Characterization of plant microRNA-encoded peptides (miPEPs) reveals molecular mechanisms from the translation to activity and specificity

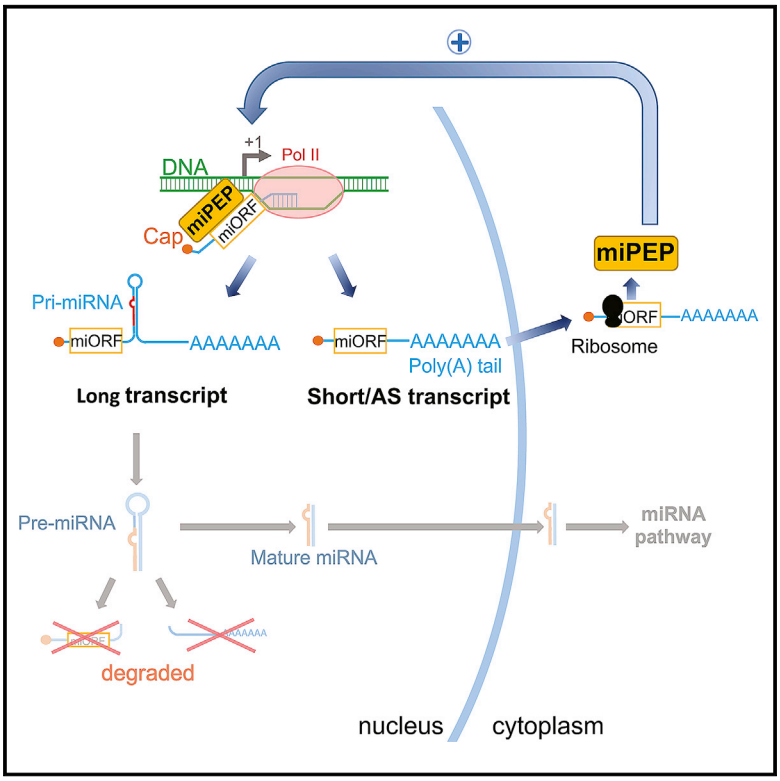
Dominique Lauressergues, Mélanie Ormancey, Bruno Guillotin, Hélène San Clemente, Laurent Camborde, Carine Duboé, Sabine Tourneur, Pierre Charpentier, Amélie Barozet, Alain Jauneau, et al.

Cell Reports

Article

Characterization of plant microRNA-encoded peptides (miPEPs) reveals molecular mechanisms from the translation to activity and specificity

Graphical abstract



Highlights

- Plant primary transcripts of miRNAs are processed in populations of transcripts
- Short transcripts and splicing variants could be translated into miPEPs
- MiPEPs specifically increase expression of their nascent pri-miRNA
- Specificity of miPEPs relies on their own encoding sequence ORF

Authors

Dominique Lauressergues, Mélanie Ormancey, Bruno Guillotin, ..., Virginie Gervais, Serge Plaza, Jean-Philippe Combier

Correspondence

combier@lrsv.ups-tlse.fr

In brief

MicroRNA-encoded peptides (miPEPs) increase transcription of their nascent pri-miRNA. Lauressergues et al. show that pri-miRNAs are transcribed in different transcripts with different fates: maturation in miRNAs or translation in miPEPs. We reveal that specificity of miPEPs for their pri-miRNA relies on the presence of their encoding ORF in the pri-miRNA.

Lauressergues et al., 2022, Cell Reports *38*, 110339
February 8, 2022
https://doi.org/10.1016/j.celrep.2022.110339

CellPress

Cell Reports

CellPress
OPEN ACCESS

Article

Characterization of plant microRNA-encoded peptides (miPEPs) reveals molecular mechanisms from the translation to activity and specificity

Dominique Lauressergues,[1,5] Mélanie Ormancey,[1,5] Bruno Guillotin,[1,5] Hélène San Clemente,[1] Laurent Camborde,[1] Carine Duboé,[1] Sabine Tourneur,[1] Pierre Charpentier,[2] Amélie Barozet,[2] Alain Jauneau,[3] Aurélie Le Ru,[3] Patrice Thuleau,[1] Virginie Gervais,[4] Serge Plaza,[1] and Jean-Philippe Combier[1,6,*]

[1]Laboratoire de Recherche en Sciences Végétales, CNRS/UPS/INP, 31320 Auzeville-Tolosane, France
[2]Micropep Technologies, 31320 Auzeville Tolosane, France
[3]Plateforme Imagerie TRI-FRAIB, CNRS/UPS, 31320 Auzeville-Tolosane, France
[4]Institut de Pharmacologie et de Biologie Structurale, CNRS/UPS, 31077 Toulouse, France
[5]These authors contributed equally
[6]Lead contact
*Correspondence: combier@lrsv.ups-tlse.fr
https://doi.org/10.1016/j.celrep.2022.110339

SUMMARY

MicroRNAs (miRNAs) are transcribed as long primary transcripts (pri-miRNAs) by RNA polymerase II. Plant pri-miRNAs encode regulatory peptides called miPEPs, which specifically enhance the transcription of the pri-miRNA from which they originate. However, paradoxically, whereas miPEPs have been identified in different plant species, they are poorly conserved, raising the question of the mechanisms underlying their specificity. To address this point, we identify and re-annotate multiple *Arabidopsis thaliana* pri-miRNAs in order to identify ORF encoding miPEPs. The study of several identified miPEPs in different species show that non-conserved miPEPs are only active in their plant of origin, whereas conserved ones are active in different species. Finally, we find that miPEP activity relies on the presence of its own miORF, explaining both the lack of selection pressure on miPEP sequence and the ability for non-conserved peptides to play a similar role, i.e., to activate the expression of their corresponding miRNA.

INTRODUCTION

MicroRNAs (miRNAs) are small RNA molecules (20–22 nt) regulating the expression of target genes at the post-transcriptional level, by inhibiting their translation or cleaving their mRNA. Because these target genes encode mainly regulatory proteins, such as transcription factors or hormone receptors, miRNAs are involved at the cross-roads of multiple biological processes in plants, such as development, reproduction, and stress responses. The miRNAs are transcribed as long primary transcripts (pri-miRNAs), capped and polyadenylated like protein coding genes (Xie et al., 2005). Pioneering work has revealed that plant pri-miRNAs encode regulatory peptides called miPEPs (Lauressergues et al., 2015), a finding recently further extended to other miRNAs (Sharma et al., 2020; Chen et al., 2020). Until now, miPEPs have been shown to be encoded by the first open reading frame (miORF) after the transcription start site (TSS) located in the 5′ region of the pri-miRNA and they enhance pri-miRNA transcription from which they originate (Lauressergues et al., 2015; Sharma et al., 2020; Chen et al., 2020). Initially characterized in *Arabidopsis thaliana*, *Medicago truncatula*, and *Glycine max* (Lauressergues et al., 2015; Couzigou et al., 2016, 2017), recent studies revealed the presence of miPEPs in other plant species (Chen et al., 2020; Zhang et al., 2020; Ormancey et al., 2021) and even in animals (Kang et al., 2020; Niu et al., 2020; Prel et al., 2021; Montigny et al., 2021; Immarigeon et al., 2021). This suggests that the number of miPEPs is probably far to be fully determined, and that miPEPs might be a common feature in both plant and animal kingdoms. In plants, miPEPs appear to be very specific as they up-regulate their associated miRNA, without disrupting the expression of other miRNAs even within the same miRNA family (Lauressergues et al., 2015; Sharma et al., 2020). Consequently, these findings raised several questions, including the understanding of the underlying mechanisms of miPEP specificity.

We show here that pri-miRNAs are in fact transcribed as populations of long canonical transcripts, but also of short transcripts that do not contain the pre-miRNA stem loop sequence and are mainly located in the cytoplasm. We provide evidence that most of pri-miRNAs are probably translated, as most of the pri-miRNAs tested were found to be associated to RPL18, a protein component of ribosomes. We further confirm that, for several pri-miRNAs, the first ORF after the TSS is translated and encodes a functional miPEP. Consistently, all the tested miPEPs derived from the first ORF were able to increase their pri-miRNA expression. In parallel, we show that non-conserved

miPEPs are only active on their pri-miRNA in the species from which they originate, while conserved miPEPs, i.e. those with less than 10% mutations between them, can be active across different plants species. Finally, we show that miPEP specificity relies on a physical interaction between the miPEP and its miORF.

RESULTS

Annotations of *Arabidopsis* pri-miRNAs

Before studying the conservation of miPEPs between species, the first step consisted in the identification of *A. thaliana* miPEPs. The study (and prediction) of miPEPs first required the identification of primary transcripts of miRNAs. Whereas the pre-miRNAs are well referenced (mirbase.org), the pri-miRNAs of *A. thaliana* are far from being fully characterized. To achieve the annotation of conserved *A. thaliana* pri-miRNAs, we crossed the data from EST sequences (www.ncbi.nlm.nih.gov/genbank/), Illumina RNA-seq data (Wang et al., 2019), rapid amplification of cDNA ends (RACE) PCR data (Xie et al., 2005), PacBio Iso-Seq data (this study), and results from our 3′ RACE PCR experiments. Corresponding sequences have been deposited (GenBank: MW_775349-MW_775424). All these approaches allowed us to detect 167 pre-miRNAs within the 326 *A. thaliana* miRNAs present in miRbase. Of these 167 pre-miRNAs, 70 were located in coding sequence (22), introns (23), 5′UTR (12), or 3′UTR (13) of coding genes (Figure 1A). In parallel, we detected two polycistronic pri-miRNAs gathering five miRNAs (miR397b and miR857 on the one hand; and miR850, miR863, and miR5026 on the other hand). The majority of the pre-miRNAs (92) are located in intergenic regions. We, therefore, focused our analysis on these 92 intergenic pri-miRNAs (and removed the polycistronic pri-miRNAs from our analysis) (Table S1), as miPEPs were initially identified from intergenic pri-miRNAs (Lauressergues et al., 2015). Interestingly, we found that the average size of pri-miRNAs was larger than previously believed, with an average length of 1,277 bases (Figure 1B). Whereas some pri-miRNAs were very short (206 bases for pri-miR865), some of them were much longer (6,026 bases for pri-miR5647). Moreover, the 5′ arm median length (363 bases) of pri-miRNAs was longer than the median length of *A. thaliana* 5′ UTR of coding genes (152 bases) (Figure 1C). In addition, the number of ORFs per kb in the different pri-miRNAs was similar even slightly higher than those found in coding genes or long non coding RNAs (lncRNAs) (Figure 1D). We found an average of 4.1 ORFs encoding more than five amino acids in length in the 5′ arm of these pri-miRNAs (Figure 1E). In 84 of the 93 pri-miRNAs of *A. thaliana* we annotated, we identified at least one ORF in the 5′ arm (Table S2). Since different studies provided experimental evidence that the first ORF after the TSS is translated and produces a miPEP (Lauressergues et al., 2015; Chen et al., 2020; Sharma et al., 2020), we focused our analyses on the first ORF detected in pri-miRNAs. The deduced peptides had an average length of 24 amino acids (molecular weight of 2.3 kDa) (Figure 1F). The average isoelectric point (pI) of these peptides was around 7.7, which was similar to the mean pI of *A. thaliana* proteins (approximately 7.5) (Figure 1G). Finally, it was not possible to identify an ORF in the 5′ arm of a few miRNAs (miR158a, 166c, 390a, 398c, 399c, 408, 843, 865), probably because of their short length (151, 141, 143, 72, 74, 129, 108, and 113 bases, respectively) (Table S1). In summary, our data suggest that most of the intergenic pri-miRNAs potentially encode for miPEPs.

Primary transcripts of miRNAs are processed in several transcripts

In parallel with the identification of pri-miRNAs, analysis of Iso-Seq and RACE-PCR data allowed us to demonstrate the transcription of a heterogeneous population of transcripts for almost all pri-miRNAs (Figures 2A and 2B). Indeed, we detected three types of transcripts. The first type corresponded with short transcripts located in the 5′ arm and excluding the miRNA or miRNA* sequence, such as miR160a (Figure 2A). The second group revealed the existence of alternatively spliced (AS) transcripts excluding miRNA or miRNA* sequence, such as miR162a, producing pri-miRNA transcript variants probably unable to produce stem-loop structures (Figure 2A). The third group corresponded to other, long transcripts, containing the full-length sequence with the entire pre-miRNA, such as miR156a, and for which our analysis did not retrieve any short or AS transcript without the miRNA sequence and having miPEP-encoding ORFs located 5′ to the pre-miRNA (Figure 2A). To go further, we performed an analysis by 5′ and 3′ RACE PCR of nine pri-miRNAs presenting short transcripts in Iso-Seq data. Thus, we found that some of them (miR160a, miR172a) potentially possessed two TSS, confirming a previous study (Xie et al., 2005), while others had only one TSS. In addition, we identified several transcripts that did not contain pre-miRNA sequence or with truncated pre-miRNA (Figure 2B). These data suggest that pri-miRNAs are transcribed either as long full-length sequences containing all the information (miRNA, miPEP sequence) or shorter or AS transcripts containing only the miPEP sequence, but without a stem loop. We can hypothesize that these different populations of transcripts have different fates: short (and possibly AS) transcripts without complete pre-miRNA could translocate to cytoplasm, be loaded by ribosomes and translated into miPEPs, whereas long transcripts, containing stem-loop structure of the pre-miRNA could be processed by DCL1-containing complex in the nucleus to finally give the mature miRNA.

To investigate this hypothesis, we performed an immunoprecipitation (IP) of the 60S ribosomal protein L18 (RPL18) (Juntawong et al., 2014) followed by a 3′ RACE PCR to identify transcripts that would be predominantly loaded into ribosomes, i.e., likely translated (Bazin et al., 2017). In six out nine cases (miR159a, miR160a, miR166b, miR171a, miR172a, and miR319b), we identified short or AS transcripts (Figures 2A and 2B). Therefore, these data indicate that short and AS transcripts that contain the ORF encoding miPEP but not the stem loop are loaded into ribosomes. In contrast, by sequencing products of RACE PCR analysis for miR156a, miR165a, and miR167a, we were able to identify only full-length transcripts associated with ribosomes, suggesting that long pri-miRNAs may be bound in ribosomes in some cases.

To better estimate the ratio of long transcripts loaded into ribosomes, we compared the expression level, relative to UBP6, of long transcripts isolated from RPL18 IP samples with that of

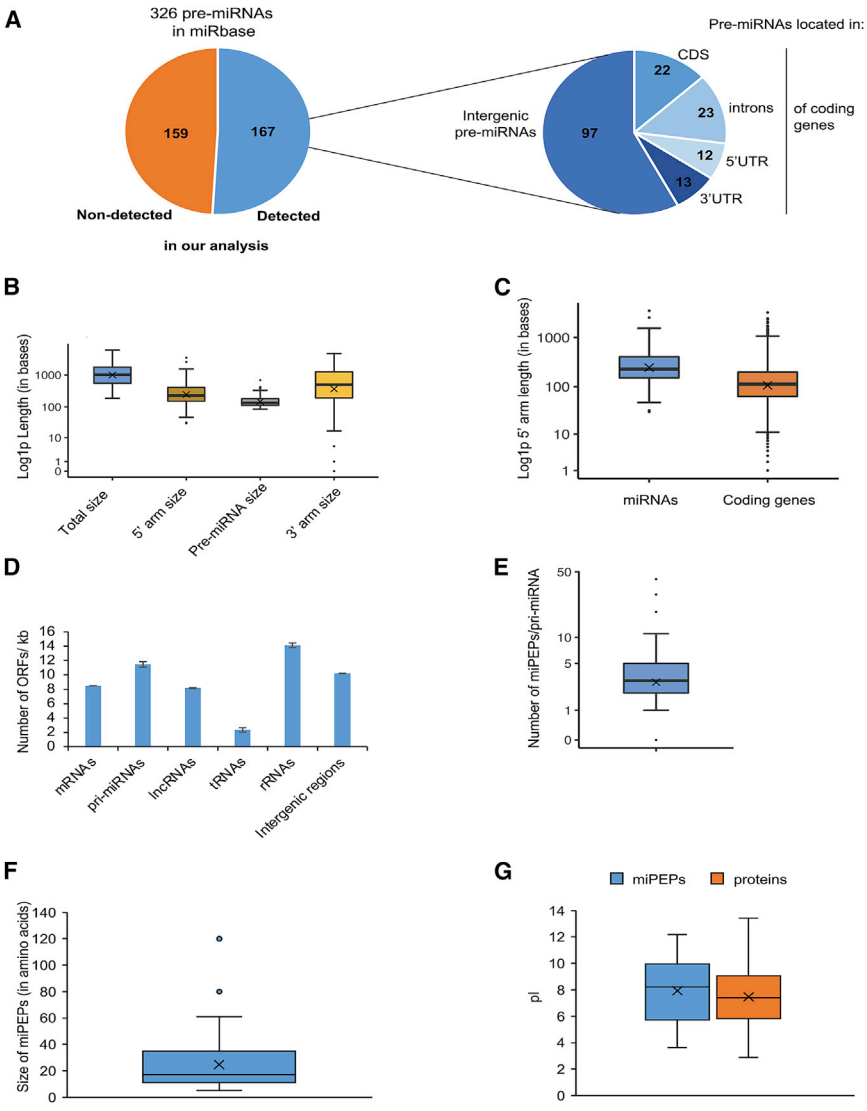


Figure 1. Characterization of pri-miRNAs of *A. thaliana*

(A) Distribution and characterization of pre-miRNAs identified in our analysis. Left panel, quantity of pre-miRNAs detected in our analysis. Right panel, localization of pre-miRNAs identified in our analysis.

(B) Mean length of the different elements forming the 93 intergenic pri-miRNAs.

(C) Size distribution of coding gene 5′ UTRs compared with intergenic pri-miRNA 5′ arms.

(D) Number of ORFs per kb in different RNA types.

(E) Distribution of the number of miPEPs identified in the 5′ arm of each intergenic pri-miRNA.

(F) MiPEP size distribution.

(G) pI distribution of miPEPs compared with that of total proteins of *A. thaliana*.

long transcripts isolated from total RNAs. To do this, we performed a qRT-PCR by amplifying the 3′ part of the pri-miRNAs which contain the stem loop (miRNA) (arrows in Figures 2A and 2B). Interestingly, for all seven tested pri-miRNAs, the data showed an under-representation of long transcripts in RPL18 IP samples compared with total RNAs (Figures 2C and S1A). In parallel, we extracted nuclei RNAs and performed the same quantification of long transcripts. All seven tested long pri-miRNAs (containing the stem loop) were enriched in nuclei (Figures 2C and S1B). These experiments reveal that a fraction of long transcripts are able to produce miPEPs and suggest that the long transcripts are enriched in nuclei (to give miRNA after maturation). Thus, we might hypothesize that short transcripts, associated to RPL18 and located in the cytoplasm, likely constitute the main source of miPEPs. To investigate the putative translatability of these short transcripts, we performed an *in vitro* transcription/translation in wheat germ extracts of three different transcripts containing HA-tagged miPEPs putatively expressed from the first ORF. Western blot analysis revealed that miPEP translation could be detected from these three short transcripts (Figure S2A).

To know whether our findings reflect a common feature of plant pri-miRNAs, we performed a RACE-PCR analysis on the previously well studied *M. truncatula* pri-miR171b (Lauressergues et al., 2015; Couzigou et al., 2017). We identified only one TSS and at least three different transcription termination sites (TTS) (Figure S2B). Whereas the two longer transcripts contained the full-length pre-miRNA, the shorter transcript held only the sequence corresponding to the miPEP as observed in *A. thaliana*. By transient expression of the *M. truncatula* pri-miR171b in tobacco leaves, in which the correct maturation of pri-miR171b occurs to form mature miR171b (Lauressergues et al., 2015), we confirmed the presence of different TTS (Figure S2B), with long transcripts containing the pre-miRNA and a short transcript holding the miPEP sequence. As mentioned above for *A. thaliana* short transcripts, *in vitro* transcription/translation of the *M. truncatula* pri-miR171b short transcript containing the HA-tagged miPEP showed the production of miPEP by this transcript (Figure S2C).

Finally, we wanted to investigate the fate of short and AS transcripts (producing miPEPs) by analyzing their expression in mutants of the dicing complex. Unfortunately, it was impossible to segregate short from long transcripts by qPCR, since they totally overlap. So we focused our analysis on AS transcripts. Using primers located in a region common to all miR157a and miR162a transcripts (FL, Figure 2A) and primers overlapping an intron (AS, Figure 2A), we were able to discriminate long transcripts and splice variants to quantify them specifically. As expected, both miR157a and miR162a long transcripts (FL), containing the stem-loop, were over accumulated in *hyl1* mutant, impaired in maturation of pri-miRNAs (Han et al., 2004)

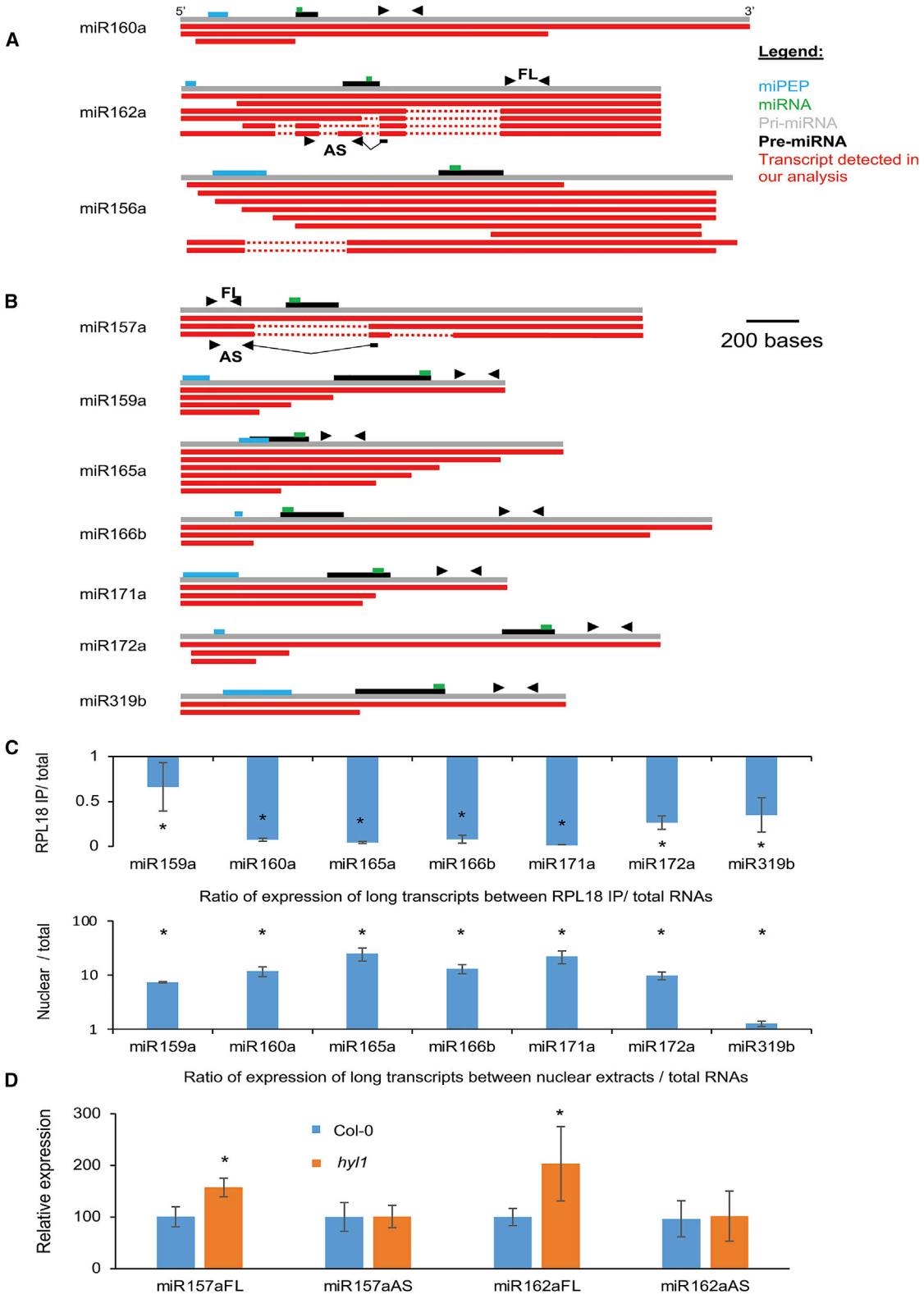


Figure 2. Primary transcripts of miRNAs are transcribed as different transcripts with different fates

(A) Examples of populations of transcripts corresponding to three representative pri-miRNAs: miR160a, for which short transcripts have been detected; miR162a, for which some AS transcripts without miRNA have been detected; and miR156a, which exemplifies other miRNAs. Color code: gray, longest transcript identified;

(legend continued on next page)

(Figure 2D). Interestingly, in both cases, the expression of AS transcripts, which do not contain the stem loop, were not affected in *hyl1* mutant, suggesting that these transcripts are not loaded by DCL1 complexes (Figure 2D).

Together, these data reveal that several transcripts are produced from a single miRNA gene, generating various classes of pri-miRNAs with different fates such as maturation into mature miRNAs or translation into miPEPs.

Expression and activity of miPEPs

Whereas a few plant peptides could be identified by MS (Wang et al., 2020), up to now MS peptide detection remains challenging. Alternatively, specific antibodies gave evidence for the existence of miPEPs (Lauressergues et al., 2015; Sharma et al., 2020). Interestingly, the detection of miPEPs was correlated with a favorable translational context of the ATG codon as revealed by the detection of GUS fusion activities, suggesting that this strategy is a good alternative approach to validate miORF translatability (Lauressergues et al., 2015; Sharma et al., 2020; Chen et al., 2020). Moreover, in each case, the first ORF was translatable, while the other ORFs were not. In this line, we generated constructs containing a GUS fusion including 2 kb of the upstream promoter region until the first ATG of miPEP for nine different pri-miRNAs randomly chosen. For all of them, GUS activity was detected in at least one tissue or part of seedlings (Figure 3A). These data showed that the first ATG of the nine pri-miRNAs was capable to initiate translation, suggesting that most of pri-miRNAs encode for miPEPs (Figure 3A).

We then randomly chose 8 of the 92 previously mentioned miPEPs (see above and Table S1), to test peptides of different lengths (9–33 amino acids) and different physical and chemical properties (hydrophilic, amphiphilic, hydrophobic; pI between 4.9 and 11.3) and we chemically synthesized them. We treated *A. thaliana* seedlings or adult plants with these miPEPs independently and harvested plants 24 h later. An up-regulation of the corresponding pri-miRNA was detected by qRT-PCR in all cases tested (8/8) (Figures 3B and 3C). We next treated *A. thaliana* seedlings with miPEP162a and miPEP396a separately and harvested them 3 h later. We were still able to detect an up-regulation of the corresponding mature miRNA (Figure 3C) and pri-miRNA (Figure 3D). In parallel, we tested the expression of eight different pri-miRNAs and showed that the expression of none of them was significantly affected by the treatment with one of the two miPEPs (Figure 3D).

Conservation and activity of miPEPs

After having characterized the pri-miRNAs from *A. thaliana*, and for some of them their potency to produce a miPEP, we studied their conservation. Indeed, although Lauressergues et al. (2015) and Morozov et al. (2019) showed a conservation of miPEP165a and miPEP156a in *Brassicaceae*, no miPEP with sequence similarity have been identified in other plant families. Whereas Couzigou et al. (2017) identified functional homologs of miPEP171b in four distant species (*M. truncatula*, *Lotus japonicus*, *Oryza sativa*, and *Solanum lycopersicum*), surprisingly, these functional homologs share no sequence conservation. A sequence analysis of the whole set of miPEPs showed no conservation within *Brassicaceae*, except for miPEP156a and miPEP165a mentioned above, and miPEP164a (Figure 4A). We first focused on the conservation of miPEPs in two related species: *A. thaliana* and *Arabidopsis lyrata*. To achieve this, we annotated *A. lyrata* pri-miRNAs from available databases and from sequence similarities with *A. thaliana* pri-miRNAs. We then compared the conservation of nucleic acid sequences between the two species for different types of sequences: coding sequences, intergenic sequences (before and after the pri-miRNA on the genome), pri- and pre-miRNAs, and miPEP-encoding sequences (miORFs) (Figure 4B). While coding and pre-miRNA sequences were strongly conserved between the two species, intergenic regions revealed a weak conservation of about 50% between the two species. Interestingly, the miORF sequences showed a strong variability between the two species, with a median conservation comparable with that of intergenic regions, suggesting a neutral selection pressure on these sequences (Figure 4B). In addition, a comparison of the protein sequences conservation between the whole set of proteins and miPEPs indicated that the latter were overall less conserved (Figure 4C). Indeed, some miPEPs were identical between the two species, while others were very poorly conserved (Figure 4D). Finally, despite this lack of sequence conservation, the number of ORFs in pri-miRNA 5′ arms was remarkably constant between the two species (Figure 4E). Taken together, these results suggest that the presence of miORFs (number of ORFs/kb) is maintained, whereas the conservation of miPEP sequence is not.

Based on these results, we wondered whether miPEPs from one species could be active on pri-miRNAs from other species. For this, we have considered two miPEPs with two different profiles of conservation within *Brassicaceae*, i.e., a strongly conserved miPEP (miPEP156a) and a poorly conserved miPEP (miPEP167a) (Figure 4F). Indeed, the 5′ arm of pri-miR156a showed 81%–97% sequence identity between *A. thaliana*, *Brassica oleracea* (cabbage) and *Brassica rapa* (turnip), while the identity for pri-miR167a was between 51% and 71% within the three plants. Similarly, whereas miPEP167a (10 amino acids) shared no sequence similarity between these three species,

blue, miPEP sequence; green, miRNA; black, pre-miRNA; red, types of transcripts identified in our analyses. Arrowheads indicate the position of primers used in (C and D). FL, full length.

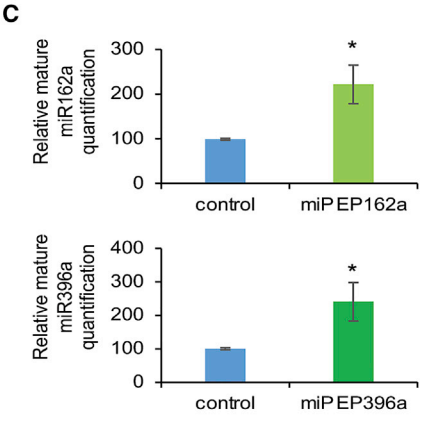
(B) Types of transcripts of pri-miRNAs showing short transcripts, as revealed by our analysis, combining RACE PCR 5′ and 3′, RNAseq data, and Iso-seq sequencing.

(C) Top: ratio of expression of long transcripts of pri-miRNAs in RPL18 IP samples versus total RNAs. Bottom: ratio of expression of long transcripts of pri-miRNAs in nuclear extract samples versus total RNAs.

(D) Relative expression of FL and AS transcripts of pri-miR157a and pri-miR162a in WT Col-0 line compared with expression in *hyl1* mutant line. Error bars represent standard errors of the means, asterisks indicate a significant difference between (C) expression in total RNAs and expression in RPL18 IP or nuclear extracts, or (D) the test condition (*hyl1*) and the control (Col-0) according to the Wilcoxon test ($n = 8$, $p < 0.05$).

B

At pri-miRNA	Induction of expression by corresponding miPEP (± SEM; * p-value<0,05)
miR156a	1.92* ± 0.47
miR162a	2.06* ± 0.42
miR162b	2.58* ± 0.32
miR163	2.05* ± 0.37
miR167a	1.58* ± 0.04
miR169l	3.03* ± 1.11
miR172b	1.99* ± 0.36
miR396a	2.03* ± 0.29



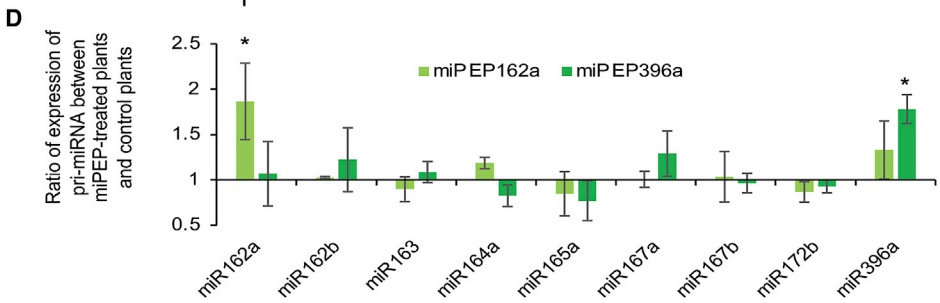


Figure 3. Expression and activity of miPEPs

(A) Expression pattern in *Arabidopsis* seedlings of several putative start codons of different miPEPs, using fusions between GUS and promoter region until the first ATG of the indicated miRNAs. Control, Col-0.

(B) Quantification by qRT-PCR of expression of several pri-miRNAs after 24 h treatment by 10 μM of the corresponding miPEP.

(legend continued on next page)

miPEP156a (33 amino acids) showed around 90% of identity (Figure 4F). We next treated *A. thaliana*, *B. oleracea*, and *B. rapa* seedlings with miPEP156a and miPEP167a from each species and quantified the expression of the corresponding pri-miRNA. This allowed us to detect an up-regulation of pri-miR156a in all species treated with the different miPEP156a peptides, showing that the 10% mismatches between these peptides are not enough to suppress their activity (Figure 4G). Conversely, no induction of pri-miR167a expression was detected when plants were treated with miPEPs, other than their own miPEP, showing that miPEP activity is correlated with sequence conservation (Figure 4G).

miORF is the molecular basis of miPEP specificity

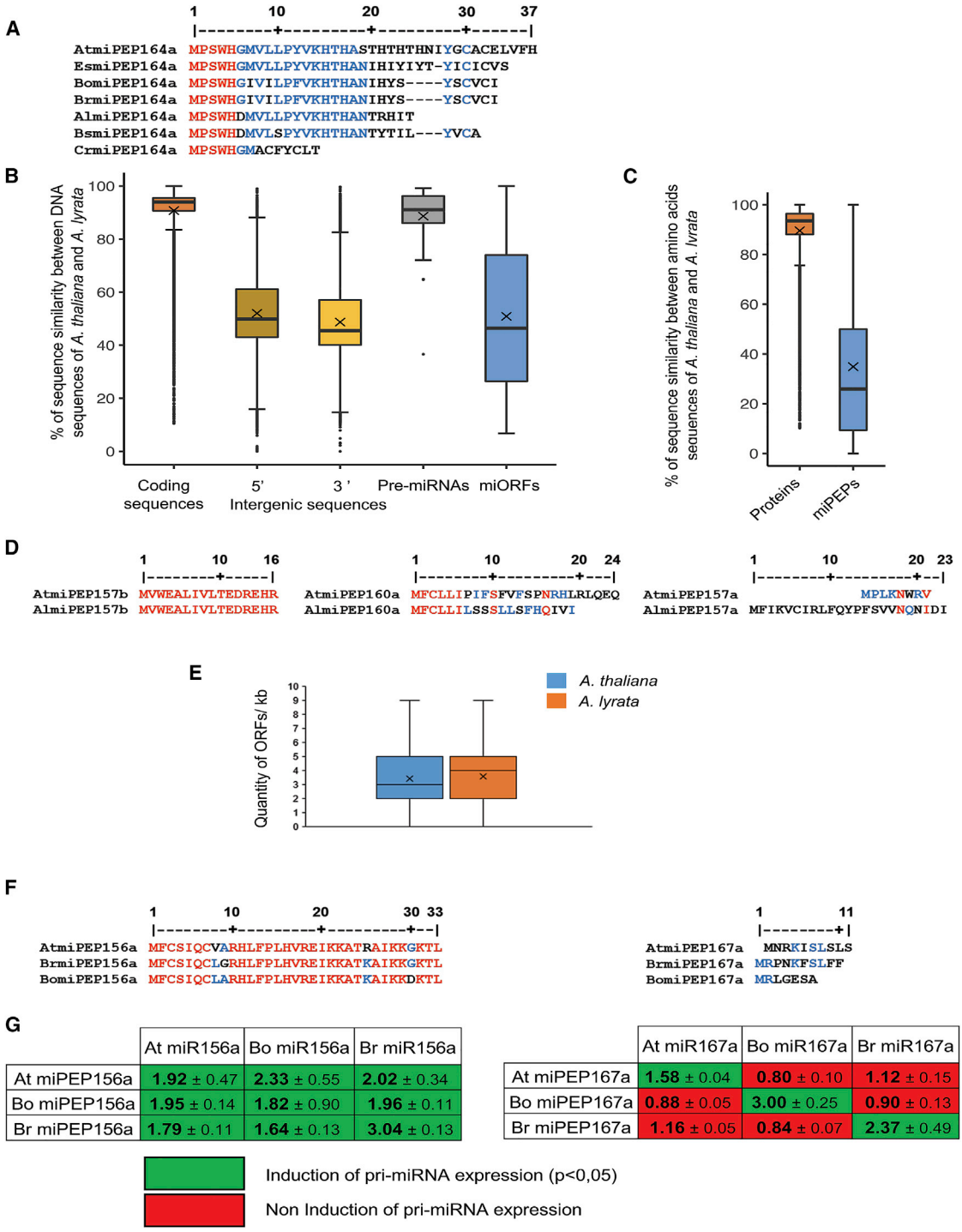
We showed here that many miPEPs are not well conserved across species. Contrasting with this, miPEPs are functional and seem to be very specific as they only up-regulate the expression of their associated miRNA, without disturbing the expression of other miRNAs, even among a same miRNA family (Lauressergues et al., 2015). Therefore, we wondered how non-conserved peptides could have similar functions in different species since they do not have high sequence identity. In order to decipher the molecular basis of this property, we used wild-type *M. truncatula* pri-miR171b that we transiently expressed in *Nicotiana benthamiana* leaves. This pri-miRNA was favored because it is one of the best characterized to date (Lauressergues et al., 2015; Couzigou et al., 2017). In addition, this heterologous context avoided potential cross-talks or unwanted feedback loops between the miRNA and its target genes, mechanisms potentially occurring in a homologous context (Feng et al., 2020). Thus, to know which components of pri-miRNA were crucial for miPEP activity, we first compared the expression levels of miR171b after expressing the miR171b precursor (pre-miR171b) or the pri-miR171b, together with miPEP171b from *M. truncatula*. While miPEP171b was able to up-regulate the expression of pri-miR171b, the pre-miR171b sequence was insensitive to the peptide (Figure 5A), indicating that the nucleotide sequences involved in miPEP response are located in the 5′ arm of the pri-miRNA. We next expressed a transcriptional fusion between the ATG of miPEP171b, under the control of its own promoter, and GUS gene or a translational fusion between the nucleotide coding sequence of miPEP171b and GUS gene (Lauressergues et al., 2015). Co-expression of miPEP171b was able to increase the expression of miPEP-GUS fusion, but not the ATG-GUS fusion (Figure 5B), showing that the miORF is important for miPEP responsiveness. Surprisingly, Sharma et al. (2020) showed that miPEP858a was able to transactivate the pri-miR858a ATG-GUS fusion (no miPEP ORF) in *A. thaliana*, suggesting a distinct mechanism by locating the miPEP858a activating sequences in the promoter region of pri-miR858a. Therefore, we questioned whether their results could be due to the fact that they were obtained in a homologous context, i.e., in the presence of the complete miR858a signaling machinery. Thus, we tested their constructs in a heterologous context. We infiltrated *N. benthamiana* leaves with ATG-GUS and ORF-GUS fusions for miPEP858a (Sharma et al., 2020), and treated plants with synthetic miPEP858a. Interestingly, we observed that miPEP858a was able to activate GUS expression of the ORF-GUS fusion, but was unable to activate GUS expression of the ATG-GUS fusion, as observed with the miR171b/miPEP171b (Figure S3).

The translation initiation codon of the miORF was next mutated in the pri-miR171b and co-expression of miPEP171b revealed an activation of expression of pri-miR171b (Figure 5C), showing that the translatability of the miORF in the pri-miRNA was not required for miPEP activity when the miPEP was independently produced.

We next degenerated the nucleotide sequence of miORF (28 mismatches on 63 nucleotides) by keeping the same amino acid sequence (degenerated miORF) (Figures 5D and 5E). Interestingly, miPEP171b was still able to activate pri-miR171b expression (Figure 5D). In parallel, we mutated the nucleotide sequence of miORF (19 mismatches on 63 nucleotides) to modify the miPEP amino acid sequence as much as possible (mutated miORF) (Figure 5E). In this case, miPEP171b was unable to improve the expression of pri-miR171b, suggesting a tight link between the peptide and its coding nucleotide sequence (Figure 5D). To test whether the miORF sequence was required for miPEP-mediated activation, we deleted it in the pri-miR171b sequence. This led to an absence of pri-miR171b induction by the miPEP171b (Figure 5F), revealing the importance of the presence of the miORF for the activity of miPEP. Interestingly, the re-introduction of the miORF171b sequence in the 3′ arm of the pri-miR171b was not able to allow an increased expression of pri-miR171b mediated by miPEP171b (Figure 5G), indicating that the role of miORF is position dependent. We next added two miORF sequences, identical to wild-type (WT) miORFs, just after the WT miORF. Interestingly, co-expression of miPEP171b increased strongly the expression of pri-miR171b (Figure 5H). Finally, we replaced the *M. truncatula* miORF171b by the *A. thaliana* miORF319a sequence in the pri-miR171b, or by a random sequence (artificial ORF), without any homology in plant genomes. We co-expressed them with an empty vector (control) or MtmiPEP171b or AtmiPEP319a or a peptide corresponding to the random sequence (artificial PEP). We first validated that the latter peptides (AtmiPEP319a and artificial PEP) had no activity on pri-miR171b expression (Figure 5I). Whereas the sequence switch removed the activity of miPEP171b as expected, it rendered pri-miR171b inducible by miPEP319a or by the artificial peptide, demonstrating a tight link between the miORF and its corresponding miPEP (Figure 5I).

Based on these observations, and given the fact that the second ORF of the *M. truncatula* pri-miR171b was presumed to be

(C) Quantification by qRT-PCR of expression of mature miR162a and miR396a after 3 h treatment by 10 μM of the indicated miPEP or irrelevant peptide.
(D) Ratio of expression of corresponding pri-miRNAs in miPEP-treated plants compared with irrelevant peptide-treated plants. Expression was quantified by qRT-PCR after 3 h of treatment by 10 μM of the indicated miPEP or irrelevant peptide. Error bars represent standard errors of the means, asterisks indicate a significant difference between the test condition (miPEP-treated plants), and the control (irrelevant peptide-treated plants) according to the Wilcoxon test ($n = 6$; $p < 0.05$). Scale bar, 5 mm (A).



	At miR156a	Bo miR156a	Br miR156a
At miPEP156a	**1.92** ± 0.47	**2.33** ± 0.55	**2.02** ± 0.34
Bo miPEP156a	**1.95** ± 0.14	**1.82** ± 0.90	**1.96** ± 0.11
Br miPEP156a	**1.79** ± 0.11	**1.64** ± 0.13	**3.04** ± 0.13

	At miR167a	Bo miR167a	Br miR167a
At miPEP167a	**1.58** ± 0.04	**0.80** ± 0.10	**1.12** ± 0.15
Bo miPEP167a	**0.88** ± 0.05	**3.00** ± 0.25	**0.90** ± 0.13
Br miPEP167a	**1.16** ± 0.05	**0.84** ± 0.07	**2.37** ± 0.49

Figure 4. Conservation of sequences and functions of miPEPs in *Brassicaceae*

(A) Alignment of miPEP164a in several *Brassicaceae* species. Al, *Arabidopsis lyrata*; At, *Arabidopsis thaliana*; Bo, *Brassica oleracea*; Br, *Brassica rapa*; Bs, *Boechera stricta*; Cr, *Capsella rubella*; Es, *Eutrema salsugineum*.

(B) Nucleotide sequence homologies of different DNA sequence types between *A. thaliana* and *A. lyrata*.

(C) Protein sequence homologies of proteins and miPEPs between *A. thaliana* and *A. lyrata*.

(D) Examples of conserved (miPEP157b) and non-conserved miPEPs (miPEP160a and miPEP157a) between *A. thaliana* and *A. lyrata*.

(E) Comparison of quantity of ORFs/kb between *A. thaliana* and *A. lyrata* pri-miRNAs.

(legend continued on next page)

not functional because it was not translated *in planta* (Lauressergues et al., 2015; Figure 6A), we tested whether the corresponding peptide might be active when expressed. For this purpose, we co-expressed the pri-miR171b together with either a control vector, a vector overexpressing miPEP171b or the peptide coded by the second ORF (miPEP171b2). Quantification of pri-miR171b revealed that both constructs had the same potency to increase the level of pri-miR171b (Figure 6B). These data suggest that the translation of any ORF present in a pri-miRNA might produce a peptide able to increase pri-miRNA expression. We validated this hypothesis by showing that other peptides from different ORFs (from second to fourth ORF) of unrelated pri-miRNAs from another plant species (*A. thaliana*), were also able to increase expression of their corresponding pri-miRNA (Figures 6C–6F).

Together, these data show that the coding sequence of the miORF itself constitutes a key element for the basis of miPEP responsiveness and specificity and explains how a non-conserved miPEP can have specific functions only in its host species.

In order to get more insights into the molecular mechanisms leading to miPEP-mediated activation of pri-miRNA expression, we tested whether these peptides could interact with their own nucleic acid sequence as suggested by the above experiments. We were able to observe that the miPEP is in close proximity of its ORF *in vivo* thanks to Förster resonance energy transfer- fluorescence lifetime imaging microscopy (FRET-FLIM) experiments (Camborde et al., 2017). For this, we fused the miPEP171b to GFP and, after validation of its functionality (Figure S4), we co-expressed it with the pri-miR171b in *N. benthamiana* leaves. GFP lifetime measurement by FRET-FLIM, after staining of the nucleic acids with Sytox orange (Camborde et al., 2017), revealed that the miPEP171b was in close proximity to the RNA of its miORF since no FRET was detected when the miORF171b was deleted from the pri-miRNA or when RNAs was degraded by RNase treatments (Figures 7A and 7B).

Finally, we questioned whether this close proximity was really an interaction and whether other molecules were required to facilitate this interaction. To do this, we performed isothermal titration calorimetry (ITC) experiments to detect interactions between the miPEP171b and its nascent RNA. Because of its strong hydrophobicity, miPEP171b was unable to be used in ITC experiments. We used instead the miPEP171b2 which behaved as miPEP171b1 (Figures 6A and 6B). Binding of RNA to the peptide is associated with an exothermic profile, where enthalpy drives exclusively the reaction (Figure S5, left panel). As the binding enthalpy is a measure of binding specificity (van der Waals interactions and hydrogen bonding), the thermodynamic profile suggests that the RNA could engage specific bonding interactions with the peptide. The ITC data reveal a weak binding (approximately 170 mM). The ITC titration of the same RNA by a scrambled peptide gave a different thermodynamic profile consisting of entropy-driven reaction that could indicate non-specific binding (Figure S5, right panel). Together, these observations support the ability of the miPEP171b2 to interact specifically with its nascent RNA.

Taken together, these data show that the coding sequence of the miORF itself constitutes a key element for the basis of miPEP responsiveness and specificity. Since the ORF encoding miPEP is important in the peptide activity, this also explains why a non-conserved miPEP fulfill specific functions only in its host species.

DISCUSSION

Whereas the precursors of miRNAs have been well identified, especially in model plants such as *A. thaliana*, their primary transcripts are still poorly known. Here, we characterized the pri-miRNAs of *A. thaliana*. We used a combination of several approaches to accurately annotate most of the pri-miRNAs. However, some of them could not be analyzed or detected, probably because our samples did not cover all the conditions of gene expression, such as plant life cycle or stress responses. Alternatively, it is possible that some *in silico* predicted and annotated miRNAs are not expressed or expressed at extremely low levels *in planta*. Interestingly, we were able to show that plant miRNAs can be located within coding genes, in UTRs or introns, as observed in animals. When possible, the annotation of pri-miRNAs gave us several information, such as the existence, for each pri-miRNA, of a population of heterogeneous transcripts, including a high occurrence of alternative splicing and different initiation/termination sites of transcription. Some of these transcripts contained the stem loop and could be maturated to generate the mature miRNAs, while the transcripts lacking the miRNA were likely translated in the cytoplasm to produce miPEPs. Whereas such complexity could explain the translatability of miPEPs, with different fates for different transcripts, its exact role remains unknown, although this situation resembles what has been described in humans (Chang et al., 2015). Following their translation, miPEPs can next diffuse into the nucleus in which they require their nascent sequence to activate transcription of their pri-miRNA. How can miPEPs specifically target their own miRNA? By performing deletion experiments, mutagenesis, and miORF nucleotide sequence swapping, we found that the miORF itself constitutes a prerequisite of miPEP responsiveness and activation. This result identifies a first molecular key to miPEP specificity. In addition, we have observed an *in vitro* association between miPEPs and their nascent ORF (RNA) and demonstrated a close proximity of these partners in nuclei, strongly suggesting a direct interaction in the nucleus.

The identification of naturally expressed miPEPs remains difficult. Indeed, mass spectrometry only allows the identification a small fraction of total peptides in an organism and, to date, only one miPEP was detected using this method in humans (Kang et al., 2020). The use of specific antibodies remains the

(F) Alignments of miPEP156a and miPEP167a in *A. thaliana*, *Brassica oleracea*, and *Brassica rapa*.
(G) *A. thaliana*, *B. oleracea*, and *B. rapa* seedlings were treated with miPEPs indicated to the left of each table and expression of corresponding pri-miRNAs (indicated on the top of tables) were measured by qRT-PCR. Green boxes illustrate significant induction of expression while red boxes show no induction of expression, according to the Wilcoxon test ($n = 6–10$; $p < 0.05$). The numbers in cases indicate the ratio of expression of corresponding pri-miRNA in miPEP-treated plants compared with irrelevant peptide-treated plants, ± standard error of the mean.

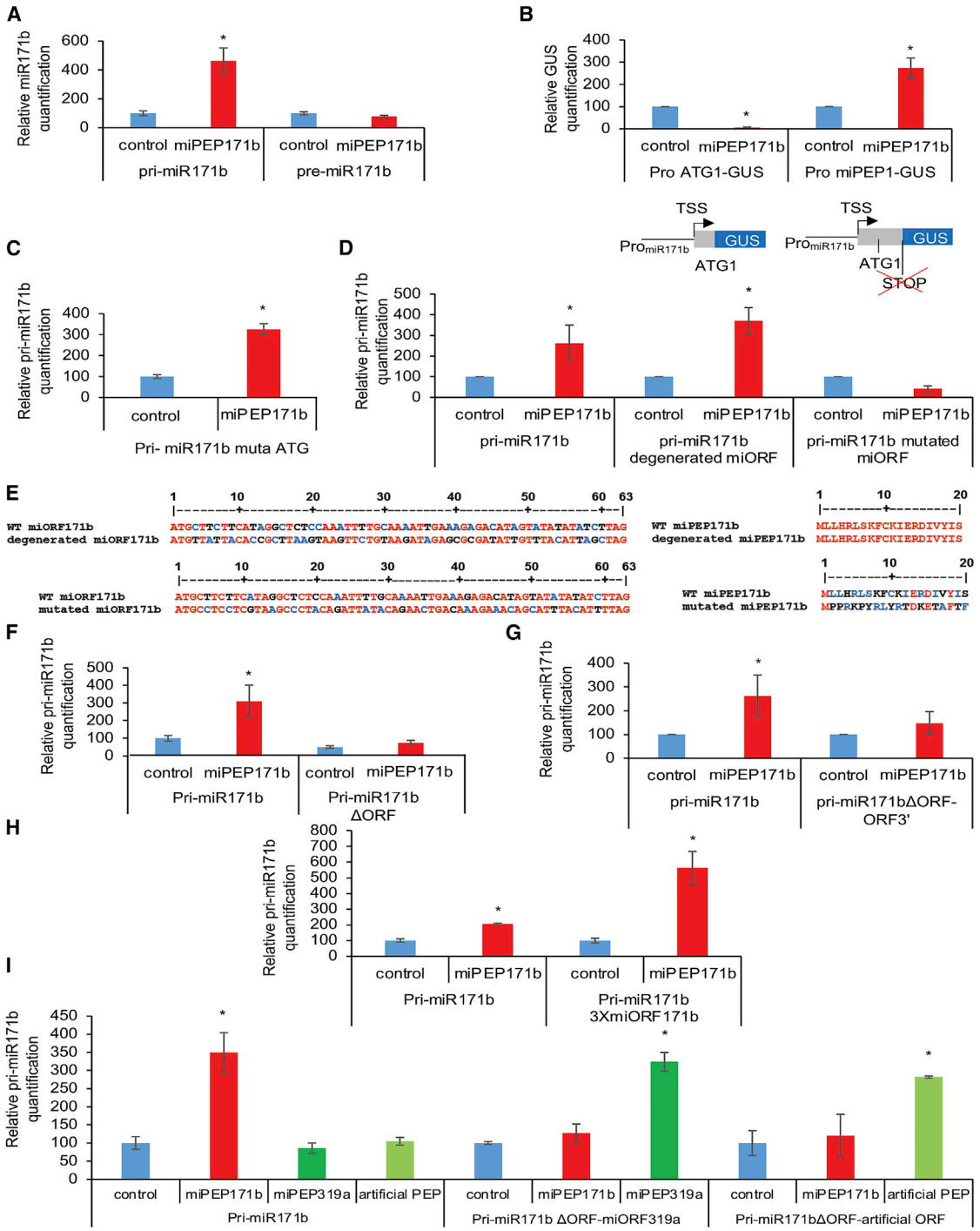


Figure 5. The miORF sequence governs specificity of miPEP for its associated pri-miRNA

Quantification by qRT-PCR of pre-miR171b, pri-miR171b, or GUS expression in *N. benthamiana* leaves after infiltration of the following constructs, together with empty vector (control) or miPEP171b, or AtmiPEP319a or artificial peptide (artificial PEP).

(A) Pri-miR171b, pre-miR171b.

(B) Pro ATG1-GUS or Pro-miPEP171b-GUS.

(C) Pri-miR171b in which the ATG of miORF was mutated.

(D) WT pri-miR171b or miR171b carrying a degenerated miORF or a mutated miORF.

(E) Nucleic acids and protein alignments of miORFs and miPEPs in the pri-miR171b degenerated miORF and mutated miORF constructs, compared with WT.

(F) WT pri-miR171b or miORF-deleted pri-miR171b (miR171b ΔORF).

(G) WT pri-miR171b or miORF-deleted pri-miR171b with miORF replaced in 3′ arm (miR171b ΔORF-orf 3′).

(legend continued on next page)

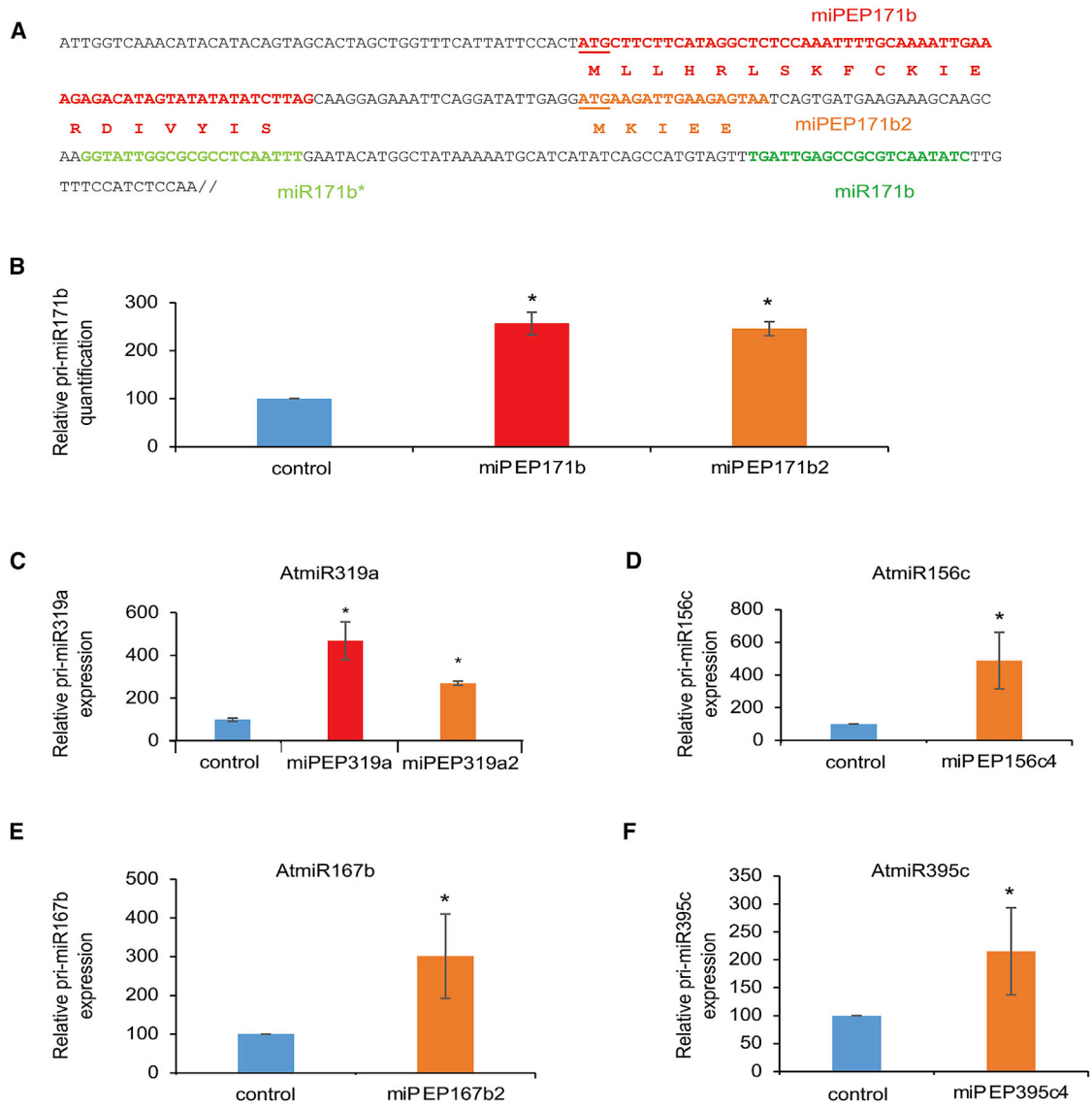


Figure 6. Further ORFs can produce active miPEPs

(A) Schematic representation of the annotated 5′ sequence of pri-miR171b of *M. truncatula* with miRNA (green), miRNA* (light green), miPEP171b (red), and miPEP171b2 (orange).

(B) Quantification by qRT-PCR of pri-miR171b expression in *N. benthamiana* leaves after co-infiltration with one of the following constructs: empty vector (control), miPEP171b (red), or miPEP171b2 (orange).

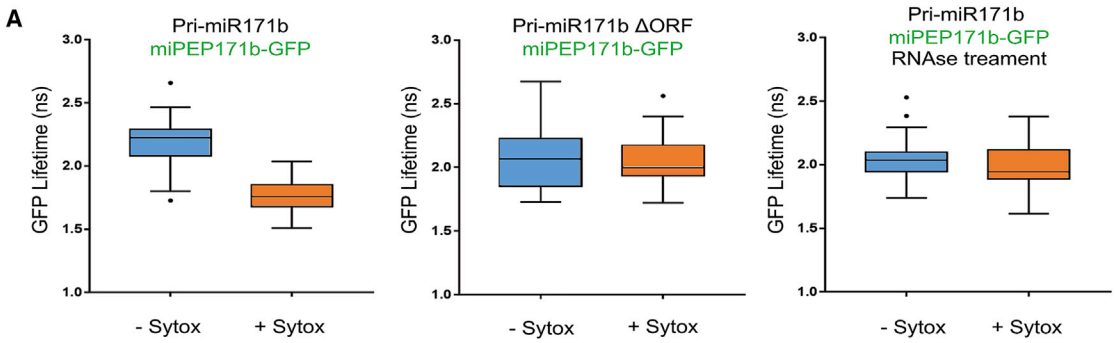
(C) Quantification by qRT-PCR of expression of pri-miR319a from *A. thaliana* in *N. benthamiana* leaves after co-infiltration with one of the following constructs: empty vector (control), miPEP319a, or miPEP319a2.

(D–F) Quantification of pri-miRNA expression after treatment of *A. thaliana* seedlings with 100 μM of the indicated miPEP for 24 h. Error bars represent SEMs, asterisks indicate a significant difference between the test condition and the control according to the Wilcoxon test (n = 8; $p < 0.05$).

best method to demonstrate miPEP synthesis (Lauressergues et al., 2015; Sharma et al., 2020), but it is not suitable for high throughput identification. Moreover, owing to the poor annotation of pri-miRNAs and the presence of several ORFs within a pri-miRNA, the use of antibodies requires the prior identification of the active ATG, by using promoter GUS-fusion analysis for example. Here, we observed that for all tested pri-miRNAs, an active ATG was detected at the first position from the

(H) WT pri-miR171b or pri-miR171b carrying three miORF171b sequences.

(I) WT pri-miR171b or pri-miR171b version in which miORF171b was replaced by miORF319a (miR171b ΔORF-miORF319a) or artificial ORF (miR171b ΔORF-artificial ORF). Transgenes were expressed in pCAMBIA 2301 under the control of 35S promoter, except for (B). Error bars represent standard errors of the means, asterisks indicate a significant difference between the test condition and the control according to the Wilcoxon test (n = 8; $p < 0.05$).



B

Donor	Acceptor (Sytox)	t [a]	sem [b]	N [c]	E [d]	p-value [e]
miPEP171b-GFP	-	2.190	0.029	41	-	-
Pri-miR171b	+	1.763	0.018	40	19	$8E^{-20}$
miPEP171b-GFP	-	2.075	0,046	30	-	-
Pri-miR171b ΔORF	+	2.045	0.025	50	1.4	0.54
miPEP171b-GFP	-	2.046	0.031	30	-	-
Pri-miR171b (RNase treatment)	+	1.995	0.026	42	2.5	0.22

Figure 7. miPEPs are physically interacting with their nascent RNAs

(A and B) FRET-FLIM measurements showing that miPEP171b interacts with pri-miR171b mRNA *in planta*.

(A) FRET-FLIM analysis of miPEP171b-GFP with pri-miR171b (left), miR171b ΔORF (middle) or pri-miR171b with a RNase treatment (right).

(B) FRET-FLIM values: ([a]) Mean lifetime in nanoseconds. For each cell, average fluorescence decay profiles were plotted and lifetimes were estimated by fitting data with exponential function using Poissonian maximum like hood estimation procedure. ([b]) Standard error of the mean. ([c]) N: total number of measured cells. ([d]) Percentage of FRET efficiency: E = 1 – (tDA/tD). ([e]) p value (Student *t* test) of the difference between the donor lifetimes in the presence or absence of an acceptor.

transcriptional start site, suggesting that the expression of miPEPs is primarily dependent on this first ORF. We finally further observed that all the miPEPs tested were able to increase expression of their pri-miRNA. In some cases, optimizing the time of treatment and the concentration was required to detect pri-miRNA up-regulation, showing that miPEP-mediated regulation was transient and dose and time dependent, as already suggested (Ormancey et al., 2020).

Sequence conservation is generally the first criterion to identify a gene of interest with an important role in plants. It is based on the hypothesis that genes allowing similar biological functions, with strong protein sequence conservation, should control important functions in plant life. Indeed, the search for proteins or genes sharing sequence similarities is often the first step to identify biological actors playing a specific role or involved in a specific pathway. It is often accepted that proteins from phylogenetically distant organisms can have homologous functions if they share significant sequence conservation. However, miPEPs seem to exhibit poor to no sequence conservation between species while retaining the same molecular activity, i.e., the enhancement of miRNA expression (Couzigou et al., 2017), raising a paradox. This property could be explained by the fact that activity of the peptide directly depends on its nucleic acid coding sequence. Therefore, if a mutation occurs in an miORF, the resulting miPEP would be modified accordingly and would still improve pri-miRNA expression. This characteristic might explain why there is apparently no selection pressure on miPEPs and the fact that non-conserved sequences can serve as driving forces for evolutionary diversification.

Limitations of the study

Because short transcripts always share their sequence with long transcripts, it is impossible to quantify their expression. Our conclusion is therefore a deduction: if the ratio of long pri-miRNA transcripts on the total pri-miRNA transcripts is lower in RPL18-IP samples and higher in nuclear samples, it means that the rest of transcripts (i.e., short transcripts) show an inverted scheme (Figure 2).

It has to be noted that FRET-FLIM experiments only demonstrate a very close proximity between the miPEP and its nascent pri-miR (less than 10 nm) (Figure 7). That is why an ITC experiment (an *in vitro* approach) was performed. Thus, the results clearly indicate that miPEPs directly interact with the pri-miRNAs from which they originate.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2022.110339.

ACKNOWLEDGMENTS

This work was funded by the French ANR project BiomiPEP (ANR-16-CE12-0018-01). It was carried out in the *Laboratoire de Recherche en Sciences Végétales*, which belongs to the *Laboratoire d'Excellence* entitled TULIP (ANR-10-LABX-41). We thank the « Plateforme GENTYANE INRAE Clermont » (Clermont-Ferrand, France) for PacBio Iso-seq sequencing. We thank PK Trivedi (CSIR, India) for sending miPEP858 constructs.

AUTHOR CONTRIBUTIONS

J.P.C. designed the research; J.P.C., D.L., B.G., S.T., S.P., and C.D. performed the molecular biology; J.P.C. and M.O. performed plant experiments; L.C., A.L., and A.J. performed imagery analyses; V.G. performed ITC; H.S.C., J.P.C., P.C., and A.B. performed bio-informatics analyzes; J.P.C., P.T., and S.P. wrote the paper.

DECLARATION OF INTERESTS

J.P.C. is a founder and a scientific advisor of Micropep Technologies, which aims to commercialize peptides in agronomy. In this study, Micropep Technologies has participated in the identification of Arabidopsis pri-miRNAs.

Received: May 26, 2021
Revised: September 16, 2021
Accepted: January 12, 2022
Published: February 8, 2022

REFERENCES

Bazin, J., Baerenfaller, K., Gosai, S.J., Gregory, B.D., Crespi, M., and Bailey-Serres, J. (2017). Global analysis of ribosome-associated noncoding RNAs unveils new modes of translational regulation. Proc. Natl. Acad. Sci. U S A *114*, E10018–E10027.

Camborde, L., Jauneau, A., Brière, C., Deslandes, L., Dumas, B., and Gaulin, E. (2017). Detection of nucleic acid-protein interactions in plant leaves using fluorescence lifetime imaging microscopy. Nat. Protoc. *12*, 1933–1950.

Chang, T.C., Pertea, M., Lee, S., Salzberg, S.L., and Mendell, J.T. (2015). Genome-wide annotation of microRNA primary transcript structures reveals novel regulatory mechanisms. Genome Res. *25*, 1401–1409.

Chen, Q.J., Deng, B.H., Gao, J., Zhao, Z.Y., Chen, Z.L., Song, S.R., Wang, L., Zhao, L.P., Xu, W.P., Zhang, C.X., et al. (2020). A miRNA-encoded small peptide, vvi-miPEP171d1, regulates adventitious root formation. Plant Physiol. *183*, 656–670.

Combier, J.P., de Billy, F., Gamas, P., Niebel, A., and Rivas, S. (2008). Trans-regulation of the expression of the transcription factor MtHAP2-1 by a uORF controls root nodule development. Genes Dev. *22*, 1549–1559.

Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res. *16*, 10881–10890.

Couzigou, J.M., André, O., Guillotin, B., Alexandre, M., and Combier, J.P. (2016). Use of microRNA-encoded peptide miPEP172c to stimulate nodulation in soybean. New Phytol. *211*, 379–381.

Couzigou, J.M., Lauressergues, D., André, O., Gutjahr, C., Guillotin, B., Bécard, G., and Combier, J.P. (2017). Positive gene regulation by a natural protective miRNA enables Arbuscular Mycorrhizal Symbiosis. Cell Host Microbe *21*, 106–112.

Feng, Y.Z., Yu, Y., Zhou, Y.F., Yang, Y.W., Lei, M.Q., Lian, J.P., He, H., Zhang, Y.C., Huang, W., and Chen, Y.Q. (2020). A Natural variant of miR397 mediates a feedback loop in circadian rhythm. Plant Physiol. *182*, 204–214.

Han, M.H., Goud, S., Song, L., and Fedoroff, N. (2004). The Arabidopsis double-stranded RNA-binding protein HYL1 plays a role in microRNA-mediated gene regulation. Proc. Natl. Acad. Sci. U S A *101*, 1093–1098.

Immarigeon, C., Frei, Y., Delbare, S.Y.N., Gligorov, D., Machado Almeida, P., Grey, J., Fabbro, L., Nagoshi, E., Billeter, J.C., Wolfner, M.F., et al. (2021). Identification of a micropeptide and multiple secondary cell genes that modulate Drosophila male reproductive success. Proc. Natl. Acad. Sci. U S A *118*, e2001897118.

Juntawong, P., Girke, T., Bazin, J., and Bailey-Serres, J. (2014). Translational dynamics revealed by genome-wide profiling of ribosome footprints in *Arabidopsis*. Proc. Natl. Acad. Sci. U S A *111*, 203–212.

Kang, M., Tang, B., Li, J., Zhou, Z., Liu, K., Wang, R., Jiang, Z., Bi, F., Patrick, D., Kim, D., et al. (2020). Identification of miPEP133 as a novel tumor-suppressor microprotein encoded by miR-34a pri-miRNA. Mol. Cancer *19*, 143.

Kramer, M.F. (2011). Stem-loop RT-qPCR for miRNAs. Curr. Protoc. Mol. Biol. Chapter 15:Unit 15.10. https://doi.org/10.1002/0471142727.mb1510s95.

Lauressergues, D., Couzigou, J.M., San Clemente, H., Martinez, Y., Dunand, C., Bécard, G., and Combier, J.P. (2015). Primary transcripts of microRNAs encode regulatory peptides. Nature *520*, 90–93.

Montigny, A., Tavormina, P., Duboé, C., San Clemente, H., Aguilar, M., Valenti, P., Lauressergues, D., Combier, J.P., and Plaza, S. (2021). Drosophila primary microRNA-8 encodes a microRNA-encoded peptide acting in parallel of miR-8. Genome Biol. *22*, 118.

Morozov, S.Y., Ryazantsev, D.Y., and Erokhina, T.N. (2019). Bioinformatics analysis of the novel conserved micropeptides encoded by the plants of family *Brassicaceae*. J. Bioinform. Syst. Biol. *2*, 066–077.

Mustroph, A., Juntawong, P., and Bailey-Serres, J. (2009). Isolation of plant polysomal mRNA by differential centrifugation and ribosome immunopurification methods. In Plant Systems Biology, D.A. Belostotsky, ed. (Humana Press), pp. 109–126. https://doi.org/10.1007/978-1-60327-563-7_6.

Niu, L., Lou, F., Sun, Y., Sun, L., Cai, X., Liu, Z., Zhou, H., Wang, H., Wang, Z., Bai, J., et al. (2020). A micropeptide encoded by lncRNA MIR155HG suppresses autoimmune inflammation via modulating antigen presentation. Sci. Adv. *6*, eaaz2059.

Ormancey, M., Le Ru, A., Duboé, C., Jin, H., Thuleau, P., Plaza, S., and Combier, J.P. (2020). Internalization of miPEP165a into *Arabidopsis* roots depends on both passive diffusion and endocytosis-associated processes. Int. J. Mol. Sci. *21*, 2266.

Ormancey, M., Guillotin, B., San Clemente, H., Thuleau, P., Plaza, S., and Combier, J.P. (2021). Use of microRNA-encoded peptides to improve agronomic traits. Plant Biotechnol. J. https://doi.org/10.1111/pbi.13654.

Prel, A., Dozier, C., Combier, J.P., Plaza, S., and Besson, A. (2021). Evidence that regulation of pri-miRNA/miRNA expression is not a general rule of miPEPs function in Humans. Int. J. Mol. Sci. *22*, 3432.

Rice, P., Longden, I., and Bleasby, A. (2000). EMBOSS: the European molecular biology open software suite. Trends Genet. *16*, 276–277.

Sharma, A., Badola, P.K., Bhatia, C., Sharma, D., and Trivedi, P.K. (2020). Primary transcript of miR858 encodes regulatory peptide and controls flavonoid biosynthesis and development in Arabidopsis. Nat. Plants. *6*, 1262–1274.

Wang, S., Quan, L., Li, S., You, C., Zhang, Y., Gao, L., Zeng, L., Liu, L., Qi, Y., Mo, B., et al. (2019). The PROTEIN PHOSPHATASE4 complex promotes transcription and processing of primary microRNAs in *Arabidopsis*. Plant Cell *31*, 486–501.

Wang, S., Tian, L., Liu, H., Li, X., Zhang, J., Chen, X., Jia, X., Zheng, X., Wu, S., Chen, Y., et al. (2020). Large-scale discovery of non-conventional peptides in maize and *Arabidopsis* through an integrated peptidogenomic pipeline. Mol. Plant *13*, 1–16.

Xie, Z., Allen, E., Fahlgren, N., Calamar, A., Givan, S.A., and Carrington, J.C. (2005). Expression of *Arabidopsis* MIRNA genes. Plant Physiol. *138*, 2145–2154.

Zhang, Q.L., Su, L.Y., Zhang, S.T., Xu, X.P., Chen, X.H., Li, X., Jiang, M.Q., Huang, S.Q., Chen, Y.K., Zhang, Z.H., et al. (2020). Analyses of microRNA166 gene structure, expression, and function during the early stage of somatic embryogenesis in Dimocarpus longan Lour. Plant Physiol. Biochem. *147*, 205–214.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit anti-HA	Sigma-Aldrich	Cat#H6908; RRID:AB_260070
Red ANTI-FLAG® M2 Affinity Gel	Sigma-Aldrich	Cat#F2426; RRID:AB_2616449
HRP-conjugated goat anti-rabbit igG	Sigma-Aldrich	Cat#AP307P; RRID:AB_11212848
Chemicals, peptides, and recombinant proteins		
See "method details" section for peptides	This paper (synthesized by sb-PEPTIDE)	N/A
Acetonitrile	Fisher Chemical	Cat#A/0638/15
Acetic acid	Fisher Chemical	Cat#A/0400/PB15
KALYS AGAR	Labover	Cat#HP696-1
Murashige and Skoog medium (MS)	Duchefa Biochemie	Cat#M0222
Mes hydrate	Sigma-Aldrich	Cat#M8250
Formaldehyde	Sigma-Aldrich	Cat#252549
Sucrose	Euromedex	Cat#200-301-B
Ficoll PM400	GE Healthcare	Cat#17-0300-10
Dextran 40	Sigma-Aldrich	Cat#31389
Hepes	Sigma-Aldrich	Cat#H3375
Magnesium Chloride ($MgCl_2$)	Sigma-Aldrich	Cat#M8266
Dithiothreitol (DTT)	Sigm-Aldrich	Cat#1114740025
Phenylmethylsulfonyl fluoride (PMSF)	Sigma-Aldrich	Cat#P7626
Plant protease inhibitors	Sigma-Aldrich	Cat#P9599
RNase inhibitor	Promega	Cat#N2511
Triton X-100	Sigma-Aldrich	Cat#T8787
Potassium chloride (KCl)	Merck-Millipore	Cat#7447-40-7
Sodium chloride (NaCl)	Euromedex	Cat#1112-A
EDTA	Euromedex	Cat#EU0007
Igepal	Sigma-Aldrich	Cat#18896
Tris base	Euromedex	Cat#26-128-3094-B
$CaCl_2$	Sigma-Aldrich	Cat#499609
SDS	Euromedex	Cat#EU0660
Tri Reagent	Sigma-Aldrich	Cat#93289
X-gluc	Euromedex	Cat#EU0700-D
Paraformaldehyde, 16% (vol/vol) solution	Electron Microscopy Sciences,	Cat#15710
Proteinase K	Thermo Fisher Scientific	Cat#25530049
Sytox Orange	Thermo Fisher Scientific	Cat#S11368
RNAse A, DNAse free	Thermo Fisher Scientific	Cat#EN0531
TBS, 10× (Tris-buffered saline solution: Tris 0.24 M; NaCl 1.37 M; KCl 26.8 mM; pH 7.5)	Euromedex	Cat#ET220
Clarity Max™ Western ECL Substrate	Biorad	Cat#1705062
4–20% Mini-PROTEAN® TGX™ Precast Protein Gels	Biorad	Cat#4561094
Nitrocellulose blotting membrane, 0.1 μm	GE Healthcare	Cat#10600000
LightCycler® 480 SYBR Green I Master	Roche	Cat#4887352001
SuperScript™ III Reverse Transcriptase 200 U/μL	Invitrogen	Cat#10368252

(*Continued on next page*)

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
Critical commercial assays		
FirstChoice RLM-RACE kit -	ThermoFischer	Cat#AM1700
RNeasy kit	Qiagen	Cat#74904
TNT® coupled wheat germ extract systems	Promega	Cat#L4140
Deposited data		
A. thaliana pri-miRNAs	This paper	GenBank: MW_775349-MW_775424
PacBio data	This paper	GenBank: SUB_9676855
Experimental models: Organisms/strains		
Arabidopsis: Col-0	N/A	N/A
Arabidopsis: 35S::HF-RPL18	Mustroph et al. (2009)	N/A
Nicotiana benthamiana	N/A	N/A
Brassica rapa: navet blanc globe à collet violet	Vilmorin	www.vilmorin-jardin.com
Brassica oleracea: chou cabus tête de pierre HF1	Vilmorin	www.vilmorin-jardin.com
Oligonucleotides		
See "method details" section for oligonucleotides	This paper	N/A
Recombinant DNA		
See "method details" section for plasmids:		
Software and algorithms		
Orthomcl	Rice et al. (2000)	
Needle	Rice et al. (2000)	
Multalin	Corpet (1988)	http://multalin.toulouse.inra.fr/multalin/
SymphoTime 64	PicoQuant	https://www.picoquant.com/products/category/software/symphotime-64-fluorescence-lifetime-imaging-and-correlation-software
Light Cycler 480	Roche	https://lifescience.roche.com/en_fr/products/lightcycler14301-480-software-version-15.html
Microcal Origin	Originlab	Microcal Origin Download - Origin - a complete graphing and data analysis software (informer.com)
R (version 4.0.0)	R	https://www.r-project.org/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Jean-Philippe Combier (combier@lrsv.ups-tlse.fr).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Primary transcripts identified in this study have been deposited in GenBank (MW775349-MW775424). PacBio Iso-Seq data have been deposited in Genbank (SUB9676855).
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this work paper is available from the Lead Contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Biological material, growth conditions, and biological assays

Nicotiana benthamiana plants were grown as described in Combier et al. (2008).

Bleach sterilized *Arabidopsis thaliana* Col-0 seeds were sawn in 24-well plates containing 1 mL of half MS medium (no sucrose), 1% KALYS AGAR. We put 10–25 seeds on the agar before cold stratification at 4°C. After 48 h, plates were placed in a growth chamber 16/8 h day/night, at 22°C, LED light 250 μmol photons $m^{-2}.s^{-1}$. Plants were treated for 1–24 h with 500 μL of half MS liquid medium containing 1–200 μM peptide. Water and an irrelevant peptides were used as controls. The experiment was repeated four to eight times independently. Alternatively, seeds were sown in pot and grown for 3 weeks before treatment with miPEPs as above. *Brassica rapa* and *Brassica oleracea* plants were cultivated as above.

METHOD DETAILS

Peptides

Peptides were synthesized by sb-PEPTIDE (www.sb-peptide.com) and dissolved at 10 mM in water (stock solution), except for MtmiPEP171b and its scrambled peptide (50% Acetonitrile, 10% acetic acid), aliquoted and conserved at –80°C. Thawed peptides were immediately used and not re-frozen to avoid peptide degradation (Ormancey et al., 2020).

Peptides were used at 1–100 μM final concentration, for 1–24 h before plant harvesting.

The peptides used in this study and their sequence are listed below.

name	Sequence
AtmiPEP156a	MFCSIQCVARHLFPLHVREIKKATRAIKKGKTL
AtmiPEP156c4	MREFWDKF
AtmiPEP162a	MSFFFLVCYSCYMLFKYDF
AtmiPEP162b	MHSLYIYEQLLVL
AtmiPEP163	MSTTQEHRS
AtmiPEP167a	MNRKISLSLS
AtmiPEP167b2	MQEETYEG
AtmiPEP169l	MRHKES
AtmiPEP172b	MCTYYYLINKYF
AtmiPEP319a	MNIHTYHHLLFPSLVFHQSSDVPNALSLHIHTYEYIIVVIDPFRITLAFR
AtmiPEP319a2	MFQTLYLFIYIHTNILLLS
AtmiPEP395c4	MTEQEEESQM
AtmiPEP396a	MTLLFPCPL
BomiPEP156a	MFCSIQCLGRHLFPLHVREIKKATKAIKKGKTL
BomiPEP167a	MRLGESA
BrmiPEP156a	MFCSIQCLARHLFPLHVREIKKATKAIKKDKTL
BrmiPEP167a	MRPNKFSLFF
MtmiPEP171b	MLLHRLSKFCKIERDIVYIS
MtmiPEP171b2	MKIEE
Scrambled miPEP171b2	EIMEK
Irrelevant peptide	RNADAGRGIP

Plasmid constructs

Plasmids were synthesized by Genecust (genecust.com) in pCAMBIA 2301. Expressions in *N. benthamiana* leaves were performed by agro-infiltration using the 35S promoter or the native miR171b promoter when indicated (Combier et al., 2008).

The different constructs used in this study and their sequence are presented in Table S1.

Expression analyses, 5′-3′ RACE PCR

Pri-miRNA quantification was performed by qRT-PCR. Levels of expression for the controls were set at 100. Expression were normalized with housekeeping genes (AtUBP6, NbEF1, BoACTINE2, BoTIP41, BrEF1, and BrGAPDH). For *N. benthamiana* agroinfiltrations, the kanamycin resistance selection gene present in plasmids and driven by the 35S promoter was used to normalize for transfection efficiency.

RNA extractions were performed using RNeasy kit (Qiagen, Hilden, Germany), except for total RNA extraction (see below), and DNase and RT were performed using 500 ng to 3 μg of RNA with Promega (Madison, WI) reagents according to the manufacturer's recommendations.

List of primers used for qPCR:

Name	Primer forward	Primer reverse
MtmiR171b 5′	CGCCTCAATTTGAATACATGG	ACGCGGCTCAATCAAACTAC
MtmiR171b 3′	GACAGGCGAAAAAGTTACCG	GCTTGACCTTATCCGAACCA
NbEF1	CAGCTTCTGCCACAGCTACA	GGTTGGTGAGTGGAGGAAAA
Kan	AGACAATCGGCTGCTCTGAT	CTCGTCCTGCAGTTCATTCA
AtUBP6	AATGCACATGCAGCAGGA	CTTTCACAGCATCAGCACCT
AtmiR156a	CTTCTCTTGCGTGCTCACTG	ACGAAGACAGGCCAAAGAGA
AtmiR156c	TCTTCTGGGGGAAGAGGTTT	CAATTTGAAAGGGGTGGCTA
AtmiR162a	CCAGCGTTTGCCTCTTGTAT	AAATCCTCAGCTTTCCCAGA
AtmiR162b	TGTTCATCAACCGATTTTCTCA	CCAGCGACTTCACTCTTTCC
AtmiR163	AGTTCCCGGTTCCTGAGAGT	TAAATCCCCAAATGGGTTCA
AtmiR167a	TGTTGTGTTTCATGACGATGG	AGAAGGGTGCGACAGTCAAC
AtmiR167b	TGACAGCCTCACTCCTTCCT	TTTCTTTCAATCGGCATGTG
AtmiR169l	CCAAGGATGACTTGCCTGAT	GCATAATAATCCAAAATCCACGA
AtmiR172b	GATCTCTTGTGCGTGCGTAA	CGCCTACAAACAACGACAGA
AtmiR319a	CGAGTCGCCAAAATTCAAAC	GCTCCCTTCAGTCCAATCAA
AtmiR395c	GGGGACTCTTGGTGTCATTC	ATAGAAAACCGCAGCAATGG
AtmiR396a	CCTCACTCCCTCTTTCCACA	AGGGTCATGTAGAGCAGACGA
AtmiR157aFL	TCTCATCAACAACCCTTTTGG	ACTCGCCAGTTTTTCAATGG
AtmiR157aAS	GGTACTTCTCTACGCCATCCA	CTTCGTTTTCAACCCTCTATGAA
AtmiR162aFL	TGTGTTTCGTTTGATCCGATT	CCTCCAGCGACTCTCACTCT
AtmiR162aAS	GGCCAATAGGCAAATCAGAG	CCTCACTTTTTTCACAAGTAGTA
Pre-rRNA	GCCCGGGTAATCTTTGAAAT	GTTCGCTCGCCGTTACTAAG
25S rRNA	TCTGACATGTGTGCGAGTCA	CACTTGGAGCTCTCGATTCC
Boactine2	GAATCCACGAGACAACATAT	AGGGAAGCAAGAATGGAAC
Botip41	ATTTGGCTGCTCTTTCACTT	AAATCGTAAGAGGAGAAACC
BomiR156a	GCTGATCTCTTTGGCCTGTC	AGAACAAGCGCCATCATTTC
BomiR167a	GTGCACAGGCATCTGATGAA	TAAGGGTGCGACAGTCAATG
BrEF1	ATACCAGGCTTGAGCATACCG	GCCAAAGAGGCCATCAGACAA
BrGAPDH	CCGCTAACTGCCTTGCTCCACTT	GCGGCTCTTCCACCTCTCCAGT
BrmiR156a	GCTGATCTCTTTGGCCTGTC	AGAACAAGCGCCATCATTTC
BrmiR167a	GTGCACAGGCATCTGATGAA	TAAGGGTGCGACAGTCAATG

The 5′ and -3′ RACE PCRs were conducted according to the manufacturer's recommendations (FirstChoice RLM-RACE kit-ThermoFischer).

The list of primers used for RACE PCR is provided in Table S1.

Stem-loop qRT-PCR for miRNAs

Primers for stem-loop RT-PCR were designed according to Kramer (2011).

Reverse transcriptase reaction were performed using 10 ng total RNAs, 1 nM stem-loop RT primer (up to four primers per reaction were used simultaneously, including reference genes (SnoR101 and 5S)), 50 U SuperScript III RT (Invitrogen), 0.25 mM dNTP, 1× First-Strand buffer, 0.01 M DTT, and 4 U RNAse inhibitor (40U/μL; Promega). We incubated 20 μL of reactions in Bio-RAD T100 Thermal Cycler in a single initial cycle for 30 min at 16°C, followed by 60 cycles for 30 s at 30°C, 30 s at 42°C, and 1 s at 50°C; and a last cycle for 5 min at 85°C.

Quantitative PCR reactions was composed of 2 μL of reverse transcriptase reaction product diluted to half, and 8 μL of reactional mix composed by 1 μL of primer mix at 10 μM (miRNA or reference gene with universal primer), 5 μL of LightCycler 480 SYBR Green I Master (Roche, Basel, Switzerland) and 2 μL of H_2O. The 10 μL reactions were incubated in 384-well plate in LightCycler 480 System

(Roche Life Science). The cycles were composed of 5 min at 95°C followed by 45 cycles for 5 s at 95°C and 10 s at 60°C, incubations of 10 s at 95°C and 20 s at 65°C for the melting curve; and the last cycle for 1 min at 40°C. All the reactions were run in technical duplicate.

Name	Primer
MAT miR162aF	CTGGCTCGATAAACCTCTG
SL RT miR162a	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACCTGGAT
Universal Primer	CCAGTGCAGGGTCCGAGGTA
Ath_MAT_5SqF	TTGCAGAATCCCGTGAACCA
ST RT 5Sq	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACTTGTGA
MAT miR396aF	CTGGCTTCCACAGCTTTCT
SL RT miR396a	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACCAGTTC

Total RNA isolation

Total RNAs from different tissues (seedlings, leaves, rosettes, flowers) were extracted with Tri Reagent (Sigma, St Louis, MO) according to the manufacturer's protocol and sequenced with PacBio (Plateforme GENTYANE INRAE Clermont-Ferrand, France).

Nuclear RNA isolation

WT *Arabidopsis* seedlings (2 weeks old) were fixed in 1% formaldehyde for 15 min under vacuum and ground in a mortar with a pestle under liquid nitrogen into fine powder. We transferred 0.5 mL of powder into a 2 mL Eppendorf tube with 1.5 mL of ice cold lysis buffer (0.44 M sucrose, 1.25% Ficoll, 2.5% dextran 40, 20 mM Hepes KOH pH 7.4, 10 mM $MgCl_2$, 5 mM DTT, 1 mM PMSF, 1% plant protease inhibitors (Sigma P9599), RNase inhibitor 8 U/mL), and vortexed for 30 s. The lysate was filtrated with two layers of miracloth in a syringe, 0.5% Triton X-100 was added. The sample was vortexed and centrifugated for 30 s at 16,000 g at 4°C. The supernatant was removed and the pellet was resuspended with 500 μL of Buffer A (0.3 M sucrose, 60 mM KCl, 15 mM NaCl, 2 mM EDTA, 15 mM Hepes, 0.5% Igepal, 5 mM DTT, 1 mM PMSF, 1% plant protease inhibitors (Sigma P9599), RNase inhibitor 8U/mL). The resuspended pellet was put on a 1 mL cushion of buffer B (same as buffer A, but 0.87 M sucrose) in a new 1.5 mL Eppendorf and centrifugated for 30 s at 16,000 g at 4°C. The supernatant was removed and the previous step was repeated on the cushion to further wash the nuclei. The supernatant was removed and 300 μL of lysis buffer (50 mM Tris-HCl pH 8, 10 mM EDTA, 25 mM $MgCl_2$, 5 mM $CaCl_2$, 1% SDS, 1 mM PMSF, 1% plant protease inhibitors, RNase inhibitor 160 U/mL) was added before vortexing. RNA were then extracted using 1.5 mL Tri Reagent according to the manufacturer's procedure.

Polysomal RNA isolation

RNA loaded into ribosomes were isolated using the modified immunopurification of polysome protocol of Mustroph et al. (2009). Briefly, two plants expressing 35S::HF-RPL18 were ground in fine powder in liquid nitrogen and polysomes were extracted following the Mustroph's protocol until polysomes elution, except that no RNAse was used. After elution in 300 μL elution buffer, RNAs were purified by adding 1,350 μL RLT buffer (provided with the RNeasy kit, Qiagen) and 675 μL of ethanol 100% to the 300 μL eluate. This mixture was loaded into a RNeasy mini spin column (Qiagen) and RNA extraction was carried out following the manufacturer's instructions.

In vitro transcription/translation

The *in vitro* transcription/translation reactions were performed according to the manufacturer's instruction (TNT coupled wheat germ extract systems; Promega). Short transcripts were amplified from gDNA, then we fused T7 promoter and HA tag. The *in vitro* transcription/translation reactions were performed on 1μg of DNA at 30°C for 90 min. Water was used as a control. List of primers used:

		name	Séquence 5'->3'
AtmiR159a	PCR1	Fwd T7 pt trans 159a	TAATACGACTCACTATAGGGTTCCAAAACATGACGTGG
		Rev HA pt trans 159a	TTAAGCGTAATCTGGAACATCGTATGGGTAAACATACTTAGAGAGAAGGAAAG
	PCR2	Fwd HA pt trans 159a	ATGTTCCAGATTACGCTTAAGGTTAATAATTAGGGTTCCTCC
		Rev pt trans 159a	TGAGAACGTGGATTAACAAAAA
	PCR3	Fwd T7 (159a)	TAATACGACTCACTATAGGGTTC

(*Continued on next page*)

Continued

		name	Séquence 5'->3'
		Rev pt trans 159a	TGAGAACGTGGATTAACAAAAA
AtmiR160a	PCR1	Fwd T7 pt trans 160a	TAATACGACTCACTATAGGGCATCCCACCCTTAATTGTTT
		Rev HA pt trans 160a	TTAAGCGTAATCTGGAACATCGTATGGGTACTGTTCTTGAAGCCTTAAATG
	PCR2	Fwd HA pt trans 160a	ATGTTCCAGATTACGCTTAACCCCAATTCCTCCACA
		Rev pt trans 160a	TACATATATAATACATATACATCTAC
	PCR3	Fwd T7 (160a)	TAATACGACTCACTATAGGGCAT
		Rev pt trans 160a	TACATATATAATACATATACATCTAC
AtmiR172b	PCR1	Fwd T7 pt trans 172b	TAATACGACTCACTATAGGGACTTGCACCTCTCACTC
		Rev HA pt trans 172b	TTAAGCGTAATCTGGAACATCGTATGGGTATTAAAAATATTTATTTATGAGATAATAGT
	PCR2	Fwd HA pt trans 172b	ATGTTCCAGATTACGCTTAAAATTAGATGCATTTATTGATATG
		Rev pt trans 172b	GTAACTTTTTCATATCAATAAATGC
	PCR3	Fwd T7 (172b)	TAATACGACTCACTATAGGGACT
		Rev pt trans 172b	GTAACTTTTTCATATCAATAAATGC
MtmiR171b	PCR1	T7-pttrans171b Fwd	TAATACGACTCACTATAGGGATTGGTCAAACATACATACA
		pttrans171b-HA Rev	TTAAGCGTAATCTGGAACATCGTATGGGTAAGATATATATACTATGTCTCTTTCA
	PCR2	HA-pttrans171b Fwd	ATGTTCCAGATTACGCTTAACAAGGAGAAATTCAGGATATTG
		pttrans171b Rev	CTTGCTTTCTTCATCACTGA
	PCR3	T7-pttrans171b-2 Fwd	TAATACGACTCACTATAGGGATT
		pttrans171b Rev	CTTGCTTTCTTCATCACTGA

Immunoblots

The *in vitro* transcription and translation reaction products were loaded and separated by SDS-PAGE (4%–20% precast gel; GE Healthcare, Chicago, IL) and transferred onto 0.1 μm nitrocellulose membrane (GE Healthcare) according to Lauressergues et al. (2015). Immunoblots were performed by using anti-HA primary antibodies (Sigma-Aldrich) at 1:1,000 (v/v) dilution. HRP-conjugated goat anti-rabbit igG (Sigma-Aldrich) was used as secondary antibody at 1:5,000 (v/v) dilution. HA-tagged translated peptides were revealed with the Clarity Max Western ECL Substrate (BioRad, Hercules, CA). It has to be noted that from one blot to another the time of exposure to visualize the HA-peptide was different, resulting in differences in the background intensity of blots.

Data analysis

Orthologous proteins between *A. thaliana* and *A. lyrata* were found using Orthomcl software (Rice et al., 2000). Nucleic acid sequences of orthologous genes, intergenic sequences, miORFs, and miRNAs were aligned using Needle (Rice et al., 2000). Alignments of miPEPs were performed using Multalin (Corpet, 1988).

FRET-FLIM

Preparation of leaf samples for FRET-FLIM experiments

Agroinfiltrated *N. benthamiana* leaf discs were fixed after 48 hours by vacuum infiltrating a paraformaldehyde solution (TBS 1× (Tris-HCl 25 mM pH7.5, NaCl 140 mM, KCl 3 mM), 4% (w/v) paraformaldehyde and 0.05 M $(CH_3)_2AsO_2Na$)) before incubation 20 min at 4°C. Discs were rinsed in TBS buffer for 10 min and then permeabilized 10 min at 37°C in a proteinase K solution (Tris-HCl 50 mM pH 7.5, NaCl 100 mM, EDTA 1 mM, SDS 0.5%, 200 μg/mL of proteinase K (Invitrogen)). Nucleic acid staining was performed by vacuum-infiltrating a 5-μM Sytox Orange (Invitrogen) solution diluted in TBS and incubating samples 30 min at room temperature in the dark. When RNAse treatment was performed, foliar discs were incubated 15 min at room temperature with 0.5μg/mL of RNAse A (Roche) before nucleic acid staining. Foliar discs were washed with and mounted on TBS before observations on an inverted microscope (Eclipse TE2000E, Nikon, Japan).

FRET/FLIM measurements

Fluorescence lifetime measurements were performed in time domain using a streak camera. The light source is a mode-locked Ti:-sapphire laser (Tsunami, model 3941, Spectra-Physics, Milpitas, CA) pumped by a 10W diode laser (Millennia Pro, Spectra-Physics) and delivering ultrafast femtosecond pulses of light with a fundamental frequency of 80 MHz. A pulse picker (model 3980, Spectra-Physics) is used to reduce the repetition rate to 2 MHz to satisfy the requirements of the triggering unit (working at 2 MHz). The experiments were carried out at $\lambda = 820$ nm (multiphoton excitation mode). All images were acquired with a 60× oil immersion lens (plan APO 1.4 N.A., IR) mounted on an inverted microscope (Eclipse TE2000E, Nikon, Japan). The fluorescence emission is directed back into the detection unit through a short pass filter $\lambda < 750$ nm) and a band pass filter (515/30 nm). The detector is a streak camera

(Streakscope C4334, Hamamatsu Photonics, Japan) coupled to a fast and high-sensitivity CCD camera (model C8800-53C, Hamamatsu). For each nucleus, average fluorescence decay profiles were plotted and lifetimes were estimated by fitting data with exponential function using a non-linear least-squares estimation procedure. Fluorescence lifetime of the donor (GFP) was experimentally measured in the presence and absence of the acceptor (Sytox Orange). FRET efficiency (E) was calculated by comparing the lifetime of the donor in the presence (τ_{DA}) or absence (τ_D) of the acceptor: $E = 1 - (\tau_{DA})/(\tau_D)$. Statistical comparisons between control (donor) and assay (donor + acceptor) lifetime values were performed by the Student *t* test. For each experiment, four leaf discs removed from two agroinfiltrated leaves were used to collect data.

ITC assays

High-performance liquid chromatography purified RNA samples (RNAmiORF171b2: 15 nt corresponding with the RNA sequence of miORF171b2 (ATGAAGATTGAAGAG, see Figure 6A), Sigma Aldrich) were resuspended in the ITC buffer (10 mM sodium phosphate pH 7, 150 mM NaCl and 10 mM $MgCl_2$). After heating to 70°C for 10 minutes, they were cooled progressively to 20°C to facilitate secondary structure formation. The 1D proton NMR spectra were recorded on a Bruker 600 MHz spectrometer between 20°C and 4°C to follow imino proton resonances. The two peptides were synthesized by using manual solid phase synthesis using Fmoc-amino acids solid phase chemistry. The purity of the final products was assessed by analytical reverse phase liquid chromatography and their integrity was checked by electro-spray mass spectrometry on a TSQ 700. ITC experiments were conducted at 20°C using an ITC200 instrument (Malvern Panalytical, Malvern, UK). The ITC experiments consisted of 20 injections of 2 μL of peptide into the RNA-containing thermostatic cell. Data were acquired and analyzed using fully automatic features in Microcal Origin Software.

QUANTIFICATION AND STATISTICAL ANALYSIS

Levels of expression for the controls were set at 100. The mean values of relative gene expression or GFP lifetime were compared by using Wilcoxon or Student *t* test, using R. Statistical details are reported in the figure legends. Error bars represent the standard error of the mean. The asterisks indicate significant differences ($p < 0.05$). n represents the number of independent samples.