Arabidopsis RNASE THREE LIKE2 Modulates the Expression of Protein-Coding Genes via 24-Nucleotide Small Interfering RNA-Directed DNA Methylation

Emilie Elvira-Matelot, Mélanie Hachet, Nahid Shamandi, Pascale Comella, Julio Sáez-Vasquez, Matthias Zytnicki, Herve H. Vaucheret

The Plant Cell, Vol. 28: 406–425, February 2016, www.plantcell.org

Arabidopsis RNASE THREE LIKE2 Modulates the Expression of Protein-Coding Genes via 24-Nucleotide Small Interfering RNA-Directed DNA Methylation[OPEN]

Emilie Elvira-Matelot,[a] Mélanie Hachet,[b] Nahid Shamandi,[a,c] Pascale Comella,[d,e] Julio Sáez-Vásquez,[d,e] Matthias Zytnicki,[b] and Hervé Vaucheret[a,1]

[a] Institut Jean-Pierre Bourgin, UMR 1318, INRA AgroParisTech CNRS, Université Paris-Saclay, 78000 Versailles, France
[b] URGI, INRA, 78000 Versailles, France
[c] Université Paris-Sud, Université Paris-Saclay, 91405 Orsay, France
[d] CNRS, UMR 5096, LGDP, 66860 Perpignan, France
[e] Université de Perpignan Via Domitia, UMR 5096, LGDP, 66860 Perpignan, France

ORCID IDs: 0000-0002-4017-6436 (E.E.-M.); 0000-0002-4643-0642 (N.S.); 0000-0002-2717-7995 (J.S.-V.)

RNaseIII enzymes catalyze the cleavage of double-stranded RNA (dsRNA) and have diverse functions in RNA maturation. *Arabidopsis thaliana* RNASE THREE LIKE2 (RTL2), which carries one RNaseIII and two dsRNA binding (DRB) domains, is a unique Arabidopsis RNaseIII enzyme resembling the budding yeast small interfering RNA (siRNA)-producing Dcr1 enzyme. Here, we show that RTL2 modulates the production of a subset of small RNAs and that this activity depends on both its RNaseIII and DRB domains. However, the mode of action of RTL2 differs from that of Dcr1. Whereas Dcr1 directly cleaves dsRNAs into 23-nucleotide siRNAs, RTL2 likely cleaves dsRNAs into longer molecules, which are subsequently processed into small RNAs by the DICER-LIKE enzymes. Depending on the dsRNA considered, RTL2-mediated maturation either improves (RTL2-dependent loci) or reduces (RTL2-sensitive loci) the production of small RNAs. Because the vast majority of RTL2-regulated loci correspond to transposons and intergenic regions producing 24-nucleotide siRNAs that guide DNA methylation, RTL2 depletion modifies DNA methylation in these regions. Nevertheless, 13% of RTL2-regulated loci correspond to protein-coding genes. We show that changes in 24-nucleotide siRNA levels also affect DNA methylation levels at such loci and inversely correlate with mRNA steady state levels, thus implicating RTL2 in the regulation of protein-coding gene expression.

INTRODUCTION

RNaseIII enzymes are defined by the presence of a highly conserved stretch of nine amino acid residues known as the RNaseIII signature motif, which catalyzes the cleavage of double-stranded RNA (dsRNA) (Blaszczyk et al., 2001; Nicholson, 2014). RNaseIII proteins vary widely in length, from 200 to 2000 amino acids, and have been subdivided into four classes based on the number and nature of the other domains carried by the proteins (Filippov et al., 2000). Class I RNaseIII enzymes are found in bacteria and bacteriophage and present one RNaseIII domain and one dsRNA binding domain (DRB). Class II RNaseIII enzymes differ from class I by the presence of a highly variable N-terminal domain extension and, in some cases, the presence of a second DRB. *Saccharomyces cerevisiae* Rtn1 and *Saccharomyces pombe* Pac1 are well-characterized members of class II RNaseIII enzymes carrying a single DRB domain, which are found in fission yeasts, whereas *Saccharomyces castellii*, *Candida albicans*, and *Kluyveromyces polysporus* Dcr1 exemplify class II RNaseIII enzymes carrying two DRB domains, which are found in budding yeasts. Class III RNaseIII enzymes, also referred to as Drosha, have two RNaseIII and two DRB domains and are found in animals. Finally, class IV RNaseIII enzymes, also referred to as Dicer, display one RNA helicase domain, a PAZ domain, two RNaseIII domains, and one or two DRB domains and are found in fungi, animals, and plants.

Class I and class II RNaseIII enzymes carrying a single DRB domain are involved in the processing of several classes of RNAs, including rRNA, small nuclear RNA, and small nucleolar RNA. Class I and class II RNaseIII enzymes are also involved in mRNA processing, leading to inhibition or activation of gene expression. Inhibition generally occurs through the cleavage of protective stem loop structures in the 5′UTR (untranslated region) (Emory and Belasco, 1990), whereas activation occurs through the cleavage of stem loop structures in the 5′UTR, which occludes ribosome binding sites (Dunn and Studier, 1975; Kameyama et al., 1991). In budding yeasts, the class II RNaseIII Dcr1 enzymes, which carry two DRB domains, also process dsRNA into small interfering RNAs (siRNAs), which modulate gene expression (Drinnenberg et al., 2009). Class III RNaseIII Drosha enzymes are involved in the first step of the processing of microRNA (miRNA) precursors in animals by cleaving the stem-loop structure ~10 nucleotides from its base. Class IV RNaseIII Dicer enzymes perform the second step of the processing of miRNA precursors in animals by cleaving the stem-loop structure ~22 nucleotides from the Drosha cleavage site. Note that in plants, which do not contain

[1] Address correspondence to herve.vaucheret@versailles.inra.fr.
The author 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) is: Hervé Vaucheret (herve.vaucheret@versailles.inra.fr).
[OPEN] Articles can be viewed online without a subscription.
www.plantcell.org/cgi/doi/10.1105/tpc.15.00540

Drosha, all steps of the processing of miRNA precursors are performed by Dicer enzymes. In animals and plants, Dicer enzymes also produce siRNAs from long, perfect dsRNA. In this case, Dicer enzymes recognize 2-nucleotide 3′ overhang structures at the end of dsRNA substrate through their PAZ domain and cleave at the distance spanning the terminus binding PAZ domain and RNaseIII active sites (21 to 24 nucleotides) (Zhang et al., 2004).

In the model plant *Arabidopsis thaliana*, there are four class IV RNaseIII enzymes called DICER-LIKE (DCL). DCL1 produces miRNAs, mostly 21 nucleotides in length, whereas DCL2, DCL3, and DCL4 produce siRNAs of 22, 24, and 21 nucleotides, respectively (Vaucheret, 2006). DCL1-dependent 21-nucleotide miRNAs are processed from the transcripts of MIR (MICRORNA) genes, which adopt a fold-back stem-loop structure and associate with ARGONAUTE1 (AGO1), AGO2, AGO7, or AGO10 to guide cleavage or translation repression of homologous mRNAs (Mallory and Vaucheret, 2010). DCL4-dependent 21-nucleotide siRNAs and DCL2-dependent 22-nucleotide siRNAs derive from dsRNA that originate either from the transcription of long inverted repeats (IRs) or from the action of the RNA-dependent RNA polymerases RDR1 or RDR6 on single-stranded RNA. Endogenous siRNAs that derive from long IR are referred to as endogenous IR-derived siRNAs (endoIR-siRNAs), whereas siRNAs that derive from the action of RDR1 or RDR6 form different classes of siRNA depending on the type of RNA from which they derive. Trans-acting siRNAs (ta-siRNAs) derive from the primary transcripts of non-protein-coding genes referred to as TAS genes. Similar to most miRNAs, 21-nucleotide ta-siRNAs associate with AGO1 to guide cleavage of homologous mRNAs. RNA quality control siRNAs (rqc-siRNAs) and coding transcript siRNAs (ct-siRNAs) derive from the aberrant forms of mRNAs encoded by protein-coding genes, which are converted into dsRNA by cellular RNA-dependent RNA polymerases (RDR) when decapping activity or 5-to-3′ and 3′-to-5′ RNA decay is impaired (Martínez de Alba et al., 2015; Zhang et al., 2015). Virus-activated siRNAs (vasiRNAs) are produced from mRNAs encoded by protein-coding genes, which are converted into dsRNA, mostly by RDR1, when plants are infected by viruses (Cao et al., 2014). Note that rqc-siRNAs, ct-siRNAs, and vasiRNAs are not found in wild-type plants grown under regular conditions. Finally, DCL3-dependent 24-nucleotide siRNAs constitute the most abundant class of endogenous siRNAs (Lu et al., 2006). Their production depends on Polymerase IV (PolIV), so they are usually referred to as p4-siRNAs (Mosher et al., 2008). A large fraction of 24-nucleotide siRNAs also requires RDR2 for the production of the DCL3 dsRNA substrate. p4-siRNAs mostly derive from transposons and repeats, and direct DNA methylation and transcriptional gene silencing (TGS), via the RNA-directed DNA methylation (RdDM) pathway. Paradoxically, TGS requires transcription by both PolIV and PolV at these loci. Briefly, PolIV transcripts are transformed into dsRNA through the action of RDR2. Then, dsRNA is processed by DCL3 into 24-nucleotide siRNAs duplexes. One strand of the duplex is loaded onto AGO4, AGO6, or AGO9 protein, which directs DNA methylation at specific DNA target loci by interacting with PolV-derived scaffold transcripts. Beside miRNAs, endoIR-siRNAs, tas-siRNAs, rqc-siRNA, ct-siRNAs, vasiRNAs, and p4-siRNAs, there are other classes of small RNAs, whose genetic requirements partially overlap with those of several of the above classes. For example, a class of 21- to 22-nucleotide siRNAs was shown to be involved in DNA methylation at a subset of specific genomic loci independent of the RdDM pathway. Instead, this class requires RDR6, DCL2, AGO2, and NERD, a member of the GW repeat protein family, which generally binds to AGO proteins. NERD is thought to bind unmethylated histone H3 lysine 9 at specific genomic target loci and direct DNA methylation via its interaction with AGO2 bound to 21- to 22-nucleotide siRNAs (Pontier et al., 2012).

Beside the four DCLs, the Arabidopsis genome potentially encodes five other RNaseIII enzymes, which are referred to as RNASE THREE LIKE (RTL) proteins. While RTL4 and RTL5 bear only one RNaseIII domain, RTL1, RTL2, and RTL3 carry one RNaseIII domain and one, two, or three conserved DRB domains, respectively. *RTL2* is widely expressed (Comella et al., 2008), as are *RTL4* and *RTL5* (AtGenExpress data). In contrast, *RTL1* is only weakly expressed in roots (Comella et al., 2008), and the expression of *RTL3* has not been detected in any tissue tested, suggesting that it is a pseudogene (Comella et al., 2008). RTL1 is the most characterized of these five enzymes. Indeed, in vivo and in vitro assays revealed that RTL1 cleaves near-perfect dsRNA before they are processed by the DCLs, resulting in decreased siRNA accumulation at 6089 out of 6102 siRNA-producing genomic loci when *RTL1* is constitutively expressed, whereas miRNA levels remain unchanged. *RTL1* expression is induced in leaves upon virus infection, suggesting that it could participate in plant antiviral defenses by cleaving viral dsRNA intermediates of replication. However, RTL1 activity is inhibited by viral suppressors of RNA silencing, indicating that viruses use the same toolbox to counteract both RNA silencing and RTL1 activity (Shamandi et al., 2015). Little is known about the four other RTL proteins. RTL2 processes, both in vivo and in vitro, the 3′external transcribed spacer (ETS) from ribosomal *45S* pre-rRNA (Comella et al., 2008). The targets of RTL4 are not known, but mutants defective in RTL4 are impaired in male and female gametophyte formation (Portereiko et al., 2006). RTL5 likely is the ortholog of maize (*Zea mays*) RNC1, which is required for the splicing of several chloroplast group II introns (Watkins et al., 2007). Structurally, RTL2 is very similar to the class II RNaseIII found in yeasts, and it shares some of their functional properties. Indeed, RTL2 processes ribosomal *45S* pre-rRNA, similar to Rnt1 in *S. cerevisiae* (Elela et al., 1996) or Dcr1 in *C. albicans* (Bernstein et al., 2012). Interestingly, RTL2 has two DRB domains, similar to *S. castellii*, *C. albicans*, or *K. polysporus* Dcr1 proteins, which are able to process dsRNAs substrates into 22- to 23-nucleotide siRNAs, although they lack PAZ and RNA helicase domains of canonical Dicer enzymes (Drinnenberg et al., 2009; Weinberg et al., 2011).

Given that RTL2 enhances the production of exogenous siRNAs when transiently overexpressed in plants (Kiyota et al., 2011), we investigated whether RTL2 could have an endogenous function in the processing of siRNAs. Unlike Dcr1, RTL2 is not able to produce siRNAs by itself. Rather, it processes certain dsRNA into molecules larger than 24 nucleotides, which are subsequently diced into siRNAs by DCL enzymes. Sequencing of the small RNA repertoire of an Arabidopsis *rtl2* loss-of-function mutant revealed a modified accumulation of siRNAs at 481 discrete loci. The vast majority of these loci produce DCL3-dependent 24 nucleotides, resulting in DNA methylation changes in the *rtl2* mutants. As

expected for loci targeted by RdDM, loci regulated by RTL2 mostly consist of transposons and nonannotated intergenic regions (IGRs). However, 13% of these loci correspond to protein-coding genes, and we show that in these cases, mRNA levels inversely correlate with DNA methylation levels, indicating that RTL2 modulates the expression of a subset of protein-coding genes. Whereas *rtl2* mutants exhibit a short root phenotype, a wild-type root phenotype is restored in the *rtl2 nrpd1a* double mutant, indicating that in wild-type plants, RTL2 prevents the production of 24-nucleotide siRNAs that trigger DNA methylation and silencing of genes involved in the regulation of root development.

RESULTS

RTL2 Cannot Substitute for DCL2, DCL3, or DCL4

To determine if, like budding yeast Dcr1, Arabidopsis RTL2 produces siRNAs by itself, we tested if RTL2 overexpression could compensate the deficiency in certain DCL enzymes. At first, a *ProUBQ10:RTL2* construct (carrying *RTL2* driven by the *UBIQUITIN10* promoter) was introduced into a *dcl2 dcl4* double mutant carrying the *L1* locus. The *L1* locus consists of a *Pro35S:GUS* transgene that spontaneously undergoes sense transgene-induced posttranscriptional gene silencing (S-PTGS) in the wild-type background. As a result, GUS activity is null, *GUS* mRNA is degraded, and 21- and 22-nucleotide *GUS* siRNAs accumulate in *L1* plants, due to the action of DCL4 and DCL2, respectively. In *dcl2* and *dcl4* single mutants, S-PTGS still functions, but in a *dcl2 dcl4* double mutant, high levels of GUS activity are detected, *GUS* mRNA accumulates, and *GUS* siRNAs are absent (Parent et al., 2015). *L1/dcl2 dcl4/ProUBQ10:RTL2* transformants exhibited levels of GUS activity similar to that of *L1/dcl2 dcl4* untransformed controls (Figure 1A), suggesting that RTL2 cannot substitute for DCL2 or DCL4. In a second round, the *ProUBQ10:RTL2* construct was introduced into a *dcl3* mutant carrying the T+S transgenes (Kanno et al., 2008). In T+S plants, the target transgene (T) expresses *GFP* under the control of an upstream enhancer while the silencer transgene (S) expresses the enhancer dsRNA under the control of the *35S* promoter. In the absence of the S transgene, the T transgene expresses *GFP* in root apical meristems. Transcriptional gene silencing of *GFP* expression due to the S transgene is associated with DNA methylation at and downstream of the T enhancer. In a *dcl3* mutant, the T+S loci do not produce 24 nucleotides, lack DNA methylation, and express *GFP* (Daxinger et al., 2009). *T+S/dcl3/ProUBQ10:RTL2* transformants expressed *GFP* similar to *T+S/dcl3* plants (Figure 1B), suggesting that RTL2 cannot substitute for DCL3.

The observation that ectopic RTL2 expression cannot complement *dcl* mutants could be due either to the incapacity of RTL2 to cleave long dsRNA or to a RTL2-mediated cleavage into RNA molecules that are not 21, 22, or 24 nucleotides long. To address this question, RNA gel analyses were performed on *L1/dcl2 dcl4/ProUBQ10:RTL2* and *T+S/dcl3/ProUBQ10:RTL2* transformants. No 21-, 22-, or 24-nucleotide siRNAs were detected in *L1/dcl2 dcl4/ProUBQ10:RTL2* (Figure 1C). Moreover, no 24-nucleotide siRNAs were detected in *T+S/dcl3/ProUBQ10:RTL2* transformants, while the level of 21- and 22-nucleotide siRNAs produced by DCL4 and DCL2 remained unchanged (Figure 1D). Together, these results indicate that ectopic RTL2 expression cannot restore S-PTGS in a *dcl2 dcl4* double mutant or TGS in a *dcl3* mutant, likely because RTL2 cannot produce 21-, 22-, or 24-nucleotide siRNAs by itself. Nevertheless, most *T+S/dcl3/ProUBQ10:RTL2* transformants accumulated an RNA molecule that is larger than an siRNA (see blue arrow in Figure 1D), suggesting that RTL2 could process dsRNA into molecules >24 nucleotides.

RTL2 Modulates the Production of siRNAs from dsRNA through Its RNaseIII and DRB Domains

The results presented above suggest that RTL2 does not produce 21-, 22-, or 24-nucleotide siRNAs per se but rather process dsRNA into molecules >24 nucleotides. Whether these molecules could be subsequently processed into 21-, 22-, or 24-nucleotide siRNAs by the DCLs is not known. However, Kiyota et al. (2011) showed that coinfiltration of *Pro35S:RTL2* and *Pro35S:GFP* constructs in leaves of the *Nicotiana benthamiana* 16c line, which carries a *Pro35S:GFP* transgene that is prone to trigger S-PTGS, enhanced the production of *GFP* siRNAs. To further analyze the function of RTL2, a *Pro35S:IR-GUS* construct carrying an inverted repeat of the 5′ end of *GUS* was infiltrated in wild-type *N. benthamiana* leaves together with *Pro35S:GFP* or *Pro35S:RTL2* constructs. The levels of 21- and 24-nucleotide *GUS* siRNAs produced by the *Pro35S:IR-GUS* construct were higher in *Pro35S:IR-GUS* + *Pro35S:RTL2* infiltrated leaves compared with *Pro35S:IR-GUS* + *Pro35S:GFP* infiltrated leaves (Figure 2B), suggesting that RTL2 stimulates the production of siRNAs during IR-PTGS. Note that an RNA product larger than a siRNA reproducibly accumulated in *Pro35S:IR-GUS* + *Pro35S:RTL2* infiltrated leaves (see blue arrow in Figures 2B to 2D), similar to the product that accumulates in *T+S/dcl3/ ProUBQ10:RTL2* transformants (see blue arrow in Figure 1D). Therefore, it is likely that RTL2 processes dsRNA into molecules larger than siRNAs and that these larger molecules are subsequently processed into 21-, 22-, or 24-nucleotide siRNAs by the DCLs.

To determine if an intact RNaseIII domain is required for RTL2 effect on siRNA production, the conserved amino acids Glu-93 and Asp-100, which are integral of the conserved catalytic site of RNaseIII enzymes and required for RNaseIII activity (Blaszczyk et al., 2001), were both mutagenized to Ala, thus creating the *RTL2mR3* construct (Figure 2A). Then, *N. benthamiana* leaves were coinfiltrated with *Pro35S:IR-GUS* and either *Pro35S:GFP*, *Pro35S:RTL2*, or *Pro35S:RTL2mR3*. The increased accumulation of 21- and 24-nucleotide siRNAs and the larger molecule observed in *Pro35S:IR-GUS* + *Pro35S:RTL2* infiltrated leaves were not found in *Pro35S:IR-GUS* + *Pro35S:RTL2mR3* infiltrated leaves (Figure 2C), indicating that a functional RNaseIII domain is required for RTL2 activity.

To further determine if the two DRB domains are required for stimulating siRNA production, the second DRB domain was deleted, thus creating the *RTL2ΔDRB* construct (Figure 2A). Then, *N. benthamiana* leaves were coinfiltrated with *Pro35S:IR-GUS* and either *Pro35S:GFP*, *Pro35S:RTL2*, or *Pro35S:RTL2ΔDRB*. The increased accumulation of 21- and 24-nucleotide siRNAs and the larger molecule observed in *Pro35S:IR-GUS* + *Pro35S:RTL2* infiltrated leaves were not found in *Pro35S:IR-GUS* + *Pro35S:RTL2ΔDRB* infiltrated leaves (Figure 2D), indicating that the second DRB domain is required for RTL2 activity.

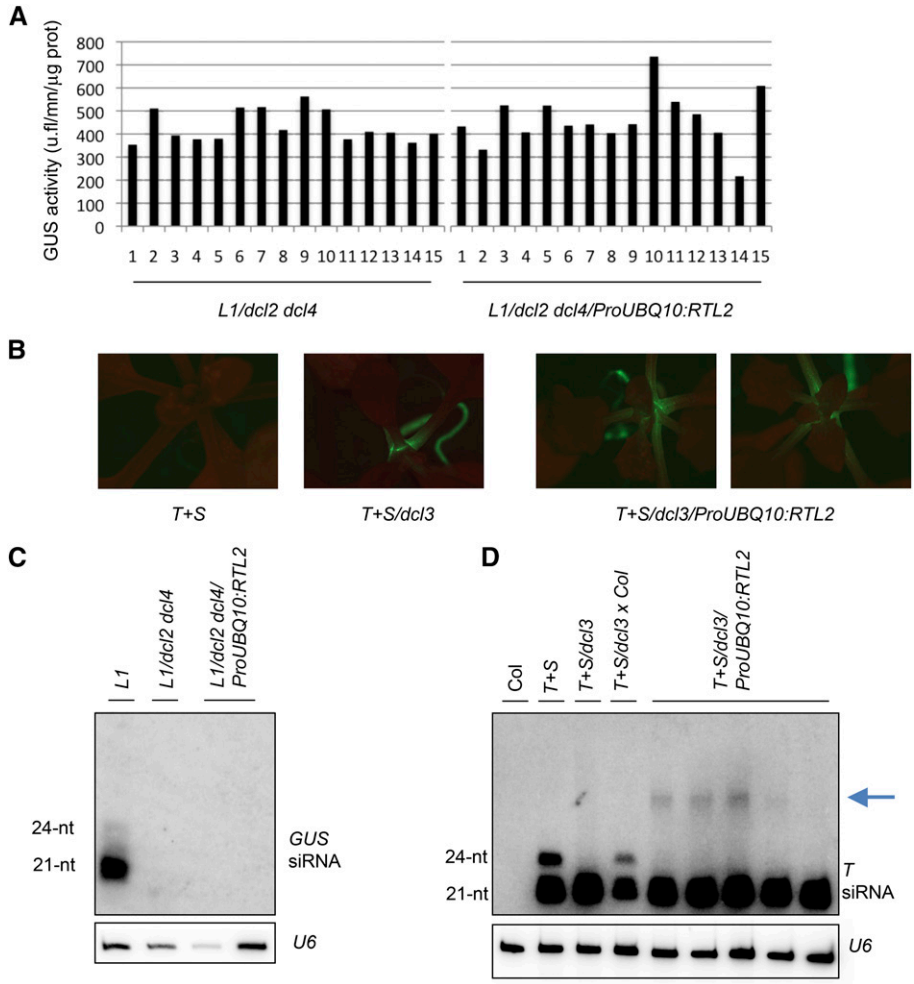

Figure 1. RTL2 Cannot Substitute for DCL2, DCL3, or DCL4.

(A) GUS activity in 15 *L1/dcl2 dcl4* untransformed controls and 15 independent *L1/dcl2 dcl4/ProUBQ10:RTL2* transformants. GUS activity is expressed as units of fluorescence/min/µg protein.

(B) GFP visualization in *T+S* and *T+S/dcl3* controls and *T+S/dcl3/ProUBQ10:RTL2* transformants using a fluorescent zoom microscope.

(C) Gel blots of total RNA from plants of the indicated genotypes were hybridized with the indicated probes. nt, nucleotides.

(D) Gel blots of total RNA from plants of the indicated genotypes were hybridized with the indicated probes. The blue arrow indicates putative RTL2 products >24 nucleotides. *U6* was used as a loading control.

To ensure that the absence of RTL2 activity of the *Pro35S:RTL2mR3* and *Pro35S:RTL2ΔDRB* constructs was not due to destabilization or impaired production of the mutant proteins, an epitope tag was added to the constructs to compare RTL2, RTL2mR3, and RTL2ΔDRB protein accumulation levels in *N. benthamiana* infiltrated leaves. At first, a Myc tag was added to the N terminus or C terminus of the wild-type RTL2 protein, and *Pro35S:Myc-RTL2* or *Pro35S:RTL2-Myc* constructs were individually infiltrated in *N. benthamiana* leaves. Protein gel blot analysis using anti-Myc antibodies revealed that the RTL2-Myc protein accumulates, whereas the Myc-RTL2 protein does not accumulate (Figure 2E), suggesting that modifying the RTL2 N terminus destabilizes the protein. Then, the *Pro35S:RTL2-Myc* construct was mutagenized at the RNaseIII site or deleted from the second DRB domain to produce the *Pro35S:RTL2mR3-Myc* and *Pro35S:RTL2ΔDRB-Myc* constructs, respectively. Protein gel blot analysis of *N. benthamiana* infiltrated leaves revealed that neither the mutation in the RNaseIII domain nor the deletion of the second DRB domain prevents the production of a protein (Figures 2E and 2F). Together, these results indicate that the absence of RTL2 activity of the *Pro35S:RTL2mR3* and *Pro35S:RTL2ΔDRB* constructs is due to the mutations imposed to these essential domains of the protein and not to destabilization or impaired production of the mutant proteins.

RTL2 Effect Likely Depends on dsRNA Sequence and/or Structure

The results presented in Figure 2B and the results by Kiyota et al. (2011) suggest that RTL2 ectopic expression stimulates the

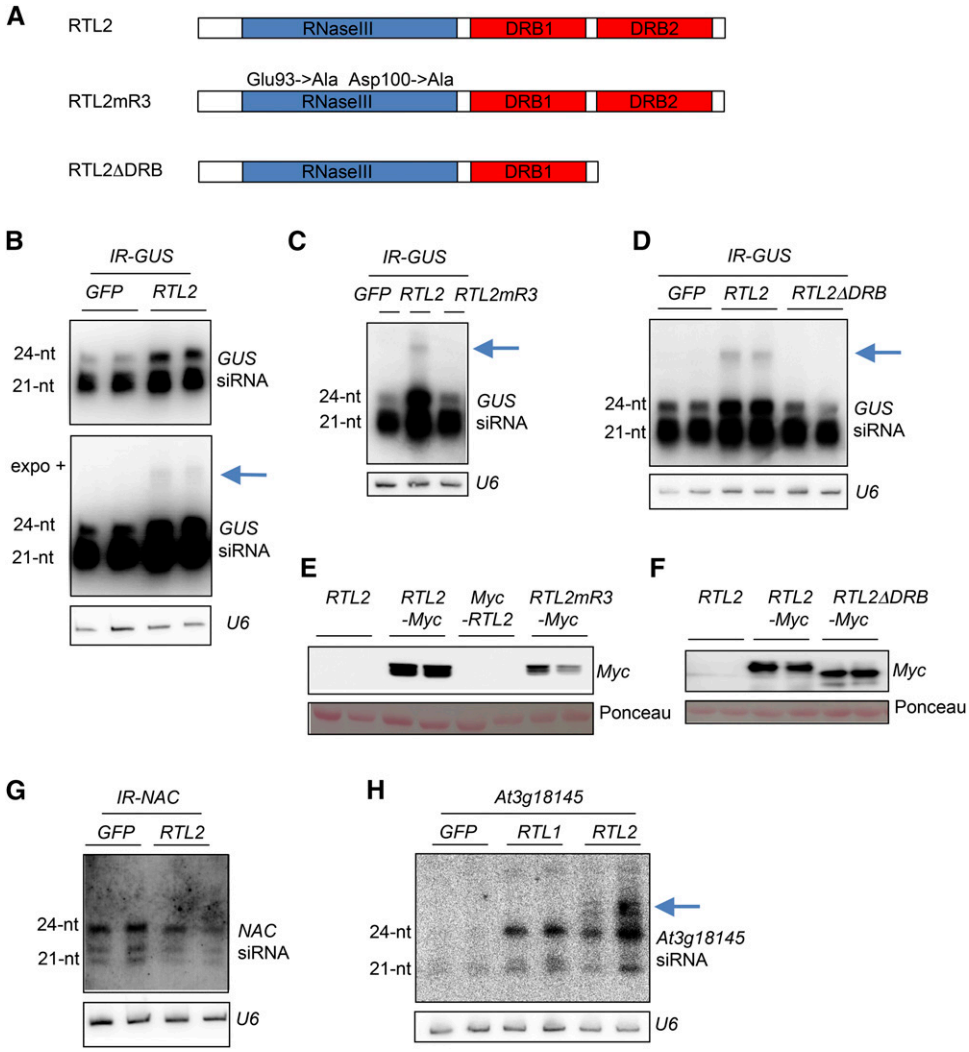

Figure 2. RTL2 Modulates siRNA Production from dsRNA Substrates through Its RNaseIII and Its DRB Domains.

(A) Schematic representation of intact RTL2, RTL2mR3 carrying two mutations in the catalytic site of the RNaseIII domain, and RTL2ΔDRB deleted from the second DRB domain.

(B) Gel blots of total RNA from *N. benthamiana* leaves coinfiltrated with the *Pro35S:IR-GUS* construct and either *Pro35S:GFP* or *Pro35S:RTL2*. Two different exposures of the same blot are shown to visualize the presence of >24-nucleotide RNA products. nt, nucleotides.

(C) Gel blots of total RNA from *N. benthamiana* leaves coinfiltrated with the *Pro35S:IR-GUS* construct and either *Pro35S:GFP*, *Pro35S:RTL2*, or *Pro35S:RTL2mR3*.

(D) Gel blots of total RNA from *N. benthamiana* leaves coinfiltrated with the *Pro35S:IR-GUS* construct and either *Pro35S:GFP*, *Pro35S:RTL2*, or *Pro35S:RTL2ΔDRB*.

(E) and **(F)** Gel blots of total proteins from *N. benthamiana* leaves coinfiltrated with *Pro35S:IR-GUS* and either *Pro35S:GFP*, *Pro35S:RTL2-myc*, *Pro35S:myc-RTL2*, or *Pro35S-RTL2mR3-myc* **(E)**, or *Pro35S:RTL2ΔDRB-myc* **(F)**. Membranes were incubated with an anti-Myc antibody. Ponceau staining serves as a loading control.

(G) and **(H)** Gel blots of total RNA from *N. benthamiana* leaves coinfiltrated with the *Pro35S:IR-NAC* or *Pro35S:At3g18145* construct and either *Pro35S:GFP*, *Pro35S:RTL1*, or *Pro35S:RTL2*. Blots were hybridized with the indicated probes and with *U6* as a loading control. The blue arrows indicate putative RTL2 products >24 nucleotides.

production of siRNAs from GUS and GFP sequences. However, RTL2 ectopic expression in *T+S/dcl3/ProUBQ10:RTL2* transformants did not affect the level of 21- and 22-nucleotide siRNAs (Figure 1D), suggesting that the stimulating effect of RTL2 on siRNA production may not be general. To address this question, the *Pro35S:RTL2* construct was coinfiltrated with another IR construct (*Pro35S:IR-NAC*), which carries a fragment of the *NAC52* transcription factor gene. Contrasting the effect of RTL2 on the *Pro35S:IR-GUS* construct, RTL2 overexpression reduced the production of siRNAs from the *Pro35S:IR-NAC* construct (Figure 2G). Given the different effects of RTL2 on different artificial IR constructs, we decided to test the effect of RTL2 on natural IRs.

At first, we used the Arabidopsis *MIR168a* and *MIR168b* genes, in which hairpins are efficiently processed into 21- and/or 22-nucleotide miR168 molecules. Infiltration of *Pro35S:MIR168a* or *Pro35S:MIR168b* constructs into *N. benthamiana* leaves resulted in the production of miR168 species, whose levels were not affected by coexpressing *Pro35S:RTL2* (Supplemental Figure 1). Then, we used the *At3g18145* gene, whose 3′UTR forms an almost perfect hairpin that is inefficiently processed by DCL1 into a single 21-nucleotide molecule (referred to as miR3440b), but which is more efficiently processed by DCL3 into a single 24-nucleotide siRNA upon cleavage of the *At3g18145* 3′UTR by RTL1 (Shamandi et al., 2015). Coinfiltration of a *Pro35S:At3g18145* construct with either *Pro35S:RTL1* or *Pro35S:RTL2* resulted in the production of an abundant 24-nucleotide species, which was not detected when infiltrating the *Pro35S:At3g18145* construct alone (Figure 2H). Two RNA molecules of larger sizes (see blue arrow) were detected in plants coinfiltrated with *Pro35S:RTL2*, but not *Pro35S:RTL1*, suggesting that RTL2 processes the *At3g18145* 3′UTR into molecules >24 nucleotides, which are subsequently processed into 24-nucleotide siRNAs, likely by DCL3. These results also indicate that RTL1 and RTL2 do not cleave *At3g18145* 3′UTR in the same way.

Because the various IR constructs tested produce dsRNA molecules that differ in sequence and structure, it is likely that the effect of RTL2 (stimulation or reduction of siRNA production) depends on the sequence or structure of its dsRNA substrates. This result and the observation that the accumulation of all sizes of siRNAs is affected by RTL2 suggest that RTL2 does not directly produce siRNAs from dsRNA. Rather, it suggests that RTL2 cleaves certain dsRNAs into molecules larger than 24 nucleotides and that these molecules are either better or worse substrates than the original dsRNA molecules for the endogenous DCLs.

Impairment of RTL2 Affects the Accumulation of siRNAs from the 45S rRNA Genes

The results presented above indicate that RTL2 overexpression modulates the processing of dsRNA precursors into siRNAs by the DCLs. However, they do not indicate if cellular RTL2 levels are critical for endogenous small RNA (sRNA) production. RNA gel blot analysis using a few RNA probes revealed that the amount of siRNAs derived from the intergenic spacer of the 45S pre-rRNA is reduced in *rtl2* whereas levels of a representative miRNA (miR159), a representative ta-siRNA (TAS4), and a representative PolIV-siRNA (SN1) are unchanged (Figure 3A), suggesting that RTL2 plays limited role in the production of endogenous sRNAs. To further examine the contribution of RTL2 to the sRNA repertoire, the sRNA population of an Arabidopsis *rtl2* null allele (Comella et al., 2008) was sequenced and compared with that of wild-type (Col) plants. A *dcl2 dcl3 dcl4* triple mutant was also sequenced as a control. Sequencing was done in triplicate. The characteristics of each library are shown in Supplemental Table 1.

Because RTL2 is required for cleavage of the 3′ETS of the 45S pre-rRNA (Comella et al., 2008) and because RNA gel blot analysis indicated a reduction of 45S siRNAs (Figure 3A), we first examined 45S siRNA accumulation. The 45S rRNA genes (encoding 18S, 5.8S, and 25S rRNA) are arranged in tandem arrays separated by an intergenic spacer (IGS). Not all rRNA genes are active. Selective activation or silencing of individual rRNA genes is regulated by epigenetic mechanisms, and DCL3-dependent siRNAs match the 45S IGS and promoter rDNA sequences to maintain the silencing state of inactive rDNA genes (Preuss et al., 2008; Tucker et al., 2010). Analysis of the size distribution of 17- to 30-nucleotide reads matching the 45S 3′ETS, IGS, and 5′ETS sequences (Supplemental Figure 2) revealed decreased accumulation of siRNA of all sizes in *rtl2* compared with Col, although not to the level observed in the *dcl2 dcl3 dcl4* triple dicer mutant (Figure 3B). Analysis of the genomic distribution of 17- to 30-nucleotide reads matching the 3′ETS, IGS, and 5′ETS revealed peaks of siRNAs near the gene promoter (GP), the spacer promoter 1 (SP1), and SP2, and within the 5′ETS, in particular close to the 18S rRNA (Figure 3C). The abundance of siRNAs from these various regions was strongly decreased in *rtl2* compared with Col plants, suggesting that the absence of cleavage of the 45S pre-rRNA by RTL2 globally reduced siRNA production.

Impairment of RTL2 Affects the Accumulation of Endogenous sRNAs at Discrete Loci

Then, we examined the distribution of sRNAs that perfectly match the Arabidopsis genome excluding rRNA and tRNA. No major difference in the global size distribution of total reads was observed between *rtl2* and Col, whereas *dcl2 dcl3 dcl4* exhibited an expected decreased accumulation of 22-, 23-, and 24-nucleotide sRNAs (Figure 4A). To further examine if RTL2 could have an effect at specific loci, the accumulation of sRNAs that match a unique position in the genome was analyzed using 100-bp sliding windows (see Methods). Sliding windows showing an at least 2-fold ratio between Col and *rtl2* were identified (Supplemental Data Set 1). Overlapping differential sliding windows were assembled into loci. A total of 463 differential loci were identified with a P value < 5E-2, including 175 loci showing a Col/*rtl2* ratio superior to two (subsequently referred to as RTL2-dependent loci) and 288 loci showing a *rtl2*/Col ratio superior to two (subsequently referred to as RTL2-sensitive loci). Among the 175 RTL2-dependent loci, 119 showed a fold ratio between Col and *rtl2* greater than five, including 50 loci that show a fold change greater than 50, pointing to an essential role of RTL2 in the production of sRNAs from these loci. Similarly, 144 out of the 288 RTL2-sensitive loci showed a fold ratio between *rtl2* and Col greater than five, including 58 loci that show a fold change greater than 50, confirming the importance of RTL2 in regulating the production of small RNA at discrete loci.

Analysis of the genomic distribution and density of reads around the differential loci revealed several types of situations (Figure 5). For the majority of the 463 loci, the differential region coincided with the entire siRNA-producing region (see a representative RTL2-dependent locus in Figure 5A and a representative RTL2-sensititive locus in Figure 5B). For the remaining loci, the differential region was adjacent to a nondifferential region or to a differential region that varied in the opposite direction (see a representative RTL2-dependent locus adjacent to a nondifferential region in Figure 5C, a representative RTL2-sensititive locus adjacent to a nondifferential region in Figure 5D, and representative RTL2-dependent loci adjacent to RTL2-sensititive loci in Figures 5E and 5F). To explain these results, one can imagine that adjacent regions that behave differently toward RTL2 define distinct transcription units and that the resulting RNAs are either

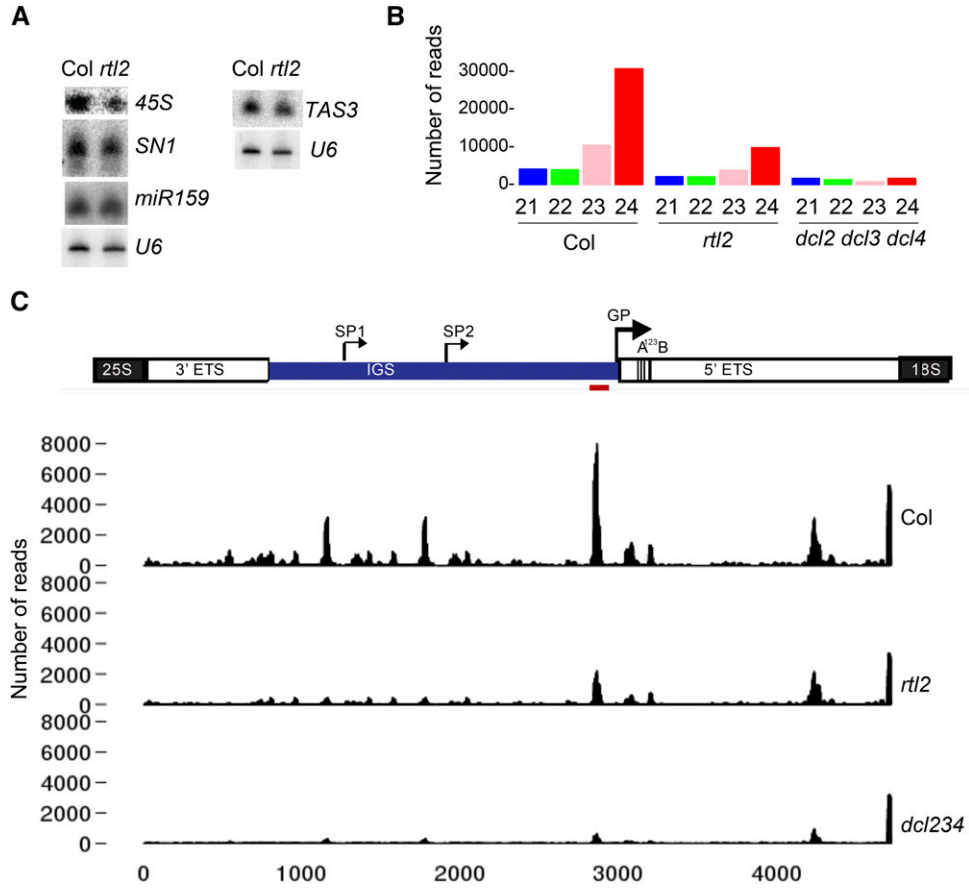

Figure 3. RTL2 Impairment Affects the Accumulation of 45S siRNAs.

(A) RNA gel blot analysis of small RNAs. Total RNA from leaves of 18-d-old seedlings of wild-type (Col) and *rtl2* plants were hybridized with the indicated probes. U6 was used as a control.

(B) Normalized size distribution of 17- to 30-nucleotide sRNAs matching the 3′ETS, IGS, and 5′ETS of the 45S rDNA in Col, *rtl2*, and *dcl2 dcl3 dcl4* plants. Sizes are represented by different colors (listed in order): 21 nucleotides (blue), 22 nucleotides (green), 23 nucleotides (pink), and 24 nucleotides (red).

(C) Normalized abundance of siRNAs spanning the 3′ETS, IGS, and 5′ETS of the 45S rDNA in Col, *rtl2*, and *dcl2 dcl3 dcl4* plants. SP1, SP2, GP, and $A^{123}B$ sequences are defined in Supplemental Figure 2. The red bar corresponds to the probe used in **(A)**.

cleaved by RTL2 (RTL2-dependent loci and RTL2-sensititive loci) or not (nondifferential loci). Alternatively, it is possible that these adjacent regions define a single transcriptional unit but that the different regions do not produce the same amount of small RNAs depending on whether the original transcribed RNA is cleaved or not by RTL2 before processing by the DCLs.

If the latter hypothesis holds true, the sliding window method used previously could have missed differential loci if they correspond to small regions that are closely adjacent to regions producing large amounts of small RNAs or if reduced and increased numbers of reads within the window compensate for each other so that the number of reads is similar in Col and *rtl2*. Therefore, we developed a method based on the analysis of reads at each position of the genome (see Methods). This method identified 18 additional loci, including eight RTL2-dependent loci and 10 RTL2-sensitive loci (Supplemental Data Set 1). As predicted, these loci were short and surrounded or adjacent to nondifferential regions producing large amounts of small RNAs that masked these differential loci in the sliding window analysis (see representative loci in Figures 5G and 5H). Overall, 481 differential loci, including 183 RTL2-dependent loci and 298 RTL2-sensitive loci were identified.

Of the 481 loci producing siRNAs that are regulated by RTL2, 252 correspond to annotated transposable elements, 61 correspond to annotated protein-coding genes (PCGs), and five correspond to annotated non-protein-coding genes (non-PCG), including one pseudogene, three MIR genes, and one gene producing a noncoding RNA of unknown function (Supplemental Table 2). The remaining 163 loci correspond to nonannotated regions, hereafter referred to as IGRs. The respective proportions of each class were not different between RTL2-dependent and RTL2-sensitive loci (Supplemental Table 2; Figure 4B).

RTL2 Does Not Produce siRNAs but Rather Processes dsRNA Substrates of the DCLs

The results of transient assays suggested that RTL2 does not directly produce siRNAs from dsRNA molecules, but rather it

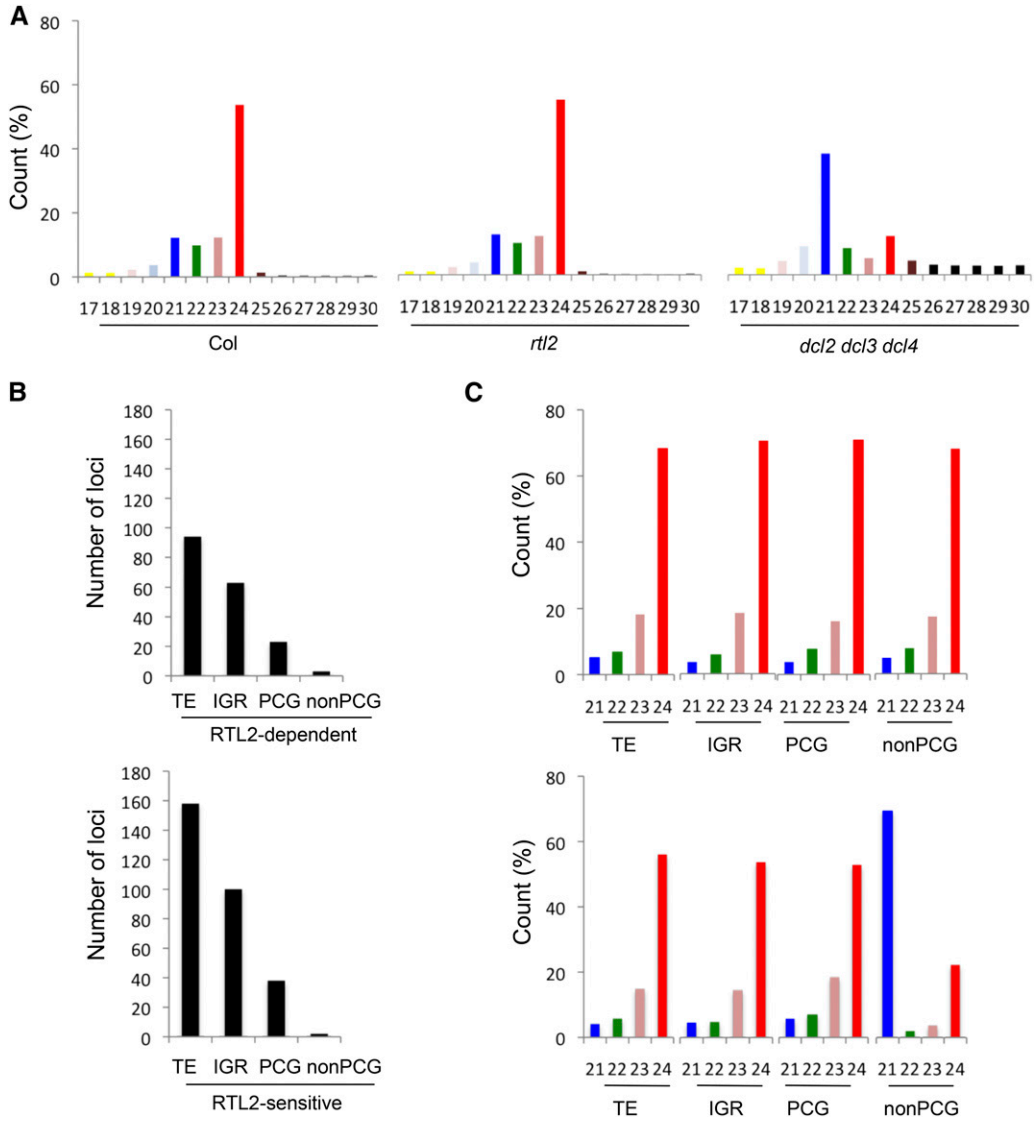

Figure 4. RTL2 Affects the Accumulation of Endogenous siRNAs at Discrete Loci.

Small RNAs from flowers of the wild type (Col), *rtl2*, and *dcl2 dcl3 dcl4* mutants were subjected to high-throughput sequencing.

(A) Size distribution and number of reads that perfectly match the Arabidopsis nuclear genome in Col, *rtl2*, or *dcl2 dcl3 dcl4*, excluding rRNA and tRNA. Color code (listed in order): <19 nucleotides (yellow), 19 nucleotides (light pink), 20 nucleotides (light blue), 21 nucleotides (blue), 22 nucleotides (green), 23 nucleotides (pink), 24 nucleotides (red), 25 nucleotides (dark red), and >25 nucleotides (black).

(B) Number of RTL2-dependent and RTL2-sensitive loci matching annotated transposable elements (TE), nonannotated IGRs, annotated PCGs, or annotated non-PCGs.

(C) Size distribution and number of reads in the different categories of RTL2-dependent (upper part) and RTL2-sensitive (lower part) loci. Color code is as in **(A)**. For **(B)** and **(C)**, data are taken from the wild-type Col libraries.

cleaves dsRNAs before they are processed by the DCLs, thus generating shorter molecules that are either better or less efficiently processed by the DCLs than the original dsRNA molecules. To determine if RTL2 actually acts upstream of the DCLs, the accumulation of RTL2-dependent siRNAs was examined in the *dcl2 dcl3 dcl4* mutant. Ninety-nine percent of the 183 RTL2-dependent loci produced siRNAs whose accumulation was reduced in *dcl2 dcl3 dcl4* (Figure 6A; Supplemental Data Set 1), indicating that these siRNAs actually require DCL2, DCL3, or DCL4 for their production. In the case of RTL2-sensitive loci, only those producing at least 10 reads per million in Col were considered. Ninety-four percent of these loci exhibited reduced siRNA accumulation in *dcl2 dcl3 dcl4* (Figure 6A; Supplemental Data Set 1). Therefore, it is likely that RTL2 acts upstream of DCL2, DCL3, and DCL4 at both RTL2-dependent and RTL2-sensitive loci. Given that RTL2 carries functional RNaseIII and DRB domains (Figure 2), it is likely that RTL2 processes certain dsRNAs before they are diced into siRNAs by the DCLs. Depending on the

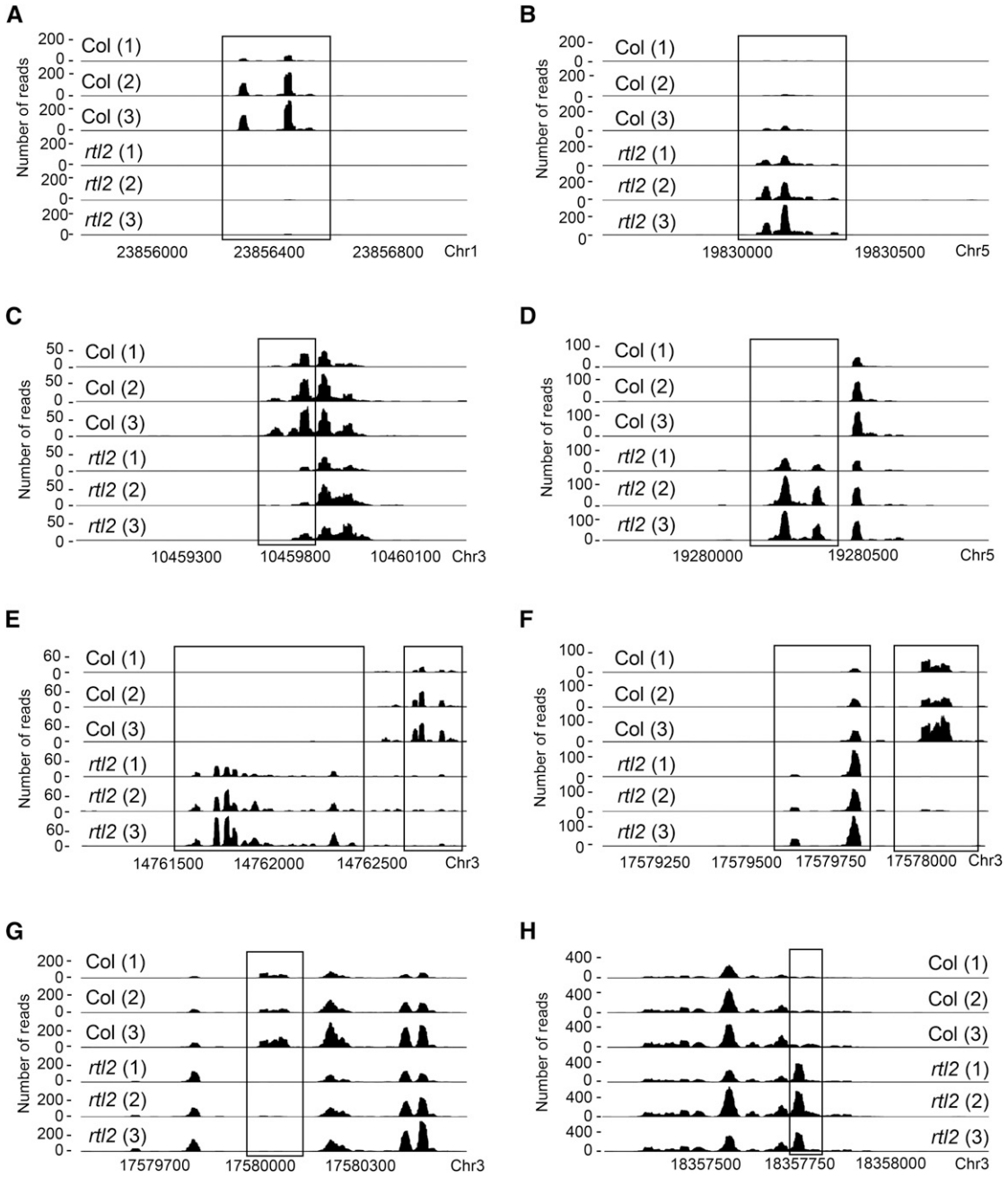

Figure 5. Abundance and Distribution of siRNAs at Representative RTL2-Dependent and RTL2-Sensitive Loci.

(A) Representative RTL2-dependent locus corresponding to the entire siRNA-producing region.
(B) Representative RTL2-sensitive locus corresponding to the entire siRNA-producing region.
(C) Representative RTL2-dependent locus adjacent to a nondifferential region.
(D) Representative RTL2-sensitive locus adjacent to a nondifferential region.
(E) and **(F)** Representative RTL2-dependent loci adjacent to RTL2-sensitive loci.
(G) Representative RTL2-dependent locus surrounded by nondifferential regions identified by the single position method.
(H) Representative short RTL2-sensitive locus adjacent to a nondifferential region identified by the single position method. The differential regions are boxed.

dsRNA considered, RTL2-mediated maturation either improves (RTL2-dependent loci) or reduces (RTL2-sensitive loci) their processing by the DCLs.

Most RTL2-Dependent and RTL2-Sensitive Loci Produce 24-Nucleotide siRNAs through PolIV

Taken as a whole, the 481 loci regulated by RTL2 produce small RNAs that are predominantly 24 nucleotides long (Figure 4C). Such predominance of 24-nucleotide siRNAs is observed when considering either RTL2-dependent or RTL2-sensitive (Figure 4C) and is consistent with transposons and IGR representing 86% of the loci regulated by RTL2 (Figure 4B). Interestingly, PCG and non-PCG loci regulated by RTL2 also predominantly produce 24-nucleotide siRNAs, and only two RTL2-sensitive MIR genes produce 21-nucleotide miRNAs (Figure 4C; Supplemental Table 2)

Because the vast majority of endogenous 24-nucleotide siRNAs require PolIV for their production, the dependence of RTL2-dependent and RTL2-sensitive siRNAs toward PolIV was examined in public databases (Law et al., 2013). Ninety-nine percent of the RTL2-dependent loci and 99.4% of the RTL2-sensitive loci produced siRNAs whose accumulation was reduced in the *nrpd1* mutant (Figure 6B; Supplemental Data Set 2), indicating that almost all loci regulated by RTL2 produce 24-nucleotide siRNAs via PolIV.

RTL2 Likely Cleaves a Subset of PolIV-RDR2-Dependent dsRNAs

It has long been known that 21-nucleotide ta-siRNAs are produced by DCL4-mediated sequential processing of long dsRNA precursors generated by RDR6 (Allen et al., 2005; Yoshikawa et al., 2005). In contrast, it has only been recently revealed that single 24-nucleotide siRNAs are produced by DCL3-mediated processing of a single short precursor, 27 to 50 nucleotides in length, generated by PolIV and RDR2 and referred to as P4R2 RNA (Blevins et al., 2015; Zhai et al., 2015). Because 24-nucleotide siRNAs coincide with P4R2 RNAs at either their 5′ or 3′ end, these short precursors likely are cut once by DCL3 to generate a single siRNA. Although this model explains the biogenesis of most 24-nucleotide siRNAs, the authors wondered if a yet to be identified RNaseIII was at play between the biosynthesis of some P4R2 RNAs by PolIV and RDR2 and their processing by DCL3 (Zhai et al., 2015). At the whole-genome level, reads 27 to 50 nucleotides in length, which likely are P4R2 RNAs, overaccumulate in *dcl234*, but not in *rtl2* (Figure 4A), ruling out that all the reads >26 nucleotides detected in *dcl234* could be RTL2 products. Rather, the finding that RTL2 regulates the production of siRNAs at only a subset of PolIV loci and sometimes within a limited region of PolIV loci (Figure 5) is more likely explained by RTL2 recognizing and cleaving a subset of P4R2 RNAs, resulting in better or worse processing by DCL3 than uncut P4R2 RNAs. To test this hypothesis, the number of reads of mature 24-nucleotide siRNAs and 27- to 50-nucleotide precursor RNAs derived from RTL2-dependent and RTL2-sensitive loci were counted in Col and *rtl2*, and the precursor/siRNA ratio was calculated (Figure 7A). Remarkably, the precursor/siRNA ratio at RTL2-dependent loci was lower in Col than in *rtl2* (0.4% versus 0.64%), whereas the precursor/siRNA ratio at RTL2-sensitive loci was higher in Col than in *rtl2* (0.3% versus 0.18%). These results therefore suggest that precursors derived from RTL2-dependent loci are better processed in Col than in *rtl2*, whereas precursors derived from RTL2-sensitive loci are better processed in *rtl2* than in Col, strongly supporting our hypothesis. To further test the hypothesis that P4R2 RNAs derived from RTL2-dependent and RTL2-sensitive loci are cleaved by RTL2, we analyzed base frequencies at the 5′ and 3′ ends of 27- to 50-nucleotide precursors in Col and *rtl2*. Indeed, previous reports indicated that P4R2 RNAs preferentially start with an A or a G and preferentially end with a U (Blevins et al., 2015; Zhai et al., 2015). Analysis of the 5′ and 3′ ends of the 27- to 50-nucleotide reads derived from the 7000 PolIV loci in our Col libraries (Figures 7B and 7C, PolIV/Col) confirmed published data. This hierarchy was unchanged in *rtl2* (Figures 7B and 7C, PolIV/*rtl2*), indicating that the absence of RTL2 does not generally affect the P4R2 RNAs derived from the 7000 PolIV loci. In contrast, the 479 PolIV loci regulated by RTL2 exhibited a different hierarchy at both 5′ and 3′ ends in Col (Figures 7B and 7C, RTL2/Col). Indeed, the 5′ ends of precursors derived from RTL2 loci exhibited a G>A>U>C hierarchy in Col, whereas an A>G>U>C hierarchy was observed for precursors derived from the 7000 PolIV loci. In addition, the 3′ ends of precursors derived from RTL2 loci exhibited an A>U>C>G hierarchy in Col, whereas a U>A>C>G hierarchy was observed for precursors derived from the 7000 PolIV loci. Note that the two RTL2-sensitive loci producing 21-nucleotide miRNA (Figure 4C) were excluded from this analysis. Remarkably, the differences observed at RTL2-regulated loci in Col were reverted by the *rtl2* mutation (Figures 7B and 7C, RTL2/*rtl2*). Indeed, the 5′ ends of precursors derived from RTL2 loci in *rtl2* exhibited the same A>G>U>C hierarchy as precursors derived from the 7000 PolIV loci, and the 3′ ends of precursors derived from RTL2 loci in *rtl2* exhibited the same U>A>C>G hierarchy as precursors derived from the 7000 PolIV loci. The observation that precursors derived from RTL2-regulated loci exhibit in *rtl2*, but not in Col, the same 5′ and 3′ end signatures as precursors derived from PolIV loci strongly suggests that precursors derived from RTL2-regulated loci are cleaved by RTL2 in Col, resulting in a truncation of their 5′ and/or 3′ ends and a general modification of 5′ and 3′ end compositions in Col.

RTL2-Dependent and RTL2-Sensitive 24-Nucleotide siRNAs Guide DNA Methylation

PolIV-dependent 24-nucleotide siRNAs generally guide DNA methylation at the locus of origin via the RdDM pathway, which requires PolV and DRM2 (Law et al., 2013). Moreover, mutations in the PolV subunit NRPE1 or in DRM2 generally reduce siRNA accumulation at these loci. To determine if RTL2-regulated loci belong to the category of PolIV-PolV-DRM2-dependent loci, the dependence of RTL2-dependent and RTL2-sensitive siRNAs toward PolV and DRM2 was examined in public databases (Law et al., 2013); 87.6% and 92.6% of RTL2-dependent and RTL2-sensitive loci produced siRNAs whose accumulation was reduced in the *nrpe1* mutant, respectively (Figure 6C; Supplemental Data Set 2). Moreover, 90.7% of the RTL2-dependent loci and 95.7% of the RTL2-sensitive loci produced siRNAs whose accumulation

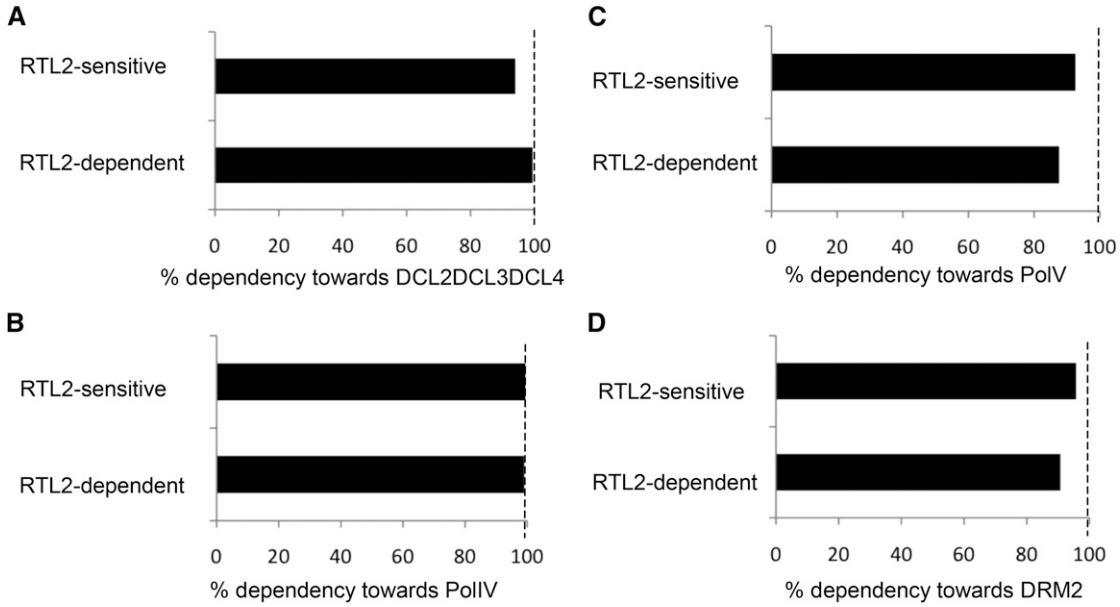

Figure 6. Most RTL2-Dependent and RTL2-Sensitive siRNAs Are Dependent on the RdDM Pathway.

Accumulation of RTL2-dependent and RTL2-sensitive siRNAs was examined in *dcl2 dcl3 dcl4* **(A)**, *nrpd1* **(B)**, *nrpe1* **(C)**, and *drm2* **(D)** mutants. Data from *nrpd1a*, *nrpe1*, and *drm2* mutants were obtained from public databases (Law et al., 2013). RTL2 siRNAs were defined as DCL2DCL3DCL4, PolIV, PolV, or DRM2 dependent when an at least 2-fold reduction of their accumulation was observed in the corresponding mutants. Results are shown as a percentage of the number of RTL2-dependent or RTL2-sensitive loci presenting a minimum of 10 reads in Col. The criteria for the dependency upon DCL2DCL3DCL4, NRPD1, NRPE1, or DRM2 are similar to those for the dependency on RTL2 (see Methods).

was reduced in the *drm2* mutant (Figure 6D; Supplemental Data Set 2). These results therefore suggest that the vast majority of the loci regulated by RTL2 produce siRNAs that guide DNA methylation via PolV and DRM2.

To determine if these 24-nucleotide siRNAs actually guide DNA methylation, representative RTL2-dependent and RTL2-sensitive loci containing at least one site for the methylation-dependent restriction enzyme McrBC were examined in Col and *rtl2*. At first, two RTL2-dependent loci (Chr2:11739900. . .11740391 and Chr5:486722. . .487047), subsequently referred to as K2-11739 and K5-486 and corresponding to one transposon and one IGR, respectively, were analyzed. McrBC digestion followed by either PCR or quantitative RT-PCR revealed that the decrease in the amount of siRNAs at these loci in *rtl2* is accompanied by a decrease in DNA methylation (Figure 8A). Although DNA methylation was not always completely eliminated in *rtl2*, these results indicate that RTL2-dependent siRNAs guide DNA methylation. Then, two RTL2-sensitive loci (Chr3:14761569. . .14762447 and Chr5:2656025. . .2656210), subsequently referred to as K3-14761 and K5-2656 and corresponding to one transposon and one IGR, respectively, were analyzed. McrBC digestion followed by either PCR or quantitative RT-PCR revealed that the increase in the amount of siRNAs at these loci in *rtl2* provokes an increase in DNA methylation (Figure 8B). Note that despite the absence of detectable siRNAs in Col, DNA methylation was not null at these loci. Moreover, DNA methylation did not reach 100% in *rtl2*. Nevertheless, these results indicate that like RTL2-dependent siRNAs, RTL2-sensitive siRNAs guide DNA methylation.

RTL2 Regulates the Expression of Protein-Coding Genes via NRPD1-Dependent 24-Nucleotide siRNA-Guided DNA Methylation

Most RTL2-dependent and RTL2-sensitive loci correspond to transposons or IGR (Figure 4B), which are well-known targets of the RdDM pathway. However, 13% of the RTL2-dependent and RTL2-sensitive loci correspond to protein-coding genes, raising the question of whether RTL2 plays a role in the regulation of protein-coding genes via a modulation of 24-nucleotide siRNA and DNA methylation levels. To test this hypothesis, we examined DNA methylation and mRNA accumulation at representative RTL2-dependent and RTL2-sensitive loci. The RTL2-dependent locus Chr3:129926. . .130082, referred to as K3-130 in Figure 9A, is internal to the transcribed region of *AT3G01345*, which encodes a putative glycoside hydrolase. This locus produces a higher number of 24-nucleotide siRNAs in Col than in *rtl2* and is more methylated in Col than in *rtl2*. Analysis of mRNA levels revealed that *AT3G01345* mRNA accumulates at lower level in Col than in *rtl2* (Figure 9A), suggesting an inverse correlation between DNA methylation and mRNA levels. The RTL2-sensitive locus Chr1:8335199. . .8336154, referred to as K1-8335 in Figure 9B, is internal to the transcribed region of *AT1G23480*, which encodes the cellulose synthase CSLA3. This locus produces a higher number of 24-nucleotide siRNAs in *rtl2* than in Col and is more methylated in *rtl2* than in Col, yet the *AT1G23480* mRNA accumulates at lower levels in *rtl2* than in Col (Figure 9B), pointing again to an inverse correlation between DNA methylation and mRNA

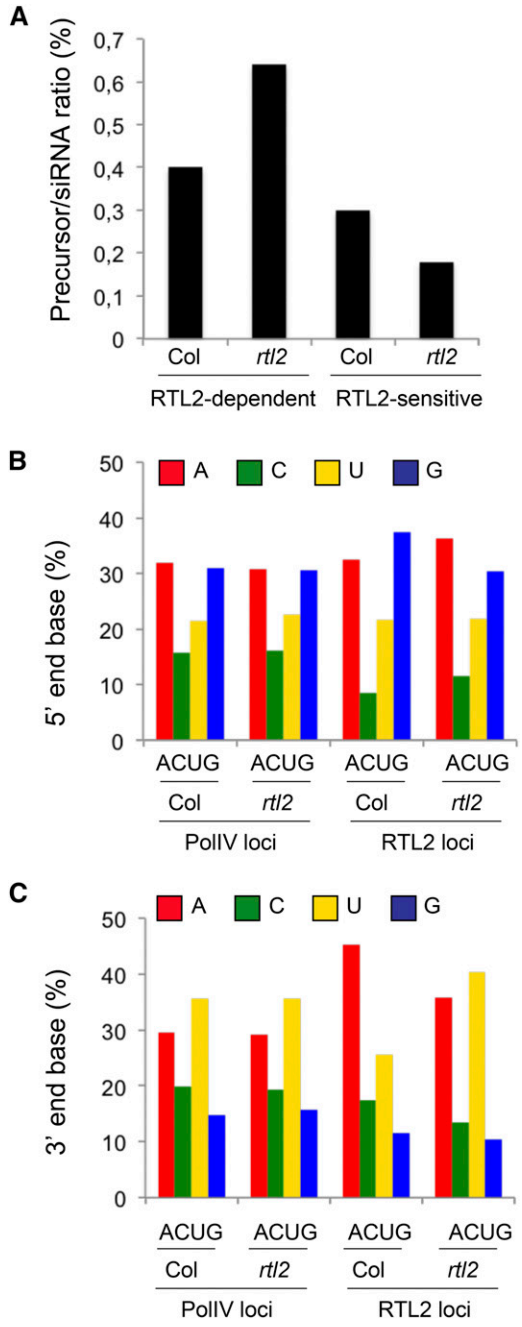

Figure 7. Precursor RNAs Derived from RTL2-Regulated Loci Differ from Canonical PolIV P4R2 Precursor RNAs.

(A) Precursor/siRNA ratio at RTL2-dependent and RTL2-sensitive loci in Col and *rtl2*.

(B) Frequencies at which the four nucleotides are present at the 5′ end of precursors derived from all PolIV loci or from PolIV loci regulated by RTL2 in Col and *rtl2*.

(C) Frequencies at which the four nucleotides are present at the 3′ end of precursors derived from all PolIV loci or from PolIV loci regulated by RTL2 in Col and *rtl2*.

levels. There is little chance that the decrease in *AT1G23480* mRNA accumulation is due to posttranscriptional gene silencing. Indeed, only two 21-nucleotide siRNAs appear at this locus in *rtl2*. Therefore, we propose that the decreased accumulation of *AT1G23480* mRNA is due to increased DNA methylation caused by increased amounts of 24-nucleotide siRNAs. To test this hypothesis, we examined *AT1G23480* mRNA accumulation and DNA methylation levels at the K1-8335 locus in the *rtl2 nrpd1* double mutant because the *nrpd1* mutation blocks the production of 24-nucleotide siRNAs. DNA methylation was increased from 83% in Col to 98% in *rtl2*, and it returned to 83% in the *rtl2 nrpd1* double mutant (Figure 10A), indicating that the increase in DNA methylation at the K1-8335 locus in *rtl2* actually depends on the RdDM pathway. Conversely, *AT1G23480* mRNA accumulation, which was decreased in *rtl2*, was almost restored to the Col level in the *rtl2 nrpd1* double mutant (Figure 10B), confirming that the increased DNA methylation in *rtl2* actually causes a decrease in mRNA accumulation. The inverse correlation between mRNA accumulation and 24-nucleotide siRNAs/DNA methylation at protein-coding genes therefore suggests that RTL2 participates in the regulation of gene expression.

RTL2 Loss-of-Function Affects Root Growth in an NRPD1-Dependent Manner

To determine if RTL2-mediated regulation of siRNA production plays a physiological role, Col and *rtl2* plants were grown on various substrates and under various growth conditions. We observed that 8-d-old *rtl2* seedlings have shorter roots (Figure 10C) when plated on saccharose-free medium. The root length of the *rtl2 nrpd1* double mutant was similar to Col (Figure 10C), indicating that the *nrpd1* mutation suppresses this developmental defect. This result therefore suggests that in wild-type plants, RTL2 prevents the production of 24-nucleotide siRNAs that, via the RdDM pathway, target DNA methylation and silencing of gene (s) involved in root development.

DISCUSSION

Arabidopsis RTL2 carries one RNaseIII and two DRB domains, which makes it structurally similar to the noncanonical Dicer enzyme Dcr1 found in budding yeasts. At first glance, their biological functions also look similar. Indeed, both Dcr1 and RTL2 are involved in rRNA maturation by cleaving the 3′ETS of the rRNA precursor (Comella et al., 2008; Bernstein et al., 2012), and both participate in the production of siRNAs (Drinnenberg et al., 2009; Weinberg et al., 2011; this work). However, they likely act in siRNA production through different mechanisms. Indeed, the noncanonical Dicer enzyme Dcr1 binds cooperatively along dsRNA molecules and cleaves them every 23 nucleotides, which corresponds to the distance between consecutive active sites (Weinberg et al., 2011). This is in contrast with the action of canonical Dicer enzymes from animals and plants, which successively remove siRNA duplexes from the dsRNA termini. It is likely that canonical Dicer enzymes recognize the end of dsRNA molecules through their PAZ domain and cleave at a distance spanning the terminus binding PAZ domain and RNaseIII active sites, resulting in siRNA molecules whose size varies between

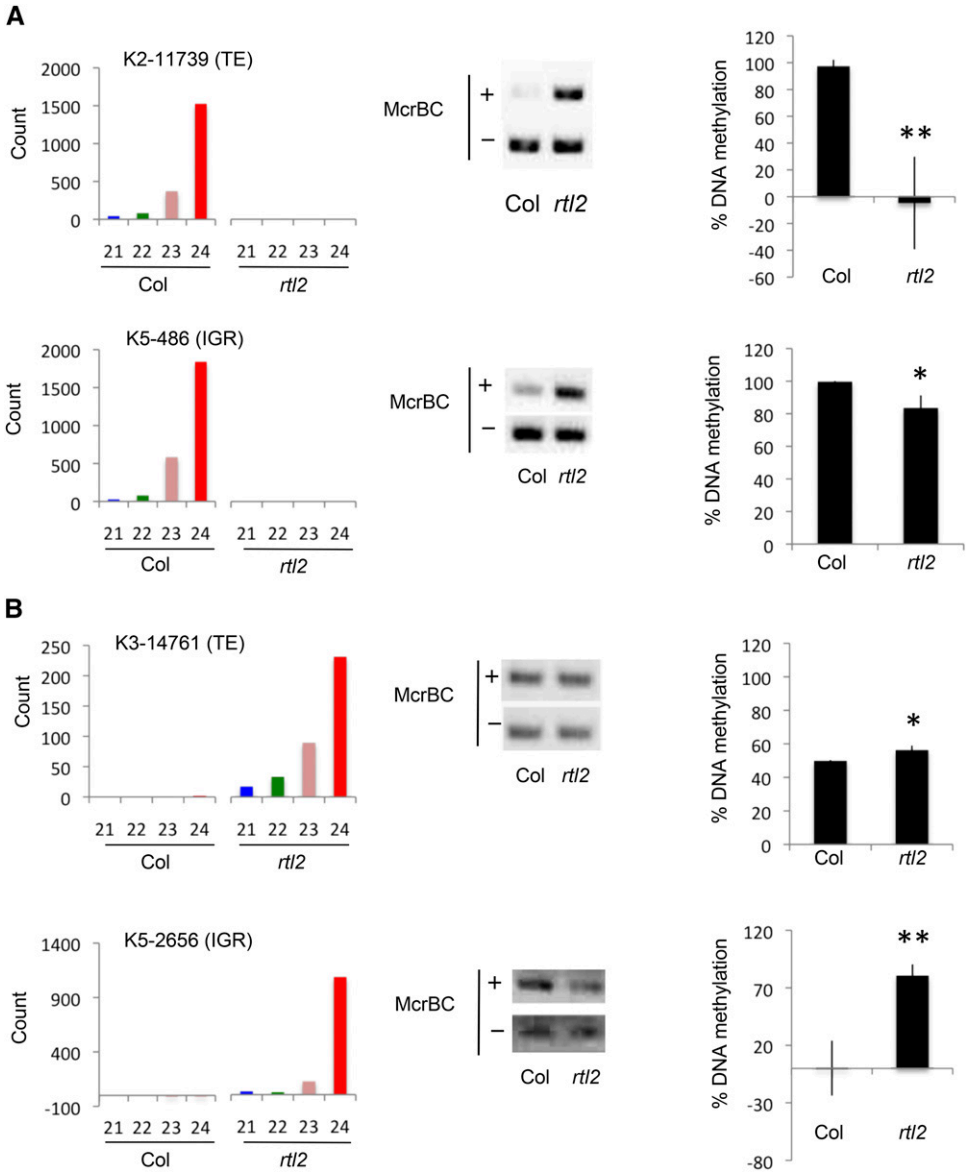

Figure 8. RTL2-Dependent and RTL2-Sensitive 24-Nucleotide siRNAs Guide DNA Methylation.

Size distribution and number of reads and DNA methylation visualized by PCR and quantitative PCR in Col and *rtl2* bulks of flowers at two selected RTL2-dependent loci **(A)** and two selected RTL2-sensitive loci **(B)**. Percentage of DNA methylation is calculated as follows: % DNA methylation = (1 − (2 ^ − Dct)) *100 with Dct = ct(McrBC) – ct(mock). Results are the average of three to four independent bulks of flowers. The different bulks were grown and collected and RNA extracted at different times. Error bars represent SE of the mean. Asterisks indicate a significant difference between Col and *rtl2* (*t* test: *P < 0.05 and **P < 0.005).

21 and 24 nucleotides depending on the Dicer enzyme considered (Zhang et al., 2004). Here, we show that, despite its structural similarities with the noncanonical Dicer enzyme Dcr1, Arabidopsis RTL2 does not produce siRNAs by itself. Nevertheless, *rtl2* mutations modify the repertoire of endogenous siRNAs. Apart from a general reduction of the amount of siRNAs derived from the *45S* pre-rRNA (Figure 3), RTL2 impairment does not result in a general change in siRNA size distribution or in a general change in siRNA genomic distribution (Figure 4A). Instead, RTL2 impairment specifically affects the abundance of siRNAs produced at a subset of loci, resulting either in increased or decreased siRNA accumulation (Figure 5). Indeed, among a total of 481 differential loci identified with a P value < 5E-2, 183 loci show a decreased accumulation of siRNAs (RTL2-dependent loci), while 298 loci show an increased accumulation of siRNAs (RTL2-sensitive loci). Similarly, increased or decreased accumulation of siRNAs was observed when overexpressing RTL2 in *N. benthamiana* together with artificial or natural dsRNA substrates. Indeed, RTL2 overexpression increased the production of siRNAs from *GFP* dsRNA (Kiyota et al., 2011), *GUS* dsRNA, and *At3g18145* 3′UTR hairpin

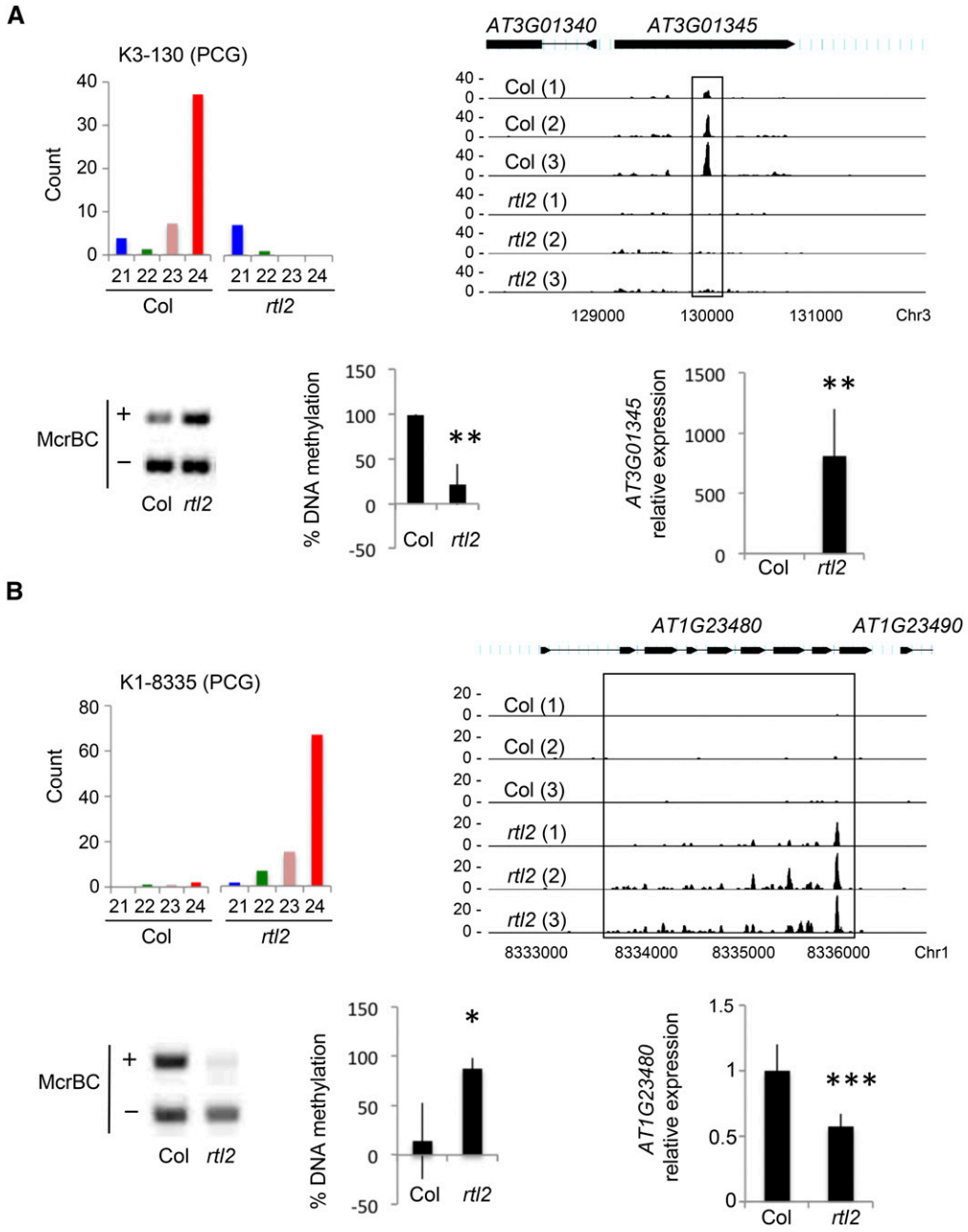

Figure 9. Modulation of siRNA Accumulation by RTL2 Affects mRNA Steady State Level.

Abundance, size, and distribution of small RNA reads, DNA methylation visualized by PCR and quantitative PCR, and mRNA steady state level analyzed by qRT-PCR and normalized to *GAPDH* expression at one selected RTL2-dependent protein-coding gene **(A)** and one selected RTL2-sensitive protein-coding gene **(B)**. The differential regions are boxed. Percentage of methylation is calculated as described in Figure 8. Results are the average of four independent bulks of flowers. The different bulks were grown and collected and RNA extracted at different times. mRNA accumulation is the average of two independent experiments on three to five flower samples. Error bars represent SE of the mean, and asterisks indicate a significant difference between Col and *rtl2* (*t* test: *P < 0.05, **P < 0.005, and ***P < 0.001).

(Figures 2B and 2H), while it decreased the production of siRNAs from *AT3G10490* dsRNA (Figure 2G). Given that this effect depends on functional RNaseIII and DRB domains (Figures 2C and 2D), we propose that RTL2 recognizes the sequence or structure of certain dsRNAs and that following cleavage, the products are either more efficiently or less efficiently processed by the DCLs than the original dsRNA.

Two RNaseIII domains are required to cleave double-stranded RNA, with one domain acting on each strand of the dsRNA substrate (Zhang et al., 2004; MacRae and Doudna, 2007). For this reason,

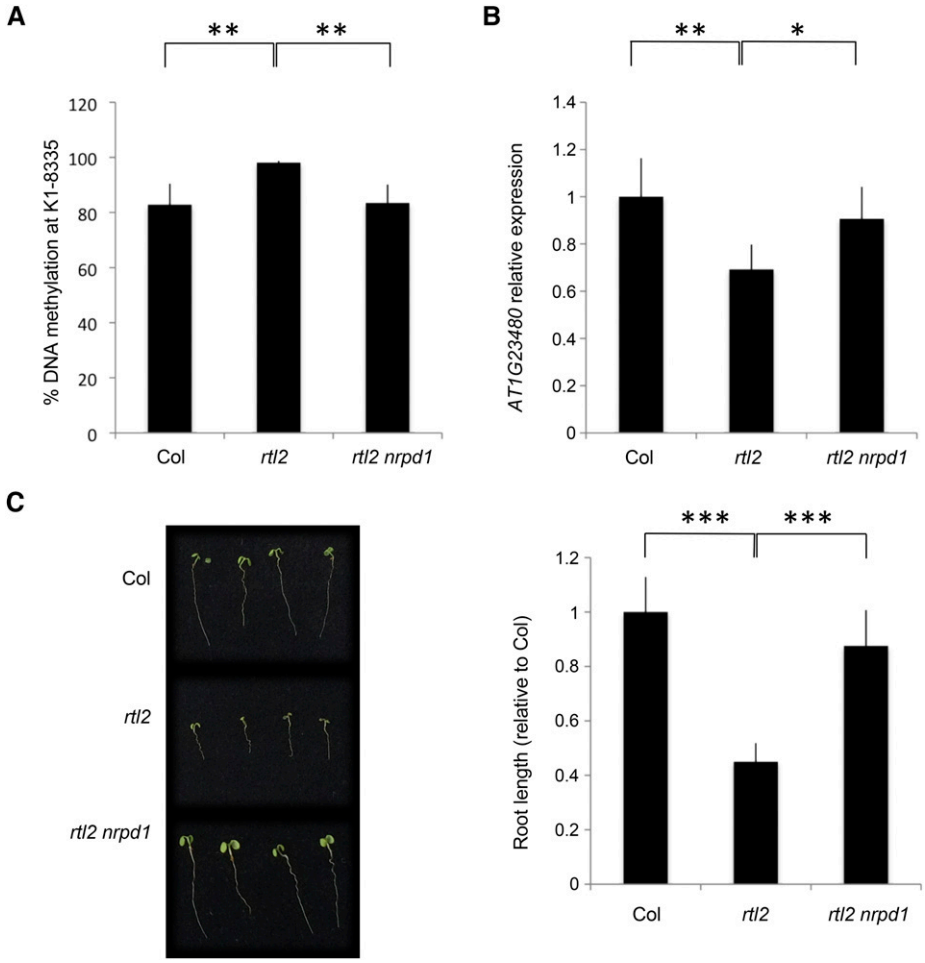

Figure 10. The *nrpd1* Mutation Suppresses Increased DNA Methylation, Reduced mRNA Accumulation, and Short Root Phenotype Triggered by the *rtl2* Mutation.

(A) DNA methylation at K1-8335 locus visualized by quantitative PCR. Percentage of methylation is calculated as in Figures 8 and 9.
(B) *AT1G23480* mRNA accumulation analyzed by qRT-PCR and normalized to *GAPDH*.
(C) Root length phenotype of representative Col, *rtl2*, and *rtl2 nrpd1* double mutant seedlings. Results are represented as the percentage of root length in Col. Root analyses were performed three times. Error bars represent SE of the mean, and asterisks indicate a significant difference between Col and *rtl2* (ANOVA test followed by Sidak's multiple tests between Col and *rtl2*) and between *rtl2* and *rtl2 nrpd1*. The asterisks represent the different levels of confidence (*P < 0.05; **P < 0.005, and ***P < 0.001).

RNaseIII enzymes that exhibit a single RNaseIII domain generally act as homodimers. In budding yeasts, Dcr1 forms homodimers that start cleaving dsRNAs from the interior and work outward every 23 nucleotides, which corresponds to the distance between consecutive active sites (Weinberg et al., 2011). Arabidopsis RTL2 also forms homodimers (Comella et al., 2008), suggesting that it could cleave dsRNA on both strands. However, the number of cleavages performed by RTL2 on a dsRNA molecule and the size of the resulting products are not known. We observed that certain hairpins give rise to RNA products larger than 25 nucleotides when they are coexpressed with RTL2 in *N. benthamiana* (Figures 2D and 2H) or when RTL2 is overexpressed in Arabidopsis *dcl3* mutants (Figure 1D). Based on their migration relative to 21- and 24-nucleotide siRNAs, the size of these products appear to vary between hairpins, suggesting that, unlike Dcr1, RTL2 dos not produce a unique size of molecules.

With the exception of two MIR genes (MIR843 and MIR863), the endogenous loci that are differentially regulated by RTL2 predominantly produce 24-nucleotide siRNAs, which are produced by DCL3. Interestingly, biochemical characterization of Arabidopsis DCL3 and DCL4 revealed that DCL3 preferentially cleaves short dsRNAs (37 to 50 nucleotides), whereas DCL4 preferentially cleaves long dsRNAs (Nagano et al., 2014). This result is consistent with DCL4 producing phased 21-nucleotide siRNAs over a long distance on *TAS* genes (Vazquez et al., 2004; Allen et al., 2005), whereas DCL3 does not. Nevertheless, DCL3-dependent 24-nucleotide siRNAs sometimes originate from long genomic regions. Two recent reports revealed that these long genomic regions produce short, often overlapping, dsRNA precursors, 27 to 50 nucleotides in length, referred to as P4R2 RNAs because they are generated through the action of PolIV and RDR2.

These short precursors are each processed into a single 24-nucleotide siRNA by DCL3, resulting in a one-precursor, one-siRNA model (Blevins et al., 2015; Zhai et al., 2015). Because our results suggest that RTL2 produce RNA molecules of a size larger than 24 nucleotides, we asked whether 27- to 50-nucleotide P4R2 RNAs were direct products of PolIV and RDR2 or products of PolIV and RDR2 that have been cleaved by RTL2. When looking at the 7000 PolIV genomic loci, we did not observed an increased accumulation of products larger than 24 nucleotides in *rtl2* compared with Col (Figure 4A). Moreover, the 5′ and 3′ end signatures of P4R2 RNAs were unchanged between Col and *rtl2* (Figures 7B and 7C), which goes against the hypothesis that all 27- to 50-nucleotide P4R2 RNAs could be products of RTL2. Rather, the finding that RTL2 regulates the production of siRNAs at only a subset of PolIV loci and sometimes within a limited region of PolIV loci (Figure 5) is more likely explained by RTL2 recognizing a specific sequence and/or structure present in only a subset of P4R2 RNAs and by the resulting cleavage product being better or less efficiently processed by DCL3 than the original uncut P4R2 RNAs. Supporting this hypothesis, we found that the precursor/siRNA ratio at RTL2-dependent loci was lower in Col than in *rtl2*, whereas the precursor/siRNA ratio at RTL2-sensitive loci was higher in Col than in *rtl2* (Figure 7A). Moreover, we found that precursors derived from RTL2-regulated loci have different 5′ and 3′ end signatures in Col and *rtl2*, whereas the 5′ and 3′ end signatures of the 7000 PolIV loci are similar in Col and *rtl2* (Figures 7B and 7C). The observation that RTL2-regulated loci have signatures in *rtl2* that resemble those of the 7000 PolIV loci strongly suggests that precursors derived from RTL2-regulated loci are cleaved by RTL2 in Col, resulting in a truncation of their 5′ and/or 3′ end, and a general modification of 5′ and 3′ end compositions in Col.

DCL3-dependent 24-nucleotide siRNAs guide DNA methylation and chromatin remodeling at transposable elements, repeats, and pseudogenes (Chan et al., 2005; Kasschau et al., 2007). Accordingly, we observed that RTL2 impairment provokes changes in DNA methylation that correlate with changes in 24-nucleotide siRNA accumulation at such loci (Figure 8). Unexpectedly, we also observed changes in DNA methylation and mRNA steady state levels at loci that produce 24-nucleotide siRNAs from protein-coding genes (Figure 9). Little is known about the role of 24-nucleotide siRNA and DNA methylation in regulating protein-coding gene expression. Previous analyses revealed that 1204 genes are found in proximity to 24-nucleotide siRNA clusters and that siRNAs originate from the transcribed region in 8.6% of them (Kasschau et al., 2007). However, even though hundreds of genes are deregulated in the *rdr2* or *dcl3* mutants, no significant overlap was found between genes producing 24-nucleotide siRNAs and genes deregulated in *rdr2* or *dcl3* (Kasschau et al., 2007). The authors concluded that 24-nucleotide siRNAs do not generally regulate protein-coding genes. However, they did not exclude the possibility that this could be the case for some of them. Another study showed that among the 70 genes that are upregulated in *rdr2*, 26 produce 24-nucleotide siRNAs (Kurihara et al., 2008), suggesting that 24-nucleotide siRNAs could participate in the regulation of protein-coding genes. We did not find these 26 genes in the list of protein-coding genes producing 24-nucleotide siRNAs that are regulated by RTL2, indicating that if these 24-nucleotide siRNAs regulate protein-coding genes, their production is not modulated by RTL2. However, we found that RTL2 impairment affects 24-nucleotide siRNA accumulation at several protein-coding genes, which results in changes in DNA methylation and mRNA steady state levels (Figure 9). Importantly, we monitored changes in both directions, i.e., we found loci at which 24-nucleotide siRNA accumulation and DNA methylation are reduced and mRNA levels increased in *rtl2* and loci at which 24-nucleotide siRNA accumulation and DNA methylation are increased and mRNA levels reduced in *rtl2*, thus implicating 24-nucleotide siRNAs in the regulation of protein-coding gene expression.

To conclude, class II RNaseIII enzymes still have biological functions and mechanisms of action that have yet to be revealed. Originally, they were known for their role in processing structural RNAs such as small nuclear RNA, rRNA, or tRNA. Then, budding yeast Dcr1 was shown to both maturate structural RNAs and process dsRNAs into siRNAs, similar to Dicer enzymes. Our work extends the diversity of class II RNaseIII functions by showing that Arabidopsis RTL2 modulates the production of siRNAs at a subset of endogenous loci, including protein-coding genes, resulting in either promotion or repression of gene expression depending on the substrate considered. This is reminiscent of recent studies showing that mutations in *Escherichia coli* RNaseIII result in increased and decreased levels of 9 and 3% of mRNAs, respectively (Stead et al., 2011). Finally, we show that the *rtl2* mutant exhibits a short root phenotype that is suppressed by the *nrpd1* mutation (Figure 10), indicating that this developmental defect is caused by de novo 24-nucleotide siRNAs produced in the absence of RTL2. Our results therefore indicate that one of the functions of RTL2 in wild-type plants is to prevent the production of 24-nucleotide siRNAs that target DNA methylation of protein-coding genes involved in root development. Whether RTL2 plays a role in other aspects of development or stress responses remains to be determined.

METHODS

Plant Material

Arabidopsis thaliana plants were in the Col-0 ecotype. The *rtl2* (GABI_568D10), *nrpd1a-5*, and *dcl2 dcl3 dcl4* mutants and the lines *L1*, *L1/dcl2 dcl4*, *T+S*, and *T+S/dcl3* have been previously described (Elmayan et al., 1998; Henderson et al., 2006; Smith et al., 2007; Comella et al., 2008; Kanno et al., 2008; Parent et al., 2015). The conditions in the controlled-environment chambers were as follows: 150 $mE \cdot m^{-2} \cdot s^{-1}$ irradiance provided by Osram Lumilux Cool Daylight L 36W/865 fluorescent tubes, 16 h of light, 60% relative humidity, 18°C night temperature, and 20°C day temperature. For root growth measurements, plants were grown vertically for 8 d in vitro on nutrient agar-solidified medium without sucrose (Estelle and Somerville, 1987).

Plasmid Constructs

The *ProUBQ10:RTL2*, *Pro35S:RTL2*, *Pro35S:Myc-RTL2*, and *Pro35S:RTL2-Myc* constructs were produced using Gateway technology (Invitrogen) as follows. *RTL2* genomic sequences starting at the ATG and ending at the stop codon were PCR amplified using *attB1RTL2F* and *attB2RTL2R* Gateway adapted oligonucleotides (Supplemental Table 3). After recombination into pDONR-207 vector through the Gateway BP recombinase reaction (Invitrogen) and verification by sequencing, final recombination into *pUB-DEST*, *pGWB2*, *pGWB17*, or *pGWB18* through

the Gateway LR recombinase reaction (Invitrogen) created *ProUBQ10:RTL2*, *Pro35S:RTL2*, *Pro35S:RTL2-Myc*, and *Pro35S:Myc-RTL2*, respectively.

Point mutations in the conserved amino acids of the RTL2 RNaseIII domain were introduced into *Pro35S:RTL2* and *Pro35S:RTL2-Myc* using a QuikChange site-directed mutagenesis kit (Stratagene). The Glu-93 and Asp-100 amino acids were replaced by an Ala using the *RTL2m-F* and *RTL2m-R* oligonucleotides (Supplemental Table 3). The *RTL2ΔDRB* sequence was amplified using the *attB1RTL2F* and *attB2RTL2-DRB1-R* oligonucleotides (Supplemental Table 3), recombined into the *pDONR207* (Gateway technology; Invitrogen), and then transferred to *pGWB2* and *pGWB17* to create *Pro35S:RTL2ΔDRB* and *Pro35S:RTL2ΔDRB-Myc*.

The *Pro35S:IR-NAC* construct was produced as follows. *NAC52* genomic sequences were PCR amplified using attB1-NAC-F and attB-NAC-R Gateway adapted oligonucleotides (Supplemental Table 3). After recombination into pDONR-207 vector through the Gateway BP recombinase reaction (Invitrogen) and verification by sequencing, final recombination into p35S:Hellsgate vector through the Gateway LR recombinase reaction (Invitrogen) created the *Pro35S-IR-NAC* construct. The *Pro35S:MIR168a* and *Pro35S:MIR168b* constructs have been described before (Vaucheret, 2009). The *Pro35S:AT3G18145* construct was produced as follows. The *AT3G18145* genomic sequence was PCR amplified using attB1-AT3G18145-F and attB-AT3G18145-R Gateway adapted oligonucleotides (Supplemental Table 3). After recombination into pDONR-207 vector through the Gateway BP recombinase reaction (Invitrogen) and verification by sequencing, final recombination into PGWB2 vector through the Gateway LR recombinase reaction (Invitrogen) created the *Pro35S:AT3G18145* construct. The *Pro35S:RTL1* construct has been described before (Shamandi et al., 2015).

Arabidopsis Transformation and *Nicotiana benthamiana* Agroinfiltration

Agrobacterium tumefaciens strains carrying plasmids of interest were grown overnight at 28°C in Luria-Bertani medium containing the appropriate antibiotics to a final OD_{600} between 1 and 2. For Arabidopsis transformation, the bacteria were pelleted and resuspended in 300 mL of infiltration medium (5% sucrose, 10 mM $MgCl_2$, and 0.015% Silwet L-77) to a final OD_{600} of 1, which was used for floral dipping (Clough and Bent, 1998). For *N. benthamiana* agroinfiltration, the bacteria were pelleted and resuspended in 1 mL of infiltration medium (10 mM $MgCl_2$, 10 mM MES pH 5.2, and 150 mM acetosyringone) to a final OD_{600} of 0.1. The solution containing the bacteria was injected into the abaxial side leaves using a 1-mL syringe, and samples were harvested 3 d after infiltration for fluorescence observation using a fluorescent zoom microscope (Axio Zoom; Zeiss) or for RNA analysis.

GUS Extraction and Activity Quantification

GUS protein was extracted and GUS activity was quantified as described before (Elmayan et al., 1998) from 12-d-old seedlings grown in vitro by measuring the quantity of 4-methylumbelliferone product generated from the substrate 4-methylumbelliferyl-β-D-glucuronide (Duchefa) on a fluorometer (Fluoroscan II; Thermo Scientific).

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, 10 mM EDTA, pH 8.0, and 20 mM β-mercaptoethanol, and RNAs were extracted two times with phenol and recovered by ethanol precipitation. Total RNA was separated on a 15% denaturing PAGE gel, stained with ethidium bromide, and transferred to nylon membrane (HybondNX; Amersham). For siRNA detection, blots were hybridized with a radioactively labeled random-primed DNA probe at 37°C in PerfectHyb Plus buffer (Sigma-Aldrich). For U6 detection, blots were hybridized with an end-labeled oligonucleotide at 50°C with hybridization buffer containing 5× SSC, 20 mM Na_2HPO_4, pH 7.2, 7% SDS, 2× Denhardt's solution, and denatured sheared salmon sperm DNA (Invitrogen). A FLA9500 phosphor imager (Typhoon) and LAS-4000 Fujifilm were used to quantify the hybridization signals. Oligonucleotides used for DNA probe amplification and for U6 oligo-probe are listed in Supplemental Table 3.

RNA Extraction and Quantitative RT-PCR

Total RNA from flowers was extracted using an RNeasy plant mini kit (Qiagen). One microgram of total RNA was digested with DNaseI (Fermentas) and converted to cDNA using the RevertAid H minus reverse transcriptase (Thermo Scientific) and oligo(dT). Gene expression was measured using SsoAdvanced Universal SYBR Green Supermix (Bio-Rad) and normalized to the housekeeping gene *GAPDH*. Primer sequences are listed in Supplemental Table 3.

Protein Extraction and Protein Gel Blot Analysis

Plant protein extraction and protein gel blot analysis were performed as previously described (Mallory et al., 2009). Briefly, 50 µg of crude protein extract was loaded on a 10% denaturing SDS-PAGE gel. Proteins were transferred to nitrocellulose membrane by semidry electroblotting. For detection, monoclonal anti-c-MYC antibody (purified mouse immunoglobulin clone 9E10 from Sigma-Aldrich; M4439) and a goat anti-mouse IgG-HRP conjugated (Santa Cruz; sc-2005) were used at 1:5000 and 1:10,000, respectively. Blots were developed using an ECL kit (Pierce Thermo; 32106), and the emitted chemiluminescence was directly quantified by an ImageQuant LAS4000 biomolecular imager (GE Helthcare).

Methylation Analysis

DNA was extracted from flowers using a DNeasy plant mini kit (Qiagen) and 250 ng DNA was mock or McrBC digested (Biolabs) overnight at 37°C. Ten percent of mock or McrBC-digested DNA was used for PCR amplification. Quantitative PCR was also performed as above. Percentage of DNA methylation was calculated as follows: % DNA methylation = (1-2 ^ − Dct) *100 with Dct = ct(McrBC) − ct(mock). Primer sequences are listed in Supplemental Table 3.

Bioinformatics

The 3′ adapters of the reads were removed with S-MART tools (Zytnicki and Quesneville, 2011) and then sequences were mapped with Bowtie (Langmead et al., 2009) to the TAIR10 assembly. For the 45S rRNA analysis, all reads that matched the genome perfectly were kept, regardless of the number of loci at which they mapped. For the rest of the analyses, only reads that matched at a unique position in the genome were kept. To identify differential loci, a sliding window method was first used. Basically, we split the genome into 100-bp-long chunks, with a 50-bp overlap. All chunks with no more than 100 (normalized) reads in any condition were discarded, and the remaining chunks were merged if they overlapped. The chunks were then trimmed in order to keep the smallest region that contained all the reads found in a chunk. We clustered the reads of all replicates to form a putative transcript. We then proceeded using the usual mRNA-seq data. We quantified the expression of each putative transcript by counting the number of reads, and we used the standard DESeq method to find differentially expressed regions. We kept regions with P value < 0.05 and fold change of 2. Regions distant by no more than 100 bp were merged. To identify shorter differential loci, we used a second method based on the analysis of reads at each position of the genome. We counted the number of reads in wild-type and mutant replicates and submitted these figures to

DESeq as if they were usual gene counts to estimate the negative binomial parameters. For consistency, we kept the library size factor found in the sliding window analysis, and we performed the standard DESeq test to obtain an adjusted P value for each position of the genome. We decided that regions with adjusted P values less than 10^{-3} were differentially expressed. To refine boundaries, the regions were extended as much as possible, so that no P value of the regions was greater than 0.5. We merged regions if they were distant by less than 100 bp. Finally, we kept all the regions that exhibited a minimum of 100 reads in at least one condition and a minimum fold change of 2, for consistency with the previous analysis. Graphs were produced with S-MART tools and ad hoc scripts. Data were normalized with respect to Col under the assumption that the 27 canonical miRNA genes (miR156, 157, 159, 160, 162, 164, 165, 166, 167, 168, 169, 170, 171, 172, 319, 390, 391, 393, 394, 395, 396, 397, 398, 399, 403, 408, and 472) (Fahlgren et al., 2007) were globally unaffected. For normalization purpose, unique and nonunique miRNA reads were taken into account. miRNA annotation was downloaded from RFAM (Burge et al., 2013). EndoIR-siRNAs were annotated using the sRNA-producing loci predicted by Lindow et al. (2007), excluding those matching RFAM miRNAs. Loci producing siRNAs in a PolIV-dependent manner were retrieved from Mosher et al. (2008). Young miRNAs were defined as DCL1-independant sRNA: We downloaded the sRNAs produced by Lu et al. (2006) and discarded all of those miRNAs that contained at least one read from this data set. To quantify the precursor/siRNA ratio, we weighted the replicates using the normalization factors previously computed and pooled the three replicates of each condition. For 5′ and 3′ end precursor analysis, we computed the nucleotidic distribution of the 5′-most (or 3′-most) position of the reads. Each position was counted once, even if the same sequence was found in several reads, or several reads started (or ended) at the same position.

Accession Numbers

Sequence data from this article can be found in the GenBank/EMBL libraries under the following accession numbers: *At3g20420* (*RTL2*), *AT3G10490* (*NAC52*), *AT3G18145*, *AT3G01345*, *AT1G23480* (*CSLA3*), and *AT1G13440* (*GAPDH*). The data reported in this article have been deposited in the Gene Expression Omnibus database (www.ncbi.nlm.nih.gov/geo) under accession number GSE71782.

Supplemental Data

Supplemental Figure 1. RTL2 does not affect the production of miR168.

Supplemental Figure 2. Sequence of the rRNA 45S 3′ETS/IGS/5′ETS.

Supplemental Table 1. Characteristics of small RNA libraries.

Supplemental Table 2. Categories of RTL2-sensitive and RTL2-dependent loci.

Supplemental Table 3. List of oligonucleotides used for plasmid constructions, small RNA detection, and qRT-PCR.

Supplemental Data Set 1. 183 and 298 loci produce small RNAs that are RTL2 dependent and RTL2 sensitive, respectively.

Supplemental Data Set 2. The 21- to 24-nucleotide reads from RTL2-regulated loci in RdDM mutants *nrpd1*, *nrpe1*, and *drm2*.

ACKNOWLEDGMENTS

We thank Hervé Ferry and Patrick Grillot for plant care, Taline Elmayan and Samantha Vernhettes for fruitful discussions, and Marjori Matzke for providing T+S/dcl3 seeds. E.E.-M. was supported by a post-doctoral fellowship from the Plant Biology Division of INRA. This work was supported by grants from the French Agence Nationale pour la Recherche (ANR-10-LABX-40 to H.V. and ANR-11-BSV6-007 to H.V., M.Z., and J.S.-V.) and the Fondation Louis D. de l'Institut de France (to H.V.).

AUTHOR CONTRIBUTIONS

E.E.-M. and H.V. designed the research. E.E.-M., N.S., P.C., and H.V. performed the research. E.E.-M., M.H., H.S., P.C., J.S.-V., M.Z., and H.V. analyzed the data. E.E.-M. and H.V. wrote the article.

Received June 16, 2015; revised December 29, 2015; accepted January 12, 2016; published January 13, 2016.

REFERENCES

Allen, E., Xie, Z., Gustafson, A.M., and Carrington, J.C. (2005). MicroRNA-directed phasing during trans-acting siRNA biogenesis in plants. Cell **121:** 207–221.

Bernstein, D.A., Vyas, V.K., Weinberg, D.E., Drinnenberg, I.A., Bartel, D.P., and Fink, G.R. (2012). *Candida albicans* Dicer (CaDcr1) is required for efficient ribosomal and spliceosomal RNA maturation. Proc. Natl. Acad. Sci. USA **109:** 523–528.

Blaszczyk, J., Tropea, J.E., Bubunenko, M., Routzahn, K.M., Waugh, D.S., Court, D.L., and Ji, X. (2001). Crystallographic and modeling studies of RNase III suggest a mechanism for double-stranded RNA cleavage. Structure **9:** 1225–1236.

Blevins, T., Podicheti, R., Mishra, V., Marasco, M., Tang, H., and Pikaard, C.S. (2015). Identification of Pol IV and RDR2-dependent precursors of 24 nucleotide siRNAs guiding de novo DNA methylation in Arabidopsis. eLife **4:** 4.

Burge, S.W., Daub, J., Eberhardt, R., Tate, J., Barquist, L., Nawrocki, E.P., Eddy, S.R., Gardner, P.P., and Bateman, A. (2013). Rfam 11.0: 10 years of RNA families. Nucleic Acids Res. **41:** D226–D232.

Cao, M., Du, P., Wang, X., Yu, Y.Q., Qiu, Y.H., Li, W., Gal-On, A., Zhou, C., Li, Y., and Ding, S.W. (2014). Virus infection triggers widespread silencing of host genes by a distinct class of endogenous siRNAs in Arabidopsis. Proc. Natl. Acad. Sci. USA **111:** 14613–14618.

Chan, S.W., Henderson, I.R., and Jacobsen, S.E. (2005). Gardening the genome: DNA methylation in *Arabidopsis thaliana*. Nat. Rev. Genet. **6:** 351–360.

Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. Plant J. **16:** 735–743.

Comella, P., Pontvianne, F., Lahmy, S., Vignols, F., Barbezier, N., Debures, A., Jobet, E., Brugidou, E., Echeverria, M., and Sáez-Vásquez, J. (2008). Characterization of a ribonuclease III-like protein required for cleavage of the pre-rRNA in the 3'ETS in Arabidopsis. Nucleic Acids Res. **36:** 1163–1175.

Daxinger, L., Kanno, T., Bucher, E., van der Winden, J., Naumann, U., Matzke, A.J., and Matzke, M. (2009). A stepwise pathway for biogenesis of 24-nucleotide secondary siRNAs and spreading of DNA methylation. EMBO J. **28:** 48–57.

Drinnenberg, I.A., Weinberg, D.E., Xie, K.T., Mower, J.P., Wolfe, K.H., Fink, G.R., and Bartel, D.P. (2009). RNAi in budding yeast. Science **326:** 544–550.

Dunn, J.J., and Studier, F.W. (1975). Effect of RNAase III, cleavage on translation of bacteriophage T7 messenger RNAs. J. Mol. Biol. **99:** 487–499.

Elela, S.A., Igel, H., and Ares, M., Jr. (1996). RNase III cleaves eukaryotic preribosomal RNA at a U3 snoRNP-dependent site. Cell **85:** 115–124.

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

Emory, S.A., and Belasco, J.G. (1990). The ompA 5′ untranslated RNA segment functions in *Escherichia coli* as a growth-rate-regulated mRNA stabilizer whose activity is unrelated to translational efficiency. J. Bacteriol. **172:** 4472–4481.

Estelle, M.A., and Somerville, C. (1987). Auxin-resistant mutants of *Arabidopsis thaliana* with an altered morphology. Mol. Gen. Genet. **206:** 200–206.

Fahlgren, N., Howell, M.D., Kasschau, K.D., Chapman, E.J., Sullivan, C.M., Cumbie, J.S., Givan, S.A., Law, T.F., Grant, S.R., Dangl, J.L., and Carrington, J.C. (2007). High-throughput sequencing of Arabidopsis microRNAs: evidence for frequent birth and death of MIRNA genes. PLoS One **2:** e219.

Filippov, V., Solovyev, V., Filippova, M., and Gill, S.S. (2000). A novel type of RNase III family proteins in eukaryotes. Gene **245:** 213–221.

Henderson, I.R., Zhang, X., Lu, C., Johnson, L., Meyers, B.C., Green, P.J., and Jacobsen, S.E. (2006). Dissecting Arabidopsis thaliana DICER function in small RNA processing, gene silencing and DNA methylation patterning. Nat. Genet. **38:** 721–725.

Kameyama, L., Fernandez, L., Court, D.L., and Guarneros, G. (1991). RNaseIII activation of bacteriophage lambda N synthesis. Mol. Microbiol. **5:** 2953–2963.

Kanno, T., Bucher, E., Daxinger, L., Huettel, B., Böhmdorfer, G., Gregor, W., Kreil, D.P., Matzke, M., and Matzke, A.J. (2008). A structural-maintenance-of-chromosomes hinge domain-containing protein is required for RNA-directed DNA methylation. Nat. Genet. **40:** 670–675.

Kasschau, K.D., Fahlgren, N., Chapman, E.J., Sullivan, C.M., Cumbie, J.S., Givan, S.A., and Carrington, J.C. (2007). Genome-wide profiling and analysis of Arabidopsis siRNAs. PLoS Biol. **5:** e57.

Kiyota, E., Okada, R., Kondo, N., Hiraguri, A., Moriyama, H., and Fukuhara, T. (2011). An Arabidopsis RNase III-like protein, AtRTL2, cleaves double-stranded RNA in vitro. J. Plant Res. **124:** 405–414.

Kurihara, Y., Matsui, A., Kawashima, M., Kaminuma, E., Ishida, J., Morosawa, T., Mochizuki, Y., Kobayashi, N., Toyoda, T., Shinozaki, K., and Seki, M. (2008). Identification of the candidate genes regulated by RNA-directed DNA methylation in Arabidopsis. Biochem. Biophys. Res. Commun. **376:** 553–557.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. **10:** R25.

Law, J.A., Du, J., Hale, C.J., Feng, S., Krajewski, K., Palanca, A.M., Strahl, B.D., Patel, D.J., and Jacobsen, S.E. (2013). Polymerase IV occupancy at RNA-directed DNA methylation sites requires SHH1. Nature **498:** 385–389.

Lindow, M., Jacobsen, A., Nygaard, S., Mang, Y., and Krogh, A. (2007). Intragenomic matching reveals a huge potential for miRNA-mediated regulation in plants. PLOS Comput. Biol. **3:** e238.

Lu, C., Kulkarni, K., Souret, F.F., MuthuValliappan, R., Tej, S.S., Poethig, R.S., Henderson, I.R., Jacobsen, S.E., Wang, W., Green, P.J., and Meyers, B.C. (2006). MicroRNAs and other small RNAs enriched in the Arabidopsis RNA-dependent RNA polymerase-2 mutant. Genome Res. **16:** 1276–1288.

MacRae, I.J., and Doudna, J.A. (2007). Ribonuclease revisited: structural insights into ribonuclease III family enzymes. Curr. Opin. Struct. Biol. **17:** 138–145.

Mallory, A., and Vaucheret, H. (2010). Form, function, and regulation of ARGONAUTE proteins. Plant Cell **22:** 3879–3889.

Mallory, A.C., Hinze, A., Tucker, M.R., Bouché, N., Gasciolli, V., Elmayan, T., Lauressergues, D., Jauvion, V., Vaucheret, H., and Laux, T. (2009). Redundant and specific roles of the ARGONAUTE proteins AGO1 and ZLL in development and small RNA-directed gene silencing. PLoS Genet. **5:** e1000646.

Martínez 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.

Mosher, R.A., Schwach, F., Studholme, D., and Baulcombe, D.C. (2008). PolIVb influences RNA-directed DNA methylation independently of its role in siRNA biogenesis. Proc. Natl. Acad. Sci. USA **105:** 3145–3150.

Nagano, H., Fukudome, A., Hiraguri, A., Moriyama, H., and Fukuhara, T. (2014). Distinct substrate specificities of Arabidopsis DCL3 and DCL4. Nucleic Acids Res. **42:** 1845–1856.

Nicholson, A.W. (2014). Ribonuclease III mechanisms of double-stranded RNA cleavage. Wiley Interdiscip. Rev. RNA **5:** 31–48.

Parent, J.S., Bouteiller, N., Elmayan, T., and Vaucheret, H. (2015). Respective contributions of Arabidopsis DCL2 and DCL4 to RNA silencing. Plant J. **81:** 223–232.

Pontier, D., et al. (2012). NERD, a plant-specific GW protein, defines an additional RNAi-dependent chromatin-based pathway in Arabidopsis. Mol. Cell **48:** 121–132.

Portereiko, M.F., Sandaklie-Nikolova, L., Lloyd, A., Dever, C.A., Otsuga, D., and Drews, G.N. (2006). NUCLEAR FUSION DEFECTIVE1 encodes the Arabidopsis RPL21M protein and is required for karyogamy during female gametophyte development and fertilization. Plant Physiol. **141:** 957–965.

Preuss, S.B., Costa-Nunes, P., Tucker, S., Pontes, O., Lawrence, R.J., Mosher, R., Kasschau, K.D., Carrington, J.C., Baulcombe, D.C., Viegas, W., and Pikaard, C.S. (2008). Multimegabase silencing in nucleolar dominance involves siRNA-directed DNA methylation and specific methylcytosine-binding proteins. Mol. Cell **32:** 673–684.

Shamandi, N., Zytnicki, M., Charbonnel, C., Elvira-Matelot, E., Bochnakian, A., Comella, P., Mallory, A.C., Lepère, G., Sáez-Vásquez, J., and Vaucheret, H. (2015). Plants encode a general siRNA suppressor that is induced and suppressed by viruses. PLoS Biol. **13:** e1002326.

Smith, L.M., Pontes, O., Searle, I., Yelina, N., Yousafzai, F.K., Herr, A.J., Pikaard, C.S., and Baulcombe, D.C. (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.

Stead, M.B., Marshburn, S., Mohanty, B.K., Mitra, J., Pena Castillo, L., Ray, D., van Bakel, H., Hughes, T.R., and Kushner, S.R. (2011). Analysis of *Escherichia coli* RNase E and RNase III activity in vivo using tiling microarrays. Nucleic Acids Res. **39:** 3188–3203.

Tucker, S., Vitins, A., and Pikaard, C.S. (2010). Nucleolar dominance and ribosomal RNA gene silencing. Curr. Opin. Cell Biol. **22:** 351–356.

Vaucheret, H. (2006). Post-transcriptional small RNA pathways in plants: mechanisms and regulations. Genes Dev. **20:** 759–771.

Vaucheret, H. (2009). AGO1 homeostasis involves differential production of 21-nucleotide and 22-nucleotide miR168 species by MIR168a and MIR168b. PLoS One **4:** e6442.

Vazquez, F., Vaucheret, H., Rajagopalan, R., Lepers, C., Gasciolli, V., Mallory, A.C., Hilbert, J.L., Bartel, D.P., and Crété, P. (2004). Endogenous trans-acting siRNAs regulate the accumulation of Arabidopsis mRNAs. Mol. Cell **16:** 69–79.

Watkins, K.P., Kroeger, T.S., Cooke, A.M., Williams-Carrier, R.E., Friso, G., Belcher, S.E., van Wijk, K.J., and Barkan, A. (2007). A ribonuclease III domain protein functions in group II intron splicing in maize chloroplasts. Plant Cell **19:** 2606–2623.

Weinberg, D.E., Nakanishi, K., Patel, D.J., and Bartel, D.P. (2011). The inside-out mechanism of Dicers from budding yeasts. Cell **146:** 262–276.

Yoshikawa, M., Peragine, A., Park, M.Y., and Poethig, R.S. (2005). A pathway for the biogenesis of trans-acting siRNAs in Arabidopsis. Genes Dev. **19:** 2164–2175.

Zhai, J., et al. (2015). A one precursor one siRNA model for Pol IV-dependent siRNA biogenesis. Cell **163:** 445–455.

Zhang, H., Kolb, F.A., Jaskiewicz, L., Westhof, E., and Filipowicz, W. (2004). Single processing center models for human Dicer and bacterial RNase III. Cell **118:** 57–68.

Zhang, X., et al. (2015). Plant biology. Suppression of endogenous gene silencing by bidirectional cytoplasmic RNA decay in Arabidopsis. Science **348:** 120–123.

Zytnicki, M., and Quesneville, H. (2011). S-MART, a software toolbox to aid RNA-Seq data analysis. PLoS One **6:** e25988.