The splicing machinery promotes RNA-directed DNA methylation and transcriptional silencing in *Arabidopsis*

Cui-Jun Zhang^{1,5}, Jin-Xing Zhou^{1,5},
Jun Liu^{1,5}, Ze-Yang Ma¹, Su-Wei Zhang¹,
Kun Dou¹, Huan-Wei Huang¹, Tao Cai¹,
Renyi Liu², Jian-Kang Zhu^{3,4} and
Xin-Jian He^{1,*}

¹National Institute of Biological Sciences, Beijing, China, ²Department of Botany and Plant Sciences, University of California Riverside, Riverside, CA, USA, ³Department of Horticulture and Landscape Architecture, Purdue University, West Lafayette, IN, USA and ⁴Shanghai Center for Plant Stress Biology, Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China

DNA methylation in transposons and other DNA repeats is conserved in plants as well as in animals. In *Arabidopsis thaliana*, an RNA-directed DNA methylation (RdDM) pathway directs *de novo* DNA methylation. We performed a forward genetic screen for suppressors of the DNA demethylase mutant *ros1* and identified a novel Zinc-finger and OCRE domain-containing Protein 1 (ZOP1) that promotes Pol IV-dependent siRNA accumulation, DNA methylation, and transcriptional silencing. Whole-genome methods disclosed the genome-wide effects of *zop1* on Pol IV-dependent siRNA accumulation and DNA methylation, suggesting that ZOP1 has both RdDM-dependent and -independent roles in transcriptional silencing. We demonstrated that ZOP1 is a pre-mRNA splicing factor that associates with several typical components of the splicing machinery as well as with Pol II. Immunofluorescence assay revealed that ZOP1 overlaps with Cajal body and is partially colocalized with NRPE1 and DRM2. Moreover, we found that the other development-defective splicing mutants tested including *mac3a3b*, *mos4*, *mos12* and *mos14* show defects in RdDM and transcriptional silencing. We propose that the splicing machinery rather than specific splicing factors is involved in promoting RdDM and transcriptional silencing.

The EMBO Journal (2013) 32, 1128–1140. doi:10.1038/emboj.2013.49; Published online 22 March 2013

Subject Categories: chromatin & transcription; RNA

Keywords: DNA methylation; RdDM; splicing; transcriptional silencing; ZOP1

Introduction

RNA-mediated repressive chromatin modifications and transcriptional silencing is a conserved mechanism that is required for maintenance of genome stability, repression of transposable elements (TEs), and regulation of genes in eukaryotes (Slotkin and Martienssen, 2007; Matzke *et al*, 2009; Law and Jacobsen, 2010). In *Arabidopsis*, an RNA-directed DNA methylation (RdDM) pathway has been characterized (Matzke *et al*, 2009; Law and Jacobsen, 2010). The transposons and other DNA repeats are transcribed by a plant-specific DNA-dependent RNA polymerase Pol IV and give rise to double-stranded RNAs by RNA-dependent RNA polymerase RDR2 (Xie *et al*, 2004; Herr *et al*, 2005; Kanno *et al*, 2005; Pontier *et al*, 2005; Ream *et al*, 2009). The double-stranded RNAs are cleaved into 24-nt small interfering RNAs (siRNAs) by a Dicer-like protein DCL3 (Xie *et al*, 2004). The Pol IV-interaction proteins, CLSY and SHH1/DTF1, are required for siRNA generation at a subset of RdDM target loci (Smith *et al*, 2007; Law *et al*, 2011; Liu *et al*, 2011). The siRNAs are loaded onto an ARGONAUTE protein AGO4, which interacts with NRPE1, the largest subunit of another DNA-dependent RNA polymerase, Pol V, and a transcription elongation factor-like protein, KTF1 (Pontes *et al*, 2006; Qi *et al*, 2006; Wierzbicki *et al*, 2008; He *et al*, 2009a). Pol V is required for the transcription of intergenic non-coding RNAs (Wierzbicki *et al*, 2008). DRD1, DMS3, and RDM1 form a tight complex, termed as DDR, that facilitates generation of Pol V-dependent RNA transcripts (Kanno *et al*, 2004, 2008; Gao *et al*, 2010; Law *et al*, 2010). The IWR1-like protein RDM4/DMS4 is a transcription factor that interacts with Pol II, Pol IV, and Pol V (He *et al*, 2009b; Kanno *et al*, 2010; Law *et al*, 2011). Pol V-dependent RNA transcripts are required for the association of AGO4 with chromatin (Wierzbicki *et al*, 2009). In the RdDM pathway, the DNA methyltransferase DRM2 is eventually recruited to RdDM target loci and catalyses DNA methylation (Cao and Jacobsen, 2002; Gao *et al*, 2010). Besides the canonical RdDM components, Pol II and its mediator (a multi-subunit regulator of Pol II) were also demonstrated to be required for RdDM and transcriptional gene silencing (Zheng *et al*, 2009; Kim *et al*, 2011). The non-coding scaffold RNAs produced by Pol II recruit AGO4 and Pol V to a subset of RdDM target loci in order to promote siRNA-mediated transcriptional gene silencing (Zheng *et al*, 2009).

The *Arabidopsis* RdDM system parallels the fission yeast RNAi-induced heterochromatin formation machinery in that they share several conserved RNAi components, including Dicer proteins, RNA-dependent RNA polymerases, Argonaute proteins, and WG/GW-containing proteins. The fission yeast RNAi-induced heterochromatin formation and transcriptional silencing is essential for the heterochromatin assembly on the *dh* and *dg* repeats of centromeric regions (Moazed, 2009; Grewal, 2010). Non-coding RNAs are transcribed from the

*Corresponding author. National Institute of Biological Sciences, Zhongguancun Life Science Park, No. 7, Science Park Road, Beijing 102206, China. Tel.: +86 10 80707712; Fax: +86 10 80707715; E-mail: hexinjian@nibs.ac.cn

⁵These authors contributed equally to this work.

Received: 3 September 2012; accepted: 7 February 2013; published online: 22 March 2013

repeat regions by Pol II and give rise to double-stranded RNAs that are processed into siRNAs by Dcr1 (Hall *et al*, 2002; Volpe *et al*, 2002). In fission yeast, Ago1, Tas3, and Chp1 are assembled into an RNAi-induced transcriptional silencing (RITS) effector complex containing siRNAs (Verdel *et al*, 2004). The RITS complex facilitates the recruitment of the RNA-dependent RNA polymerase complex (RDRC, containing Rdp1, cid12, and Hrr1). RDRC increases the production of double-stranded RNA precursors, reinforcing siRNA generation (Motamedi *et al*, 2004). The processing of RNA precursors into siRNAs ultimately recruits histone H3K9 dimethylase Clr4 to the centromeric repeat regions and promotes histone H3K9 dimethylation and heterochromatin formation (Buhler *et al*, 2006; Grewal, 2010).

The recognition and removal of introns from pre-mRNA is catalysed by spliceosome complexes in eukaryotes (Sharp, 1994). The major spliceosome complex is composed of five small nuclear ribonucleoproteins (snRNPs: U1, U2, U4, U6, and U5) and numerous non-snRNP splicing factors (Zhou *et al*, 2002). After the core proteins of snRNPs are assembled in the cytoplasm, they are transported to the nucleolus-adjacent Cajal body for assembly and modification (Cioce and Lamond, 2005; Morris, 2008). To assemble a mature spliceosome on pre-mRNA, U1 snRNP and U2AF (U2 snRNP auxiliary factor) bind to the 5' splice site and the branchpoint of the 3' splice site, respectively. Thus, the U4/U6-U5 tri-snRNP is added to the intermediate spliceosome, which finally leads to the displacement of U1 and U4 snRNPs and the formation of a catalytic mature spliceosome (Matlin and Moore, 2007). Assembly of snRNPs and spliceosome complexes is facilitated by many non-snRNP splicing factors.

Recent studies revealed that the spliceosome proteins and splicing factors affect not only pre-mRNA splicing but also Pol II transcription elongation, mRNA polyadenylation, and telomerase RNA biogenesis and processing, indicating that splicing-related proteins have multiple functions (Kaida *et al*, 2010; Shukla *et al*, 2011; Tang *et al*, 2012). Several specific splicing factors are also required for RNAi-induced heterochromatin formation and transcriptional silencing in fission yeast (Bayne *et al*, 2008; Chinen *et al*, 2010). The underlying mechanism remains to be elucidated. In *Arabidopsis*, the splicing factor SR45 was recently found to be required for RdDM, but an indirect role of SR45 via the splicing of RdDM regulator genes was not precluded in this report (Ausin *et al*, 2012).

The Cajal body is a dynamic nuclear structure involved in snRNP assembly, rRNA processing, and telomerase RNP formation (Pikaard, 2006; Morris, 2008). The immunolocalization of RdDM proteins revealed that AGO4, RDR2, and DCL3 localize in the Cajal body (Li *et al*, 2006, 2008). The Cajal body mutant *colin* disrupts not only the formation of Cajal bodies but also the concentrated AGO4 foci in the Cajal body (Li *et al*, 2008). It is possible that proteins other than AGO4 in Cajal bodies are required for assembly and recruitment of the AGO4 effector complex in RdDM.

To increase our understanding of the RdDM mechanism in *Arabidopsis*, in the current study we searched for new RdDM mutants by screening for suppressors of *ros1* from the previously described EMS-mutagenized library in the *ros1* mutant background (Liu *et al*, 2011). The results showed that most of the identified mutants are new alleles of known RdDM components (unpublished data). In addition,

the screen identified a previously uncharacterized gene encoding Zinc finger (ZnF) and OCRE domain-containing Protein 1 (ZOP1). We demonstrated that ZOP1 is a novel nucleic acid-binding protein that is required for both RdDM and pre-mRNA splicing. Moreover, we found that the four other development-defective splicing mutants that were tested (*mac3a3b*, *mos4*, *mos12*, and *mos14*) also show defects in siRNA accumulation and DNA methylation. This study reveals a novel function of the splicing machinery in siRNA accumulation and RdDM.

Results

Identification and characterization of the *zop1* mutant

Previous studies suggest that the *RD29A* promoter-driven luciferase (*RD29A-LUC*) transgene, endogenous *RD29A*, and the 35S promoter-driven *NPTII* (*35S-NPTII*) transgene are highly expressed in wild-type plants under stress conditions, whereas loss-of-function mutation in the DNA demethylase gene *ROS1* silences the expression of all three genes (Gong *et al*, 2002). The previous forward genetic screens for suppressors of *ros1* identified most of the RdDM components and other important proteins involved in transcriptional gene silencing (He *et al*, 2009c; Liu *et al*, 2011). The RdDM pathway is required for the silencing of the *RD29A-LUC* transgene as well as of the endogenous *RD29A* in *ros1*, but is dispensable for the silencing of the *35S-NPTII* transgene from the same construct. In the current study, a new mutant, *zop1*, was identified as a suppressor of *ros1* (Figure 1A). Like the previously identified RdDM mutants, *zop1* suppresses the silencing of *RD29A-LUC* and endogenous *RD29A* but not of *35S-NPTII* in the *ros1* background, although suppression of silencing is less with *zop1* than with *nrpe1* (Figure 1A and B). The DNA methylation of both transgenic and endogenous *RD29A* promoters was tested by bisulphite sequencing. Similar to the results of previous reports, heavy DNA methylation in the *ros1* mutant occurs in all three cytosine contexts (CG, CHG, and CHH) at both transgenic and endogenous *RD29A* promoters. The high DNA methylation is substantially reduced at CHG and CHH sites in both *ros1zop1* and *ros1nrpe1* (Figure 1C and D).

We mapped the *zop1* mutation by using a F2 segregation population from the cross between the *ros1zop1* mutant in the C24 background and a homozygous *ros1* mutant (Salk_045303) in the Col-0 background. The plants that emitted high bioluminescence were selected from the F2 population for map-based cloning. The *zop1* mutation was mapped to a ~273-kb interval on Chromosome 1 (Supplementary Figure S1A). A single-nucleotide G to A substitution in AT1G49590 was identified by deep sequencing of the whole *ros1zop1* genome (Supplementary Figure S1B). The substitution creates a premature stop codon and truncates the protein at Arg18 (Supplementary Figure S1B). The construct harbouring the full AT1G49590 genomic sequence in-frame to the 3 × Flag tag was transformed into *ros1zop1* for a complementation test. The results show that the AT1G49590 transgene is able to restore the *RD29A-LUC* silencing (Supplementary Figure S1C). Moreover, the transgene complements the developmental defects of *ros1zop1* (Supplementary Figure S1D). The results suggest that AT1G49590 is the *ZOP1* gene that is not only required for

Figure 1 *RD29A-LUC* transgene silencing and DNA methylation are affected by *zop1*. (A) The expression of *RD29A-LUC* was measured by luminescence imaging in the wild type, *ros1*, *ros1zop1*, and *ros1nrpe1*. (B) The RNA transcripts of *LUC*, *RD29A*, and *NPTII* were detected by real-time RT-PCR in the indicated genotypes. The relative expression levels of the genes are shown. Total RNA was extracted from 2-week old seedlings with or without cold treatment (48 h, 4°C). (C, D) The DNA methylation at both transgene (C) and endogenous (D) *RD29A* promoters was determined by bisulphite sequencing. The percentage of methylated cytosines at CG, CHG, and CHH sites is separately shown in the charts.

RdDM and transcriptional gene silencing but also for proper development.

ZOP1 contains an N-terminal C2H2-type ZnF domain and a predicted octamer repeat (OCRE) domain (Supplementary Figure S1B and S2), and its name reflects its status as a zinc finger- and OCRE domain-containing protein. Previous studies reported that the OCRE-containing proteins are likely to be involved in nucleic acid binding and RNA metabolism (Callebaut and Mornon, 2005). However, the function of the ZnF and OCRE domains in ZOP1 remains to be elucidated. ZOP1 and its homologues are conserved from unicellular green algae to various angiosperms (Supplementary Figure S2), whereas there is no ZOP1 homologue in animals and fungi.

The *zop1* mutation impairs RdDM at endogenous genome targets

We measured the effect of the *zop1* mutation on accumulation of Pol IV-dependent siRNAs from *RD29A-LUC* transgene promoter and endogenous genome target loci. The 24-nt siRNA from the transgene promoter is blocked in *ros1nrpd1* (He *et al*, 2009c), suggesting that it is a Pol IV-dependent siRNA. Our small RNA northern blot analysis indicates that the accumulation of the *RD29A* promoter siRNA is partially reduced in *ros1zop1* as well as in *ros1nrpe1* (Figure 2A).

Moreover, we found that the *zop1* mutation markedly reduces the accumulation of the Pol IV- and Pol V-dependent siRNAs, including *AtSN1* siRNA and siRNA1003 (Figure 2A; Supplementary Table S1). Cluster4 siRNA and siRNA02 were previously recognized as Pol IV-dependent but Pol V-independent siRNAs (Mosher *et al*, 2008; Zheng *et al*, 2009), but our results indicate that both siRNAs are weakly reduced by *nrpe1* in *ros1nrpe1* (Figure 2A; Supplementary Table S1). Although Cluster4 siRNA and siRNA02 are partially dependent on Pol V, the dependence is much lower than that of previously characterized Pol IV- and Pol V-dependent siRNAs including *AtSN1* siRNA and siRNA1003. The reduction of Cluster4 siRNA and siRNA02 in *ros1zop1* is similar to that in *ros1nrpe1* (Figure 2A; Supplementary Table S1). Interestingly, we found that the accumulation of *AtSN1* siRNA and Cluster4 siRNA seems to be weakly reduced by *ros1* (Figure 2A; Supplementary Table S1). This effect may be caused by the feedback effect of *ros1* on RdDM.

The effect of *zop1* on Pol IV-dependent siRNA accumulation was measured by small RNA deep sequencing. The results indicate that 201 658 Pol IV-dependent 24-nt siRNA reads are uniquely matched on the *Arabidopsis* nuclear genome (Figure 2B; Supplementary Table S2). These siRNAs are reduced to 18 654 reads (9.3%) and 115 604 reads (57.3%) in *ros1nrpd1* and *ros1nrpe1*, respectively.

Figure 2 Effect of *zop1* on small RNA accumulation and DNA methylation. (A) Effect of *zop1* on small RNA accumulation was determined by small RNA northern blotting. The accumulation of 24-nt Pol IV-dependent siRNAs, 21-nt ta-siRNA255, and miRNA171 was detected. The ethidium bromide-stained small RNA gel is shown as a loading control. (B) Effect of *zop1* on Pol IV-dependent siRNA accumulation as determined by small RNA deep sequencing. Numbers of uniquely matched Pol IV-dependent siRNA reads in *ros1*, *ros1nrpd1*, *ros1zop1*, and *ros1nrpe1*. siRNA abundance in each genotype is expressed as a percentage of siRNAs in *ros1*. (C, D) The DNA methylation of *AtSN1* and *MEA-ISR* was determined by bisulphite sequencing in the wild type, *ros1*, *ros1zop1*, and *ros1nrpe1*. CG, CHG, and CHH methylation in each ecotype is shown in the charts. (E) The diagrams indicate the numbers of genes and TEs that show reduced CHG methylation in *ros1nrpd1* and *ros1zop1* relative to *ros1*. The number of overlapping loci and specific loci is shown. (F) The numbers of loci that show reduced CHH methylation. Source data for this figure is available on the online supplementary information page.

The siRNAs in *ros1zop1* are reduced to 146555 reads (72.7%), which is comparable to that in *ros1nrpe1* (Figure 2B; Supplementary Table S2). Plotting of the Pol IV-dependent 24-nt siRNA distribution in *ros1*, *ros1zop1*, and *ros1nrpd1* (Supplementary Figure S3A–E) indicates that unlike *nrpd1*, *zop1* only partially affects Pol IV-dependent siRNA accumulation. The effect of *zop1* on siRNA accumulation is similar to that of *nrpe1*, indicating that the *zop1* mutation may influence a downstream step at the RdDM pathway.

The effect of *zop1* on the DNA methylation level of endogenous RdDM genomic targets was tested by bisulphite sequencing analysis. The results indicate that the DNA methylation of *AtSN1* A and *MEA-ISR* at CHG and CHH sites is substantially reduced in *ros1zop1* as well as in *ros1nrpe1*, whereas the DNA methylation at CG sites is not affected by either *zop1* or *nrpe1* (Figure 2C and D). The effect of *zop1* on the DNA methylation of *AtSN1* was further confirmed by chop-PCR. The results suggest that the *AtSN1* DNA methylation is partially reduced in *ros1zop1* compared to that in WT and *ros1*, but the reduction is less than that in *ros1nrpd1* and *ros1nrpe1* (Supplementary Figure S4). Moreover, the suppressive effect of *zop1* on DNA methylation was demonstrated by chop-PCR at one more RdDM target *IGN23* (Supplementary Figure S4).

To determine the genome-wide effect of *zop1* on DNA methylation, we performed whole-genome bisulphite sequencing in *ros1*, *ros1nrpd1*, and *ros1zop1*. The results show that *zop1* as well as *nrpd1* reduces DNA methylation especially at CHG and CHH sites at a whole-genome level (Supplementary Figure S5; Supplementary Table S3). The genome-wide effect of *nrpd1* on DNA methylation determined by this study is consistent with the previous study (Wierzbicki *et al*, 2012). The reduction of CHH methylation in *ros1zop1* is clearly less than that in *ros1nrpd1*, whereas the reduction of CHG methylation in *ros1zop1* is comparable to that in *ros1nrpd1* (Supplementary Table S3). The identified DNA methylation differences caused by *zop1* and *nrpd1* were measured by sequence-specific bisulphite sequencing analysis at three randomly selected loci (AT5G35540, AT1G54750, and AT1G14247) and the results indicated that our whole-genome bisulphite sequencing results are reliable (Supplementary Figure S6A–C).

In the *ros1nrpd1* mutant, there are 3522 genes and 3799 TEs that show reduced CHG methylation. Among them, 1211 genes (34.4%) and 1015 TEs (26.7%) also have reduced CHG methylation in *ros1zop1* (Figure 2E; Supplementary Tables S4 and S5). ZOP1 may act on these loci through an RdDM-dependent pathway. Moreover, there are 2672 genes and 1683 TEs whose CHG methylation is reduced in *ros1zop1* but not in

ros1nrpd1, suggesting that the CHG methylation at these loci may be established and maintained through an RdDM-independent pathway (Figure 2E; Supplementary Tables S4 and S5). There are 1807 genes and 5709 TEs that have reduced CHH methylation in *ros1nrpd1* relative to *ros1*, in which only 277 genes (15.3%) and 410 TEs (7.2%) have reduced CHH methylation in both *ros1nrpd1* and *ros1zop1* (Figure 2F; Supplementary Tables S4 and S5). There are 606 genes and 563 TEs that show reduced CHH methylation in *ros1zop1* but not in *ros1nrpd1*. Totally, the *zop1* mutation can reduce CHH methylation at 883 genes and 973 TEs, which are much less than 1807 genes and 5709 TEs that are affected by *nrpd1* (Figure 2F; Supplementary Tables S4 and S5). The result is consistent with the weak reduction of the overall CHH methylation level in *ros1zop1* (Supplementary Table S3). In the 5709 TEs that show NRPD1-dependent CHH methylation, only less than one tenth of these loci (410/5709) whose CHH methylation is markedly reduced by *zop1* (Figure 2F; Supplementary Tables S4 and S5), suggesting that the effect of *zop1* on CHH methylation is much less than *nrpd1*.

Like *nrpd1*, *zop1* reduces DNA methylation substantially in euchromatic regions and only slightly in centromeric regions (Supplementary Figure S5). At the promoter regions of protein-coding genes, the DNA methylation level is generally reduced at CHG and CHH sites by both *zop1* and *nrpd1*, although the reduction of CHH methylation caused by *zop1* is less than that caused by *nrpd1* (Supplementary Figure S7A). At TEs, the DNA methylation is reduced by *nrpd1* especially at CHH sites, whereas the reduction of DNA methylation caused by *zop1* is significant at CHG sites rather than at CHH sites (Supplementary Figure S7B). The results suggest that ZOP1 may act on DNA methylation through an uncharacterized mechanism that is different from the canonical RdDM pathway.

The transcriptional silencing at endogenous RdDM genome target loci is associated with DNA hypermethylation.

Our quantitative RT-PCR results indicate that the silencing of *AtSN1 A*, *AtGP1*, and *solo LTR* is substantially released in either *ros1nrpd1* or *ros1nrpe1* and is released to a much lesser extent in *ros1zop1* (Supplementary Figure S8A and B). Unlike the previously characterized RdDM mutants including *nrpd1* and *nrpe1*, *zop1* has no effect on *ROS1* expression (Supplementary Figure S8A). The release of silencing at *AtGP1* and *solo LTR* was also found in the *zop1* single mutants (Supplementary Figure S8B), suggesting that the effect of *zop1* on silencing is unrelated to the *ros1* mutation in *ros1zop1*. The Pol V-dependent RNA transcripts *AtSN1 B*, *IGN5 B*, and *IGN23* from RdDM target loci are blocked in *ros1nrpe1* but are not affected in *ros1zop1* (Supplementary Figure S8C), demonstrating that ZOP1 is not required for accumulation of Pol V-dependent RNA transcripts.

ZOP1 is a novel pre-mRNA splicing factor

ZOP1 is a ZnF and OCRE domain-containing protein. The OCRE domain-containing proteins are usually involved in RNA processing (Callebaut and Mornon, 2005). We performed RNA deep sequencing to determine the possible effect of *zop1* on RNA processing. From the *ros1* and *ros1zop1* RNA libraries, we obtained 25.5 million and 27.7 million RNA reads, respectively. Among them, 14.5 million and 13.5 million reads were mapped to the *Arabidopsis* genome. The intron-retention events were analysed in *ros1* and *ros1zop1*, and 361 intron-retention events in 215 genes were identified in *ros1zop1* but not in *ros1* ($P < 0.01$, intron read coverage $> 80\%$) (Supplementary Table S6). In contrast, only 31 intron-retention events in 27 genes were identified in *ros1* but not in *ros1zop1* (Supplementary Table S7). The intron-retention events in *ros1zop1* were confirmed by RT-PCR using intron-flanking primers (Figure 3A and B), suggesting that *zop1* may cause defects in pre-mRNA splicing. Moreover, we found that the defects can be complemented by

Figure 3 Pre-mRNA splicing is affected by *zop1*. (A) The splicing defects in *ros1zop1* at three indicated loci are shown. In the gene diagrams, bars and lines represent exons and introns, respectively. Filled bars are encoding regions, whereas blank bars are 5' UTR or 3' UTR. Normalized reads for the three genes are plotted for *ros1zop1* as well as for *ros1*. The intron-retention events in the three genes are labelled with frames. (B) The identified intron-retention events were confirmed by semiquantitative RT-PCR. The primers crossing introns were used to test the intron-retention events in *ros1zop1* as well as in the wild type and *ros1*. The amplification of an intron-containing fragment of *ACT7* as a control suggests the absence of DNA contamination. No RT represents amplification of *ACT7* from total RNA without reverse transcription. Source data for this figure is available on the online supplementary information page.

the *ZOP1* transgene in the *ros1zop1* mutant background (Figure 3B). The results confirmed that *zop1* causes defects in pre-mRNA splicing. Additionally, we found that the mature mRNAs are still present in *ros1zop1* even some mRNAs are probably mildly reduced (Figure 3B). Together, the results indicate that ZOP1 is a protein related to pre-mRNA splicing.

According to RNA deep sequencing in *ros1* and *ros1zop1*, the RNA transcripts of hundreds of genes were affected by *zop1* (Supplementary Table S8). A possible explanation for the involvement of ZOP1 in Pol IV-dependent siRNA accumulation and DNA methylation is that the splicing defects caused by *zop1* impair the accumulation of the mature mRNAs encoding typical RdDM components. However, our RNA deep sequencing result shows that no RdDM component-encoding gene is affected in *ros1zop1* (Supplementary Table S8). Furthermore, our quantitative RT-PCR results confirmed that the expression of major RdDM components is not reduced in *ros1zop1* (Supplementary Figure S9). Thus, the effect of *zop1* on RdDM is unlikely to be the result of indirect regulation of RdDM component-encoding genes. The splicing factor ZOP1 is likely to directly contribute to DNA methylation and transcriptional gene silencing.

To better understand the function of ZOP1, we carried out affinity purification of the ZOP1-Flag fusion protein from the *ZOP1-Flag* transgenic plants. The copurified proteins were analysed by mass spectrometry. Many peptides corresponding to components of spliceosome complexes and splicing factors were identified from the ZOP1-Flag copurified proteins (Supplementary Figure S10; Supplementary Table S9), suggesting the association between ZOP1 and these splicing proteins. In the list of ZOP1 copurified proteins, STA1 is the homologue of a yeast splicing factor PRP6 that interacts with both the U4/U6 and U5 snRNPs and facilitate the formation of U4/U6-U5 tri-snRNPs (Makarov *et al*, 2000). Our coimmunoprecipitation assay confirmed that ZOP1 can specifically associate with STA1 (Figure 4A), indicating that ZOP1 is likely to be a novel splicing factor. Subcellular localization of ZOP1 was determined by immunolocalization assay with the Flag antibody in the *ZOP1-Flag* transgenic plants. The result shows that ZOP1 is a nuclear protein (Supplementary Figure S11), which is consistent with its roles in pre-mRNA splicing as well as in DNA methylation and transcriptional silencing.

Because ZOP1 belongs to a member of OCRE domain-containing proteins that are usually involved in nucleic acid binding, we carried out electrophoretic mobility shift assays (EMSAs) to determine whether ZOP1 is capable of binding to nucleic acids. The results show that the recombinant His-tagged ZOP1 protein can bind double-stranded DNA as well as double-stranded RNA but not single-stranded DNA or RNA (Supplementary Figure S12A). ZOP1 can bind both blunt-end double-stranded RNA (bdsRNA) and 5' overhanging double-stranded RNA (odsRNA) (Supplementary Figure S12A). ZOP1 binds to double-stranded DNA or RNA in a dose-dependent manner, and the binding is reduced by competition with unlabelled double-stranded DNA or RNA (Supplementary Figure S12B and C), suggesting that the nucleic acid-binding ability of ZOP1 is specific. To identify the domain that is required for the nucleic acid binding of ZOP1, we generated a series of truncated ZOP1 proteins and used these proteins in EMSA (Supplementary Figure S12D). The results show that the previously uncharacterized C-terminal domain but not

the ZnF and OCRE domains is responsible for ZOP1 binding to both double-stranded DNA and double-stranded RNA (Supplementary Figures S12D and E). Positively charged amino acids are rich in the ZOP1 C-terminal domain (Supplementary Figure S2), which is consistent with its nucleic acid-binding ability.

The relationship between ZOP1 and other RdDM-related proteins

AGO4 protein levels are reduced in the RdDM mutants that are directly responsible for primary siRNA biogenesis, including *nrpd1*, *rdr2*, and *dcl3* (Li *et al*, 2006; Supplementary Figure S13). Our western blotting result indicates that neither *zop1* nor *nrpe1* affects AGO4 protein levels (Supplementary Figure S13), suggesting that like NRPE1, ZOP1 is not directly involved in primary siRNA biogenesis, but how ZOP1 is involved in DNA methylation remains to be elucidated. We used coimmunoprecipitation to determine whether ZOP1 interacts with the proteins that promote DNA methylation. The results suggest that ZOP1 cannot interact with the tested canonical RdDM components, including AGO4, NRPD1, NRPE1, DRM2, and RDM4 (Figure 4B; Supplementary Figure S14A–D). Coimmunoprecipitation indicates, however, that ZOP1 can interact *in vivo* with Pol II (Figure 4B). Furthermore, the interaction between ZOP1 and Pol II was tested in the presence of RNase and DNase. The result demonstrated that the interaction is insensitive to RNase and DNase (Supplementary Figure S15). Pol II was previously demonstrated to be involved in transcriptional silencing (Zheng *et al*, 2009), suggesting that the function of ZOP1 in transcriptional silencing is related to Pol II.

The localization pattern of ZOP1 in nuclei was investigated by immunofluorescence assay. Anti-Flag and anti-Myc antibodies did not produce any visible signals in the nuclei of wild-type plants (Supplementary Figure S16), but the anti-Flag antibody produced specific ZOP1 signals in the nuclei of *ZOP1-Flag* transgenic plants (Figure 4C). The results show that the ZOP1 is present in the form of either condensed nucleolus-adjacent foci or dispersed nucleoplasmic speckles (Figure 4C–F). The condensed nucleolus-adjacent foci of ZOP1 include the signals that overlap with U2B in the Cajal bodies (Figure 4C). The Cajal body, an snRNP assembly centre, was previously reported to be required for the assembly of the AGO4 effector complex of RdDM (Li *et al*, 2008). U2B, a component of U2 snRNP, was used as a marker for the Cajal body, and AGO4 colocalizes with U2B in the Cajal body (Li *et al*, 2006). Our results suggest that ZOP1 can colocalize with AGO4 at the Cajal body. We tested the relationship between ZOP1 and other canonical RdDM components by immunofluorescence assay. The *ZOP1-Myc* transgenic plants were crossed to *NRPE1-Flag*, *DRM2-Flag*, and *NRPD1-Flag* transgenic plants, respectively. Anti-Flag and anti-Myc antibodies were used to detect the signals in the nuclei of their F1 plants. The results indicated that ZOP1 partially colocalizes with NRPE1 as well as with DRM2 but not with NRPD1 in the nucleolus-adjacent foci (Figure 4D–F). NRPE1 and DRM2 directly associate with chromatin at RdDM target loci. Partial colocalization between ZOP1 and the two RdDM components is consistent with the finding that ZOP1 shares a large number of chromatin targets with the RdDM pathway (Figure 2E; Supplementary Tables S4 and S5).

Figure 4 Detection of the interaction between ZOP1 and related proteins. (A) Detection of the interaction between ZOP1 and STA1 by coimmunoprecipitation assay. Total protein extracts from STA1-Flag transgenic plants were immunoprecipitated with anti-Flag antibody. Immunoprecipitated proteins were tested by western blotting. (B) Coimmunoprecipitation testing the interaction between ZOP1 and NRPB1 or AGO4. Total protein extracts from wild-type and *ZOP1-Flag* transgenic plants were immunoprecipitated with anti-Flag antibody, and the precipitated proteins were subjected to western blotting with the antibodies aFlag, aNRPB1, and aAGO4. (C) Immunolocalization of ZOP1-Flag and U2B in nuclei of *ZOP1-Flag* transgenic plants using anti-Flag and the anti-U2B antibodies. Nuclei were counterstained with DAPI. Colocalization between ZOP1-Flag and U2B generates yellow signals. The percentage of the nuclei with indicated patterns is shown. (D) Immunolocalization of ZOP1-Myc and NRPE1-Flag in nuclei of plants expressing both *ZOP1-Myc* and *NRPE1-Flag* transgenes. (E) Immunolocalization of ZOP1-Myc and DRM2-Flag in nuclei of plants expressing both *ZOP1-Myc* and *DRM2-Flag* transgenes. (F) Immunolocalization of ZOP1-Myc and NRPD1-Flag in nuclei. Source data for this figure is available on the online supplementary information page.

To investigate the genetic relationship between ZOP1 and known RdDM proteins, we crossed *ros1zop1* to *ros1nrpd1*, *ros1nrpe1*, and *ros1ago4*, and generated *ros1zop1nrpd1*, *ros1zop1nrpe1*, and *ros1zop1ago4*. The luminescence analysis shows that the combination of *zop1* with *nrpd1*, *nrpe1*, or *ago4* in the *ros1* background can enhance the expression of the *RD29A-LUC* transgene (Figure 5A). Moreover, we found that *zop1* can also enhance the expression of endogenous *RD29A* in *ros1zop1nrpd1*, *ros1zop1nrpe1*, and *ros1zop1ago4* (Figure 5B). These results suggest that *zop1* has an additive

effect with the RdDM mutants *nrpd1*, *nrpe1*, and *ago4*. Involvement of ZOP1 in transcriptional silencing is at least partially independent of the previously characterized RdDM pathway.

To evaluate the function of ZOP1 on the silencing of endogenous RdDM targets, we measured the transcript levels of endogenous RdDM target loci in *ros1zop1*, *ros1nrpd1*, *ros1nrpe1*, and *ros1ago4* as well as in *ros1nrpd1zop1*, *ros1nrpe1zop1*, and *ros1ago4zop1*. The results show that the *zop1* mutation induces the transcript level of a typical RdDM target

Figure 5 Effect of the *zop1* mutation on transcriptional silencing in the *ros1nrpd1*, *ros1nrpe1*, and *ros1ago4* backgrounds. (A) The expression of the *RD29A-LUC* transgene was evaluated by luminescence imaging in the indicated genotypes at the top panel. The averages of three independent experiments with standard deviations are shown at the bottom panel. (B) Relative expression levels of endogenous *RD29A*, *SDC*, and *AtGP1* were determined by quantitative RT-PCR.

SDC in the *ros1nrpd1* and *ros1ago4* backgrounds, whereas the *zop1* mutation has no effect on *SDC* in the *ros1nrpe1* background (Figure 5B). Moreover, the *zop1* mutation reduces the transcript levels of *AtGP1* in the *ros1nrpd1*, *ros1nrpe1*, and *ros1ago4* backgrounds (Figure 5B). Combination of the *zop1* mutation with the RdDM mutants is likely to complicate the transcriptional regulation of RdDM target loci. But we have now known that ZOP1 is likely to act in transcriptional silencing at least partially through an RdDM-independent pathway.

Other development-defective splicing mutants affect RdDM

To investigate the possible role of other splicing-related proteins in RdDM and transcriptional silencing, we surveyed four other splicing mutants (*mac3a3b*, *mos4*, *mos12*, and *mos14*) that show defects in both pre-mRNA splicing and development (Monaghan *et al*, 2009; Xu *et al*, 2011, 2012). MAC3A and MAC3B are the homologues of the conserved yeast PRP19 splicing factor, and MOS4 is a component of the PRP19 complex, an evolutionarily conserved spliceosome-associated complex (Monaghan *et al*, 2009). MOS12 associates with the MOS4 and is required for proper pre-mRNA splicing (Xu *et al*, 2012). MOS14 is a nuclear-import receptor for serine/arginine-rich splicing factors (Xu *et al*, 2011). We first tested whether the splicing sites affected by *zop1* are also differentially spliced in these four splicing mutants. The results show that in the five indicated *zop1*-affected splicing sites, two of them in AT2G38025 and

AT3G04400 are affected in *mac3a3b*, *mos4*, *mos12*, but not in *mos14*, whereas the other three splicing sites in AT1G67090, AT1G23310, and AT3G24503 are correctly spliced in all the four tested splicing mutants (Supplementary Figure S17). The splicing defects of *SNC1* and *RPS4* caused by *mos14* are not present in *ros1zop1* (Supplementary Figure S17). Affected splicing sites in the four splicing mutants (*mac3a3b*, *mos4*, *mos12*, and *mos14*) are much different from that in *zop1*.

By small RNA northern blotting, we found that in the four the tested splicing mutants (*mac3a3b*, *mos4*, *mos12*, and *mos14*), the Pol IV-dependent 24-nt siRNAs, including *solo* LTR siRNA and siRNA1003, are reduced (Figure 6A). The results demonstrate that like ZOP1, these splicing-related proteins are required for siRNA accumulation. The reduction of siRNAs in *mac3a3b* is comparable to that in *nrpe1* and is greater than that in *mos4*, *mos12*, or *mos14* (Figure 6A). The *mos4* mutant has the least effect on siRNA accumulation, whereas the effects of *mos12* and *mos14* are intermediate (Figure 6A). Together, the four splicing proteins contribute to accumulation of Pol IV-dependent siRNAs. Interestingly, we found that both miRNA171 and ta-siRNA255 are reduced in *mac3a3b*, *mos4*, *mos12*, and *mos14*, whereas they are not affected in *nrpd1* and *nrpe1* as expected (Figure 6A). These splicing proteins are likely to be involved in biogenesis of miRNAs and ta-siRNAs. It will be interesting to determine how miRNAs and ta-siRNAs are affected in these splicing mutants in future.

Bisulphite sequencing indicated that in the four splicing mutants (*mac3a3b*, *mos4*, *mos12*, and *mos14*), the DNA

Figure 6 The splicing mutants *mac3a3b*, *mos4*, *mos12*, and *mos14* affect siRNA accumulation, DNA methylation, and transcriptional gene silencing. (A) Effects of the indicated splicing mutants on accumulation of 24-nt siRNAs, 21-nt microRNA171, ta-siRNA255, and Met-tRNA were determined by small RNA northern blotting. Met-tRNA signals are shown as a loading control. The quantitative results of small RNAs are indicated. (B) Effects of the splicing mutants on DNA methylation of *AtSN1* and the newly identified RdDM target *AT5G35540* were detected by bisulphite sequencing. (C) Effects of the splicing mutants on transcriptional silencing of *solo LTR* and *AtGP1* were determined by real-time RT-PCR. Source data for this figure is available on the online supplementary information page.

methylation levels at the CHH sites of *AtSN1* are weakly reduced relative to those in wild type, and the reduction is less than that in *nrpd1* (Figure 6B). Moreover, the reduction of DNA methylation in these mutants was also confirmed at the newly identified RdDM target *AT5G35540* (Figure 6B). The results suggest that not only ZOP1 but also other splicing proteins including MAC3A, MAC3B, MOS4, MOS12, and MOS14 are involved in regulation of DNA methylation.

Furthermore, quantitative RT-PCR showed that *mac3a3b*, *mos4*, and *mos14* partially induce the transcript levels of both *solo LTR* and *AtGP1*, whereas *mos12* induces the transcript level of *solo LTR* but has no effect on *AtGP1* (Figure 6C). Generally, all four of the tested splicing mutants partially release the silencing of RdDM targets, although the effect is less than that in the RdDM mutants *nrpd1* and *nrpe1* (Figure 6C). The weak effect of these splicing mutants on the silencing of RdDM targets is comparable to that of *zop1*. Together, these results suggest that the fully functional splicing machinery is probably involved in Pol IV-dependent siRNA accumulation, DNA methylation, and transcriptional silencing in *Arabidopsis*.

We carried out quantitative RT-PCR to test whether the transcript levels of major RdDM component-encoding genes are reduced in the splicing mutants *mac3a3b*, *mos4*, *mos12*, and *mos14*. The results indicated that *DCL3* is weakly

reduced in all the four splicing mutants, whereas *NRPE1* and *RDM1* are weakly reduced in *mos12* and *mos14* (Supplementary Figure S18). The other RdDM component-encoding genes are generally not reduced in the four splicing mutants (Supplementary Figure S18). Further studies are required to determine whether the effect of these splicing mutants on RdDM is completely due to the indirect effect of the mutants on the transcript levels of RdDM component-encoding genes.

Discussion

We have identified a novel plant-specific ZnF and OCRE domain-containing protein, ZOP1. The protein contains an OCRE domain that is usually related to RNA processing (Callebaut and Morion, 2005). By affinity purification of ZOP1, we copurified many splicing-related proteins (Supplementary Figure S10; Supplementary Table S9), suggesting a possible role of ZOP1 in pre-mRNA splicing. RNA deep sequencing results indicated that mutation of *ZOP1* affects the pre-mRNA splicing of hundreds of protein-coding genes (Supplementary Table S6). Coimmunoprecipitation assay demonstrated that ZOP1 can associate with the PRP6-like protein STA1 (Figure 4A). PRP6 is an snRNP-associated protein that facilitates the assembly of U4/U6-U5 tri-snRNPs

(Makarov *et al*, 2000). The association between ZOP1 and STA1 suggests that ZOP1 may act during the assembly of U4/U6-U5 tri-snRNPs. The core snRNPs are preliminarily assembled in the cytoplasm and then transported to the nuclear Cajal body, in which many splicing-related proteins are assembled on the core snRNPs (Cioce and Lamond, 2005; Morris, 2008). Our immunolocalization assay revealed that the ZOP1 signals overlap with the Cajal body (Figure 4C), which is consistent with the view that ZOP1 functions in snRNP assembly. In addition, ZOP1 signals also appear in the nucleoplasm, suggesting that ZOP1 may continue to associate with snRNPs after snRNPs move out of the Cajal body and may form mature spliceosome complexes on pre-mRNA targets. The detailed function of ZOP1 in pre-mRNA splicing remains to be studied.

Ausin *et al* (2012) recently reported that the previously characterized splicing factor SR45 affects the RdDM pathway in *Arabidopsis*, but they did not determine whether SR45 affects RdDM directly or indirectly through the splicing of RdDM factor genes. In our study, we demonstrated that the novel pre-mRNA splicing factor ZOP1 is involved in both pre-mRNA splicing and RdDM in *Arabidopsis*. To test whether ZOP1 is involved in regulation of RdDM factor genes, we compared gene expression in *ros1* and *ros1zop1* by RNA deep sequencing. The result indicated that no RdDM factor-encoding gene is markedly reduced by *zop1* (Supplementary Table S8). Our quantitative RT-PCR results confirmed that the transcript levels of all major RdDM factor-encoding genes in *ros1zop1* are not reduced compared to those in *ros1* (Supplementary Figure S9). Furthermore, the *zop1* mutation can even repress the transcript levels of a subset of RdDM target loci in the presence of the RdDM mutations, *nprp1*, *nrpe1*, and *ago4* (Figure 5B). These results suggest that ZOP1 is directly involved in RdDM regulation and is different from canonical RdDM regulators. In addition to ZOP1, five other tested splicing proteins MAC3A, MAC3B, MOS4, MOS12, and MOS14 are also involved in RdDM and transcriptional silencing (Figure 6). It is possible that the splicing machinery rather than specific splicing factors is involved in promoting RdDM and transcriptional silencing.

The two RNA-binding proteins FPA and FCA are involved in RNA 3' processing and FLC chromatin silencing (Horniyk *et al*, 2010; Liu *et al*, 2010). FPA and FCA can also promote *de novo* DNA methylation and chromatin silencing at some specific RdDM target loci (Baurle *et al*, 2007). Depression of the RdDM target *AtSN1* in the *fpa* mutant is caused by defective RNA 3' end formation at an upstream RNA polymerase II-dependent gene (Horniyk *et al*, 2010), suggesting that FPA-mediated RNA 3' processing contributes to the silencing of *AtSN1*. Our study indicates that RNA splicing factors promote DNA methylation and chromatin silencing at some RdDM target loci. It is possible that many RNA processing proteins are shared by both pre-mRNAs and non-coding RNAs. Although the detailed role of non-coding RNA processing in chromatin silencing remains to be elucidated, our results have demonstrated that siRNAs and DNA methylation are all affected in the splicing mutants, suggesting that the proper status of non-coding RNAs may play an important role during the silencing of RdDM target loci.

Splicing factors were previously reported to be involved in RNA-induced chromatin silencing in fission yeast (Bayne

et al, 2008; Bernard *et al*, 2010; Chinen *et al*, 2010). The role of the splicing machinery in the *Arabidopsis* RdDM pathway resembles that of the splicing factors in RITS of fission yeast (Bayne *et al*, 2008), but the mechanism remains elusive. Recent work revealed that RNA processing activities that utilize both non-coding and coding RNAs are involved in histone H3K9 methylation and chromatin silencing in fission yeast (Reyes-Turcu *et al*, 2011; Zofall *et al*, 2012). Given the contribution of the splicing machinery to RdDM and chromatin silencing in *Arabidopsis*, the involvement of RNA processing in chromatin silencing is likely to be an evolutionarily conserved mechanism from fungi to plants. We propose that the splicing machinery, in addition to contributing to pre-mRNA splicing on protein-coding genes, may function in non-coding RNA processing on RdDM target loci, by which it facilitates recruitment of RdDM effector to chromatin and contributes to *de novo* DNA methylation and chromatin silencing at non-coding regions of the genome.

ZOP1 is a pre-mRNA splicing factor that can associate with several known splicing proteins (Supplementary Table S9; Figure 4A). Although no association between ZOP1 and canonical RdDM components was found (Figure 4B; Supplementary Figure S14), the association between ZOP1 and Pol II was demonstrated by coimmunoprecipitation (Figure 4B). Previous reports suggested that Pol II is required for recruitment of AGO4 to RdDM target loci on chromatin (Zheng *et al*, 2009). The coupling of Pol II and the splicing machinery by ZOP1 may facilitate recruitment of AGO4 to RdDM target loci on chromatin. The Cajal body is required for assembling AGO4 effector and snRNPs (Li *et al*, 2006; Pikaard, 2006), and ZOP1 may be involved in the assembly of snRNPs and AGO4 effector in the Cajal body. Our EMSA results indicated that ZOP1 is capable of binding to double-stranded RNAs (Supplementary Figure S12). Double-stranded RNAs are present in the second structures of snRNAs and the base pairing between snRNPs and pre-mRNAs (Matlin and Moore, 2007). The siRNAs bound by AGO4 can form double-stranded RNAs with Pol V-dependent scaffold non-coding RNAs (Wierzbicki *et al*, 2008). Thus, the double-stranded RNA-binding ability of ZOP1 may be helpful in both pre-mRNA splicing and RdDM. The double-stranded DNA binding ability of ZOP1 described in this study suggests that ZOP1 may directly associate with chromatin *in vivo*. ZOP1 partially colocalizes with NRPE1 and DRM2 at nucleoplasmic speckles, indicating that ZOP1 is likely to act on the chromatin regions that are occupied by NRPE1 and DRM2. However, further studies are required to clarify the details of ZOP1 function *in vivo*.

Co-transcriptional splicing and Pol II transcription elongation are coordinated by the cooperation of the splicing machinery and Pol II (Shukla *et al*, 2011). A recent finding indicated that inclusion of introns in transgenes can increase the transcript levels of the transgenes in *Arabidopsis* (Christie *et al*, 2011), which is consistent with the possible role of the splicing machinery in the regulation of RNA transcript levels. The transcript level of the RdDM target *AtGP1* was partially derepressed by the *zop1* mutation in *ros1zop1* (Supplementary Figure S8A and B). When the *zop1* mutation was combined with RdDM mutations in *ros1zop1nprp1*, *ros1zop1nrpe1*, and *ros1zop1ago4*; however, the transcript level of *AtGP1* was repressed by the *zop1* mutation (Figure 5B). The results suggest that ZOP1 may have two

inverse functions at RdDM targets. One is to contribute to the silencing of RdDM targets, and the other is to promote RNA transcript levels at RdDM targets in association with Pol II (and probably Pol III). The two inverse functions of ZOP1 may partially explain the weak effect of *zop1* on silencing of RdDM targets in *ros1zop1* compared to that in *ros1nrpd1*, *ros1nrpe1*, and *ros1ago4* (Supplementary Figure S8A and B).

In conclusion, this study has found a novel function of the splicing machinery in *de novo* DNA methylation. Further studies are needed to identify additional splicing-related proteins that are required for DNA methylation and to elucidate how these proteins are involved in DNA methylation and transcriptional silencing.

Materials and methods

Plant materials, mutant screening, and cloning

Both wild-type C24 and *ros1* mutant plants carry a homozygous *RD29A* promoter-driven luciferase transgene and a 35S promoter-driven *NPTII* transgene. The *ros1* mutant plants with the two transgenes were mutagenized with ethyl methanesulphonate and screened for suppressors of *ros1*. Plants that emitted a high luminescence were crossed to the *ros1* mutant in the Col-0 background (Salk_045303). The selfed F2 plants were used for map-based cloning of the mutant. Deep sequencing (Illumina) was carried out to detect the mutation in the localized genome interval of the mutant. The *ZOP1* genomic sequence was cloned into the modified plant expression vector pCambia1305 and transformed into *ros1zop1* for a complementation test.

Analysis of RNA transcripts and small RNA accumulation

Total RNA was extracted from the indicated plant materials using Trizol reagent (Invitrogen). After contaminating DNA was removed by DNase, total RNA was used for semiquantitative RT-PCR and real-time RT-PCR. The oligo-dT or sequence-specific reverse primers were used for reverse transcription. The single-stranded cDNA was amplified by Ex-Taq DNA polymerase (Takara) for semiquantitative RT-PCR and quantitative RT-PCR. For quantitative RT-PCR, the results shown were based on at least three replications. Small RNA was extracted as previously described and separated on a 15% polyacrylamide gel at 200 V for 4 h. The separated small RNA was transferred onto hybond-N+ membranes (Amersham) for small RNA hybridization. The probes of DNA oligonucleotides and PCR products were radiation labelled by [γ - 32 P]ATP and [α - 32 P]dCTP, respectively. Small RNA hybridization was carried out in PerfectHyb buffer (Sigma) overnight at 38°C. The DNA oligonucleotides that were used are listed in Supplementary Table S10.

DNA methylation assay

Genomic DNA of indicated samples was extracted by CTAB and purified by phenol:chloroform (1:1). The DNA in the supernatant was precipitated by ethanol. DNA methylation was tested by bisulphite sequencing and chop-PCR. For bisulphite sequencing, 2 μ g of genomic DNA was sodium bisulphite-converted and purified using the EpiTect Bisulfite Kit (Qiagen). The purified DNA was amplified and cloned for bisulphite sequencing. For each ecotype, at least 15 samples were sequenced. For chop-PCR, genomic DNA was digested by DNA methylation-sensitive restriction enzymes (*HaeIII*), followed by amplification of the digested DNA at indicated DNA loci. The DNA oligonucleotides used for DNA methylation assay are listed in Supplementary Table S10.

Electrophoretic mobility shift assay

The full length of *ZOP1* and its truncated forms were cloned in-frame to the N-terminal His tag in the pET28a vector, and the constructs were transformed into *E. coli* strain BL21 (Invitrogen) for expression. The bacterially expressed His fusion proteins were purified by Ni-NTA His Bind Resin (Novagen) and used for EMSA. EMSA was carried out as described previously with minor modifications (He *et al*, 2009a). The complemented single-stranded oligonucleotides were annealed to generate double-stranded DNA and RNA oligonucleotides. All probes were end labelled by

[γ -32P]ATP with T4 polynucleotide kinase. The labelled probes were purified with G-25 columns (GE Healthcare). The binding assay was carried out in buffer containing 25 mM HEPES (pH 7.6), 50 mM KCl, 0.1 mM EDTA (pH 8.0), 12.5 mM MgCl₂, 1 mM DTT, 0.5% (w/v) BSA, and 5% (w/v) glycerol. The binding reaction was incubated at 25°C for 30 min. The reaction mixtures were separated on 4% non-denaturing polyacrylamide gels at 200 V for 2 h, and the gels were exposed to X-ray film for analysis.

Immunolocalization

Protoplasts were isolated from young *Arabidopsis* leaves as described (Yoo *et al*, 2007), and nuclei were fixed in 4% formaldehyde and applied to slides. Immunostaining was performed as previously described (Onodera *et al*, 2005). After nuclei were blocked with 3% BSA in PBS, primary antibodies were incubated with the nuclei overnight at 4°C. Primary antibodies were diluted as follows: rabbit or mouse anti-Flag (1:200), rabbit or mouse anti-cMyc (1:200), and mouse anti-U2B (1:50, lifespan). Secondary anti-mouse TRITC (Invitrogen) and anti-rabbit FITC (Invitrogen) were used at 1:200 dilutions. Chromatin was counterstained with DAPI in mounting medium. Images were acquired by SPINNING DISK confocal microscopy, analysed with Velocity software, and processed with Adobe Photoshop (Adobe Systems).

Analysis of RNA deep sequencing results

Total RNA was extracted from 1-month-old seedlings of *ros1* and *ros1zop1*. The mRNA was purified for RNA-seq library construction and whole transcriptome analysis. The *Arabidopsis* genome sequences and annotated gene models were downloaded from TAIR10 (www.arabidopsis.org). Tophat v1.3.1 was used to align the raw reads to genome sequences. Cufflinks (v1.1.0 <http://cufflinks.cbcb.umd.edu/>) was performed to assemble transcripts and calculate transcript abundances. Differences in RNA transcript levels between *ros1* and *ros1zop1* were identified by using the Cuffdiff. The presence of more than three FPKM in total exons of a gene was the criterion for gene detection. The cutoff of differential expression depended on the *P*-value and *ln* (FC) (*ln* of Fold Change). For intron-retention analysis, CoverageBed was performed to find reads located in intron regions, and Exact Testes in R was used to identify reliable expressed introns. The introns with >80% read coverage and *P*-values <0.01 were regarded as intron-retention events. This method was also used for the analysis of intron deletion events in *ros1zop1*.

Data processing and analysis of whole-genome bisulphite sequencing results

Raw *Arabidopsis* sequence data provided by Illumina were mapped to the modified Tair10 reference genome using Bismark (Krueger and Andrews, 2011). According to our previously sequenced *Arabidopsis* C24 genome, ~461 666 SNP sites were identified compared to Tair10 genome. Because *ros1*, *ros1zop1*, and *ros1nrpd1* are in the C24 background, Tair10 genome was modified according to the identified SNPs. Sequences that mapped to more than one position were removed to retain only reads that mapped uniquely. The methylation level of each cytosine site was represented by the percentage of the number of reads reporting a C relative to the total number of reads reporting a C or T. Only sites with at least five-fold coverage were included in the results. Gene annotations were downloaded from The *Arabidopsis* Information Resource (TAIR). The DNA methylation of annotated protein-coding genes and TEs was calculated. DNA methylation at CG, CHG, and CHH sites was separately evaluated.

Analysis of small RNA deep sequencing

Small RNA raw data were processed and mapped to the modified *Arabidopsis* TAIR10 reference genome. Only perfectly matched 24-nt siRNAs were extracted for analysis in *ros1*, *ros1nrpd1*, *ros1zop1*, and *ros1nrpe1*. The small RNAs in 100-nt-long windows were counted along the chromosomes, and the counts were adjusted according to the total library size for the comparison and graphic presentation. When the number of combined 24-nt small RNA reads in 100-bp windows in WT is significantly higher than that in *nrpd1*, the siRNAs in the windows are defined as Pol IV-dependent siRNAs.

Affinity purification, mass spectrometry, and coimmunoprecipitation

A 3-g quantity of flowers from *ZOP1-Flag* transgenic plants and wild-type control plants were used to prepare protein extracts as described previously (Pontes *et al*, 2006). In brief, the flowers were ground in liquid nitrogen and homogenized in 15 ml of protein extraction buffer (50 mM Tris (pH 7.6), 150 mM NaCl, 5 mM MgCl₂, 10% glycerol, 0.1% NP-40, 0.5 mM DTT, 1 mM PMSF, and one protease inhibitor cocktail tablet/50 ml (Roche)). Cell debris was removed by three centrifugations. The supernatant of each sample was incubated with Anti-Flag M1 agarose (Sigma, A4596) at 4°C for 2.5 h and washed five times with the protein extraction buffer. The affinity-purified proteins were eluted from the agarose with 3 × Flag peptides (Sigma). The eluted proteins were run on a 12% SDS-PAGE gel and subjected to silver staining with the ProteoSilver Silver Stain Kit (Sigma, PROT-SIL1).

For mass-spectrometric analysis, proteins on SDS-PAGE gels were de-stained and digested in-gel with sequencing grade trypsin (10 ng/μl trypsin, 50 mM ammonium bicarbonate, pH 8.0) at 37°C overnight. The digested peptides were eluted on a capillary column and sprayed into an LTQ mass spectrometer equipped with a nano-ESI ion source (Thermo Fisher Scientific, USA). Identified peptides were searched in the IPI (International Protein Index) *Arabidopsis* protein database on the Mascot server (Matrix Science Ltd, UK).

For coimmunoprecipitation, the protein extracts were incubated with protein A agarose beads conjugated with the indicated antibody, washed five times, and boiled in SDS-PAGE sample buffer. The boiled extracts were run on a 12% SDS-PAGE gel for western blotting. For coimmunoprecipitation analysis between *ZOP1* and *NRPD1*, *NRPE1*, or *AGO4*, the *ZOP1-Flag* or *ZOP1-Myc* transgenic plants were crossed to *NRPD1-Flag*, *NRPE1-Flag*, and *Myc-AGO4* transgenic plants, respectively. The offspring plants expressing both fusion proteins were subjected to coimmunoprecipitation.

References

- Ausin I, Greenberg MV, Li CF, Jacobsen SE (2012) The splicing factor SR45 affects the RNA-directed DNA methylation pathway in *Arabidopsis*. *Epigenetics* **7**: 29–33
- Baurle I, Smith L, Baulcombe DC, Dean C (2007) Widespread role for the flowering-time regulators FCA and FPA in RNA-mediated chromatin silencing. *Science* **318**: 109–112
- Bayne EH, Portoso M, Kagansky A, Kos-Braun IC, Urano T, Ekwall K, Alves F, Rappsilber J, Allshire RC (2008) Splicing factors facilitate RNAi-directed silencing in fission yeast. *Science* **322**: 602–606
- Bernard P, Drogat J, Dheur S, Genier S, Javerzat JP (2010) Splicing factor Spf30 assists exosome-mediated gene silencing in fission yeast. *Mol Cell Biol* **30**: 1145–1157
- Buhler M, Verdel A, Moazed D (2006) Tethering RITS to a nascent transcript initiates RNAi- and heterochromatin-dependent gene silencing. *Cell* **125**: 873–886
- Callebaut I, Morion JP (2005) OCRE: a novel domain made of imperfect, aromatic-rich octamer repeats. *Bioinformatics* **21**: 699–702
- Cao X, Jacobsen SE (2002) Role of the *Arabidopsis* DRM methyltransferases in de novo DNA methylation and gene silencing. *Curr Biol* **12**: 1138–1144
- Chinen M, Morita M, Fukumura K, Tani T (2010) Involvement of the spliceosomal U4 small nuclear RNA in heterochromatic gene silencing at fission yeast centromeres. *J Biol Chem* **285**: 5630–5638
- Christie M, Croft LJ, Carroll BJ (2011) Intron splicing suppresses RNA silencing in *Arabidopsis*. *Plant J* **68**: 159–167
- Cioce M, Lamond AI (2005) Cajal bodies: a long history of discovery. *Annu Rev Cell Dev Biol* **21**: 105–131
- Gao Z, Liu HL, Daxinger L, Pontes O, He X, Qian W, Lin H, Xie M, Lorkovic ZJ, Zhang S, Miki D, Zhan X, Pontier D, Lagrange T, Jin H, Matzke AJ, Matzke M, Pikaard CS, Zhu JK (2010) An RNA polymerase II- and AGO4-associated protein acts in RNA-directed DNA methylation. *Nature* **465**: 106–109
- Gong Z, Morales-Ruiz T, Ariza RR, Roldán-Arjona T, David L, Zhu JK (2002) ROS1, a repressor of transcriptional gene silencing in *Arabidopsis*, encodes a DNA glycosylase/lyase. *Cell* **111**: 803–814

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (<http://www.embojournal.org>).

Acknowledgements

We thank Yuelin Zhang and Xin Li (Department of Botany, University of British Columbia) for the splicing mutants *mac3a3b*, *mos4*, *mos12*, and *mos14*, and She Chen (National Institute of Biological Sciences, Beijing, China) for the technique support of mass spectrometry. This work was supported by the National Basic Research Program of China (973 Program) (2012CB910900) and the 973 Program (2011CB812600) from the Chinese Ministry of Science and Technology.

Author contributions: Xin-Jian He designed experiments. Cui-Jun Zhang performed the experiments shown in Figures 1, 2A, 2C, 2D, 3, 4A, 4B, 6A, and 6C, and Supplementary Figures S4, S8–S10, S12–S15, S17, and S19. Jin-Xing Zhou performed the experiments shown in Figures 5 and 6B and Supplementary Figures S1 and S6. Jun Liu identified the *ZOP1* gene by map-based cloning. Ze-Yang Ma performed immunofluorescence experiments shown in Figures 4C–F and Supplementary Figures S11 and S16. Su-Wei Zhang and Kun Dou helped with luminescence assay and *in vivo* coimmunoprecipitation assay. Huan-Wei Huang, Tao Cai, and Renyi Liu performed statistic analysis on small RNA deep sequencing data and genome-wide DNA methylation data. Cui-Jun Zhang, Jian-Kang Zhu, and Xin-Jian He analysed results. Cui-Jun Zhang and Xin-Jian He wrote the manuscript. Xin-Jian He supervised the work.

Conflict of interest

The authors declare that they have no conflict of interest.

- Grewal SI (2010) RNAi-dependent formation of heterochromatin and its diverse functions. *Curr Opin Genet Dev* **20**: 134–141
- Hall IM, Shankaranarayana GD, Noma K, Ayoub N, Cohen A, Grewal SI (2002) Establishment and maintenance of a heterochromatin domain. *Science* **297**: 2232–2237
- He XJ, Hsu YF, Zhu S, Wierzbicki AT, Pontes O, Pikaard CS, Liu HL, Wang CS, Jin H, Zhu JK (2009a) An effector of RNA-directed DNA methylation in *Arabidopsis* is an ARGONAUTE 4- and RNA-binding protein. *Cell* **137**: 498–508
- He XJ, Hsu YF, Zhu S, Liu HL, Pontes O, Zhu J, Cui X, Wang CS, Zhu JK (2009b) A conserved transcriptional regulator is required for RNA-directed DNA methylation and plant development. *Genes Dev* **23**: 2717–2722
- He XJ, Hsu YF, Pontes O, Zhu J, Lu J, Bressan RA, Pikaard C, Wang CS, Zhu JK (2009c) NRPD4, a protein related to the RPB4 subunit of RNA polymerase II, is a component of RNA polymerases IV and V and is required for RNA-directed DNA methylation. *Genes Dev* **23**: 318–330
- Herr AJ, Jensen MB, Dalmay T, Baulcombe DC (2005) RNA polymerase IV directs silencing of endogenous DNA. *Science* **308**: 118–120
- Hornik C, Terzi LC, Simpson GG (2010) The spen family protein FPA controls alternative cleavage and polyadenylation of RNA. *Dev Cell* **18**: 203–213
- Kaida D, Berg MG, Younis I, Kasim M, Singh LN, Wan L, Dreyfuss G (2010) U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature* **468**: 664–668
- Kanno T, Bucher E, Daxinger L, Huettel B, Böhmendorfer G, Gregor W, Kreil DP, Matzke M, Matzke AJ (2008) A structural-maintenance-of-chromosomes hinge domain-containing protein is required for RNA-directed DNA methylation. *Nat Genet* **40**: 670–675
- Kanno T, Bucher E, Daxinger L, Huettel B, Kreil DP, Breinig F, Lind M, Schmitt MJ, Simon SA, Gurazada SG, Meyers BC, Lorkovic ZJ, Matzke AJ, Matzke M (2010) RNA-directed DNA methylation and plant development require an IWR1-type transcription factor. *EMBO Rep* **11**: 65–71

- Kanno T, Huettel B, Mette MF, Aufsatz W, Jaligot E, Daxinger L, Kreil DP, Matzke M, Matzke AJ (2005) Atypical RNA polymerase subunits required for RNA-directed DNA methylation. *Nat Genet* **37**: 761–765
- Kanno T, Mette MF, Kreil DP, Aufsatz W, Matzke M, Matzke AJ (2004) Involvement of putative SNF2 chromatin remodeling protein DRD1 in RNA-directed DNA methylation. *Curr Biol* **14**: 801–805
- Kim YJ, Zheng B, Yu Y, Won SY, Mo B, Chen X (2011) The role of Mediator in small and long noncoding RNA production in *Arabidopsis thaliana*. *EMBO J* **30**: 814–822
- Krueger F, Andrews SR (2011) Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**: 1571–1572
- Law JA, Ausin I, Johnson LM, Vashisht AA, Zhu JK, Wohlschlegel JA, Jacobsen SE (2010) A protein complex required for polymerase V transcripts and RNA-directed DNA methylation in *Arabidopsis*. *Curr Biol* **20**: 951–956
- Law JA, Jacobsen SE (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet* **11**: 204–220
- Law JA, Vashisht AA, Wohlschlegel JA, Jacobsen SE (2011) 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
- Li CF, Henderson IR, Song L, Fedoroff N, Lagrange T, Jacobsen SE (2008) Dynamic regulation of ARGONAUTE4 within multiple nuclear bodies in *Arabidopsis thaliana*. *PLoS Genet* **4**: e27
- Li CF, Pontes O, El-Shami M, Henderson IR, Bernatavichute YV, Chan SW, Lagrange T, Pikaard CS, Jacobsen SE (2006) An ARGONAUTE4-containing nuclear processing center colocalized with Cajal bodies in *Arabidopsis thaliana*. *Cell* **126**: 93–106
- Liu F, Marquardt S, Lister C, Swiezewski S, Dean C (2010) Targeted 3' processing of antisense transcripts triggers *Arabidopsis* FLC chromatin silencing. *Science* **327**: 94–97
- Liu J, Bai G, Zhang C, Chen W, Zhou J, Zhang S, Chen Q, Deng X, He XJ, Zhu JK (2011) An atypical component of RNA-directed DNA methylation machinery has both DNA methylation-dependent and -independent roles in locus-specific transcriptional gene silencing. *Cell Res* **21**: 1691–1700
- Makarov EM, Makarova OV, Achsel T, Luhrmann R (2000) The human homologue of the yeast splicing factor prp6p contains multiple TPR elements and is stably associated with the U5 snRNP via protein-protein interactions. *J Mol Biol* **298**: 567–575
- Matlin AJ, Moore MJ (2007) Spliceosome assembly and composition. *Adv Exp Med Biol* **623**: 14–35
- Matzke M, Kanno T, Daxinger L, Huettel B, Matzke AJ (2009) RNA-mediated chromatin-based silencing in plants. *Curr Opin Cell Biol* **21**: 367–376
- Moazed D (2009) Small RNAs in transcriptional gene silencing and genome defence. *Nature* **457**: 413–420
- Monaghan J, Xu F, Gao M, Zhao Q, Palma K, Long C, Chen S, Zhang Y, Li X (2009) Two Prp19-like U-box proteins in the MOS4-associated complex play redundant roles in plant innate immunity. *PLoS Pathog* **5**: e1000526
- Morris GE (2008) The Cajal body. *Biochim Biophys Acta* **1783**: 2108–2115
- Mosher RA, Schwach F, Studholme D, Baulcombe DC (2008) PolIVb influences RNA-directed DNA methylation independently of its role in siRNA biogenesis. *Proc Natl Acad Sci USA* **105**: 3145–3150
- Motamedi MR, Verdel A, Colmenares SU, Gerber SA, Gygi SP, Moazed D (2004) Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. *Cell* **119**: 789–802
- Onodera Y, Haag JR, Ream T, Costa Nunes P, Pontes O, Pikaard CS (2005) Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell* **120**: 613–622
- Pikaard CS (2006) Cell biology of the *Arabidopsis* nuclear siRNA pathway for RNA-directed chromatin modification. *Cold Spring Harb Symp Quant Biol* **71**: 473–480
- Pontes O, Li CF, Costa Nunes P, Haag J, Ream T, Vitins A, Jacobsen SE, Pikaard CS (2006) The *Arabidopsis* chromatin-modifying nuclear siRNA pathway involves a nucleolar RNA processing center. *Cell* **126**: 79–92
- Pontier D, Yahubyan G, Vega D, Bulski A, Saez-Vasquez J, Hakimi MA, Lerbs-Mache S, Colot V, Lagrange T (2005) Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases IV in *Arabidopsis*. *Genes Dev* **19**: 2030–2040
- Qi Y, He X, Wang XJ, Kohany O, Jurka J, Hannon GJ (2006) Distinct catalytic and non-catalytic roles of ARGONAUTE4 in RNA-directed DNA methylation. *Nature* **443**: 1008–1012
- Ream TS, Haag JR, Wierzbicki AT, Nicora CD, Norbeck AD, Zhu JK, Hagen G, Guilfoyle TJ, Pasa-Tolić L, Pikaard CS (2009) Subunit compositions of the RNA-silencing enzymes Pol IV and Pol V reveal their origins as specialized forms of RNA polymerase II. *Mol Cell* **33**: 192–203
- Reyes-Turcu FE, Zhang K, Zofall M, Chen E, Grewal SI (2011) Defects in RNA quality control factors reveal RNAi-independent nucleation of heterochromatin. *Nat Struct Mol Biol* **18**: 1132–1138
- Sharp PA (1994) Split genes and RNA splicing. *Cell* **77**: 805–815
- Shukla S, Kavak E, Gregory M, Imashimizu M, Shutinoski B, Kashlev M, Oberdoerffer P, Sandberg R, Oberdoerffer S (2011) CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature* **479**: 74–79
- Slotkin RK, Martienssen R (2007) Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet* **8**: 272–285
- Smith LM, Pontes O, Searle I, Yelina N, Yousafzai FK, Herr AJ, Pikaard CS, Baulcombe DC (2007) An SNF2 protein associated with nuclear RNA silencing and the spread of a silencing signal between cells in *Arabidopsis*. *Plant Cell* **19**: 1507–1521
- Tang W, Kannan R, Blanchette M, Baumann P (2012) Telomerase RNA biogenesis involves sequential binding by Sm and Lsm complexes. *Nature* **484**: 260–264
- Verdel A, Jia S, Gerber S, Sugiyama T, Gygi S, Grewal SI, Moazed D (2004) RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* **303**: 672–676
- Volpe TA, Kidner C, Hall IM, Teng G, Grewal SI, Martienssen RA (2002) Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297**: 1833–1837
- Wierzbicki AT, Haag JR, Pikaard CS (2008) Non-coding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. *Cell* **135**: 635–648
- Wierzbicki AT, Ream TS, Haag JR, Pikaard CS (2009) RNA polymerase V transcription guides ARGONAUTE4 to chromatin. *Nat Genet* **41**: 630–634
- Wierzbicki AT, Cocklin R, Mayampurath A, Lister R, Rowley MJ, Gregory BD, Ecker JR, Tang H, Pikaard CS (2012) Spatial and functional relationships among Pol V-associated loci, Pol IV-dependent siRNAs, and cytosine methylation in the *Arabidopsis* epigenome. *Genes Dev* **26**: 1825–1836
- Xie Z, Johansen LK, Gustafson AM, Kasschau KD, Lellis AD, Zilberman D, Jacobsen SE, Carrington JC (2004) Genetic and functional diversification of small RNA pathways in plants. *PLoS Biol* **2**: e104
- Xu F, Xu S, Wiermer M, Zhang Y, Li X (2012) The cyclin L homolog MOS12 and the MOS4-associated complex are required for the proper splicing of plant resistance genes. *Plant J* **70**: 916–928
- Xu S, Zhang Z, Jing B, Gannon P, Ding J, Xu F, Li X, Zhang Y (2011) Transportin-SR is required for proper splicing of resistance genes and plant immunity. *PLoS Genet* **7**: e1002159
- Yoo SD, Cho YH, Sheen J (2007) *Arabidopsis* mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nat Protoc* **2**: 1565–1572
- Zheng B, Wang Z, Li S, Yu B, Liu JY, Chen X (2009) Intergenic transcription by RNA polymerase II coordinates Pol IV and Pol V in siRNA-directed transcriptional gene silencing in *Arabidopsis*. *Genes Dev* **23**: 2850–2860
- Zhou Z, Licklider LJ, Gygi SP, Reed R (2002) Comprehensive proteomic analysis of the human spliceosome. *Nature* **419**: 182–185
- Zofall M, Yamanaka S, Reyes-Turcu FE, Zhang K, Rubin C, Grewal SI (2012) RNA elimination machinery targeting meiotic mRNAs promotes facultative heterochromatin formation. *Science* **335**: 96–100