Current Biology 20, 37–41, January 12, 2010 ©2010 Elsevier Ltd All rights reserved DOI 10.1016/j.cub.2009.10.076

Report

RNA Secondary Structural Determinants of miRNA Precursor Processing in *Arabidopsis*

Liang Song,[1,2,3] Michael J. Axtell,[1,2,3] and Nina V. Fedoroff[1,2,3,4,*]
[1]The Huck Institutes of the Life Sciences
[2]Department of Biology
[3]Plant Biology Graduate Program
Pennsylvania State University, University Park, PA 16802, USA

Summary

MicroRNAs (miRNAs) are excised from hairpin structures within primary miRNAs (pri-miRNAs). Most animal pri-miRNAs are processed by two cleavages, the first at a loop-distal site ~11 nucleotides (nt) from the end of the hairpin and the second ~22 nt beyond the first [1–3]. To identify RNA structural determinants of miRNA processing in plants, we analyzed the functional consequences of changing the secondary structure of the lower (loop-distal), middle (miRNA:miRNA*), and upper (loop-proximal) stems of the hairpin in two different pri-miRNAs. Closing bulges immediately below the loop-distal cleavage sites increased the accumulation of accurately cleaved precursor miRNAs but decreased the abundance of the mature miRNAs. A pri-miRNA variant with an unpaired lower stem was not processed, and variants with a perfectly paired middle or upper stem were processed normally. Bioinformatic analysis of pri-miRNA structures, together with physical mapping of initial cleavage sites and in vitro processing of pri-miRNA, reveals that the first, loop-distal cleavage is often at a distance of ~15 nt from an unpaired region. Hence, a common determinant of the rate and location of the initial pri-miRNA cleavage is an imperfectly base-paired duplex of ~15 nt between the miRNA:miRNA* duplex and either a less structured region of the lower stem or its end.

Results and Discussion

Overexpression of Primary MicroRNAs 171a and 167a Causes Developmental Abnormalities

Overexpression of miR171, which is both deeply conserved and highly expressed, from a cauliflower mosaic virus 35S promoter (*35S:MIR171a*) causes multiple developmental defects (see Figure S1A available online). The predominant defects are fewer cauline and rosette leaves and reduced shoot branching (Figures S1A and S1B). These defects are not attributable to miR171-triggered secondary small interfering RNAs (siRNAs), because they occurred in the *sgs3* background, which is deficient in secondary siRNA production (Figure S1A) [4]. Moreover, the phenotypes were positively correlated with the amount of mature miR171 (Figure S1C). We used the number of cauline leaves as a quantifiable trait in analyzing the RNA secondary structural variants described below.

We also analyzed plants overexpressing pri-miR167a, whose microRNA (miRNA) is cleaved from the 5′ arm of the hairpin, unlike pri-miR171a, whose miRNA is cleaved from the hairpin's 3′ arm. As reported by others [5, 6], *35S:MIR167a* plants have shorter siliques as a result of reduced fertility (Figure S1D). We used silique length to quantify the effect of overexpressing pri-miR167a structural variants.

Phenotypic Analysis of Plants Expressing Primary miRNA Structural Variants

We closed bulges and repaired mismatches in pri-miR171a to make two variants: *MIR171a-Closed Bulges Lower stem* (*MIR171a-CBL*), which lacked bulges immediately below the loop-distal cleavage site, and *MIR171a-Closed Bulges Middle stem* (*MIR171a-CBM*), which lacked the middle-stem bulge in the miRNA:miRNA* duplex (Figure 1A). All T1 *MIR171a* overexpressers had fewer cauline leaves than vector-transformed controls (Figure 1B), indicating that no variants were completely defective in miR171a function. *35S:MIR171a-WT* plants had significantly fewer cauline leaves than *35S:MIR171a-CBL* plants in two independent transformations ($p < 0.01$, Mann-Whitney test; Figure 1B; Figure S2A). The phenotypic abnormalities of *35S:MIR171a-WT* plants were less severe than those of *35S:MIR171a-CBM* plants; however, the difference was not statistically significant in one of two independent transformations ($p > 0.05$, Figures 1B; $p < 0.01$, Figure S2A). These observations suggest that bulges in the lower stem immediately below the loop-distal DCL1 cleavage site are important for miRNA processing, whereas the bulge in the miRNA sequence may be less important.

We made two similar variants of pri-miR167a, *MIR167a-CBL* and *MIR167a-CBM*, as well as two additional ones: *MIR167a-Abolished Lower stem* (*MIR167a-AL*) and *MIR167a-Closed Bulges Upper stem* (*MIR167a-CBU*) (Figure 1C). More of the T1 plants developed shorter siliques than vector-transformed controls for all *35S:MIR167a* variants except *35S:MIR167a-AL*, indicating that the lower stem of the hairpin is critical for processing. However, as observed for pri-miR171a, the abnormalities of *35S:MIR167a-CBL* plants were less severe than those of *35S:MIR167a-WT* plants in three independent transformations ($p < 0.001$, Figure 1D; $p < 0.05$, Figure S2B), suggesting inefficient silencing of miRNA targets in the *CBL* lines. By contrast, the severity of the silique phenotype was similar in the *35S:MIR167a-CBU*, *35S:MIR167a-CBM*, and *35S:MIR167a-WT* T1 plants ($p > 0.05$, Figure 1D). Thus, both increasing and decreasing base pairing in the lower stem immediately below the loop-distal cleavage site appears to interfere with miRNA biogenesis, whereas eliminating bulges in the upper and middle stems does not. The difference between *35S:MIR167a-CBM* and *35S:MIR171a-CBM* suggests that the consequence of increasing base pairing in the middle stem may vary among *MIRNA*s. In summary, the results of the phenotypic analyses identify the secondary structure of the lower stem as critical for miRNA biogenesis, whereas the secondary structure of the upper and middle stems is less or not important.

*Correspondence: nvf1@psu.edu
[4]Present address: US Department of State, 2101 C Street NW, Washington, DC 20520, USA

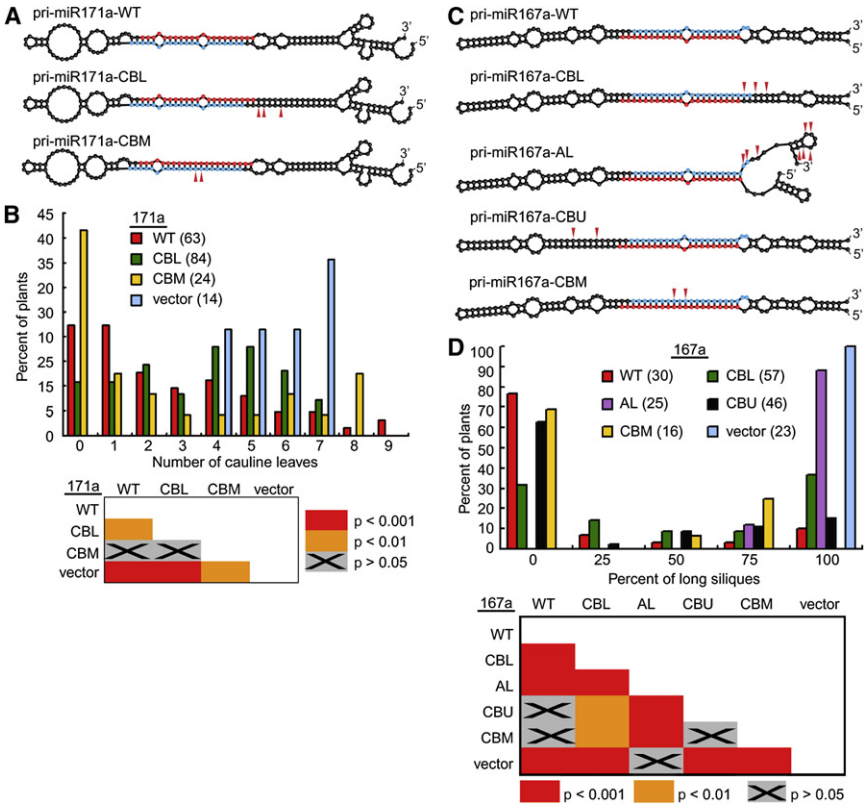

Figure 1. The Importance of the Lower Stem for *Arabidopsis* Primary miRNA Function

(A) A diagram of pri-miR171a structural variants. Mutated nucleotides (nt) are indicated by red arrowheads. The following abbreviations are used: CBL, Closed Bulges Lower stem; CBM, Closed Bulges Middle stem. Red nucleotides represent mature miR171a, and blue nucleotides represent miR171a*.

(B) Top: distribution of cauline leaf number in the indicated T1 transgenic plants. The numbers of plants analyzed are shown in parentheses. Bottom: pairwise p value cutoffs for significant differences in cauline leaf number distributions in *35S:MIR171a* T1 populations (Mann-Whitney test).

(C) As in (A) for pri-miR167a structural variants. The following abbreviations are used: AL, Abolished Lower stem; CBU, Closed Bulges Upper stem.

(D) As in (B) for the silique length phenotype of pri-miR167a-overexpressing plants.

The Lower Stem Is Critical for miRNA Processing

We then focused on plants overexpressing primary miRNAs (pri-miRNAs) with altered lower stems. The abundance of the pri-, pre-, and mature miRNAs was examined. *35S:MIRNA-CBL* plants accumulated less pri-miRNA than *35S:MIRNA-WT* plants for both *MIRNA*s (Figure 2A). Plants expressing *35S:MIR167a-AL* and *35S:MIR167a-WT* accumulated similar amounts of pri-miRNA (Figure 2A). Thus, complete disruption of the lower stem did not affect pri-miRNA abundance, whereas eliminating lower-stem mismatches reduced its accumulation. RNA blots detected species with sizes consistent with pre-miR171a and pre-miR167a specifically in samples derived from *CBL* plants (Figures 2B and 2C). Hence, *35S:MIRNA-CBL* plants accumulated less pri-miRNA and more precursor miRNA (pre-miRNA) than *35S:MIRNA-WT* plants, suggesting that DCL1 cleaves the more extensively base-paired substrate more efficiently than the mispaired substrate. However, RNA blots also showed that *35S:MIR171a-CBL* and *35S:MIR167a-CBL* plants accumulated less mature miRNA than their wild-type counterparts (Figures 2B and 2C), indicating either that the second cleavage that releases miRNA:miRNA* from the pre-miRNA is slower or that processing accuracy has decreased [7]. *35S:MIR167a-AL* and vector control plants accumulated similar amounts of miR167 and miR167a* (Figure 2C), suggesting that pri-miR167a-AL cannot be processed.

pri-miRNA secondary structural changes could affect cleavage accuracy, rate, or both. Animal Dicer processes pre-miRNAs to release the miRNA:miRNA* duplex; the cleavage site is determined by a measurement of ~22 nt [8]. If DCL1 cleaves plant pre-miRNAs by a similar mechanism, the identity of the released miRNA:miRNA* duplex would be determined solely by the position of the first DCL1 cleavage site in pri-miRNA. To assess the accuracy of the first DCL1 cleavage in *MIRNA* structural variants, we used 5′ RNA ligase-mediated rapid amplification of cDNA ends (5′-RLM-RACE) to detect 3′ remnants of cleaved pri-miRNAs. 5′-RLM-RACE results directly support the previous conclusions that both perfectly base-paired *CBL* variants are cleaved more efficiently than the wild-type pri-miRNAs and that pri-miR167a-AL is not processed by DCL1 (Figure S3). The most frequently occurring 5′ ends reflect accurate cleavage at the loop-distal site in all samples except for *35S:MIR167a-AL* (Figures 3A and 3B), indicating that processing accuracy at the loop-distal cleavage site was not affected in the *CBL* variants. This implies that the *35S:MIRNA-CBL* lines accumulate less miRNA and miRNA* because they process pre-miRNA less efficiently. We also detected more random single cleavage sites in *35S:MIRNA-WT* than in *35S:MIRNA-CBL* samples (Figures 3A and 3B). Because the wild-type transgenes are highly expressed but processed less efficiently than the *CBL* variants, the unique 5′ ends are probably due to nonspecific degradation of the transgenic pri-miRNA. Similarly, the random distribution of unique 5′ ends in pri-MIR167a-AL suggests that DCL1 was unable to process this precursor and that the remnants derived from random cleavage and degradation (Figure 3A).

The results obtained with both *CBL* variants suggest that unpaired residues immediately below the loop-distal DCL1 cleavage site reduce the rate of the loop-distal cleavage but are necessary for the efficient loop-proximal cleavage to liberate the miRNA:miRNA* duplex. In plants, both pri-miRNA and pre-miRNA are processed by DCL1 [9, 10]. The increased base pairing in the lower stem may alter the affinity of the DCL1 complex for the RNA substrates, accelerating the first cleavage but perhaps retarding rearrangement of the complex for the second cleavage. Alternatively, the more highly duplexed *CBL* variants may be cleaved by the DCL4-DRB4 complex rather than the DCL1-HYL1-SE complex. Biogenesis of some newly evolved miRNAs depends on *DCL4* instead of *DCL1*, possibly because of the long, near-perfectly base-paired structure of the pri-miRNA hairpins [11] and the greater affinity of DRB4 than HYL1 for double-stranded RNA [12].

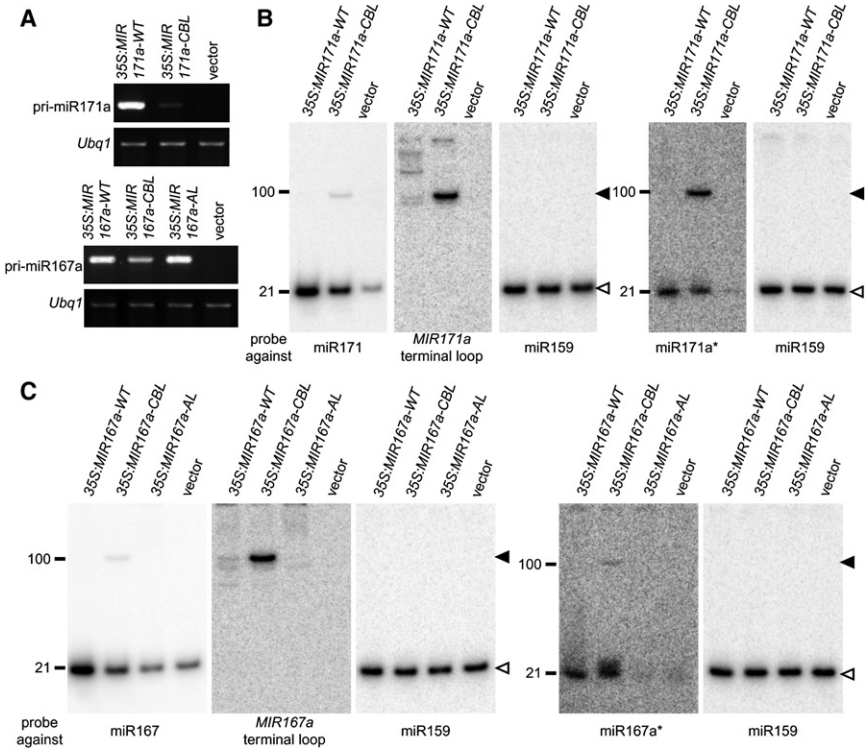

Figure 2. *CBL* Variants of *MIR167a* and *MIR171a* Accumulate Precursor miRNA

(A) Reverse transcriptase-polymerase chain reaction to detect unprocessed transgenic pri-miR171a (top) or unprocessed transgenic pri-miR167a (bottom) in the indicated transgenic lines. *Ubq1* was used as a loading control.

(B) RNA blots for the indicated species in the indicated transgenic lines. The expected mobility of pre-miR171a is indicated by solid arrowheads; mature miR171 or miR171a* is indicated by hollow arrowheads. miR159 was used as a loading control. The blot in the first three panels was reprobed twice, and the blot in the last two panels was reprobed once with indicated oligos.

(C) As in (B) for *MIR167a* variants.

RNA Secondary Structural Motifs in pri-miRNA Processing

A second cluster of cleavage sites was observed downstream from the loop-distal cleavage site in pri-miR171a constructs, which had a longer hairpin than the endogenous pri-miR171a because of sequences introduced during cloning (Figure 3B). The transgene-specific cluster occurred in a stem region 14–16 nt from a multibranched loop (Figure 3B). In vitro processing of pri-miR171a with the long hairpin by a recombinant DCL1-HYL1-SE complex released an approximately 40 nt 5′ fragment and an artificial miRNA (Figure 3C). This is consistent with the extra cleavage sites identified by 5′-RLM-RACE, indicating that the extended hairpin contains sufficient structural information for an extra DCL1 cleavage. Because all of the frequently occurring 5′ ends of pri-miR167a and pri-miR171a are located at a distance of about 15 nt from a less-structured region of the hairpin, we infer that the initial DCL1 cleavage is determined by the presence of a relatively unstructured segment in the pri-miRNA hairpin at a distance of ~15 nt from the cleavage site.

To ask whether such RNA secondary structural motifs are common in *Arabidopsis* pri-miRNA hairpins, we analyzed a filtered set of 82 *Arabidopsis MIRNA* loci selected from miRBase 13.0 [13]. These 82 loci were those with strong experimental evidence of canonical *MIRNA* processing (see Supplemental Experimental Procedures). The length of the stem regions within these hairpins was highly diverse, with heterogeneity apparent within all stem regions (Figure S4). This size heterogeneity suggests that, by contrast to what is observed in animals, pri-miRNA processing is not always initiated by a consistent distance from the end of the lower stem [3]. To identify common secondary structural features, we counted the number of hairpins in which the stem ends or a "big loop" (defined as a loop contained within a single stem arm that contains no less than three consecutive unpaired nucleotides) begins at each position in the predicted hairpins. A clear peak of stem ends and/or big loops is observed in the lower stem 15–17 nt from the loop-distal cleavage site, adjacent to a valley in the distribution at 12–13 nt (Figure 4). The frequencies of stem ends and big loops observed at positions 15–17 in the lower stem are highly unlikely to have occurred by chance ($p < 0.001$). In 41 of the 82 analyzed pri-miRNA hairpins, the loop-distal DCL1 cleavage site resides ~15 nt from an unpaired region in the lower stem. Stem ends and big loops are more evenly distributed in the upper stem, indicating that the upper stem lacks such a high-frequency secondary structural motif.

Conclusions

We have demonstrated that the lower stem is critical for miRNA biogenesis. This is consistent with observations described in the accompanying papers by Werner et al. and Mateos et al. in this issue of *Current Biology* [14, 15]. We also discovered that closing bulges immediately below the loop-distal DCL1 cleavage site has opposite impacts on the efficiency of first and second DCL1 cleavages. We speculate that during evolution, mutations in the lower stem immediately below the initial DCL1 cleavage site may fine-tune the amount of miRNAs. Our analyses also suggest that the first DCL1 cleavage in miRNA processing often occurs at a distance of ~15 nt from either the end of the hairpin or an internal unstructured region. In agreement with this "+15 rule," Werner et al. [14] show that changing the position of bulged nucleotides in the lower stem of pri-miR172a changes the position of the lower DCL1 cleavage site. The +15 mechanism may be important in the evolution of *MIRNA*s. Transcripts of evolutionarily young *MIRNA* genes derived from inverted gene duplication have long hairpins [11, 16]. These transcripts are processed by DCL1 or DCL4 into heterogeneous small RNAs whose ends are roughly in 21 nt phase with the initial DCL cleavage site [11, 16]. Random mutations, particularly short indels, may introduce internal loops and therefore new initial DCL cleavage sites, changing the sequences of small RNAs liberated from the hairpins. However, the +15 rule is not sufficient to completely specify DCL1 processing accuracy, because aberrant cleavages in the pri-miR171a's extended lower hairpin occur at roughly equal frequency at distances of 14, 15, and 16 nt from the less-structured internal loop, whereas canonical cleavages occur at

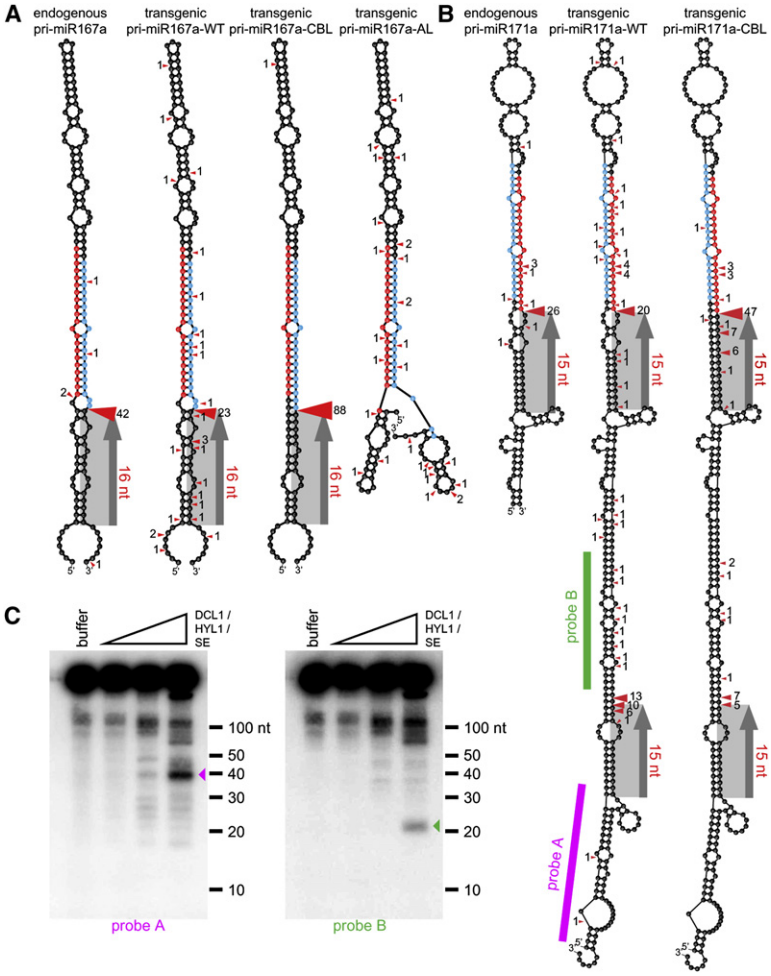

Figure 3. *CBL* Variants Are Accurately Processed at the Loop-Distal Cleavage Site

(A) 5′ positions and frequencies of the cleavage sites in pri-miR167a. Identified sites are indicated by red arrowheads, and the number of times each was observed is given by the adjacent number. Red nucleotides represent mature microRNA (miRNA), and blue nucleotides represent miRNA*.
(B) As in (A) for pri-miR171a variants. Magenta and green bars represent the positions of probes in (C) that are complementary to the nucleotides in the 5′ arm of the hairpin.
(C) In vitro processing of pri-miR171a by a recombinant DCL1-HYL1-SE complex. RNA substrate has the same sequence as the transgenic pri-miR171a-WT shown in (B) except for two extra Gs at the 5′ end introduced by the T7 promoter. The panels are of the same blot hybridized with different oligonucleotide probes to detect the 5′ fragment released by the initial DCL1 cleavage (magenta arrowhead) and the artificial miRNA (green arrowhead) liberated by two DCL1 cleavages in the extended hairpin.

precisely +15 and +16 nt in pri-miR171a and pri-miR167a, respectively. Moreover, computational analysis indicates that there are accurately processed pri-miRNAs that do not conform to the +15 rule.

Supplemental Information

Supplemental Information includes Supplemental Experimental Procedures, four figures, and three tables and can be found with this article online at doi:10.1016/j.cub.2009.10.076.

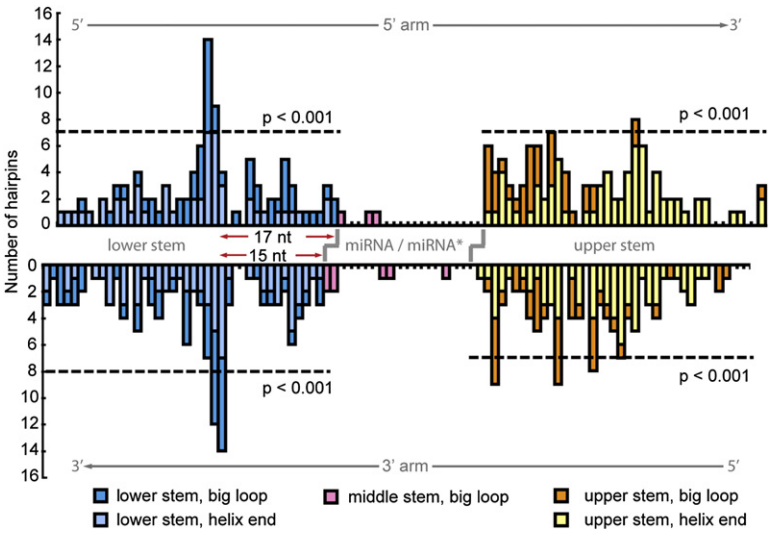

Figure 4. Unpaired Regions in the Lower Stem Are Frequently Located ~15 Nucleotides from the miRNA: miRNA* Duplex

Frequencies of primary miRNAs in which the hairpin ends or a big loop (≥3 unpaired nt within the stem) begins at the indicated hairpin positions either in the 5′ (top) or 3′ (bottom) arm of the hairpin. Blue, pink, and yellow bars represent helix ends and big loops in the lower, middle, and upper stems, respectively. p values are estimated by 100,000 simulated random distributions of features in each region.

Acknowledgments

We thank H. Ma, T. Nakagawa, and R. Singer for sharing seeds and plasmids. We also thank D. Weigel, J. Carrington, J. Palatnik, and S. Werner for sharing data prior to publication. This study was supported by award MCB-0640186 to N.V.F. and by award MCB-0718051 to M.J.A. from the United States National Science Foundation.

Received: July 19, 2009
Revised: October 3, 2009
Accepted: October 29, 2009
Published online: December 14, 2009

References

1. Lee, Y., Jeon, K., Lee, J.T., Kim, S., and Kim, V.N. (2002). MicroRNA maturation: Stepwise processing and subcellular localization. EMBO J. *21*, 4663–4670.
2. Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. Cell *75*, 843–854.
3. 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.
4. Peragine, A., Yoshikawa, M., Wu, G., Albrecht, H.L., and Poethig, R.S. (2004). *SGS3* and *SGS2/SDE1/RDR6* are required for juvenile development and the production of trans-acting siRNAs in *Arabidopsis*. Genes Dev. *18*, 2368–2379.
5. Ru, P., Xu, L., Ma, H., and Huang, H. (2006). Plant fertility defects induced by the enhanced expression of microRNA167. Cell Res. *16*, 457–465.
6. Wu, M.F., Tian, Q., and Reed, J.W. (2006). *Arabidopsis* microRNA167 controls patterns of *ARF6* and *ARF8* expression, and regulates both female and male reproduction. Development *133*, 4211–4218.
7. Mi, S., Cai, T., Hu, Y., Chen, Y., Hodges, E., Ni, F., Wu, L., Li, S., Zhou, H., Long, C., et al. (2008). Sorting of small RNAs into *Arabidopsis* argonaute complexes is directed by the 5′ terminal nucleotide. Cell *133*, 116–127.
8. Jinek, M., and Doudna, J.A. (2009). A three-dimensional view of the molecular machinery of RNA interference. Nature *457*, 405–412.
9. Kurihara, Y., and Watanabe, Y. (2004). *Arabidopsis* micro-RNA biogenesis through Dicer-like 1 protein functions. Proc. Natl. Acad. Sci. USA *101*, 12753–12758.
10. Park, W., Li, J., Song, R., Messing, J., and 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.
11. Rajagopalan, R., Vaucheret, H., Trejo, J., and Bartel, D.P. (2006). A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*. Genes Dev. *20*, 3407–3425.
12. Pouch-Pélissier, M.N., Pélissier, T., Elmayan, T., Vaucheret, H., Boko, D., Jantsch, M.F., and Deragon, J.M. (2008). SINE RNA induces severe developmental defects in *Arabidopsis thaliana* and interacts with HYL1 (DRB1), a key member of the DCL1 complex. PLoS Genet. *4*, e1000096.
13. Griffiths-Jones, S., Saini, H.K., van Dongen, S., and Enright, A.J. (2008). miRBase: Tools for microRNA genomics. Nucleic Acids Res. *36*(*database issue*), D154–D158.
14. Werner, S., Wollmann, H., Schneeberger, K., and Weigel, D. (2010). Structure determinants for accurate processing of miR172a in *Arabidopsis thaliana*. Curr. Biol. *20*, this issue, 42–48.
15. 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*, this issue, 49–54.
16. 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.

Note Added in Proof

While this manuscript was in press, a paper by Cuperus et al., discussing the importance of the loop-distal region in the *MIR390* hairpin for miRNA biogenesis, was accepted for publication. The citation details are as follows: Cuperus, J.T., Montgomery, T.A., Fahlgren, N., Burke, R.T., Townsend, T., Sullivan, C.M., and Carrington, J.C. (2009). Identification of *MIR390a* precursor processing-defective mutants in Arabidopsis by direct genome sequencing. Proc. Natl. Acad. Sci. USA *106*. Published online December 14, 2009. 10.1073/pnas.0913203107.