dramatically increased in *cly1234* as well as in *nrpd1* and *nrpe1* (Fig. 8d–g). Therefore, consistent with their redundant roles in RdDM, the four CLSY proteins have redundant roles in the transcriptional silencing of transposons.

The four CLSY members antagonistically regulate DNA methylation

Because the DNA hypermethylation in *cly4* mutants depends on RdDM, which in turn requires the redundant function of the four CLSY proteins, we suspected that the DNA demethylation function of one CLSY is antagonized by the other CLSY proteins. To examine this possibility, we examined methylation levels of three *cly4* hyper-DMRs using individual locus bisulfite sequencing. At Pm36, the non-CG hypermethylation of *cly4* was suppressed in *cly1234* mutant plants, while the CG hypermethylation did not change in the quadruple mutant (Fig. 9a). At the other two *cly4* hyper-DMRs, ID2 and

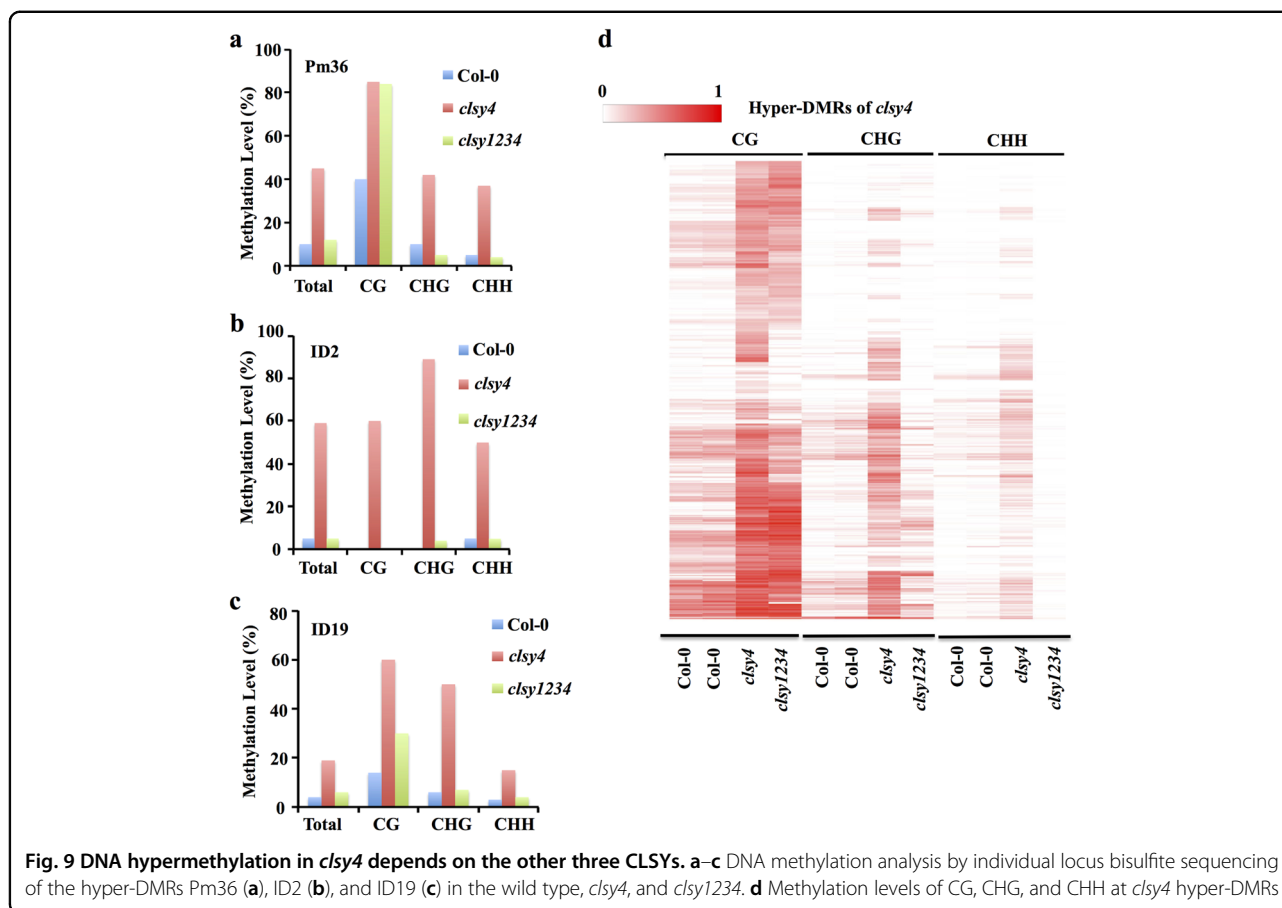


ID19, hypermethylation was reduced or abolished in all three contexts in the quadruple mutant (Fig. 9b, c). In addition, Integrated Genome Viewer (IGV) examination of randomly selected *clsy4* hyper-DMRs from whole-genome bisulfite sequencing results also showed a loss of DNA hypermethylation in *clsy1234* compared to *clsy4* (Supplementary Fig. S12a–f). These results revealed that, at some RdDM target loci, CLSY4-dependent DNA demethylation counteracts the DNA methylation mediated by the other three CLSY proteins, demonstrating an important role of the CLSY proteins in regulating DNA methylation dynamics.

On a whole-genome scale, CHG and CHH methylation at *clsy4* hyper-DMRs was significantly reduced in *clsy1234* compared to *clsy4*, as shown by boxplot analysis (Supplementary Fig. S12g–j). Moreover, heatmap analysis showed that the increased CHG and CHH DNA methylation at most *clsy4* hyper-DMRs was lost in *clsy1234*, whereas the CG hypermethylation at *clsy4* hyper-DMRs

largely remained in *clsy1234* (Fig. 9d). These genome-wide patterns further supported the conclusion that CLSY4 antagonizes the other CLSYs in balancing DNA methylation levels at specific genomic loci, although CLSY4 also positively contributes to DNA methylation at some other loci in a way that is genetically redundant with the other CLSYs.

We further investigated the relationship among the four CLSYs by comparing all *clsy* single mutants with *clsy1234*. At *clsy1*, *clsy2*, or *clsy3* hyper-DMRs, DNA hypermethylation was largely abolished in *clsy1234*, as shown by heatmap analyses (Supplementary Fig. S13a–c). In addition, at these *clsy* single mutant hyper-DMRs, cytosine methylation levels and especially CHG and CHH methylation levels were reduced in *clsy1234* (Supplementary Fig. S14a–c), indicating a general antagonistic relationship among the four CLSY members in regulating DNA methylation dynamics. Consistent with this inference, the hyper-DMRs in *clsy123*, *clsy134*, and *clsy234* were also



largely reduced in the *clsy1234* mutant (Supplementary Fig. S13d–f), with CHG and CHH methylation reduced more than CG methylation (Supplementary Fig. S15a–c).

Discussion

Arabidopsis CLSY1 was first identified in a genetic screen for cellular factors required for gene silencing induced by double-stranded RNAs and was demonstrated to function with RDR2 and NRPD1 in the production of 24-nt siRNA at endogenous loci²⁰. However, DNA methylome analysis of multiple RdDM mutants showed that the methylation level at RdDM target loci was only weakly reduced in *clsy1* mutants²¹, suggesting that other chromatin-remodeling proteins likely work together with CLSY1 in siRNA biogenesis. Later, CLSY1 and its three paralogs, CLSY2, CLSY3, and CLSY4, were identified as Pol IV-associated proteins, suggesting that these proteins probably function redundantly in siRNA production in the RdDM pathway³⁴. Here we profiled the DNA methylomes of single, triple, and quadruple *clsy* mutants and found that the DNA methylation of RdDM target loci is indeed cooperatively controlled by the four CLSY proteins. We found that 24-nt siRNAs were almost

eliminated in the *clsy1234* mutant. Although the detailed molecular mechanism underlying the involvement of the CLSYs in Pol IV function remains to be elucidated, chromatin enrichment of DTF1/SHH1, the cofactor important for Pol IV recruitment and 24-nt siRNA production at a subset of RdDM loci^{18,19}, was reduced in *clsy1234* at several of the tested genomic regions.

In this study, CLSY4 (a close homolog of CLSY1) was originally identified in our genetic screen for DNA demethylation factors. More than 1000 hyper-DMRs were found in *clsy4* mutant plants, indicating that CLSY4 is important for genome-wide DNA demethylation. *ROS1* expression was not reduced in *clsy4* mutants, indicating that CLSY4 does not function indirectly in demethylation by being an RdDM factor that controls *ROS1* expression. The biochemical function of CLSY4 in DNA demethylation is unknown, and we have not observed a physical association between ROS1 and any of the four CLSYs (unpublished data). Nevertheless, we found that the nucleosome occupancy was reduced in *clsy4* at several tested genomic loci targeted by active DNA demethylation. ATP-dependent chromatin-remodeling proteins generally function in nucleosome sliding, eviction, and

histone exchange on chromatin⁴¹. It is possible that active DNA demethylation at some genomic loci may somehow involve CLSY-mediated nucleosome movement. In support of this notion of a chromatin remodeling functioning in active DNA demethylation, Ikeda et al. have documented that the FACT histone chaperone component SSRP1 plays an important role in DME-mediated active DNA demethylation and gene imprinting⁴².

With the use of *clsy4nrpd1* and *clsy1234* mutants, our genetic analysis demonstrated that DNA hypermethylation in *clsy4* depends on the RdDM pathway, a pattern resembling that of the *ros1* mutant. It is possible that, besides CLSY4, the other three CLSYs also directly regulate DNA demethylation at discrete genomic regions, because hundreds of hyper-DMRs were identified in the single and triple *clsy* mutants. Those hyper-DMRs also depended on the RdDM pathway, because DNA hypermethylation was largely abolished in the *clsy1234* mutant, in which DNA methylation and siRNA accumulation at RdDM target loci was eliminated. Therefore, the dual function of CLSY proteins and especially of CLSY4, whose mutation results in the highest number of hyper-DMRs, in DNA methylation via RdDM and in active DNA demethylation mediated by the ROS1 family of DNA demethylases represents a new mechanism underlying the dynamic regulation of DNA methylation. DNA methylation and demethylation activities were previously determined to be coordinated by a “methylstat”, a methylation-sensitive cis-element at the *ROS1* promoter^{30,31}. The current results suggest that chromatin remodeling by the four CLSYs is another mechanism for the coordination of DNA methylation and demethylation.

Mutations in the CLSYs led to increased DNA methylation at certain genomic sites but decreased DNA methylation at other sites. The opposite DNA methylation phenotypes in the *clsy* mutants raises the possibility that other epigenetic marks, like histone modifications and pre-existing DNA methylation together with nucleosome occupancy regulated by CLSYs, may collectively determine whether a region undergoes RdDM or active DNA demethylation. The chromatin status regulated by the CLSYs is only one of the factors determining the dominance of RdDM or active DNA demethylation. How RdDM and active DNA demethylation differ in their requirements for chromatin environments warrants additional study.

Materials and methods

Plant materials and growth conditions

All Arabidopsis plants used in this study were in the Col-0 background. The *clsy1-3* mutant was reported previously²⁰. All other mutants (Supplementary Table S3) were either ordered from the Arabidopsis Biological

Resource Center (<http://www.arabidopsis.org>) or made in this study. The oligos used for genotyping and making higher order mutants are listed in Supplementary Table S4. Plants that were subjected to DNA methylation and small RNA analyses were 2-week-old seedlings grown on 1/2 strength Murashige Skoog (MS) medium supplemented with 0.7% agar and 1% sucrose. The seedlings were grown in growth chambers (Percival, Inc.) at 22 °C with 16 h of light and 8 h of darkness.

Chop-PCR and genetic screen

The T-DNA mutants (Supplementary Table S1) were grown in soil for 5 weeks. Four plants of each genotype were selected for DNA extraction by the CTAB method. About 2 µg of DNA of each plant was diluted and digested with HhaI or HpaII in a 20-µl reaction mixture for 8 h. The PCR reaction used 1 µl of enzyme-digested DNA as the template. Agarose (1%) gels were used to compare the amplified DNA signal, with the wild type and *ros1-4* as controls. Candidate mutants were then further genotyped and confirmed.

Individual locus bisulfite sequencing

After the genomic DNA was isolated from 4-week-old Arabidopsis plants using the Plant DNeasy Mini Kit (Qiagen), about 100 ng of genomic DNA was subjected to sodium bisulfite treatment and purification using the BisulFlash DNA Modification Kit (Epigentek) and following the manufacturer's protocol. A 1 µl volume of bisulfite-converted DNA was used for PCR reaction with oligos designed for specific target regions (Supplementary Table S4). The PCR products were ligated into the pMD18-T vector (Takara, Japan). The ligation product was transformed into DH5α competent cells, and at least 15 single clones were sequenced for each type of transformant. The sequencing results were aligned and analyzed using CyMate (<http://www.cymate.org>).

Whole-genome bisulfite sequencing and data analysis

Genomic DNA was extracted from 2-week-old seedlings grown in 1/2 strength MS medium and was sent to the Beijing Genomics Institute (Shenzhen, China) for bisulfite treatment, library construction, and high-throughput sequencing. Clean reads were generated by trimming adaptor and low-quality sequences ($q < 20$) and then were mapped to the Arabidopsis genome (TAIR 10) using BSMAP (Bisulfite Sequence Mapping Program) and allowing two mismatches. Differentially methylated regions (DMRs) were identified as previously described¹⁶. First, the cytosines with a read coverage < 5 were filtered out. The DNA methylation level in every 200-bp window with a step size of 50 bp was compared between control and mutant plants using Fisher's exact test. *P*-values calculated from the tests

were then adjusted using the Benjamini-Hochberg method to control for the false discovery rate (FDR). Windows with an adjusted P -value < 0.05 were then tested for the number of DMCs (differentially methylated cytosines), which were defined as cytosines with a P -value < 0.01 in Fisher's exact test. The DMR-identified windows with at least 7 DMCs and ≥ 1.5 -fold change in DNA methylation levels were combined to generate the final list of DMRs if the gap length between two windows was ≤ 100 bp.

All of the figures related to DNA methylomes were generated using R (<https://www.r-project.org/>). The heatmaps were generated using "heatmap.2" from "gplots" library, with the following parameters: Colv = FALSE, dendrogram = "row", trace = "none", and labRow = "NA". The methylation data were not normalized for generating the heatmaps.

Genome-wide small RNA sequencing and analysis

Total RNA was isolated from 2-week-old seedlings with TRIzol reagent (Sigma) and the standard protocol. Total RNA was separated on PAGE gels and < 35 -nt fractions were cut and purified for small RNA library preparation and sequencing at BGI (Shenzhen, China). After sequencing, adapter sequences were trimmed, and clean reads from 18 to 30 nt were mapped to the Arabidopsis genome (TAIR10) using Bowtie with parameter "-v 0 -k 10"¹⁶. Reads were mapped to annotated structural RNAs including tRNAs, rRNAs, snRNAs, and snoRNAs were excluded. Read counts were normalized to Reads Per Ten Million (RPTM) based on total mapped reads. The "hits-normalized abundance" (HNA) values were calculated by dividing the normalized abundance (in RPTM) for each small RNA hit, where a hit was simply defined as the number of loci at which a given sequence perfectly matched the genome¹⁶, and the values were used for boxplot generation.

Chromatin immunoprecipitation assay

The chromatin immunoprecipitation was conducted using 2-week-old seedlings of the wild type Col-0, DTF1-MYC/*clsy1234*, and DTF1-MYC/Col-0 grown in 1/2 strength MS solid medium as previously described⁴³. The antibody used for DTF1 ChIP assays was anti-MYC (16–219, Millipore). The DNA was diluted in 50 μ l of TE, and a 1.5 μ l volume was used for each qPCR reaction. The relative enrichment to input was calculated from three biological replicates.

Nuclei were extracted from 2-week-old Arabidopsis seedlings grown in 1/2 strength MS medium as previously described³⁸ and were digested with Micrococcal Nuclease (MNase; NEB) before chromatin immunoprecipitation was conducted using antibodies against H3 (Abcam, ab1791) as previously reported⁴⁴.

Small RNA extraction and northern blotting

After total RNA was extracted from 2-week-old Arabidopsis seedlings with TRIzol reagent (Sigma) and the standard protocol, the small RNA fraction was precipitated by the PEG method. In brief, equal volumes of a PEG8000 solution (20% PEG 8000, 1 M NaCl) were added to the total RNA. After centrifugation at 16000 g at 4 °C, an equal volume of isopropanol and 0.1 volume of 3 M NaAc were added to the supernatant. Following another centrifugation at 16000 g at 4 °C, the small RNA pellet was re-suspended in DEPC-treated water. For each sample, small RNAs from about 100 μ g of total RNA were separated on a 17% polyacrylamide gel, which was electrotransferred to a Hybond N+ membrane (GE Lifesciences). Membranes were cross-linked, incubated for 2 h at 80 °C, and hybridized overnight at 38 °C with ³²P-labelled DNA probes or oligonucleotides (listed in Supplementary Table S4) in PerfectHyb buffer (Sigma). The washed membranes were then exposed in the presence of X-ray film at -80 °C for 7 days.

Data access

The sequencing data generated in this study have been deposited in the NCBI Gene Expression Omnibus (GEO; <http://www.ncbi.nlm.nih.gov/geo/>) and are accessible through accession number GSE110338.

Acknowledgements

We thank David C. Baulcombe for the seeds of the *clsy1-3* mutant. This work was supported by the Chinese Academy of Sciences to J.-K. Z., and by the National Science Foundation of China (31671340), the Natural Science Foundation of Jiangsu (BK20170027), the Innovative Project of State Key Laboratory of Crop Genetics and Germplasm Enhancement (ZZ2017009), Fundamental Research Funds for the Central Universities (KYZ201503, KJJQ201801), and the Jiangsu Collaborative Innovation Center for Modern Crop Production to D.-L. Y.

Author details

¹Shanghai Center for Plant Stress Biology and Center of Excellence in Molecular Plant Sciences, Chinese Academy of Science, Shanghai 200032, China. ²State Key Laboratory of Crop Genetics and Germplasm Enhancement, Nanjing Agricultural University, Nanjing 210095, China. ³Department of Horticulture and Landscape Architecture, Purdue University, West Lafayette, 47907 IN, USA

Author contributions

Experimental design and data analysis: D.L.Y. and J.K.Z. Experimental performance: D.L.Y., G.Z., D.X., D.M. Bioinformatics analysis: L.Y., L.Z., K.T., and L.W. Contribution of reagents/materials/analysis tools: J.L., C.G.D., and C.D. Manuscript writing: D.L.Y., H.Z., and J.K.Z.

Conflict of interest

The authors declare that they have no conflict of interest.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary Information accompanies the paper at (<https://doi.org/10.1038/s41421-018-0056-8>).

Received: 14 June 2018 Revised: 3 August 2018 Accepted: 7 August 2018
Published online: 16 October 2018

References

- Slotkin, R. K. & Martienssen, R. Transposable elements and the epigenetic regulation of the genome. *Nat. Rev. Genet.* **8**, 272–285 (2007).
- Vaillant, I. & Paszkowski, J. Role of histone and DNA methylation in gene regulation. *Curr. Opin. Plant. Biol.* **10**, 528–533 (2007).
- He, X. J., Chen, T. & Zhu, J. K. Regulation and function of DNA methylation in plants and animals. *Cell Res.* **21**, 442–465 (2011).
- Law, J. A. & Jacobsen, S. E. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat. Rev. Genet.* **11**, 204–220 (2010).
- Robertson, K. D. DNA methylation and human disease. *Nat. Rev. Genet.* **6**, 597–610 (2005).
- Lang, Z. et al. Critical roles of DNA demethylation in the activation of ripening-induced genes and inhibition of ripening-repressed genes in tomato fruit. *Proc. Natl Acad. Sci. USA* **114**, E4511–E4517 (2017).
- Zhang, H. & Zhu, J. K. RNA-directed DNA methylation. *Curr. Opin. Plant. Biol.* **14**, 142–147 (2011).
- Pikaard, C. S., Haag, J. R., Pontes, O. M., Blevins, T. & Cocklin, R. A transcription fork model for Pol IV and Pol V-dependent RNA-directed DNA methylation. *Cold Spring Harb. Symp. Quant. Biol.* **77**, 205–212 (2012).
- Matzke, M. A. & Mosher, R. A. RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nat. Rev. Genet.* **15**, 394–408 (2014).
- Cuerda-Gil, D. & Slotkin, R. K. Non-canonical RNA-directed DNA methylation. *Nat. Plants* **2**, 16163 (2016).
- Zhang, X., Henderson, I. R., Lu, C., Green, P. J. & Jacobsen, S. E. Role of RNA polymerase IV in plant small RNA metabolism. *Proc. Natl Acad. Sci. USA* **104**, 4536–4541 (2007).
- Mosher, R. A., Schwach, F., Studholme, D. & Baulcombe, D. C. PolIVb influences RNA-directed DNA methylation independently of its role in siRNA biogenesis. *Proc. Natl Acad. Sci. USA* **105**, 3145–3150 (2008).
- Huang, Y. et al. Ancient Origin and Recent Innovations of RNA Polymerase IV and V. *Mol. Biol. Evol.* **32**, 1788–1799 (2015).
- Zhai, J. et al. A One Precursor One siRNA Model for Pol IV-Dependent siRNA Biogenesis. *Cell* **163**, 445–455 (2015).
- Blevins, T. et al. Identification of Pol IV and RDR2-dependent precursors of 24 nt siRNAs guiding de novo DNA methylation in Arabidopsis. *eLife* **4**, e09591 (2015).
- Yang, D. L. et al. Dicer-independent RNA-directed DNA methylation in Arabidopsis. *Cell Res.* **26**, 66–82 (2016).
- Ye, R. et al. A dicer-independent route for biogenesis of siRNAs that direct DNA methylation in Arabidopsis. *Mol. Cell* **61**, 222–235 (2016).
- Law, J. A. et al. Polymerase IV occupancy at RNA-directed DNA methylation sites requires SHH1. *Nature* **498**, 385–389 (2013).
- Zhang, H. et al. DTF1 is a core component of RNA-directed DNA methylation and may assist in the recruitment of Pol IV. *Proc. Natl Acad. Sci. USA* **110**, 8290–8295 (2013).
- Smith, L. M. et al. An SNF2 protein associated with nuclear RNA silencing and the spread of a silencing signal between cells in Arabidopsis. *Plant Cell* **19**, 1507–1521 (2007).
- Stroud, H., Greenberg, M. V., Feng, S., Bernatavichute, Y. V. & Jacobsen, S. E. Comprehensive analysis of silencing mutants reveals complex regulation of the Arabidopsis methylome. *Cell* **152**, 352–364 (2013).
- Tang, K., Lang, Z., Zhang, H. & Zhu, J. K. The DNA demethylase ROS1 targets genomic regions with distinct chromatin modifications. *Nat. Plants* **2**, 16169 (2016).
- Gong, Z. et al. ROS1, a repressor of transcriptional gene silencing in Arabidopsis, encodes a DNA glycosylase/lyase. *Cell* **111**, 803–814 (2002).
- Gehring, M. et al. DEMETER DNA glycosylase establishes MEDEA polycomb gene self-imprinting by allele-specific demethylation. *Cell* **124**, 495–506 (2006).
- Agius, F., Kapoor, A. & Zhu, J. K. Role of the Arabidopsis DNA glycosylase/lyase ROS1 in active DNA demethylation. *Proc. Natl Acad. Sci. USA* **103**, 11796–11801 (2006).
- Ortega-Galisteo, A. P., Morales-Ruiz, T., Ariza, R. R. & Roldan-Arjona, T. Arabidopsis DEMETER-LIKE proteins DML2 and DML3 are required for appropriate distribution of DNA methylation marks. *Plant Mol. Biol.* **67**, 671–681 (2008).
- He, X. J. et al. An effector of RNA-directed DNA methylation in Arabidopsis is an ARGONAUTE 4- and RNA-binding protein. *Cell* **137**, 498–508 (2009).
- Gao, Z. et al. An RNA polymerase II- and AGO4-associated protein acts in RNA-directed DNA methylation. *Nature* **465**, 106–109 (2010).
- Mathieu, O., Reinders, J., Caikovski, M., Smathajitt, C. & Paszkowski, J. Transgenerational stability of the Arabidopsis epigenome is coordinated by CG methylation. *Cell* **130**, 851–862 (2007).
- Lei, M. et al. Regulatory link between DNA methylation and active demethylation in Arabidopsis. *Proc. Natl Acad. Sci. USA* **112**, 3553–3557 (2015).
- Williams, B. P., Pignatta, D., Henikoff, S. & Gehring, M. Methylation-sensitive expression of a DNA demethylase gene serves as an epigenetic rheostat. *PLoS. Genet.* **11**, e1002142 (2015).
- Qian, W. et al. A histone acetyltransferase regulates active DNA demethylation in Arabidopsis. *Science* **336**, 1445 (2012).
- Qian, W. et al. Regulation of active DNA demethylation by an alpha-crystallin domain protein in Arabidopsis. *Mol. Cell* **55**, 361–371 (2014).
- Law, J. A., Vashisht, A. A., Wohlschlegel, J. A. & Jacobsen, S. E. SHH1, a homeodomain protein required for DNA methylation, as well as RDR2, RDM4, and chromatin remodeling factors, associate with RNA polymerase IV. *PLoS. Genet.* **7**, e1002195 (2011).
- Li, X. et al. Antisilencing role of the RNA-directed DNA methylation pathway and a histone acetyltransferase in Arabidopsis. *Proc. Natl Acad. Sci. USA* **109**, 11425–11430 (2012).
- Law, J. A. et al. A protein complex required for polymerase V transcripts and RNA-directed DNA methylation in Arabidopsis. *Curr. Biol.* **20**, 951–956 (2010).
- Chodavarapu, R. K. et al. Relationship between nucleosome positioning and DNA methylation. *Nature* **466**, 388–392 (2010).
- Wierzbicki, A. T., Ream, T. S., Haag, J. R. & Pikaard, C. S. RNA polymerase V transcription guides ARGONAUTE4 to chromatin. *Nat. Genet.* **41**, 630–634 (2009).
- Huettel, B. et al. Endogenous targets of RNA-directed DNA methylation and Pol IV in Arabidopsis. *EMBO J.* **25**, 2828–2836 (2006).
- Zemach, A. et al. The Arabidopsis nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. *Cell* **153**, 193–205 (2013).
- Ryan, D. P. & Owen-Hughes, T. Snf2-family proteins: chromatin remodellers for any occasion. *Curr. Opin. Chem. Biol.* **15**, 649–656 (2011).
- Ikeda, Y. et al. HMG domain containing SSRP1 is required for DNA demethylation and genomic imprinting in Arabidopsis. *Dev. Cell.* **21**, 589–596 (2011).
- Saleh, A., Alvarez-Venegas, R. & Avramova, Z. An efficient chromatin immunoprecipitation (ChIP) protocol for studying histone modifications in Arabidopsis plants. *Nat. Protoc.* **3**, 1018–1025 (2008).
- Zhu, Y., Rowley, M. J., Bohmdorfer, G. & Wierzbicki, A. T. A SWI/SNF chromatin-remodeling complex acts in noncoding RNA-mediated transcriptional silencing. *Mol. Cell* **49**, 298–309 (2013).