

# The *Phaseolus vulgaris* miR159a precursor encodes a second differentially expressed microRNA

Cecilia Contreras-Cubas · Fernando A. Rabanal ·  
Catalina Arenas-Huertero · Marco A. Ortiz ·  
Alejandra A. Covarrubias · José Luis Reyes

Received: 26 November 2010 / Accepted: 30 October 2011 / Published online: 15 November 2011  
© Springer Science+Business Media B.V. 2011

**Abstract** Plant microRNAs originate from a stem-loop structured single-stranded RNA precursor. Each stem-loop is processed to generate a mature microRNA that is recruited to an ARGONAUTE-containing multi-protein complex to direct silencing of its target mRNA. Here we report that the conserved plant miR159a precursor produces a second 21-nt long RNA with the properties of a microRNA. Its presence in different plant species is supported by its conservation in the stem-loop position and

expression as determined by northern blot analysis. We show that successive processing by DCL1 produces this novel microRNA from the same precursor as miR159a. In contrast to the low levels observed in other plant models for the equivalent of miR159.2, in *P. vulgaris*, the accumulation of miR159.2 is easily detectable and when compared to miR159a, their expression patterns are distinct in different organs and growth conditions. Further evidence of the functionality of miR159.2 comes from its association with silencing complexes as demonstrated by co-immunoprecipitation experiments using an AGO1-specific antibody and processing of an artificial GFP reporter construct containing a complementary target sequence. These results indicate that the second small RNA corresponds to a microRNA, at least partially independent of miR159 activity, and that in plants a miRNA precursor may encode multiple regulatory small RNAs.

**Electronic supplementary material** The online version of this article (doi:10.1007/s11103-011-9847-0) contains supplementary material, which is available to authorized users.

C. Contreras-Cubas · F. A. Rabanal · C. Arenas-Huertero ·  
M. A. Ortiz · A. A. Covarrubias · J. L. Reyes (✉)  
Depto. Biología Molecular de Plantas, Instituto de  
Biotecnología, Universidad Nacional Autónoma de México, A.P.  
62210, Cuernavaca, Morelos, Mexico  
e-mail: jlreyes@ibt.unam.mx

C. Contreras-Cubas  
Doctorado en Ciencias Biomédicas, Universidad Nacional  
Autónoma de México, A.P. 62210, Cuernavaca, Morelos,  
Mexico

F. A. Rabanal · M. A. Ortiz  
Licenciatura en Ciencias Genómicas, Universidad Nacional  
Autónoma de México, A.P. 62210, Cuernavaca, Morelos,  
Mexico

**Present Address:**  
F. A. Rabanal  
Gregor Mendel Institute of Molecular Plant Biology, Austrian  
Academy of Sciences, Vienna, Austria

**Present Address:**  
C. Arenas-Huertero  
The Rockefeller University, 1230 York Ave., New York, NY  
10021, USA

**Keywords** *Phaseolus vulgaris* · Pre-miRNA · Plant  
microRNA · ARGONAUTE 1 · miRNA processing

## Introduction

MicroRNAs (miRNAs) are small RNA molecules, 20–24 nucleotides in length, recognized as important regulators of gene expression in animals and plants (Bartel 2004; Voinnet 2009). In plants, each miRNA precursor (pre-miRNA) is processed by the Dicer-like enzyme DCL1 through two consecutive cleavage reactions to generate a single small duplexed-RNA containing the miRNA and its partially complementary strand or miRNA\*, leaving a 5'-phosphate and a two-nucleotide overhang at the 3'-end (Kurihara and Watanabe 2004; Reinhart et al. 2002). This processing is aided by a number of protein factors, including

*HYPONASTIC LEAVES 1* (HYL1) and *SERRATE* (SE) that help increasing cleavage accuracy by DCL1 (Dong et al. 2008), and *HUA-ENHANCER 1* (HEN1) that modifies the 3'-end of small RNAs, including miRNAs and siRNAs, through its RNA methyl transferase activity (Yu et al. 2005). The miRNA strand is then selectively recruited into silencing complexes, where the AGO1 protein recognizes the miRNA and uses it as a specificity guide to target complementary mRNAs for silencing, either by cleavage and degradation or by translational inhibition (Baumberger and Baulcombe 2005; Vaucheret 2008).

The miR159 family targets MYB transcription factor mRNAs and is conserved in many plant species including angiosperms, mosses and lycopods (Axtell et al. 2007; Rhoades et al. 2002). Its precursor is unusually long and additional small RNAs originating from the precursor have been reported through large-scale sequencing in *Arabidopsis* (Fahlgren et al. 2007; Rajagopalan et al. 2006) and genome-scale analyses (Li et al. 2011; Zhang et al. 2010). Recently, two groups have independently described the particular processing of the miR159 precursor (and its relative miR319) where a stepwise mechanism allows DCL1 to cleave the precursor in a loop-to-stem direction where the additional small RNAs are considered as intermediates in the maturation of miR159/miR319 (Addo-Quaye et al. 2009; Bologna et al. 2009). Because of their low abundance, in *Arabidopsis* these intermediate molecules were only detected by high-throughput sequencing (Fahlgren et al. 2007; Rajagopalan et al. 2006), but not by hybridization with probes specifically designed for northern blot experiments (this study, data not shown). Our past work with small RNAs from the legume *Phaseolus vulgaris* (common bean) led us to identify miR159.2 as a small RNA present in different legumes and induced in response to water deficit conditions (Arenas-Huertero et al. 2009). In common bean, miR159.2 is encoded by the same precursor as miR159a. In contrast to the low abundance of the equivalent small RNA found in *Arabidopsis*, *P. vulgaris* miR159.2 is abundant and easily detectable by northern blot, suggesting it is a functional miRNA. It has been recently shown that a number of plant precursors produce additional miRNAs different from the canonical miRNAs already described, however evidence to indicate so for *Arabidopsis* miR159.2 was limited (Zhang et al. 2010). Here we characterize it as a conserved miRNA and provide evidence of its expression in different plant species and with the properties of a miRNA, including its modification at the 3'-end, consistent with it being a substrate for HEN1-directed methylation. We show that the stepwise precursor processing already described in other plants occurs in common bean and further that the expression of miR159.2 is selective, as other fragments of the precursor do not accumulate in vivo. Interestingly, the pattern of accumulation for miR159.2 does not mirror that of

miR159, suggesting that each miRNA within this precursor has its own regulation mechanisms and may be involved in independent response pathways. Finally, even though co-immunoprecipitation with AGO1 indicates that miR159.2 is recruited to silencing complexes, suggesting that *pvu*-miR159.2 is functional in common bean, our efforts to identify target mRNAs have been thwarted due to our limited knowledge of genome sequence information for common bean. Similarly, others have also failed to identify targets for the equivalent of miR159.2 in *Arabidopsis* (Axtell et al. 2007). Nonetheless, its expression in the *Nicotiana benthamiana* heterologous system results in the accurate processing of a GFP reporter transcript, consistent with *pvu*-miR159.2 being a functional microRNA. In addition, its distinct accumulation pattern suggests that its role may be independent from that of *pvu*-miR159a in legumes and other species as well, including rice and tobacco.

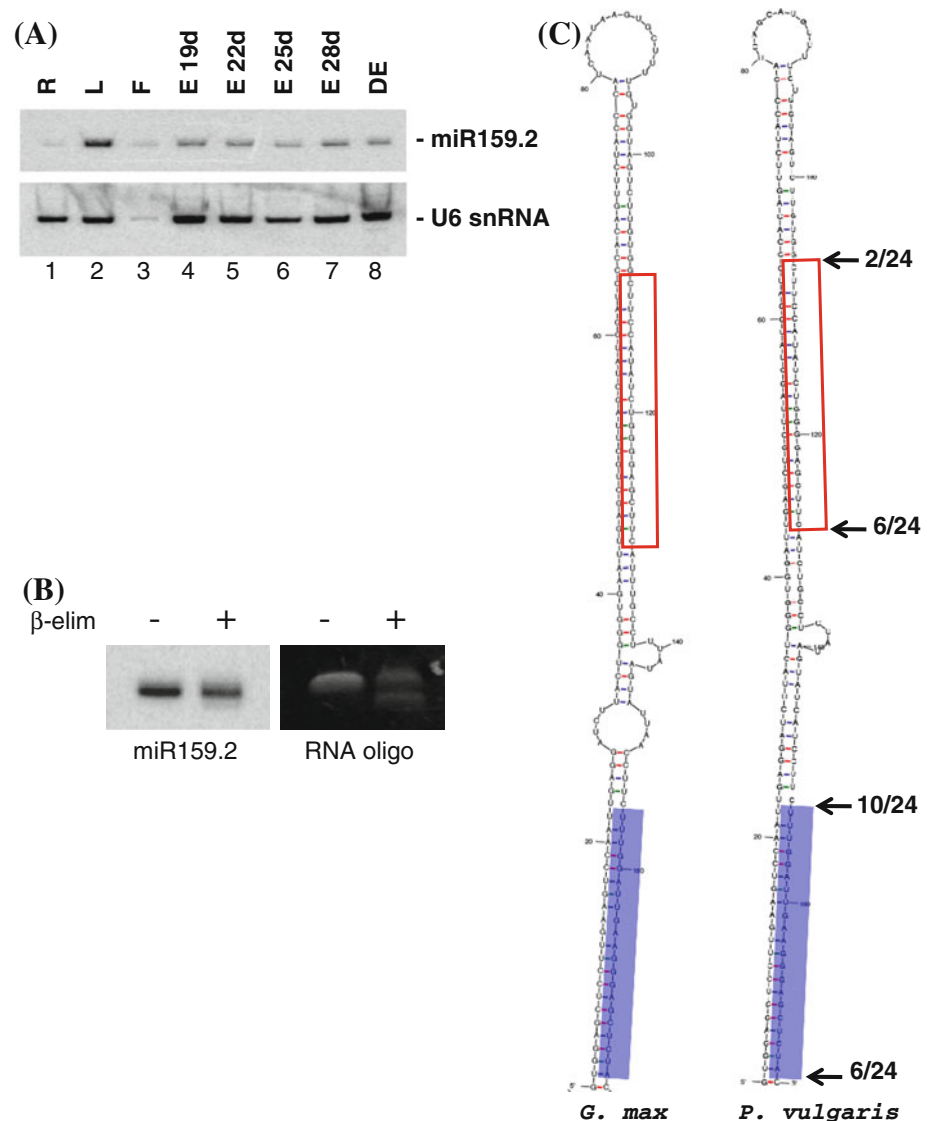
## Results

### A second small RNA in the precursor of miR159

In the course of small RNA characterization from the legume *P. vulgaris*, we identified a 21-nt long RNA that is readily detected using northern blot analysis of different organ samples (hereafter miR159.2, Fig. 1a; Arenas-Huertero et al. 2009). No other RNA species could be detected in the same blot (data not shown), indicating that the small RNA is not a degradation product of a larger molecule. We also determined that miR159.2 is modified at its 3'-end as it is resistant to  $\beta$ -elimination (Fig. 1b), consistent with it being the product of HEN1-directed methylation (Yu et al. 2005). These data established that miR159.2 is a real small RNA and potentially processed by the miRNA pathway machinery in common bean.

A search of public nucleotide databases identified an EST from *Glycine max*, a closely related legume species, containing an identical 21-nt sequence. Further inspection revealed that the novel small RNA sequence was part of the same stem-loop structure encoding soybean miR159a contained within the EST that we initially identified (miRBase, ver. 16.0, Griffiths-Jones 2004). Based on the soybean EST sequence, we obtained the corresponding sequence from *P. vulgaris* by PCR amplification. In both *G. max* and *P. vulgaris*, miR159a and miR159.2 are located in the 3' arm of the stem-loop structure, separated by an imperfect stem of around 20 base-pairs (Fig. 1c). To address whether the arrangement present in the legume lineage is conserved in other species, we analyzed pre-miR159a sequences annotated in miRBase ([www.mirbase.org](http://www.mirbase.org)), as well as ESTs from different plant species found in NCBI that contained the highly conserved miR159a

**Fig. 1** miR159.2 is encoded in the *pvu*-miR159a precursor. **a** RNA samples from *P. vulgaris* roots (*R*), leaves (*L*), flowers (*F*), developing embryos (*E*) of 19, 25, 28 days post-flowering, or dry embryos (*DE*) from mature seeds were used in northern blot analysis to detect the expression of *pvu*-miR159.2. U6 snRNA hybridization was used as loading control. **b** Leaf RNA samples from *P. vulgaris* were subjected to  $\beta$ -elimination treatment (+) or mock-treated (–), blotted and detected by hybridization using a *pvu*-miR159.2 probe. As control, an RNA oligonucleotide was treated in parallel and resolved in the same gel, and the ethidium bromide staining is shown. **c** Secondary structure of the *G. max* and *P. vulgaris* miR159a precursor as predicted by mFold. The location of miR159a (filled blue boxes) and miR-159.2 (red boxes) is indicated. Arrows indicate the position of processing as determined by 5'RACE, and numbers indicate number of clones obtained at each position (see text for details)



sequence and able to conform to the secondary structure typical of a pre-miRNA as determined by the mfold program (Zuker 2003). Using the T-Coffee program (Notre-dame et al. 2000) for alignment of the collected sequences from diverse dicotyledoneous and monocotyledoneous plant species, we identified conserved blocks of sequence similarity (Fig. 2a). We observed sequence conservation around the regions encoding miR159a and miR159a\*, where precursor secondary structure and function constrain sequence variation, as previously observed for conserved miRNA precursors. In addition, we observed another two blocks of sequence conservation, one corresponding to miR159.2, and a second one located at the position corresponding to its complementary sequence. Other positions along the loop and in other double-stranded regions of the precursor are less conserved, suggesting a lower degree of selective pressure. The overall sequence of miR159.2 itself has been conserved during plant diversification with a

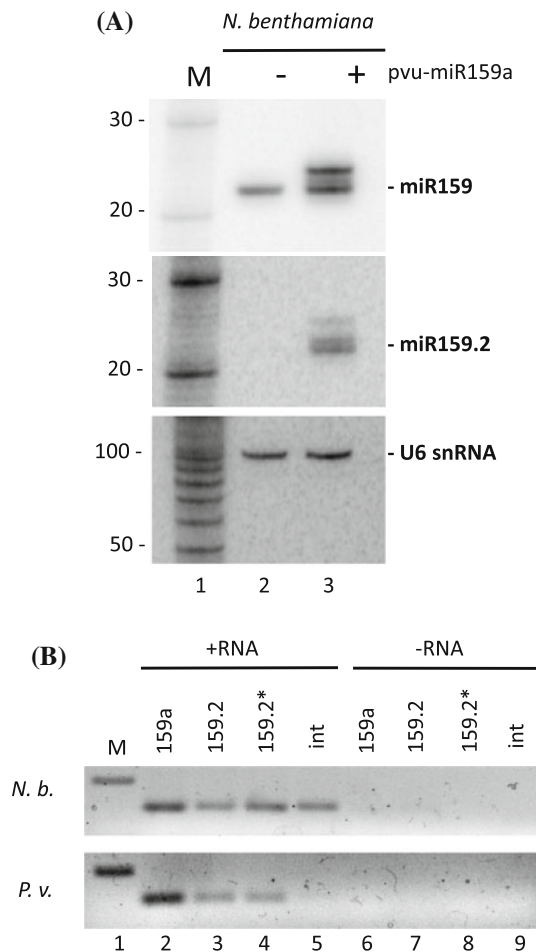
limited number of changes that still allow easy alignment of the small RNA sequences (Fig. 2b). Together, these observations suggest that pre-miR159a encodes a second microRNA, displaying conservation across plant species, both in sequence and in position along the stem-loop structure.

The miR159a precursor is sequentially processed

miR159a and miR-159.2 are separated by 21–24 nts in most of the precursors analyzed, with the notable exception of pre-*ath*-miR159a, where the spacer sequence corresponds to 29 nts (see Fig. 2a). To determine if both RNAs are indeed originating from the same precursor we followed two complementary strategies. First, we used *Agrobacterium*-mediated infiltration to express the *pvu*-miR159a precursor under the control of the 35S promoter in *N. benthamiana* leaves followed by northern blot



analysis to monitor expression of the transgene. Endogenous *nbe*-miR159 was detected in non-infiltrated leaves with the same probe used to detect *pvu*-miR159a (Fig. 3a). In comparison, infiltrated leaves showed an increased accumulation of miR159a and a secondary product consistent with the expression and processing of *pvu*-pre-miR159a (Fig. 3a, lane 3). Next, we tested whether *pvu*-miR159.2 could be detected in the same RNA samples. Figure 3a shows that *pvu*-miR159.2 is efficiently processed from the introduced precursor in *N. benthamiana*. Because the sequence in common bean miR159.2 is sufficiently



**Fig. 3** Processing of the miR159a precursor in *P. vulgaris*. **a** *N. benthamiana* leaves were infiltrated with an *Agrobacterium* strain carrying a 35S: pre-*pvu*-miR159a construct and after 2 days the leaves were collected for RNA isolation. Total RNA (10  $\mu$ g) for each sample was analyzed by northern blot to detect miR159 (upper panel), miR159.2 (middle panel) or U6 snRNA (bottom panel). Molecular size markers (*M*) are shown to the left in each panel. **b** Stem-looped RT-PCR analysis was performed in the same samples as in **a** to detect miR159a (lane 2), miR159.2 (lane 3), miR159.2\* (lane 4) and the intermediate region (int, lane 5) between mir159a and mir159.2 (*N.b.*, upper panel). Alternatively, total RNA from *P. vulgaris* leaves was used to detect the same small RNAs (*P.v.*, bottom panel). The ‘- RNA’ reactions were used for reverse transcriptase and PCR negative controls. The *M* lane indicates 100 bp molecular size marker

different from that of *Nicotiana* miR159.2, we did not observe a hybridization signal in non-infiltrated leaves (lane 2, middle panel in Fig. 3a).

Similar results were obtained by using a more sensitive approach based on stem-looped RT-PCR designed to amplify small RNAs. When the precursor was expressed using the viral 35S promoter in *N. benthamiana* transient expression experiments, we detected miR159a and miR159.2, as well as the small RNA sequence separating them (called ‘int’, for intermediate in Fig. 3b, lane 5). A similar RT-PCR experiment using *P. vulgaris* total RNA failed to detect the intermediate RNA, while miR159a and miR159.2 were again easily detectable, suggesting they are more abundant than the intermediate fragment in vivo (Fig. 3b, bottom panel). Put together, these results indicate that a single RNA precursor is specifically processed in vivo to generate two different small RNAs, miR159 and miR159.2.

In a second experiment, we took advantage of the previous studies of the miR159a and miR319 families and their precursor processing in *Physcomitrella patens* and *A. thaliana* (Addo-Quaye et al. 2009; Bologna et al. 2009). It was shown that these precursors are processed in a loop-to-stem direction, where the equivalent RNA of that of *pvu*-miR159.2 is generated as a processing intermediate during miR159a maturation. Using a modified 5’RACE approach to determine the processing pattern of *P. vulgaris* pre-miR159a present in total RNA samples obtained from leaves, we mapped cleavage sites along the precursor consistent with a loop-to-stem processing of the precursor (shown as arrows in Fig. 1c, see figure legend), and precisely generating the RNA molecules corresponding to miR159a and miR159.2.

In *A. thaliana* the processing intermediates are not detectable by northern blot (this work, data not shown) and only a small number of reads were recovered from previously reported deep-sequencing experiments (Fahlgren et al. 2007; Rajagopalan et al. 2006), consistent with the very low abundance of these fragments. In contrast, miR159.2 is readily detectable by northern blot in common bean as well as other legumes (Arenas-Huertero et al. 2009). These results are supported by the analysis of the small RNA sequences obtained from a recent high-throughput small RNA sequencing project from different *P. vulgaris* organs as well as from other legumes (Nakano et al. 2006). This analysis showed that miR159a is the most abundant small RNA originating from its precursor; interestingly, miR159.2 is as abundant as miR159 in some organs, such as symbiotic nodules (123 and 127 per 100,000 reads, respectively, see Supplemental Fig. 1). In contrast, the same analysis indicated that neither the intermediate sequence nor the corresponding miRNAs\* could be detected except for a few reads (a maximum of 13 reads in developing seeds). The same database was searched for small RNAs derived from the *G. max*

pre-miR159a. Similar to what we found for *P. vulgaris*, the most abundant sRNA was miR159a, followed by miR159.2 (236 and 222 in developing seeds, respectively), with very few reads corresponding to other regions of the miR159a precursor (see Supplemental Fig. 2). Furthermore, in common bean developing seeds miR159.2 is more abundant than miR159 (Supplemental Fig. 1). These results indicate that in contrast to what was observed in Arabidopsis, miR159.2 readily accumulates in *P. vulgaris* and *G. max*. The miR159.2 accumulation levels, together with the fact that other less-conserved small RNAs derived from the same precursor are rarely present in vivo, suggest that this miRNA is functional in legumes.

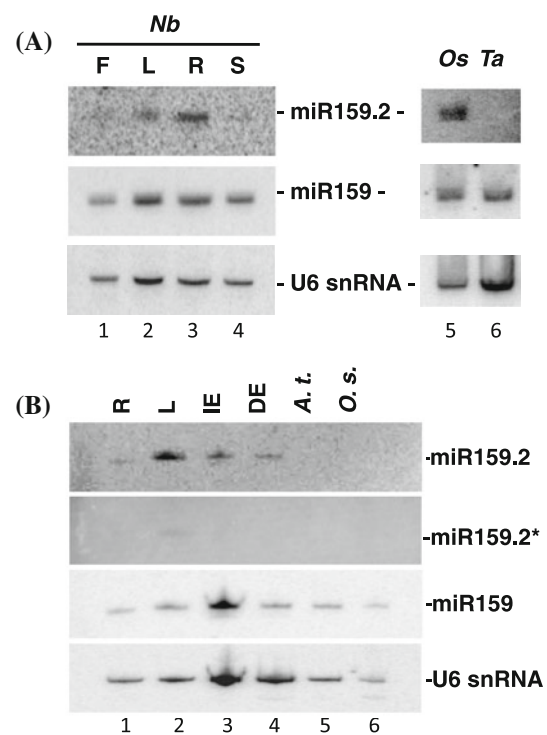
miR159.2 is expressed in different plant species

Because miR159.2 accumulates in common bean but not in Arabidopsis, and because it was easily detected in high-throughput sequencing of soybean small RNAs (Nakano et al. 2006), we decided to further explore whether it could be detected at all in Arabidopsis as well as other species that show conservation in the sequence for miR159.2 (Fig. 2b). We performed northern blot analysis of Arabidopsis leaf RNA samples from transgenic plants overexpressing the viral suppressor P1/HC-Pro from Cucumber Mosaic Virus (CuMV) which were previously shown to over-accumulate miRNAs, miRNAs\* and siRNAs when compared to wild type plants (Mallory et al. 2002). The Arabidopsis sequence equivalent to miR159.2 is different from that of *P. vulgaris* at 5 nucleotide positions (Fig. 2b), thus northern blot hybridizations were performed with an Arabidopsis-specific probe. As expected, *ath*-miR159 accumulated to a larger extent in the presence of P1/HC-Pro, however the corresponding *ath*-miR159.2 could not be detected (data not shown). This result is consistent with cloning frequencies found in deep-sequencing experiments where *ath*-miR159.2 originating from pre-miR159b, 159c or the related pre-miR319 RNAs showed extremely low abundance levels (Fahlgren et al. 2007; Rajagopalan et al. 2006), together with a very low abundance of the *ath*-miR159.2\* strand (Rajagopalan et al. 2006). To investigate whether this phenomenon was also present in other plant species, the miR159.2 accumulation pattern was obtained from *N. benthamiana* using as probe an oligonucleotide specific for *Solanum lycopersicum* miR159.2, a related species for which we could identify the pre-miR159a sequence. A low level of *nbe*-miR159.2 could be detected in leaf and root but not in stems or flowers (Fig. 4a). In addition, we performed northern blot on RNA samples from *Oryza sativa* and *Triticum aestivum* seedlings, which according to our analysis share the same sequence for miR159.2 (Fig. 2b). While we could detect *osa*-miR159.2, *tae*-miR159.2 was not detected (Fig. 4a). At

this point we cannot rule out that *tae*-miR159.2 is expressed in other tissues or under different growth conditions. Northern blot experiments and high-throughput sequencing of small RNAs have also shown the expression of *osa*-miR159.2 in rice (Lacombe et al. 2008; Nobuta et al. 2007). In sum, these results indicate that miR159.2 accumulates to detectable levels in different plants and thus it could be functional in at least those species.

Differential expression patterns of miR159 and miR159.2

We observed in *N. benthamiana* that the accumulation of miR159.2 did not reflect that of miR159a in the different organs examined (Fig. 4a). To analyze the relationship between expression of miR159a and miR159.2 in *P. vulgaris*, we performed northern blot analysis on RNA samples

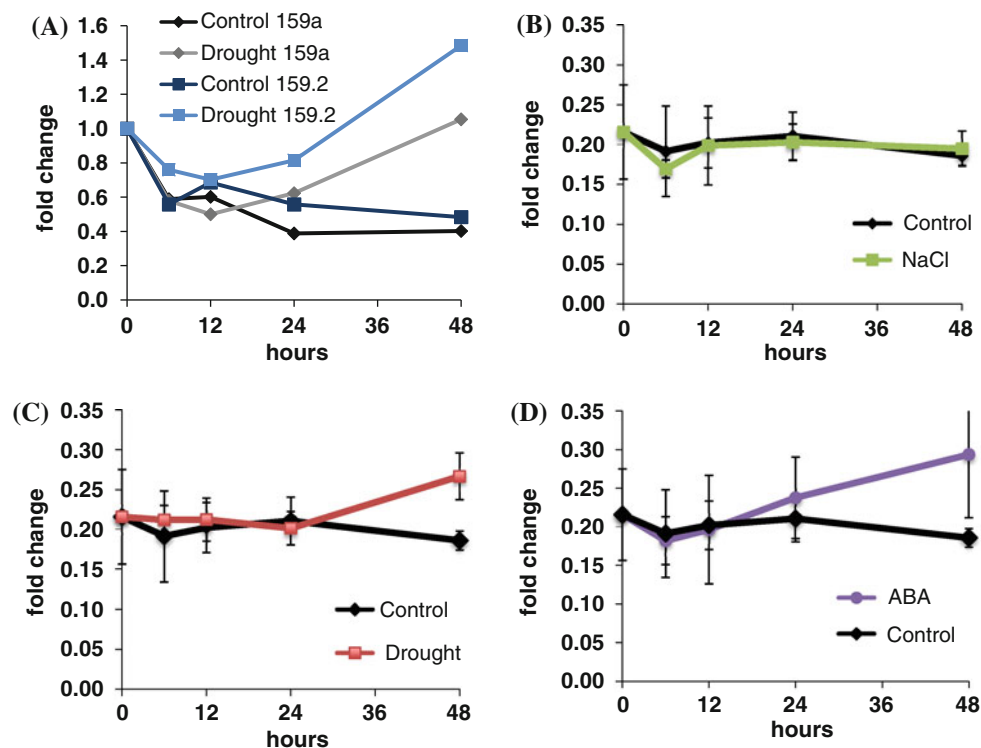


**Fig. 4** miR159.2 is expressed under different conditions. **a** Total RNA (20 µg) from *N. benthamiana* flower buds (F), leaves (L), roots (R) or stems (S) was analyzed by northern blot to detect *Nb*-miR159.2 using as probe an DNA oligonucleotide complementary to the proposed *S. lycopersicum* miR159.2. RNA samples from *Oryza sativa* (Os, root), *Triticum aestivum* (Ta, seedlings) were analyzed by northern blot to detect miR159.2 expression using probes complementary to the sequences for miR159.2 shown in Fig. 2b. The same blots were reprobed to detect miR159a and U6 snRNA as loading controls. **b** Total RNA (10 µg) from root (R), leaves (L), immature embryo (IE), or dry embryo (DE) from *P. vulgaris* or seedlings from *A. thaliana* (A.t.) or *O. sativa* (O.s.) was analyzed by northern blot to detect *pvu*-miR159.2 using as probe a DNA oligonucleotide complementary to *pvu*-miR159.2. The same blots were reprobed to detect miR159.2\*, miR159a and U6 snRNA as loading control

obtained from different organs. Figure 4b shows that miR159a could be detected in all samples tested, including *Arabidopsis* and *Oryza* seedlings (Fig. 4b). In contrast, miR159.2 could only be detected in *P. vulgaris*, where it showed higher abundance in leaves (L in Fig. 4b, lane 2). The small RNA corresponding to miR159.2\* could only be detected in leaf RNA at very low levels (lane 2, Fig. 4b). We have previously reported the accumulation of miR159.2 in response to NaCl, drought and ABA treatments (Arenas-Huerta et al. 2009) as well as for miR159 in common bean seedlings (Reyes et al. 2010). Under those conditions, we observed unique accumulation patterns for both small RNAs.

Because miR159.2 is induced by stress, we further characterized the abundance of miR159 and miR159.2 in a time course of seedling response to NaCl, ABA or drought treatments. Four-day old seedlings were transferred to soil and treated by irrigating with water to full field-capacity (control samples), 1/12 of that amount of water (drought), 200 mM NaCl or 100  $\mu$ M ABA for up to 2 days. Under control growth conditions the levels of miR159 and miR159.2 showed a slight decrease during the course of the experiment, however we detected an induction for both small RNAs upon drought conditions (Fig. 5a). After

2 days of treatment, miR159.2 showed a larger increase in accumulation than miR159a. To establish whether any of the two miRNAs were differentially accumulating with respect to each other, we determined the starting level of each miRNA and defined it as an expression ratio (miR159.2/miR159). Any variations in this ratio from the starting value (set at 1 for the 0 h time point) would indicate that one or the other miRNA is being differentially accumulated during the course of the experiment. In this way, we determined that the ratio of the levels of miR159 and miR159.2 did not significantly change under control conditions in three independent experiments (Fig. 5b–d, black line). In a similar way, NaCl treatment did not affect the miR159.2/miR159 ratio (Fig. 5b). By contrast, ABA and drought treatments showed an increase in the accumulation ratio for miR159.2/miR159, indicating that there is more relative accumulation for miR159.2 than for miR159 at 48 h of treatment (Fig. 5c, d). These results indicate that the accumulation of miR159.2 is not completely dependent on that of miR159 under particular growth conditions, and further that miR159.2 and/or miR159 maybe responding to specific factors regulating its abundance and potential function.



**Fig. 5** miR159a and miR159.2 are differentially expressed in *P. vulgaris* seedlings under stress conditions. Four-day old seedlings were subjected to standard growth conditions (control), irrigated with 1/12 of the standard amount of water (Drought), 200 mM NaCl (NaCl) or 100  $\mu$ M ABA (ABA). Seedlings (three per time point) were collected at 6, 12, 24 or 48 h and RNA isolated and analyzed by northern blotting to detect miR159a or miR159.2. **a** Accumulation of

miR159a (diamonds) and 159.2 (squares) under control (dark color) or drought (light color) conditions. Values were set to 1.0 for the starting time point and normalized using U6 snRNA as loading control. **b–d** Accumulation of miR159a and 159.2 expressed as the ratio miR159a/miR159.2. Values correspond to the average of three independent experiments with standard deviation. **b** NaCl, **c** drought and **d** ABA addition

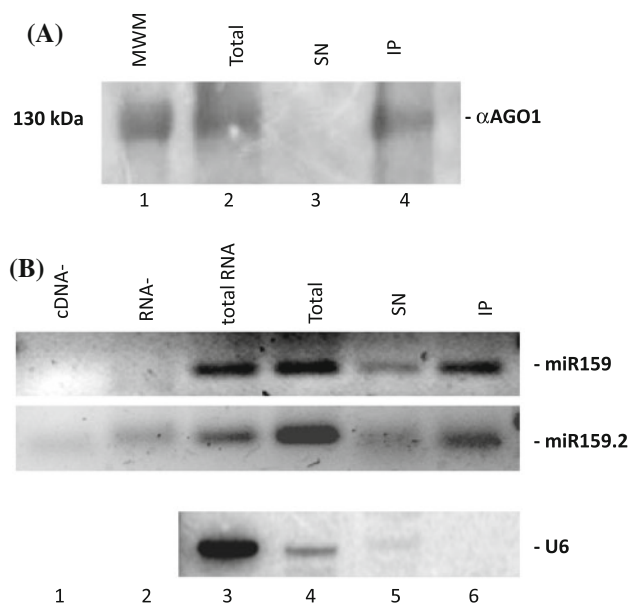
The distinct expression pattern for each of these small RNAs suggests that processing and/or half-life is not directly related to the expression levels of its precursor. Together with the differential expression seen in different organs and plant species, these results suggest that both miRNAs play different roles in different organs and/or growth conditions, possibly regulating distinct targets in response to different situations.

#### miR159.2 is recruited to AGO1-containing complexes

To address the functionality of miR159.2, we investigated whether it is recruited to silencing complexes. For that aim, AGO1 immunoprecipitation assays were performed using an antibody raised against a peptide at the N-terminus of Arabidopsis AGO1 (Qi et al. 2005). The specificity of the antibody was confirmed by a western blot experiment that detected a ca. 110 kDa protein in *P. vulgaris* whole plant extracts, and efficiently immunoprecipitated the same protein from total leaf protein extracts (Fig. 6a). RNA samples obtained from AGO1 co-immunoprecipitated material showed the presence of miR159a but not unrelated RNAs such as U6 snRNA, as determined by stem-looped RT-PCR and northern blot hybridization, respectively (IP, lane 6 in Fig. 6b). In the same co-immunoprecipitated RNA population, we observed the presence of *pvu*-miR159.2, indicating that this miRNA is associated to silencing complexes in vivo (Fig. 6b, middle panel, lane 6). As a negative control, the abundant U6 snRNA was detected by northern blot in whole extract RNA, total and supernatant fractions but not in the immunoprecipitated material (Fig. 6b, bottom panel). Even a prolonged exposure of the U6 snRNA blot failed to detect any signal in the IP RNA sample (data not shown).

#### Cleavage of target mRNAs by miR159.2

To further address the functionality of miR159.2, we have tried to identify regulatory targets by performing 5'RACE assays to search for cleaved mRNAs based on predicted transcripts, including pyruvate dehydrogenase (PDH) (Arenas-Huertero et al. 2009) and ERG, a predicted target using the psRNATarget web application (Dai et al. 2010). Unfortunately, we could not confirm cleavage for any of the suspected transcripts (data not shown). More recently, a global analysis of cleavage products commonly known as degradome analysis was published for soybean (Song et al. 2011). The assay is based on high-throughput sequencing of mRNA tags derived from cleaved transcripts, among which miRNA targets are found. Based on sequence complementarity to *gma*-miR159.2 we predicted potential targets found in this dataset and tested by 5'RACE analysis in common bean. We analyzed 3 candidates: a RAS-related

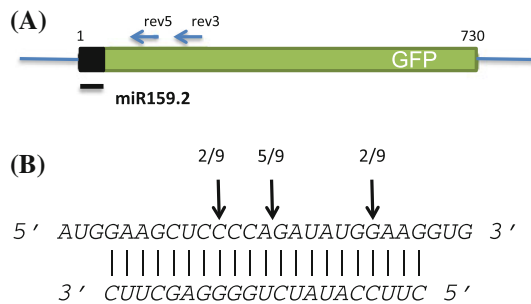


**Fig. 6** miR159.2 is associated with AGO1. **a** Total protein from whole-cell extract (Total) or after antibody incubation and removal (SN, for supernatant) or the immunoprecipitated material (IP) was resolved by PAGE and analyzed by western blotting using the  $\alpha$ -athAGO1 antibody. A single band of approximately 130 kDa was detected. **b** RNA obtained from total extract, supernatant or immunoprecipitated material was analyzed by stem-looped RT-PCR to detect miR159 or miR159.2. The total RNA lane corresponds to an RNA sample (100 ng) obtained from whole leaf tissue. The '- cDNA' and '- RT' lanes correspond to control reactions for the PCR and reverse transcriptase steps, respectively. Northern blot was used to detect U6 snRNA as specificity control for the immunoprecipitation assay. Longer exposure of the blot failed to show any signal for U6 snRNA in the IP lane (not shown)

protein, PLATZ-related transcription factor, and lipoxygenase-related genes, but we could not recover amplification fragments for any of them.

We used an alternative assay to test for potential functionality of *pvu*-miR159.2 where we performed infiltration assays in *N. benthamiana* leaves using a vector construct that included the precursor to *pvu*-miR159a under control of the viral 35S promoter and a GFP reporter gene modified to include a site complementary to miR159.2 at the 5'-end of the coding region (black box in Fig. 7a) under control of the NOS promoter. Two days after infiltration we collected leaves for RNA isolation and 5'RACE amplification. Our results are shown in Fig. 7b. Preferential cleavage was observed at the expected position between nucleotides 10 and 11 of the base-paired region to the miRNA, while other less-abundant cleavage products were recovered as well. While we did not identify the endogenous regulatory target, these results indicate that *pvu*-miR159.2 can be accurately processed, incorporated into silencing complexes and functional in a heterologous system.





**Fig. 7** miR159.2 directs cleavage of a GFP reporter transcript. **a** Diagram of the GFP construct used showing the location of the sequence complementary to miR159.2 (*black box*) and primers used for 5'RACE experiments. **b** Diagram showing miR159.2 base-paired to the GFP construct and location of fragments with the frequency of recovery (indicated by *arrows*)

## Discussion

Our initial characterization of small RNAs present in the legume *Phaseolus vulgaris* led us to the discovery of a second small RNA in the precursor of miR159 identified by cloning and Northern blot analysis (Arenas-Huertero et al. 2009). We observed sequence conservation around the regions encoding miR159a and miR159a\*, where precursor secondary structure and function constrain sequence variation, as previously observed for conserved miRNA precursors. In addition, we observed another two blocks of sequence conservation, one corresponding to miR159.2, and the other sequence located in the opposite side of the precursor. The overall sequence of miR159.2 itself has been conserved during plant diversification with a limited number of changes that still allow comparison of the small RNA sequences (Fig. 2b) and (Axtell et al. 2007; Li et al. 2011). More recently, this observation was further extended to understand the phylogenetic relationships between miR159 and miR319 precursors that confirms the evolutionary conservation of both sequence and position of miR159.2 in the precursor by using a total 231 sequences comprising 71 plant species (Li et al. 2011). Together, these observations suggest that pre-miR159a encodes a second microRNA, displaying conservation across plant species, both in sequence and in position along the stem-loop structure.

Sequential processing of the miR159a precursor results in accumulation of *pvu*-miR159.2

miR159a and miR-159.2 are separated by 21–24 nts in most of the precursors analyzed, with the notable exception of pre-*ath*-miR159a, where the spacer sequence corresponds to 29 nts (see Fig. 2a). This arrangement along the stem-loop structure suggested to us that DCL1 is responsible for producing the second miRNA during maturation of miR159a. Consistently, it has been determined that the

*Arabidopsis* precursor for miR319/159 is processed in a loop-to-stem manner to produce mature miR319/159, generating the equivalent to miR159.2 as a processing intermediate (Addo-Quaye et al. 2009; Bologna et al. 2009). We did confirm this mode of maturation using a 5'RACE approach, however as noted already, processing in *Arabidopsis* does not result in accumulation of miR159.2 while it does in *Phaseolus*. Furthermore, when expressed in *N. benthamiana* leaves, *pvu*-miR159.2 was efficiently processed from the introduced precursor and the cellular machinery in *N. benthamiana* recognized the precursor to accumulate two different miRNAs.

It is intriguing that the precursor in *Arabidopsis* does not accumulate a second miRNA while a similar RNA in legumes does. The distance between miR159a and miR-159.2 is roughly equivalent to 20 or 21 base pairs in most of the pre-miR159a precursors discussed here. Only in *Arabidopsis* the distance is slightly longer (~24 bp), suggesting that this particular length might interfere with miR-159.2 processing and/or accumulation. Alternatively, secondary structure features might influence its processing and/or stability. This hypothesis is consistent with previous published results where point mutations introduced into the related *ath*-pre-miR319a that extended the base-pairing region of the small RNA equivalent to miR159.2 led to its increased accumulation (Addo-Quaye et al. 2009; Bologna et al. 2009). The secondary structure observed in *pvu*-miR159.2 naturally includes more extensive base-pairing in the precursor, suggesting that this feature contributes to its accumulation. Among the conserved families of plant microRNAs, miR159a possesses one of the longest precursors known to date. This property can be accounted for by the presence of a second microRNA. The other long stem-loop precursor is present in the miR169 family, and recent evidence suggests the presence of a second miRNA as well (Zhang et al. 2010). Comparison of different miR159a precursor sequences reveals an extended occurrence of miR159.2 (Li et al. 2011), thus it seems likely that selective pressure to produce two miRNAs has maintained this arrangement during plant diversification and evolution. Different lifestyles have subsequently adopted a very low or non-existent expression of miR159.2 such as in *A. thaliana* or in the other extreme, high levels of expression, as in the case of *P. vulgaris*, suggesting a functional role in this latter species. In fact, detailed analysis of small RNAs derived from miR319/miR159 precursors in diverse species such as soybean, rice, tobacco or *Populus* indicates that under defined conditions small RNAs other than the annotated miRNA can be more abundant, in particular miR159.2, as revealed by high-throughput sequencing experiments (Supplementary Figures 1 and 2; Li et al. 2011).

The possibility of generating multiple small RNAs from this precursor is important to consider because of the

previously reported use of *Arabidopsis* pre-miR159a and the related pre-miR319 as backbones to design artificial miRNAs (Niu et al. 2006; Schwab et al. 2006). Overexpression of amiRNAs from the *pvu*-premiR159 would result in increased accumulation of miR159.2 and its associated activity. This is not the case when using the *ath*-miR159 precursor as our own attempts to detect *ath*-miR159.2 from an artificial construct introduced into *Nicotiana* leaves have failed (data not shown). Thus, use of amiRNAs from *ath*-miR159 would not cause unwanted production of other potentially functional small RNAs, while the *pvu*-miR159a precursor can potentially be used to introduce two distinct miRNAs expressed from the same precursor.

#### Differential expression of miR159 and miR159.2

Because the accumulation of miR159a and miR159.2 does not preserve a direct relationship and we observed distinct accumulation patterns under different growth conditions in *P. vulgaris* and in different *N. benthamiana* organs (Fig. 4), it is possible that an as yet unidentified precursor contributes to small RNA accumulation. However, we think this is unlikely because a Southern blot analysis using the sequence of pre-*pvu*-miR159a as probe showed a single band of hybridization, consistent with a single copy of the gene being present in common bean (data not shown). Consistently, the miR159.2 sequence in the soybean genome is located in only two loci, in agreement with the whole genome duplication that occurred in *Glycine* after it diverged from *Phaseolus* (Schmutz et al. 2010). Thus, we suggest that the differential expression observed could be due to an unequal processing of the precursor or, alternatively, to a selective recruitment to downstream effector complexes. In either case, this could be modulated in response to changing environmental conditions to alter levels of miR159 and miR159.2, accordingly. Such scenario would be consistent with the differential accumulation observed during stress treatments in common bean seedlings (Fig. 5), where particular conditions may require distinct quantities of miR159a and miR159.2 for their individual functions.

miR159.2 is recruited to functional AGO1-containing complexes

RNA:ARGONAUTE protein co-immunoprecipitation experiments previously reported in *Arabidopsis* showed that there is a preference for a U at the 5'-end of miRNAs recruited to AGO1 (Mi et al. 2008). The same study also showed that the few molecules of *ath*-miR159.2 obtained (A at the 5'-end), preferentially associated with AGO1. In common bean, miR159.2 starts with a C, a nucleotide also found in miR169 and miR395 family members, known to

associate with AGO1. It is likely that the identity of the nucleotide at the 5'-end influences the recruitment of the small RNA to different silencing complexes, possibly contributing to the different abundance observed for miR159.2 in *Arabidopsis* compared to legumes. The differential expression observed for miR159.2 and miR159, together with the co-immunoprecipitation of *pvu*-miR159.2 with AGO1 (Figs. 5, 6, respectively), offer the possibility that recruitment to downstream effector complexes could be altered under particular cellular conditions, thus modulating the activity of the two miRNAs originating from the same precursor. It is feasible that this arrangement is more widespread in other plant miRNA precursors or even in animal systems as indicated by recent evidence (Zhang et al. 2010). However, it will have to be established for individual precursors whether the resulting small RNAs correspond to functional miRNAs or processing by-products.

We could not confirm cleavage for any of the predicted transcripts (data not shown). One explanation is that we have not found the target for miR159.2 or, alternatively, that miR159.2 is not functional within silencing complexes. However, using a transient expression system we detected cleavage of a reporter GFP construct using the introduced *pvu*-miR159.2 precursor in *N. benthamiana* leaves. This result indicates that the heterologous system is capable to correctly process the precursor and further to load miR159.2 into functional RISC to target a complementary sequence in GFP. A third possibility is that regulation by miR159.2 occurs by inhibiting translation of the targeted transcript, in which case we would not detect cleaved fragments. The forthcoming availability of genomic and transcriptomic data for *P. vulgaris* will help us to identify the regulatory target for miR159.2.

An appealing concept emerging from this study is that multiple products originating from a single stem-loop structure could have a broader impact in the regulation of gene expression than previously anticipated. Our results strongly suggest that miR159.2 is functional in common bean and further, that its accumulation is at least partially different from that of miR159, the other miRNA encoded in the same precursor, suggesting they may have regulatory functions independent of each other. This arrangement is conserved in different plant species, some of which accumulate miR159.2 to detectable levels, indicating that similar mechanisms may operate in other plants as well.

## Experimental procedures

### Biological materials

Seedlings of *P. vulgaris* (L) var. Negro-Jamapa or Pinto-Villa, *Glycine max*, *Zea mays*, *Oryza sativa*, *Triticum*

*aestivum*, *Arabidopsis thaliana* Col-0, or different organ samples of *Nicotiana benthamiana* were frozen and ground to powder in liquid nitrogen prior to RNA isolation.

#### RNA isolation and northern blotting

Total RNA was prepared using the Trizol reagent (Invitrogen, Carlsbad, CA), according to manufacturer's directions. Total RNA (10–20 µg) was separated in a 15% polyacrylamide/7 M urea/0.5× TBE gel and transferred to Hybond-N+ membrane (Amersham, Piscataway, NJ). Hybridizations were carried out with the Ultra-Hyb Oligo solution (Ambion, Austin, TX) using as probes DNA oligonucleotides labeled with P<sup>32</sup>-γ-ATP at their 5'-end using T4 polynucleotide kinase (Fermentas). Two washes using 2xSSC/0.1%SDS were carried for 30 min at 42°C. Autoradiographs were analyzed using a Typhoon PhosphorImager (Amersham).

#### Sequence alignments

Sequences for miR159a precursors were obtained from miRBase v. 16.0 (Griffiths-Jones 2004), except for *M. truncatula* (gi: 7566673), *A. hypogae* (gi: 115596364) and *S. lycopersicum* (gi: 116645971) that correspond to EST sequences found using the BLAST program at NCBI based on the mature *ath*-miR159a sequence. For these sequences, we confirmed their folding into stable stem-loop structures (data not shown) using the mFold program (Zuker 2003). Pre-*pvu*-miR159a was PCR-amplified from *P. vulgaris* genomic DNA using oligonucleotides based on the *G. max* sequence (gma-MIR159: MI0001773, miRBase (Griffiths-Jones 2004), and cloned into the pENTR-SD-TOPO vector (Invitrogen) to obtain pENTR-*pvu*-miR159a. Sequence alignments were performed using the web interface of T-Coffee ([www.tcoffee.org](http://www.tcoffee.org), Notredame et al. 2000) applying the default parameters.

#### *Agrobacterium* infiltration

The pre-*pvu*-miR159a sequence was transferred from pENTR-*pvu*-miR159a to the pBA-DC plasmid (Zhang et al. 2006) using Gateway technology to obtain pBA-*pvu*-miR159a. This construct was subsequently used to co-express a reporter GFP construct. The BAR gene present in pBA-*pvu*-pre-miR159a was excised from the vector using BglIII and StuI restriction enzymes and replaced by a GFP reporter construct (BglIII-blunt) that contains a sequence complementary to miR159.2 at the 5' end of the coding sequence, introduced using PCR techniques (see Supplemental Table 1 and Fig. 7 for sequences).

The *Agrobacterium tumefaciens* C58 strain carrying the pBA-*pvu*-miR159a plasmid was used to infiltrate *N. benthamiana* leaves. Two days post-infiltration, leaves

were collected and immediately frozen in liquid nitrogen prior to RNA isolation with Trizol reagent as described above.

#### β-elimination

RNA samples or an RNA oligonucleotide were subjected to β-elimination as previously described (Martin et al. 2007). Briefly, the RNA sample (20 µg) was dissolved in 17.5 µl borax buffer, pH 8.6 (4.375 mM borax, 50 mM boric acid) and 2.5 µl of 200 mM NaIO<sub>4</sub> was added. The mixture was incubated for 10 min at room temperature in the dark. Two-µl glycerol was added and the mixture incubated for an additional 10 min at room temperature. After drying under vacuum, samples were dissolved in 50 µl borax buffer pH 9.5 (33.75 mM borax, 50 mM boric acid) and incubated for 90 min at 45°C. The RNA was then purified by ethanol precipitation, and analyzed by polyacrylamide gel electrophoresis and northern blot hybridization.

#### 5'RACE

To identify processing intermediates in the generation of small RNAs we used a modified 5'-end rapid amplification of cDNA ends protocol, using the RLM-RACE kit (Ambion) as indicated by the manufacturer, except that the RNA sample was subjected to adapter ligation without any prior treatment. For PCR amplification of the 5'RACE product we used nested oligonucleotides complementary to the *pvu*-miR159a precursor RNA (see Supplemental Table 1). In the case of the GFP reporter construct, we used oligonucleotides complementary to GFP designed for nested PCR reactions (see Supplemental Table 1 and Fig. 7a).

#### Immunopurification of AGO1 complexes

For immunopurification of AGO1-containing complexes, we used an antibody generated against the N-terminus of Arabidopsis AGO1 (Qi et al. 2005) kindly provided by Yijun Qi (National Institute of Biological Sciences, Beijing, China). The immunoprecipitation assay was performed as previously described (Qi and Mi 2010) with some modifications. For each immunoprecipitation assay 400 mg of N<sub>2</sub>-ground tissue from adult *P. vulgaris* leaves were used. A 1:50 antibody dilution was used and incubation of whole-cell extract with the antibody and protein A-agarose beads (Roche, Indianapolis, IN) was performed overnight at 4°C. RNA from the resulting immunoprecipitated material bound to beads was purified by phenol extraction for further analysis. Alternatively, protein from immunopurified material or total extract was obtained and analyzed by SDS-PAGE and western blotting, using the α-AGO1 antibody at a 1:1000 dilution.

## RT-PCR for microRNA detection

An RT-PCR approach designed to specifically detect miRNA presence was performed as previously described (Varkonyi-Gasic et al. 2007). RNA samples obtained from immunoprecipitated material or from whole tissues were subjected to reverse transcription (RT) using Super Script III reverse transcriptase (Invitrogen) and 1  $\mu$ M final concentration of specific stem-loop primer for miR159a, miR159.2, miR159.2\* and Intermediate (int) sequence (for primer sequence details see Supplemental Table 1). A pulsed RT-step was performed as follows: 16°C for 30 min followed by 60 cycles of 30°C for 30 s, 42°C for 30 s and 50°C for 1 s. Inactivation of reverse transcriptase was carried out at 85°C for 5 min.

For semi-quantitative PCR, 2  $\mu$ l of the obtained cDNA were used for each condition. MicroRNA-specific primer and universal primer (10  $\mu$ M final concentration) were used (for sequence details see Supplemental Table 1). The PCR products after 30 amplification cycles were subjected to electrophoresis on a 3% agarose gel.

**Acknowledgments** We thank members of the Covarrubias group for stimulating discussions throughout the development of this work. CC-C is a PhD student from Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México and a recipient of a studentship from Consejo Nacional de Ciencia y Tecnología, México (203284). This work was supported by grants from DGAPA-UNAM (IN-222509) and CONACyT (J-48740) to JLR.

## References

- 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
- Arenas-Huerta C, Perez B, Rabanal F, Blanco-Melo D, De la Rosa C, Estrada-Navarrete G, Sanchez F, Covarrubias AA, Reyes JL (2009) Conserved and novel miRNAs in the legume *Phaseolus vulgaris* in response to stress. *Plant Mol Biol* 70:385–401
- Axtell MJ, Snyder JA, Bartel DP (2007) Common functions for diverse small RNAs of land plants. *Plant Cell* 19:1750–1769
- Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116:281–297
- Baumberger N, Baulcombe DC (2005) Arabidopsis ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs. *Proc Natl Acad Sci USA* 102:11928–11933
- 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
- Dai X, Zhuang Z, Zhao PX (2010) Computational analysis of miRNA targets in plants: current status and challenges. *Brief Bioinform* 12:115–121
- Dong Z, Han MH, Fedoroff N (2008) The RNA-binding proteins HYL1 and SE promote accurate in vitro processing of pri-miRNA by DCL1. *Proc Natl Acad Sci USA* 105:9970–9975
- Fahlgrén N, Howell MD, Kasschau KD, Chapman EJ, Sullivan CM, Cumbie JS, Givan SA, Law TF, Grant SR, Dangel JL et al (2007) High-throughput sequencing of Arabidopsis microRNAs: evidence for frequent birth and death of MIRNA genes. *PLoS ONE* 2:e219
- Griffiths-Jones S (2004) The microRNA Registry. *Nucleic Acids Research* (32 Database issue), D109–D111
- Kurihara Y, Watanabe Y (2004) Arabidopsis micro-RNA biogenesis through Dicer-like 1 protein functions. *Proc Natl Acad Sci USA* 101:12753–12758
- Lacombe S, Nagasaki H, Santi C, Duval D, Piegu B, Bangratz M, Breitler JC, Guiderdoni E, Brugidou C, Hirsch J et al (2008) Identification of precursor transcripts for 6 novel miRNAs expands the diversity on the genomic organisation and expression of miRNA genes in rice. *BMC Plant Biol* 8:123
- Li Y, Li C, Ding G, Jin Y (2011) Evolution of MIR159/319 microRNA genes and their post-transcriptional regulatory link to siRNA pathways. *BMC Evol Biol* 11:122
- Mallory AC, Reinhart BJ, Bartel D, Vance VB, Bowman LH (2002) A viral suppressor of RNA silencing differentially regulates the accumulation of short interfering RNAs and micro-RNAs in tobacco. *Proc Natl Acad Sci USA* 99:15228–15233
- Martin R, Arenas C, Daros JA, Covarrubias A, Reyes JL, Chua NH (2007) Characterization of small RNAs derived from Citrus exocortis viroid (CEVd) in infected tomato plants. *Virology* 367:135–146
- 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
- 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
- Niu QW, Lin SS, Reyes JL, Chen KC, Wu HW, Yeh SD, Chua NH (2006) Expression of artificial microRNAs in transgenic *Arabidopsis thaliana* confers virus resistance. *Nat Biotechnol* 24:1420–1428
- Nobuta K, Venu RC, Lu C, Belo A, Vemaraju K, Kulkarni K, Wang W, Pillay M, Green PJ, Wang GL et al (2007) An expression atlas of rice mRNAs and small RNAs. *Nat Biotechnol* 25:473–477
- Notredame C, Higgins DG, Heringa J (2000) T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J Mol Biol* 302:205–217
- Qi Y, Mi S (2010) Purification of *Arabidopsis argonaute* complexes and associated small RNAs. *Methods Mol Biol* 592:243–254
- Qi Y, Denli AM, Hannon GJ (2005) Biochemical specialization within Arabidopsis RNA silencing pathways. *Mol Cell* 19:421–428
- 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
- Reyes JL, Arenas-Huerta C, Sunkar R (2010) Cloning of stress-responsive microRNAs and other small RNAs from plants. *Methods Mol Biol* 639:239–251
- Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, Bartel DP (2002) Prediction of plant microRNA targets. *Cell* 110:513–520
- Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J et al (2010) Genome sequence of the palaeopolyploid soybean. *Nature* 463:178–183
- Schwab R, Ossowski S, Riester M, Warthmann N, Weigel D (2006) Highly specific gene silencing by artificial microRNAs in Arabidopsis. *Plant Cell* 18:1121–1133

- Song QX, Liu YF, Hu XY, Zhang WK, Ma B, Chen SY, Zhang JS (2011) Identification of miRNAs and their target genes in developing soybean seeds by deep sequencing. *BMC Plant Biol* 11:5
- Varkonyi-Gasic E, Wu R, Wood M, Walton EF, Hellens RP (2007) Protocol: a highly sensitive RT-PCR method for detection and quantification of microRNAs. *Plant Methods* 3:12
- Vaucheret H (2008) Plant ARGONAUTES. *Trends Plant Sci* 13:350–358
- Voinnet O (2009) Origin, biogenesis, and activity of plant microRNAs. *Cell* 136:669–687
- Yu B, Yang Z, Li J, Minakhina S, Yang M, Padgett RW, Steward R, Chen X (2005) Methylation as a crucial step in plant microRNA biogenesis. *Science* 307:932–935
- Zhang X, Yuan YR, Pei Y, Lin SS, Tuschl T, Patel DJ, Chua NH (2006) Cucumber mosaic virus-encoded 2b suppressor inhibits *Arabidopsis Argonaute1* cleavage activity to counter plant defense. *Genes Dev* 20:3255–3268
- 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
- Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31:3406–3415