Multiple RNA recognition patterns during microRNA biogenesis in plants

Nicolás G. Bologna,[1,4] Arnaldo L. Schapire,[1] Jixian Zhai,[2] Uciel Chorostecki,[1] Jerome Boisbouvier,[3] Blake C. Meyers,[2] and Javier F. Palatnik[1,5]

[1]*IBR (Instituto de Biología Molecular y Celular de Rosario), CONICET and Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, 2000 Rosario, Argentina;* [2]*Department of Plant & Soil Sciences, and Delaware Biotechnology Institute, University of Delaware, Newark, Delaware 19711, USA;* [3]*Institut de Biologie Structurale Jean-Pierre Ebel CNRS-CEA-UJF, 38027 Grenoble Cedex, France*

MicroRNAs (miRNAs) derive from longer precursors with fold-back structures. While animal miRNA precursors have homogenous structures, plant precursors comprise a collection of fold-backs with variable size and shape. Here, we design an approach to systematically analyze miRNA processing intermediates and characterize the biogenesis of most of the evolutionarily conserved miRNAs present in *Arabidopsis thaliana*. We found that plant miRNAs are processed by four mechanisms, depending on the sequential direction of the processing machinery and the number of cuts required to release the miRNA. Classification of the precursors according to their processing mechanism revealed specific structural determinants for each group. We found that the complexity of the miRNA processing pathways occurs in both ancient and evolutionarily young sequences and that members of the same family can be processed in different ways. We observed that different structural determinants compete for the processing machinery and that alternative miRNAs can be generated from a single precursor. The results provide an explanation for the structural diversity of miRNA precursors in plants and new insights toward the understanding of the biogenesis of small RNAs.

[Supplemental material is available for this article.]

In multicellular organisms, microRNAs (miRNAs) are a class of small RNAs of ~21 nt that originate from endogenous loci and regulate other target RNAs by base complementarity (Voinnet 2009). MiRNAs are distinguished from other small RNAs by their unique biogenesis that involves the precise excision of the stem of a fold-back precursor (Meyers et al. 2008; Bologna et al. 2013). Although the current evidence indicates that miRNAs have emerged and specialized independently in animals and plants, their biogenesis depends on the recognition of structural cues located in the miRNA precursors (Axtell et al. 2011; Cuperus et al. 2011).

Most plant miRNAs are encoded as independent transcriptional units. They are transcribed by RNA polymerase II and then capped, spliced, and polyadenylated (Xie et al. 2005). Precursor processing occurs in the specialized D-bodies located in the plant nuclei (Fang and Spector 2007; Fujioka et al. 2007; Song et al. 2007). In *Arabidopsis thaliana* and other plants, the core component of the miRNA processing machinery is the RNase III DICER-LIKE 1 (DCL1), which produces the cuts on the fold-back precursors that release the mature miRNAs (Park et al. 2002; Reinhart et al. 2002; Kurihara and Watanabe 2004; Liu et al. 2012).

In addition to DCL1, several proteins contribute to the processing of plant precursors. DCL1 interacts with the double-strand RNA (dsRNA) binding protein HYPONASTIC LEAVES 1 (HYL1) (Han et al. 2004; Vazquez et al. 2004). SERRATE and components of the CAP binding complex, which are required for splicing, also contribute to the biogenesis of miRNAs (Lobbes et al. 2006; Yang et al. 2006; Gregory et al. 2008; Laubinger et al. 2008). Furthermore, mutations in *CPL* phosphatases (Manavella et al. 2012b), *DAWDLE* (Yu et al. 2008), *TOUGH* (Ren et al. 2012), *SICKLE* (Zhan et al. 2012), and *HASTY* (Bollman et al. 2003), also compromise the biogenesis of miRNAs in plants.

Plant miRNA precursors constitute a wide range of structures, and their fold-back lengths vary between 50 and 900 nt (Bologna et al. 2009; Cuperus et al. 2011). In contrast, animal miRNA precursors have a typical structure comprising a stem of ~three helical turns, a small terminal loop, and long single-stranded RNA (ssRNA) sequences flanking the fold-back structure (Han et al. 2006). The animal microprocessor, containing the RNase III Drosha and the dsRNA-binding protein DGCR8/Pasha, recognizes the ssRNA-dsRNA junction of the precursor and produces a first cleavage 11 bp away, defining the base of the miRNA (Han et al. 2006; Kim et al. 2009). The second cleavage is then performed in the cytoplasm by Dicer recognizing the free 3′ end of the precursor stem and cutting 22 nt (~two helical turns) from this end (Saito et al. 2005).

Part of the miRNA precursors in plants have a stem of ~15 nt below the miRNA/miRNA* followed by a large internal loop, which serves as a structural signal recognized by the processing machinery (Cuperus et al. 2010; Mateos et al. 2010; Song et al. 2010; Werner et al. 2010). However, this structural determinant is not found in all plant miRNA precursors (Mateos et al. 2010). Furthermore, the biogenesis of the evolutionarily conserved miR319 and miR159 starts with a first cut next to the loop and continues with three additional cuts in a loop-to-base direction until the miRNAs are finally released (Addo-Quaye et al. 2009; Bologna et al. 2009). Further plant precursors have been shown to release other small RNAs in addition to the miRNAs (Kurihara and Watanabe 2004; Zhang et al. 2010), although the underlying processing mechanisms are unknown.

[4]Present address: Department of Biology, Swiss Federal Institute of Technology (ETH), 8092 Zurich, Switzerland.
[5]Corresponding author
E-mail palatnik@ibr-conicet.gov.ar
Article published online before print. Article, supplemental material, and publication date are at http://www.genome.org/cgi/doi/10.1101/gr.153387.112.

Although previous work has characterized the biogenesis of a specific subset of miRNA precursors, a genome-wide view of the processing of plant miRNAs is still missing. Here, we developed and applied an approach to systematically analyze miRNA processing intermediates. We could determine the processing mechanism of most of the evolutionarily conserved miRNAs in *Arabidopsis thaliana*. We found that the biogenesis of plant miRNAs is more complex than previously thought, and we identified four processing mechanisms. Furthermore, we observed that miRNAs belonging to the same family can be processed by different mechanisms. The complexity of processing pathways found in ancient miRNAs is mirrored in young recently evolved small RNAs, in which competing RNA determinants might release alternative small RNAs. The results obtained here might explain the wide range of sizes and shapes observed in the plant miRNA precursors.

Results and Discussion

Specific parallel amplification of RNA ends (SPARE)

Previous approaches to identify processing intermediates of individual precursors in plants have been based on a modified 5′-rapid amplification of cDNA ends (RACE) (German et al. 2008; Addo-Quaye et al. 2009; Bologna et al. 2009; Mateos et al. 2010; Song et al. 2010; Werner et al. 2010). We sought to modify this approach to systematically identify processing intermediates of plant miRNAs. To do this, we combined (1) a specific reverse transcription primed by a mixture of oligos matching miRNA precursors, (2) a modified 5′ RACE-PCR coupled to high-throughput deep-sequencing, and (3) bioinformatics tools. We called this approach SPARE (Fig. 1A), as it is a modification of previously described methods (e.g., PARE) to detect poly(A)+ mRNA derived fragments by deep-sequencing (Addo-Quaye et al. 2008; German et al. 2008).

We designed specific oligos that hybridize to the 3′ arm of 169 miRNA precursors, 30 nt below the miRNA/miRNA* duplex (Supplemental Table 1). Of these sequences, 90 oligos were designed against miRNAs present in a broad range of species, at least across angiosperms (designated as "conserved miRNAs"), while the remaining 79 were prepared against evolutionarily younger miRNAs present in fewer species more related to *Arabidopsis* (designated as "young miRNAs") (Cuperus et al. 2011) (Supplemental Table1).

Briefly, after depleting rRNA, uncapped RNAs were ligated with an RNA adaptor. A reverse transcription is performed using a mixture of the precursor-specific oligos that also contain a generic

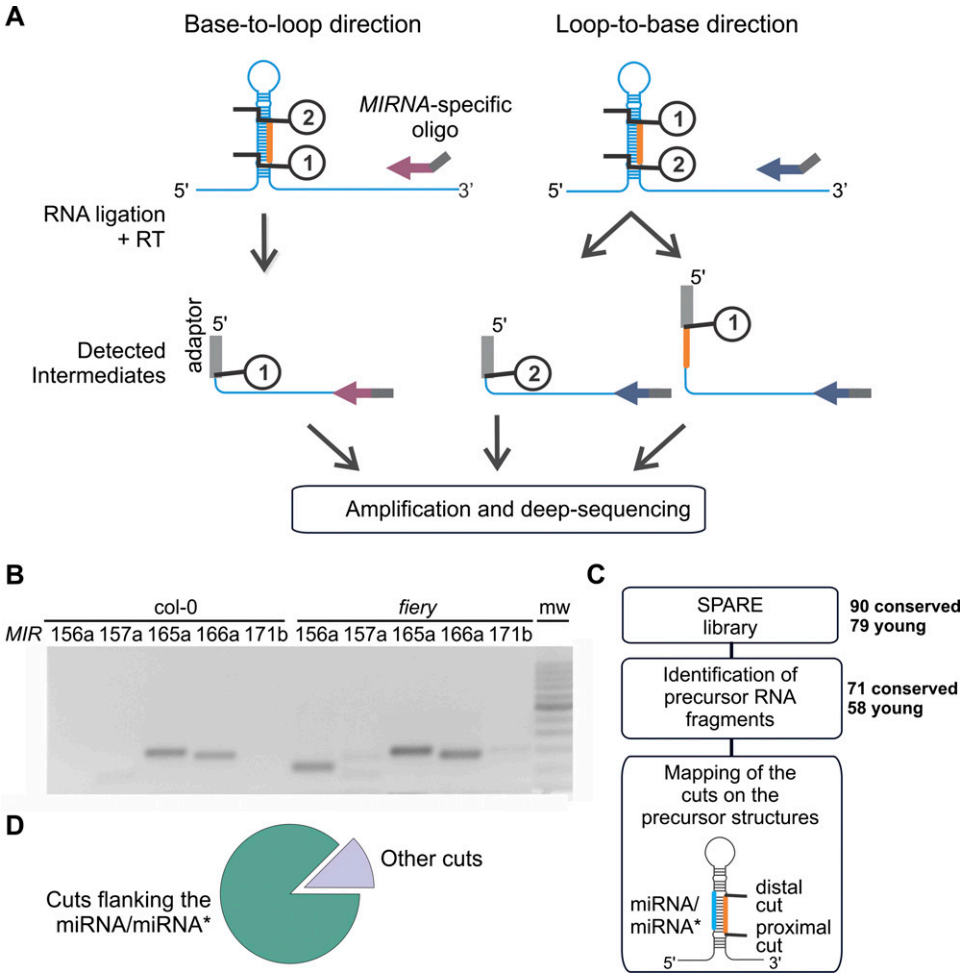

Figure 1. Construction and analysis of a miRNA precursor intermediates library (SPARE). (*A*) Scheme illustrating the SPARE library construction used to map processing intermediates and the expected results depending on the processing direction of the miRNA precursors; (1) indicates the first cleavage reaction and (2) the second one. (*B*) Modified 5′ RACE PCR results from wild-type and *fiery1* mutants on selected miRNA precursors. (*C*) Scheme of the procedure to analyze the SPARE data. (*D*) Frequency of detected cuts flanking the miRNA/miRNA* region (green) and other positions along the precursor (purple).

adaptor tail (Fig. 1A). Precursor intermediates are next amplified using oligos against the RNA adaptor and the common tail of precursor-specific oligos and subjected to deep-sequencing (Fig. 1A; see Methods for details). For our studies, we used *Arabidopsis* plants defective in *FIERY1*, which leads to reduced XRN activities, and therefore have an increased accumulation of miRNA processing intermediates (Gy et al. 2007). A pilot experiment designed against five precursors showed that the modified 5′ RACE-PCR produced the same amplification products in wild-type and *fiery1* plants, although the detection of intermediates for some miRNAs was improved in *fiery1* (e.g., *MIR156a* and *MIR171b*) (Fig. 1B). These results prompted us to use *fiery1* in our studies.

Next, we applied the SPARE method to a mixture of samples of seedlings, leaves, and inflorescences of *fiery1* plants. To analyze the results, we arbitrarily defined a miRNA precursor as the sequence between the miRNA and the miRNA*, plus 30 additional bases below the miRNA/miRNA*. We obtained reads from 18 to 40 nt in length. As contamination with small RNAs have been shown as a problem in PARE or degradome libraries (Addo-Quaye et al. 2009), we only focused on sequences longer than 25 nt. These longer sequences also allowed us to assign the reads to specific members of the same family. Using these stringent conditions, we obtained more than 30,000 reads corresponding to miRNA processing intermediates in *Arabidopsis thaliana* (Supplemental Table 2).

A precursor was considered "detected" if more than three reads corresponded to its sequence. We identified RNA fragments corresponding to 129 precursors, 71 of them of conserved miRNAs and 58 young miRNAs (Fig. 1C; Supplemental Table 1). Precursors that were not detected might correspond to miRNAs expressed under specific growth conditions or cells not represented in the tissues we analyzed, but we cannot rule out that some sequences in miRBase do not correspond to bona fide miRNAs (Axtell 2008; Axtell and Bowman 2008; Fahlgren et al. 2010).

Due to the relative position of the oligos, this method allows the detection of the first cleavage position only if the precursor is processed in a base-to-loop direction, or all the cleavage intermediates if the biogenesis proceeds in a loop-to-base pathway (Fig. 1A). We analyzed the sequence of the processing intermediates and mapped the corresponding cuts along the secondary structure of the precursors (Fig. 1C). We found that >75% of the sequence reads corresponded to the cleavages flanking the miRNA/miRNA* duplexes (proximal side and distal side, located adjacent to the lower stem and the upper stem, respectively) along the precursors of conserved miRNAs. This frequency further increased to 87% if a flexibility of one base is allowed flanking the miRNA/miRNA* ends (Fig. 1D; Supplemental Table 3), which is in agreement with a certain degree of heterogeneity observed in the sequences of the small RNAs obtained by sequencing (Nakano et al. 2006; Lu et al. 2010; Jeong and Green 2012).

Next, we applied the SPARE approach to wild-type plants (Supplemental Table 2). The results obtained were similar to those of *fiery1*, and most of the reads mapped to the flanks of the miRNA/miRNA* duplexes (Supplemental Fig. 1B–E; Supplemental Table 3). However, we detected a higher number of reads for some miRNA precursors in *fiery1*, which is consistent with the previous findings showing that the miRNA precursor intermediates accumulate in this mutant (Gy et al. 2007). Altogether, the results show that our method detects precursor intermediates that are processed to give rise to miRNAs as detected by small RNA sequencing. In principle, additional cuts that do not match the flanking positions of the miRNA/miRNA* might correspond to misprocessing reactions, precursor decay, or productive processing reactions that start farther away from the miRNA/miRNA* such as those occurring during miR319 biogenesis (Addo-Quaye et al. 2009; Bologna et al. 2009).

Identification of precursors processed in a base-to-loop direction and structural determinants characterization

The design of the SPARE library determines that only the first cut is detected in precursors cleaved in a base-to-loop processing (Fig. 1A). We then focused on those precursors with reads located only at the proximal side of the miRNA/miRNA* (Fig. 2A). We detected 32 precursors of conserved miRNAs with reads representing the proximal end of the miRNA/miRNA* without detecting cuts in the upper part of the duplex (Fig. 1A; Table 1; Supplemental Table 3).

Approximately 85% of the reads detected in these precursors matched to the exact proximal site of the miRNA/miRNA* (Fig. 2A,B; Supplemental Fig. 1B), which could be extended to ~90% if we include in this analysis the two positions at each side of the proximal site (Fig. 2B, inset). Our result is consistent with the analysis of small RNA libraries that has shown that miRNA variants shifted one or two bases are also easy to detect (Nakano et al. 2006; Lu et al. 2010; Jeong and Green 2012) and further confirms that many of these variants are likely generated by certain variability during the precursor processing. Interestingly, slight differences in the sequence of the mature miRNAs might have relevant implications in the selection of target RNAs (Palatnik et al. 2007; Jeong and Green 2012).

Genetic analyses have shown that positioning of the initial DCL1 processing event in many precursors is dependent on a lower stem of ~15 bp below the miRNA/miRNA* duplex followed by a large internal loop (Cuperus et al. 2010; Mateos et al. 2010; Song et al. 2010; Werner et al. 2010). We looked at the secondary structures of the 32 precursors analyzed in this section and found that all of them have a clear lower stem (see below) (Fig. 2C), with the exception of certain *MIR166* family members (Supplemental Fig. 2). However, the miR166/miR166* duplexes are specifically recognized by AGO10 (Zhu et al. 2011), and therefore, it might be interesting to study whether other internal loops in the *MIR166* precursors have also additional functions during the miRNA biogenesis.

Previous structural determinants were identified by averaging the secondary structure of all plant precursors, so we thought that other features might be uncovered if we analyze the precursors experimentally validated to be processed in a base-to-loop direction. To do this, the proximal side of the miRNA/miRNA* was defined as +1, and we analyzed the secondary structure of each nucleotide from positions −25 to +40 (Fig. 2C). A lower stem of 15 nt, as well as a structured region containing the miRNA/miRNA*, could be identified in both, the average of all conserved miRNAs and the experimentally validated base-to-loop precursors (Fig. 2C). However, the signal for structural determinants was clearer in the base-to-loop precursors group, which allowed us to search for additional features prevalent in these precursors. We observed that the bases immediately below the miRNA/miRNA* tend to be unpaired (positions −2 to −1), separating two double-stranded segments, one harboring the lower stem and the other the miRNA/miRNA* (Fig. 2C). Furthermore, positions −4 to −3 and the proximal three bases of the lower stem (positions −15 to −13) were paired in most cases, likely providing a clamp at the ends of the structured region of the lower stem. The transition from the lower stem to the single-stranded sequences below this region was sharper in the precursors validated as being processed by a base-to-loop mechanism than in the average of all conserved miRNAs (Fig. 2C). In turn, the structured region corresponding to

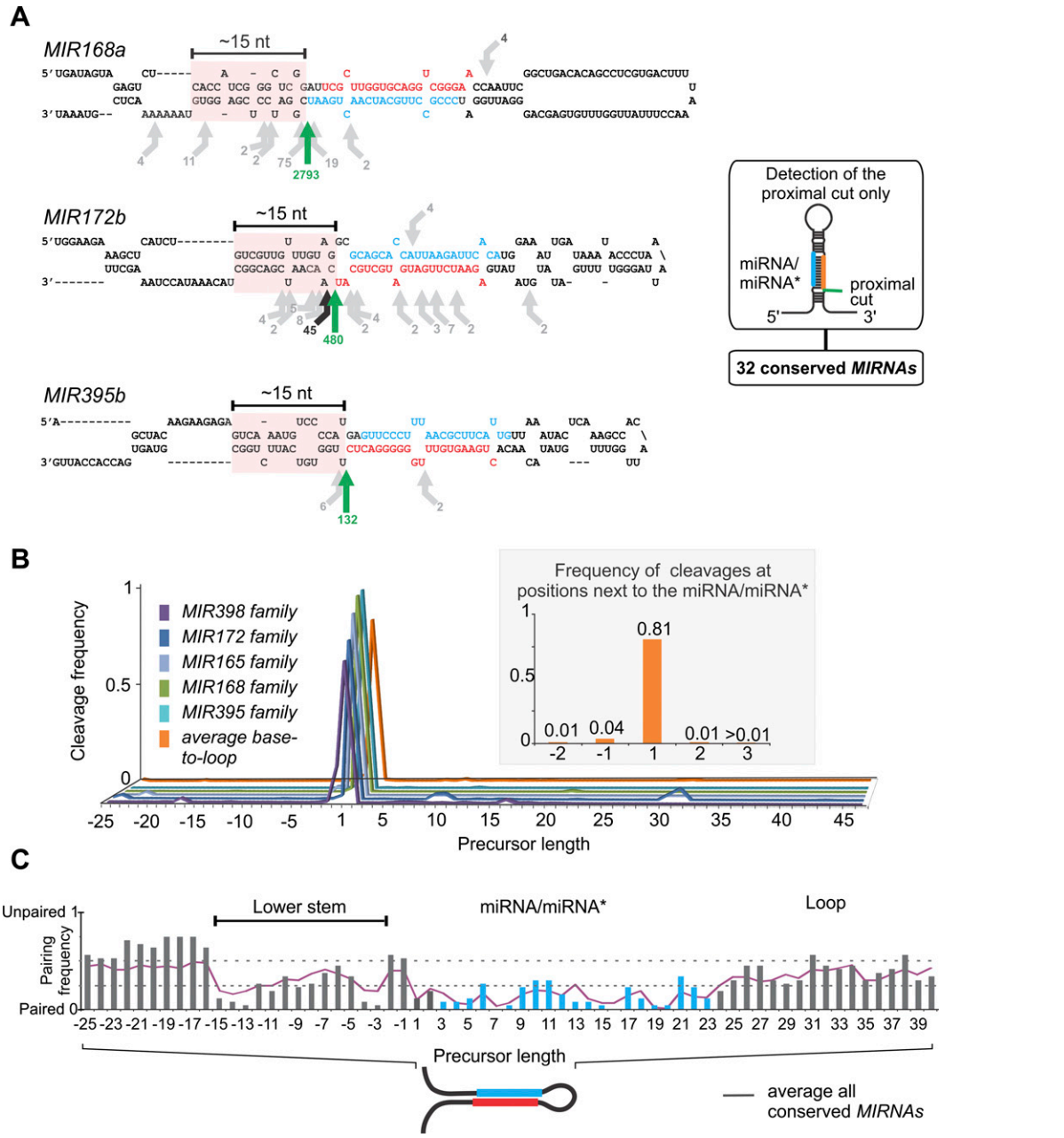

Figure 2. Identification and characterization of miRNA precursors processed in a base-to-loop direction. (*A*) Scheme showing the secondary structure of *MIR168a*, *MIR172b*, and *MIR395b*. The arrows indicate the positions and number of reads of the precursor cuts identified. Green arrows show the most abundant cleavage site detected, which also corresponds to the proximal site of the miRNA/miRNA*. Black arrows show other cleavage sites of at least 5% abundance of the total reads, while other minor cuts are shown in gray. A lower stem structured region of ~15 nt *below* the proximal cut is highlighted with a pink box. The miRNAs are indicated in red and the miRNAs* in blue. The *inset* on the *right* shows the typical cleavage pattern detected in the SPARE library for these precursors. (*B*) Distribution of cuts along the precursor sequence for specific miRNA families and all conserved miRNAs detected as being processed in a base-to-loop direction. The proximal side of the miRNA/miRNA* is defined as +1. The *inset* shows the frequency of cuts at the proximal site of the miRNA/miRNA* and the two positions next to it. (*C*) Secondary structure of the precursors detected to be processed in a base-to-loop direction. The structures were obtained from mfold (Zuker 2003; http://www.bioinfo.rpi.edu/applications/mfold/), and the matches in each position were considered as 0, while bulges and mismatches were considered as 1. The secondary structure considering all conserved miRNAs is indicated as a purple line.

the miRNA/miRNA* tends to accumulate unpaired bases at the central part (9 to 11), and certain positions (6, 17, and 21) (Fig. 2C). Overall, our results show that the precursors processed in a base-to-loop direction are more uniform than previously thought and that at least some of the precursors not detected as base-to-loop likely have other specific RNA determinants.

Sequential base-to-loop processing

The *MIR169* family has 14 members, being the largest *MIRNA* family of *Arabidopsis thaliana*. The miRNAs of this family could be classified in four groups according to their mature sequences (Fig. 3A). All these small RNA sequences are detected in vivo, with

Table 1. Processing mechanisms of *Arabidopsis MIRNAs*

MIRNA	Structured regions	Mechanism
Conserved *MIRNA*s		
MIR156a	Upper stem	Short loop-to-base
MIR156b	Upper stem	Short loop-to-base
MIR156c	Upper stem	Short loop-to-base
MIR156d	Upper stem	Short loop-to-base
MIR156h	Upper stem	Short loop-to-base
MIR159b	Upper stem	Sequential loop-to-base
MIR160a	Upper stem	Short loop-to-base
MIR160b	Upper stem	Short loop-to-base
MIR160c	Upper stem	Short loop-to-base
MIR162b	Upper and lower stem	Probable short loop-to-base
MIR164b	Lower stem	Short base-to-loop
MIR164c	Lower and upper stem	Short base-to-loop
MIR165a	Lower stem	Short base-to-loop
MIR165b	Lower stem	Short base-to-loop
MIR166a	Unclear	Unclear
MIR166b	Upper stem	Unclear
MIR166e	Unclear	Unclear
MIR166f	Upper stem	Unclear
MIR167a	Lower stem	Short base-to-loop
MIR167b	Lower stem	Short base-to-loop
MIR167d	Lower stem	Short base-to-loop
MIR168a	Lower stem	Short base-to-loop
MIR168b	Lower stem	Short base-to-loop
MIR169a	Lower stem	Short base-to-loop
MIR169b	Lower stem	Sequential base-to-loop
MIR169d	Lower stem	Sequential base-to-loop
MIR169e	Lower stem	Sequential base-to-loop
MIR169f	Lower stem	Sequential base-to-loop
MIR169g	Lower stem	Sequential base-to-loop
MIR169j	Lower and upper stem	Sequential base-to-loop
MIR169l	Lower and upper stem	Sequential base-to-loop
MIR169m	Lower and upper stem	Sequential base-to-loop
MIR169n	Lower and upper stem	Sequential base-to-loop
MIR170	Lower stem	Short base-to-loop
MIR171a	Lower stem	Short base-to-loop
MIR171b	Upper stem	Short loop-to-base
MIR171c	Upper stem	Short loop-to-base
MIR172a	Lower stem	Short base-to-loop
MIR172b	Lower stem	Short base-to-loop
MIR172d	Lower stem	Short base-to-loop
MIR172e	Lower stem	Short base-to-loop
MIR319a	Upper stem	Sequential loop-to-base
MIR319b	Upper stem	Sequential loop-to-base
MIR319c	Upper stem	Sequential loop-to-base
MIR390a	Lower stem	Short base-to-loop
MIR390b	Lower stem	Short base-to-loop
MIR391	Lower stem	Short base-to-loop
MIR393a	Lower stem	Short base-to-loop
MIR393b	Lower stem	Short base-to-loop
MIR394a	Lower and upper stem	Sequential base-to-loop
MIR394b	Lower and upper stem	Sequential base-to-loop
MIR395a	Lower stem	Short base-to-loop
MIR395b	Lower stem	Short base-to-loop
MIR395c	Lower stem	Short base-to-loop
MIR396a	Lower stem	Unclear
MIR396b	Lower stem	Short base-to-loop
MIR397a	Lower stem	Short base-to-loop
MIR398b	Lower stem	Short base-to-loop
MIR398c	Lower stem	Short base-to-loop
MIR399b	Lower stem	Short base-to-loop
MIR399c	Lower stem	Short base-to-loop
MIR408	Lower stem	Short base-to-loop
MIR827	Lower stem	Short base-to-loop
Young *MIRNA*s		
MIR158a	Lower stem	Short base-to-loop
MIR161 (miR161.1)	Lower and upper stem	Short base-to-loop
MIR161 (miR161.2)	Lower and upper stem	Short base-to-loop

(*continued*)

miR169a being the most abundant one (Fig. 3B). While we identified *MIR169a* as a miRNA precursor processed in a base-to-loop direction displaying a clear structured lower stem (Fig. 3C; Table 1), we did not detect any cuts flanking the miRNA/miRNA* in the rest of the precursors of this family (Fig. 3C; Supplemental Figs. 1E, 3A).

A more detailed inspection of the data showed that the cleavage sites in the precursors of *MIR169b/d/e/f/g* were located ~21 nt below the proximal side of the miRNA/miRNA* duplex (Fig. 3C; Supplemental Fig. 3A). This is the expected distance between two DCL1 cleavages, suggesting these precursors are processed by more than two cuts of the enzyme. The fact that we detected only one main cut indicates that the processing direction is base-to-loop (Fig. 1A; Fig. 3C). Inspection of the predicted secondary structure of these precursors revealed the existence of an internal loop followed by an ~15-nt lower stem below the detected cuts. These results suggest that the miRNA processing complex recognizes the cleavage site of the first cut in a similar way to the short precursors processed by a base-to-loop mechanism, by counting a 15-nt stem region above the internal loop. However, in contrast to the canonical base-to-loop precursors, these *MIR169* precursors suffered three dicing events, and the miRNA is released by the second and third cuts (Fig. 3C). In agreement with this possibility, small RNAs have been cloned that correspond to the fragments cleaved by the first and second cleavage sites (Fig. 3C; Zhang et al. 2010). The secondary structure of *MIR394* precursors and their cleavage pattern suggest that the *MIR394* family is also processed by a long base-to-loop mechanism (Supplemental Fig. 3B; Table 1).

Interestingly, there is a third set of *MIR169* precursors (*MIR169h-n*) that contains a lower stem but also a structured upper stem (Table 1; Fig. 3C; Supplemental Fig. 3A). These members of the *MIR169* family are detected at low frequency in small RNA databases and in our library (Fig. 3B,C; Supplemental Fig. 3A). Still, we found most of the cuts located ~21 nt below the miRNA/miRNA* (Fig. 3C; Supplemental Fig. 3A), suggesting that they are also processed by a sequential base-to-loop

Table 1. *Continued*

MIRNA	Structured regions	Mechanism
MIR163	Lower and upper stem	Sequential base-to-loop
MIR400	Lower and upper stem	Short loop-to-base
MIR403	Lower stem	Short base-to-loop
MIR447a	Lower stem	Sequential base-to-loop
MIR447b	Lower stem	Sequential base-to-loop
MIR472	Lower stem	Sequential base-to-loop
MIR771	Lower stem	Short base-to-loop
MIR779	Lower and upper stem	Sequential loop-to-base
MIR824	Lower stem	Short base-to-loop
MIR825	Lower and upper stem	Short loop-to-base
MIR862	Lower and upper stem	Sequential base-to-loop
MIR864	Lower stem	Short base-to-loop

mechanism, with the possibility that the cuts continue along the structured upper stem after the miRNAs are released. Altogether, these results show the existence of another miRNA precursor processing mechanism, which is a modification of the two-step base-to-loop pathway, and requires several cleavage reactions. Furthermore, it might be interesting to study whether *MIR169a* and *MIR169b-h*, which have a different biogenesis, also have distinct properties as small RNAs.

Identification of short precursors processed in a loop-to-base direction

The precursors described above have one major cleavage site detected in our library, which is expected for the base-to-loop processing mechanism. In principle, precursors processed in a loop-to-base direction have all their cleavage sites detected by our approach. Therefore, to find miRNAs whose biogenesis starts from the loop, we searched for precursors with cuts at both sides of the miRNA/miRNA* (Figs. 1A, 4A). We found 16 precursors of conserved miRNAs with detectable cuts (>5%) in the distal side of the miRNA/miRNA* (Table 1; Fig. 4A,B; Supplemental Fig. 1C). With the exception of two miRNAs (*MIR396a* and *MIR162b*), these precursors did not have any obvious structured region below the miRNA (Fig. 4C). Among these miRNAs, we also detected members of the *MIR319*/*MIR159* families that have long upper stems and are processed by four DCL1 cuts in a loop-to-base direction (Fig. 4B; Supplemental Fig. 1D; Addo-Quaye et al. 2009; Bologna et al. 2009), confirming that our method allows the discrimination of the different processing directions.

We were interested in finding new miRNAs processed in a loop-to-base direction. Twelve precursors were identified that harbor a short upper stem above the miRNA/miRNA*, therefore differing from the long *MIR319* and *MIR159* precursors. These newly identified precursors were processed by only two cuts, instead of the four observed in the case of *MIR319* and *MIR159*. Among these miRNAs, we found the members of the *MIR156* and *MIR160* families (Table 1; Fig. 4A,B), known to regulate SPL and ARF transcription factors, respectively (for review, see Jones-Rhoades et al. 2006; Axtell and Bowman 2008). We also observed that the cuts of members of the *MIR156* family always release miRNA/miRNA* duplexes with asymmetric bulges and containing miRNAs* of ~22 nt and miRNAs of ~20 nt (Fig. 4A), in good agreement with previous results characterizing these small RNAs and their ability to generate secondary siRNA (Manavella et al. 2012c).

The short precursors processed in a loop-to-base direction have a structured terminal region (Fig. 4C) that has a homogeneous size of ~42 nt that includes a short loop, in contrast to the same region in the precursors processed in a base-to-loop direction, which is quite variable (Fig. 4D). Previous analysis of the *MIR319* precursors revealed that they have a conserved region of the stem above the miRNA/miRNA*, which can generate other small RNAs (Axtell et al. 2007) and is necessary for the biogenesis of the miR319 (Bologna et al. 2009). We did not observe any sequence conservation in the short loop-to-base precursors outside the miRNA/miRNA*. These observations suggest that there are differences in the *MIR319*/*MIR159* precursors and those identified here, although the data show that plant miRNAs are frequently generated by a loop-to-base processing mechanism.

Mixed processing of members of the *MIR170*/*MIR171* family

In general, we found that different members of a miRNA family share their biogenesis pathway (Table 1; Figs. 2B–4B). This observation is not surprising since miRNA families are thought to expand by duplication events of an ancestral miRNA gene (Allen et al. 2004; Maher et al. 2006). However, we noticed that the processing signatures of the members of certain families could vary from each other. In addition to the *MIR169* family (Fig. 3C), we found that the *MIR170*/*MIR171* precursors also have different processing signatures. A detailed inspection of the processing intermediates for *MIR170* and *MIR171a* revealed cuts only at the lower end of the miRNA/miRNA*, indicative of a base-to-loop biogenesis pathway (Fig. 5; Supplemental Fig. 4A) and in agreement with previous data on *MIR171a* (Song et al. 2010). In contrast, we found that *MIR171b* and *MIR171c* precursors had cleavage sites at both ends of the miRNA/miRNA* duplex, which is compatible with a loop-to-base processing mechanism (Fig. 5; Supplemental Fig. 4A). That individual miR170/171 miRNAs have potentially different biogenesis mechanisms prompted us to study this family in more detail.

The *MIR170*/*MIR171* family is composed of four miRNAs in *Arabidopsis thaliana* and is known to regulate Scarecrow-like transcription factors. The family members differ in the nucleotide sequences of their mature miRNA (Fig. 5A), and therefore, their expression in plants can be validated by deep-sequencing small RNAs (Fig. 5B). We overexpressed different family members in *Arabidopsis thaliana*. Plants harboring a *35S:MIR171a* transgene had darker and fewer leaves and branches than wild type (Fig. 5C–E; Song et al. 2010). To quantify the phenotypes caused by miR171 overexpression, we counted the number of cauline leaves in at least 50 independent primary transgenic plants, as has been done previously (Fig. 5E; Song et al. 2010). Overexpression of *MIR171b* caused a similar phenotype to *MIR171a*, as expected by their similar sequences (Fig. 5A,C,E).

Next, we analyzed the structural determinants required for the processing of *MIR171a* and *MIR171b* precursors. We observed that *MIR171a* has a 15-nt lower stem, while *MIR171b* has a structured upper stem, in agreement with their processing direction starting from the base and the loop, respectively (Fig. 5C; Supplemental Fig. 4A). First, we deleted the region below the miRNA/miRNA* in both precursors (Fig. 5D). We found that the biogenesis

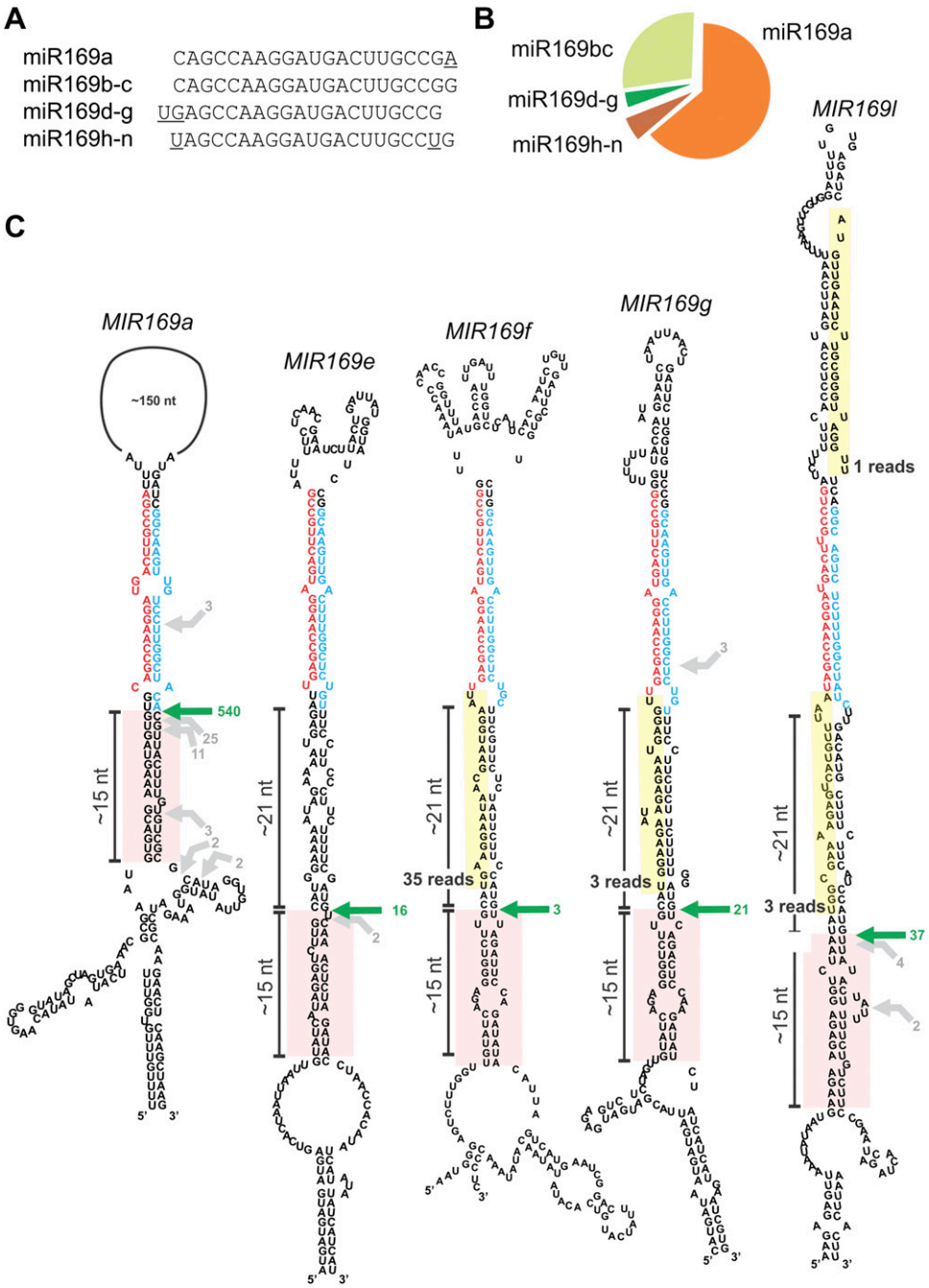

Figure 3. Single and sequential processing of *MIR169* family members. (*A*) Sequences of miR169 small RNAs. (*B*) Relative abundance of the small RNAs determined by deep-sequencing small RNAs from wild-type plants. (*C*) Scheme showing precursors of selected *MIR169* family members. The arrows indicate the positions and number of reads corresponding to the miRNA precursor cleavage sites identified. The most abundant cut is indicated by a green arrow. Gray arrows show other less abundant cleavage sites. A lower stem structured region of ~15 nt *below* the first cut is highlighted with a pink box. Additional small RNAs detected by deep-sequencing small RNAs are indicated in yellow. The miRNAs are indicated in red and the miRNAs* in blue.

of miR171a was completely abolished in this mutant precursor, and no obvious phenotype was detected among 50 independent transgenic plants (Fig. 5E). In contrast, the biogenesis of miR171b was unaffected after removing the lower region (Fig. 5D–F), and the resulting transgenic plants had a typical miR171 overexpression phenotype (Fig. 5E). Then, we eliminated most of the region above the miRNA/miRNA* in both precursors. In this case, miR171a overexpression was unaffected by the modification, while the biogenesis of miR171b was completely impaired in all the transgenic plants analyzed (Fig. 5D–F). Determination of the *MIR171* precursors by RT-qPCR showed that they were all expressed (Supplemental Fig. 4B), confirming that the processing of the mutant precursors was affected.

These results experimentally link the processing direction found in the SPARE library, the structural features in the precursors'

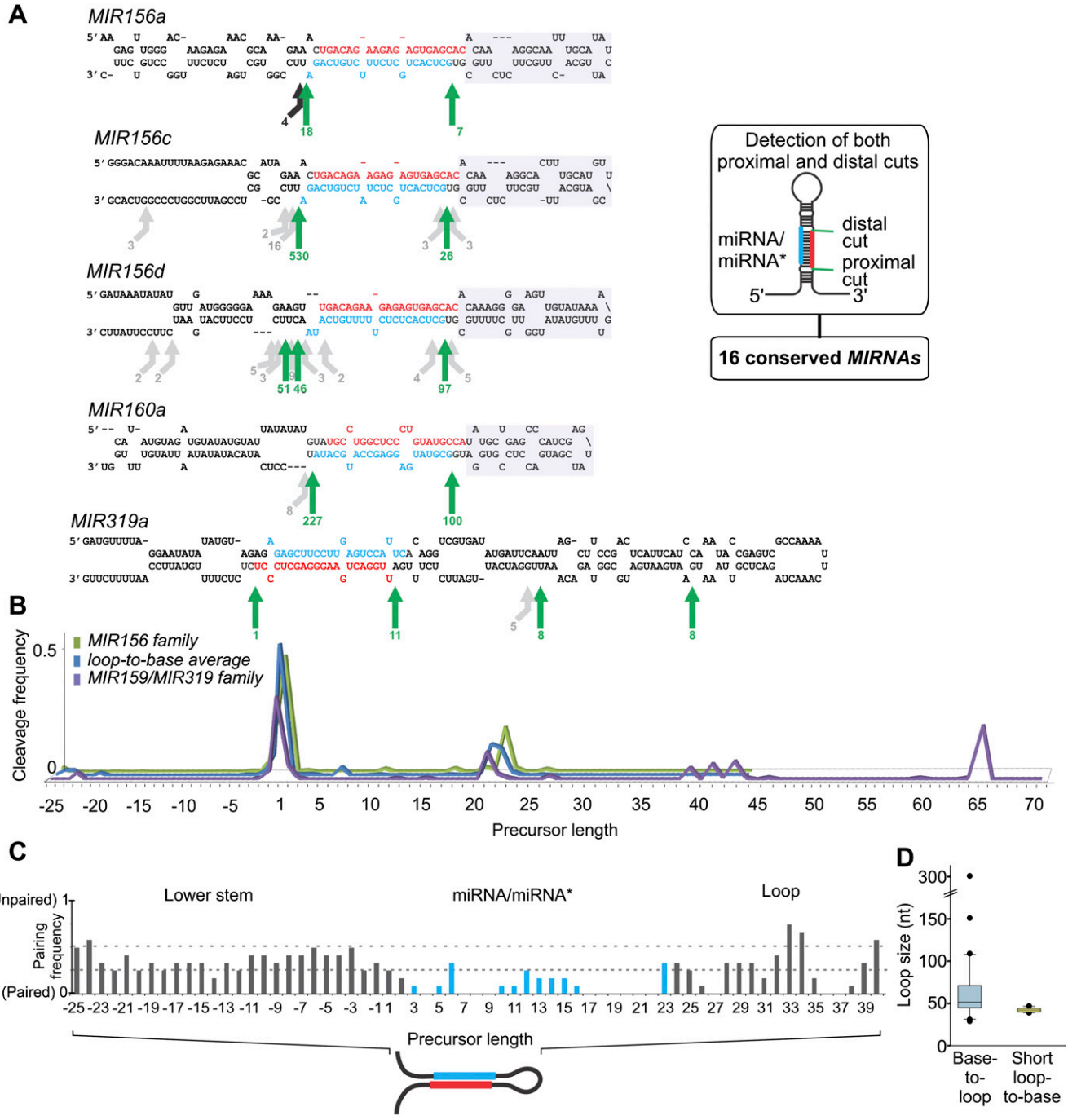

Figure 4. Loop-to-base processing of plant miRNAs. (*A*) Scheme showing the secondary structure of *MIR156a/c/d*, *MIR160a*, and *MIR319a*. The arrows indicate the positions and number of reads corresponding to the miRNA precursor cleavage sites identified. Green arrows show the most abundant cleavage site detected, which also match to the proximal and distal sides of the miRNA/miRNA* duplex of *MIR156* and *MIR160* and four DCL1 cuts of *MIR319*. Black arrows show other cleavage sites of at least 5% abundance of the total reads, while other minor cuts are shown in gray. A gray box highlights the structured upper stem of *MIR156* and *MIR160*. The miRNAs are indicated in red and the miRNAs* in blue. (*B*) Distribution of cuts along the precursor sequences of the average of all conserved miRNAs detected as loop-to-base and individual miRNA families. The proximal side of the miRNA/miRNA* was defined as +1. Note that *MIR159/319* precursors have two additional cuts. (*C*) Secondary structure of the precursors detected to be processed in a loop-to-base direction. The structures were obtained from mfold, and the matches in each position were considered as 0, while bulges and mismatches were considered as 1. (*D*) Size of the terminal region in base-to-loop (light blue) and short loop-to-base (yellow) precursors.

secondary structure, and the RNA domains required for the biogenesis of the miRNAs. While *MIR171a* is processed in a base-to-loop direction and harbors a lower stem necessary for the biogenesis of the miRNA, *MIR171b* is processed in a loop-to-base direction and contains a structured upper stem necessary for the biogenesis of the miRNA. Analysis of the predicted secondary structure of *MIR171* from other plant species such as rice (Supplemental Fig. 4C) revealed that some of them have structural features similar to

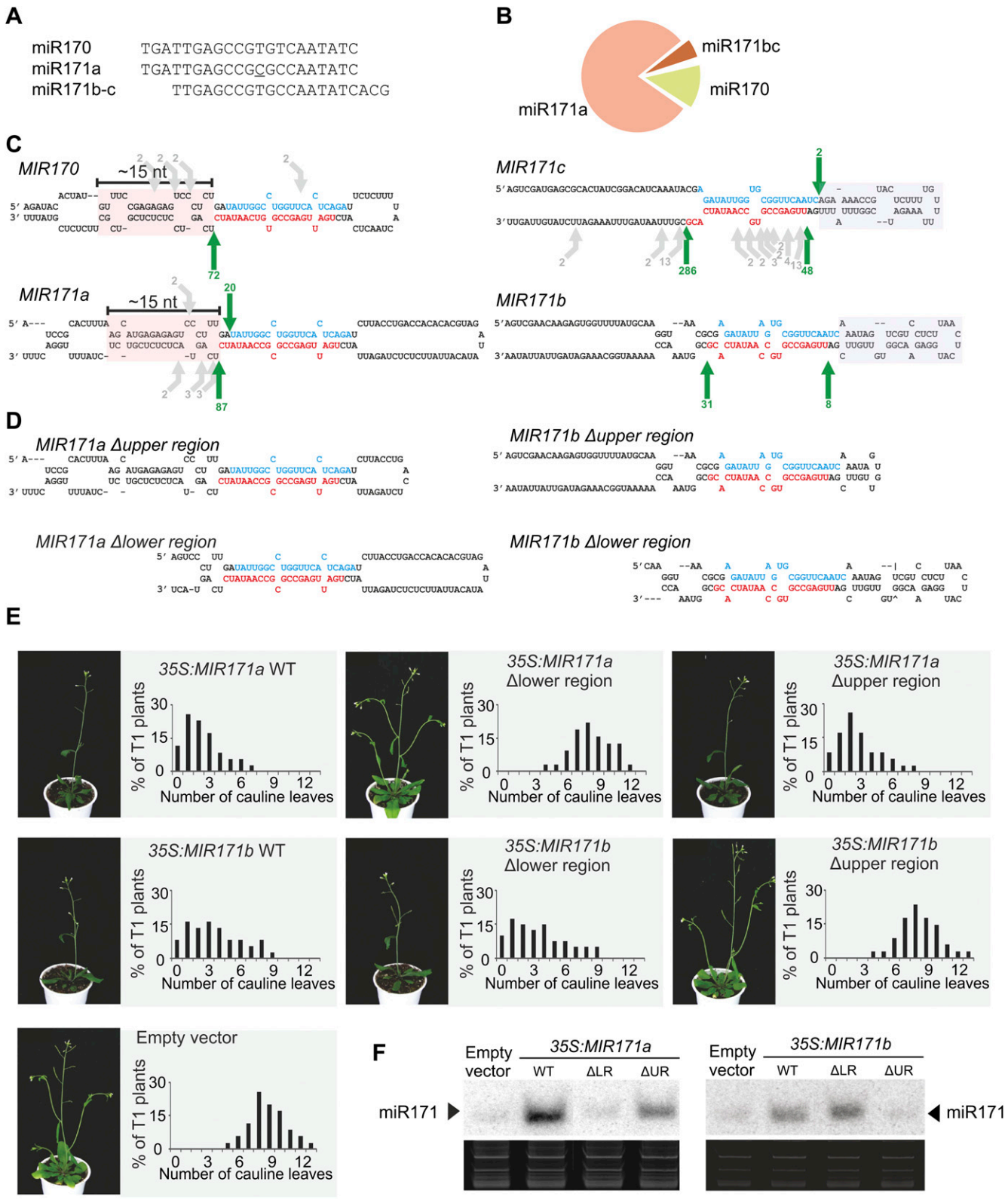

Figure 5. Mixed processing of *MIR170/171* family members. (*A*) Sequences of miR170/171 miRNAs. (*B*) Relative abundance of the small RNAs by deep-sequencing small RNAs from wild-type plants. (*C*) Scheme showing the secondary structure and cleavage sites detected in the family members. Structured regions in the lower stem (*MIR170*, *MIR171a*) and upper stem (*MIR171b/c*) are highlighted with pink and gray boxes, respectively. The most abundant cleavage sites detected, which also release the miRNA/miRNA*, are indicated by green arrows. Gray arrows show other cleavage sites that are less abundant. The miRNAs are indicated in red and the miRNAs* in blue. (*D*) Scheme of *MIR171a/b* mutant precursors. (*E*) Phenotypes of transgenic *Arabidopsis* plants transformed with wild-type and mutant *MIR171a/b* precursors under the control of the 35S promoter. A representative transgenic plant overexpressing each construct is shown on the *left*. The panels show the cauline leaves' number distribution observed in at least 50 independent transgenic plants overexpressing each construct. (*F*) Small RNA blots showing the accumulation of miR171 in transgenic *Arabidopsis* seedlings expressing the different mutant precursors. (UR) Upper region; (LR) lower region. Each sample is a pool of at least 20 independent transgenic plants.

MIR170/*MIR171a*, while others to *MIR171*b/c, suggesting that the mixed processing of the *MIR171* family members is widespread in plants. Interestingly, miR170/171a and miR171b/c are offset by three bases (Fig. 5A), and miR171a* can be incorporated into AGO1 complexes and regulate another target *SUVH8* (Manavella et al. 2012a); it is tempting to speculate that perhaps this shift in their sequences appeared as a consequence of the diversification of the processing pathways of the family.

Processing of young miRNA precursors

We next looked at the biogenesis of young miRNAs that are present in *Arabidopsis* and related species. Young miRNAs are usually expressed at low levels and, therefore, are difficult to detect (Rajagopalan et al. 2006; Fahlgren et al. 2007, 2010; Addo-Quaye et al. 2008). We found ~10,000 reads corresponding to processing intermediates of 58 different young miRNAs (Table 1; Fig. 6; Supplemental Table 3). We observed that *MIR158a* and *MIR824* precursors had detectable cuts in the proximal side of the miRNA/miRNA* and a ~15-nt lower stem, consistent with a base-to-loop processing mechanism (Fig. 6A,B). *MIR163* was also processed in a base-to-loop direction; however, the miRNA was released after the second and third cuts (Fig. 6C), in agreement with previous studies (Kurihara and Watanabe 2004). The size of the terminal region of these precursors is fairly different; *MIR158a* has a length of 22 nt, and *MIR824* and *MIR163* have 569 nt and 254 nt, respectively, confirming that precursors processed by a base-to-loop mechanism have variable loop sizes (Figs. 4D, 6A,C). Although no introns have been reported to occur in the loops of these precursors (Kurihara and Watanabe 2004; Szarzynska et al. 2009), we cannot rule out that large loops could be reduced by a splicing reaction, as has been shown to occur in rice nat-miRNAs (Lu et al. 2008).

We found that *MIR400* and *MIR779* reads were flanking the miRNA/miRNA* consistently with a loop-to-base processing direction (Fig. 6D,E). Our data also showed that *MIR779* is processed by several cuts. We think that this processing activity on the miRNA precursors is determined by its secondary structure, which in the case of *MIR779*, *MIR319*, and *MIR159* consists of an imperfect long dsRNA (Bologna et al. 2009). While *MIR400* has evolved from the inverted duplication of their targets encoding PPR proteins (Allen et al. 2004; Rajagopalan et al. 2006; Fahlgren et al. 2007), *MIR779* seems to have evolved from random sequences (Felippes et al. 2008). In both cases, however, the precursors contain a structured terminal loop of ~40 nt. Therefore, the loop-to-base processing mechanism can be responsible for the biogenesis of small RNAs with a different evolutionary origin. Altogether, our results show that the different mechanisms and RNA structural determinants recognized during the processing of ancient miRNAs can also be used during the biogenesis of miRNAs that evolved recently.

Partial cleavage and misprocessing of plant precursors

The design of the SPARE library determines that the detected DCL1 cuts correspond to the 3′ arm of the miRNA precursors (Fig. 1A). Still, a proportion of the reads are mapped to the 5′ arm of the precursors. These cuts might correspond to other pathways competing with the biogenesis of miRNAs, such as precursor decay. However, we also found cuts on the 5′ arm at the expected DCL1 cleavage sites of several miRNA precursors such as *MIR171a/b* (Fig. 5C), *MIR393b* (Fig. 7A), and *MIR166b* (Fig. 7B).

Detection of cuts on the 5′ arm at the flanks of the miRNA/miRNA* implicate the existence of intermediates containing a terminal loop with the miRNA or miRNA* as a single-stranded RNA extension (Fig. 7A,B, see scheme on the right). To confirm these results and study the existence of these partial intermediates in more detail, we turned to another approach based on self-ligation of the RNA followed by an RT-PCR, which detects both ends of the

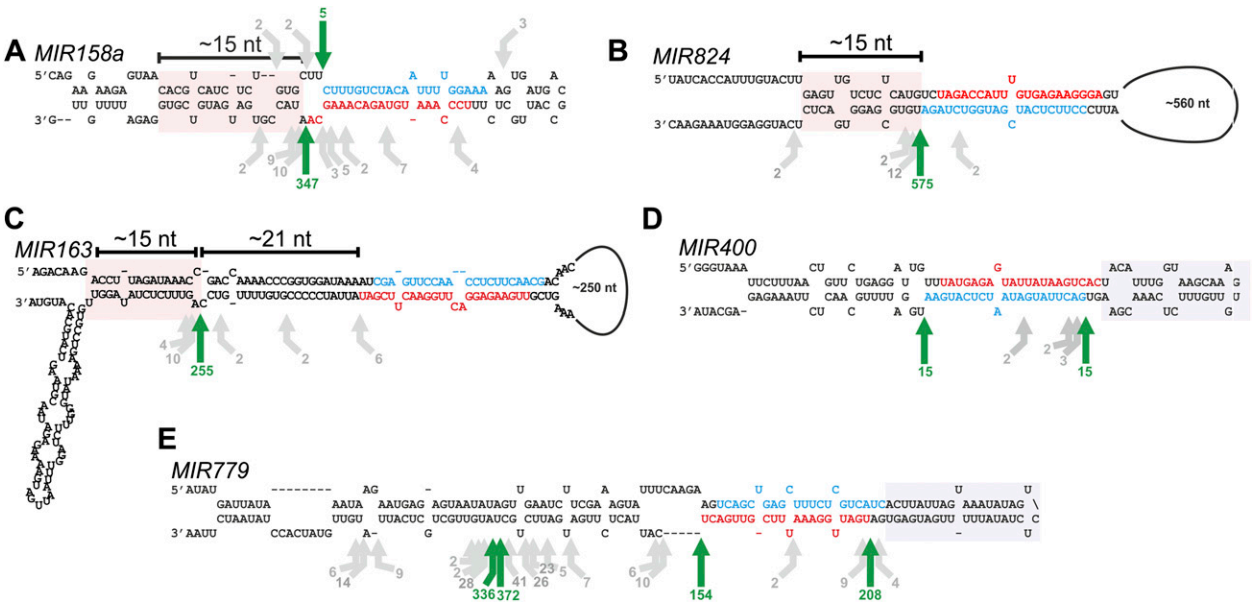

Figure 6. Processing of young miRNAs. Scheme showing the precursors of *MIR158a* (*A*), *MIR824* (*B*), *MIR163* (*C*), *MIR400* (*D*), and *MIR779* (*E*). The arrows indicate the positions and number of reads of the precursor cuts identified. Green arrows show the most abundant cleavage site detected, which also matches to the proximal and distal sides of the miRNA/miRNA*. Gray arrows show other cleavage sites that are less abundant. Structured regions in the lower stem (*MIR158a*, *MIR824*, and *MIR163*) and the upper stem (*MIR400* and *MIR779*) are highlighted with pink and gray boxes, respectively. The miRNAs are indicated in red and the miRNAs* in blue.

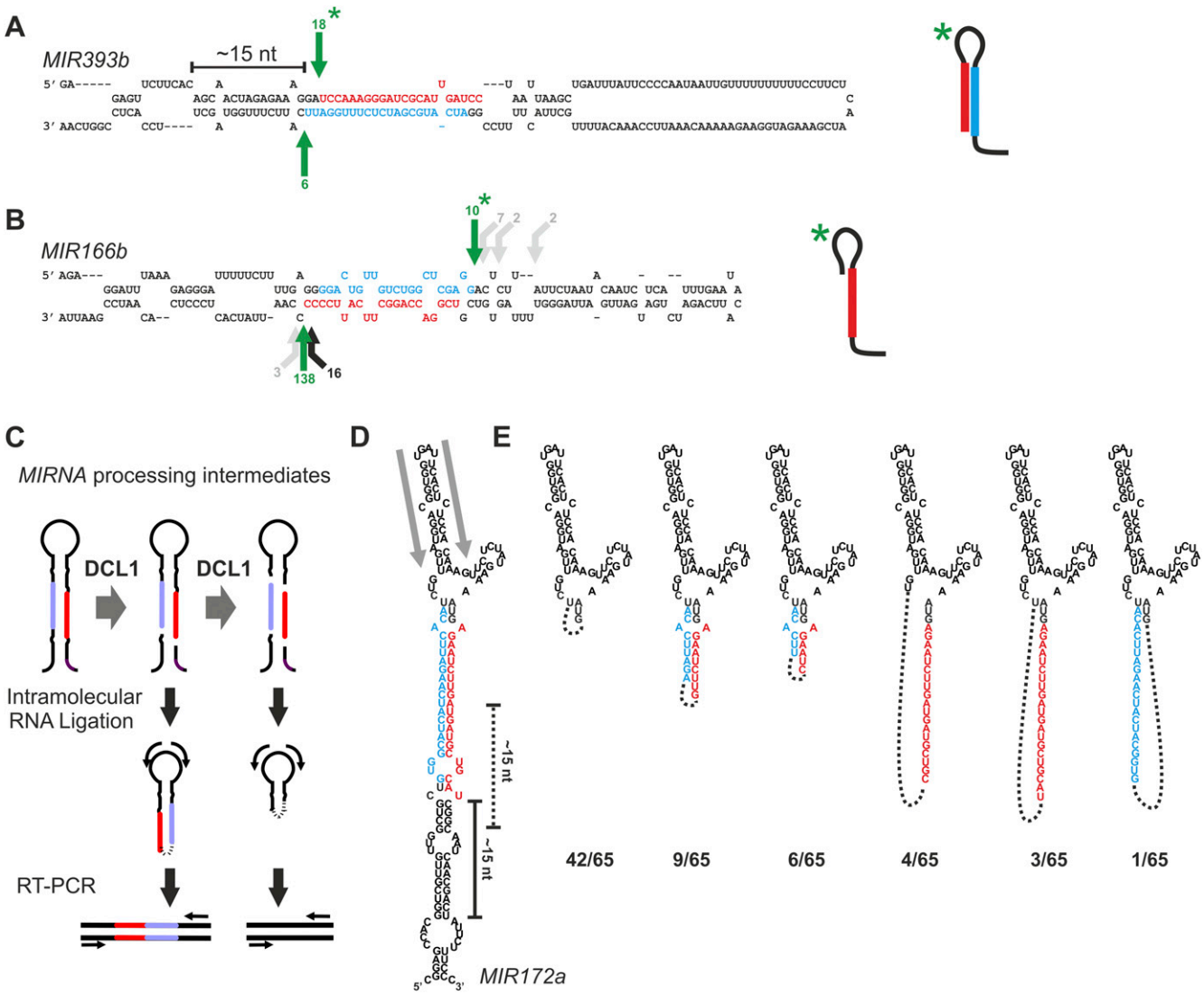

Figure 7. Partial cleavage and misprocessing of plant precursors. (*A*,*B*) Processing intermediates of *MIR393b* (*A*) and *MIR166b* (*B*) detected by the SPARE library. Cuts generating intermediates with single-strand cuts are indicated with an asterisk, and the scheme of these precursors are indicated on the *right*. (*C*) Scheme illustrating the cycle RT-PCR method used to map processing intermediates and the expected results for a precursor processed in a base-to-loop direction. Scheme showing the *MIR172a* precursor and the primers used (*D*), and the processing intermediates detected for *MIR172a* (*E*). The miRNAs are indicated in red and the miRNAs* in blue.

RNA molecule (Fig. 7C; Basyuk et al. 2003). We applied this approach to plants overexpressing a *MIR172a* precursor whose processing has been studied in detail (Mateos et al. 2010; Werner et al. 2010), using oligos located in the loop region (Fig. 7D).

Using this technique, we detected the terminal loop of *MIR172a* as the most common intermediate (Fig. 7E). Several clones were also detected in which the miRNA* has been released from the precursor, while the miRNA continued being part of it (Fig. 7E). These results are consistent with the partial DCL1 cleavage found in our library. RNase III enzymes, such as DCL1, contain one processing center with two RNA cleavage sites that generate the products with 2-nt 3′ overhangs (Zhang et al. 2004). The functional implications of this partial cleavage remain to be elucidated; however, similar intermediates produced by partial cleavage have also been detected in animal cells (Bracken et al. 2011; Gurtan et al. 2012).

We also detected cuts in the central region of miR172/miR172*. Analysis of these intermediates showed that they have 2-nt 3′ overhangs, indicating that they are likely the result of the misprocessing of a *MIR172a* precursor by RNase type III enzymes (Fig. 7E). We then looked at the secondary structure of the precursor and found that there is a 3-nt bulge ~15 nt below one of these cuts, suggesting that the cleavage might be the result of a misprocessing by DCL1 (Fig. 7D,E). Previous work has already shown that DCL1 can guide the cleavage of additional sequences that contain a bulge followed by a 15-nt stem (Song et al. 2010); however, in this case, this additional cut would be a competing reaction inactivating the miRNA.

Multiple RNA recognition patterns by the plant miRNA processing machinery

That our library often detected cuts not involved or required for the release of the small RNAs (Figs. 2–6) and that the cycle RT-PCR on *MIR172a* confirmed that some of these cuts contained the 2-nt 3′

overhangs (Fig. 7E) prompted us to study the possibility of misprocessing in vivo of miRNA precursors in more detail. *MIR825* is a young *MIRNA* that seems to be processed by a short loop-to-base mechanism (Fig. 8A). However, we also observed the abundant reads for this precursor correspond to a cut inside the miRNA/miRNA* (Fig. 8A, see orange arrow). This additional cleavage site is located above a 15-nt lower stem, suggesting the processing machinery can recognize this precursor also from a base-to-loop direction (Fig. 8A). An alternative small RNA from the other strand of the precursor can also be detected by deep-sequencing small RNAs, albeit to a lower level than miR825 (Fig. 8A). The cleavage pattern and small RNA data therefore suggests that *MIR825* is being processed via a dual mechanism, one starting from the loop and another from the base, which in turn results in the generation of the two different miRNAs. Similarly, miR161.1 and miR161.2 are also overlapping miRNAs generated from the same precursor, which likely result from the recognition of alternative structural determinants at the base of the precursor (Supplemental Fig. 5A). Furthermore, we found that *MIR472* is processed through a sequential base-to-loop mechanism, according to the reads located 21 nt below miRNA/miRNA* and the presence of a 15-nt lower stem (Supplemental Fig. 5B). However, another abundant cleavage site inside the miRNA/miRNA* region was also detected, which is likely the consequence of the recognition of a second internal loop (Supplemental Fig. 5B). Nonproductive cuts consistent with the recognition of alternative structural determinants were also found in conserved miRNA precursors, such as *MIR166f* (Supplemental Fig. 5C).

To study the recognition of alternative structure determinants alongside one precursor, we focused on the *MIR319a* precursor, which is processed by four DCL1 cuts starting below the loop (Fig. 8B; Addo-Quaye et al. 2009; Bologna et al. 2009). We generated different mutants of this precursor and expressed them in *Arabidopsis thaliana* using the 35S promoter (Fig. 8B,C). As we previously reported, deleting the terminal region of this precursor completely abolished the biogenesis of this miRNA (short *MIR319*) (Fig. 8B–D; Bologna et al. 2009). Next, we mapped the processing intermediates of this mutant precursor by 5′ RACE PCR and found that the cuts were mainly located in the miRNA/miRNA* region (Fig. 8B; Bologna et al.

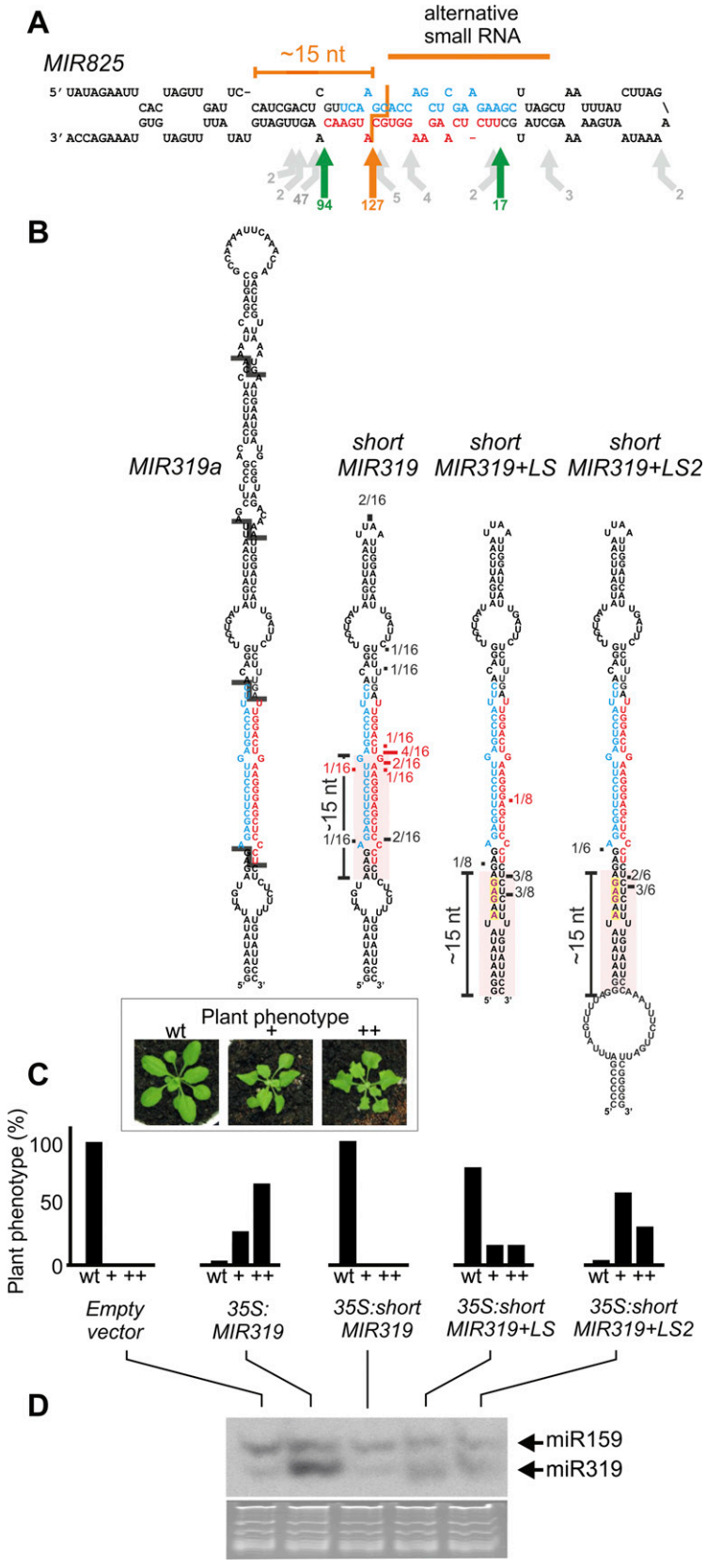

Figure 8. (Legend on next page)

2009). The positions of these cuts were above an internal bulge followed by a 15- to 16-nt stem, suggesting that the miRNA processing machinery is recognizing an alternative structural determinant in the absence of a terminal region (Fig. 8B, pink box).

To test this possibility, we generated two additional derivatives of the *MIR319a* precursor. We first generated a short *MIR319* precursor with a lower stem of ~15 nt, below the miRNA (short *MIR319*+LS) (Fig. 8B). Next, we further extended the lower stem below the large internal loop (short *MIR319*+LS2) (Fig. 8B). These two precursors partially recovered the activity of *MIR319a* in vivo (Fig. 8C) and expressed a functional miRNA (Fig. 8D). The precursor with the bulge in the lower stem (short *MIR319*+LS2) seemed to be processed with a slightly higher efficiency (Fig. 8C,D). Mapping of the processing intermediates of these precursors revealed that the cuts were mainly located below the miRNA/miRNA*, which is consistent with a switch to a base-to-loop processing mechanism. These results confirm the existence of different competing RNA domains in the same molecule that can be recognized by the miRNA processing machinery, with the stronger of them guiding the biogenesis pathway, although in some cases one precursor might generate two different small RNAs from the recognition of alternative determinants.

Conclusions

We have developed an approach to systematically analyze miRNA processing intermediates in *Arabidopsis thaliana* that could be easily applied to other experimental systems. Using this strategy we characterized four processing pathways that generate most of the miRNAs in *Arabidopsis* (summarized in Table 1). We found that these pathways present several characteristics:

1. A short base-to-loop pathway, which involves the recognition of a bulge followed by ~15 nt of a lower stem (Mateos et al. 2010; Song et al. 2010; Werner et al. 2010; this work). Although this stem segment might contain bulges, we found that the transition region from the single-strand sequences to the lower stem is rather sharp, and three paired bases usually mark the beginning of the precursor stem.
2. A long base-to-loop pathway in which the first cut proceeds in a similar way to the shorter version, but then three cuts are required until the miRNA is released. This mechanism generates additional small RNAs from the precursor, although they usually accumulate at low levels.
3. A short loop-to-base mechanism, where the processing machinery is guided by an upper stem segment, and two cuts release the mature miRNA. The terminal region of these precursors has a conserved length (~42 nt) and a small loop.
4. A long loop-to-base mechanism in which four sequential DCL1 cuts process the miRNA precursor (Addo-Quaye et al. 2009; Bologna et al. 2009). These precursors have a long conserved stem segment that can generate other small RNAs (Rajagopalan et al. 2006; Axtell et al. 2007) and is essential for the biogenesis of the miRNAs (Bologna et al. 2009).

We found that the complexity of the miRNA processing pathways occurs for both ancient and evolutionarily young sequences, and that members of the same family can be processed in different ways, suggesting that the structural determinants can change or evolve with time. The processing mechanisms are not mutually exclusive, as we found that different RNA domains can compete for the miRNA processing machinery in one molecule. In this process, alternative miRNAs can be generated from the same precursor.

Plants' miRNA precursors are highly variable in comparison with their cognate precursors in animals (Bologna et al. 2009; Cuperus et al. 2011). The results obtained here provide the explanation that diverse RNA structures can serve as miRNA precursors in plants, as they can be recognized in at least four different ways. It has been shown that animal miRNA precursors can be recognized from the loop in vitro by Drosha (Han et al. 2006). However, processing of animal miRNAs is compartmentalized in the nuclei and the cytoplasm, and after the initial cleavage by Drosha, the pre-miRNA is transported to the cytoplasm by Exportin-5 (Lund et al. 2004). It is tempting to speculate that the structural requirements for their nuclear export limit the structural diversity of the animal precursor and the alternative processing pathways of the animal miRNA precursors.

In principle, the precise RNA structure of a precursor might affect the final activity of the miRNA, as it has been shown for the asymmetric bulges located in a miRNA/miRNA* duplex (Manavella et al. 2012c). AGO complexes containing miRNAs processed from these duplexes are more prone to generate secondary siRNAs from their targets (Manavella et al. 2012c). It will be interesting to know whether the recognition of the different precursor structures by the processing machinery also confers specific features to the mature miRNAs.

Methods

Plant material and growth

All plant material was from *Arabidopsis thaliana* ecotype Col-0 and the *fiery1* mutant. Inflorescence tissue and leaf were harvested from plants grown in soil in a growth chamber with 16 h of light for 5 wk. Seedlings were grown at 23°C under the same 16 h, long-day conditions and were harvested after 2 wk. Seeds of *fiery1* (SALK_020882) were obtained from the *Arabidopsis* Biological Resource Center (ABRC). Total RNA was isolated using TRIzol reagents (Invitrogen).

SPARE library

Total RNA (1 μg) depleted of ribosomal RNA was isolated using the RiboMinus Plant Kit for RNA-seq (Invitrogen). The RNA was ligated to the RNA oligonucleotide adaptor (5′-GUUCAGAGUUCUACA GUCCGAC-3′) using T4 RNA ligase (Fermentas). The ligated products were purified and used as a template in 10 multiplex reverse transcription reactions using ~18 different precursor-specific oligos in each reaction (Supplemental Table 1).

Figure 8. Switching structural determinants in plant miRNA precursors. (*A*) Scheme of *MIR825*. Cuts releasing the miRNA/miRNA* are indicated by the green arrows. Gray arrows show other less abundant cleavage sites. A frequent cut located at 15 nt of a lower stem is indicated with an orange arrow, as well as an alternative small RNA detected at low frequency. (*B*) Scheme of *MIR319a* precursor and mutants with a deleted terminal region and the addition of a structured lower stem. Structured regions in the lower stems are highlighted with pink boxes. The cleavage sites analyzed by 5′ RACE PCR are indicated by red and black lines. The length of the arrows and the number of sequenced clones indicate the relative cloning frequency of the intermediates. (*C*) Phenotypes caused by overexpression of the wild-type and mutant *MIR319* precursors. Note that wild-type leaves are flat but become crinkled when miR319 is overexpressed. (*Inset*) Transgenic plants were classified according to the shape of their leaves. At least 50 independent transgenic plants were scored in each case. (*D*) Small RNA blots of transgenic *Arabidopsis* inflorescences expressing the different precursors. Each sample is a pool of 25 independent transgenic plants. The miRNAs are indicated in red and the miRNAs* in blue.

First-strand cDNA synthesis was carried out using SuperScript III Reverse Transcriptase (Invitrogen). Each precursor-specific oligo also contained a 5′ generic adaptor tale with the sequence 5′-AG CAGAAGACGGCATACGA-3′. Then, mixes of the 10 multiplex reactions were amplified by PCR using the generic P5 primer 5′-AA TGATACGGCGACCACCGACAGGTTCAGAGTTCTACAGTCCGA-3′ and P7 primer 5′-CAAGCAGAAGACGGCATACGA-3′. The conditions for PCR were 18 cycles of 94°C for 20 sec, 60°C for 20 sec, and 72°C for 20 sec. PCR products were gel-purified and subjected to SBS sequencing.

Transgenes

Mutated versions of the *MIR171* and *MIR319* precursor were generated by PCR, as described previously (Bologna et al. 2009). The precursors were expressed from the 35S promoter using the CHF3 binary vector (Jarvis et al. 1998). See Supplemental Table 4 for a list of binary plasmids and the sequences of the mutant precursors used in this study.

RNA expression analysis

To perform small RNA blots, 4 to 8 μg of total RNA was resolved on 17% polyacrylamide gels under denaturing conditions (7 M urea). Antisense DNA oligos to miR319 or miR171 were end-labeled as probes with [γ-32P] ATP using T4 polynucleotide kinase (Fermentas). Hybridizations were performed overnight at 38°C using Perfect Hyb buffer (Sigma). See Supplemental Figure 4 for details about the determination of *MIR171* precursors by RT-qPCR.

Cycle RT-PCR cleavage site mapping of miRNA precursors

Aliquots of 5 μg of low molecular weight RNA isolated using the mirVana miRNA Isolation Kit (Ambion) were self-ligated with T4 RNA ligase (Fermentas); the reaction volume was 100 μl, and the reaction was performed overnight at 14°C. After the ligation reaction, RNAs were precipitated with ethanol, 3M NaAC, and glycogen, and further resuspended in water. Next, first-strand cDNA synthesis was carried out using SuperScript III Reverse Transcriptase (Invitrogen) with *MIR172a* loop-specific primer 5′-TGAATC ACCACCGTCCATCAAC-3′. PCR reactions were performed using MIR172A loop-specific primers 5′-TGAATCACCACCGTCCATC AAC-3′ and 5′- CTCTCCACAAAGTTCTCTATG-3′. The PCR products were resolved on 3% agarose gels and detected by ethidium bromide staining, cloned into pBlueScript vector, and sequenced.

5′ RACE cleavage site mapping of miRNA precursors

For construction of libraries, polyadenylated RNA molecules were isolated using the PolyATtract kit (Promega). Ligation of an RNA adaptor, reverse transcription, and 5′ RACE were performed according to the procedure described by Bologna et al. (2009). Two nested vector-specific reverse oligonucleotides (RACE vector-specific oligo 5′-GTGCGCAATGAAACTGATGC-3′ and RACE vector-specific oligo nested 5′- CGAAACCGATGATACGAACG-3′) were used for 5′ RACE for plants overexpressing mutated *MIR319* precursors. The PCR products were resolved on 3% agarose gels, detected by ethidium bromide staining, cloned into pBlueScript vector, and sequenced.

Bioinformatic analysis

The secondary structure of the precursors was calculated using the program MFOLD (Zuker 2003) with default parameters to a temperature of 37° C. The proximal end of the miRNA/miRNA* duplex was defined as position +1. We analyzed the secondary structure, and we considered positions that were matched as 0, while those unpaired were considered as +1, and an average for all the precursors was made. In the precursors' schemes, some large bulges and secondary hairpins were simplified and/or indicated by a straight line. We have constructed and implemented a bioinformatic pipeline using in-house scripts and publicly available data from miRBase to assist in the analysis of the deep-sequencing libraries. We have developed a web-tool for the analysis of SPARE data using MySQL as a database. Small RNA sequences were obtained from *Arabidopsis* next-gen sequence DB (http://mpss.udel.edu/at_sRNA/index.php; Nakano et al. 2006) and miRBase database (http://www.mirbase.org/; Kozomara and Griffiths-Jones 2011).

Data access

Deep-sequencing data with the SPARE results are accessible through the NCBI Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo) under accession number GSE46429.

Acknowledgments

We thank Julieta Mateos for the plasmids with the *MIR171a* and *MIR171b* precursors, and Carla Schommer, Alexis Maizel, Christopher Brosnan, and members of the J.F.P. laboratory for comments on the manuscript. This work was supported by grants from HFSP to J.F.P. and J.B., HHMI and Agencia Nacional de Promocion Cientifica y Tecnica to J.F.P., and fellowships from CONICET to N.G.B., U.C., and A.L.S. J.F.P. is a member of the same institution. Research in the Meyers' lab was supported by the U.S. National Science Foundation award no. 0701745.

References

Addo-Quaye C, Eshoo TW, Bartel DP, Axtell MJ. 2008. Endogenous siRNA and miRNA targets identified by sequencing of the *Arabidopsis* degradome. *Curr Biol* **18:** 758–762.

Addo-Quaye C, Snyder JA, Park YB, Li YF, Sunkar R, Axtell MJ. 2009. Sliced microRNA targets and precise loop-first processing of MIR319 hairpins revealed by analysis of the *Physcomitrella patens* degradome. *RNA* **15:** 2112–2121.

Allen E, Xie Z, Gustafson AM, Sung G-H, Spatafora JW, Carrington JC. 2004. Evolution of microRNA genes by inverted duplication of target gene sequences in *Arabidopsis thaliana*. *Nat Genet* **36:** 1282–1290.

Axtell MJ. 2008. Evolution of microRNAs and their targets: Are all microRNAs biologically relevant? *Biochim Biophys Acta* **1779:** 725–734.

Axtell MJ, Bowman JL. 2008. Evolution of plant microRNAs and their targets. *Trends Plant Sci* **13:** 343–349.

Axtell MJ, Snyder JA, Bartel DP. 2007. Common functions for diverse small RNAs of land plants. *Plant Cell* **19:** 1750–1769.

Axtell MJ, Westholm JO, Lai EC. 2011. Vive la difference: Biogenesis and evolution of microRNAs in plants and animals. *Genome Biol* **12:** 221.

Basyuk E, Suavet F, Doglio A, Bordonne R, Bertrand E. 2003. Human let-7 stem-loop precursors harbor features of RNase III cleavage products. *Nucleic Acids Res* **31:** 6593–6597.

Bollman KM, Aukerman MJ, Park MY, Hunter C, Berardini TZ, Poethig RS. 2003. HASTY, the *Arabidopsis* ortholog of exportin 5/MSN5, regulates phase change and morphogenesis. *Development* **130:** 1493–1504.

Bologna NG, Mateos JL, Bresso EG, Palatnik JF. 2009. A loop-to-base processing mechanism underlies the biogenesis of plant microRNAs miR319 and miR159. *EMBO J* **28:** 3646–3656.

Bologna NG, Schapire AL, Palatnik JF. 2013. Processing of plant microRNA precursors. *Brief Funct Genomics* **12:** 37–45.

Bracken CP, Szubert JM, Mercer TR, Dinger ME, Thomson DW, Mattick JS, Michael MZ, Goodall GJ. 2011. Global analysis of the mammalian RNA degradome reveals widespread miRNA-dependent and miRNA-independent endonucleolytic cleavage. *Nucleic Acids Res* **39:** 5658–5668.

Cuperus JT, Montgomery TA, Fahlgren N, Burke RT, Townsend T, Sullivan CM, Carrington JC. 2010. Identification of *MIR390a* precursor processing-defective mutants in *Arabidopsis* by direct genome sequencing. *Proc Natl Acad Sci* **107:** 466–471.

Cuperus JT, Fahlgren N, Carrington JC. 2011. Evolution and functional diversification of *MIRNA* genes. *Plant Cell* **23:** 431–442.

Fahlgren N, Howell MD, Kasschau KD, Chapman EJ, Sullivan CM, Cumbie JS, Givan SA, Law TF, Grant SR, Dangl JL, et al. 2007. High-throughput sequencing of *Arabidopsis* microRNAs: Evidence for frequent birth and death of *MIRNA* genes. *PLoS ONE* **2:** e219.
Fahlgren N, Jogdeo S, Kasschau KD, Sullivan CM, Chapman EJ, Laubinger S, Smith LM, Dasenko M, Givan SA, Weigel D, et al. 2010. MicroRNA gene evolution in *Arabidopsis lyrata* and *Arabidopsis thaliana*. *Plant Cell* **22:** 1074–1089.
Fang Y, Spector DL. 2007. Identification of nuclear dicing bodies containing proteins for microRNA biogenesis in living *Arabidopsis* plants. *Curr Biol* **17:** 818–823.
Felippes FF, Schneeberger K, Dezulian T, Huson DH, Weigel D. 2008. Evolution of *Arabidopsis thaliana* microRNAs from random sequences. *RNA* **14:** 2455–2459.
Fujioka Y, Utsumi M, Ohba Y, Watanabe Y. 2007. Location of a possible miRNA processing site in SmD3/SmB nuclear bodies in *Arabidopsis*. *Plant Cell Physiol* **48:** 1243–1253.
German MA, Pillay M, Jeong DH, Hetawal A, Luo S, Janardhanan P, Kannan V, Rymarquis LA, Nobuta K, German R, et al. 2008. Global identification of microRNA-target RNA pairs by parallel analysis of RNA ends. *Nat Biotechnol* **26:** 941–946.
Gregory BD, O'Malley RC, Lister R, Urich MA, Tonti-Filippini J, Chen H, Millar AH, Ecker JR. 2008. A link between RNA metabolism and silencing affecting *Arabidopsis* development. *Dev Cell* **14:** 854–866.
Gurtan AM, Lu V, Bhutkar A, Sharp PA. 2012. In vivo structure-function analysis of human Dicer reveals directional processing of precursor miRNAs. *RNA* **18:** 1116–1122.
Gy I, Gasciolli V, Lauressergues D, Morel JB, Gombert J, Proux F, Proux C, Vaucheret H, Mallory AC. 2007. *Arabidopsis* FIERY1, XRN2, and XRN3 are endogenous RNA silencing suppressors. *Plant Cell* **19:** 3451–3461.
Han MH, Goud S, Song L, Fedoroff N. 2004. The *Arabidopsis* double-stranded RNA-binding protein HYL1 plays a role in microRNA-mediated gene regulation. *Proc Natl Acad Sci* **101:** 1093–1098.
Han J, Lee Y, Yeom KH, Nam JW, Heo I, Rhee JK, Sohn SY, Cho Y, Zhang BT, Kim VN. 2006. Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell* **125:** 887–901.
Jarvis P, Chen LJ, Li H, Peto CA, Fankhauser C, Chory J. 1998. An *Arabidopsis* mutant defective in the plastid general protein import apparatus. *Science* **282:** 100–103.
Jeong DH, Green PJ. 2012. Methods for validation of miRNA sequence variants and the cleavage of their targets. *Methods* **58:** 135–143.
Jones-Rhoades MW, Bartel DP, Bartel B. 2006. MicroRNAS and their regulatory roles in plants. *Annu Rev Plant Biol* **57:** 19–53.
Kim VN, Han J, Siomi MC. 2009. Biogenesis of small RNAs in animals. *Nat Rev Mol Cell Biol* **10:** 126–139.
Kozomara A, Griffiths-Jones S. 2011. miRBase: Integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res* **39:** D152–D157.
Kurihara Y, Watanabe Y. 2004. *Arabidopsis* micro-RNA biogenesis through Dicer-like 1 protein functions. *Proc Natl Acad Sci* **101:** 12753–12758.
Laubinger S, Sachsenberg T, Zeller G, Busch W, Lohmann JU, Ratsch G, Weigel D. 2008. Dual roles of the nuclear cap-binding complex and SERRATE in pre-mRNA splicing and microRNA processing in *Arabidopsis thaliana*. *Proc Natl Acad Sci* **105:** 8795–8800.
Liu C, Axtell MJ, Fedoroff NV. 2012. The helicase and RNaseIIIa domains of *Arabidopsis* Dicer-Like1 modulate catalytic parameters during microRNA biogenesis. *Plant Physiol* **159:** 748–758.
Lobbes D, Rallapalli G, Schmidt DD, Martin C, Clarke J. 2006. SERRATE: A new player on the plant microRNA scene. *EMBO Rep* **7:** 1052–1058.
Lu C, Jeong DH, Kulkarni K, Pillay M, Nobuta K, German R, Thatcher SR, Maher C, Zhang L, Ware D, et al. 2008. Genome-wide analysis for discovery of rice microRNAs reveals natural antisense microRNAs (nat-miRNAs). *Proc Natl Acad Sci* **105:** 4951–4956.
Lu T, Lu G, Fan D, Zhu C, Li W, Zhao Q, Feng Q, Zhao Y, Guo Y, Huang X, et al. 2010. Function annotation of the rice transcriptome at single-nucleotide resolution by RNA-seq. *Genome Res* **20:** 1238–1249.
Lund E, Guttinger S, Calado A, Dahlberg JE, Kutay U. 2004. Nuclear export of microRNA precursors. *Science* **303:** 95–98.
Maher C, Stein L, Ware D. 2006. Evolution of *Arabidopsis* microRNA families through duplication events. *Genome Res* **16:** 510–519.
Manavella P, Koenig D, Rubio-Somoza I, Burbano HA, Becker C, Weigel D. 2012a. Tissue-specific silencing of *Arabidopsis thaliana SU(VAR)3-9 HOMOLOG8* by miR171a*. *Plant Physiol* **161:** 805–812.
Manavella PA, Hagmann J, Ott F, Laubinger S, Franz M, Macek B, Weigel D. 2012b. Fast-forward genetics identifies plant CPL phosphatases as regulators of miRNA processing factor HYL1. *Cell* **151:** 859–870.
Manavella PA, Koenig D, Weigel D. 2012c. Plant secondary siRNA production determined by microRNA-duplex structure. *Proc Natl Acad Sci* **109:** 2461–2466.
Mateos JL, Bologna NG, Chorostecki U, Palatnik JF. 2010. Identification of microRNA processing determinants by random mutagenesis of *Arabidopsis* MIR172a precursor. *Curr Biol* **20:** 49–54.
Meyers BC, Axtell MJ, Bartel B, Bartel DP, Baulcombe D, Bowman JL, Cao X, Carrington JC, Chen X, Green PJ, et al. 2008. Criteria for annotation of plant microRNAs. *Plant Cell* **20:** 3186–3190.
Nakano M, Nobuta K, Vemaraju K, Tej SS, Skogen JW, Meyers BC. 2006. Plant MPSS databases: Signature-based transcriptional resources for analyses of mRNA and small RNA. *Nucleic Acids Res* **34:** D731–D735.
Palatnik JF, Wollmann H, Schommer C, Schwab R, Boisbouvier J, Rodriguez R, Warthmann N, Allen E, Dezulian T, Huson D, et al. 2007. Sequence and expression differences underlie functional specialization of *Arabidopsis* microRNAs miR159 and miR319. *Dev Cell* **13:** 115–125.
Park W, Li J, Song R, Messing J, Chen X. 2002. CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana*. *Curr Biol* **12:** 1484–1495.
Rajagopalan R, Vaucheret H, Trejo J, Bartel DP. 2006. A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*. *Genes Dev* **20:** 3407–3425.
Reinhart BJ, Weinstein EG, Rhoades MW, Bartel B, Bartel DP. 2002. MicroRNAs in plants. *Genes Dev* **16:** 1616–1626.
Ren G, Xie M, Dou Y, Zhang S, Zhang C, Yu B. 2012. Regulation of miRNA abundance by RNA binding protein TOUGH in *Arabidopsis*. *Proc Natl Acad Sci* **109:** 12817–12821.
Saito K, Ishizuka A, Siomi H, Siomi MC. 2005. Processing of pre-microRNAs by the Dicer-1–Loquacious complex in *Drosophila* cells. *PLoS Biol* **3:** e235.
Song L, Han MH, Lesicka J, Fedoroff N. 2007. *Arabidopsis* primary microRNA processing proteins HYL1 and DCL1 define a nuclear body distinct from the Cajal body. *Proc Natl Acad Sci* **104:** 5437–5442.
Song L, Axtell MJ, Fedoroff NV. 2010. RNA secondary structural determinants of miRNA precursor processing in *Arabidopsis*. *Curr Biol* **20:** 37–41.
Szarzynska B, Sobkowiak L, Pant BD, Balazadeh S, Scheible WR, Mueller-Roeber B, Jarmolowski A, Szweykowska-Kulinska Z. 2009. Gene structures and processing of *Arabidopsis thaliana* HYL1-dependent pri-miRNAs. *Nucleic Acids Res* **37:** 3083–3093.
Vazquez F, Gasciolli V, Cre P. 2004. The nuclear dsRNA binding protein HYL1 is required for microRNA accumulation and plant development, but not posttranscriptional transgene silencing. *Current* **14:** 346–351.
Voinnet O. 2009. Origin, biogenesis, and activity of plant microRNAs. *Cell* **136:** 669–687.
Werner S, Wollmann H, Schneeberger K, Weigel D. 2010. Structure determinants for accurate processing of miR172a in *Arabidopsis thaliana*. *Curr Biol* **20:** 42–48.
Xie Z, Allen E, Fahlgren N, Calamar A, Givan SA, Carrington JC. 2005. Expression of *Arabidopsis* MIRNA genes. *Plant Physiol* **138:** 2145–2154.
Yang L, Liu Z, Lu F, Dong A, Huang H. 2006. SERRATE is a novel nuclear regulator in primary microRNA processing in *Arabidopsis*. *Plant J* **47:** 841–850.
Yu B, Bi L, Zheng B, Ji L, Chevalier D, Agarwal M, Ramachandran V, Li W, Lagrange T, Walker JC, et al. 2008. The FHA domain proteins DAWDLE in *Arabidopsis* and SNIP1 in humans act in small RNA biogenesis. *Proc Natl Acad Sci* **105:** 10073–10078.
Zhan X, Wang B, Li H, Liu R, Kalia RK, Zhu JK, Chinnusamy V. 2012. *Arabidopsis* proline-rich protein important for development and abiotic stress tolerance is involved in microRNA biogenesis. *Proc Natl Acad Sci* **109:** 18198–18203.
Zhang H, Kolb FA, Jaskiewicz L, Westhof E, Filipowicz W. 2004. Single processing center models for human Dicer and bacterial RNase III. *Cell* **118:** 57–68.
Zhang W, Gao S, Zhou X, Xia J, Chellappan P, Zhang X, Jin H. 2010. Multiple distinct small RNAs originate from the same microRNA precursors. *Genome Biol* **11:** R81.
Zhu H, Hu F, Wang R, Zhou X, Sze SH, Liou LW, Barefoot A, Dickman M, Zhang X. 2011. *Arabidopsis* Argonaute10 specifically sequesters miR166/165 to regulate shoot apical meristem development. *Cell* **145:** 242–256.
Zuker M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **31:** 3406–3415.

Received December 10, 2012; accepted in revised form June 12, 2013.