A Link between RNA Metabolism and Silencing Affecting Arabidopsis Development

Brian D. Gregory,1,2 Ronan C. O’Malley,1,2 Ryan Lister,1,2 Mark A. Urich,1 Julian Tonti-Filippini,3 Huaming Chen,2 A. Harvey Millar,3 and Joseph R. Ecker1,2,*

1Plant Biology Laboratory
2Genomic Analysis Laboratory
The Salk Institute for Biological Studies, La Jolla, CA 92037, USA
3ARC Centre of Excellence in Plant Energy Biology, The University of Western Australia, Crawley, WA 6009, Australia
*Correspondence: ecker@salk.edu
DOI 10.1016/j.devcel.2008.04.005

SUMMARY

MicroRNAs (miRNAs) and small interfering RNAs (siRNAs) are abundant endogenous small RNAs (smRNAs) that control transcript expression through posttranscriptional gene silencing. Here, we show that concomitant loss of XRN4/EIN5, a 5’-3’ exoribonuclease, and ABH1/CBP80, a subunit of the mRNA cap binding complex, results in Arabidopsis plants manifesting myriad developmental defects. We find that ABH1/CBP80 is necessary to obtain proper mature miRNA levels, which suggests this protein affects the miRNA-mediated RNA silencing pathway. Additionally, we show that XRN4/EIN5 affects the levels of a smRNA class that is processed from both sense and antisense strands of ~130 endogenous transcripts that apparently are converted to double-stranded RNA (dsRNA) and subsequently processed. We find that the parent transcripts of these smRNAs accumulate in an uncapped form upon loss of XRN4/EIN5, which suggests that uncapped endogenous transcripts can become smRNA biogenesis substrates. Overall, our results reveal unexpected connections between RNA metabolism and silencing pathways.

INTRODUCTION

RNA silencing represents a pathway that controls gene expression transcriptionally and posttranscriptionally (Baulcombe, 2004). In RNA silencing, production of double-stranded RNA (dsRNA) or self-complementary fold-back structures gives rise to small RNAs (smRNAs) through the activity of DICER or DICER-LIKE (DCL) RNase III-type ribonucleases (Jones-Rhoades et al., 2006). These smRNAs comprise the sequence-specific effectors of RNA silencing pathways that direct the negative regulation or control of genes, repetitive sequences, viruses, and mobile elements (Almeida and Allshire, 2005; Tomari and Zamore, 2005). In plants, these smRNAs are composed of microRNAs (miRNAs) and several classes of endogenous small interfering RNAs (siRNAs), which are differentiated from one another by their distinct biogenesis pathways and the classes of genomic loci from which they arise (Baulcombe, 2004). miRNAs are a class of smRNAs that are 20–24 nucleotides in length and arise from much longer primary transcripts that form characteristic stem-loop structures (Bartel, 2004). In Arabidopsis, the stem-loop precursor of miRNAs is processed by DCL1 ribonuclease to generate a miRNA/miRNA* duplex with a 2 nucleotide (nt) 3’ overhang. The miRNA* is derived from the opposite strand of the stem-loop structure and pairs imperfectly with the miRNA (Bartel, 2004; Jones-Rhoades et al., 2006). The miRNA is then incorporated into an RNA-induced silencing complex (RISC) that has the AGONAUTE1 (AGO1) protein at its core (Baumberger and Baulcombe, 2005; Qi et al., 2005). Plant miRNAs have imperfect but extensive complementarity to their miRNA targets, and typically direct cleavage of these transcripts (Jones-Rhoades and Bartel, 2004; Llave et al., 2002). The targets of many plant miRNAs are mRNAs encoding transcription factors (Jones-Rhoades and Bartel, 2004), and the importance of miRNA-mediated regulation of a number of these target transcripts for proper development is well established (Mallory and Vaucheret, 2006; Willmann and Poethig, 2007). Additionally, plants containing mutations in genes encoding proteins involved in miRNA biogenesis or function (AGO1, DCL1, HEN1, HYL1, SERRATE, and HST) exhibit myriad dramatic developmental abnormalities, exemplifying the importance of this class of smRNAs in growth and differentiation (Boutet et al., 2003; Lobbes et al., 2006; Park et al., 2002; Prigge and Wagner, 2001; Vaucheret et al., 2004; Vazquez et al., 2004a; Yang et al., 2006). Endogenous siRNAs are a class of smRNAs that arise from long dsRNA, which are formed as a product of an RNA-dependent RNA polymerase (RDR), by convergent transcription, or transcription of repetitive elements. siRNAs typically perform autosilencing, in that they target DNA or transcripts corresponding to (or homologous in sequence to) the loci from which they are processed (Baulcombe, 2004). However, the trans-acting siRNAs (tasiRNAs), which are processed from noncoding RNAs known as TRANS-ACTING siRNA (TAS) genes, are the exception in that they posttranscriptionally downregulate protein-coding transcripts from unrelated loci in a fashion reminiscent of the miRNA-directed RNA silencing pathway (Allen et al., 2005; Peragine et al., 2004; Vazquez et al., 2004b; Yoshikawa et al., 2005). During the biogenesis of tasiRNAs, a segment
of the TAS transcript that is defined by miRNA-mediated RISC cleavage is converted by RDR6 to dsRNA, which is successively cleaved by DCL4 into 21 nt siRNAs (Adenot et al., 2006; Allen et al., 2005; Fahlgren et al., 2006; Garcia et al., 2006).

Another class of plant endogenous siRNAs is the heterochromatic siRNAs, which are smRNAs that are mostly 24 nt in length and associated with DNA methylation. The concerted activity of plant-specific DNA-dependent RNA polymerases, PolIVa and PolIVb, correlates with the accumulation of 24 nt heterochromatic siRNAs via dsRNA formation by RDR2 and DCL3-mediated processing (Chan et al., 2005; Herr et al., 2005; Onodera et al., 2005; Xie et al., 2004). A fraction of these siRNAs associate with AGO4 to form a silencing complex thought to direct sequence-specific methylation events (Chan et al., 2005; Qi et al., 2006). Subsequently, this siRNA-directed DNA methylation can result in maintained transcriptional gene silencing at loci from which the smRNAs arise (Herr et al., 2005; Kanno et al., 2005; Onodera et al., 2005; Pontier et al., 2005).

The 5’ cap structure and the 3’ poly(A) tail are the two boundary marks that define the extreme borders of a eukaryotic mRNA. In the eukaryotic cell nucleus, the 5’ cap is recognized by the nuclear mRNA cap-binding complex (CBC). CBC is a heterodimeric complex that consists of a small (CBP20) and a large (CBP80) protein subunit (Mazza et al., 2001) that plays numerous roles in RNA metabolism (Aguilera, 2005). The Arabidopsis homolog of CBP80 is encoded by the ABH1 gene. Plants harboring a genetic lesion in ABH1 (abh1-1 mutant plants) manifest an ABA-hypersensitive regulation of seed germination phenotype that suggests a link between mRNA metabolism and ABA signaling. Interestingly, the characterization of abh1-1 mutant plants resulted in the discovery that Arabidopsis CBC is entirely composed of the ABH1/CBP80-AtCBP20 heteroduplex (Hugouvieux et al., 2001).

In Arabidopsis, the 5′-3′ exoribonuclease XRN4/EIN5 (hereafter referred to as EIN5) acts as an mRNA-degrading enzyme that is involved in the decay of specific transcripts that include the 3′ products of miRNA-mediated cleavage (Kastenmayer and Green, 2000; Souret et al., 2004). Two of its other mRNA targets, EBF1 and EBF2, encode F box proteins that target the ubiquitin/proteasome-meditated turnover of EIN3, a key transcription factor mediating gene expression regulated by the phytohormone ethylene. These specific EIN5 targets account for the ethylene insensitive (Ein−) phenotype exhibited by e1n5 mutant plants (Omedo et al., 2006; Potuschak et al., 2006).

Here, through analysis of the developmental defects manifested on e1n5-6 abh1-1 double mutant plants, we demonstrate unexpected roles in RNA silencing pathways for two proteins involved in general RNA metabolism, ABH1/CBP80 (hereafter referred to as ABH1) and EIN5. First, we find that loss of ABH1 decreases the levels of mature miRNAs, suggesting that this protein functions in the miRNA-mediated RNA silencing pathway. Additionally, we show that EIN5 affects the abundance of a distinct class of mostly 21 nt smRNAs that in many cases emanate from the entire length of endogenous functionally annotated transcripts, which often accumulate in an uncapped form in e1n5 mutant plants. Taken together, our results suggest that an additional fate for endogenous uncapped transcripts is shuffling into an RNA silencing pathway where they become smRNA-biogenesis substrates.

## RESULTS

### Ethylene Insensitivity of ein5 Mutant Plants Is Suppressed by the abh1 Mutation

The position that ABH1 occupies on mRNA molecules suggests that it may counteract the function of 5′-3′ exoribonucleases such as EIN5 (Aguilera, 2005). Therefore, in order to determine if mutation of ABH1 can suppress the ethylene insensitivity of ein5 mutant plants, we generated double mutants of e1n5 and abh1. We found that loss of ABH1 (abh1-1) almost completely suppresses the Ein− phenotype of e1n5 mutant plants (Figures 1A–1E). Previously, we and others have demonstrated that the ethylene sensitivity of e1n5 mutant plants is a consequence of EBF1 and EBF2 mRNA accumulation (Omedo et al., 2006; Potuschak et al., 2006). To characterize the effect of ABH1 on EBF1 and EBF2 mRNA levels, we performed northern blot analysis using total RNA from 3-day-old etiolated seedlings of wild-type Columbia (Col-0) ecotype, abh1-1, e1n5-6, and e1n5-6 abh1-1 grown in hydrocarbon-free air or 10 parts per million (ppm) ethylene for various time periods (Figure 1F). We found that abh1-1 seedlings maintained levels of EBF1 and EBF2 mRNA similar to those of wild-type Col-0 (Figure 1F). As expected, e1n5-6 seedlings accumulated a significantly increased level of both mRNAs in air and upon ethylene treatment (Figure 1F). Conversely, e1n5-6 abh1-1 seedlings exhibited levels of both EBF1 and EBF2 mRNA similar to those of wild-type Col-0 (Figure 1F), suggesting that loss of ABH1 restores proper maintenance of these two mRNAs in e1n5 mutant plants. Taken together, these results reveal that ABH1 is required for the accumulation of increased levels of EBF1 and EBF2 mRNAs in the absence of EIN5, and abh1 mutation can act as a genetic suppressor of the hormone response phenotype manifested by e1n5 mutant plants.

### e1n5-6 abh1-1 Plants Manifest Developmental Defects Similar to Those Observed for miRNA Pathway Mutants

During analysis of the ethylene response of e1n5-6 abh1-1 mutant plants, we noticed they manifested severe developmental defects. Previously, abh1-1 and e1n5-6 single mutant plants had been found to present a serrated leaf margin phenotype (Hugouvieux et al., 2002; Omedo et al., 2006), which we have determined is strongly enhanced in e1n5-6 abh1-1 double mutant plants (Figure 2Q). In addition, a small percentage of abh1-1 (~7%) and e1n5-6 (~1%) single mutant seedlings demonstrated fused cotyledons when compared with wild-type Col-0 plants, which never displayed this phenotype. This phenotype was also greatly enhanced in e1n5-6 abh1-1 double mutant seedlings, where we observed over 16% of seedlings with fused cotyledons (Figures 2A–2P). e1n5-6 abh1-1 double mutant flowers also manifested defects not observed for the other three genotypes, including extranumerary petals and bending of the gynoecium (Figure S1, see the Supplemental Data available with this article online). Finally, abh1-1 and e1n5-6 single mutant plants were also found to manifest an altered phyllotactic pattern, where the internode length is severely decreased and results in multiple fruits emanating from the same node, which was observed at a far greater frequency in e1n5-6 abh1-1 double mutant plants (Figures 2R–2U). In fact, we observed numerous nodes from which more than three fruits emanated on double mutant plants, while...
this phenotype was never manifested by plants of the other genotypes (Figure 2V). Interestingly, Prigge and Wagner (2001) previously determined that mutation of the *SERRATE* gene in *Arabidopsis* resulted in plants that presented similar phenotypes to those observed for *ein5-6 abh1-1*. Subsequently, it was determined that *SERRATE* is required for proper miRNA biogenesis (Lobbes et al., 2006; Yang et al., 2006). Therefore, the developmental defects manifested by *ein5-6 abh1-1* (and to a lesser extent *abh1-1* and *ein5-6*) suggested that loss of ABH1 and EIN5 might affect the miRNA-mediated RNA silencing pathway.

**Loss of ABH1 and EIN5 Disrupts the miRNA-Mediated RNA Silencing Pathway**

To determine if EIN5, ABH1, or both have effects on the miRNA-mediated RNA silencing pathway, we carried out an unbiased analysis of the transcriptome using whole-genome tiling microarrays (Chekanova et al., 2007; Kapranov et al., 2002; Figure S2 and Tables S1–S6). Interestingly, tiling microarray analysis revealed that the levels of a number of primary *MIRNA* transcripts were significantly increased in *ein5-6 abh1-1* (18 increased) and *abh1-1* (19 increased) compared with *ein5-6* and wild-type Col-0 plants (Figure 3A and Table S7). To establish that the stem-loop structure containing the miRNA/miRNA* duplex was part of the primary *MIRNA* transcript upregulated in our tiling array analysis and to provide validation for this class of mRNAs, we performed reverse transcription quantitative polymerase chain reaction (RT qPCR) with primer sets homologous to sequences just upstream and downstream of this structural moiety for a subset of the primary *MIRNA* transcripts that were statistically upregulated from our microarray studies (Table S7).

We found that *ein5-6 abh1-1* and *abh1-1* mutant plants accumulated significantly more *MIRNA158a, MIRNA164b, MIRNA167a*, and *MIR168a* transcripts that contained the stem-loop structural moiety than did wild-type Col-0 or *ein5-6* plants (Figure 3B), suggesting that it is the loss of ABH1 that results in accumulation of this class of mRNAs (Figures 3A and 3B and Table S7). Furthermore, tiling array analysis revealed that a number of miRNA-target mRNAs accumulated in *ein5-6 abh1-1* double
mutant plants to levels far exceeding those observed for the three other genotypes, two examples of which can be seen in Figure 3C (see also Table S5). We validated these two examples, CUC1 and HAP2C, using RT qPCR (Figure 3D). As expected based on the tiling microarray results, we determined that ein5-6 abh1-1 double mutant plants accumulated increased levels of the miRNA-target transcripts CUC1 and HAP2C compared with the other three genotypes (Figure 3D). These results provided further validation of our tiling microarray analysis and suggest that the concomitant loss of ABH1 and EIN5 results in the accumulation of miRNA-target mRNAs. Previously, it was determined that these two classes of mRNAs (primary MIRNA transcripts and miRNA-target mRNAs) were significantly increased in Arabidopsis plants lacking proteins involved in miRNA biogenesis and function (Llave et al., 2002; Lobbes et al., 2006). Overall, these results suggest that concomitant loss of ABH1 and EIN5 disrupts miRNA-mediated RNA silencing.

ABH1 Affects the miRNA-Mediated RNA Silencing Pathway

Next, we wanted to obtain a genome-wide view of the effects of ABH1, EIN5, or both on the smRNA populations of Arabidopsis. In order to accomplish this, we employed deep sequencing of smRNA samples from wild-type Col-0, abh1-1, ein5-6, and ein5-6 abh1-1 immature flower buds using an Illumina Genetic Analyzer (GA) (Figures S3 and S4). A total of 3,264,170 (1,036,593 unique), 4,446,687 (1,400,713 unique), 3,173,518 (1,030,343 unique), and 2,910,019 (959,915 unique) smRNAs were identified from the wild-type Col-0, abh1-1, ein5-6, and ein5-6 abh1-1 sequencing libraries, respectively (Figure S4). We found that the majority (~86%) of smRNAs in all genotypes sequenced were 21–24 nt in size (data not shown). Upon focusing our analysis on the sequencing data encompassing the 21 nt size class of smRNAs, we noticed that this smRNA size class was overall underrepresented (~20%) in abh1-1 single mutant
plants compared with wild-type Col-0 (Figure S5A). These findings indicate that ABH1 is required to obtain proper levels of a class of smRNAs that are 21 nt in length. Conversely, we found that ein5-6 abh1-1 double mutant plants accumulated higher levels of 21 nt smRNAs than the other three genotypes (Figure S5A), which suggests that EIN5 negatively regulates a class of 21 nt smRNAs. To obtain a genome-wide view of the 21 nt smRNA population, we first normalized the number of sequence reads at all genomic locations corresponding to 21 nt smRNAs by the total reads sequenced for the matching library, then parsed the values into 250 kilobase (kb) bins, and plotted these locations across the five Arabidopsis nuclear chromosomes. We found that for abh1-1 and ein5-6 abh1-1 plants, there was a reduction on average of ~60% in the levels of highly abundant 21 nt smRNAs, which correspond mostly to miRNAs (Figure 4A). These results suggest that ABH1 is required to obtain proper miRNA levels. Interestingly, we also observed that ein5-6 abh1-1, and to a lesser extent ein5-6, mutant plants accumulated 21 nt smRNA clusters that mapped to a number of different gene-rich locations throughout the Arabidopsis genome compared with the other two genotypes (Figure 4A, red arrows), suggesting that EIN5 negatively affects a population of 21 nt smRNAs (see below). To determine if these two proteins regulate the levels of specific classes of 21 nt smRNAs, we used our sequencing data encompassing another size class of smRNAs, 24 nts, as a control. This class of smRNAs has been demonstrated to consist mostly of heterochromatic siRNAs (Kasschau et al., 2007; Rajagopalan et al., 2006). We found that on
a genome-wide scale, the levels of 24 nt smRNAs were not significantly different between the four sequenced genotypes (Figure S5B), indicating that EIN5 and ABH1 specifically regulated certain classes of 21 nt smRNAs.

To validate our sequencing data and further characterize ABH1’s effect on miRNA levels, we performed smRNA-enriched northern blot analysis on samples from wild-type Col-0, abh1-1, ein5-6, and ein5-6 abh1-1 immature flower buds. These experiments revealed that abh1-1 and ein5-6 abh1-1 plants accumulated 62%–83% less miRNA for all those interrogated (miRNA156, 158, 159, 164, 167, 169, and 390) compared with wild-type Col-0 and ein5-6 (Figures 4B and 4C and Figure S6). Furthermore, this reduction in the levels of miRNAs upon loss of ABH1 function was corroborated by our smRNA sequencing data when the number of reads for specific miRNAs was normalized to the total sequenced read numbers (Figure S7A). Taken together, these results suggest that ABH1 is required to obtain proper miRNA levels. Notably, the smRNA sequencing data...
demonstrated that the normalized levels of some well-characterized 24 nt siRNAs were not reduced by the loss of ABH1 function (Figure S7B), which suggests that this protein is specific in its effect on miRNA levels.

To test if the reduction in miRNA levels observed for abh1-1 and ein5-6 abh1-1 plants affected miRNA-mediated target mRNA cleavage, we employed a modified 5′-rapid amplification of cDNA ends (5′-RACE) protocol (Llave et al., 2002). Using this methodology we found that for abh1-1 and ein5-6 abh1-1 plants, there was decreased miRNA-mediated cleavage of the target mRNAs interrogated compared with those of wild-type Col-0 and ein5-6 (Figures 4D and 4E). Overall, these results demonstrate that ABH1 is required to obtain wild-type levels of miRNAs, and suggest it has effects on the miRNA-mediated RNA silencing pathway through the reduction of this class of smRNAs.

Next, we wanted to determine if mutation of ABH1 could enhance the developmental defects manifested by plants containing a hypomorphic allele of AGO1, because AGO1 is known to function in the miRNA-mediated RNA silencing pathway. To do this, we generated double mutants between abh1-8 (a null allele) and plants containing a hypomorphic allele of AGO1, ago1-38. A hypomorphic allele of AGO1 was used because plants harboring null mutations in this gene do not survive. We then analyzed the abh1-8 ago1-38 double mutant plants for novel growth defects or enhancement of developmental abnormalities manifested in abh1-8 or ago1-38 single mutant plants (Figure 5). We found that approximately 15% of abh1-8 ago1-38 double mutant seedlings exhibited fused cotyledons, which was more than double the amount manifested by either abh1-8 or ago1-38 single mutant plants (Figures 5B–5E, 5G–5J, 5L–5O, and 5P). Furthermore, none of the wild-type Col-0 seedlings analyzed exhibited this developmental abnormality (Figures 5A and 5P). Additionally, we found that approximately 12% of abh1-8 ago1-38 double mutant seedlings never developed a root, even after 2 weeks of growth (Figures 5E, 5J, 5O, and 5Q). Interestingly, this developmental defect was never observed for wild-type Col-0, abh1-8, or ago1-38 single mutant plants (Figures 5A–5C, 5F–5H, 5K–5M, and 5Q). We also found that abh1-8 ago1-38 double mutant plants presented an enhancement of the altered phyllotaxy wherein the internode length is severely decreased, resulting in multiple fruits emanating from the same node, as also demonstrated by both abh1-8 and ago1-38 single mutant plants. Specifically, ~16% of all abh1-8 ago1-38 double mutant plants manifested more than three fruits emanating from the same node, which is more than double the amount witnessed for wild-type Col-0, abh1-8, or ago1-38 single mutant plants (Figures 5R and 5S). These results demonstrate that the abh1-8 mutation is able to enhance the developmental defects manifested by Arabidopsis plants harboring a hypomorphic mutant allele of AGO1, which suggests that like AGO1, ABH1 also affects the miRNA-mediated RNA silencing pathway.

EIN5 Affects the Levels of a Class of smRNAs Processed from Functionally Annotated mRNAs

As noted previously, we observed that in the absence of EIN5 function, clusters of 21 nt smRNAs accumulate in euchromatic regions of the genome (Figure 4A). Therefore, we filtered the sequenced smRNAs for 21 nt smRNAs present at least twice in abh1-1, ein5-6, or ein5-6 abh1-1, but not found in Col-0. These three lists were further filtered for functionally annotated mRNAs from which at least two independent smRNAs were processed, which identified 67, 253, and 235 such transcripts from abh1-1, ein5-6, and ein5-6 abh1-1 plants, respectively (Figure 6A and Tables S8–S10). We found that 156 of these transcripts overlap between the ein5-6 and ein5-6 abh1-1 lists, with 23 of them also overlapping with those found for abh1-1 (Table S11). Since the smRNAs observed (Figure 4A) were only present in the absence of EIN5, we focused on the 133 transcripts that overlapped between the lists from the ein5-6 and ein5-6 abh1-1 smRNA libraries (Figure 6A, Tables S11 and S12). We found that clusters of mostly 21 nt smRNAs were generated from both sense and antisense strands of these 133 functionally annotated transcripts (Figures 6B and 6C), suggesting that these mRNAs had been converted to dsRNA and then processed in 21 nt increments in the absence of functional EIN5 (Figures 6B and 6C). In many cases, these EIN5-affected smRNAs are processed from the entire length of the parental transcript in multiple 21 nt registers, which suggests there is not an ordered processing event or “common” mRNA terminus that is consistently required for biogenesis of these smRNAs (Figure 6B and Figure S8). There were also numerous examples where the EIN5-affected smRNAs accumulated to significantly higher levels in ein5-6 abh1-1 compared with ein5-6 plants (Figures 6B and 6D, Figure S9, and Table S12). In fact, these smRNAs compose ~19% of the entire 21 nt smRNA population of ein5-6 abh1-1 double mutant plants (Figure S10), which suggests that concomitant loss of ABH1 and EIN5 results in the enhanced processing or stability (or both) of these smRNAs. Interestingly, we observed that often low levels of EIN5-affected smRNAs were present in wild-type plants (Figures 6B and 6D, Figure S9, and Table S12), which indicates that this class of smRNAs is indeed processed under normal conditions and is not just a byproduct of unnatural consequences (i.e. ein5 mutation). Overall, these results suggest that the 5′-3′ ribonuclease EIN5 negatively regulates the levels of a class of endogenous smRNAs that are processed from functionally annotated Arabidopsis gene transcripts.

Endogenous Uncapped Transcripts Likely Act as smRNA Biogenesis Substrates

It had been previously reported that EIN5 likely antagonizes RDR-dependent RNA silencing of exogenously introduced transgenes through degradation of “aberrant” mRNAs that may lack a 5′ terminal cap moiety and serve as biogenesis substrates for smRNA production (Gazzani et al., 2004). Many of the features exhibited by our endogenous EIN5-affected smRNA-generating transcripts (Figure 6) suggest that they may also lack a 5′ terminal cap moiety, and therefore are regulated in a similar manner to what was previously witnessed for transgene mRNAs. To detect uncapped 5′ RNA ends, we used a modified 5′-RACE protocol (Llave et al., 2002). Full-length, decapped mRNA corresponding to the EIN5-affected smRNA-generating transcripts (verified by cloning and sequencing of the cDNA; Figure S11) consistently accumulated in ein5-6 and ein5-6 abh1-1 plants (Figure 7A). In order to generate a more comprehensive, genome-wide view of all 133 smRNA-generating transcripts, we designed an Illumina GA sequencing-based 5′-RACE assay termed genome-wide mapping of uncapped
Figure 5. The abh1 Mutation Enhances Developmental Defects Associated with Plants Containing a Hypomorphic Genetic Lesion in AGO1, ago1-38

(A–E) One-week-old seedlings of wild-type Col-0 (A), abh1-8 (B), ago1-38 (C), and abh1-8 ago1-38 (D and E) double mutant plants, the latter of which manifests a cotyledon fusion phenotype and has no root (E).

(F–J) Ten-day-old seedlings of the same genotypes as in (A–E).

(K–O) Two-week-old seedlings of the same genotypes as in (A–E).

(P) The percentage of wild-type Col-0, abh1-8, ago1-38, and abh1-8 ago1-38 double mutant seedlings that manifest the cotyledon fusion defect, which was scored using 1-week-old plants.

(Q) The percentage of wild-type Col-0, abh1-8, ago1-38, and abh1-8 ago1-38 double mutant seedlings that do not develop a root, which was scored using 2-week-old plants.

(R) Inflorescence stems of plants of the indicated genotypes.

(S) The percentage of wild-type Col-0, abh1-8, ago1-38, and abh1-8 ago1-38 double mutant inflorescence stems that manifest the developmental defect wherein multiple fruits (two, three, or more than three) emanate from the same node.
and cleaved transcripts (GMUCT) (see Figure S12 and Experimental Procedures for details). We then applied GMUCT to unmodified, poly-adenylated RNA samples from wild-type Col-0 and ein5-6 immature flower buds. As expected for this modified 5’-RACE protocol, we were able to identify known miRNA-mediated target mRNA cleavage sites from the obtained sequencing data (Figure S13), thereby providing validation of this methodology. Using this novel assay we determined that at least 46% of the EIN5-affected siRNA-generating transcripts accumulate in an uncapped form in the absence of EIN5 function (Figures 7B and 7C). Overall, we validated ~50% of the total population of EIN5-affected smRNA-generating transcripts as accumulating in an uncapped form in ein5-6 mutant plants through the combination of 5’-RACE and GMUCT (Figure 7B), with two of these mRNAs being validated by both methodologies. These results demonstrate that EIN5 is responsible for the removal of the uncapped forms of these 133 transcripts. Thus, we have identified a class of endogenous EIN5-affected mRNAs that are processed into smRNAs by an RNA silencing pathway similar to one previously observed to regulate aberrant transgenic RNAs (Gazzani et al., 2004; Figure 7). More specifically, the endogenous mRNAs also accumulate in an uncapped form, which is likely the defining feature that shunts them into an RNA silencing pathway where they are processed into mostly 21 nt smRNAs.

**DISCUSSION**

The mRNA CBC is known to increase the efficiency of pre-mRNA splicing and non-sense-mediated decay (NMD) (an mRNA surveillance pathway that generally eliminates messenger RNAs that prematurely terminate translation) through direct interaction with proteins involved in these two important processes (Aguilera, 2005). Therefore, we hypothesize that because ABH1 is a subunit of the CBC, its role in the miRNA-mediated mRNA silencing pathway could be to increase the efficiency of processing of these smRNAs from the initial primary MIRNA transcript. Specifically, we suggest that CBC bound to the 5’ end of primary MIRNA transcripts interacts with one of the proteins of the miRNA biogenesis complex (SERRATE, DCL1, and/or HYL1), thus increasing the likelihood of proper processing of these important smRNAs. The overlapping phenotypic similarities between abh1 and se mutant plants (Bezerra et al., 2004; Figure 2 and Figure 5) offer a tantalizing suggestion that ABH1 may interact directly with SERRATE to ensure proper and efficient processing of primary MIRNA transcripts into the biologically active miRNA products.

Interestingly, consistent with our findings and providing evidence that a role for ABH1 in the miRNA-mediated RNA silencing pathway may also be conserved in animals, Parry et al. (2007)...
A recent performed a whole-genome RNA interference (RNAi) screen in *C. elegans* that suggested a function for the worm homolog of ABH1 (F37E3.1) specifically in the *let-7* miRNA-mediated RNA silencing pathway. Upon performing smRNA northern blots using RNA isolated from young adult worms that were fed an RNAi knockdown construct targeting F37E3.1, only a slight reduction in mature *let-7* levels compared with those of wild-type animals was detected. Based on these results, this group concluded that the miRNA-debilitating gene inactivation of F37E3.1 likely abrogated gene function downstream of the expression and processing of *let-7*. But here, we demonstrate that *Arabidopsis* ABH1 is intimately required for accumulation of mature miRNA levels, which suggests a role for this protein either in or upstream of miRNA processing. These discrepancies may result from differences in the miRNA biogenesis pathways between the two organisms (*C. elegans* and *Arabidopsis*) examined in these studies, or more likely are due to differences in methodologies employed: RNAi knockdown (Parry et al., 2007) versus stable genetic mutation (abh1 mutant plants [this study]). The latter suggestion is supported by the fact that even complete and stable loss of ABH1 in *Arabidopsis* caused an ~66% reduction in mature miRNA levels, rather than a complete loss (Figure 4 and Figures S6 and S7). Therefore, transient and incomplete reduction of F37E3.1 levels by RNAi likely only causes a slight decrease in mature miRNA levels, which would result in the differing interpretations of the results. Nevertheless, in combination these studies suggest an unexpected role for ABH1 in the miRNA-mediated RNA silencing pathway of eukaryotic organisms.

Additionally, we demonstrate that EIN5 affects the levels of a class of mostly 21 nt smRNAs that are processed from functionally annotated *Arabidopsis* gene transcripts (Figure 6). We find that the smRNA-generating mRNAs accumulate in an uncapped form in the absence of EIN5 function (Figure 7), which suggests that this may be the defining feature of these transcripts that shuttles them into the RNA silencing pathway, much like was previously demonstrated to occur for capless mRNAs derived from introduced transgenes (Gazzani et al., 2004). An alternative hypothesis would be that loss of EIN5 function allows cryptic antisense transcripts specific to the 133 EIN5-affected smRNA-generating loci to accumulate, thereby resulting in dsRNA formation and subsequent DICER processing into smRNAs. We strongly disfavor this latter hypothesis because our tiling microarray analysis did not reveal upregulated antisense transcripts corresponding to any of the 133 EIN5-affected smRNA-generating loci. Additionally, none of the EIN5-affected smRNAs map to the intronic regions of transcripts from which they are processed (Figure 6B), which strongly suggests that the mature mRNA molecule itself is acting as the template for synthesis of the antisense strand, likely in an RDR-dependent manner. Furthermore, we observe an enhancement in the levels of many of the EIN5-affected smRNAs in ein5-6 abh1-1 double mutant plants compared with those of any of the other three genotypes used in this study (Figures 6B and 6D, Figure S9, and Table S12), which suggests that the processing and/or stability of these smRNAs is enhanced by the absence of both ABH1 and EIN5 function. While these results led us to prefer the latter explanation, wherein enhanced processing of the EIN5-affected smRNAs results from an increased availability of their biogenesis substrate, these models are not mutually exclusive. Specifically, based on the data from the ein5-6 abh1-1 double mutant, we hypothesize that the loss of ABH1 function likely reduces protection of the 5’ cap structure from removal by decapping enzymes, while at the same time, the lack of functional EIN5 stabilizes these uncapped RNA intermediates. Overall, these results support a model in which the 133 transcripts that accumulate in an uncapped form in the absence of EIN5 function are shunted into an RDR-dependent RNA silencing pathway where they are processed into smRNAs. Furthermore, the findings presented herein for the single and double mutants suggest an integration of EIN5 and ABH1 functions in protecting these 133 endogenous transcripts from this proposed fate.

Recently, it was demonstrated that the terminal 5’ nucleotide of *Arabidopsis* smRNAs specifies the AGO protein with which they associate. Specifically, it was found that smRNAs with a 5’ terminal adenosine preferentially associate with AGO2 and AGO4, smRNAs with a 5’ terminal cytosine preferentially associate with AGO5, and smRNAs with a 5’ terminal uridine
preferentially associate with AGO1 (Mi et al., 2008). From deep smRNA sequencing, we find that the majority of 21 nt smRNAs from immature flower buds begin with a 5′ adenosine (Figure S14 and Table S13), and this is also observed for the entire population of sequenced smRNAs (data not shown). Therefore, the majority of smRNAs in these plant organs associate with AGO2 and AGO4. Guanine is the next most abundant 5′ terminal nucleotide for 21 nt smRNAs, while thymine and cytosine are the least favored 5′ terminal nucleotides for 21 nt smRNAs (Figure S14 and Table S13). Interestingly, the distribution of 5′ terminal nucleotides among the 21 nt EIN5-affected smRNAs is not significantly different from that in the total population of 21 nt smRNAs (Figure S14 and Table S13). Overall, these results suggest that EIN5-affected smRNAs are quite diversified at the 5′ terminal nucleotide position, which likely allows them to associate with a number of different Arabidopsis AGO proteins.

As noted above, we observed numerous cases of EIN5-affected smRNAs that correspond to gene transcripts that are also detectable in wild-type plants (Figures 6B and 6D, Figure S9, and Table S12), which indicates that the processing of this class of smRNAs is not simply a byproduct of genetic mutation of EIN5 and ABH1. An intriguing possibility is that these smRNAs may have regulatory functions, and that they might be processed more readily and/or accumulate in response to specific physiological or environmental conditions in which the function of EIN5 is negatively regulated. In this model, the expansion of EIN5-affected smRNAs could thus provide an additional population of smRNAs to regulate target mRNAs with sites of sequence complementarity in both cis and trans. Although evidence that EIN5-affected smRNAs are able to regulate mRNAs posttranscriptionally is still lacking, this tantalizing possibility warrants further examination. Overall, these findings suggest that there is still much to learn concerning RNA silencing and the regulation of these posttranscriptional regulatory pathways.

EXPERIMENTAL PROCEDURES

Plant Materials

The Col-0 ecotype of Arabidopsis was used in this study. The abh1-1 and ein5-6 mutants were previously identified (Hugouvieux et al., 2001, 2002; Olmedo et al., 2006), abh1-8 (this study) was recovered from the Salk T-DNA collection (Alonso et al., 2003), ago1-38 comes from an ethyl methanesulfonate (EMS)-mutagenized population, and contains a Gly to Arg substitution at position 186 (seeds kindly provided by E. Meyerowitz, California Institute of Technology).

Genetic Analysis of Double Mutants

Double mutant plants were generated by genetic crosses, and homozygous lines were identified by PCR-based genotyping. All primers used for PCR-based genotyping of mutant plants used in this study can be found in Table S14. In the case of ein5-6 genotyping, the amplified PCR products were subsequently cleaved with MboII. In the case of ago1-38 genotyping, the amplified PCR products were subsequently cleaved with EcoNI.

RNA Analyses

For all experiments performed herein, immature flower bud clusters were collected for RNA isolation using the RNeasy Plant Mini Kit (QiAGEN, Valencia, CA). Low-molecular-weight RNA was purified from total RNA, and northern blots for mature miRNAs were performed as previously described (Peragine et al., 2004). Northern blot analysis of EBF1 and EBF2 mRNA levels was performed as previously described (Olmedo et al., 2006). Transcripts were quantified by RT qPCR using the comparative threshold cycle method (ΔΔCt, primers listed in Table S14), using Actin 2 (At3g18780) as the endogenous reference.

5′-RACE

5′-RACE and cloning of 5′-RACE products was carried out using the GeneRacer Kit (Invitrogen, Carlsbad, CA) as previously described (Kasschau et al., 2003). Briefly, there were no modifications made to the total RNA used for the 5′-RACE experiments. The 5′-adapter molecule was immediately ligated onto the total RNA population, thereby isolating only mRNAs or mRNA fragments with a 5′ end with a free 5′ phosphate. 5′-RACE products of miRNA-directed cleavage (Figure S7) were quantified by RT qPCR using the comparative threshold cycle method (ΔΔCt, primers listed in Table S14), using Actin 2 (At3g18780) as the endogenous reference.

Microarray Analysis

Tiling microarray analysis was carried out as previously described (Chekanova et al., 2007). For more detailed methodology see the Supplemental Experimental Procedures.

smRNA Library Construction

smRNA libraries were constructed as per the manufacturer’s instructions (Illumina Inc., San Diego, CA). For more detailed methodology see the Supplemental Experimental Procedures and Figure S3.

GUMCT Sequencing Library Construction

Briefly, GUMCT libraries were constructed by adapting a modified 5′-RACE protocol (see 5′-RACE section above) for sequencing on an Illumina GA. For more detailed methodology see the Supplemental Experimental Procedures and Figure S12.

High-Throughput Sequencing

smRNA and GUMCT libraries were sequenced using the Illumina GA as per the manufacturer’s instructions (Illumina Inc., San Diego, CA), except that sequencing of GUMCT libraries was performed for 50 cycles to yield longer sequences that are more amenable to unambiguous mapping to the Arabidopsis genome sequence.

Mapping smRNA Reads

Sequence information was extracted from the image files with the Illumina Firecrest and Bustard software applications. Prior to alignment of the smRNA reads, a custom Perl script was used to identify the first seven bases of the 3′ adaptor sequence, and the read was truncated up to the junction with the adaptor sequence. No further analysis was performed on reads that did not contain adaptor sequence, as those reads lacking an adaptor cannot be precisely sized. The reads were then mapped to the Arabidopsis Col-0 reference genome (TAIR v.7) with the Illumina Eland application. Eland aligns 32 base pair or shorter reads with up to two mismatches to a reference genome. However, only perfect matches were accepted because shorter reads have a higher tendency to falsely map than longer reads. Since Eland does not return positions for reads that map to multiple positions in the reference genome, we utilized the Basic Local Alignment Search Tool (BLAST) to map these nonunique reads, using a word size of 10 and expectation value of 10. Once again, only perfect matches were accepted, because these shorter reads have a higher tendency to falsely map than longer reads.

Mapping GUMCT Reads

In order to avoid omitting reads mapping within unannotated transcripts, 36 base sequence reads were aligned to the Arabidopsis reference genome sequence (TAIR v.7) with the Eland application.

ACCESSION NUMBERS

All raw microarray data (CEL files) for expression analyses (wild-type Col-0, abh1-1, ein5-6, and ein5-6 abh1-1), as well as smRNA (Col-0, abh1-1, ein5-6, and ein5-6 abh1-1 libraries) and GUMCT (wild-type Col-0 and ein5-6 libraries) sequence data from our analyses, were deposited in Gene Expression Omnibus under the accession GSE11070.
SUPPLEMENTAL DATA

The Supplemental Data include fourteen figures, fourteen tables, Supplemental Experimental Procedures, and Supplemental References and can be found with this article online at http://www.developmentalcell.org/cgi/content/full/14/6/841/DC1/.

ACKNOWLEDGMENTS

The authors thank Angelita C. Garcia for technical assistance; Kim Emerson for assistance with preparing digital artwork; and Jennifer C. Fletcher, Eva C. Ziegelhofer, and Elliott M. Meyerowitz for agoT-38 seeds. B.D.G. is a Damon Runyon Fellow supported by the Damon Runyon Cancer Research Foundation (DRG-1909-06). R.L. is supported by a Human Frontier Science Program Long-term Fellowship. J.T.-F. is supported by Hackett and Shacklock Scholarships from the University of Western Australia. The development of the visualization and genome profiling browser AnnoJ was supported by grants from the Australian Research Council to A.H.M. (CE0561495 and DP0771156).

This work was supported by grants from the NSF (MCB-0516888) and DOE (DE-FG02-04ER15517) to J.R.E.

Received: March 7, 2008
Revised: April 1, 2008
Accepted: April 23, 2008
Published online: May 15, 2008

REFERENCES


Developmental Cell

Roles of EIN5 and ABH1 in RNA Silencing


A Link between RNA Metabolism and Silencing Affecting Arabidopsis Development
Brian D. Gregory, Ronan C. O’Malley, Ryan Lister, Mark A. Urich, Julian Tonti-Filippini, Huaming Chen, A. Harvey Millar, and Joseph R. Ecker

SUPPLEMENTAL METHODS

Microarray Analysis
15 µg of total RNA isolated using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA) from immature flower bud clusters was used for tiling microarray analysis as previously described (Chekanova et al., 2007; Zhang et al., 2006). Two biological replicates were performed for wild-type Col-0, abh1-1, ein5-6, and ein5-6 abh1-1 mutant plants. To do this, oligo(dT)-primed targets prepared from wild-type (Col-0), abh1-1, ein5-6 and ein5-6 abh1-1 plants were hybridized to Arabidopsis tiling arrays. The tiling arrays used for this analysis were the Affymetrix® GeneChip® Arabidopsis Tiling 1.0F and GeneChip® Arabidopsis Tiling 1.0R Arrays, which are housed in the Gene Expression Omnibus (GEO) under the accessions GPL1979 and GPL1980, respectively (Zhang et al., 2006). We then employed the TileMap algorithm (Ji and Wong, 2005), which utilizes a two-state hidden Markov model based on probe-level t statistics, to identify genomic regions showing statistically significant expression changes between the mutant genotypes and wild-type Col-0 plants (Fig. S2 and Tables S1-S6). The Tilemap tiling array analysis was performed as previously described (Chekanova et al., 2007), with the exception that a posterior probability value of 0.6 was used.

smRNA library construction
Total RNA was isolated using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA). Immediately following RNA precipitation, the flow through from the anion-exchange chromatography column was further precipitated in another 2.5 volumes of 100% ethanol (smRNA fraction). The smRNA fraction was further purified by a phenol-chloroform extraction and an additional ethanol precipitation. smRNAs were resolved by electrophoresis of 2.5 µg of the smRNA fraction and 7.5 µg of total RNA on 15% polyacrylamide gels containing 7 M urea in TBE buffer (45 mM Tris-borate, pH 8.0, and 1.0 mM EDTA). A gel slice containing RNAs of 15 to 35 nucleotides (based on the 10 base pair ladder size standard (Invitrogen, Carlsbad, CA)) was excised and eluted in 0.2 M NaCl rotating at room temperature for 4 hours. The eluted RNAs were precipitated using ethanol and resuspended in diethyl pyrocarbonate–treated deionized water. Gel-purified smRNA molecules were ligated sequentially to 5’ and 3’ RNA oligonucleotide adapters using T4 RNA ligase (10 units/µL) (Promega, Madison, WI). The 5’ RNA adapter (5’ - GUUCAGAGUUCUACAGUCCGACGAUC - 3’) possessed 5’ and 3’ hydroxyl groups. The 3’ RNA adapter (5’ - pUCGUAUGCCGUCUUCUGCUUGidT - 3’) possessed a 5’ mono-phosphate
and a 3' inverted deoxythymidine (idT). All oligonucleotides (including RNA and DNA) were provided by Illumina (Illumina Inc., San Diego, CA). The smRNAs were first ligated to the 5' RNA adapter. The ligation products were gel eluted and ligated to the 3' RNA adapter as described above. The final ligation products were then used as templates in a reverse transcription (RT) reaction using the RT-primer (5' - CAAGCAGAAGACGGCATACGA - 3') and Superscript II reverse transcriptase (Invitrogen Inc., Carlsbad, CA). This was followed by a limited (16 cycle) PCR amplification step using the PCR reverse (5' – AATGATACGGGCCACCACCGAGATCTACAGGTTCCGA - 3') and forward (5' - CAAGCAGAAGACGGCATACGA - 3') primers and Phusion hot-start high fidelity DNA polymerase (New England Biolabs, Cambridge, MA). The amplification products were electrophoresed on a 6% polyacrylamide gel in TBE buffer, eluted in 0.2 M NaCl rotating at room temperature for 4 hours, precipitated using ethanol, and resuspended in nuclease-free water. A schematic of this procedure is presented in Figure S3.

**GMUCT sequencing library construction**

Total RNA from wild-type Col-0 and ein5-6 immature flower buds was isolated using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA) and treated with RNase-free DNasel (Qiagen) for 25 min at room temperature. Following an ethanol precipitation, poly(A)+ RNA was column purified from 50 µg of the total RNA samples using Oligotex mRNA kit (Qiagen). The poly(A)+ RNA was then depleted for 18S and 28S rRNA molecules in two sequential Ribominus (Invitrogen, Carlsbad, CA) reactions as per manufacturer’s instructions, using 6 plant-specific biotinylated LNA oligonucleotide rRNA probes supplied by the manufacturer (Invitrogen, Carlsbad, CA). The rRNA-depleted poly(A)+ RNA was then ligated to the Illumina 5' smRNA adapter using T4 RNA ligase (10 units/µL) (Promega, Madison, WI). The 5' RNA adapter (5' - GUUCAGAGGUUCACAGUCCGACGAUC - 3') possessed 5' and 3' hydroxyl groups. The ligated rRNA-depleted poly(A)+ RNA was purified by phenol:chloroform extraction and ethanol precipitation. Oligo(dT) was then used to prime cDNA synthesis using SuperScript II reverse transcriptase (200 units/µL) (Invitrogen, Carlsbad, CA). The cDNA was then subjected to second strand synthesis using Oligo(dT) and the PCR primer specific for the Illumina 5' smRNA adapter (5' – AATGATACGGGCCACCACCGAGATCTACAGGTTCCGA - 3'). The dscDNA was ethanol precipitated, and then fragmented by sonication to 50-500 bp with a Bioruptor (Diagenode Sparta, NJ), followed by end repair and ligation of the genomic DNA (gDNA) adapters provided by Illumina (Illumina, San Diego, CA) as per manufacturer’s instructions for gDNA library construction. 100-200 ng of adapter-ligated dscDNA of 195-235 bp was isolated by agarose gel electrophoresis, and subsequently purified using a gel extraction kit (Qiagen, Valencia, CA). This was followed by a limited (16 cycle) PCR amplification step using the PCR primer specific for the Illumina 5' smRNA adapter (5' – AATGATACGGGCCACCACCGAGATCTACAGGTTCCGA - 3') and
primer 2.1 that is specific for the gDNA 3’ adapter using Phusion hot-start high fidelity DNA polymerase (New England Biolabs, Cambridge, MA). Using this primer combination for the limited amplification enriched for the desired products that had an Illumina 5’ smRNA adapted on the 5’ end and the 3’ gDNA adapter on the 3’ end. The enriched library was purified with the PCR purification kit (Qiagen, Valencia, CA) and quantity and quality examined by spectrophotometry, gel electrophoresis, and limited sequencing of cloned library molecules. A schematic of this procedure is presented in Figure S12.

SUPPLEMENTAL REFERENCES


SUPPLEMENTAL FIGURES

**Figure S1.** *ein5* mutant plant developmental defects are enhanced by *abh1* mutation. (A-D) Flowers of wild type Col-0 (A), *abh1-1* (B), *ein5-6* (C), and *ein5-6 abh1-1* (D) plants. *ein5-6 abh1-1* double mutant flowers also manifested defects not observed for either single mutant or wild-type Col-0; including extranumerary petals and bending of the gynoecium.
Figure S2. Venn diagram representation of tiling microarray analysis demonstrating the up and down-regulated expression changes in the Arabidopsis transcriptome in response to mutation of \textit{ABH1} and/or \textit{EIN5}. Some of the more interesting up-regulated targets for \textit{abh1-1} (primary MIRNA transcripts (Fig. 3)) and \textit{ein5-6 abh1-1} (primary MIRNA transcripts and miRNA target mRNAs (Fig. 3)) double mutant plants are highlighted in Figure 3.
Figure S3. Schematic of deep sequencing of the smRNA-component of the transcriptome using an Illumina Genetic Analyzer. See Supplemental Methods for details on methodology.
Figure S4. Deep sequencing of smRNAs from wild-type Col-0, 
abh1-1, ein5-6, and ein5-6 abh1-1 double mutant plants using an 
Illumina Genetic Analyzer. (A) The total number of smRNAs identified from 
Illumina smRNA libraries of wild-type Col-0, 
abh1-1, ein5-6, and ein5-6 abh1-1 double mutant plants. (B) The 
number of smRNAs with completely unique 5' ends identified from Illumina 
smRNA libraries of wild-type Col-0, 
abh1-1, ein5-6, and ein5-6 abh1-1 double mutant plants.
Figure S5. Deep sequencing of smRNA populations from wild-type Col-0, abh1-1, ein5-6, and ein5-6 abh1-1 mutant plants demonstrates that ABH1 is required for proper accumulation of numerous highly abundant 21 nt smRNAs, while concomitant loss of ABH1 and EIN5 results in the accumulation of distinct 21 nt smRNAs. (A) The percent of total reads (with wild-type Col-0 being set at 100%) corresponding to 21 nt smRNAs from wild-type Col-0, abh1-1, ein5-6, and ein5-6 abh1-1 double mutant plants. abh1-1 mutant plants accumulate significantly fewer 21 nt smRNAs compared to wild-type Col-0, whereas ein5-6 abh1-1 double mutant plants demonstrate an expansion of this class of smRNAs. (B) The top panels demonstrate the normalized abundance of 24 nt smRNAs (y axis, left-side scale) in a sliding 250 kb window in wild-type (Col-0) (blue line), abh1-1 (pink line), ein5-6 (yellow line), and ein5-6 abh1-1 (green line) double mutant plants. The bottom panels demonstrate the total lengths (base pairs) of all repeats in a sliding 250 kb window.
**Figure S6.** ABH1 is required for proper miRNA processing. SmRNA northern blot analysis of samples from indicated genotypes with DNA probes complementary to miRNA156, 159, and 169. U6 is shown as a loading control.
Figure S7. ABH1 is required for proper miRNA processing. (A) Normalized values of specific miRNAs to total read numbers (Figure S4) obtained by smRNA sequencing with an Illumina GA for the indicated genotypes. (B) Normalized values of specific heterochromatic smRNAs (24 nt smRNA sequences) for the indicated genotypes. In Figure S7A, we show data for miRNA390 to serve as control for a miRNA that was also validated as being reduced in the absence of ABH1 function by smRNA northern blotting (Figure 4B-C). Specifically, the comparison of results presented in Figure S7A to those in Figure 4B-C demonstrates that the quantitation of miRNA390 levels from the sequencing data are very similar to those obtained from northern blot analysis. Therefore, quantitation of the effects of ABH1 and/or EIN5 on the levels of other miRNAs for which northern blots are not available (miRNA 173 and 822 shown) can also be extracted from the smRNA sequencing data. Overall, the collection of smRNA sequences we present here provides a vast set of data from which the effects of ABH1 and/or EIN5 on various populations of smRNAs can be determined quantitatively.
Figure S8. The number of smRNA reads with a 5' terminus at each position is plotted for smRNAs processed from the 3' end of At3g46550 (bp 920 to 933). Bars above the axis represent sense reads; below represent antisense reads. The tick marks denote a single nucleotide position within the transcript. The two asterisks denote phased siRNAs originating from the sense (nt 933) and antisense (930) strands, while taking into account the 2 nt overhang that is likely to be a byproduct of processing of these smRNAs. Most strikingly, this analysis demonstrated that EIN5-affected smRNAs are processed in multiple 21 nt registers. This 13 nt stretch of At3g46550 is used as a model for what is also demonstrated for the processing of EIN5-affected smRNAs from the other 133 loci (Supplemental Tables S11-12). Specifically, we find that for all EIN5-affected smRNA-generating loci processing of this class of smRNAs occurs from multiple 21 nt registers for some of which we can identify phased siRNAs. Overall, this analysis suggests that there is not an ordered processing event or “common” mRNA terminus that is consistently required for biogenesis of these smRNAs.
Figure S9. SmRNA Northern blot analysis of EIN5-affected smRNAs. (A) The
graph shows the total amount of smRNA reads emanating from all positions of
the At5g65630 transcript. The asterisk above one of the graph bars marks the
smRNA that is interrogated in the blot in panel B. Interestingly, low levels of one
of these smRNAs can be detected in the wild-type Col-0 sample, which suggests
that they are processed in wild-type Arabidopsis plants. (B) RNA gel-blot
analyses of smRNA-enriched RNA isolated from unopened flower clusters of
wild-type Col-0, abh1-1, ein5-6, and ein5-6 abh1-1 double mutant plants with
DNA probes complementary to the smRNA marked by an asterisk in A. ein5-6
abh1-1 double mutant plants accumulate sufficient levels of this EIN5-affected
smRNA to be detected by our Northern blot analysis. It is likely that the levels of
this smRNA in ein5-6 mutant plants are below the limit of detection in that these
single mutant plants accumulate ~5-fold lower levels than double mutant plants.
Figure S10. EIN5-affected smRNAs become a significant proportion of the 21 nt smRNA population in ein5-6 abh1-1 double mutant plants. The graph shows the percent of total 21 nt smRNA-generating sites occupied by EIN5-affected smRNAs in ein5-6 abh1-1 double mutant plants (left bar). Additionally, the graph shows the percent of all 21 nt smRNA sequenced reads occupied by EIN5-affected smRNAs in ein5-6 abh1-1 double mutant plants (right bar). Surprisingly, this class of smRNAs makes up ~19% of the total population of 21 nt smRNAs sequenced from ein5-6 abh1-1 double mutant plants, which suggests that the expansion of EIN5-affected smRNAs could dilute the other populations of 21 nt smRNAs.
Figure S11. Transcripts from which EIN5-affected smRNAs are processed accumulate in an uncapped form. (A) Products from 5' RACE reactions (Fig. 6A) were cloned and sequenced. All six clones sequenced from the 5' RACE reactions for \textit{At2g35945} contained sequence abutting the 5' RNA adapter that was within 35 bp up or downstream of the annotated 5' end of this transcript. Additionally, all 14 clones sequenced from the 5' RACE reactions for \textit{At3g46550} contained sequence abutting the 5' RNA adapter that was within 71 bp up or downstream of the annotated 5' end of this transcript. Interestingly, the cloned sequence that mapped furthest from the annotated 5' end was homologous to sequence that started 71 bp of upstream of this site. (B) This is an example of one of the cloned 5' RACE products for \textit{At3g46550}, where the RNA adapter was directly ligated to RNA that is homologous to this transcript starting 43 bp upstream of the annotated 5' end. In fact, 3/14 cloned products consisted of this exact same adapter-ligated product.
Figure S12. Schematic of extremely deep sequencing of genome-wide mapping of uncapped and cleaved transcript (GMUCT) libraries using an Illumina Genetic Analyzer. See supplemental materials and methods for details on methodology.
Figure S13. An example of a miRNA-mediated target mRNA cleavage site (screenshots from the smRNAome database) that was determined using the GMUCT method, which provides validation that this methodology is identifying the expected substrates. W and C indicate signal from Watson and Crick strands, respectively. Multi-colored bars represent independent Illumina sequencing reads. For the sequencing data the color coding is as follows: green = T, red = A, yellow = C, and blue = G.
Figure S14. The relative frequencies of each of the four nucleotides (A, U, C, G) as the 5’ terminal nucleotide within the total population of 21 nt small RNAs from wild-type Col-0, *abh1-1*, *ein5-6*, *ein5-6 abh1-1*, and EIN5-affected smRNA generating loci specifically (all 133 loci were used for this analysis). This analysis demonstrated that there is not a significant difference in the 5’ terminal nucleotide distribution within the total 21 nt smRNA populations from wild-type Col-0, *abh1-1*, *ein5-6*, *ein5-6 abh1-1* plants. Furthermore, the 5’ terminal nucleotide distribution within the 21 nt EIN5-affected smRNA population is not significantly different from that of the total population of 21 nt smRNAs.