The Plant Cell, Vol. 29: 1248–1261, June 2017, www.plantcell.org © 2017 ASPB.

LARGE-SCALE BIOLOGY ARTICLE

Evolutionary Footprints Reveal Insights into Plant MicroRNA Biogenesis

Uciel Chorostecki,[a,b,1] Belen Moro,[a,b,1] Arantxa M.L. Rojas,[a] Juan M. Debernardi,[a] Arnaldo L. Schapire,[a] Cedric Notredame,[c,d] and Javier F. Palatnik[a,e,2]

[a] Instituto de Biología Molecular y Celular de Rosario, CONICET, and Universidad Nacional de Rosario, Rosario 2000, Argentina
[b] Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Rosario 2000, Argentina
[c] Centre for Genomic Regulation, The Barcelona Institute of Science and Technology, Barcelona 08003, Spain
[d] Universitat Pompeu Fabra, Barcelona 08003, Spain
[e] Centro de Estudios Interdisciplinarios, Universidad Nacional de Rosario, Rosario 2000, Argentina

ORCID IDs: 0000-0003-2229-6853 (U.C.); 0000-0002-2810-3533 (B.M.); 0000-0003-4316-4518 (A.M.L.R.); 0000-0002-4591-7061 (J.M.D.); 0000-0001-7996-5224 (J.F.P.)

MicroRNAs (miRNAs) are endogenous small RNAs that recognize target sequences by base complementarity and play a role in the regulation of target gene expression. They are processed from longer precursor molecules that harbor a fold-back structure. Plant miRNA precursors are quite variable in size and shape, and are recognized by the processing machinery in different ways. However, ancient miRNAs and their binding sites in target genes are conserved during evolution. Here, we designed a strategy to systematically analyze *MIRNAs* from different species generating a graphical representation of the conservation of the primary sequence and secondary structure. We found that plant *MIRNAs* have evolutionary footprints that go beyond the small RNA sequence itself, yet their location along the precursor depends on the specific *MIRNA*. We show that these conserved regions correspond to structural determinants recognized during the biogenesis of plant miRNAs. Furthermore, we found that the members of the miR166 family have unusual conservation patterns and demonstrated that the recognition of these precursors in vivo differs from other known miRNAs. Our results describe a link between the evolutionary conservation of plant *MIRNAs* and the mechanisms underlying the biogenesis of these small RNAs and show that the *MIRNA* pattern of conservation can be used to infer the mode of miRNA biogenesis.

INTRODUCTION

MicroRNAs (miRNAs) are small RNAs of 20 to 22 nucleotides that originate from endogenous loci and regulate other target RNAs by base complementarity in animals and plants (Rogers and Chen, 2013; Bologna and Voinnet, 2014). They have emerged and specialized independently in both kingdoms, which likely explains differences in their biogenesis and action modes (Axtell et al., 2011; Cui et al., 2017). miRNAs are transcribed as longer precursors harboring an imperfect fold-back structure, with the small RNA embedded in one of its arms. These precursors contain spatial cues that are recognized during the biogenesis of the small RNAs (Bologna and Voinnet, 2014; Ha and Kim, 2014).

A typical animal miRNA primary transcript harbors a fold-back structure that consists of an ~35-bp stem and a terminal loop that is flanked by single-stranded RNA (ssRNA) segments (Ha and Kim, 2014). These transcripts are processed by the microprocessor, a complex that contains the RNase type III Drosha, which recognizes the transition of the ssRNA and the double-stranded region (dsRNA) of the stem loop, and produces a first cut ~11 bp away of this ssRNA-dsRNA junction (Ha and Kim, 2014). The resulting pre-miRNA is exported to the cytoplasm where Dicer performs the second cut ~22 nucleotides away from the first cleavage site, releasing a miRNA/miRNA* duplex (Ha and Kim, 2014). The miRNA is finally incorporated into an AROGNAUTE (AGO) complex, which is responsible for the activity of the small RNA, while the miRNA* is generally degraded (Axtell et al., 2011; Bologna et al., 2013a; Bologna and Voinnet, 2014; Ha and Kim, 2014).

Plant miRNA precursors are much more variable in size and shape than their animal counterparts, and they are completely processed in the nucleus by a complex harboring DICER-LIKE1 (DCL1) (Axtell et al., 2011; Rogers and Chen, 2013; Bologna and Voinnet, 2014). That plant miRNAs can be processed in different ways likely explains the lack of features common to all precursors (Bologna et al., 2013b). Rather, plant miRNA precursors can be classified into several groups. One group harbors plant miRNA precursors with an ~15- to 17-nucleotide stem below the miRNA/miRNA* (lower stem), which specifies the position of the first cut by DCL1 (Mateos et al., 2010; Song et al., 2010; Werner et al., 2010; Bologna et al., 2013b; Zhu et al., 2013). A second cut by DCL1, ~21 nucleotides away from the first cleavage site, releases the miRNA/miRNA*. These precursors are processed in a base-to-loop direction resembling the processing of animal miRNAs.

[1] These authors contributed equally to this work.
[2] Address correspondence to palatnik@ibr-conicet.gov.ar.
The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Javier F. Palatnik (palatnik@ibr-conicet.gov.ar).
www.plantcell.org/cgi/doi/10.1105/tpc.17.00272

However, another group of plant miRNA precursors are processed by a first cleavage below the terminal loop, and from there processing continues toward the base of the precursor (Addo-Quaye et al., 2009; Bologna et al., 2009, 2013b). These loop-to-base processed precursors have a structured dsRNA region above of the miRNA/miRNA* (upper stem), which is recognized by the processing machinery (Addo-Quaye et al., 2009; Bologna et al., 2009, 2013b; Kim et al., 2016). In addition, some plant miRNA precursors are processed sequentially by several cuts instead of the usual two found in animals (Kurihara and Watanabe, 2004; Addo-Quaye et al., 2009; Bologna et al., 2009, 2013b; Zhang et al., 2010).

Since the discovery of plant *MIRNAs*, it has been pointed out that conservation in distant species is only clearly seen in the miRNA/miRNA* region (Reinhart et al., 2002). The conservation of the actual miRNA sequences can be readily explained by the conserved recognition sites of cognate target genes (Reinhart et al., 2002; Allen et al., 2004; Jones-Rhoades and Bartel, 2004), a characteristic that has been exploited to predict miRNA target genes (Jones-Rhoades and Bartel, 2004; Chorostecki et al., 2012). However, the precursor of the ancient miR319 harbors a second conserved region in the precursor stem above the miRNA/miRNA* (Palatnik et al., 2003; Axtell and Bartel, 2005; Warthmann et al., 2008; Addo-Quaye et al., 2009; Bologna et al., 2009; Li et al., 2011; Sobkowiak et al., 2012), showing that additional conserved sequences exist in at least certain *MIRNAs*.

Here, we performed a global analysis of *MIRNA* sequences in different plant species. We designed a graphic representation that displays quantitative information on the conservation of the primary sequences and secondary structures. We found evolutionary footprints in plant *MIRNAs* that go beyond the miRNA/miRNA* region and reveal conservation of miRNA processing. Precursors processed in a loop-to-base or base-to-loop direction by two or more cuts all have distinct evolutionary footprints, suggesting that the miRNA processing pathway can be inferred from the conservation pattern of a *MIRNA*. As a proof of principle, we used this approach to identify new miRNA processing determinants and found that the evolutionarily conserved miR166 miRNAs require just a few bases outside the miR166/miR166* region for their biogenesis, demonstrating that their precursor recognition differs from other known miRNAs. The results describe a strong link between the evolutionary conservation of plant *MIRNAs* and the mechanisms underlying the biogenesis of the small RNAs.

RESULTS AND DISCUSSION

Identification of Plant miRNA Precursors in Different Species

MIRNAs that encode similar or identical small RNAs are usually grouped into a single family (Meyers et al., 2008). There are 29 families of miRNAs conserved at least in dicotyledonous plants (Cuperus et al., 2011; Chávez Montes et al., 2014), which are represented by 96 different precursors in *Arabidopsis thaliana* (miRBASE, release 21) (Figure 1), although the exact number might vary depending on whether miR156/157, miR165/166, miR170/171, or miR159/miR319 are considered part of a single family or superfamily (Meyers et al., 2008; Cuperus et al., 2011). Considering that miRNA precursors of the same family can be processed in different ways (Bologna et al., 2013b), we analyzed the conservation of orthologous *MIRNAs*, instead of grouping all different members of each miRNA family.

Reciprocal BLAST was used to identify putative orthologous genes to the Arabidopsis miRNAs in the genomes of 30 dicotyledonous and 6 monocotyledonous species available in the Phytozome database, version 11 (https://phytozome.jgi.doe.gov) (Figure 1). Starting with the 96 Arabidopsis *MIRNAs*, we identified 2112 putative orthologous sequences in other species (Figure 1; Supplemental Data Set 1). This large group of sequences will not cover exhaustively all miRNA precursors corresponding to the conserved miRNA families in the 36 angiosperms analyzed, but it should provide enough sequence information to allow a general analysis of their sequence conservation.

Visualization of *MIRNA* Primary Sequence and Secondary Structure of Different Species

The putative orthologous sequences of each Arabidopsis conserved *MIRNA* were used to perform a multiple alignment using T-Coffee (Chang et al., 2014), and 96 different alignments were generated (Supplemental Files 1 and 2 and Supplemental Data Set 2). We analyzed separately dicots alone (1886 precursors; Supplemental File 1; see Figure 2A for an example with *MIR172a*) or dicots together with monocots (2112 precursors; Supplemental File 2). The secondary structure of each *MIRNA* sequence was also predicted using RNAfold (Lorenz et al., 2011) (Supplemental File 3). To visualize the complex data obtained, we generated a representation of the miRNA precursors based on Circos (Krzywinski et al., 2009) (Supplemental Files 4 and 5; Figures 2B to 2G). In this representation, the inner ring shows a histogram of the frequency and distribution of paired (Figures 2B to 2G, green bars) and unpaired bases (Figures 2B to 2G, purple bars) for every position in the precursor, which therefore quantitatively indicates the conservation of secondary structures in different species. At the same time, the outer Circos data shows the nucleotide sequence of the Arabidopsis precursor maintaining the color conservation of the multiple sequence alignment consensus (Supplemental Files 4 and 5; Figures 2B to 2G). The Circos-based studies were performed in dicots (Figures 2B to 2D; Supplemental File 4) and dicots together with monocots (Figures 2E to 2G; Supplemental File 5), as we had done for the T-Coffee alignments (Supplemental Files 1 and 2).

An inspection of the *MIRNA* alignments revealed that the plant precursors have evolutionary conserved regions that go beyond the miRNA/miRNA*. However, the length and relative position of these footprints varied among the different *MIRNAs* (Supplemental File 1). We looked in more detail into the alignment of *MIR172a* (Figure 2A), whose precursor structure-function relationship has already been studied experimentally in detail (Mateos et al., 2010; Werner et al., 2010). In this case, the miR172/miR172* region was conserved as expected, but there were additional conserved regions next to miRNA/miRNA* (Figure 2A). The Circos analysis of *MIR172a* revealed conserved regions that generate a dsRNA segment of ~15 nucleotides below the miRNA/miRNA* (Figures 2B and 2E, pink line). Furthermore, sequences below this lower stem or above the miRNA/miRNA* tended to be ssRNAs in the different

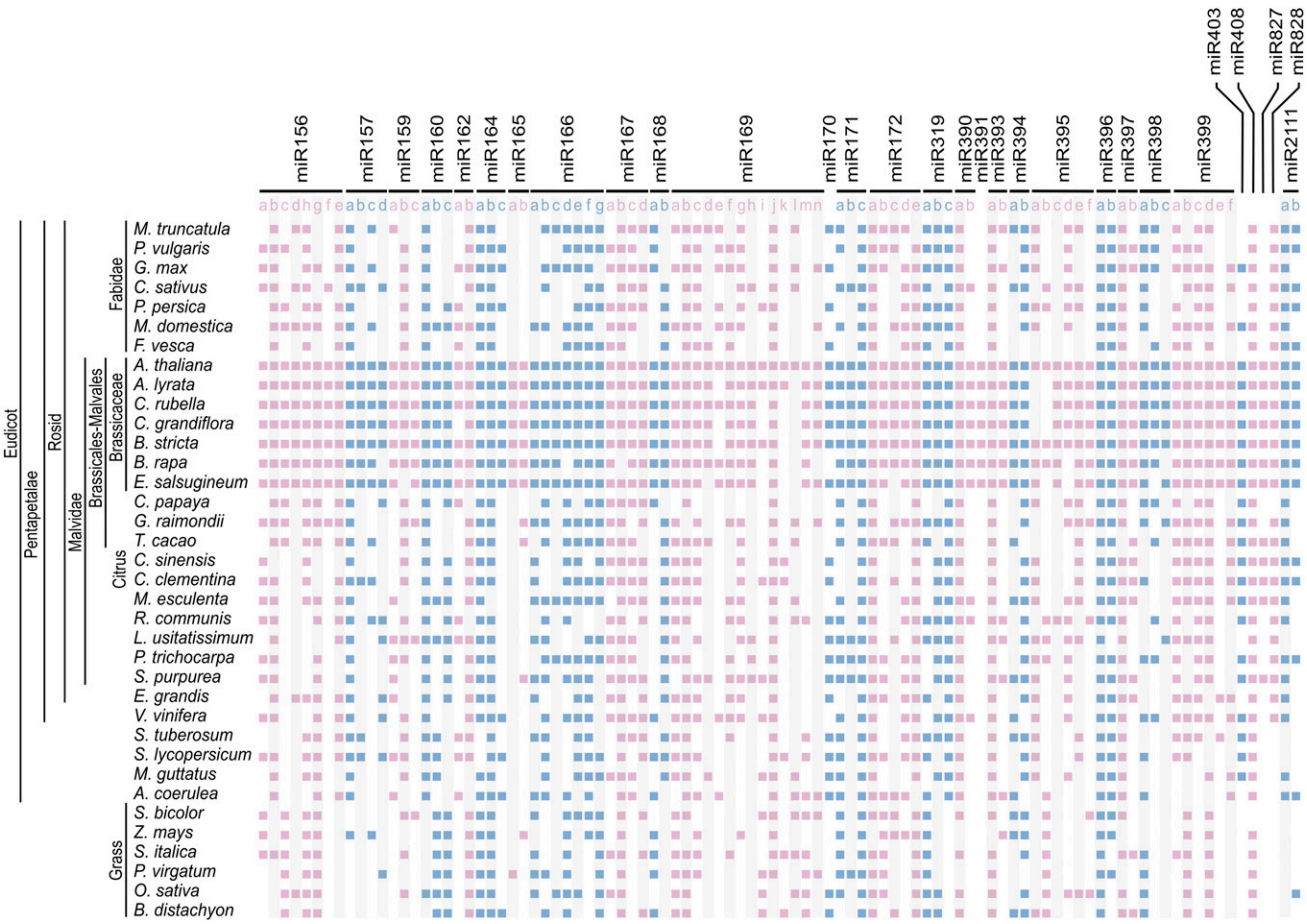

Figure 1. Identification of Arabidopsis *MIRNA* Orthologs from Angiosperms.

Representation of putative orthologs detected in dicotyledonous and monocotyledonous species for 96 *MIRNAs*.

species (Figures 2B and 2E, see purple bars). This visualization of the miR172a precursor obtained after our sequence analysis represents the model of the base-to-loop processing of plant miRNAs fairly well, which requires an ~15- to 17-nucleotide stem below the miRNA/miRNA* that is recognized by a DCL1 complex to produce the first cut (Mateos et al., 2010; Song et al., 2010; Werner et al., 2010; Bologna et al., 2013b; Zhu et al., 2013). miR393a and miR390a precursors are also known to have a dsRNA region below the miRNA/miRNA* that is important for their processing (Cuperus et al., 2010; Bologna et al., 2013b). Analysis of their sequences revealed the presence of a conserved ~15- to 17-nucleotide stem below the miRNA/miRNA* (Figures 2C, 2D, 2F, and 2G). Therefore, the analysis of the *MIRNA* sequences of different species can identify conserved regions that are coincidental with the structural determinants necessary for the precursor processing.

As shown by the Circos-based visualization of miRNA precursors (Figures 2B to 2G; Supplemental File 4) and the *MIRNA* alignments (Figure 2A; Supplemental File 1), conservation of primary sequences often co-occur with conservation of secondary structures. We also analyzed the existence of compensatory mutations in the *MIRNA* sequence alignments, identifying positions in which the precursor secondary structure is conserved, despite changes in the primary sequence (Supplemental Figure 6). These results are in good agreement with the experimental data showing that the fold-back structure of the precursor is recognized during miRNA biogenesis (Cuperus et al., 2010; Mateos et al., 2010; Song et al., 2010; Werner et al., 2010). The overall analysis generated similar results in dicots alone (Supplemental File 4) or dicots and monocots (Supplemental File 5), although for a few *MIRNAs*, such as *MIR390a* (Figures 2D and 2G), there was more divergence in monocots (Figures 2D and 2G). Therefore, we focused on the analysis of the 30 dicotyledonous species.

Precursor Sequence Conservation Correlates with Processing Direction

It has been shown that several plant miRNA precursors are processed by two DCL1 cuts in a loop-to-base direction (Bologna et al., 2013b). Furthermore, the Arabidopsis miR171a precursor is processed from the base to the loop (Song et al., 2010; Bologna et al., 2013b), while miR171b and miR171c are processed starting from the terminal loop toward the base (Bologna et al., 2013b). We analyzed the sequence conservation of *MIR171a* and *MIR171c* and found strikingly different patterns of conservation (Figures 3A

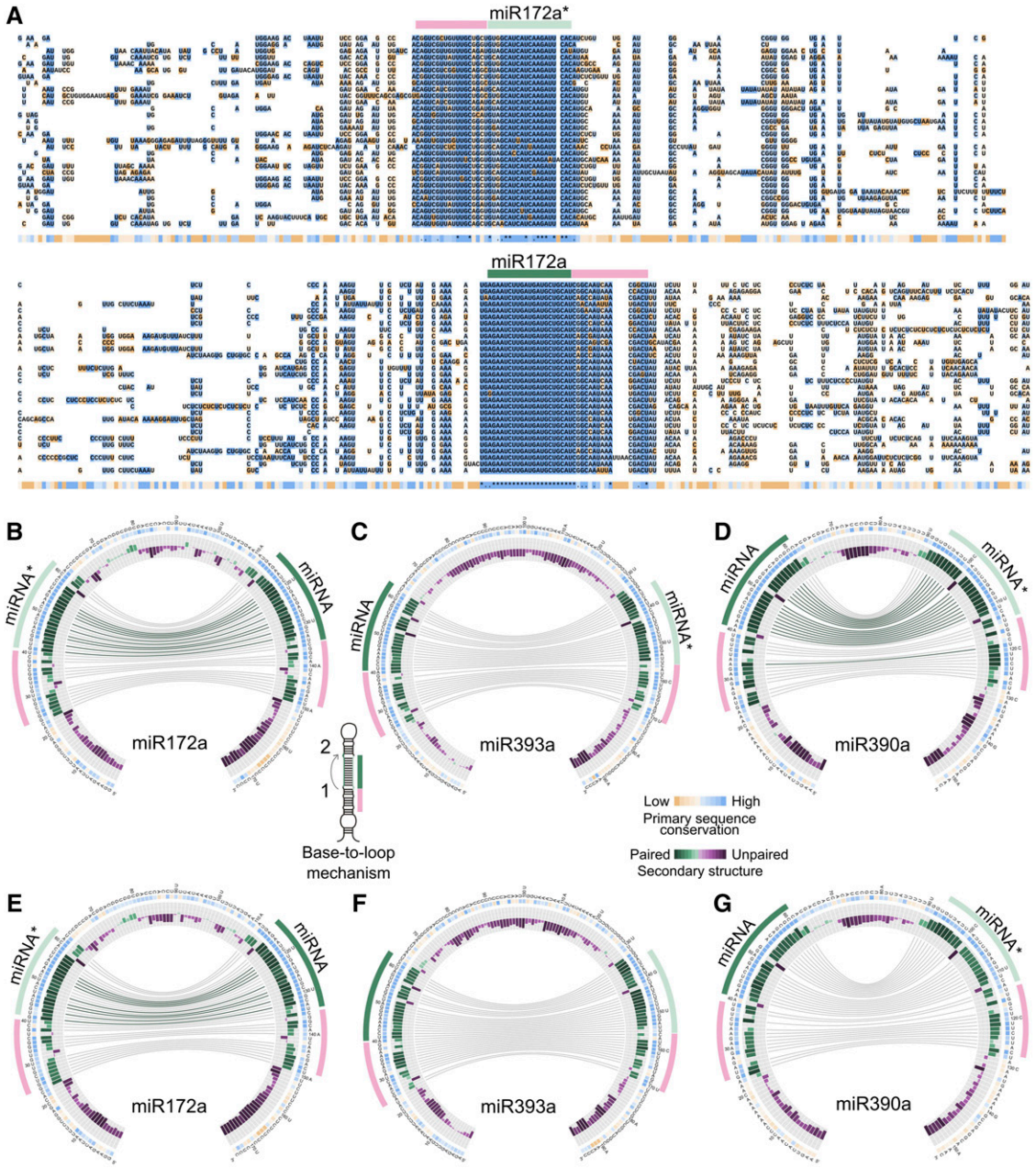

Figure 2. Circos Representation of miRNA Precursors Processed in a Base-to-Loop Direction.

(A) Alignment of miR172a precursors from *Arabidopsis lyrata* (top), *Malus domestica*, *Medicago truncatula*, *Solanum lycopersicum*, *Brassica rapa* FPsc, *Eucalyptus grandis*, *Capsella grandiflora*, *Prunus persica*, *Citrus sinensis*, *Linum usitatissimum*, *Citrus clementina*, *Glycine max*, *Vitis vinifera*, *Ricinus communis*, *Salix purpurea*, *Boechera stricta*, *Cucumis sativus*, *Aquilegia coerulea*, *Mimulus guttatus*, *Manihot esculenta*, *Eutrema salsugineum*, *Carica papaya*, *A. thaliana*, *Capsella rubella*, *Theobroma cacao*, *Populus trichocarpa*, *Phaseolus vulgaris*, *Gossypium raimondii*, *Fragaria vesca*, and *Solanum tuberosum* (bottom).

(B) to **(G)** Circos representation of miR172a (**[B]** and **[E]**), miR393a (**[C]** and **[F]**), and miR390a (**[D]** and **[G]**) precursors in dicots (**[B]** to **[D]**) and dicots and monocots (**[E]** to **[G]**). Conserved sequences are indicated with the same color code as the alignment **(A)**. Green bars indicate bases that tend to form dsRNA regions that are quantitatively indicated by the height of the bars. Connecting lines refer to bases that are interacting in the secondary structure of the precursors, green lines refer to bases that interact 100%, while gray lines show bases interacting in at least 50% of the species. Purple bars refer to bases that tend to be ssRNA regions. The miRNA is indicated with green line, while the miRNA* is light green. A conserved region that corresponds to an ~15-nucleotide lower stem is indicated with a pink line. The reference sequence is the Arabidopsis miRNA precursor. Note that sequences below the lower stem and the loop are mostly unpaired (purple bars). The inset (right) shows a scheme of a precursor processed by a base-to-loop mechanism.

and 3B). The precursor of miR171a had a conserved dsRNA region below the miRNA/miRNA* region (Figure 3A) like miR172 (Figure 2B). In contrast to the precursor of mi172a and miR171a, in the case of the miR171c, there was a conserved region above the miRNA/miRNA*, which determines a dsRNA segment (Figure 3B, pink line). That members of the miR171 family have different patterns of conservation supported our strategy that sought to compare orthologous *MIRNAs* rather than grouping different members of the same family. Other miRNA precursors processed from the loop with two cuts, such as *MIR160a*, also have a conserved dsRNA segment above the miRNA/miRNA* (Figure 3C, pink line). Furthermore, the sequence alignments of *MIR160a* (Figure 3D) or *MIR171c* (Supplemental File 1) of different species displayed conserved regions between the miRNA and miRNA*, in contrast to the alignments of *MIR172a* (Figure 2A) or *MIR171a* (Supplemental File 1) that showed additional conserved regions outside the miRNA and miRNA*.

To quantify the degree of sequence conservation, we turned to phastCons (Siepel et al., 2005). We analyzed the conservation in two contiguous regions of 15 nucleotides below the miRNA in the precursor (L1 and L2, L1 being the region adjacent to the miRNA sequence) and one region above the miRNA (U). The regions next to the miRNA* (L1*, L2*, and U*) were also analyzed. In base-to-loop precursors, the L1/L1* regions were more highly conserved than the U/U* (P value < 1.6e-15, Wilcoxon nonparametric test) and the L2/L2* (P value < 2.1e-11) (Figure 3E), as expected from the known importance of the dsRNA region immediately below the miRNA/miRNA* for the precursor processing. By contrast, in loop-to-base precursors, the U/U* regions were more highly conserved than the L1/L1* regions (P value < 2.4e-07, Wilcoxon nonparametric test) and the L2/L2* (P value < 1.2e-14) (Figure 3F). Overall, these results confirmed that precursors processed in different directions have distinct patterns of sequence conservation.

Next, we analyzed the conservation of young *MIRNAs*, which have emerged recently in evolution and are present only in Brassicaceae species. We selected young *MIRNAs* processed in a base-to-loop direction. We observed again that L1/L1* were more conserved than L2/L2* (P value < 3.4e-5, Wilcoxon nonparametric test) and the U/U* (P value < 0001) (Figure 3G). In the young *MIRNAs*, however, we did not observe a statistical difference in the conservation of the miRNA/miRNA* and the L1/L1* regions (Figure 3G). Previous analysis of the young *MIR824* in Arabidopsis ecotypes revealed selection of stable precursor sequences (de Meaux et al., 2008). Our results show the importance of the selection of both the miRNA and specific processing determinants during early events of miRNA evolution. However, during a longer period of time, it would be expected that the L region will diversify more than the miRNA/miRNA* duplex.

miRNA Biogenesis Shapes Precursor Conservation Pattern

The previous analysis focused on precursors processed by two DCL1 cuts. However, plant miRNA precursors can be processed sequentially by three or more cuts (Figures 4A to 4D, upper panels). The miR319 and miR159 precursors are sequentially processed in a loop-to-base direction by four DCL1 cuts, which generates additional small RNAs (Addo-Quaye et al., 2009; Bologna et al., 2009, 2013b; Zhang et al., 2010). The Circos analysis revealed an extended conservation of the secondary structure of these precursors, which generates a dsRNA region of ~80 nucleotides (Figure 4D). This region correlated with the region spanning the four cleavage sites (Figure 4D, green line) and a dsRNA segment above the first cut (Figure 4D, pink line).

In contrast to miR319 and miR159, the miR394 family and most miR169 family members are processed sequentially by three cuts starting at the base of the precursor (Bologna et al., 2013b). The analysis of these miRNA precursors revealed that they have a conserved dsRNA region of ~35 nucleotide below the miRNA and miRNA* (Figure 4C, green line, Supplemental File 1) that corresponds to the region spanning the first two cuts from DCL1 and an ~15-nucleotide dsRNA stem below the first cleavage site (Figure 4C, pink line). Quantitative analysis using phastCons for these *MIRNAs* revealed that both regions below the miRNA/miRNA* (L1/L1* and L2/L2*) were more highly conserved than the region above the miRNA/miRNA* (U/U*) (P value < 1.8e-05, Wilcoxon nonparametric test). By contrast, in the case of *MIR319* and *MIR159*, the two contiguous regions above the miRNA/miRNA* (U1/U1* and U2/U2*) were more conserved than the region below (L1/L1*) (P value < 6.5e-07). Overall, the results show that precursors processed by more than two cuts have correspondingly longer conserved regions than those processed only by two cuts. The extension in the conserved sequence corresponded to an ~21-nucleotide dsRNA segment for each additional cut in the processing of the precursor, which is the approximate distance between two DCL1 cuts.

The results show that there is variation in the sequence conservation of plant *MIRNAs* but that the pattern of sequence conservation can be linked to the processing mechanism of the miRNA precursors (Figures 4A to 4D). Moreover, this analysis might also be applied to other RNAs or systems. We analyzed the pattern of conservation of animal *MIRNAs* and observed that they have an extended conservation below and above the miRNA/miRNA* (Figure 4E). We think that this conservation might also be linked to the biogenesis of animal miRNAs. While a lower stem below the miRNA/miRNA* is necessary for the first cut by DROSHA (Han et al., 2006), the region above the miRNA/miRNA* might be important for the export from the nucleus to the cytoplasm (Yi et al., 2003; Lund et al., 2004; Zeng and Cullen, 2004).

Divergence of Precursor Length Correlates with Processing Direction

Next, we analyzed the distance between the miRNA and the miRNA* in the precursors of different species. Analysis of the miR160a precursor, which is processed from the loop to the base revealed that the distance between miR160 and miR160* remained fairly constant in different species with ~37 nucleotides and ranging from 36 to 40 nucleotides (Figure 5A, right panel). By contrast, the miR172a precursor, which is processed from the base to the loop, displayed a larger variability, and the region between miR172a and miR172a* varied from 38 to 114 nucleotides (Figure 5A, left panel). A general analysis showed strikingly different patterns of conservation regarding the distance between the miRNA and the miRNA* in different species (Figure 5B). Precursors processed in a base-to-loop direction like miR172a (Cuperus et al., 2010; Mateos et al., 2010; Song et al., 2010;

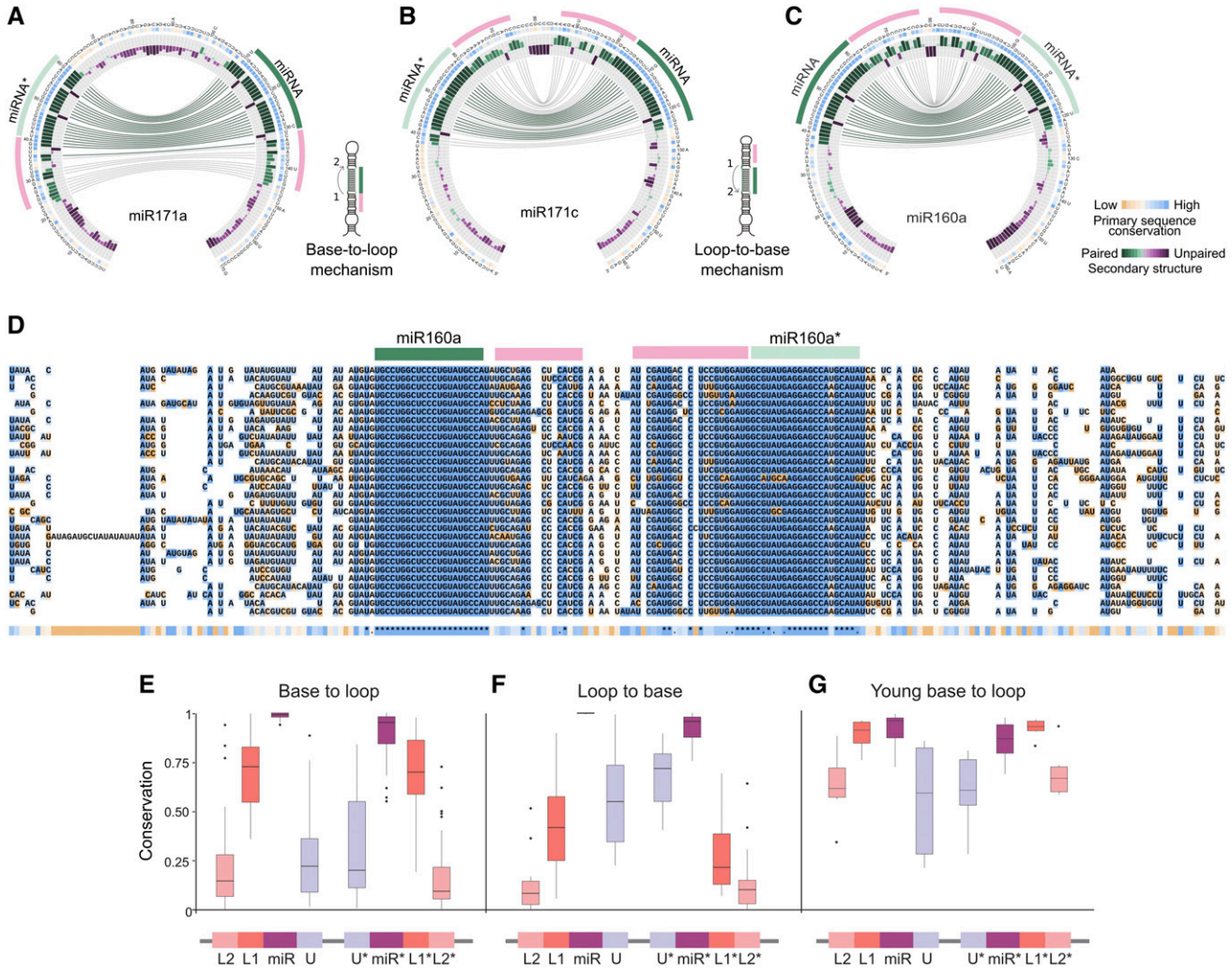

Figure 3. Conservation and Divergence of Precursors Processed in Different Directions.

(A) to **(C)** Circos representation of miR171a **(A)**, miR171c **(B)**, and miR160a **(C)**. miR171a is processed from the base, while miR171c and miR160a are processed from the loop. Note the different position of the additional conserved regions (pink line) in the precursors according their processing direction. The insets show schemes of precursors processed in base-to-loop or a loop-to-base direction.

(D) Alignment of miR171c precursors from *A. lyrata* (top), *M. domestica*, *M. truncatula*, *S. lycopersicum*, *B. rapa* FPsc, *E. grandis*, *C. grandiflora*, *P. persica*, *C. sinensis*, *L. usitatissimum*, *C. clementina*, *G. max*, *V. vinifera*, *R. communis*, *S. purpurea*, *B. stricta*, *C. sativus*, *A. coerulea*, *M. guttatus*, *M. esculenta*, *E. salsugineum*, *C. papaya*, *A. thaliana*, *C. rubella*, *T. cacao*, *P. trichocarpa*, *P. vulgaris*, *G. raimondii*, *F. vesca*, and *S. tuberosum* (bottom).

(E) and **(F)** Box plot showing the conservation of different precursor regions using phastCons for precursors processed in base-to-loop **(E)** or loop-to-base **(F)** direction.

(G) Analysis using young *MIRNAs* processed in base-to-loop direction. The band inside the box represents the median, the bottom and top of the box are the first (Q1) and third (Q3) quartiles, dots are outliers, upper whisker denotes min (max(x), Q3 + 1.5 * (Q3 − Q1)), and lower whisker denotes max(min(x), Q1 − 1.5 * (Q3 − Q1)).

Werner et al., 2010; Bologna et al., 2013b; Zhu et al., 2013) have variable distances between the miRNA and the miRNA* (Figure 5B). By contrast, miRNA precursors experimentally validated to be processed in a loop-to-base direction (Addo-Quaye et al., 2009; Bologna et al., 2009, 2013b) displayed uniform miRNA-miRNA* distances (Figure 5B, yellow and orange boxes). It has been shown that the terminal region of the precursors processed by two cuts can be largely modified without impairing the miRNA biogenesis (Mateos et al., 2010; Song et al., 2010; Werner et al., 2010), while deletions in the terminal region of precursors processed from the loop significantly affect their processing (Bologna et al., 2009, 2013a; Kim et al., 2016). We also extended the terminal region of the miR171b and miR319a precursors and analyzed their importance in vivo. We found that both precursors were not processed after extending their terminal region (Figures 5C and 5D; Supplemental Figure 1). Overall, the data show that there is an agreement between the conservation of the precursor length and its importance during miRNA processing. We also noted that

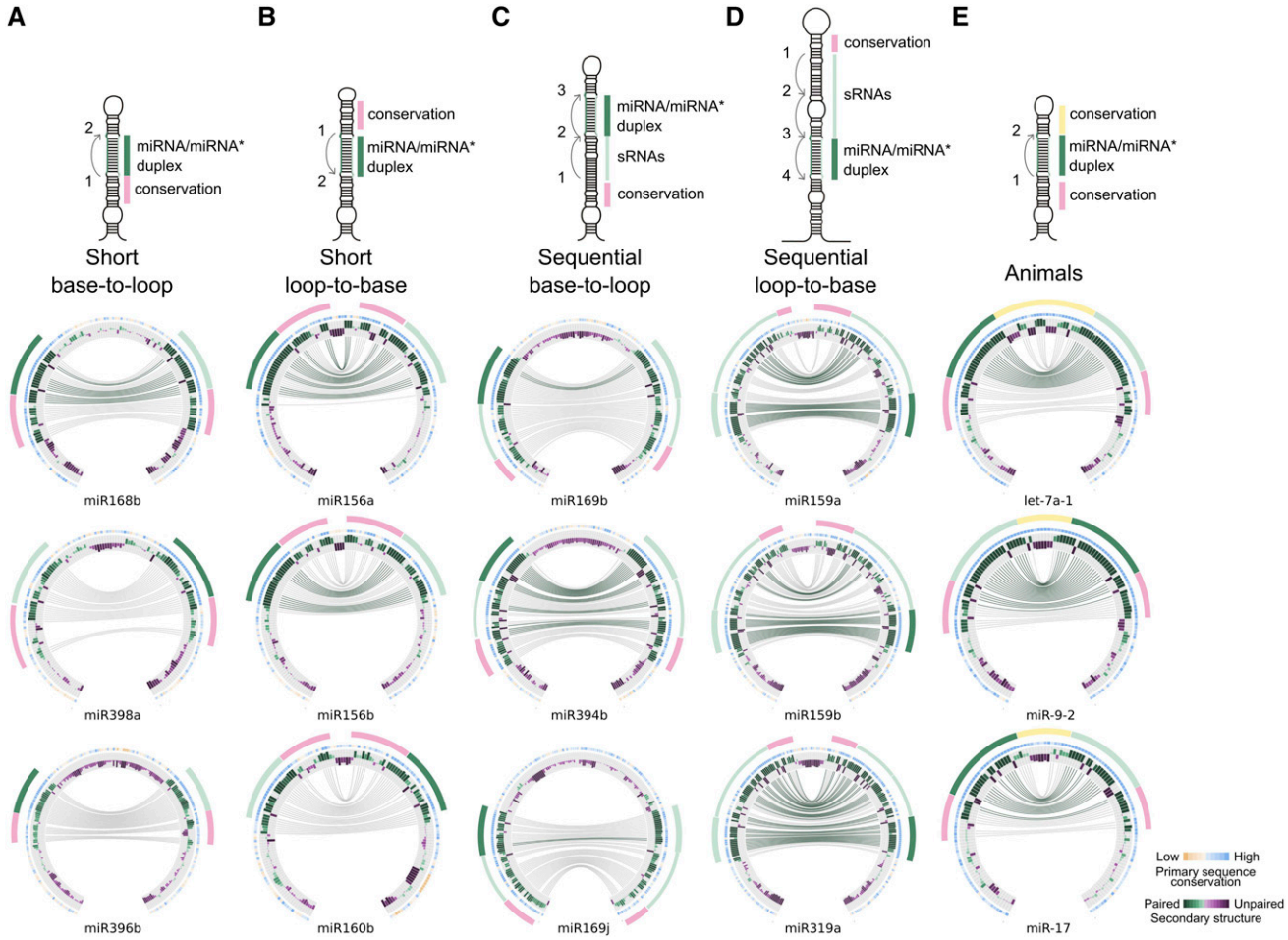

Figure 4. miRNA Biogenesis Pathways Tune *MIRNA* Conservation during Evolution.

(A) to **(D)** Circos representation of precursors processing through different directions: base to loop **(A)**, loop to base **(B)**, and sequentially processed precursors from the base **(C)** or the loop **(D)**. The miRNA is indicated by a green line, while the miRNA* is light green. Other small RNAs are indicated by a narrow green line. A conserved region that corresponds to a lower or upper stem is indicated by a pink line. The conservation of the distal region in animals is indicated in yellow. Top: Schemes representing the different processing pathways.

(E) Circos representation of animal MIRNAs (analyzed species include *Bos taurus*, *Canis familiaris*, *Equus caballus*, *Gallus gallus*, *Gorilla gorilla*, *Homo sapiens*, *Macaca mulatta*, *Monodelphis domestica*, *Mus musculus*, *Ornithorhynchus anatinus*, *Petromyzon marinus*, *Sus scrofa*, and *Xenopus tropicalis*).

miR319a was more tolerant to modifications in the precursor length than miR171b (Figures 5C and 5D) and that miR319 precursors were slightly more variable in their length in different species compared with miR171b and miR160 (Figure 5B; Supplemental Figure 1).

MIR166 MIRNAs Have Specific Patterns of Sequence Conservation

The observation that the conservation of sequence and RNA structure of plant *MIRNAs* correlates with their processing mechanism (Figures 2 to 4) prompted us to explore whether the conservation pattern of certain miRNAs deviates from the pattern expected based on known processing pathways. We generated a global view of the sequence conservation for all experimentally validated precursors processed from the loop or the base by comparing the relative conservation above and below each miRNA/miRNA* (Figure 6A). As expected, precursors processed in a base-to-loop direction were more highly conserved below the miRNA/miRNA*(Figure 6A, light-blue dots), while the precursors processed in a loop-to-base fashion were more highly conserved above the miRNA/miRNA* (Figure 6A, blue dots) (Student's *t* test, $P < 0.05$).

We noticed that members of the miR165/166 family of miRNAs displayed an uneven conservation pattern (Figure 6A, orange dots). The precursor of miR165a, which was more highly conserved below the miRNA/miRNA*, showed a conserved ~15-nucleotide stem below the miRNA/miRNA* (Figure 6B), as expected from the base-to-loop processing mechanism. However, other members of the same family, such as *MIR166b* and *MIR166e*, showed a short conserved stem below the miRNA/miRNA* followed by a large internal loop and a second dsRNA segment (Figures 6C and 6D). The latter precursors also have

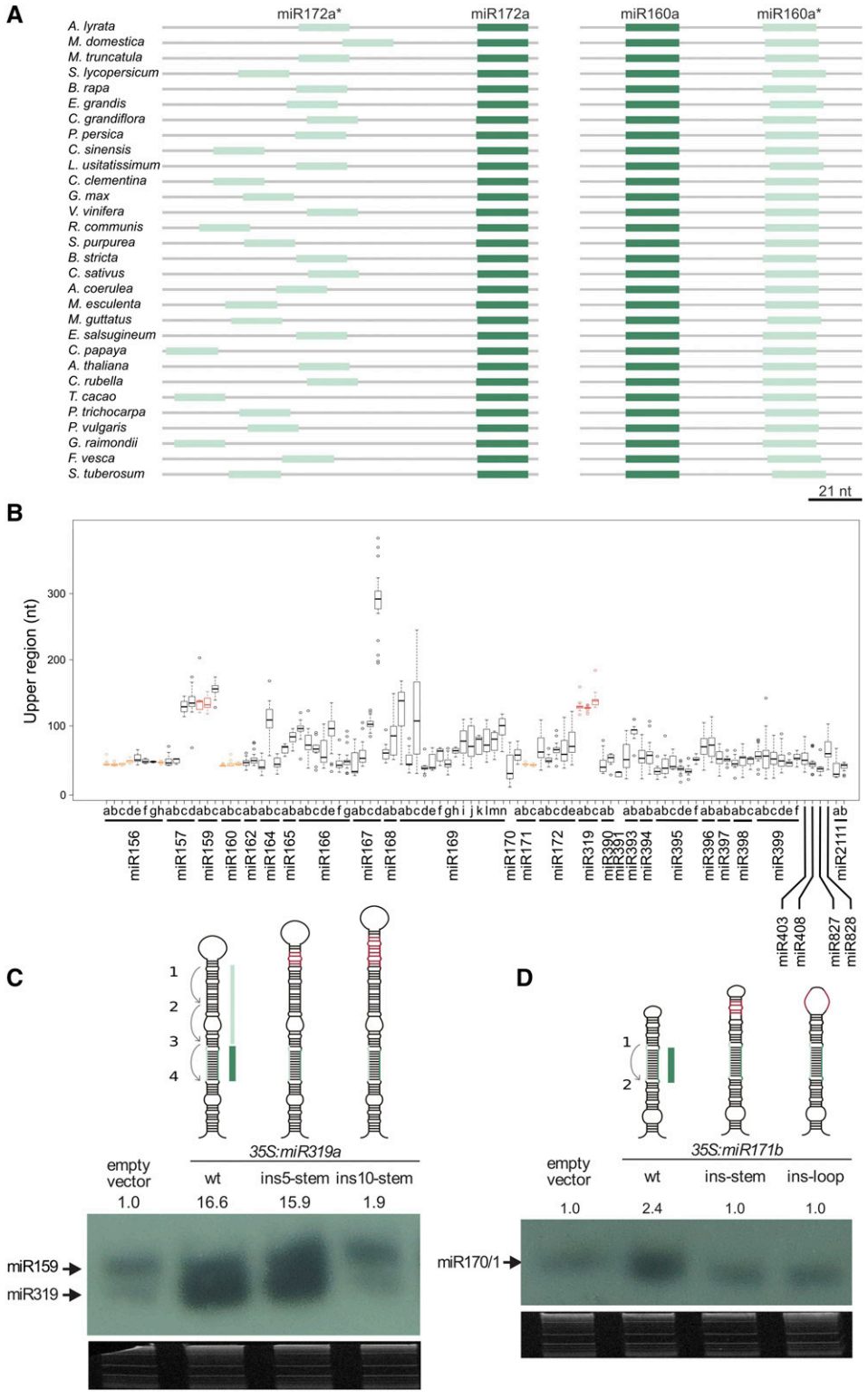

Figure 5. Precursor Length Divergence in Different Species.

(A) Schemes showing the relative positions of the miR172a and miR172a*, and miR160a and miR160* in the precursor sequences of different species. Note the changes in the relative positions for miR172* and the conservation for miR160*.

a short but conserved stem above the miRNA/miRNA* duplex (Figures 6C and 6D). The secondary structure of the Arabidopsis miR166 precursors correlated with the conserved structured regions observed in the sequence conservation analysis (Supplemental File 3; Figure 7A). These results suggest that different processing mechanisms might account for the biogenesis of miR165/miR166 miRNAs but, most importantly, that several miR166 precursors have conservation patterns that do not match the known processing mechanisms so far described in plants (Figure 4, upper panels).

Specific Sequence Requirements for miR166 Processing

To study the sequence requirements for the unusually conserved miR165/166 family members, we focused on *MIR166b*. This precursor harbors a conserved dsRNA region of ~6 nucleotides above the miRNA/miRNA* region, although the structured region in the Arabidopsis precursor is longer (Figures 6C and 7A). Below the miRNA/miRNA*, there is a short stem of ~4 nucleotides, a large internal loop followed by an additional structured segment (Figures 6C and 7A). Analysis of processing intermediates revealed a single intermediate for miR166b precursor, as well as for the precursors miR165a and miR166a (Figures 7B and 7C), which corresponds to a first cut at the base of the precursor (Figure 7A, red lines).

To study the relative importance of the miR166b precursor sequences, we generated a truncated *MIR166b* harboring only the short conserved dsRNA region of 6 nucleotides above the miRNA/miRNA* (Figure 7A, miR166bΔUS). We introduced miR165a, miR166b, and miR166bΔUS precursors in Arabidopsis plants under the control of the *35S* promoter. Analysis of the primary transcript levels by RT-qPCR revealed that all precursors were expressed in plants (Figure 7H). We detected a higher accumulation of the mature miRNA by small RNA gel blots from the unusually conserved miR166b precursor than the miR165a. Furthermore, miR166b and miR166bΔUS precursors accumulated similar levels of mature miRNA (Figure 7D), suggesting that a longer dsRNA region above the miRNA/miRNA*, which is not conserved during evolution, is not required for its processing. Then, we deleted the sequences below the conserved short stem of 4 nucleotides below the miRNA/miRNA* (Figure 7A, miR166bΔLS). We found that *35S:miR166bΔLS* lines accumulated less small RNA than *35S:miR166b*-expressing plants (Figure 7D), indicating that this lower stem had some quantitative effects on the accumulation of the miRNA, but it was not essential. This result was surprising because known precursors processed from base to the loop require an ~15-nucleotide dsRNA region below the miRNA/miRNA*, and deletions or point mutations in this region completely impair their processing (Cuperus et al., 2010; Mateos et al., 2010; Song et al., 2010; Werner et al., 2010). Previous analysis on the miR172a precursor have shown that loop sequences that are nonessential for miRNA biogenesis can enhance the processing efficiency and have been claimed to stabilize the precursor (Werner et al., 2010), which might also be the case for the deleted lower stem region of miR166b.

Finally, we prepared a mini miR166b precursor leaving only the few conserved bases next to the miR166b/miR166b* (Figures 6C and 7A, miR166bΔLSΔUS). Furthermore, the miR166bΔLSΔUS precursor was expressed and processed in plants producing the mature miRNA, albeit to lower levels than the wild-type precursor (Figures 7D to 7H). Longer exposure of the small RNA gel blot for miR166 allowed the detection of processing intermediates, which are consistent with the accumulation of a stem-loop after the first cleavage reaction (Figures 7C and 7E, yellow arrows). These intermediates accumulate at higher levels in the precursors lacking the extended dsRNA regions, suggesting that these stem segments, which are not essential for the miRNA biogenesis, might recruit processing factors that will aid to the biogenesis of miR166b, including the second cleavage reaction. We also scored the phenotypes of primary transgenic plants overexpressing the wild-type and mutant precursors and found a correlation between the mature miRNAs and the developmental defects observed (Figures 7F and 7G). Most importantly, the mini miR166b precursor was processed in vivo and caused developmental defects, confirming that the processing of miR166b does not required an ~15- to 17-nucleotide dsRNA region below or above the miRNA/miRNA* as seen in other plant precursors.

Summary and Conclusions

Here, we systematically analyzed *MIRNAs* in different species. We developed a strategy to visualize the conservation of the primary sequence and secondary structure of *MIRNAs*. A general description of plant *MIRNA* sequences revealed regions of sequence conservation that go beyond the miRNA/miRNA* and that evolutionary footprints can be linked to mechanistic processes occurring during miRNA biogenesis. The approach described here can be used as a practical tool to characterize the constraints of known processing determinants or to provide insights into new mechanisms. Furthermore, the representation allows a quantitative visualization of the conservation of the primary and secondary structures. It is known that single point mutations at specific positions modify the RNA secondary structure and impair the precursor processing (Cuperus et al., 2010; Mateos et al., 2010; Song et al., 2010; Werner et al., 2010), which can explain the conservation at the primary sequence of the structural determinants for miRNA biogenesis.

Figure 5. (continued).

(B) Box plot showing the length (in nucleotides) between the miRNA and the miRNA* of conserved precursors. Short loop-to-base precursors experimentally validated are shown in orange and sequential loop-to-base precursors are shown in red. The band inside the box represents the median, the bottom and top of the box are the first (Q1) and third quartiles (Q3), dots are outliers, upper whisker denotes min(max(x), Q3 + 1.5 * (Q3 − Q1)), and lower whisker denotes max(min(x), Q1 − 1.5 * (Q3 − Q1)).

(C) and **(D)** Small RNA gel blot of transgenic lines overexpressing miR319a **(C)** and miR171b **(D)** and mutant precursors. At least 20 independent transgenic seedlings were pooled in each sample. The relative quantification of the miRNA bands is indicated. Top: Schemes representing the precursor mutants analyzed.

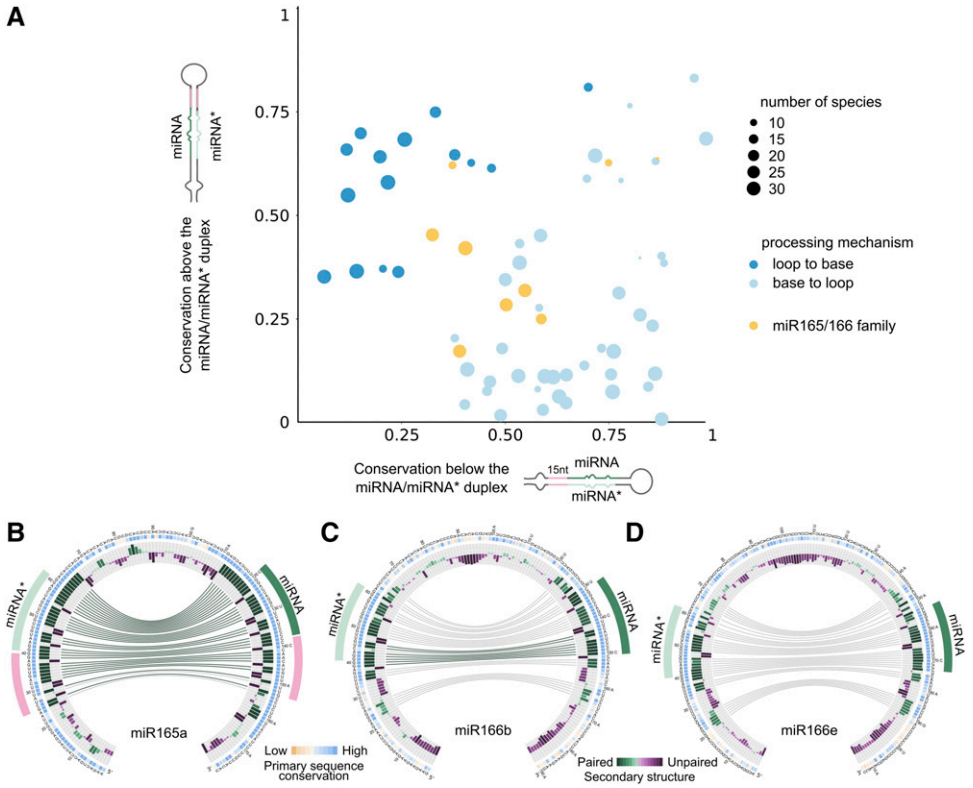

Figure 6. miR166 Family Members Have Unusual Patterns of Conservation.

(A) Scatterplot showing mean phastCons conservation score for the sequences above and below the miRNA/miRNA*. We considered 15 nucleotides on each arm of the precursor above the miRNA/miRNA* and 15 nucleotides below this region in both arms. The vertical axis represents the average phastCons conservation above the miRNA/miRNA* and the horizontal axis the average phastCons conservation below this region.
(B) to **(D)** Circos representation of miR165a **(B)**, miR166b **(C)**, and miR166e **(D)** precursors.

Extensive biochemical and genetic studies have allowed the characterization of structural determinants that promote the biogenesis of plant miRNAs (Addo-Quaye et al., 2009; Bologna et al., 2009; Cuperus et al., 2010; Mateos et al., 2010; Song et al., 2010; Werner et al., 2010). Our work revealed that the experimentally validated structural determinants can be visualized as clearly conserved regions in the *MIRNAs*. We observed conserved regions corresponding to ~15- to 17-bp stems below or above the miRNA/miRNA* depending on the direction of the precursor processing, from the base to the loop or the loop to the base, respectively. Our analysis focused on dicots and monocots, but conserved features corresponding to the lower stem of the miR390 precursor can be found in angiosperms and liverworts (Xia et al., 2017).

miR166 miRNAs fulfill key biological roles in the control of the shoot apical meristem and leaf polarity; thus, overexpression of miR165/166 affects the shoot meristem and leaf polarity (reviewed in Holt et al., 2014). *MIR166* precursors have been identified in a wide range of plant species, including mosses (Floyd and Bowman, 2004; Barik et al., 2014). Previous studies have shown that the miRNA/miRNA* duplex of miR165/miR166 have a specific structure that allows the loading into AGO10 (Zhu et al., 2011). The data obtained here suggest that the entire process of miR166 biogenesis might have specific features. For example, the fact that at least some miR166 precursors require only a few bases adjacent to the miRNA/miRNA* region is different from other known miRNAs. The processing of the miR166c precursor has been studied in detail in vitro and was shown to have a base-to-loop processing mechanism; however, in the same system, miR166a and miR166b precursors were not processed (Zhu et al., 2013). These results are consistent with different processing mechanisms acting on miR165/miR166 family members. Furthermore, we cannot overlook the recruitment of specific co-factors for the processing of these precursors in plants.

The structure of the miR319 precursors is unusual as it has a long fold-back with an additional block of sequence conservation below the loop (Palatnik et al., 2003; Axtell and Bartel, 2005; Warthmann et al., 2008; Addo-Quaye et al., 2009; Bologna et al., 2009; Li et al., 2011; Sobkowiak et al., 2012). In this regard, we would like to propose that miR319 and miR166 are two extreme examples: In the long miR319 precursor, the structural determinants for its processing are separated from the miRNA/miRNA* region generating an additional block of conservation, while in the miR166b precursor, the processing determinants are partially overlapping with the miRNA/miRNA* region.

The conservation of the miRNA itself during evolution can be explained by its function in the regulation of conserved cognate target sequences (reviewed in Cui et al., 2017). The results presented here show that conservation of the precursor processing

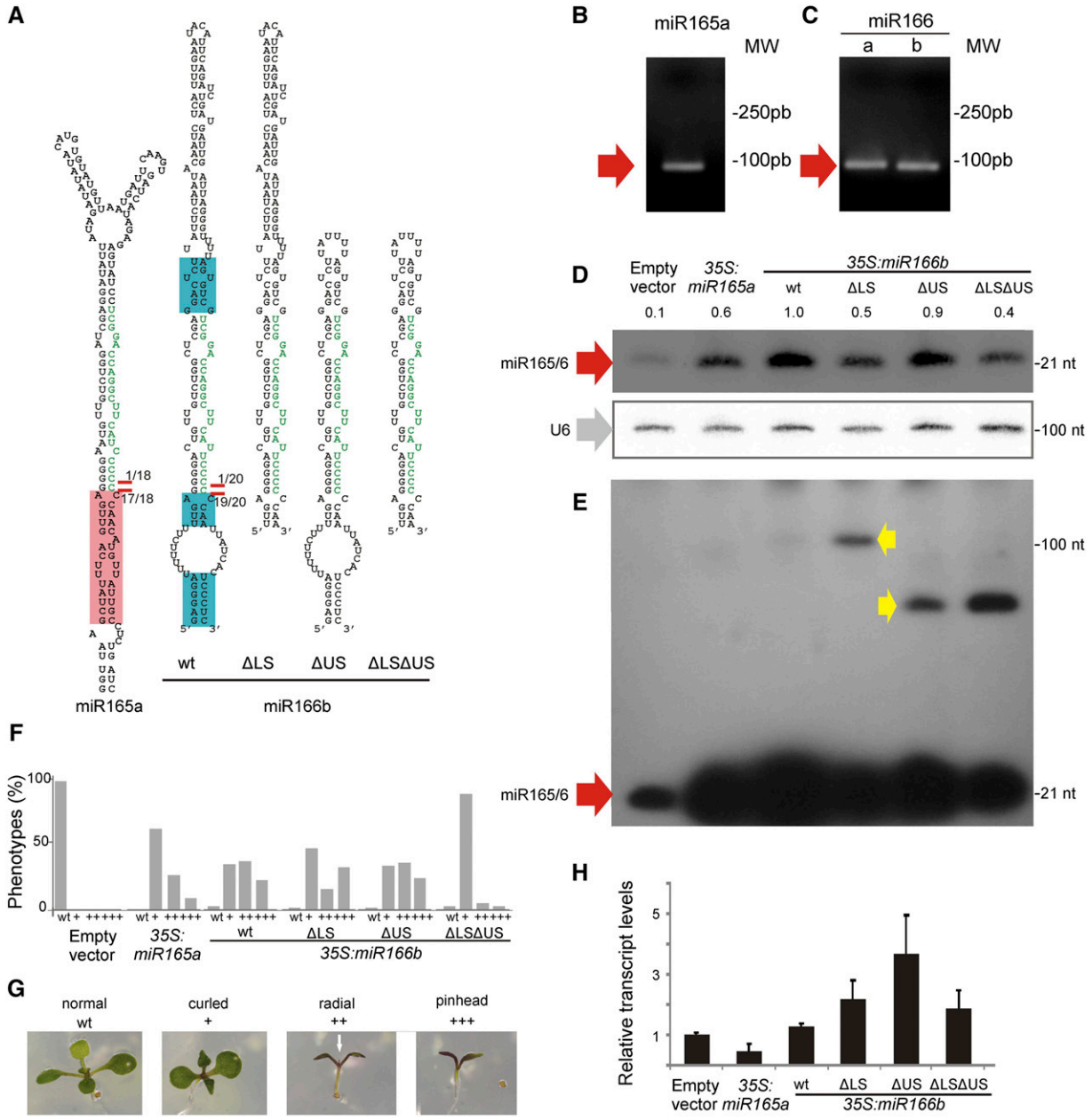

Figure 7. Sequence Requirements for miR166 Processing in Plants.

(A) Predicted secondary structures of miR165a, miR166b, and mutant miR166b precursors. Red lines and number indicate the number of processing intermediates sequenced in each position. The pink box in miR165a precursor indicates the conserved dsRNA stem of 15 nucleotides below the miRNA/miRNA*. The blue boxes in miR166b precursor indicate the conserved sequences in different species.

(B) and **(C)** Agarose gels after a modified RACE-PCR to identify processing intermediates of miR165a, miR166a, and miR166b precursors. The red arrow indicates the only DNA fragment recovered.

(D) and **(E)** Small RNA gel blot of transgenic lines overexpressing miR165a, miR166b, and mutant miR166b precursors. At least 20 independent transgenic seedlings were pooled in one sample. The U6 equal loading is shown below. The blot in **(E)** is a longer exposure of **(D)**. Red arrows indicate the miRNA while yellow arrows indicate processing intermediates detected in the blot.

(F) Phenotypic defects caused by miR165/miR166 overexpression. Distribution of phenotypic defects in plants overexpressing miR165a, miR166b, and mutant miR166b precursors. At least 50 independent primary transgenic plants were analyzed in each case.

(G) Photos of the typical developmental defects caused by the overexpression of miR165/166. The scale was used to quantify the phenotypic defects on **(F)**. The white arrow indicates a radial leaf-like organ.

(H) Primary miRNA quantification by RT-qPCR of the precursors showed in **(D)**. Error bars indicate the SE (biological triplicates).

mechanism cannot be uncoupled from the miRNA sequence and that the conservation of structural determinants can be already identified in young *MIRNAs* present only in a group of related species. Studies performed in animals have shown that a group of precursors whose processing is posttranscriptionally regulated require the binding of accessory proteins to their terminal loops, which are conserved during evolution (Michlewski et al., 2008). Furthermore, certain miRNA loops in animals can be incorporated into Argonaute complexes (Okamura et al., 2013). We think that the approach developed here can be further used to identify and predict mechanistic processes that are specific for a group of miRNA precursors as an alternative to time-consuming experimental approaches.

METHODS

Identification of Plant miRNA Precursor Orthologs

MIRNA sequences belonging to 96 evolutionarily conserved miRNAs present in *Arabidopsis thaliana* were downloaded from miRBASE release 21 (http://www.mirbase.org/). We extended the *MIRNA* sequences to 150 nucleotides outside of the miRNA and miRNA*. Plant genome sequences from 30 dicotyledonous and 6 monocotyledonous species were downloaded from Phytozome, version 11 (https://phytozome.jgi.doe.gov) for the identification of orthologs. We identified putative orthologous genes using a reciprocal BLAST hit method using in-house scripts and the NCBI Blast+ package (Altschul et al., 1990). Reciprocal BLAST was performed by running a BLAST comparison of the *MIRNA* from Arabidopsis against the genome of each of the dicotyledonous species. The highest-scoring sequence was used to run a BLAST comparison back to the Arabidopsis database. If this returned the sequence originally used as the highest scorer, then the two sequences were considered putative orthologs. The animal sequences were downloaded from the Ensembl database (Kersey et al., 2016). *Homo sapiens MIRNAs* obtained from miRbase were used to identify orthologous sequences from 13 species: *Bos taurus*, *Canis familiaris*, *Equus caballus*, *Gallus gallus*, *Gorilla gorilla*, *Macaca mulatta*, *Monodelphis domestica*, *Mus musculus*, *Ornithorhynchus anatinus*, *Petromyzon marinus*, *Sus scrofa*, and *Xenopus tropicalis.*

Multiple Sequence Alignments and RNA Secondary Structure Analysis

Multiple sequence alignments were performed using the command-line version of T-Coffee (version 11.00.8cbe486) (Notredame et al., 2000). We used the slow_pair global pairwise alignment method to build the library, recommended for distantly related sequences. Then we used the +evaluate flag to color the precursor alignments according to its conservation level. Secondary structure prediction from individual precursors in different species were made using RNAfold (Vienna RNA package version 2.1.9) (Lorenz et al., 2011).

Circos Visualization of miRNA Precursors

We used Circos (Krzywinski et al., 2009) to make a representation of the miRNA precursors. We put together data from different plant species including multiple alignments with T-Coffee, secondary structure with RNAFold, and miRNA information. The outer ring of the Circos plots shows the nucleotide sequence of Arabidopsis precursor, and the color conservation for each position in the consensus of the multiple sequence alignment according to its conservation level output from T-Coffee (Supplemental Data Set 3). We have omitted gaps and only the bases within the Arabidopsis precursor are represented. The inner Circos ring shows a histogram of the frequency distribution of paired and unpaired base for that base in the precursor. The degree of conservation of the secondary structure for each miRNA precursor was calculated using structure information in bracket notation from RNAfold. The lines with different colors show the interaction of base pairs in the precursor (green lines mean that these two bases interact in all the analyzed species and the gray lines mean that these two bases interact in at least half of the species).

Sequence Conservation Analysis

We used Phast (v1.4) for identifying evolutionarily conserved elements in a multiple alignment, given a phylogenetic tree (Siepel et al., 2005). PhyloFit was used to compute phylogenetic models for conserved and nonconserved regions among species, and these models and HMM transition parameters were used with phastCons to compute base-by-base conservation scores of aligned miRNAs precursors. Using this score, we analyzed the conservation in two contiguous regions of 15 nucleotides below the miRNA in the precursor (L1 and L1) and one region above the miRNA (U). We also considered the cognate regions next to the miRNA* (L1*, L2*, and U*). All statistical tests and plots were performed using the R statistical software package (https://www.r-project.org/). The Wilcoxon Signed-rank test was computed in R with default parameters and used in PhastCons comparisons between different regions of precursors. For the analysis of miRNAs processed from the base to the loop, we used miR164b miR164c, miR165a, miR167a, miR167b, miR167d, miR168a, miR168b, miR169a, miR170, miR171a, miR172a, miR172b, miR172c, miR172d, miR172e, miR390a, miR390b, miR393a, miR393b, miR395a, miR395b, miR395c, miR396a, miR396b, miR397a, miR398b, miR398c, miR399b, miR399c, miR403, and miR827 precursors. For the analysis of miRNAs processed from the loop to the base, we used miR156a, miR156b, miR156c, miR156d, miR156e, miR156f, miR156g, miR156 h, miR160a, miR160b, miR160c, miR162a, miR162b, miR171b, and miR171c precursors. And for the analysis of young miRNAs, we used miR158a, miR158b, miR161, miR771, and miR824 precursors.

Plant Material

All plants used in this work are Arabidopsis, accession Col-0. Seedlings were grown on agar plates with Murashige and Skoog media at continuous light at 100 μmol photons m^{-2} s^{-1} and 22°C. Described phenotypes were scored in at least 50 independent primary transgenic plants.

Transgenes and Precursor Analysis

MIR165a, *MIR166b*, *MIR319a*, and *MIR171b* were obtained from Arabidopsis genomic DNA. Site-directed mutagenesis, plant transformation, and scoring of phenotypes were performed as described previously (Bologna et al., 2013b; Zhu et al., 2013). The exact precursor sequences and vectors used here are described in Supplemental Table 1. Cleavage site mapping by modified 5′ RACE PCR was performed as described previously (Bologna et al., 2009), using 10-d-old Col-0 seedlings. The PCR products were resolved on 3% agarose gels and detected by UV exposure of the ethidium bromide.

Small RNA Analysis

Seedlings were collected and processed with TRIzol (Invitrogen). RNA gel blots were performed with 6 to 12 μg of total RNA resolved on 17% polyacrylamide denaturing gels (7 M urea). At least 20 independent transgenic plants were pooled together in one sample. For miR171, miR319 and miR165/166 antisense oligos were 5′end-labeled with [γ-^{32}P]ATP using T4 polynucleotide kinase (Fermentas). Hybridizations were performed as described previously (Bologna et al., 2009). The relative miRNA

accumulation in the small RNA gel blots was measured using GelQuant. NET software provided by biochemlabsolutions.com.

Primary miRNA transcript levels were determined by RT-qPCR. Total RNA (40 ng) was treated with RQ1 RNase-free DNase (Promega). The first-strand cDNA synthesis was performed using M-MLV reverse transcriptase (Invitrogen). PCR was performed in a Mastercycler ep realplex thermal cycler (Eppendorf) using SYBR Green I to monitor double-stranded DNA synthesis. The relative transcript level was determined for each sample, normalized to the PROTEIN PHOSPHATASE2A cDNA level (Czechowski et al., 2005). For the primary miRNA, we used primer sequences against the CHF3 transcribe regions as described previously (Supplemental Table 2) (Bologna et al., 2013b).

Accession Numbers

Sequence data from this article can be found in the GenBank/EMBL data libraries under the following accession numbers: *MIR165a*, At1g01183; *MIR166b*, At3g61897; *MIR171b*, At1g11735; and *MIR319a*, At4g23713.

Supplemental Data

Supplemental Figure 1. Insertions affect the processing of miR171b and miR319a precursors.

Supplemental Table 1. List of binary vectors used in this work.

Supplemental Table 2. List of oligonucleotide primers used for RT-qPCR.

Supplemental Data Set 1. Precursor sequences identified and used in this work.

Supplemental Data Set 2. T-Coffee alignments.

Supplemental Data Set 3. Information to generate Circos-based visualizations.

Supplemental File 1. T-Coffee alignments of putative orthologs in dicotyledonous species of the 96 Arabidopsis precursors analyzed in this work.

Supplemental File 2. T-Coffee alignments of putative orthologs in dicotyledonous and monocotyledonous species of the 96 Arabidopsis precursors analyzed in this work.

Supplemental File 3. RNAfold secondary structure predictions of the miRNA precursors analyzed in this work.

Supplemental File 4. Circos-based representation of 96 precursor orthologs in dicotyledonous species.

Supplemental File 5. Circos-based representation of 96 precursor orthologs in dicotyledonous and monocotyledonous species.

Supplemental File 6. T-Coffee alignments with compensatory mutations.

ACKNOWLEDGMENTS

We thank Detlef Weigel, Alexis Maizel, Evan Floden, Nicolas Bologna, Carla Schommer, and members of the J.F.P. lab for discussions and comments on the manuscript. This work was supported by Bunge and Born and IUBMB Wood-Whelan fellowships to U.C. and CONICET fellowships to A.M.L.R., A.L.S., B.M., and J.M.D. J.F.P. is a member of the same institution. Most of the work was supported by grants from the Argentinian Ministry of Science (PICT-2012-1780 and PICT-2015-3557 to J.F.P.) and was also supported by the Plan Nacional (BFU2011-28575 to C.N.); the Spanish Ministry of Economy and Competitiveness, 'Centro de Excelencia Severo Ochoa 2013–2017' (SEV-2012-0208); and the Center for Genomic Regulation Funding for open access.

AUTHOR CONTRIBUTIONS

U.C. performed the bioinformatic analysis and designed the CIRCOS-based visualization. B.M. and A.M.L.R. performed the miRNA experiments. J.M.D. and A.L.S. generated experimental tools. U.C., B.M., C.N., and J.F.P. analyzed the data. U.C., B.M., and J.F.P. designed the experiments. U.C. and J.F.P. wrote the article.

Received April 10, 2017; revised May 26, 2017; accepted May 26, 2017; published May 26, 2017.

REFERENCES

Addo-Quaye, C., Snyder, J.A., Park, Y.B., Li, Y.F., Sunkar, R., and Axtell, M.J. (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, A.M., Sung, G.H., Spatafora, J.W., and Carrington, J.C. (2004). Evolution of microRNA genes by inverted duplication of target gene sequences in *Arabidopsis thaliana*. Nat. Genet. **36:** 1282–1290.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. **215:** 403–410.

Axtell, M.J., and Bartel, D.P. (2005). Antiquity of microRNAs and their targets in land plants. Plant Cell **17:** 1658–1673.

Axtell, M.J., Westholm, J.O., and Lai, E.C. (2011). Vive la différence: biogenesis and evolution of microRNAs in plants and animals. Genome Biol. **12:** 221.

Barik, S., SarkarDas, S., Singh, A., Gautam, V., Kumar, P., Majee, M., and Sarkar, A.K. (2014). Phylogenetic analysis reveals conservation and diversification of micro RNA166 genes among diverse plant species. Genomics **103:** 114–121.

Bologna, N.G., and Voinnet, O. (2014). The diversity, biogenesis, and activities of endogenous silencing small RNAs in Arabidopsis. Annu. Rev. Plant Biol. **65:** 473–503.

Bologna, N.G., Mateos, J.L., Bresso, E.G., and Palatnik, J.F. (2009). A loop-to-base processing mechanism underlies the biogenesis of plant microRNAs miR319 and miR159. EMBO J. **28:** 3646–3656.

Bologna, N.G., Schapire, A.L., and Palatnik, J.F. (2013a). Processing of plant microRNA precursors. Brief. Funct. Genomics **12:** 37–45.

Bologna, N.G., Schapire, A.L., Zhai, J., Chorostecki, U., Boisbouvier, J., Meyers, B.C., and Palatnik, J.F. (2013b). Multiple RNA recognition patterns during microRNA biogenesis in plants. Genome Res. **23:** 1675–1689.

Chang, J.M., Di Tommaso, P., and Notredame, C. (2014). TCS: a new multiple sequence alignment reliability measure to estimate alignment accuracy and improve phylogenetic tree reconstruction. Mol. Biol. Evol. **31:** 1625–1637.

Chávez Montes, R.A., de Fátima Rosas-Cárdenas, F., De Paoli, E., Accerbi, M., Rymarquis, L.A., Mahalingam, G., Marsch-Martínez, N., Meyers, B.C., Green, P.J., and de Folter, S. (2014). Sample sequencing of vascular plants demonstrates widespread conservation and divergence of microRNAs. Nat. Commun. **5:** 3722.

Chorostecki, U., Crosa, V.A., Lodeyro, A.F., Bologna, N.G., Martin, A.P., Carrillo, N., Schommer, C., and Palatnik, J.F. (2012). Identification of new microRNA-regulated genes by conserved targeting in plant species. Nucleic Acids Res. **40:** 8893–8904.

Cui, J., You, C., and Chen, X. (2017). The evolution of microRNAs in plants. Curr. Opin. Plant Biol. **35:** 61–67.

Cuperus, J.T., Fahlgren, N., and Carrington, J.C. (2011). Evolution and functional diversification of MIRNA genes. Plant Cell **23:** 431–442.

Cuperus, J.T., Montgomery, T.A., Fahlgren, N., Burke, R.T., Townsend, T., Sullivan, C.M., and Carrington, J.C. (2010). Identification of MIR390a precursor processing-defective mutants in Arabidopsis by direct genome sequencing. Proc. Natl. Acad. Sci. USA **107:** 466–471.

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

de Meaux, J., Hu, J.Y., Tartler, U., and Goebel, U. (2008). Structurally different alleles of the ath-MIR824 microRNA precursor are maintained at high frequency in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA **105:** 8994–8999.

Floyd, S.K., and Bowman, J.L. (2004). Gene regulation: ancient microRNA target sequences in plants. Nature **428:** 485–486.

Ha, M., and Kim, V.N. (2014). Regulation of microRNA biogenesis. Nat. Rev. Mol. Cell Biol. **15:** 509–524.

Han, J., Lee, Y., Yeom, K.H., Nam, J.W., Heo, I., Rhee, J.K., Sohn, S.Y., Cho, Y., Zhang, B.T., and Kim, V.N. (2006). Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. Cell **125:** 887–901.

Holt, A.L., van Haperen, J.M., Groot, E.P., and Laux, T. (2014). Signaling in shoot and flower meristems of *Arabidopsis thaliana*. Curr. Opin. Plant Biol. **17:** 96–102.

Jones-Rhoades, M.W., and Bartel, D.P. (2004). Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. Mol. Cell **14:** 787–799.

Kersey, P.J., et al. (2016). Ensembl Genomes 2016: more genomes, more complexity. Nucleic Acids Res. **44:** D574–D580.

Kim, W., Kim, H.E., Jun, A.R., Jung, M.G., Jin, S., Lee, J.H., and Ahn, J.H. (2016). Structural determinants of miR156a precursor processing in temperature-responsive flowering in Arabidopsis. J. Exp. Bot. **67:** 4659–4670.

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J., and Marra, M.A. (2009). Circos: an information aesthetic for comparative genomics. Genome Res. **19:** 1639–1645.

Kurihara, Y., and Watanabe, Y. (2004). Arabidopsis micro-RNA biogenesis through Dicer-like 1 protein functions. Proc. Natl. Acad. Sci. USA **101:** 12753–12758.

Li, Y., Li, C., Ding, G., and Jin, Y. (2011). Evolution of MIR159/319 microRNA genes and their post-transcriptional regulatory link to siRNA pathways. BMC Evol. Biol. **11:** 122.

Lorenz, R., Bernhart, S.H., Höner Zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P.F., and Hofacker, I.L. (2011). ViennaRNA Package 2.0. Algorithms Mol. Biol. **6:** 26.

Lund, E., Güttinger, S., Calado, A., Dahlberg, J.E., and Kutay, U. (2004). Nuclear export of microRNA precursors. Science **303:** 95–98.

Mateos, J.L., Bologna, N.G., Chorostecki, U., and Palatnik, J.F. (2010). Identification of microRNA processing determinants by random mutagenesis of Arabidopsis MIR172a precursor. Curr. Biol. **20:** 49–54.

Meyers, B.C., et al. (2008). Criteria for annotation of plant microRNAs. Plant Cell **20:** 3186–3190.

Michlewski, G., Guil, S., Semple, C.A., and Cáceres, J.F. (2008). Posttranscriptional regulation of miRNAs harboring conserved terminal loops. Mol. Cell **32:** 383–393.

Notredame, C., Higgins, D.G., and Heringa, J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. J. Mol. Biol. **302:** 205–217.

Okamura, K., Ladewig, E., Zhou, L., and Lai, E.C. (2013). Functional small RNAs are generated from select miRNA hairpin loops in flies and mammals. Genes Dev. **27:** 778–792.

Palatnik, J.F., Allen, E., Wu, X., Schommer, C., Schwab, R., Carrington, J.C., and Weigel, D. (2003). Control of leaf morphogenesis by microRNAs. Nature **425:** 257–263.

Reinhart, B.J., Weinstein, E.G., Rhoades, M.W., Bartel, B., and Bartel, D.P. (2002). MicroRNAs in plants. Genes Dev. **16:** 1616–1626.

Rogers, K., and Chen, X. (2013). Biogenesis, turnover, and mode of action of plant microRNAs. Plant Cell **25:** 2383–2399.

Siepel, A., et al. (2005). Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res. **15:** 1034–1050.

Sobkowiak, L., Karlowski, W., Jarmolowski, A., and Szweykowska-Kulinska, Z. (2012). Non-canonical processing of Arabidopsis pri-miR319a/b/c generates additional microRNAs to target one RAP2.12 mRNA isoform. Front. Plant Sci. **3:** 46.

Song, L., Axtell, M.J., and Fedoroff, N.V. (2010). RNA secondary structural determinants of miRNA precursor processing in Arabidopsis. Curr. Biol. **20:** 37–41.

Warthmann, N., Das, S., Lanz, C., and Weigel, D. (2008). Comparative analysis of the MIR319a microRNA locus in Arabidopsis and related Brassicaceae. Mol. Biol. Evol. **25:** 892–902.

Werner, S., Wollmann, H., Schneeberger, K., and Weigel, D. (2010). Structure determinants for accurate processing of miR172a in *Arabidopsis thaliana*. Curr. Biol. **20:** 42–48.

Xia, R., Xu, J., and Meyers, B.C. (2017). The emergence, evolution, and diversification of the miR390-TAS3-ARF pathway in land plants. Plant Cell **29:** 1232–1247.

Yi, R., Qin, Y., Macara, I.G., and Cullen, B.R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes Dev. **17:** 3011–3016.

Zeng, Y., and Cullen, B.R. (2004). Structural requirements for pre-microRNA binding and nuclear export by Exportin 5. Nucleic Acids Res. **32:** 4776–4785.

Zhang, W., Gao, S., Zhou, X., Xia, J., Chellappan, P., Zhou, X., Zhang, X., and 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, S.H., Liou, L.W., Barefoot, A., Dickman, M., and Zhang, X. (2011). Arabidopsis Argonaute10 specifically sequesters miR166/165 to regulate shoot apical meristem development. Cell **145:** 242–256.

Zhu, H., Zhou, Y., Castillo-González, C., Lu, A., Ge, C., Zhao, Y.T., Duan, L., Li, Z., Axtell, M.J., Wang, X.J., and Zhang, X. (2013). Bidirectional processing of pri-miRNAs with branched terminal loops by Arabidopsis Dicer-like1. Nat. Struct. Mol. Biol. **20:** 1106–1115.