Small Silencing RNAs in Plants Are Mobile and Direct Epigenetic Modification in Recipient Cells

Attila Molnar,* Charles W. Melnyk,* Andrew Bassett, Thomas J. Hardcastle, Ruth Dunn, David C. Baulcombe†

A silencing signal in plants with an RNA specificity determinant moves through plasmodesmata and the phloem. To identify the mobile RNA, we grafted *Arabidopsis thaliana* shoots to roots that would be a recipient for the silencing signal. Using mutants that block small RNA (sRNA) biogenesis in either source or recipient tissue, we found that transgene-derived sRNA as well as a substantial proportion of the endogenous sRNA had moved across the graft union, and we provide evidence that 24-nucleotide mobile sRNAs direct epigenetic modifications in the genome of the recipient cells. Mobile sRNA thus represents a mechanism for transmitting the specification of epigenetic modification and could affect genome defense and responses to external stimuli that have persistent effects in plants.

In flowering plants, RNA silencing is a non–cell-autonomous process: It spreads both to neighboring cells and systemically over long distances (*1*–*3*). Mobile silencing operates in a nucleotide sequence–specific manner consistent with the involvement of RNA as a component of the signal, but until now the mobile RNA has not been detected. The 24-nucleotide (nt) and 21-nt species of small RNA (sRNA) have both been implicated in mobile silencing, but it could not be ruled out that a long sRNA precursor is the form that moves (*4*, *5*). Conversely, analyses with grafted plants suggested that long sRNA precursors are mobile (*6*), although the data did not rule out sRNAs. Here, we describe a similar grafting approach with *Arabidopsis thaliana* combined with high-throughput sequencing to detect 21- to 24-nt sRNA molecules that had moved across the graft union. Silencing pathway mutants allow us to demonstrate that the mobile sRNA is synthesized in the source rather than the recipient tissue. We show that mobile 24-nt sRNAs from three separate genomic loci can direct epigenetic modifications in the recipient cells.

Transgene-specific sRNAs are mobile. The grafting experiments used transgenic *A. thaliana* plants expressing a green fluorescent protein (GFP) transgene (G) as a reporter of RNA silencing (*7*) and a GFP-derived hairpin RNA [GF inverted-repeat (GF-IR)] to silence GFP (Fig. 1A). GFP-expressing plants are green under ultraviolet light, so that the movement of the silencing signal can be monitored through the loss of green fluorescence (Fig. 1, B and C, and fig. S1, A to F). In test grafts we ascertained that movement of the GFP silencing signal was more efficient from shoot to root than vice versa, consistent with previous findings of the source-to-sink movement of viruses and assimilates (*8*, *9*) (fig. S1, C and E). Subsequent experiments to analyze

Department of Plant Sciences, University of Cambridge, Cambridge CB2 3EA, UK.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: dcb40@cam.ac.uk

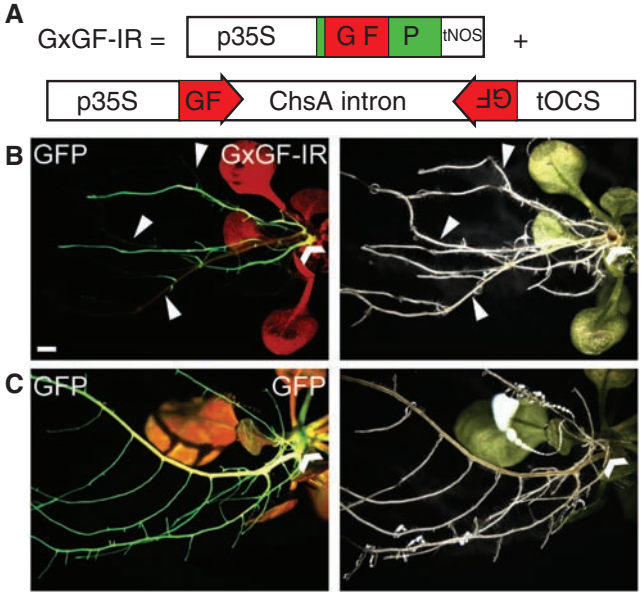
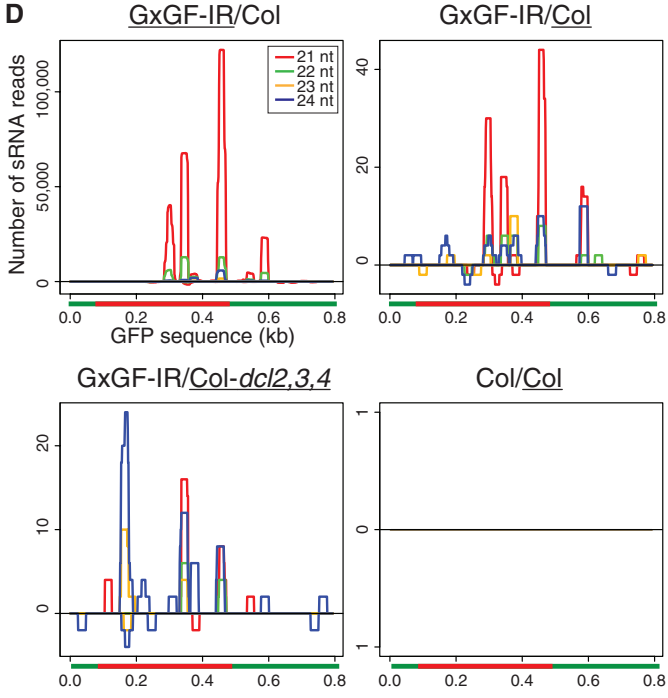


Fig. 1. Movement of transgene sRNA. (**A**) Schematic diagram of the GFP and GF-IR transgene constructs. The cauliflower mosaic virus 35S promoter (p35S) directed the expression of the GFP and GF-IR transgenes. The GF-IR construct contained a chalcone synthase A (ChsA) intron. The GFP and GF-IR transgenes contained nopaline synthase (tNOS) or octopine synthase (tOCS) terminators, respectively. (**B** and **C**) GFP-silenced [GxGF-IR (B)] or GFP-expressing (C) shoots were grafted onto GFP-expressing roots of 7-day-old *Arabidopsis* seedlings, and ultraviolet fluorescence (left) and bright-field images (right) were taken 5 weeks after grafting. Arrowheads indicate emerging silenced roots; chevrons indicate graft junction. Scale bar, 0.5 cm.



(**D**) sRNA sequence libraries generated from grafted *Arabidopsis* were aligned at the first nucleotide position to the GFP coding sequence; the positive or negative *y* axis shows the number of reads at each position on either plus or minus strand. The line below the *x* axis represents positions of the GFP coding sequence (green) and the inverted-repeat (GF) region (red). See fig. S1 for an independent biological replicate, and table S1 for a summary of sRNA library details.

mobile silencing were therefore carried out to assay shoot-to-root movement.

sRNA cDNA libraries were generated from signal source and recipient tissues 5 weeks after grafting. (The sampled tissue in grafted samples is underlined below using shoot/root notation.) The receiving root tissue was from nontransgenic Columbia (Col) genotype plants or from Col-*dcl2,3,4* triple mutant plants that are unable to produce 22-, 23-, and 24-nt sRNAs from double-stranded RNA precursors (*10*). In the signal donor GxGF-IR shoots, the lack of GFP expression was associated with the accumulation of GFP-derived sRNAs. All size classes of silencing RNAs (21, 22, and 24 nt) were present, indicating that multiple DCL proteins can act on the precursor of GFP silencing (Fig. 1D). There was a bias toward the accumulation of sense strand–specific, GFP-derived sRNAs from localized "hotspot" regions of the GFP RNA. This nonrandom distribution of the sRNA could be due to differential processing and/or stability of sRNA from different regions of the GFP transgenes (Fig. 1D), and it is likely a mixture of primary and secondary sRNAs that were produced by an RNA-dependent RNA polymerase on the full-length GFP RNA (*11*).

The grafted Col roots had GFP-specific sRNAs that were 21 to 24 nt in size, although they were less abundant than in the source tissue by three orders of magnitude (table S1). These RNAs were predominantly from the same hotspots of the sense GFP strand as in the source tissue (Fig. 1D and fig. S1G). A similar population of sRNAs was present in GxGF-IR/Col-*dcl2,3,4* roots, although the proportion of 24-nt RNA relative to the 21-nt sRNAs was higher than in the source tissue or in the Col receiving tissue (Fig. 1D and fig. S1G). These 22- to 24-nt sRNAs could not have been produced in the recipient tissue because of the absence of both DCL2 and DCL3, and hence they must have translocated from the source tissue. These mobile RNAs corresponded to the "GF" and "P" region in both Col and Col-*dcl2,3,4* roots, indicating that both primary and secondary sRNAs are mobile (Fig. 1D and fig. S1G).

Endogenous mobile small RNAs. To find out whether endogenous sRNAs are mobile, we analyzed the size profile of the sRNA libraries from various wild-type and mutant genotypes. The wild-type genotype used in this analysis was the GxGF-IR line in the C24 genotype background (*7*). The RNA silencing mutants were Col-*dcl2,3,4* in the Col background and a mutant (*sde4-2*) in C24 that lacks the functional RNA polymerase IV (Pol IV) required for 24-nt sRNA biogenesis (*12*). Col is the most widely used *Arabidopsis* genotype and its genome has been sequenced (*13*); C24 and Col are more genetically diverse than most other pairwise comparisons of *Arabidopsis* genotypes (*14*). Because transgene-derived sRNAs accumulate to similar levels in the sRNA libraries compared, the transgenes are not relevant to the analysis of the endogenous sRNAs, and these genotypes are referred to below as Col, C24, Col-*dcl2,3,4*, and C24-*sde4*.

The grafted Col roots accumulated more 23- and 24-nt sRNAs than 21-nt sRNAs, according to an analysis of either total or nonredundant sequence reads (Fig. 2A). The 21-nt species include microRNAs (miRNAs) and small interfering RNAs (siRNAs), whereas the 23- and 24-nt class includes the siRNAs associated with chromatin silencing (*15*, *16*). In the Col-*dcl2,3,4*/Col-*dcl2,3,4* roots, the 23- and 24-nt sRNAs were reduced to the background level because of the lack of DCL3. However, the 23- and 24-nt sRNAs in the C24/Col-*dcl2,3,4* roots were near wild-type levels, indicating that the endogenous sRNAs are mobile and that they account for a substantial proportion of the endogenous sRNA. These mobile sRNAs are likely

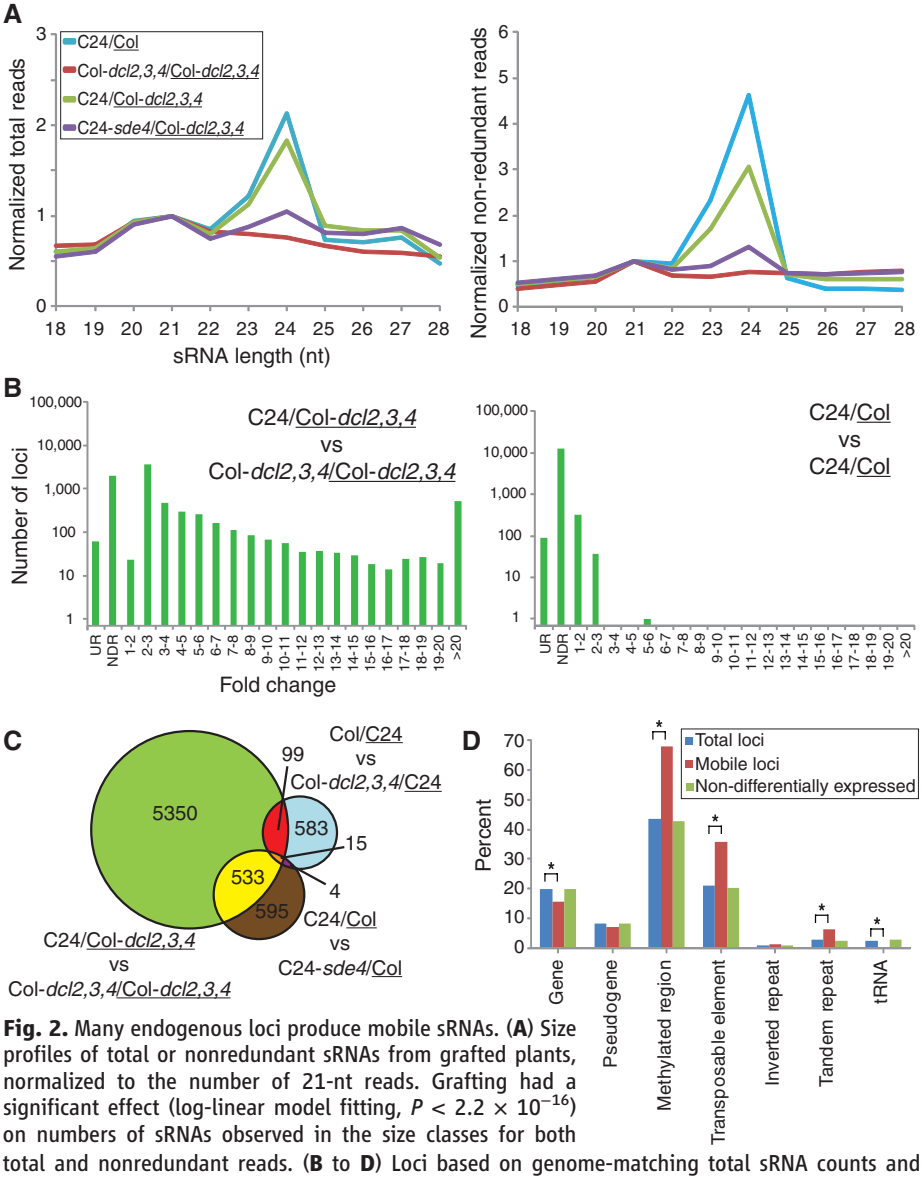


Fig. 2. Many endogenous loci produce mobile sRNAs. (**A**) Size profiles of total or nonredundant sRNAs from grafted plants, normalized to the number of 21-nt reads. Grafting had a significant effect (log-linear model fitting, $P < 2.2 \times 10^{-16}$) on numbers of sRNAs observed in the size classes for both total and nonredundant reads. (**B** to **D**) Loci based on genome-matching total sRNA counts and defined as having a >90% likelihood of expression above background levels, with at least 100 sRNA sequences in one or more of the sRNA libraries used for the pairwise comparison. Mobile sRNA loci are defined as having a >90% likelihood of differential representation. (B) Observed relative changes in mobile sRNA loci overrepresented in C24/Col-*dcl2,3,4* versus Col-*dcl2,3,4*/Col-*dcl2,3,4*. Replicate C24/Col comparisons are shown as a control. Underrepresented (UR) and non–differentially represented (NDR) sRNA loci are shown for comparison. (C) Venn diagram illustrating the extent to which the mobile sRNA loci identified overlap in the three pairwise comparisons. Of 20,679 loci identified from all comparisons, 7179 were associated with mobility and are represented here. (D) Genome features associated with total sRNA, mobile sRNA, and non–differentially represented sRNA loci from the Col-*dcl2,3,4*/C24 versus Col/C24 pairwise comparison were assigned on the basis of annotated features of the *Arabidopsis* genome. Asterisks mark differences between total sRNA and mobile sRNA loci that are statistically significant (Fisher's exact test, $P < 10^{-8}$).

produced by the sequential action of Pol IV and DCL3 in the source tissue, because there were very few 23- and 24-nt sRNAs in C24-*sde4*/Col-*dcl2,3,4* roots (Fig. 2A and table S2A). Moreover, levels of 23- and 24-nt sRNAs in grafted roots of wild-type Col or C24 plants were lower if they were grafted to Col-*dcl2,3,4* or C24-*sde4* shoots rather than wild-type shoots (fig. S2, A and B).

The mobile 23- and 24-nt sRNA populations are likely to be derived from many different genomic loci because they include a high proportion of the nonredundant sequence reads (Fig. 2A, table S2B). This interpretation is reinforced by pairwise comparisons based on the representation of genomic sRNA in data sets from grafted roots (table S3). In each pair the shoots differed in that only one was competent to generate 23- and 24-nt sRNAs that would migrate into the root. Thus, the mobile sRNA loci would be more abundantly represented in C24/Col-*dcl2,3,4* than in Col-*dcl2,3,4*/Col-*dcl2,3,4* data sets; the same would be true for Col/C24 (relative to Col-*dcl2,3,4*/C24) and for C24/Col (relative to C24-*sde4*/Col) sample data sets (Fig. 2B and fig. S3, A and B). The degree of differential representation would be high if all of the locus-specific sRNA were mobile, and would be correspondingly lower if the receiving tissue contained a mixture of RNAs that had either moved from the shoot or were produced in situ. Our analysis focused on loci with more than 100 sRNA reads in at least one data set and with a >90% chance of differential representation based on the analysis of replicates.

In these pairwise comparisons the observed differential locus representation was predominantly by a factor of 2 to 11, which is greater than the difference between biological replicates (Fig. 2B and fig. S3A). Of the 20,679 total sRNA loci, 7179 loci were identified in one or more of the comparisons as producing mobile sRNA (Fig. 2C). In a parallel analysis considering the sRNAs only if they matched a unique site in the Col genome, 795 of 2900 sRNA loci were differentially represented (fig. S3, B and C). It is therefore likely that a high proportion of the genomic sRNA loci produce mobile 23- and 24-nt sRNAs, with a substantial proportion being loci from which most of the sRNAs migrate from the shoot to the root. The sRNA data from the 100 loci with highest posterior likelihoods of producing mobile sRNAs identified in each of these pairwise comparisons are shown in tables S5 to S7.

Although most of the mobile sRNA loci did not correspond to genes (Fig. 2D), a notable exception involved the phosphoribosylanthranilate isomerase (PAI) tryptophan biosynthetic genes of the C24 genotype. C24 has a tail-to-tail inverted repeat (*PAI1-PAI4*) that substitutes for a single copy of *PAI1* in Col (fig. S4A) (*17*). The inverted-repeat DNA has dense cytosine methylation and might silence the other unlinked C24 *PAI* loci through sRNA-directed DNA methylation, although the predicted *PAI* sRNAs have not previously been detected (*17*, *18*).

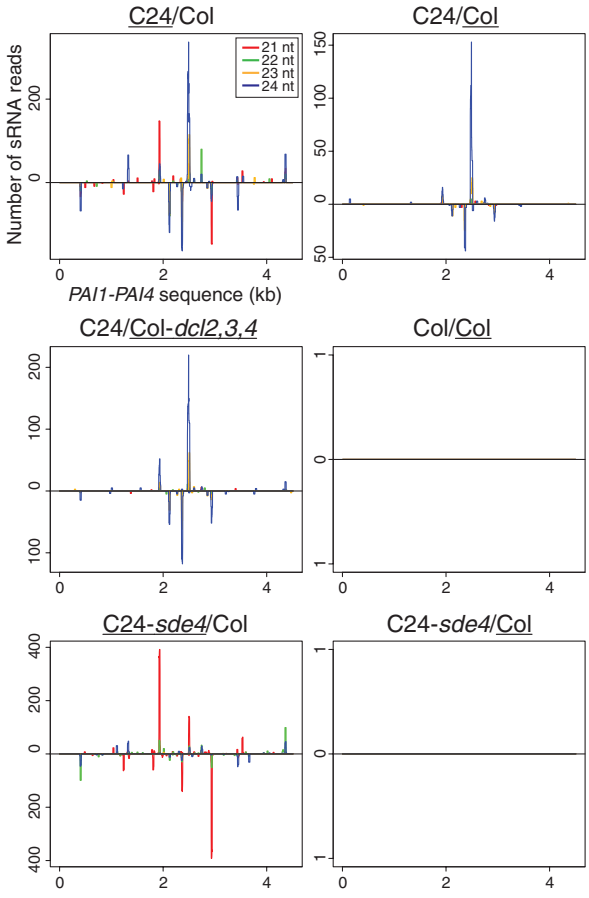


Fig. 3. Movement of PAI inverted-repeat sRNAs. sRNA sequences from grafted *Arabidopsis* were aligned at the first nucleotide position to the *PAI1-PAI4* coding sequence; the positive or negative *y* axis shows the number of reads at each position on either plus or minus strand. See fig. S4 for an independent biological replicate, and table S1 for a summary of sRNA library details.

We detected 21- to 24-nt sRNAs originating from the C24 *PAI1-PAI4* locus that are absent from the Col data sets (Fig. 3 and fig. S4B). The 24-nt *PAI* siRNAs are present in the C24/Col and C24/Col-*dcl2,3,4* data sets, but, consistent with their mobility between the shoot and root, they are absent in C24-*sde4*/Col. This analysis of the *PAI* locus additionally illustrates a preference for mobility of 24-nt rather than 21-nt sRNAs at this locus. In the shoot samples of a C24-*sde4* plant, the *PAI* 24-nt sRNAs were much less abundant than in the wild-type C24 (Fig. 3 and fig. S4B), but the 21-nt species that do not require Pol IV were unaffected or slightly increased. However, these 21-nt PAI sRNAs were almost absent from a C24-*sde4*/Col data set, which indicates that (unlike the 24-nt siRNAs) they are not mobile. This analysis therefore suggests that the bias toward 24-nt species in the mobile sRNAs is not only because they are more numerous than the 21-nt species: There might also be a link between the Pol IV pathway for biogenesis of the 24-nt sRNAs and the transport pathway that mobilizes sRNA between cells and in the phloem.

Mobile sRNAs direct DNA methylation. In Fig. 2D we show that the mobile sRNA loci are typical of genomic loci producing 24-nt sRNAs (*16*). Genes and pseudogenes are associated with only a minority of these loci; the majority are associated with transposons and methylated DNA in genome-wide analyses of *Arabidopsis* chromatin (*19*). Prompted by the particularly strong correlation with methylated DNA, we addressed the possibility that mobile sRNA could direct DNA methylation in the recipient tissue.

We first looked for genomic regions that produce mobile 23- and 24-nt sRNAs and at which DNA methylation in roots is reduced relative to the wild type in a Col-*dcl2,3,4* mutant. We selected eight loci with Col-*DCL2,3,4*-dependent or C24-specific DNA methylation in nongrafted samples that was detectable by digestion of genomic DNA with the methylation-dependent enzyme Mcr BC (fig. S5A). At three of these loci (referred to as mbl D, E, and K), there was hypermethylation in the C24/Col-*dcl2,3,4* root relative to Col-*dcl2,3,4*/Col-*dcl2,3,4*. This result is consistent with a sRNA signal from the shoot moving into the root and guiding DNA methylation (Fig. 4 and fig. S5B). When the silencing signal source had an *SDE4* mutation, the production of 24-nt sRNAs at loci mbl D, E, and K was suppressed (Fig. 4A), and, when grafted to a Col-*dcl2,3,4* root, it was unable to rescue DNA methylation at these loci (Fig. 4B and fig. S5B). It is therefore likely that the mobile sRNA responsible for DNA methylation at these three loci is the 24-nt size class produced by Pol IV.

Loci mbl D and E corresponded to methylated upstream or promoter regions of MuDR-type transposable elements (AT3TE73255 and

A

		C24/ Col	Col-*dcl2,3,4*/ Col-*dcl2,3,4*	C24/ Col-*dcl2,3,4*	C24-*sde4*/ Col-*dcl2,3,4*
21nt	mbl D	9	0	15	0
	mbl E	8	0	13	0
	mbl K	9	0	10	0
22nt	mbl D	6	0	9	0
	mbl E	7	0	7	0
	mbl K	7	0	4	0
23/24nt	mbl D	388	1	269	1
	mbl E	359	1	229	1
	mbl K	232	1	170	0

B

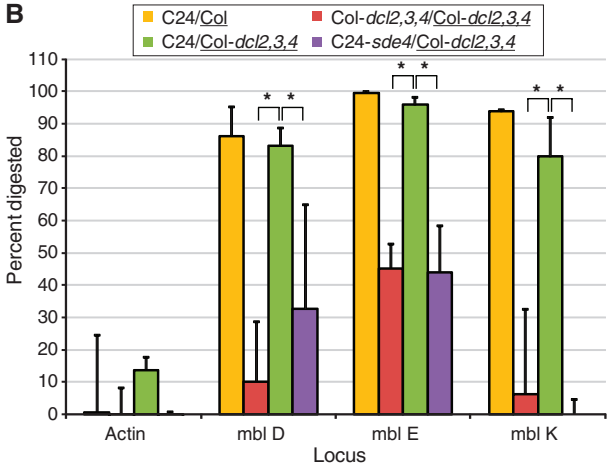


Fig. 4. Endogenous mobile sRNAs direct root DNA methylation. (**A**) sRNA sequence reads corresponding to three mobile sRNA loci that associate with methylated genomic DNA. The sRNA read number for each data set was normalized to 1,000,000 genome-matching sRNA reads. (**B**) DNA digestion by the DNA methylation-dependent enzyme Mcr BC was used to assess the degree of methylation at mobile loci mbl D, E, and K in samples from grafted plants. Asterisks show statistically significant ($P < 0.01$, Student's t test) differences in root DNA methylation that can be attributed to DCL2, 3, 4 or SDE4 function in the shoot. Error bars indicate 95% confidence intervals from two to four biological replicates.

AT2TE04060, respectively), whereas locus mbl K was in a methylated region downstream of a Harbinger-type transposable element (AT5TE00870). The effect of the mobile sRNA on DNA methylation was greater with loci mbl D and K than with locus mbl E (Fig. 4B and fig. S5B). This difference could be because loci mbl D and K have predominantly asymmetric C methylation (CHH) observed in epigenome maps (*20*) that is a diagnostic feature of RNA-directed DNA methylation, whereas locus mbl E has a higher level of symmetric CG methylation that might be independent of sRNA. Bisulfite sequencing of locus mbl E confirmed a partial restoration of CHH and CHG methylation (fig. S6).

Discussion. Our data clearly show that DCL-generated 24-nt sRNAs are mobile in plants. It is likely that these mobile 24-nt sRNAs are produced by DCL3 (*15*) and that they are a subset of the 24-nt sRNAs that have been associated with epigenetic effects (*12*, *16*, *21*, *22*). Long-range mobility of 24-nt RNAs is consistent with findings that sRNA of this size is present in phloem sap samples (*23*) and also with molecular analyses showing a correlation between the presence of 24-nt sRNA and systemic silencing in *Nicotiana benthamiana* (*24*). However, movement of sRNA is not an exclusive property of the 23- and 24-nt sRNA size class; we also found that 22-nt sRNAs produced by DCL2 (*15*, *25*) are mobile (Fig. 1D and fig. S1G). Unfortunately, we cannot use our approach to assay the mobility of 21-nt sRNAs because of the unavailability of viable mutants that completely suppress 21-nt sRNAs. However, given the presence of 21-nt GFP sRNAs in the transgene grafting experiments (Fig. 1D and figs. S1G and S7) and other recent data, we consider it likely that there are mobile sRNAs in all size classes. Mobility is influenced by features of the genomic locus, including the cell type in which the sRNAs accumulate, in addition to features of individual sRNA. The *PAI* locus will help to unravel these complexities because it produces mobile 23- and 24-nt species and nonmobile 21- and 22-nt sRNAs.

In proposing biological roles for mobile 24-nt sRNAs, we have taken into account the possibility that they move into the shoot and to the shoot meristem, where they may introduce epigenetic effects (Fig. 1 and fig. S7) (*1*, *11*). Our finding that very rare sRNAs such as those in the transgene assay (Fig. 1 and fig. S1) can have an effect may also be relevant, because it implies the existence of an amplification mechanism that may also affect endogenous sRNAs. On the basis of these considerations, we propose two possible roles of the mobile 24-nt siRNAs moving into the shoot meristem. These RNAs could allow the meristem to respond to even transient activation of transposons in somatic cells and thereby reinforce epigenetic silencing of these potentially damaging genetic elements in the cells that give rise to the next generation. It could also be that the mobile sRNAs mediate responses to external stimuli and that they initiate epigenetic changes in the meristem that influence competency to flower or adaptation to stress (*26*, *27*). In the developing seed or pollen (*28*, *29*), the mobile sRNAs could induce transgenerational epigenetic changes that adapt progeny to future stress.

References and Notes

1. O. Voinnet, D. C. Baulcombe, *Nature* **389**, 553 (1997).
2. O. Voinnet, P. Vain, S. Angell, D. C. Baulcombe, *Cell* **95**, 177 (1998).
3. J. C. Palauqui, T. Elmayan, J. M. Pollien, H. Vaucheret, *EMBO J.* **16**, 4738 (1997).
4. P. Dunoyer, C. Himber, O. Voinnet, *Nat. Genet.* **37**, 1356 (2005).
5. B. D. Pant, A. Buhtz, J. Kehr, W. R. Scheible, *Plant J.* **53**, 731 (2008).
6. C. A. Brosnan *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 14741 (2007).
7. T. Dalmay, A. J. Hamilton, E. Mueller, D. C. Baulcombe, *Plant Cell* **12**, 369 (2000).
8. S. M. Leisner, R. Turgeon, S. H. Howell, *Plant Cell* **5**, 191 (1993).
9. T. J. Lough, W. J. Lucas, *Annu. Rev. Plant Biol.* **57**, 203 (2006).
10. I. R. Henderson *et al.*, *Nat. Genet.* **38**, 721 (2006).
11. F. Schwach, F. E. Vaistij, L. Jones, D. C. Baulcombe, *Plant Physiol.* **138**, 1842 (2005).
12. A. J. Herr, M. B. Jensen, T. Dalmay, D. C. Baulcombe, *Science* **308**, 118 (2005); published online 3 February 2005 (10.1126/science.1106910).
13. Arabidopsis Genome Initiative, *Nature* **408**, 796 (2000).
14. M. Nordborg *et al.*, *PLoS Biol.* **3**, e196 (2005).
15. Z. Xie *et al.*, *PLoS Biol.* **2**, e104 (2004).
16. R. A. Mosher, F. Schwach, D. Studholme, D. C. Baulcombe, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 3145 (2008).
17. S. Melquist, B. Luff, J. Bender, *Genetics* **153**, 401 (1999).
18. S. Melquist, J. Bender, *Genes Dev.* **17**, 2036 (2003).
19. X. Zhang *et al.*, *Cell* **126**, 1189 (2006).
20. R. Lister *et al.*, *Cell* **133**, 523 (2008).
21. Y. Onodera *et al.*, *Cell* **120**, 613 (2005).
22. X. Zhang, I. R. Henderson, C. Lu, P. J. Green, S. E. Jacobsen, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 4536 (2007).
23. A. Buhtz, F. Springer, L. Chappell, D. C. Baulcombe, J. Kehr, *Plant J.* **53**, 739 (2008).
24. A. J. Hamilton, O. Voinnet, L. Chappell, D. C. Baulcombe, *EMBO J.* **21**, 4671 (2002).
25. A. Deleris *et al.*, *Science* **313**, 68 (2006); published online 1 June 2006 (10.1126/science.1128214).
26. I. Bäurle, L. Smith, D. C. Baulcombe, C. Dean, *Science* **318**, 109 (2007).
27. O. Borsani, J. Zhu, P. E. Verslues, R. Sunkar, J.-K. Zhu, *Cell* **123**, 1279 (2005).
28. R. A. Mosher *et al.*, *Nature* **460**, 283 (2009).
29. R. K. Slotkin *et al.*, *Cell* **136**, 461 (2009).

30. We thank K. Rutherford for data assembly and the Genomics and Bioinformatics core facilities (CRUK Cambridge Research Institute) for Illumina sequencing. Supported by Commonwealth scholarship CACR 2006-115 and Natural Sciences and Engineering Research Council of Canada scholarship PGS D2 358506-2008 (C.W.M.), a Royal Society research professorship (D.C.B.), the Gatsby Charitable Foundation, the European Union Sixth Framework Programme Integrated Project SIROCCO (grant LSHG-CT-2006-037900), and European Research Council advanced investigator grant 233325 REVOLUTION. RNA sequencing raw data are available from Gene Expression Omnibus (GSE20663).

Supporting Online Material

www.sciencemag.org/cgi/content/full/science.1187959/DC1
Materials and Methods
Figs. S1 to S7
Tables S1 to S7
References

4 February 2010; accepted 26 March 2010
Published online 22 April 2010;
10.1126/science.1187959
Include this information when citing this paper.