Synergistic action of the Arabidopsis spliceosome components PRP39a and SmD1b in promoting posttranscriptional transgene silencing

Jérémie Bazin, Emilie Elvira-Matelot, Thomas Blein, Vincent Jauvion, Nathalie Bouteiller, Jun Cao, Martin Crespi, Hervé Vaucheret

To cite this version:

Jérémie Bazin, Emilie Elvira-Matelot, Thomas Blein, Vincent Jauvion, Nathalie Bouteiller, et al.. Synergistic action of the Arabidopsis spliceosome components PRP39a and SmD1b in promoting posttranscriptional transgene silencing. The Plant cell, 2023, 35 (6), pp.1917-1935. 10.1093/plcell/koad091 . hal-04298436

HAL Id: hal-04298436

https://hal.inrae.fr/hal-04298436v1

Submitted on 22 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Synergistic action of the *Arabidopsis* spliceosome components PRP39a and SmD1b in promoting post-transcriptional transgene silencing

Jérémie Bazin[1], Emilie Elvira-Matelot[2], Thomas Blein[1], Vincent Jauvion[2], Nathalie Bouteiller[2], Jun Cao[3,#], Martin D. Crespi[1] and Hervé Vaucheret[2]

[1] Institute of Plant Sciences Paris-Saclay (IPS2), CNRS, INRA, Universités Paris-Sud, Evry, Paris-Diderot, Sorbonne Paris-Cité, Paris-Saclay, Bâtiment 630, 91405 0rsay, France
[2] Institut Jean-Pierre Bourgin, UMR1318, INRAE, AgroParisTech, Université Paris-Saclay, RD10, 78000 Versailles, France
[3] Department of Molecular Biology, Max Planck Institute for Developmental Biology, Spemannstrasse 35, D-72076 Tübingen, Germany

[#] Current address: Dow AgroSciences, Indianapolis, IN 46268, USA.

Corresponding authors: Hervé Vaucheret (herve.vaucheret@inrae.fr) and Martin Crespi (martin.crespi@ips2.université-paris-saclay.fr)

Material distribution: "The authors responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) are: Hervé Vaucheret (herve.vaucheret@inrae.fr) and Martin Crespi (martin.crespi@ips2.université-paris-saclay.fr)

Short title: Splicing factors PRP39a and SmD1b promote transgene PTGS

Synopsis

Arabidopsis PRP39a and SmD1b, two conserved components of the spliceosome, participate in splicing and synergistically promote RNA silencing by preventing aberrant RNA degradation in the nucleus.

Abstract

Besides regulating splicing, the conserved spliceosome component SmD1b was shown to promote posttranscriptional silencing of sense transgenes (S-PTGS). Here, we show that the conserved spliceosome component PRP39a also plays a role in S-PTGS. However, PRP39a and SmD1b actions appear distinct in both splicing and S-PTGS. Indeed, RNA-seq analysis of *prp39a* and *smd1b* mutants identified different sets of deregulated mRNAs and non-coding RNAs, both at expression level and alternative splicing genome-wide. Moreover, double mutant analyses involving *prp39a* or *smd1b* and RNA quality control (RQC) mutants revealed genetic interactions of SmD1b and PRP39a with distinct nuclear RQC machineries, suggesting synergistic rather than redundant roles in the RQC/PTGS interplay. Supporting this hypothesis, a *prp39a smd1b* double mutant exhibited enhanced suppression of S-PTGS compared with single mutants. Because no major changes in the expression of PTGS or RQC components or in small RNA production were identified in *prp39a* and *smd1b* mutants, and because *prp39a* and *smd1b* mutations do not alter PTGS triggered by inverted-repeat transgenes directly producing dsRNA (IR-PTGS), PRP39a and SmD1b seem to synergistically promote a step specific to S-PTGS. We propose that, independent of their specific roles in splicing, PRP39a and SmD1b limit 3’-to-5’ and/or 5’-to-3’ degradation of transgene aberrant RNAs in the nucleus, thus favoring the export of aberrant RNAs to the cytoplasm where their transformation into dsRNA initiates S-PTGS.

Introduction

RNA interference (RNAi) is a conserved mechanism that regulates endogenous gene expression and acts as a defense mechanism against invasive sequences of endogenous (transposons) or exogenous (virus, transgenes) origins. RNAi is initiated by double-stranded RNA (dsRNA) formed by the transcription of an inverted repeat, the copying of single-stranded RNA (ssRNA) by RNA-dependent RNA polymerases (RDRs), or convergent transcription on both strands of the DNA. These dsRNA are processed by DICER or DICER-like (DCL) proteins into microRNAs (miRNAs) or small interfering RNAs (siRNAs). These small RNAs are loaded onto ARGONAUTE (AGO) proteins to mediate sequence-specific gene silencing. Generally, miRNAs silence complementary mRNAs at the post-transcriptional level, while siRNAs can act at either transcriptional or post-transcriptional level to silence homologous endogenous genes, transgenes, transposons and viruses (Bartel, 2004; Baulcombe, 2004; Lippman and Martienssen, 2004).

In plants, post-transcriptional RNAi triggered by loci exhibiting an inverted-repeat (IR) structure is referred to as IR-PTGS. This process does not require RDR6 because dsRNAs are formed by the folding of self-complementary RNAs. In contrast, plant post-transcriptional RNAi triggered by loci that are not supposed to produce dsRNA is referred to as sense PTGS (S-PTGS). It is assumed that such loci produce a type of aberrant RNAs (abRNA) that are prone to be converted to dsRNA by RDR6. However, because RDR6 resides in cytoplasmic foci referred to as siRNA-bodies (Jouannet et al., 2012), abRNAs could only be transformed into dsRNA by RDR6 if they escape degradation by specialized RNA quality control (RQC) pathways that normally intercept and remove abRNAs in the nuclear and cytoplasmic compartments. Therefore, it is assumed that loci undergoing S-PTGS produce large amounts of abRNAs that somehow saturate the RQC machinery, thus allowing a sufficient amount of abRNAs to reach siRNA-bodies to initiate S-PTGS. Supporting this hypothesis, *Arabidopsis dpc1, dcp2, hen2, ski2, ski3, upf1, upf3, vcs, xrn3* and *xrn4* mutations, which compromise RQC at different steps, promote or enhance RDR6-dependent PTGS of both transgenes (Gazzani et al., 2004; Gy et al., 2007; Gregory et al., 2008; Thran et al., 2012; Moreno et al., 2013; Martinez de Alba et al., 2015; Yu et al., 2015) and endogenous genes (Martinez de Alba et al., 2015; Zhang et al., 2015).

Following the transformation of abRNAs into dsRNA by RDR6, DCL4 and DCL2 process dsRNAs into 21- and 22-nt siRNAs, respectively. These siRNAs form a complex with AGO1 and guide silencing of complementary mRNAs. Once S-PTGS is initiated, the biogenesis of 21- and 22-nt siRNAs usually extends along the entire length of the mRNA in a process known as transitivity (Vaistij et al., 2002; Mlotshwa et al., 2008; Parent et al., 2015). Furthermore, once induced in cells where RQC is impaired or saturated, S-PTGS moves systemically throughout the plant. The existence of systemic PTGS signals was first demonstrated by grafting non-silenced transgenic scions onto silenced transgenic rootstocks carrying the same transgene (Palauqui et al., 1997), and subsequent grafting experiments with mutant rootstocks and scions has identified several plant genes that are required for transmission of systemic PTGS from rootstocks and/or reception of systemic PTGS in scions (Brosnan et al., 2007; Taochy et al., 2017; Taochy et al., 2019). Nevertheless, full systemic S-PTGS has only been observed in the cases of transgenes, suggesting that transgenes differ from endogenous genes in a way that makes them particularly prone to trigger S-PTGS all over the plant. Therefore, it is likely that if RQC is impaired or saturated in one cell, the entry of abRNAs into siRNA-bodies and the triggering of S-PTGS remains confined to this cell and its surrounding cells in the case of endogenous genes, thus preventing the deleterious effects that could result from systemic S-PTGS of essential genes. Supporting this hypothesis, impairing RQC constitutively causes the death of plants, which can be suppressed by impairing RDR6 (Martinez de Alba et al., 2015; Zhang et al., 2015). In contrast, transgenes, which obviously are dispensable because they are not present naturally, generally are silenced systemically, suggesting that transgenes are treated like exogenous invaders such as viruses, which need to be systemically silenced all over the plant.

Genetic screens for S-PTGS-deficient mutants have been instrumental in identifying S-PTGS components, whereas genetic screens for S-PTGS-enhanced mutants have been instrumental in identifying RQC components that counteract S-PTGS. However, little is known about factors that counteract the action of RQC and favor S-PTGS. We previously reported that the conserved spliceosome component SmD1b, which localizes in the nucleoplasm, promotes transgene S-PTGS, even when triggered by intron-less transgenes, and has no effect on transgene IR-PTGS (Elvira-Matelot et al., 2016). Because S-PTGS is restored in double mutants impaired in SmD1b and the RQC factor XRN3, which also localizes in the nucleoplasm, we proposed that SmD1b protects transgene abRNAs from degradation by XRN3. Thus, beside its role in splicing, a spliceosome component plays a role in the interplay between RQC and S-PTGS in the nucleus (Elvira-Matelot et al., 2016). Here, we show that another conserved spliceosome component, PRP39a, promotes S-PTGS, even when triggered by intron-less transgenes, while it has no effect on IR-PTGS. S-PTGS is restored in *prp39a hen2* and *smd1b xrn3* double mutants, partially in *smd1b hen2* double mutants, but not in the *prp39a xrn3*, suggesting distinct interactions of PRP39a and SmD1b with the HEN2- and XRN3-dependent RQC pathways, respectively. Impairing simultaneously PRP39a and SmD1b further suppresses S-PTGS demonstrating a synergistic action of these proteins. These proteins also affect the alternative splicing of different pre-mRNA targets genome-wide. Hence, PRP39a and SMD1b define two nuclear pathways with dual roles in splicing and S-PTGS regulation.

Results

SGS15 encodes an ortholog of the yeast splicing factor PRP39

The Arabidopsis *2a3* line carries a transgene consisting of the *NIA2* gene under the control of the 35S promoter. This line undergoes S-PTGS of the *35S:NIA2* transgene and of the endogenous *NIA1* and *NIA2* genes in each generation, resulting in a chlorotic phenotype (Elmayan et al., 1998b). Fast-neutron mutagenesis followed by a screen for plants that remain green yield several S-PTGS-deficient mutants that have been described previously (Adenot et al., 2006; Elvira-Matelot et al., 2016). Here we describe two additional mutants identified during this screen. Based on allelic tests, these two mutants define the *SGS15* locus. In addition to PTGS-deficiency, the *sgs15-1 and sgs15-2* mutants exhibit a late flowering phenotype (Figure 1A). Whole-genome sequencing revealed a deletion in *sgs15-1*, removing part or all of six protein-coding genes (At1g04020, At1g040340, At1g04040, At1g04050, At1g04060, and At1g04080), while *sgs15-2* exhibits a 725 kb inversion on chromosome 1, which disrupts At1g01950 and At1g04080 (Figure 1B). At1g04080 is the only common gene mutated in both *sgs15-1 and sgs15-2,* suggesting that it is responsible for late flowering and S-PTGS defect. At1g04080 was previously described as PRP39a, one of the two Arabidopsis genes encoding a protein homologous to the yeast splicing factor PRP39 (Wang et al., 2007; Kanno et al., 2017). Homozygous *prp39a-1* plants, which harbor a T-DNA insertion in At1g04080 (SAIL_1249A03), exhibited a late flowering phenotype similar to *sgs15-1 and sgs15-2* (Wang et al., 2007; Kanno et al., 2017) and crosses between *prp39a-1* and *sgs15-1 and sgs15-2* yield late flowering F1 plants, suggesting that mutations in At1g04080 are responsible for the late flowering phenotype of *prp39a-1 sgs15-1* and *sgs15-2* mutants. Introduction of the *2a3* locus into the *prp39a-1* mutant by crossing, yield plants that do not undergo *NIA* S-PTGS. Moreover, crosses between *2a3*/*prp39a-1* and *2a3*/*sgs15-1* or *2a3*/*sgs15-2* yield F1 plants that do not undergo *NIA* S-PTGS, indicating that mutations in At1g04080 are also responsible for *NIA* S-PTGS defect. To definitely confirm this, we transformed the *sgs15-2* mutant with a construct carrying genomic sequences of the At1g04080 gene comprised between the ATG and the stop codon under the control of the pUBQ10 promoter and fused to GFP, hereafter referred to as *pUBQ10:PRP39a-GFP*. At first, we transformed *sgs15-2* mutant plants from which the *2a3* locus has been segregated away. Among 30 *sgs15-2/ pUBQ10:PRP39a-GFP* transformants, 27 developed like wild-type plants with regards to flowering time, confirming that the deletion of At1g04080 is responsible for the late flowering phenotype of the *sgs15-2* mutant. We also transformed *sgs15-2* mutant plants carrying the *2a3* locus. Among 55 transformants *sgs15-2/ pUBQ10:PRP39a-GFP*, 51 exhibited *NIA* S-PTGS (Figure 1C). Altogether, these results indicate that *sgs15-1* and *sgs15-2* are *prp39a* alleles, hereafter referred to as *prp39a-7* and *prp39a-8*, which exhibit a late flowering phenotype, similar to the previously identified *prp39a* mutants (Wang et al., 2007; Kanno et al., 2017). They also indicate that PRP39a participate to PTGS.

PRP39a localizes in nucleoplasm but is excluded from the nucleolus

We explored where PRP39a is expressed in the cell using transgenic *sgs15-2* mutants carrying a *pUBQ10:PRP39a-GFP* construct that perfectly complements the late flowering and PTGS deficient phenotypes. Confocal analysis revealed a PRP39a-GFP diffuse signal in the nucleoplasm, and excluded from the nucleolus (Figure 2). This pattern resembles that of SmD1b, another splicing factor identified through the 2a3 genetic screen (Elvira-Matelot et al., 2016), although SmD1b also associates to nucleocytoplasmic speckles. (Figure 2). Co-localization experiments performed in *N. benthamiana* using XRN2, XRN3 and XRN4, which localize in the nucleolus, nucleoplasm and cytoplasm, respectively, confirmed the nucleoplasmic localization of PRP39a (Figure 2A).

The abundance and splicing of certain endogenous RNA is differently affected in prp39a and smd1b mutants

In yeast, PRP39 is part of the U1 snRNP complex, which plays important roles in splicing. In Arabidopsis, *prp39a* mutants were previously identified in a forward genetic screen for mutants deficient in pre-mRNA splicing (Kanno et al., 2017). Moreover, Kanno et al., 2017 reported RNA-seq data and analyses of differentially expressed genes (DEGs) and differentially alternatively splicied (DAS) events, confirming the role of Arabidopsis PRP39a in splicing. RNAseq performed on *prp39a* in our lab revealed little overlap with the DEGs identified by Kanno et al 2017, but 220 common DAS (Supplemental Figure S1). Considering that Kanno et al., 2017 used different mutant alleles, developmental stages and DAS analysis methods, this significant overlap between DAS support a major role of PRP39a in AS regulation.

Because *prp39a* and *smd1b* mutants were both recovered from the *2a3* PTGS screen, RNAseq was performed in parallel on *prp39a* and *smd1b* mutants to compare the impact of these two conserved spliceosomal components on RNA splicing and RNA abundance. Both *prp39a* and *smd1b* mutations affect the abundance of numerous transcripts, including mRNAs and a number of non-coding RNAs such as long non-coding RNAs (lncRNA) and natural antisense RNAs (NAT) (Figure 3A). DEGs include intron-less and intron-containing transcripts. The proportion of intron-containing compared to intron-less genes was not significantly different in DEGs than it is in all detected transcripts in the dataset. On the contrary, differentially expressed lncRNAs and differentially expressed NATs were significantly enriched for intron-containing transcript ($p<0.001$, Fisher exact test) suggesting a potential role of lncRNA splicing in the control of their abundance (Supplemental Figure S2). Although both proteins are splicing regulators, limited overlap was observed between DEGs in both mutants (Figure 3B). Clustering analysis of DEG expression identified cluster of genes with distinct expression patterns in *prp39a* and *smd1b* (Figure 3C). Gene ontology enrichment analysis of mRNA clusters (Table 1) showed that genes specifically upregulated in *smd1b* (Cluster 3) are enriched for GO categories associated with defense responses such as response to chitin (p.hyper < 5E-16), response to jasmonic acid (p.hyper < 2.3E-6) and glucosinolate metabolism (p.hyper < 2.5E-5). Gene clusters specifically up-regulated in *prp39a* were enriched for growth and development GO terms such as auxin response (p.hyper < 2.6E-7), regulation of growth (p.hyper < 4.6E-6) and negative regulation of flower development (p.hyper < 24E-4). Consistently with the late flowering phenotype in the mutant, the master floral regulator *FLC* accumulated in *prp39a* seedlings compared to WT (Supplemental Figure S3A). Opposite to *smd1b*, specifically down-regulated genes (Cluster 5) in *prp39a* were associated with plant defense responses (Response to bacterium, p.hyper < 3.6E-08; immune system process, p.hyper < 3E-5).

To evaluate the impact of *prp39a* and *smd1b* on mRNA differential splicing, we determined transcript isoforms abundance using the AtRTD2 transcript database (Zhang et al., 2017) and we computed the difference of percentage of spliced-in (ΔPSI) of each possible splicing event genome-wide using SUPPA2 (Trincado et al., 2018). These data showed that *prp39a* had a stronger impact on alternative splicing than *smd1b* when compared to Col-0, with 356 and 142 high confidence DAS (differentially alternatively spliced) genes respectively (ΔPSI >0.2, p.adj < 0.001). Both mutations affected all type of alternative splicing events with intron retention being the most abundant (Figure 4A). Manual evaluation of mRNA-seq read coverage confirmed intron retention events in the *prp39a* background (three examples of AS events are shown in Supplemental Figure S3B). Analysis of DAS genes showed that *prp39a* seems to preferentially affect transcripts with smaller exons and higher number of introns and that DAS in *prp39a* and *smd1b* have significantly smaller introns that non-DAS transcript (Supplemental Figure S4A). GC content analysis of differential introns and their flanking exons indicated that both mutants preferentially target introns having low GC content in their 3’ flanking exons (Supplemental Figure S4B). As previously presented for transcript abundance (DEGs), both mutations affect splicing of different group of transcripts and DAS genes were largely different from DEGs genome-wide (Figure 4B). To further compare the impact of both mutations on splicing, we compared ΔPSI for each individual intron retention event in both mutants. The analysis shows that Intron Retention (IR) changes were not globally correlated in both mutants (as compared to wild type) except for a small group of introns (Figure 4C). Overall, PRP39a and SmD1b appear as AS regulators acting on different splicing targets.

To determine whether changes in splicing or expression of genes involved in RNA silencing could explain the effect of *prp39a* or *smd1b* on PTGS, we compared DEG and DAS with a list of manually curated genes from the literature, which have or could have function in TGS, PTGS or RQC. Only very few DAS or DEG overlapped with this gene set (Supplemental Figure S5, Supplemental Dataset S2 and S3) and no key player of TGS, PTGS or RQC was found to be mis-expressed or mis-spliced in *prp39a* or *smd1b*. This further suggests that these proteins affect PTGS in ways other than regulating the splicing or expression of known RNA silencing genes.

The effect of prp39a on transgene S-PTGS is not specific to intron-containing transgenes but rather depends on the strength of the silencing locus

To ensure that S-PTGS is actually compromised in *2a3/prp39a* plants, RNAs were extracted from wild-type Col, *2a3* silenced plants, and *2a3/prp39a, 2a3/smd1b* and *2a3/sgs3* seedlings and hybridized with a *NIA2* probe. Whereas *2a3* plants accumulated *NIA2* siRNAs, *2a3/prp39a* mutants lacked *NIA2* siRNAs and accumulated *NIA2* mRNA, similar to *2a3/smd1b* and *2a3/sgs3* mutants (Figure 5A), indicating that the *prp39a* mutation actually affects *NIA* S-PTGS. Nevertheless, reminiscent of the *smd1b* mutation (Elvira-Matelot et al., 2016), the *prp39a* mutations do not totally suppress *NIA* S-PTGS. Indeed, ~60% of either *2a3/prp39a-1, 2a3/prp39a-6* or *2a3/prp39a-7* plants and ~40% of the *2a3/smd1b* plants eventually undergo *NIA* S-PTGS at each generation (Figure 5B).

Similar to the *smd1b* mutant, the *prp39a* mutant was recovered from the S-PTGS genetic screen based on the *2a3* line, but not from the S-PTGS genetic screen based on the *L1* line, which carries a silenced *p35S:GUS* transgene. To test if *prp39a* has an effect on *GUS* S-PTGS, the *L1* locus was introduced into *prp39a-1* by crossing. Similar to *L1/smd1b, L1/prp39a* plants still exhibited *GUS* S-PTGS (Figure 5C), explaining why *smd1b* and *prp39a* mutations were not recovered from the *L1* genetic screen.

The *L1* and *2a3* loci differ by their genomic location and structure, but also by the presence of introns in the *2a3* line. Indeed, the *35S:NIA2* transgene carried at the *2a3* locus derives from a genomic fragment containing the three introns of the Arabidopsis *NIA2* gene. To determine if the basis of the different behavior of the *prp39a* mutation towards *L1* and *2a3* is due to the presence of introns, the *159* locus was introduced into the *prp39a* mutant by crossing. The *159* locus carries the same *p35S:GUS* transgene as *L1* except for the presence of a plant intron within the *GUS* sequence (Vancanneyt et al., 1990). Like *L1*, line *159* exhibits high GUS activity at early stages of development, low GUS activity at later stages in 100% of the population (Figures 5C and 5D), and accumulates high levels of *GUS* siRNAs when S-PTGS is triggered (Elvira-Matelot et al., 2016). However, the timing of silencing in line *159* is delayed compared with *L1*, indicating that the *159* locus is a weaker silencing inducer than the *L1* locus. Analysis of *159, 159/prp39a, 159/smd1b* and *159/sgs3* seedlings revealed that efficient *GUS* S-PTGS at the *159* locus requires PRP39a (Figure 5D).

To determine if splicing defects where responsible for S-PTGS impairment in *prp39a* plants, we first examined if the endogenous *NIA1* or *NIA2* genes exhibited splicing defects in *prp39a*. Like in *smd1b*, no intron retention or intron skipping was observed in *prp39a* (Supplemental

Figure S6). Then, we examined if the *prp39a* mutation causes transgene splicing defects in *159/prp39a* and *2a3/prp39a* plants that could affect S-PTGS. However, the *159* and *2a3* transgenes did not show detectable changes in their splicing patterns (Figure 5E), suggesting that the *prp39a* mutation does not compromise transgene splicing. Thus, S-PTGS impairment in *prp39a* does not appear to result from perturbed splicing of *NIA* endogenous genes or *NIA* or *GUS* transgenes, suggesting that PRP39a acts in S-PTGS independent of its role in splicing.

Despite the absence of effect of *prp39a* on *NIA* and *GUS* splicing, the fact that *prp39a* affects S-PTGS at the intron-containing loci *2a3* and *159* but not at the intron-less locus *L1* raised the possibility that PRP39a could contribute to transgene S-PTGS only if the splicing machinery is attracted to transgene RNA because of the presence of an intron. Alternatively, PRP39a could be required to promote S-PTGS only in the case of transgene loci that produce few aberrant RNAs, making it dispensable for transgene loci like *L1*, which exhibit very strong S-PTGS due to the production of high levels of aberrant RNAs (Parent et al, 2015). To discriminate between these two hypotheses, we used the *Hc1* locus, which carries the same *p35S:GUS* intron-less transgene as *L1*, but triggers *GUS* S-PTGS in only 20% of the population at each generation whereas *L1* triggers *GUS* S-PTGS with 100% efficiency (Elmayan et al., 1998b; Gy et al., 2007; Martinez de Alba et al., 2011; Martinez de Alba et al., 2015). Similar to *Hc1/smd1b* and *Hc1/sgs3,* none of the *Hc1/prp39a* plants triggered *GUS* S-PTGS (Figure 5F), indicating that, like SmD1b, PRP39a promotes S-PTGS of intron-less transgenes, but only at loci that trigger S-PTGS with low efficiency. To confirm this, we used the two-components GxA system, where “G” represents a *35S:GFP* transgene and “A” a virus-derived *35S:PVX-GFP* amplicon, which causes *GFP* S-PTGS. Similar to *GxA/smd1b* plants, *GxA/prp39a* plants exhibited GFP expression (Figure 5G), indicating that, similar to *smd1b*, *prp39a* affects *GFP*, *GUS* and *NIA2* S-PTGS, but only from weak S-PTGS loci.

Finally, we tested whether *prp39a* affects IR-PTGS. For this, we crossed *prp39a* to the line *JAP3*, which produces *PDS* dsRNA in companion cells of the phloem, resulting in IR-PTGS of the endogenous *PDS* mRNA in a layer of 15 cells around the veins. *JAP3/smd1b* plants were also generated, and *JAP3/jmj14* plants were used as controls impaired in *PDS* IR-PTGS (Le Masson et al., 2012). *JAP3/prp39a* and *JAP3/smd1b* plants exhibited *PDS* silencing similar to *JAP3* control plants (Figure 5H) and still accumulated *PDS* siRNAs (Figure 5I), indicating that *prp39a* and *smd1b* do not impair IR-PTGS.

Altogether, these results indicate that PRP39a participates in a step that is specific to S-PTGS. The *prp39a* mutation has a significant effect on most S-PTGS loci (*Hc1, 2a3, GxA, 159*), with the exception of the very strong S-PTGS loci *L1*, indicating that PRP39a facilitates S-PTGS but is not a core component of the S-PTGS pathway, similarly to SmD1b (Elvira-Matelot et al., 2016).

PRP39b does not appear to play a role in S-PTGS

In Arabidopsis, SmD1 is encoded by two genes, *SmD1a* and *SmD1b*, which act redundantly, but *SmD1b* plays a major role, likely because it is more expressed than *SmD1a* (Elvira-Matelot et al., 2016). Yeast PRP39 has two orthologs in Arabidopsis, At1g04080 and At5g46400, which are referred to as PRP39a and PRP39b, respectively (Supplemental Figure S7A). At5g46400/PRP39b is less expressed than At1g04080/PRP39a (Kanno et al., 2017; Supplemental Figure S7B), and a lack of redundancy between PRP39a and PRP39b was previously shown for splicing activity (Kanno et al., 2017). To determine if PRP39b plays a role in S-PTGS, a T-DNA insertion line in the PRP39b gene (SAIL_157_B08), hereafter referred to as *prp39b-2*, was crossed to lines *159* and *2a3*. Unlike *159/prp39a*, *159/prp39b* plants undergo *GUS* S-PTGS, similar to *159* control plants (Figure 5C). Moreover, all *2a3/prp39b-2* plants undergo *NIA* S-PTGS, similar to *2a3* control plants. Together, these results suggest that, unlike PRP39a, PRP39b does not contribute to S-PTGS.

PRP39a does not contribute to the endogenous small RNA repertoire Because many of the PTGS-deficient mutants previously identified in our screen are impaired in components of the cellular machinery producing endogenous siRNAs (Elmayan et al., 1998b; Fagard et al., 2000; Mourrain et al., 2000; Morel et al., 2002; Boutet et al., 2003; Jauvion et al., 2010; Le Masson et al., 2012), we examined the accumulation of representative endogenous small RNAs (miRNAs, ta-siRNAs and p4-siRNAs) in *prp39a* mutants. The accumulation of miR775, TAS1-, TAS2- and TAS3-tasiRNAs, and siR1003 was not affected in *prp39a* mutants (Supplemental Figure S8A). To further determine the impact of *prp39a* on sRNA biogenesis, we performed sRNA sequencing on *prp39a* and Col-0 seedlings. Accumulation of 21- to 24-nt siRNAs from introns, exons, promoter and intergenic regions did not show any global changes in *prp39a* as compared to Col-0 (Supplemental Figure S8B). We then used ShortStack to detect siRNA clusters *de novo* and analyzed their differential expression. The expression pattern of siRNA clusters was strongly correlated between *prp39a* and Col-0

(Supplemental Figure S8C). In addition, among the 891 21-22nt long and the 23179 24nt-long siRNA clusters detected in our dataset, only 19 and 96 were differentially regulated, respectively (Supplemental Figure S8C). Together these data indicate that PRP39a does not generally contribute to the production or distribution of small RNAs genome-wide, but likely affects transgene S-PTGS at a different step.

PRP39a and SmD1b interplay with distinct nuclear RQC factors

The results presented above suggest that, like SmD1b, PRP39a is not essential for IR-PTGS and is not a core component of S-PTGS, but somehow facilitates S-PTGS. Because RQC does not interplay with IR-PTGS (Moreno et al, 2013) but limits S-PTGS by degrading part of the transgene aberrant RNAs that provoke the activation of S-PTGS (Gy et al., 2007; Moreno et al., 2013; Lange et al., 2014; Martinez de Alba et al., 2015), we propose that, like SmD1b, the nuclear factor PRP39a facilitates S-PTGS by interplaying with the degradation of transgene aberrant RNAs by the nuclear RQC machinery, thus favoring their transformation in dsRNAs that trigger S-PTGS. To test this hypothesis, *prp39a rqc* double mutants were generated by crossing *159/prp39a* plants with mutants impaired in nuclear RQC. 5'-to-3' nuclear degradation of abRNAs is performed by XRN3, while 3'-to-5' nuclear degradation of abRNAs is performed by the nuclear exosome. Most nuclear exosome components are shared with the cytoplasmic exosome, but HEN2 encodes an RNA helicase that resides in the nucleoplasm and specifically associates with nuclear exosome (Lange et al, 2014). Moreover, *hen2* mutations stimulate S-PTGS, like does *xrn3* (Gy et al, 2007; Lange et al, 2014; Hematy et al, 2016). Therefore, *159/prp39a* plants were crossed with *hen2* and *xrn3*. In parallel, *159/smd1b* plants were crossed with *hen2* mutants to complete the analysis previously initiated by crossing *159/smd1b* with *xrn3* (Elvira-Matelot et al., 2016). Lastly, *159* plants were crossed with *hen2* and *xrn3* to determine if the *159* locus behaves as expected with regards to RQC. *GUS* PTGS was enhanced in *159/hen2* and *159/xrn3* compared with *159* (Figure 6), indicating that S-PTGS at the *159* locus is actually counteracted by RQC. *GUS* PTGS was restored in *159/prp39a hen2* but not *159/prp39a xrn3* (Figure 6). In contrast, *GUS* S-PTGS was restored in *159/smd1b xrn3* and only partially in *159/smd1b hen2* (Figure 6), indicating that PRP39a and SmD1b have distinct ways to limit the action of nuclear RQC against the aberrant RNAs and favor their conversion to dsRNA by RDR6 to initiate S-PTGS.

PRP39a and SmD1b have additive effect in promoting S-PTGS

The results presented above indicate that PRP39a and SmD1b promote S-PTGS by interplaying differently with RQC. To test if PRP39a and SmD1b have additive effects on S-PTGS, lines *2a3/prp39a* and *L1/prp39a* were crossed with lines *2a3/smd1b* and *L1/smd1b* to generate *prp39a smd1b* double mutants. *NIA* S-PTGS was reduced to ~20% in *2a3/prp39a smd1b* plants, which less than *2a3/prp39a* (~60%) and *2a3/smd1b* (~40%) single mutants. Moreover, the rate of of *GUS* S-PTGS was reduced in the *L1/prp39a smd1b* double mutant compared with the *L1/prp39a* and *L1/smd1b* single mutants (Figure 5C). Together, these results support the hypothesis of a synergistic effect of these two mutations in promoting S-PTGS.

PRP39a does not bind to transgene RNA

The absence of a detectable effect of *prp39a* on the splicing of transgene *GUS* and *NIA* RNAs, and the ability of *prp39a* to suppress S-PTGS triggered by intron-less transgenes, suggest that PRP39a facilitates S-PTGS independently of its role in splicing. To determine if, like SmD1b (Elvira-Matelot et al., 2016), PRP39a directly interacts with transgene RNA, the *159/prp39a* line was transformed with the *pUBQ10:PRP39a-GFP* construct, and complemented transformants that developed like wild-type plants and lacked GUS activity were selected to perform RNA immunoprecipation assays (RIP). The nuclei extract (Input) of the *159* line and one *159/prp39a/pUBQ10:PRP39a-GFP* transgenic line that triggered *GUS* S-PTGS as efficiently as the *159* line were used for RIP using anti-GFP antibodies to detect transgene RNAs bound to PRP39a-GFP.

A *159/smd1b/pUBQ10:SmD1b-GFP* line was included as positive control because SmD1b was previously shown to bind *GUS* RNA (Elvira-Matelot et al., 2016). Specific pairs of primers that amplify *GUS* mRNA (Figure 7) were used to perform Reverse Transcription followed by quantitative real-time PCR on the RIP and the Input samples. As positive controls, we also included three primer sets targeting pre-mRNA transcripts identified as differentially spliced in the *prp39a* RNA-seq analysis. Results showed an enrichment for the three pre-mRNAs targeted by PRP39a, but did not show any enrichment of *GUS* RNA in *159/prp39a/pUBQ10:PRP39a-GFP* plants (Figure 7). To determine the efficiency of the PRP39-GFP immunoprecipitation in the RIP buffer, we performed a western blot of the input (nuclei extract), unbound and IP (ImmunoPrecipitated) fraction in two *159/prp39a/pUBQ10:PRP39a-GFP* transgenic lines using GFP antibody and an IGG negative control. Results showed no signal in the unbound fraction and a stronger signal in the IP fraction when GFP antibodies were used. On the other hand, IgG IP did not recover any protein signal. Together these results indicate that PRP39a-GFP was quantitatively recovered during IP and that the IP did not produce any detectable background. Together, these results indicated that PRP39a does not seem to bind to RNAs produced by silenced transgenes.

PRP39a largely reduces RNA degradation by the nuclear exosome

The results presented above indicate that whereas SmD1b binds to transgene abRNAs to prevent their degradation by RQC, PRP39a does not bind to transgene abRNAs but somehow prevents 3'-to-5' transgene abRNA degradation by the HEN2-containing nuclear exosome. The exact nature of the abRNA that initiate *GUS* S-PTGS is still not known; however, we previously identified an uncapped RNA antisense to *GUS* (referred to as *SUG*). *SUG* RNA is detected in *rdr6*, indicating that it is not an RDR6 product. *SUG* accumulates at higher level in *rdr6* mutants than in wildtype plants, indicating that, like *GUS* mRNA, it is targeted by *GUS* siRNAs. *SUG* accumulates at even higher levels in *rdr6 xrn3 xrn4* mutants, indicating that it is also a target of RQC (Parent et al, 2015). We therefore examined *SUG* accumulation using strand specific qRT-PCR in *159, 159 prp39a, 159 smd1b* and *159 rdr6*. Similar to *159 rdr6, 159 smd1b* accumulated high levels of *SUG* (Supplemental Figure S9). Given the sensitivity of the uncapped *SUG* RNA to XRN-mediated degradation (Parent et al, 2015), this result supports the proposed role of SmD1b in binding transgene abRNA to protect them against 5'-to-3' degradation (Figure 6A). In contrast, *SUG* level was only slightly increased in *159 prp39a*. This result also is consistent with the fact that PRP39a does not bind transgene RNA (Figure 7), and likely limits the action of the nuclear exosome on aberrant RNAs indirectly.

To test this hypothesis, we examined the accumulation of known HEN2 targets (Lange et al., 2014) in *prp39a*, assuming that, in the absence of PRP39a, these targets should be more efficiently degraded than in wild type plants. Approximately half of the known HEN2 targets (41 out of 97) that we could detect in both RNA-seq dataset were down-regulated in *prp39a* (Figure 8A), supporting this hypothesis. We further examined the expression of a subset of HEN2 targets in *prp39a, hen2* and *prp39a hen2* seedlings. As previously shown (Lange et al., 2014), we observed accumulation of short transcripts derived from mRNAs, 3'extension of mRNAs, and incompletely spliced transcripts in *hen2* (Figure 8). In *prp39a*, these transcripts were only marginally affected, but the introduction of the *prp39a* mutation in the *hen2* mutants background strongly affected the accumulation of these exosome targets, indicating that PRP39a has a protecting effect on RNAs targeted by the nuclear exosome.

Because the binding of SmD1b to transgene abRNAs likely protects them against degradation by RQC, and in particular XRN3-mediated 5'-to-3' degradation, we examined if impairing SmD1b could have a global effect on endogenous XRN3 targets. For this, we examined the accumulation of a subset of known XRN3 targets (Kurihara et al., 2012) in *smd1b*, *xrn3* and *smd1b xrn3*. In our conditions, not all known XRN3 targets RNA accumulated in *xrn3*; however, we observed that introducing *smd1b* mutation in the *xrn3* background did not affect the accumulation of the XRN3 targets (Supplemental Figure S10), suggesting that SmD1b does not have a general protective effect on XRN3 substrates, further supporting the hypothesis that PRP39a and SmD1b have distinct actions on PTGS.

Discussion

RNA interference (RNAi) is a conserved silencing mechanism initiated by dsRNA. In plants, RNAi induced by dsRNA directly resulting from the folding of a self-complementary RNA is referred to as IR-PTGS. However, plant RNAi can also be initiated by (trans)genes that do not directly produce dsRNA, and is referred to as S-PTGS. It is generally assumed that (trans)genes that undergo S-PTGS produce aberrant RNAs (abRNAs) that can be transformed into dsRNA by the RNA-dependent RNA polymerase RDR6. The exact nature of these abRNAs and what makes them specific substrates for RDR6 as well as how they escape degradation by nuclear and cytoplasmic RQC before they reach cytoplasmic siRNA-bodies remains unknown. Recently, we reported that the conserved spliceosome component SmD1b binds transgene RNA and promotes S-PTGS, even when it is triggered by intron-less transgenes (Elvira-Matelot et al., 2016). Because IR-PTGS is not affected in *smd1b* mutants and because S-PTGS is restored in double mutants impaired in SmD1b and RQC factors, we proposed that SmD1b protects transgene aberrant RNAs from degradation by nuclear RQC pathways. Hence, beside its role in splicing, a spliceosome component can play a role in the interplay between RQC and S-PTGS in the nucleus (Elvira-Matelot et al., 2016).

Here, we describe that another spliceosome component, PRP39a, was identified in the same genetic screen for S-PTGS-deficient mutants as SmD1b. Like SmD1b, PRP39a has dual role in splicing and S-PTGS, and suppresses S-PTGS even when it is triggered by intron-less transgenes (Figure 5). However, SmD1b and PRP39a have distinct roles in both splicing and

S-PTGS. RNA-seq analysis of *prp39a* and *smd1b* mutants identified different sets of deregulated RNAs, both at expression level and alternative splicing. Since the relative abundance of alternatively spliced isoforms is mainly influenced by the binding of trans-acting factors such as RNA binding protein and splicing factors on cis-regulatory element in pre-mRNA, it is not surprising that mutants in distinct splicing factors affect splicing of different set of mRNAs. In addition, PRP39 has been shown to be linked to the U1-snRNP component of the spliceosome, which plays a critical role in 5' splice site recognition (Li et al., 2017), whereas SmD1 is a core component of the splicing machinery, suggesting that these two proteins play central but distinct roles in splicing.

Given that *prp39a* and *smd1b* mutants i) do not affect IR-PTGS, ii) affect S-PTGS triggered by intron-less transgenes (Figure 5), and iii) do not modify in the expression or splicing of RQC or PTGS components (Supplemental Figures S3 and S4), PRP39a and SmD1b likely promote a step of PTGS that is specific to the S-PTGS pathway, independently of their role in splicing. Because S-PTGS involves abRNAs that must escape degradation by RQC before being transformed into dsRNA by RDR6, and because PRP39a and SmD1b reside in the nucleoplasm (Figure 2), we examined the interplay between the splicing factors PRP39a and SmD1b and nucleoplasmic RQC factors involved in 5'-to-3' RNA degradation (XRN3) or 3'-to-5' RNA degradation (HEN2). Double mutant analyses revealed distinct genetic interactions of PRP39a and SmD1b with nuclear RQC machineries. Indeed, S-PTGS, was restored in *smd1b xrn3* and *prp39a hen2* mutants, partially in *smd1b hen2,* but not in *prp39a xrn3* (Figure 6A), suggesting that SmD1b mostly counteracts the action of 5'-to-3' nuclear RNA degradation, whereas PRP39a only counteracts the action of 3'-to-5' nuclear RNA degradation. RIP-qPCR experiments also revealed that SmD1b binds transgene RNAs whereas PRP39a does not (Figure 7). Therefore, it is likely that the binding of SmD1b to transgene RNA limits its degradation by XRN3 and partly by the exosome, whereas PRP39a seems to indirectly limit the degradation of transgene RNAs by the nuclear exosome but not by XRN3. To test whether PRP39a generally counteracts the nuclear exosome, we analyzed the accumulation of known endogenous HEN2 targets in a *prp39a* mutant, and found that the *prp39a* mutation reduces the abundance of almost half of known HEN2 targets (Figure 8), supporting this hypothesis. Nevertheless, we cannot rule out the possibility that PRP39a promotes the production of HEN2 targets.

The way PRP39a interplays with the exosome remains to be determined. The structure of the NEXT complex in humans revealed a connection with splicing factors (Falk et al., 2016). Moreover, recent results suggest that the splicing factor SRSF3 may help in the recognition and degradation of certain mRNAs by the nuclear exosome (Mure et al., 2018). Because the Arabidopsis splicing factor PRP39a limits the degradation of certain mRNAs by the nuclear exosome, our results suggest that a connection between splicing factors and RNA degradation also exist in plants, but opposite to that observed in humans. PRP39 is part of the U1 snRNP complex, which beside its role in splicing, is also involved in 3’ end formation by preventing premature transcription termination and cleavage/polyadenylation in vertebrates (Kaida et al., 2010). It is not known if U1 snRNP components participate to 3’ end formation in plants, but Kanno et al., 2017, reported that, in Arabidopsis, PRP39a is highly co-expressed with several cleavage and polyadenylation factors. Given that transgene RNAs lacking a polyA tail are good substrates for RDR6 (Baeg et al., 2017) and highly prone to undergo PTGS (Luo and Chen, 2007), it is tempting to speculate that PRP39a promotes S-PTGS by influencing 3’ end formation and indirectly limiting the activity of the nuclear exosome. Still, PRP39a does not fully prevent the action of the nuclear exosome. Whether the limited effect of the *prp39a* mutation on the accumulation of nuclear exosome targets is due to specific targeting of PRP39a to certain genes only or to the redundant activity of PRP39b remains to be determined. However, PRP39b is expressed at a lower level than PRP39a, and the single *prp39b* mutation has no effect on S-PTGS (Figure 5D). Moreover, PRP39a and PRP39b were previously shown to not act redundantly in splicing (Kanno et al., 2017). Therefore, PRP39a may be the only plant PRP39 protein interplaying between splicing, RQC and PTGS.

In conclusion, we propose that PRP39a and SmD1b promote S-PTGS by limiting RQC-mediated degradation of transgene aberrant RNAs using two distinct but complementary manners (Figure 9). SmD1b binds to transgene abRNAs, thus limiting their degradation by XRN3 and partly by the exosome, whereas PRP39a does not bind transgene RNAs but indirectly limits the action of the nuclear exosome, and consequently the degradation of transgene abRNAs. Supporting this hypothesis, a *prp39a smd1b* double mutant exhibits enhanced synergistic suppression of S-PTGS (Figures 5B and 5D), likely because XRN3 and the nuclear exosome degrade transgene aberrant RNAs more efficiently when both PRP39a and SmD1b are absent. Nevertheless, S-PTGS still occurs to some extent in a *prp39a smd1b* double mutant, indicating that additional nuclear factors are at work to prevent the complete degradation of transgene aberrant RNAs by RQC.

Methods

Plant material and growth conditions

All *Arabidopsis thaliana* plants are in the Columbia accession. Transgenic lines *2a3, 159, L1* and *Hc1*, and mutants *hen2 (sop3-2), sgs3-1, smd1b-1, xrn3-3* have been previously described (Elmayan et al., 1998a; Mourrain et al., 2000; Gy et al., 2007; Moreno et al., 2013; Lange et al., 2014; Elvira-Matelot et al., 2016). The T-DNA insertion mutants *prp39a-1* (SAIL_1249A03) and *prp39b-2* (SAIL_157_B08) were obtained from NASC (Alonso et al., 2003). Plants were grown on Bouturage media (Duchefa) under long-day conditions (16 hours of light at an intensity of 160 µmol/m²/s at 21°C and 70% relative humidity, 8 hours dark at 18°C and 70% relative humidity), then transferred to soil after two weeks and grown in controlled growth chambers under the same long-day conditions.

Plasmid constructs

The *pUBQ10:PRP39a-GFP* construct was generated using the Gateway technology (Invitrogen) as follows. The PRP39a genomic fragment was amplified from the first ATG to the penultimate codon to omit the stop codon and cloned into the pENTR/D TOPO vector using pENTR/D TOPO cloning kit (Invitrogen). PRP39a was then transferred into the plant expression vector pB7FWG2 (Karimi et al., 2002) using the LR clonase II kit (Invitrogen). The *p35S:XRN2-RFP, p35S:XRN3-RFP and p35S:XRN4-RFP* constructs have been described previously (Lorkovic et al., 2008; Moreno et al., 2013).

Arabidopsis transformation and Nicotiana benthamiana agro-infiltration

Agrobacterium strains carrying plasmids of interest were grown overnight at 28°C in 3 ml LB medium containing the appropriate antibiotics to a final OD600 between 1 and 2. For Arabidopsis transformation, the bacteria were pelleted and resuspended in 300 ml of infiltration medium (5% sucrose, 10mM MgCl2, 0,015% silwet L-77) to a final OD600 of 1, which was used for floral dipping. For *Nicotiana benthamiana* agro-infiltration, the bacteria were pelleted and resuspended in 1 ml of infiltration medium (10 mM $MgCl_2$, 10 mM MES pH 5.2, 150 mM acetosyringone) to a final OD600 of 0.1. The solution containing the bacteria was injected into the abaxial side leaves using a 1-ml syringe and samples were observed in a confocal microscope 3 days after infiltration.

Imaging and image analysis

After agro-infiltration, fluorescent cells were imaged by confocal microscopy (Leica TCS SP2, Leica Microsystem, Wetzlar, Germany) with excitation at 488 nm and fluorescence emission signal between 495 and 530 nm for GFP fusions, and excitation at 543 nm and emission signal between 555 and 620 nm for DsRed or RFP fusions. The Leica confocal software was used for image acquisition and for the quantification of fluorescence profiles. Sequential scans were performed when necessary. Spectral profiles were calculated for five cells. Data processing was performed using ImageJ (http://rsbweb.nih.gov/ij/).

RNA extraction and RNA gel blot analysis

For RNA gel blot analyses, frozen tissue was homogenized in a buffer containing 0.1 M NaCl, 2% SDS, 50 mM Tris-HCl pH 9.0, 10mM EDTA pH 8.0 and 20 mM beta-mercaptoethanol and RNAs were extracted two times with phenol and recovered by ethanol precipitation. To obtain high molecular weigh (HMW) RNA fraction, resuspended RNAs were precipitated overnight in 2M LiCl at 4°C and recovered by centrifugation. For low molecular weigh (LMW) RNA analysis, total RNA was separated on a 15% denaturing PAGE gel, stained with ethidium bromide, and transferred to nylon membrane (HybondNX, Amersham). LMW RNA and U6 hybridizations were at 50°C with hybridization buffer containing 5X SSC, 20mM Na_2HPO_4 pH 7.2, 7% SDS, 2X Denhardt’s Solution and denatured sheared salmon sperm DNA (Invitrogen). HMW RNA hybridization was at 37°C in PerfectHyb Plus buffer (Sigma). Blots were hybridized with a radioactively labeled random-primed DNA probes for *GUS* mRNA and *GUS* siRNAs, and an end-labeled oligonucleotide probe for U6 detection.

RT-PCR analysis

Total RNA was prepared from plantlets at different developmental stages using Trizol (Invitrogen). The DNAse1 treatment (Thermo Scientific) was performed according to the manufacturer’s protocols. For reverse transcription with SuperScriptII (Invitrogen), 2 µg of total DNAse-treated RNA was used. One microliter of the resulting cDNA solution was used for RT-PCR or RT-qPCR analyses. The latter was done using standard protocols and a complete list of RT-qPCR primers is available in Supplemental Table 1. Each cDNA sample was precisely calibrated and verified for two constitutive genes, AT1G13320 and AT4G26410 (Czechowski et al., 2005). For RT-PCR, the amplification was performed as follows: one cycle of 4 min at 98°C, 26 cycles of 30 s at 98°C, 30 s at 59°C, and 1 min at 72°C. The products were separated on agarose gel and stained with ethidium bromide. RT-qPCR was performed using a Roche Light Cycler 480 standard protocol (40 cycles, 60°C annealing).

RNA Immunoprecipitation

Eleven-day-old plants grown in Petri dishes were irradiated three times with UV using a Crosslinker CL-508 (Uvitec) at 0.400 J/cm^2. Briefly, fixed material was ground in liquid nitrogen, homogenized, and nuclei isolated and lysed according to (Gendrel et al., 2005).RNA IP was basically performed as described by (Carlotto et al., 2015). The nuclei extract (Input) was used for the immunoprecipitation performed by Direct ChIP Protocol of the Diagenode IP-Star SX-86 Compact robot, using 50 µl of Dynabeads – Protein A (Novex 10008D; Life Technologies, USA) and anti-GFP (632381; Clontech, USA) antibodies. Beads were washed twice for 5 min at 4°C with Wash Buffer 1 (150 mM NaCl, 1% Triton, 0.5% NP-40, 1 mM EDTA, 20 mM Tris-HCl pH7.5) and twice with Wash Buffer 2 (20 mM Tris-HCl pH 8) instead of the Diagenode assigned buffers, and finally resuspended in 100 µl Proteinase K buffer (100mM Tris-HCl pH 7.4, 50 mM NaCl, 10 mM EDTA). After Proteinase K (AM2546; Ambion, USA) treatment, beads were removed with a magneto, and the supernatants were transferred to a 2-ml tube. Each RNA sample was extracted from 800 µl (8 IP-Star tubes) of RNA IP product using 1 ml of TriReagent (Sigma-Aldrich T9424, USA) as indicated by the manufacturer. 80 µl of nuclei extracts were used for Input RNA extraction. The IP and the Input samples were treated with DNase, and random hexamers were used for subsequent RT. Quantitative real-time PCR reactions were performed using specific primers. Results were expressed as a percentage of cDNA detected after IP, taking the Input sample as 100%.

GUS extraction and activity quantification

GUS protein was extracted and GUS activity was quantified as described before (Gy et al., 2007) from cauline leaves of flowering plants by measuring the quantity of 4-methylumbelliferone product generated from the substrate 4-methylumbelliferyl-b-D-glucuronide (Duchefa) on a fluorometer (Fluoroscan II; Thermo Scientific).

Transcriptome analysis by RNA sequencing

Total RNA was extracted using RNeasy plant mini kit (Qiagen) from whole 14-day-old Col-0, *prp39a-1* (SAIL_1249A03), *smd1b-1* (Elvirat-Matelot et al., 2016) and Col-0 plants grown on ½ MS medium. Three independent biological replicates were produced per genotype. For each biological repetition and each point. After RNA extraction, polyA RNAs were purified using Dynabeads mRNA direct micro kit (Ambion). Libraries were constructed using the Tru-Seq Stranded mRNA Sample Prep kit (Illumina®). Sequencing was carried out at the POPS Transcriptomic Platform, Institute of Plant Sciences Paris-Saclay in Orsay, France. The Illumina HiSeq2000 technology was used to perform paired-end 100-bp sequencing. A minimum of 30 millions of paired-end reads by sample were generated. RNA-seq preprocessing included trimming library adapters and quality controls with Trimmomatic (Bolger et al., 2014). Paired-end reads with Phred Quality Score Qscore > 20 and read length > 30 bases were kept, and ribosomal RNA sequences were removed with SortMeRNA (Kopylova et al., 2012). Processed reads were aligned using Tophat2 with the following arguments : --max-multihits 1 -i 20 --min-segment-intron 20 --min-coverage-intron 20 --library-type fr-firststrand --microexon-search -I 1000 --max-segment-intron 1000 --max-coverage-intron 1000 --b2-very-sensitive. Reads overlappings exons per genes were counted using the FeatureCounts function of the Rsubreads package using the GTF annotation files from the Araport11 repository (https://www.araport.org/downloads/Araport11_Release_201606/annotation/Araport11_GFF3_genes_transposons.201606.gff.gz). Significance of differential gene expression was estimated using DEseq2 (Love et al., 2014), and the FDR correction of the p-value was used during pairwise comparison between genotypes. A gene was declared differentially expressed if its adjusted p-value (FDR) was ≤ 0.01 and its absolute fold change was ≥ 1.5.

Hierarchical clustering analysis was performed in R on scaled normalized read count data. Heatmap was plotted using pheatmap package (https://cran.r-project.org/web/packages/pheatmap/index.html). Gene ontology enrichment analysis was done using the systemPipeR package (Backman and Girke, 2016).

Transcript level quantification was performed using pseudo-alignment counts with kallisto (Bray et al., 2016) on AtRTD2 transcripts sequences (https://ics.hutton.ac.uk/atRTD/RTD2/AtRTDv2_QUASI_19April2016.fa) with a K-mer size of 31-nt. Differential AS events in the AtRTD2 database were detected using SUPPA2 with default parameters (Trincado et al., 2018). Intron length, GC content was performed using in-house R script.

Small RNAseq library construction, sequencing and analysis

Small RNA (sRNA) libraries were constructed from 1µg of total RNA treated with DNaseI (Thermo Scientific) using the NEBNext® Multiplex Small RNA Library Prep Set for Illumina® kit (New England Biolabs) according to manufacturer's instructions. Libraries were sequenced on a NextSeq 500 Sequencing System (Illumina) using 75-nt single-end reads. The sequencing adapters were removed using cutadapt (v2.10) and sequence matching Arabidopsis ribosomal or transfer RNA were discarded using bowtie (v1.3.1). The reads of length 19 to 25nt, were then mapped on TAIR10 genome with the help of ShortStack (v3.8.5) without mismatch (--mismatches 0), keeping all primary multi-mapping (--bowtie_m all) and correcting for multi-mapped reads according to the uniquely mapped reads (--mmap u). For each gene annotation in Araport11 (coding genes, non-coding RNA and miRNA precursor), 19-nt to 25-nt long read density was counted with ShortStack. sRNA accumulation in each library was normalized using the median of ratio of the corresponding miRNA counts inside DESeq2 (v1.34.0). Differential sRNA accumulation and log2 fold changes between genotypes were computed using DESeq2. FDR correction of the p-value was used.

Acknowledgements

We thank Hervé Ferry and Philippe Maréchal for plant care, Florian Bardou for his initial contribution to this work, and Taline Elmayan for fruitful discussion. This work was supported by the Agence Nationale de la Recherche ANR-16-CE12-0032 (to HV and MC).

Authors contributions

HV and MC designed the experiments. All authors contributed to the production and analysis of the results. HV, JB and MC wrote the paper.

Conflict of interest

The authors declare that they have no conflict of interest.

Accession Numbers

Gene sequence data from this article can be found in the GenBank/EMBL libraries under the following accession numbers: PRP39a (At1g04080), PRP39b (At5g46400).

RNAseq data generated in this study are accessible through NCBI's Gene Expression Omnibus (GSE158145).

References

Adenot, X., Elmayan, T., Lauressergues, D., Boutet, S., Bouche, N., Gasciolli, V., and Vaucheret, H. (2006). DRB4-dependent TAS3 trans-acting siRNAs control leaf morphology through AGO7. Curr Biol **16,** 927-932.

Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C.C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H., Geralt, M., Hazari, N., Hom, E., Karnes, M., Mulholland, C., Ndubaku, R., Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel, D., Carter, D.E., Marchand, T., Risseeuw, E., Brogden, D., Zeko, A., Crosby, W.L., Berry, C.C., and Ecker, J.R. (2003). Genome-wide insertional mutagenesis of Arabidopsis thaliana. Science **301,** 653-657.

Backman, T.W., and Girke, T. (2016). systemPipeR: NGS workflow and report generation environment. BMC Bioinformatics **17,** 388.

Baeg, K., Iwakawa, H.O., and Tomari, Y. (2017). The poly(A) tail blocks RDR6 from converting self mRNAs into substrates for gene silencing. Nat Plants **3,** 17036.

Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. Cell **116,** 281-297.

Baulcombe, D. (2004). RNA silencing in plants. Nature **431,** 356-363.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics **30,** 2114-2120.

Boutet, S., Vazquez, F., Liu, J., Beclin, C., Fagard, M., Gratias, A., Morel, J.B., Crete, P., Chen, X., and Vaucheret, H. (2003). Arabidopsis HEN1: a genetic link between endogenous miRNA controlling development and siRNA controlling transgene silencing and virus resistance. Curr Biol **13,** 843-848.

Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol **34,** 525-527.

Brosnan, C.A., Mitter, N., Christie, M., Smith, N.A., Waterhouse, P.M., and Carroll, B.J. (2007). Nuclear gene silencing directs reception of long-distance mRNA silencing in Arabidopsis. Proc Natl Acad Sci U S A **104,** 14741-14746.

Carlotto, N., Wirth, S., Furman, N., Ferreyra Solari, N., Ariel, F., Crespi, M., and Kobayashi, K. (2015). The chloroplastic DEVH-box RNA helicase INCREASED SIZE EXCLUSION LIMIT 2 involved in plasmodesmata regulation is required for group II intron splicing. Plant Cell Environ.

Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K., and Scheible, W.R. (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. Plant Physiol **139,** 5-17.

Elmayan, T., Balzergue, S., Beon, F., Bourdon, V., Daubremet, J., Guenet, Y., Mourrain, P., Palauqui, J.C., Vernhettes, S., Vialle, T., Wostrikoff, K., and Vaucheret, H. (1998a). Arabidopsis mutants impaired in cosuppression. Plant Cell **10,** 1747-1758.

Elmayan, T., Balzergue, S., Beon, F., Bourdon, V., Daubremet, J., Guenet, Y., Mourrain, P., Palauqui, J.C., Vernhettes, S., Vialle, T., Wostrikoff, K., and Vaucheret, H. (1998b). Arabidopsis mutants impaired in cosuppression. Plant Cell **10,** 1747-1758.

Elvira-Matelot, E., Bardou, F., Ariel, F., Jauvion, V., Bouteiller, N., Le Masson, I., Cao, J., Crespi, M.D., and Vaucheret, H. (2016). The Nuclear Ribonucleoprotein SmD1 Interplays with Splicing, RNA Quality Control, and Posttranscriptional Gene Silencing in Arabidopsis. Plant Cell **28,** 426-438.

Fagard, M., Boutet, S., Morel, J.B., Bellini, C., and Vaucheret, H. (2000). AGO1, QDE-2, and RDE-1 are related proteins required for post-transcriptional gene silencing in plants, quelling in fungi, and RNA interference in animals. Proc Natl Acad Sci U S A **97,** 11650-11654.

Falk, S., Finogenova, K., Melko, M., Benda, C., Lykke-Andersen, S., Jensen, T.H., and Conti, E. (2016). Structure of the RBM7-ZCCHC8 core of the NEXT complex reveals connections to splicing factors. Nat Commun **7,** 13573.

Gazzani, S., Lawrenson, T., Woodward, C., Headon, D., and Sablowski, R. (2004). A link between mRNA turnover and RNA interference in Arabidopsis. Science **306,** 1046-1048.

Gendrel, A.V., Lippman, Z., Martienssen, R., and Colot, V. (2005). Profiling histone modification patterns in plants using genomic tiling microarrays. Nat Methods **2,** 213-218.

Gregory, B.D., O'Malley, R.C., Lister, R., Urich, M.A., Tonti-Filippini, J., Chen, H., Millar, A.H., and Ecker, J.R. (2008). A link between RNA metabolism and silencing affecting Arabidopsis development. Dev Cell **14,** 854-866.

Gy, I., Gasciolli, V., Lauressergues, D., Morel, J.B., Gombert, J., Proux, F., Proux, C., Vaucheret, H., and Mallory, A.C. (2007). Arabidopsis FIERY1, XRN2, and XRN3 Are Endogenous RNA Silencing Suppressors. Plant Cell **19,** 3451-3461.

Jauvion, V., Elmayan, T., and Vaucheret, H. (2010). The conserved RNA trafficking proteins HPR1 and TEX1 are involved in the production of endogenous and exogenous small interfering RNA in Arabidopsis. Plant Cell **22,** 2697-2709.

Jouannet, V., Moreno, A.B., Elmayan, T., Vaucheret, H., Crespi, M.D., and Maizel, A. (2012). Cytoplasmic Arabidopsis AGO7 accumulates in membrane-associated siRNA bodies and is required for ta-siRNA biogenesis. EMBO J **31,** 1704-1713.

Kaida, D., Berg, M.G., Younis, I., Kasim, M., Singh, L.N., Wan, L., and Dreyfuss, G. (2010). U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. Nature **468,** 664-668.

Kanno, T., Lin, W.D., Fu, J.L., Chang, C.L., Matzke, A.J.M., and Matzke, M. (2017). A Genetic Screen for Pre-mRNA Splicing Mutants of Arabidopsis thaliana Identifies Putative U1 snRNP Components RBM25 and PRP39a. Genetics **207,** 1347-1359.

Kopylova, E., Noe, L., and Touzet, H. (2012). SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics **28,** 3211-3217.

Kurihara, Y., Schmitz, R.J., Nery, J.R., Schultz, M.D., Okubo-Kurihara, E., Morosawa, T., Tanaka, M., Toyoda, T., Seki, M., and Ecker, J.R. (2012). Surveillance of 3' Noncoding Transcripts Requires FIERY1 and XRN3 in Arabidopsis. G3 (Bethesda) **2,** 487-498.

Lange, H., Zuber, H., Sement, F.M., Chicher, J., Kuhn, L., Hammann, P., Brunaud, V., Berard, C., Bouteiller, N., Balzergue, S., Aubourg, S., Martin-Magniette, M.L., Vaucheret, H., and Gagliardi, D. (2014). The RNA helicases AtMTR4 and HEN2 target specific subsets of nuclear transcripts for degradation by the nuclear exosome in Arabidopsis thaliana. PLoS Genet **10,** e1004564.

Le Masson, I., Jauvion, V., Bouteiller, N., Rivard, M., Elmayan, T., and Vaucheret, H. (2012). Mutations in the Arabidopsis H3K4me2/3 demethylase JMJ14 suppress posttranscriptional gene silencing by decreasing transgene transcription. Plant Cell **24,** 3603-3612.

Li, X., Liu, S., Jiang, J., Zhang, L., Espinosa, S., Hill, R.C., Hansen, K.C., Zhou, Z.H., and Zhao, R. (2017). CryoEM structure of Saccharomyces cerevisiae U1 snRNP offers insight into alternative splicing. Nat Commun **8,** 1035.

Lippman, Z., and Martienssen, R. (2004). The role of RNA interference in heterochromatic silencing. Nature **431,** 364-370.

Lorkovic, Z.J., Hilscher, J., and Barta, A. (2008). Co-localisation studies of Arabidopsis SR splicing factors reveal different types of speckles in plant cell nuclei. Exp Cell Res **314,** 3175-3186.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol **15,** 550.

Luo, Z., and Chen, Z. (2007). Improperly terminated, unpolyadenylated mRNA of sense transgenes is targeted by RDR6-mediated RNA silencing in Arabidopsis. Plant Cell **19,** 943-958.

Martinez de Alba, A.E., Jauvion, V., Mallory, A.C., Bouteiller, N., and Vaucheret, H. (2011). The miRNA pathway limits AGO1 availability during siRNA-mediated PTGS defense against exogenous RNA. Nucleic Acids Res **39,** 9339-9344.

Martinez de Alba, A.E., Moreno, A.B., Gabriel, M., Mallory, A.C., Christ, A., Bounon, R., Balzergue, S., Aubourg, S., Gautheret, D., Crespi, M.D., Vaucheret, H., and Maizel, A. (2015). In plants, decapping prevents RDR6-dependent production of small interfering RNAs from endogenous mRNAs. Nucleic Acids Res **43,** 2902-2913.

Mlotshwa, S., Pruss, G.J., Peragine, A., Endres, M.W., Li, J., Chen, X., Poethig, R.S., Bowman, L.H., and Vance, V. (2008). DICER-LIKE2 plays a primary role in transitive silencing of transgenes in Arabidopsis. PLoS One **3,** e1755.

Morel, J.B., Godon, C., Mourrain, P., Beclin, C., Boutet, S., Feuerbach, F., Proux, F., and Vaucheret, H. (2002). Fertile hypomorphic ARGONAUTE (ago1) mutants impaired in post-transcriptional gene silencing and virus resistance. Plant Cell **14,** 629-639.

Moreno, A.B., Martinez de Alba, A.E., Bardou, F., Crespi, M.D., Vaucheret, H., Maizel, A., and Mallory, A.C. (2013). Cytoplasmic and nuclear quality control and turnover of single-stranded RNA modulate post-transcriptional gene silencing in plants. Nucleic Acids Res **41,** 4699-4708.

Mourrain, P., Beclin, C., Elmayan, T., Feuerbach, F., Godon, C., Morel, J.B., Jouette, D., Lacombe, A.M., Nikic, S., Picault, N., Remoue, K., Sanial, M., Vo, T.A., and Vaucheret, H. (2000). Arabidopsis SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural virus resistance. Cell **101,** 533-542.

Mure, F., Corbin, A., Benbahouche, N.E.H., Bertrand, E., Manet, E., and Gruffat, H. (2018). The splicing factor SRSF3 is functionally connected to the nuclear RNA exosome for intronless mRNA decay. Sci Rep **8,** 12901.

Palauqui, J.C., Elmayan, T., Pollien, J.M., and Vaucheret, H. (1997). Systemic acquired silencing: transgene-specific post-transcriptional silencing is transmitted by grafting from silenced stocks to non-silenced scions. Embo J **16,** 4738-4745.

Parent, J.S., Jauvion, V., Bouche, N., Beclin, C., Hachet, M., Zytnicki, M., and Vaucheret, H. (2015). Post-transcriptional gene silencing triggered by sense transgenes involves uncapped antisense RNA and differs from silencing intentionally triggered by antisense transgenes. Nucleic Acids Res **43,** 8464-8475.

Taochy, C., Yu, A., Bouche, N., Bouteiller, N., Elmayan, T., Dressel, U., Carroll, B.J., and Vaucheret, H. (2019). Post-transcriptional gene silencing triggers dispensable DNA methylation in gene body in Arabidopsis. Nucleic Acids Res **47,** 9104-9114.

Taochy, C., Gursanscky, N.R., Cao, J., Fletcher, S.J., Dressel, U., Mitter, N., Tucker, M.R., Koltunow, A.M.G., Bowman, J.L., Vaucheret, H., and Carroll, B.J. (2017). A Genetic Screen for Impaired Systemic RNAi Highlights the Crucial Role of DICER-LIKE 2. Plant Physiol **175,** 1424-1437.

Thran, M., Link, K., and Sonnewald, U. (2012). The Arabidopsis DCP2 gene is required for proper mRNA turnover and prevents transgene silencing in Arabidopsis. Plant J **72,** 368-377.

Trincado, J.L., Entizne, J.C., Hysenaj, G., Singh, B., Skalic, M., Elliott, D.J., and Eyras, E. (2018). SUPPA2: fast, accurate, and uncertainty-aware differential splicing analysis across multiple conditions. Genome Biol **19,** 40.

Vaistij, F.E., Jones, L., and Baulcombe, D.C. (2002). Spreading of RNA targeting and DNA methylation in RNA silencing requires transcription of the target gene and a putative RNA-dependent RNA polymerase. Plant Cell **14,** 857-867.

Vancanneyt, G., Schmidt, R., O'Connor-Sanchez, A., Willmitzer, L., and Rocha-Sosa, M. (1990). Construction of an intron-containing marker gene: splicing of the intron in transgenic plants and its use in monitoring early events in Agrobacterium-mediated plant transformation. Mol Gen Genet **220,** 245-250.

Wang, C., Tian, Q., Hou, Z., Mucha, M., Aukerman, M., and Olsen, O.A. (2007). The Arabidopsis thaliana AT PRP39-1 gene, encoding a tetratricopeptide repeat protein with similarity to the yeast pre-mRNA processing protein PRP39, affects flowering time. Plant Cell Rep **26,** 1357-1366.

Yu, A., Saudemont, B., Bouteiller, N., Elvira-Matelot, E., Lepere, G., Parent, J.S., Morel, J.B., Cao, J., Elmayan, T., and Vaucheret, H. (2015). Second-Site Mutagenesis of a Hypomorphic argonaute1 Allele Identifies SUPERKILLER3 as an Endogenous Suppressor of Transgene Posttranscriptional Gene Silencing. Plant Physiol **169,** 1266-1274.

Zhang, X., Zhu, Y., Liu, X., Hong, X., Xu, Y., Zhu, P., Shen, Y., Wu, H., Ji, Y., Wen, X., Zhang, C., Zhao, Q., Wang, Y., Lu, J., and Guo, H. (2015). Plant biology. Suppression of endogenous gene silencing by bidirectional cytoplasmic RNA decay in Arabidopsis. Science **348,** 120-123.

Figure legends

Figure 1: Mutations in PRP39a cause late flowering and suppress S-PTGS

Photographs of 40-day-old plants of wild-type Col, *prp39a-7 (sgs15-1)* and *prp39a-8 (sgs15-2)* mutants.

B, Structure of the top of chromosome 1 in wild-type Col, *prp39a-7 (sgs15-1)* and *prp39a-8 (sgs15-2)*. A deletion removed the 5’ half of the *PRP39a* gene in *prp39a-7 (sgs15-1),* while an inversion disrupted the 5’ end of the *PRP39a* gene in *prp39a-8 (sgs15-2).*C, Photographs of 20-day-old *2a3/prp39a-8 (sgs15-2)* plants, which resist to *NIA* S-PTGS, and *2a3/prp39a-8 (sgs15-2) /pUBQ10:PRP39a* transformants, which undergo *NIA* S-PTGS.

Figure 2: PRP39a subcellular localization

A, Confocal images of *N. benthamiana* leaves infiltrated with *pUBQ10:PRP39a-GFP* reveal the localization of PRP39a in the nucleoplasm. *pUBQ10:XRN2-RFP*, *pUBQ10:XRN3-RFP* and *pUBQ10:XRN4-RFP* were used as controls for nucleolar, nucleoplasmic and cytoplasmic compartments. B, Confocal images of Arabidopsis root expressing *pUBQ10:PRP39a-GFP or pUBQ10:SmD1b-GFP* reveal distinct nuclear localization

Figure 3. Different subsets of mRNA and lncRNA are deregulated in prp39a and smd1b as compared to Col-0 wild type

A, Number of up and down regulated genes of each class of RNA for prp39a and smd1b as compared to wildtype. B, Overlap of up and down-regulated lincRNA, NAT and mRNA genes in prp39a and smd1b as compared to Col-0. C, Hierarchical clustering of the normalized RNA abundance in Col-0, prp39a, smd1b. Only differentially expressed NAT, lincRNA and mRNA are represented. Co-expression clusters of mRNA genes showing distinct expression profiled are highlighted with colors.

Figure 4: Changes in RNA abundance in *smd1b* and *prp39a* compared to Col-0

A, Number of each class of AS showing significant change in *prp39a* or *smd1b* compared to Col-0. B, Comparison of DEG and DAS genes in *prp39a* or *smd1b* compared to Col-0. C, Comparison of Δpsi of *prp39a* relative to Col-0 with Δpsi *of smd1b* relative to Col-0 for all IR events. Significant changes in *prp39a* vs Col-0, *smd1b* vs Col-0 or both are highlighted in red, blue and green, respectively.

Figure 5: Comparative effect of *prp39a* and *smd1b* mutations on different PTGS systems

A, RNA gel blot analyses of *NIA* mRNA and siRNAs in the indicated genotypes. *25S* rRNA ethidium bromide staining served as loading controls for high molecular weight RNA blots. *U6* snRNA hybridization served as loading controls for low molecular weight RNA blots. Part of these blots was previously shown in Elvira-Matelot et al, 2016. Original scans can be seen in Supplemental Figure S11. B, Kinetics of *NIA* S-PTGS in the indicated genotypes. Seeds were sown in vitro and transferred to soil after 10 days. The percentage of silenced plants was determined by scoring the number of chlorotic plants (n = 96 plants for each genotype). C, Kinetics of *GUS* S-PTGS at the intron-less *L1* locus in the indicated genotypes. Seeds were sown in vitro and harvested for GUS analysis at 5, 11 and 17 days after germination. GUS is expressed as arbitrary units of fluorescence per mn per ug of protein. (n = 24 plants for each genotype at each timepoint).D, Kinetics of *GUS* S-PTGS at the intron-containing *159* locus in the indicated genotypes. Seeds were sown in vitro and harvested for GUS analysis at 8 and 13 days after germination. GUS is expressed as arbitrary units of fluorescence per mn per ug of protein. (n = 8 plants for each genotype at each timepoint). E, Analysis of *NIA2* RNA splicing by RT-PCR using primers spanning an intron. Transgene cassette organization is shown above the gel picture, red arrows mark the relative position of RT-PCR primers. F, Efficiency of *GUS* S-PTGS at the intron-less *Hc1* locus in the indicated genotypes. Seeds were sown in vitro and transferred to soil after 10 days. The percentage of silenced plants was determined by GUS activity at the flowering stage (n = 96 plants for each genotype). G, Photographs of plants carrying the GxA transgene loci in the indicated genotypes. H, Photographs of plants carrying the JAP3 transgene locus in the indicated genotypes.I, RNA gel blot analyses of *PDS* siRNAs in the indicated genotypes. miR390 hybridization served as loading control.

Figure 6: Interplay between *prp39a* and *smd1b* mutations and RQC mutations

Time course analysis of *GUS* S-PTGS at the 159 locus in the indicated genotypes. Seeds were sown in vitro and harvested for GUS analysis at 5, 11 and 17 days after germination. GUS activity is expressed as arbitrary units (A.U) of fluorescence per min per ug of protein, and is represented as boxplots (n = 24 plants for each genotype at each time-point). Results for each time-point were analyzed separately by a one-way analysis of variance (ANOVA) followed by Tukey’s post hoc test. Different letters indicate statistical differences between genotypes ($p < 0.05$) at the corresponding developmental time point.

Figure 7: Transgene RNA immunoprecipitation in *159/prp39a/pUBQ10:PRP39a-GFP* and *159/smd1b/pUBQ10:SmD1b-GFP* plants

Lines exhibiting a full restoration of *GUS* S-PTGS were used for immunoprecipitation using GFP antibodies. A, RIP-qRT-PCR analysis using *GUS* and *NPTII* primers. As positive control for PRP39a-GFP RIP, we included three primer sets targeting pre-mRNA transcripts identified as differentially spliced in the *prp39a* RNA-seq analysis. Data represent the mean percentage of input of 3 biological replicates. Error bars show standard error of the mean. Significant enrichment over the IgG RIP was determined using a student-t test (* p.val <0.05). B, Western blot analysis of the PRP39a-GFP immunoprecipitation, the membranes was probed with GFP antibodies (α-GFP). Input, unbound and IP fraction from immunoprecipitation with GFP antibodies or normal Rabbit IgG was loaded on gel. GFP was performed on two independent pUBI:PRP39-GFP transgenic lines. Mock IP using normal rabbit IgG was performed on one single transgenic line (Line1).

Figure 8: PRP39a generally limits the degradation of HEN2 targets by the nuclear exosome

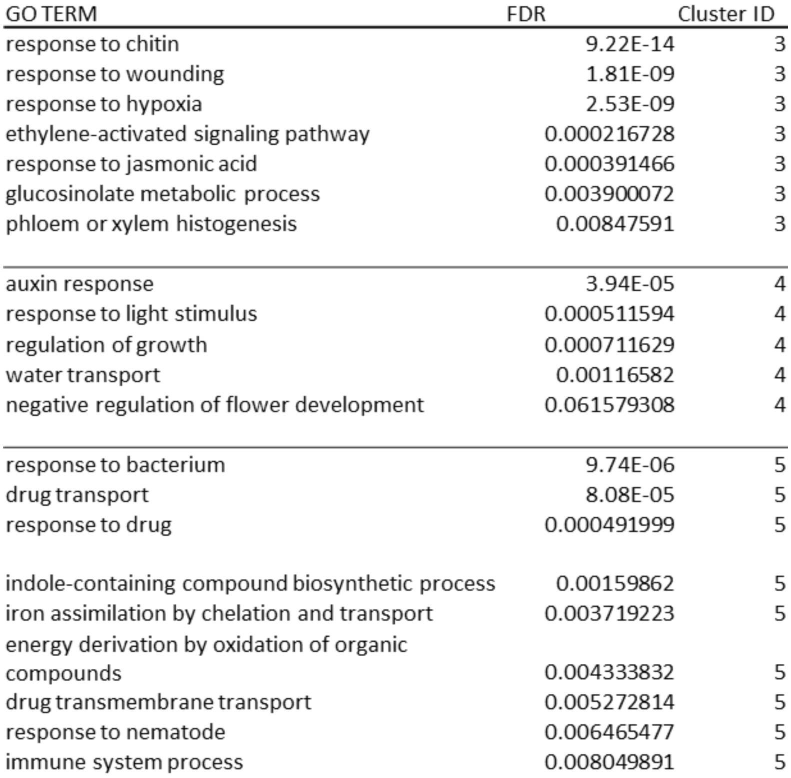
A, Overlap between down-regulated and up-regulated genes related to WT in *prp39a* and *hen2* mutants, respectively. Only genes detected in both RNA-seq were kept for the analysis.

B, Relative expression of selected HEN2 targets determined by RT-qPCR in WT plants and *prp39a, hen2* and *prp39a hen2* mutants. The red lines show the relative position of qPCR amplicons. On the gene model : dark and light blue boxes mark CDS and exons, respectively. Light blue thick lines show introns and black thin lines show the surrounding genomic DNA. Results are shown as box-plots of log2 fold change of 3 to 4 biological replicates, individual data point are also plotted. Significant difference between genotypes was determined using One-way Anova with post-hoc pair-wise Tukey HSD test (p<0.05). For each gene two pair of primers have been used. One measuring the steady state abundance of the transcript and one measuring the aberrant RNA (3' extension, unspliced transcript, see Lange et al 2014 for details).

Figure 9: Tentative model for PRP39a and SmD1b action in S-PTGS

Transgene loci that undergo S-PTGS likely produce aberrant RNAs (abRNA). Degradation of these abRNAs by nuclear RQC 5'-to-3, and 3'-to-5' RNA degradation pathways limits the amount of abRNAs that enter cytoplasmic siRNA-bodies where RDR6 transform them into dsRNA that are processed by DCL2 and DCL4 into siRNAs that are loaded on AGO1 to promote the degradation of regular mRNA. By binding to transgene abRNAs, SmD1b likely impairs XRN3-mediated 5'-to-3 degradation and partly limits exosome-mediated 3'-to-5' degradation. In contrast, PRP39a does not bind transgene abRNAs, but somehow limits the action of HEN2 or other components of the nuclear exosome, without altering the activity of XRN3 on these abRNAs. In the absence of PRP39a and SmD1b, transgene abRNAs are more sensitive to 3’-to-5’ and 5’-to-3’ degradation, respectively, thus reducing the probability that a sufficient amount of abRNAs reach siRNA-bodies to activate S-PTGS.

Table 1. Gene ontology analysis of gene clusters described in Figure 3C



GO TERM	FDR	Cluster ID
response to chitin	9.22E-14	3
response to wounding	1.81E-09	3
response to hypoxia	2.53E-09	3
ethylene-activated signaling pathway	0.000216728	3
response to jasmonic acid	0.000391466	3
glucosinolate metabolic process	0.003900072	3
phloem or xylem histogenesis	0.00847591	3
auxin response	3.94E-05	4
response to light stimulus	0.000511594	4
regulation of growth	0.000711629	4
water transport	0.00116582	4
negative regulation of flower development	0.061579308	4
response to bacterium	9.74E-06	5
drug transport	8.08E-05	5
response to drug	0.000491999	5
indole-containing compound biosynthetic process	0.00159862	5
iron assimilation by chelation and transport	0.003719223	5
energy derivation by oxidation of organic compounds	0.004333832	5
drug transmembrane transport	0.005272814	5
response to nematode	0.006465477	5
immune system process	0.008049891	5

Supplemental Data

Supplemental Figure S1 supporting Figure 3: Overlap between our mRNAseq analysis of *prp39a* and Col-0 and that of Kanno et al., 2017

A) overlap between DEG. B) overlap between DAS.

Supplemental Figure S2 supporting Figure 3: Analysis of transcript classes affected in *prp39a* and *smd1b* mutants compared to Col-0.

A) Number of intron-containing and intron-differentially expressed genes within each class of transcript (Structural RNA = tRNA and rRNA)

B) Ratio of intron vs intron-less transcripts within each class differentially expressed or detected genes.

Supplemental Figure S3 supporting Figure 3: A, Relative mRNA abundance fold change calculated from polyA RNA-seq of FLC (AT5G10140) in smd1b and prp39a as compared to Col-0.

(* : FDR<0.05). B, : Normalized mRNA-seq read coverage on three genes showing signifcantly retained introns in prp39a as compared to Col-0. Red boxes show the position of retained introns detected by SUPPA2**.**

Supplemental Figure S4 supporting Figure 4: Analysis of exon-intron properties in DAS genes from in *prp39a* and *smd1b* mutants compared to Col-0.

A) Mean introns length, exons length and mean number of introns per DAS genes in *prp39a* and *smd1b* compared to Col-0.

B) GC content of flanking 5’ and 3’ exons and introns showing significant intron retentions in *prp39a* and *smd1b* compared to Col-0 (“*prp39a* vs Col-0 introns diff”, “*smd1b* vs Col-0 introns”) compared to non-retained introns (“Introns non diff”)

Significant differences were calculated with an ANOVA and post hoc Tukey HSD test using **A**, non-DAS genes (Not DAS) and **B**, non-retained introns (“Introns non diff”)
as reference group

Supplemental Figure S5 supporting Figure 4: Differential expression and splicing of genes encoding silencing and RQC components in *prp39a* and *smd1b* mutants compared to Col-0.

Overlap of DEG and DAS genes in *smdb1* and *prp39a* compared to Col-0 among genes curated from the literature whose function in TGS, PTGS or RQC has been experimentally demonstrated (listed in Supplemental Dataset S2).

Supplemental Figure S6 supporting Figure 4: *NIA1* and *NIA2* expression profile in *prp39a* and *smd1b* mutants compared to Col-0.

RNAseq analysis of *NIA1* and *NIA2* expression in *prp39a* and *smd1b* compared to Col-0 does not reveal defects in intron splicing.

Supplemental Figure S7: Characterization of the *PRP39b* gene

A) Alignment of PRP39a and PRP39b proteins. Conserved amino acids are boxed in black and similar amino acids are boxed in grey.

B) ATH1 array expression profiles of *PRP39a* and *PRP39b* genes.

C) Effect of *prp39a* and *prp39b* mutations on S-PTGS triggered by the *159* or *2a3* locus. The percentage of plants showing *GUS* S-PTGS n the indicated genotypes was determined by quantitative GUS activity measurements (n = 96 plants for each genotype). The percentage of silenced plants showing *NIA* S-PTGS in the indicated genotypes was determined by scoring the number of chlorotic plants (n = 96 plants for each genotype).

Supplemental Figure S8 supporting Figure 5: Effect of *prp39a* mutations on the accumulation of endogenous small RNAs.

A. RNA gel blot analysis of representative endogenous small RNAs in wild-type Col, *prp39a-7 (sgs15-1)* and *prp39a-8 (sgs15-2)* mutants. *U6* snRNA hybridization served as loading control. B Fraction of 19-nt to 25-nt long siRNA-seq reads mapped to the whole CDS, exons, 3' and 5'Untranslated Region (3'UTR, 5'UTR), Transcripts, introns and Gene promoters (1kb upstream). C, Normalized abundance of 21-22nt and 24nt siRNA reads on siRNA clusters detected by ShortStack. Red dots mark clusters which show a statistically significant difference of siRNA abundance (DESeq2, FDR <0.01, 3 bioreps).

Supplemental Figure S9 supporting Figure 5: Accumulation of transgene aberrant RNA SUG measured by strand specific RT-qPCR in plants *159, 159 prp39a, 159 smd1b, 159 sgs3*.

Results are mean of 3 biological replicates. Error bars show standard error of the replicates. Significant difference compared to the 159 line was determined using a student's-t test (* p.val <0.05). Primers and the strand specific PCR strategy is described in Parent et al., 2015

Supplemental Figure S10 supporting Figure 8: SmD1b does not generally limits the degradation of XRN3 targets.

Relative expression of selected XRN3 targets (as defined in Kurihara et al., 2012) determined by qRT-qPCR in 159 plants and 159 *smd1b,159 xrn3* and 159 *smd1b xrn3.* Results are displayed as boxplot of log2 fold change of 3 to 4 biological replicates, individual data point are also plotted. Significant difference between genotypes was determined using One-way Anova with post-hoc pair-wise Tukey HSD test (p<0.05).

Supplemental Figure S11: original blots for Figure 5B

RNA gel blot analyses of *NIA* mRNA and siRNAs in the indicated genotypes. *25S* rRNA ethidium bromide staining served as loading controls for high molecular weight RNA blots. *U6* snRNA hybridization served as loading controls for low molecular weight RNA blots. The part of the blots previously shown in Elvira-Matelot et al, 2016 are indicated in blue, while the parts of the blots shown in Figure 5B are indicated in red.

Supplemental Figure S12: original blots for Figure 5I

RNA gel blot analyses of *PDS* siRNAs in the indicated genotypes. *U6* snRNA and miR390 hybridization served as loading controls.

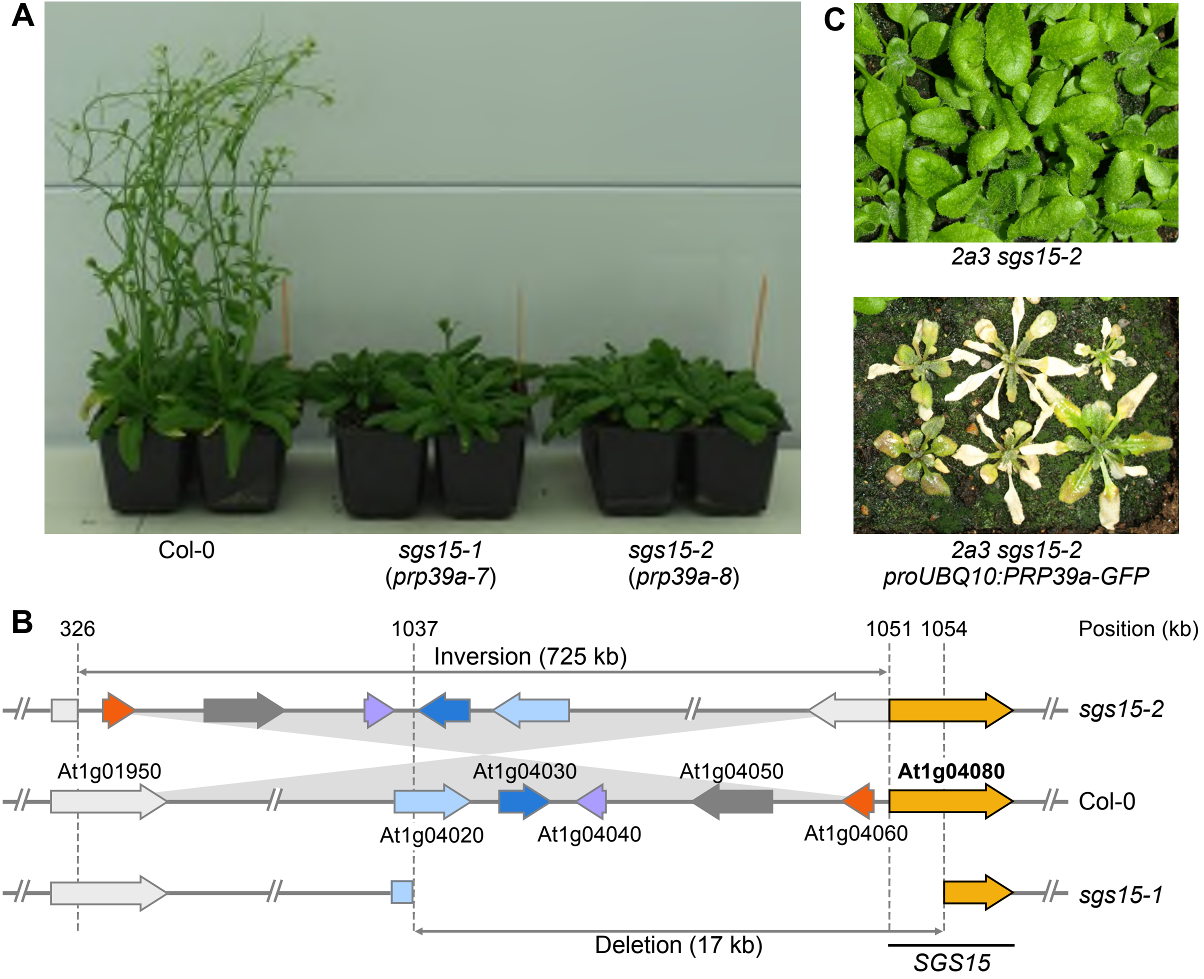
Supplemental Dataset S1: Primers used in this study

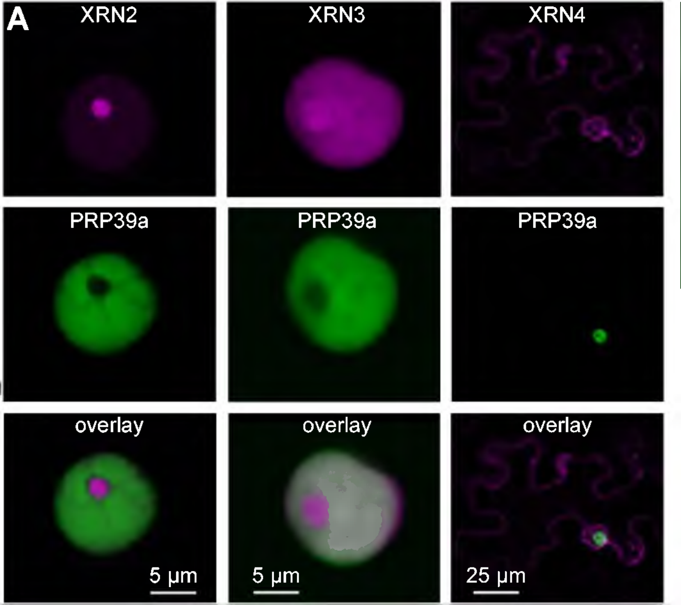
Supplemental Dataset S2: List of curated PTGS and RQC genes analyzed in Supplemental Figure S5.

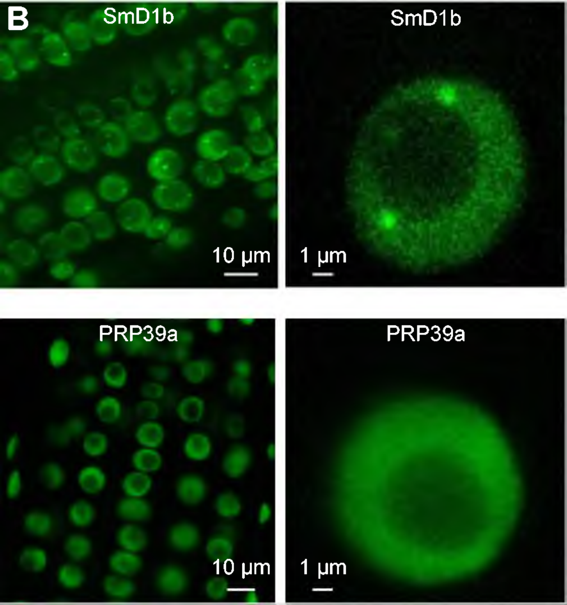
DEG and DAS genes between *prp39a-7* and Col-0 were compared to a set of genes manually curated from the literature to be involved in PTGS, or RQC. Column "prp39a_vs_col","smd1b_vs_col" and "PTGS or RQC related genes". 0 and 1 means presence /absence of the the gene in the considered gene set. When available the gene Name in indicated in the Name column and the biochemical pathway is described in the pathway column. FDR, False Discovery Rate. dPSI (percent spliced index difference). Event show the unique ID of the AS event as defined by SUPPA2.

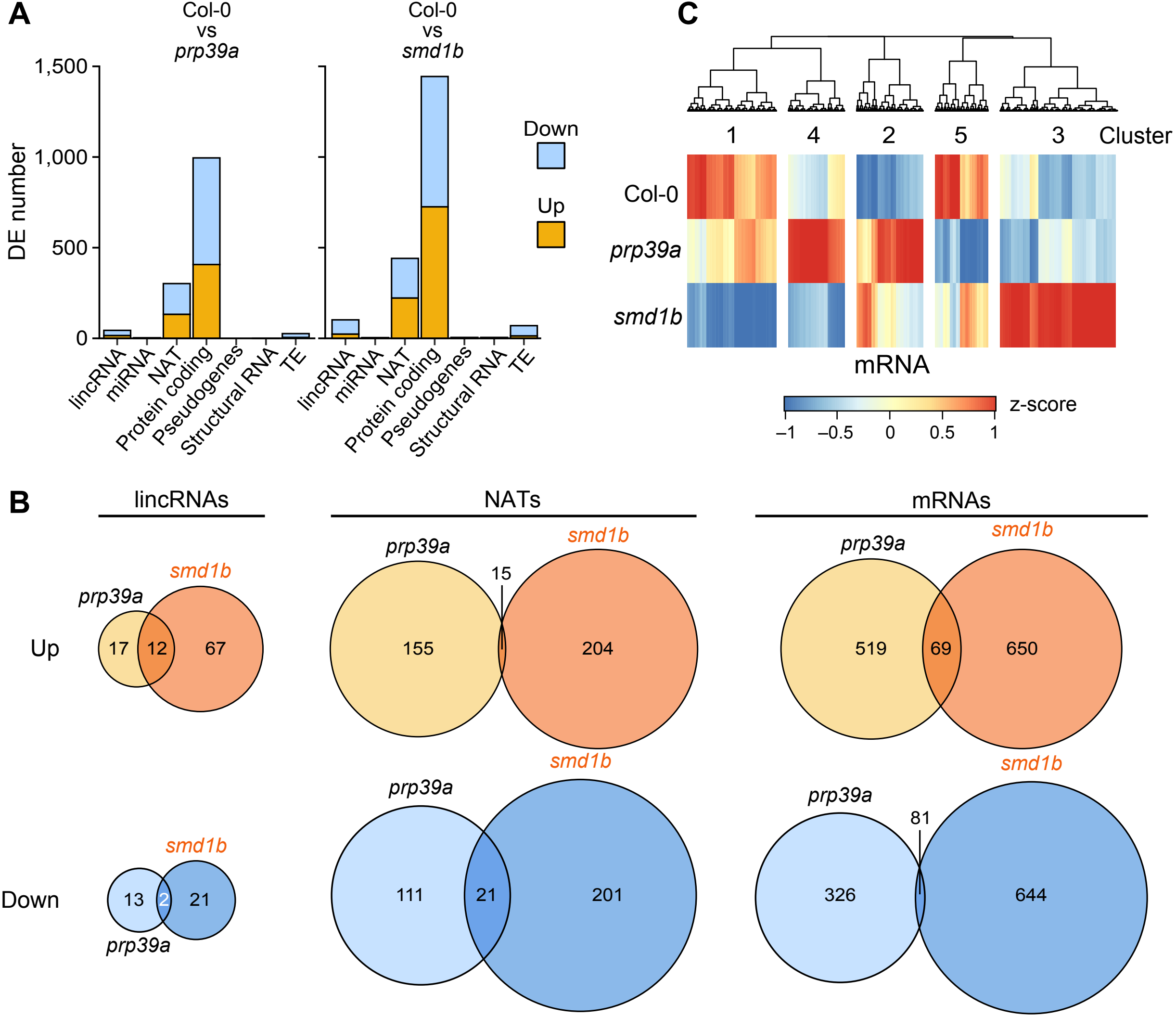
Supplemental Dataset S3: List of Differentially Expressed (DEG) and Differentially Spliced Genes in prp39a-7 and smd1b as compared to Col-0 WT seedlings.

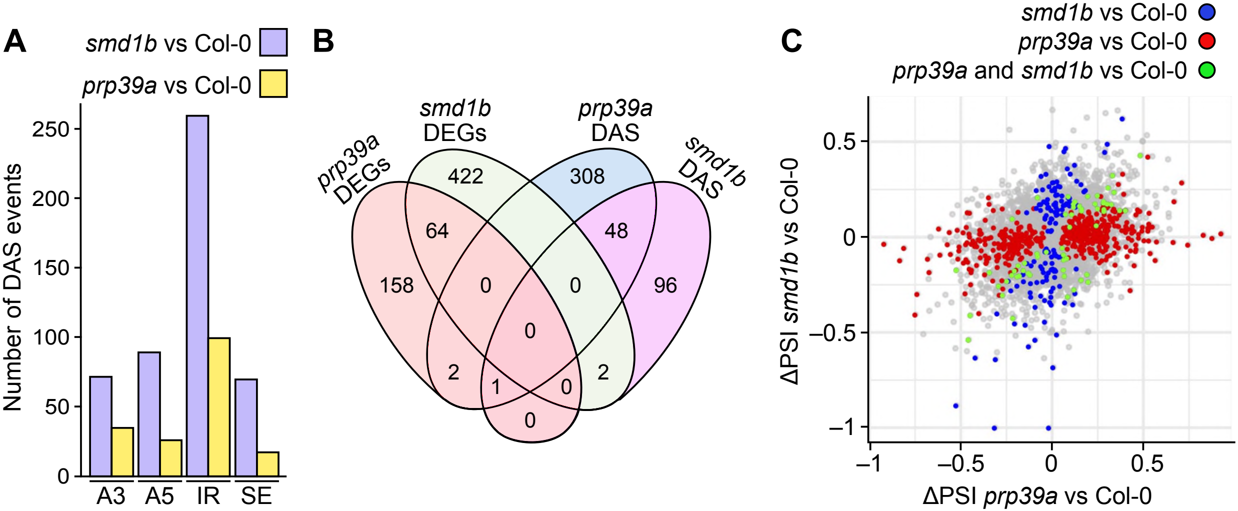
Supplemental Dataset S4: GUS activity in individual plants for Figures 5C, 5D and 6

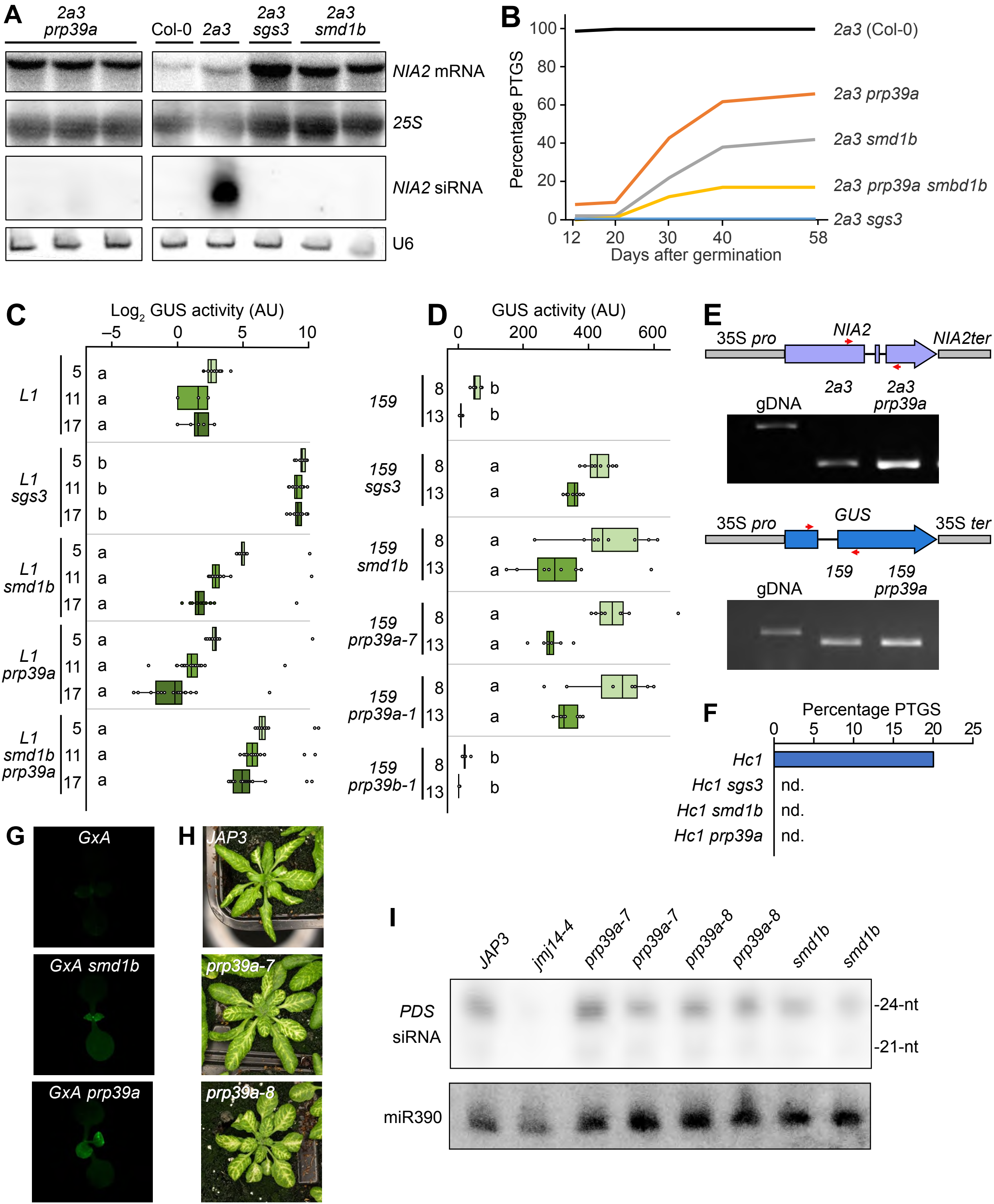

A
Col-0
sgs15-1
(*prp39a-7*)
sgs15-2
(*prp39a-8*)
C
2a3 sgs15-2
2a3 sgs15-2
proUBQ10:PRP39a-GFP
B
326
1037
1051
1054
Position (kb)
Inversion (725 kb)
sgs15-2
At1g01950
At1g04030
At1g04050
At1g04080
Col-0
At1g04020
At1g04040
At1g04060
sgs15-1
Deletion (17 kb)
SGS15

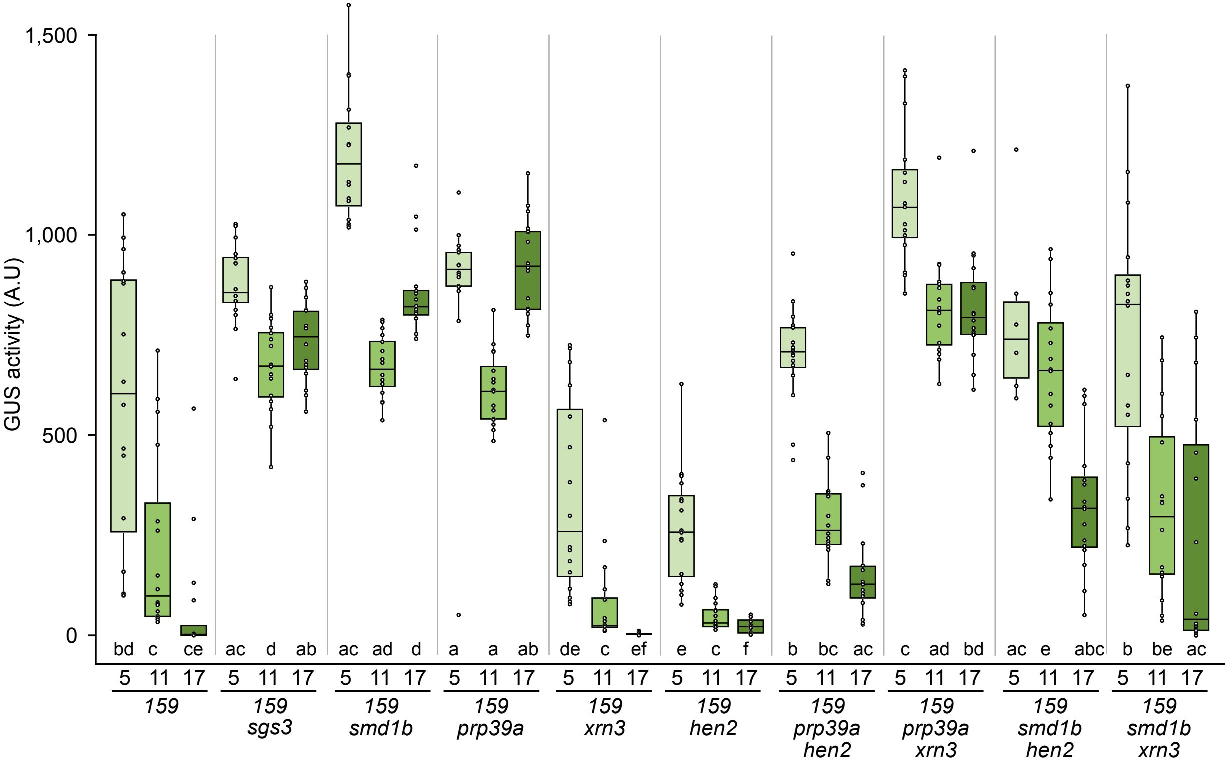

A
XRN2
XRN3
XRN4
PRP39a
PRP39a
PRP39a
overlay
overlay
overlay
5 μm
5 μm
25 μm

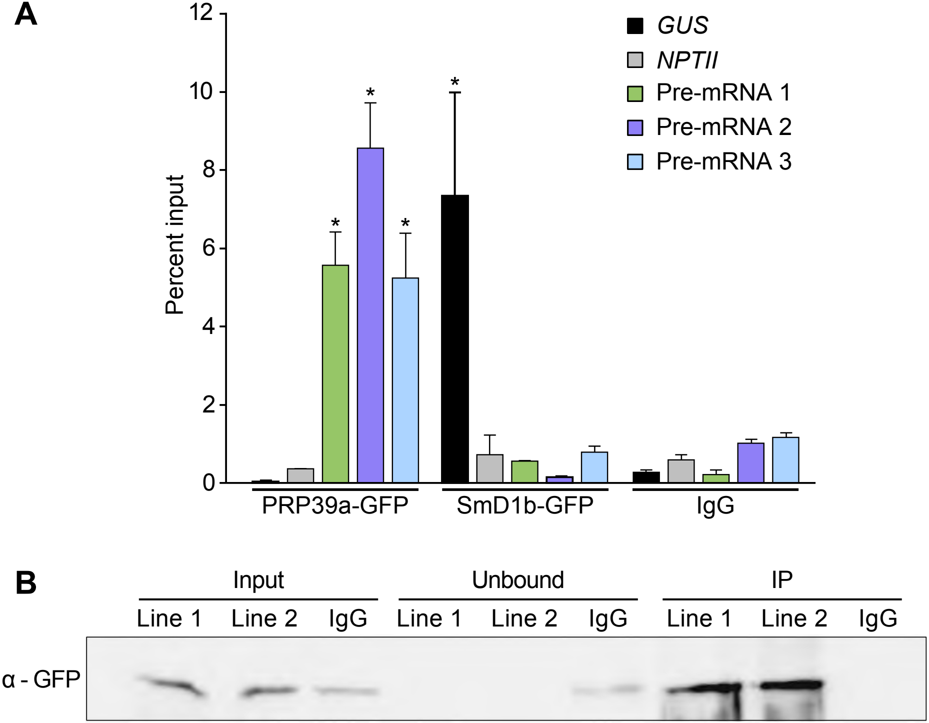

B
SmD1b
10 µm
SmD1b
1 µm
PRP39a
10 µm
PRP39a
1 µm

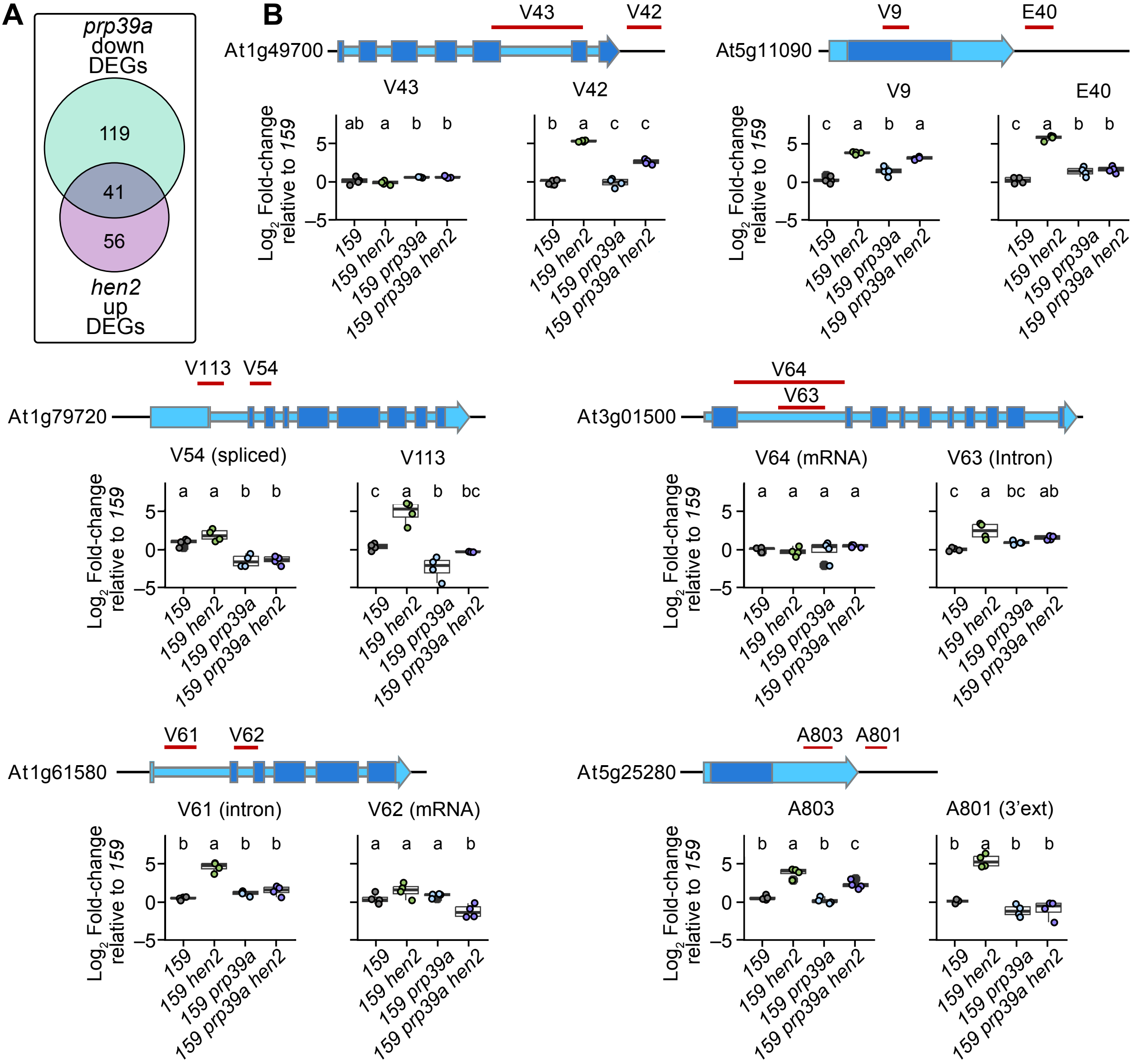

A
Col-0 vs prp39a
Col-0 vs smd1b
DE number
1,500
1,000
500
0
lincRNA
miRNA
NAT
Protein coding
Pseudogenes
Structural RNA
TE
Down
Up
B
lincRNAs
NATs
mRNAs
Up
Down
prp39a
smd1b
17
12
67
155
15
204
519
69
650
13
2
21
111
21
201
326
81
644
C
1
4
2
5
3
Cluster
Col-0
prp39a
smd1b
mRNA
z-score
−1
−0.5
0
0.5
1

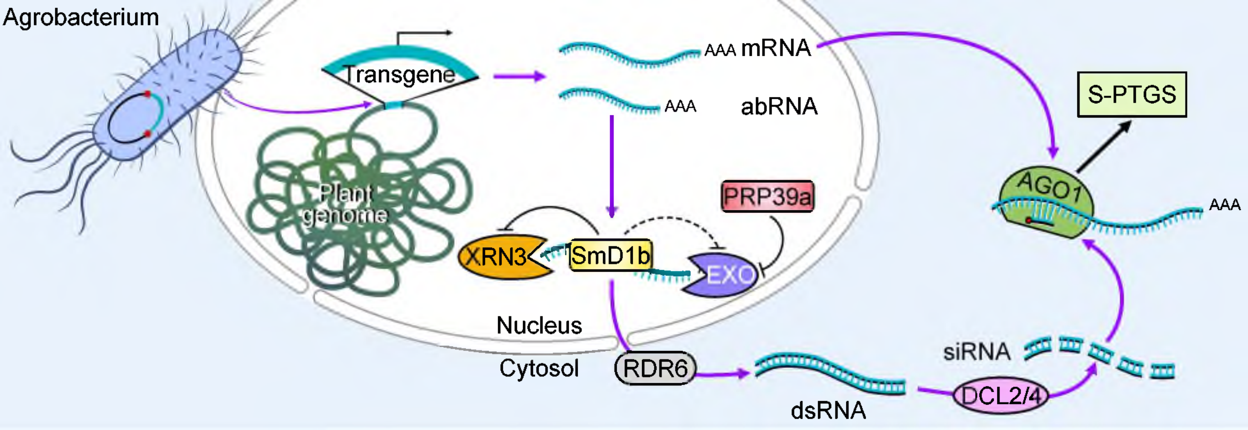

A
smd1b vs Col-0
prp39a vs Col-0
Number of DAS events
0
50
100
150
200
250
A3
A5
IR
SE
B
prp39a DEGs
smd1b DEGs
prp39a DAS
smd1b DAS
422
308
64
48
158
0
0
96
0
2
1
0
2
0
C
smd1b vs Col-0
prp39a vs Col-0
prp39a and smd1b vs Col-0
ΔPSI smd1b vs Col-0
0.5
0
−0.5
−1
−1
−0.5
0
0.5
ΔPSI prp39a vs Col-0


A
2a3 prp39a
Col-0
2a3
2a3 sgs3
2a3 smd1b
NIA2 mRNA
25S
NIA2 siRNA
U6
B
Percentage PTGS
100
80
60
40
20
0
12
20
30
40
58
Days after germination
2a3 (Col-0)
2a3 prp39a
2a3 smd1b
2a3 prp39a smbd1b
2a3 sgs3
C
Log₂ GUS activity (AU)
–5
0
5
10
L1
L1 sgs3
L1 smd1b
L1 prp39a
L1 smd1b prp39a
5
11
17
a
b
D
GUS activity (AU)
0
200
400
600
159
159 sgs3
159 smd1b
159 prp39a-7
159 prp39a-1
159 prp39b-1
8
13
a
b
E
35S pro
NIA2
NIA2ter
gDNA
2a3
2a3 prp39a
35S pro
GUS
35S ter
gDNA
159
159 prp39a
F
Percentage PTGS
0
5
10
15
20
25
Hc1
Hc1 sgs3
Hc1 smd1b
Hc1 prp39a
nd.
nd.
nd.
G
GxA
GxA smd1b
GxA prp39a
H
JAP3
prp39a-7
prp39a-8
I
JAP3
jmj14-4
prp39a-7
prp39a-7
prp39a-8
prp39a-8
smd1b
smd1b
PDS siRNA
-24-nt
-21-nt
miR390


GUS activity (A.U)
1,500
1,000
500
0
bd c ce
ac d ab
ac ad d
a a ab
de c ef
e c f
b bc ac
c ad bd
ac e abc
b be ac
5 11 17
159
5 11 17
159
sgs3
5 11 17
159
smd1b
5 11 17
159
prp39a
5 11 17
159
xrn3
5 11 17
159
hen2
5 11 17
159
prp39a
hen2
5 11 17
159
prp39a
xrn3
5 11 17
159
smd1b
hen2
5 11 17
159
smd1b
xrn3


A
12
10
8
6
4
2
0
Percent input
*
*
*
*
GUS
NPTII
Pre-mRNA 1
Pre-mRNA 2
Pre-mRNA 3
PRP39a-GFP
SmD1b-GFP
IgG
B
Input
Unbound
IP
Line 1
Line 2
IgG
Line 1
Line 2
IgG
Line 1
Line 2
IgG
α - GFP


A
prp39a down DEGs
119
41
56
hen2 up DEGs
B
At1g49700
V43
V42
At5g11090
V9
E40
At1g79720
V113
V54
V54 (spliced)
At3g01500
V64
V63
V64 (mRNA)
V63 (Intron)
At1g61580
V61
V62
V61 (intron)
V62 (mRNA)
At5g25280
A803
A801
A801 (3'ext)
Log2 Fold-change relative to 159
5
0
−5
159
159 hen2
159 prp39a
159 prp39a hen2
V43: ab a b b
V42: b a c c
V9: c a b a
E40: c a b b
V54 (spliced): a a b b
V113: c a b bc
V64 (mRNA): a a a a
V63 (Intron): c a bc ab
V61 (intron): b a b b
V62 (mRNA): a a a b
A803: b a b c
A801 (3'ext): b a b b


Agrobacterium
Transgene
AAA mRNA
AAA
abRNA
Plant genome
PRP39a
XRN3
SmD1b
EXO
Nucleus
Cytosol
RDR6
dsRNA
DCL2/4
siRNA
AGO1
AAA
S-PTGS