

Rapid sequence and functional diversification of a miRNA superfamily targeting calcium signaling components in seed plants

Komal Attri¹ , Zijie Zhang¹ , Atinder Singh¹ , Robert A. Sharrock²  and Zhixin Xie¹ 

¹Department of Biological Sciences, Texas Tech University, Lubbock, TX 79409, USA; ²Department of Plant Sciences and Plant Pathology, Montana State University, Bozeman, MT 59717, USA

Author for correspondence:
Zhixin Xie
Email: zhixin.xie@ttu.edu

Received: 4 April 2022
Accepted: 20 April 2022

New Phytologist (2022) **235**: 1082–1095
doi: 10.1111/nph.18185

Key words: Arabidopsis, autoinhibited Ca²⁺-ATPase (ACA), calcium signaling, calmodulin-like (CML), EF hand (EF-h), miR391, posttranscriptional gene silencing (PTGS), spermatophytes.

Summary

- MicroRNA (miRNA)-directed posttranscriptional gene silencing (miR-PTGS) is an integral component of gene regulatory networks governing plant development and responses to the environment. The sequence homology between Sly-miR4376, a miRNA common to Solanaceae and reported to target *autoinhibited Ca²⁺-ATPase 10 (ACA10)* messenger RNA (mRNA) in tomato, and Arabidopsis miR391 (Ath-miR391), previously annotated as a non-conserved member of the deeply conserved miR390 family, has prompted us to revisit the function of Ath-miR391, as well as its regulatory conservation.
- A combination of genetic, molecular, and bioinformatic analyses revealed a hidden conservation for miR-PTGS of *ACA10* homologs in spermatophytes.
- We found that the Arabidopsis *ACA10* mRNA undergoes miR391-directed cleavage *in vivo*. Furthermore, transgenic overexpression of miR391 recapitulated the *compact inflorescence (cif)* phenotypes characteristic of *ACA10* loss-of-function mutants, due to miR391-directed PTGS of *ACA10*. Significantly, comprehensive data mining revealed robust evidence for widespread PTGS of *ACA10* homologs directed by a superfamily of related miRNAs sharing a conserved sequence core. Intriguingly, the *ACA*-targeting miRNAs in Poaceae also direct PTGS for calmodulin-like proteins which are putative Ca²⁺ sensors.
- The PTGS of *ACA10* homologs is therefore directed by a miRNA superfamily that is of ancient origin and has undergone rapid sequence diversification associated with functional innovation.

Introduction

MicroRNAs (miRNAs) are small regulatory RNAs expressed from defined genomic noncoding RNA loci in a distinct biogenesis pathway (Bologna & Voinnet, 2014; Song *et al.*, 2019). Mature miRNAs typically of 21- to 22-nucleotide (nt) in length act as sequence determinants for posttranscriptional gene silencing (PTGS) of homologous transcripts in both plants and animals. In plants, an essential role of miRNA-directed PTGS (miR-PTGS) in development is well established. First, Arabidopsis loss-of-function mutants defective in miRNA biogenesis or function are embryonic lethal (Schauer *et al.*, 2002; Vaucheret, 2008). Secondly, perturbation of a specific miRNA-target regulatory module, caused by either genetic lesions or transgenic manipulation often leads to profound phenotypic alterations as have been shown in numerous cases (Chen, 2009). Furthermore, interference of miR-PTGS by virus-encoded suppressors of RNA silencing (VSR) has been shown to largely account for viral disease symptoms in the infected hosts, which may be viewed as pleiotropic developmental defects that can often be recapitulated

by transgenic expression of a VSR (Burgan & Havelda, 2011). These observations are consistent with the fact that regulatory targets for some of the most conserved plant miRNAs include members of several plant-specific transcription factor (TF) families known to be integral parts of the regulatory networks governing development (Riechmann *et al.*, 2000; Jones-Rhoades *et al.*, 2006).

Several remarkable features are known to characterize the interaction between miRNAs and their targets in plants. First, the miRNA-target interaction exhibits high sequence specificity which often involves a near-perfect base-pairing between the two, with few mismatches (mm) tolerable only at certain positions (Bologna & Voinnet, 2014). This feature explains why a plant miRNA may have only a few cellular targets. Another remarkable feature is the target site location. Both protein-coding and non-coding transcripts may serve as miRNA targets. In cases of protein-coding targets, the miRNA-interacting site is commonly found within the coding region, with a few exceptions in which a target site is found in either 5' or 3' untranslated region (UTR) (Jones-Rhoades *et al.*, 2006). Additionally, the interaction

between a miRNA and its cellular target often results in a cleavage of the target, catalyzed by the ribonucleolytic activities of an Argonaute (AGO) family protein in miRNA-associated ribonucleoprotein complexes (Vaucheret, 2008). Intriguingly, in some cases the miRNA-directed cleavage triggers production of 21-nt phased, secondary small interfering RNAs (phasiRNAs) from the cleaved target in a distinct pathway that requires RNA-dependent RNA Polymerase 6 (RDR6) and Dicer-like 4 (DCL4) (Fei *et al.*, 2013). A 22-nt miRNA arising from a pre-miRNA harboring an asymmetric bulge within the miRNA:miRNA* duplex was shown to be both necessary and sufficient to trigger phasiRNA biogenesis (Chen *et al.*, 2010; Cuperus *et al.*, 2010), although the underlying mechanism remains unknown. Some of the phasiRNAs may direct PTGS of homologous transcripts that would otherwise not be directly targeted by the 'trigger' miRNA *per se*.

miR391, a 21-nt miRNA first identified in Arabidopsis and common to other Brassicaceae species but with no experimentally validated target, was annotated as a member of the deeply conserved miR390 family due to an apparent sequence similarity (Xie *et al.*, 2005). This annotation has been challenged by findings from biochemical studies in Arabidopsis. miR390 which bears a 5' terminal adenosine (5'-A) was shown to functionally associate with AGO7, whereas miR391 which bears a 5' terminal uridine (5'-U) was found to be mainly associated with AGO1 (Mi *et al.*, 2008; Montgomery *et al.*, 2008). Interestingly, Arabidopsis miR391 (Ath-miR391) also shares extensive sequence homology to a more recently identified miR4376 (Sly-miR4376) in tomato (*Solanum lycopersicum* L.), a 22-nt miRNA thought to be conserved only in Solanaceae (Wang *et al.*, 2011). Sly-miR4376 was shown to target the 5' UTR of messenger RNA (mRNA) encoding an autoinhibited Ca²⁺-ATPase (ACA) which is closely related to Arabidopsis ACA10, and elicit phasiRNA production (Wang *et al.*, 2011). It has been noticed that Ath-miR391, Sly-miR4376, and several other plant miRNAs including miR3627 and miR5225 in dicots and miR1432 in monocots all share partial sequence homology with miR390 and may be evolutionarily related (Xia *et al.*, 2013; Chavez Montes *et al.*, 2014).

The Ca²⁺-ATPases, also known as calcium-transporting pumps, are a core component of calcium signaling which is known to play an essential role in cellular life (Clapham, 2007). The Ca²⁺-permeable channels and Ca²⁺-ATPases represent two major players operating the influx and efflux of the Ca²⁺ ions, respectively, in defining the cytoplasmic Ca²⁺ levels. Key components of cellular calcium signaling also include a variety of Ca²⁺-sensing proteins that contain the characteristic calcium binding helix-loop-helix motifs known as the EF hand (EF-h), also dubbed as the professional Ca²⁺ chelator (Clapham, 2007; Gifford *et al.*, 2007). In addition to calmodulins (CaM) which are well known Ca²⁺ sensors, plant genomes encode a remarkable repertoire of calmodulin like proteins (CMLs) which contain a diverse number of EF-h motifs and may also function as Ca²⁺ sensors (Zhu *et al.*, 2015). Of note, an ACA family Ca²⁺-transporter is autoinhibited under low cytosolic Ca²⁺ concentrations. Upon elevated Ca²⁺ levels, CaM-binding to the

N-terminal cytosolic regulatory domain of a plant ACA releases the autoinhibition and activates its ATP-dependent Ca²⁺-transporting activity (Dodd *et al.*, 2010). Consistent with a universally essential role of calcium signaling, genetic studies have revealed an important role of ACA family members including the Arabidopsis ACA10 (AtACA10) in plant development, immunity, and responses to the environment (Dodd *et al.*, 2010; Yang *et al.*, 2017). We have been particularly intrigued by the presence of a putative miR391 target site in the 5' UTR of AtACA10 mRNA, and several CMLs reported as the predicted targets for miR1432 in rice (*Oryza sativa* L.) (Lu *et al.*, 2008; Sunkar *et al.*, 2008) and maize (*Zea mays* L.) (Zhang *et al.*, 2009). These observations have prompted us to examine the unknown conservation for miR-PTGS of calcium signaling components in plants.

Here we report Ath-miR391-directed PTGS of ACA10 in Arabidopsis. Moreover, comprehensive mining of small RNA data and expressed sequence tags (ESTs) further revealed robust evidence for a previously unrecognized conservation for miR-PTGS of ACA family members in seed plants. Our data also revealed a lineage-specific acquisition of CML targets for ACA-targeting miRNAs in Poaceae, a regulatory innovation in miR-PTGS which may have been implicated in the rapid sequence diversification of the ACA-targeting miRNAs.

Materials and Methods

Plant materials and growth conditions

The Arabidopsis *compact inflorescence 1 (cif1-1)* and *aca10-1* mutants in Nossen (No-0) genetic background have been previously described (George *et al.*, 2008). Arabidopsis plants including wild-type accessions of Columbia (Col-0), Landsberg erecta (Ler), and No-0 ecotypes, as well as mutants and transgenic lines were grown in commercial soil mix (SunGrow LC-1) in a growth chamber with a daily cycle of 16 h light at 24°C and 8 h dark at 22°C. Rice (*Oryza sativa* L. ssp. *japonica* cv. Nipponbare) seedlings were grown in bench top containers filled with the same soil mix and kept in a glasshouse.

Mapping internal cleavage of mRNA

Mapping of miRNA-directed cleavage in the predicted mRNA targets was done following the modified RNA ligase-mediated rapid amplification of complementary DNA (cDNA) 5' end (RLM-5' RACE) procedure (Llave *et al.*, 2002), using either poly (A)⁺-enriched (Arabidopsis Col-0) or total RNA extracts (rice) prepared from young seedlings. Sequence information for both gene-specific primers and adapter-specific primers used in RLM-5' RACE is provided in Supporting Information Table S1.

Plasmid constructs

A set of three constructs featuring a length variation, and inclusion or exclusion of a vector-derived 'leader' sequence, in the precursor transcripts of *MIR391* transgene were made. Targeted regions of the endogenous *MIR391* locus were PCR amplified

from genomic DNA extracted from wild-type *Arabidopsis* (Col-0) using DNeasy Plant Mini Kit (Qiagen). A short (S), 149 bp *MIR391* fragment was amplified using primer pairs with an engineered *EcoR*I or *Nco*I restriction site in the forward primer, and a *Bam*H I site in the reverse primer to facilitate the subsequent cloning into the pRTL2 vector (Restrepo *et al.*, 1990). Similarly, a longer (L), 225 bp extended *MIR391* fragment was also amplified with engineered *Nco*I and *Bam*H I sites at the 5' and 3' end, respectively (see Table S1 for primer sequences). Notice that designation of the constructs incorporated the restriction sites (*EcoR*I, *Nco*I, and *Bam*H I) used for cloning and the length (Short or Long) of the insert (Fig. 2a). Thus the construct harboring the 149 bp *MIR391* fragment inserted using *EcoR*I and *Bam*H I was designated as *35S::MIR391_{EB}*, or *EB* for short. Likewise, those harboring either the 149 bp or the 225 bp *MIR391* fragment inserted using *Nco*I and *Bam*H I were designated as *35S::MIR391_{NBS}* and *35S::MIR391_{NBL}* (or *NBS* and *NBL* for short), respectively (Fig. 2a). The expression cassettes were then excised from pRTL2 using *Pst*I and inserted into the binary vector pCB302 (Xiang *et al.*, 1999).

Arabidopsis transformation and analysis of transgenic lines

Transgenic *Arabidopsis* overexpressing miR391 was made in both No-0 and Ler ecotypes by *Agrobacterium*-mediated transformation using the vacuum infiltration method (Clough & Bent, 1998). Seeds from primary transformants were grown in soil under selection for BASTA (phosphinothricin) resistance. Putative transgenic T1 plants were further confirmed for the presence of transgene by a diagnostic PCR from genomic DNA (see Table S1 for primer information). Seeds from selected transgenic lines were collected for advanced generations. For each of the three constructs, and only for those in No-0 genetic background, progenies (T2 and T3) from multiple independent T1 lines were grown for phenotypic and expression analysis. Nomenclature of the transgenic lines followed the designation of corresponding constructs.

Gene expression analysis

Analysis of miRNA expression was done by small RNA (sRNA) blot assays using radio-labeled oligonucleotide probes, essentially as described (Xie, 2010). Modifications included preparation of low molecular weight (LMW) RNA extracts from rosette stage seedlings using mirPremier microRNA isolation kit (Sigma), and a lower amount (3.25 µg LMW RNA per lane) of sample loading.

Radioactive signals from a probed blot were first captured in a storage phosphor screen, then visualized and analyzed on a phosphorimager (Bio-Rad, Hercules, CA, USA). For analysis of mRNA levels by reverse transcription quantitative polymerase chain reaction (RT-qPCR) or RT-semi-qPCR, total RNA extracts were prepared from rosette stage *Arabidopsis* seedlings (T2) using Spectrum plant total RNA kit (Sigma), followed by DNase treatment using TURBO DNA-free kit (ThermoFisher Scientific, Waltham, MA, USA). Up to 4.5 µg of total RNA was

used in each oligo d(T)-primed RT reaction using the SuperScript™ III first-strand synthesis system (ThermoFisher Scientific). A parallel set of RT reactions with no reverse transcriptase added (control RT reactions) were also run as a negative control, which was used in the subsequent semi-qPCR assays for evaluation of residual genomic DNA contamination. *ACTIN1* (At2g37620) was used as an internal control. The qPCR reactions with PowerTrack™ SYBR green master mix (Applied Biosystems, Waltham, MA, USA) were run using the ABI 7500 system. *PROTEIN PHOSPHATASE 2A SUBUNIT A3* (*PP2AA3*; At1g13320) was used as an internal control (Czechowski *et al.*, 2005).

Sequence collection and computational analysis

Sequence information for mature miRNAs and associated precursors was mainly collected from miRBase (Kozomara *et al.*, 2019) and the sRNAanno-associated database (Chen *et al.*, 2021). Additional miRNA sequence information was also collected from other published works for *Brassica oleracea* var. *italica* (Li *et al.*, 2017), *Chrysanthemum morifolium* (Xia *et al.*, 2015), *Eucommia ulmoides* (Wang *et al.*, 2016), *Eutrema halophilum* (Zhang *et al.*, 2013), *Hordeum vulgare* (Schreiber *et al.*, 2011), *Monotropa hypopitys* (Shchennikova *et al.*, 2016), *Nelumbo nucifera* (Hu *et al.*, 2016), *Picea glauca* (Liu & El-Kassaby, 2017), *Solanum lycopersicum* (Wang *et al.*, 2011), *Solanum melongena* (Yang *et al.*, 2013), and *Triticum aestivum* (Pandey *et al.*, 2014; Ma *et al.*, 2015).

Identification of ESTs for ACA and EF-h/CML targets were done by reiterative BLAST search using either cognate mature miRNA or homologous target sequences as query at the National Center for Biotechnology Information's (NCBI's) web interface.

ESTs encoding putative miRNA-targeted ACA family members were generally recovered from species for which a dedicated effort for full-length cDNA sequencing has been made. Examples for some of the published effort include *Brachypodium distachyon* (Mochida *et al.*, 2013), *Brassica oleracea* var. *viridis* (Araki *et al.*, 2013), *Cucumis melo* (Clepet *et al.*, 2011), *Eucalyptus camaldulensis* (Hirakawa *et al.*, 2011), *Eucommia ulmoides* (Suzuki *et al.*, 2012), *Hordeum vulgare* (Tanaka *et al.*, 2013), *Medicago sativa* (Aziz *et al.*, 2005), *Populus nigra* (Nanjo *et al.*, 2007), *Solanum pennellii* (McDowell *et al.*, 2011), and *Eutrema halophilum* (formerly *Thellungiella halophila*) (Taji *et al.*, 2008).

The miRNA-target sequence complementarity was assessed using a position-weighted scoring matrix known as TargetFinder (Fahlgren & Carrington, 2010), which typically yields TF scores ranging from 0 to 3.5 for authentic miRNA targets. For multiple sequence alignment (MSA)-based conservation analysis involving noncoding (i.e. 5' UTR) sequences, sections of sequences with a defined length and features (e.g. 5' UTR or open reading frame (ORF)) were first manually compiled from the collected entries as the raw input sequences. Subsections of sequences with desired lengths were then computationally extracted using extractalign (<https://www.bioinformatics.nl/cgi-bin/emboss/extractalign>), one of the open web-based tools provided by the European Bioinformatics Institute (EMBL-EBI), for use as input sequences. MSA

was executed using the T-Coffee web server (Tommaso *et al.*, 2011). MSA of cDNA sequences encoding the EF-h/CML family members was done using TranslatorX (Abascal *et al.*, 2010), a tool for DNA sequence alignment guided by the corresponding amino acid sequence. Pictograms were generated using the WebLogo online server (Crooks *et al.*, 2004). Whenever applicable, presentation of miRNA and ACA target sequence information in the Supporting Information tables generally followed the Angiosperm Phylogeny Group classification (Byng *et al.*, 2016).

Results

AtACA10 mRNA undergoes miR391-directed cleavage *in vivo*

Based on the apparent sequence similarity, Ath-miR391 was considered a member of the deeply conserved miR390 family in plants (Xie *et al.*, 2005). However, substantial sequence similarity also exists between Ath-miR391 and the Sly-miR4376 (Fig. 1a), raising the possibility of Ath-miR391 being an ortholog of Sly-miR4376. Such a proposal, although consistent with the presence of a putative target site in the 5' UTR of AtACA10 mRNA (Fig. 1b), lacked experimental substantiation. As a first step to address this issue, we mined evidence for Ath-miR391-directed cleavage of AtACA10 mRNA using existing datasets known as degradome or PARE libraries (Addo-Quaye *et al.*, 2008; German *et al.*, 2008). Specifically, numerous Arabidopsis degradome libraries from sequence read archive (SRA) repository at NCBI were searched for predicted ACA10-derived tags. This effort yielded inconclusive results due to either none or very low reads

that were indistinguishable from the background (data not shown). We reasoned that degradome tags derived from either 5' UTR or 5' proximal coding region could be intrinsically under-represented due to oligo (dT)-primed reverse transcription often involved in library preparation. We therefore resorted to the conventional RLM-5' RACE for a closer check. A control 5' RACE reaction for miR171-directed cleavage of *Scarecrow-like27* (*SCL27*) mRNA (Llave *et al.*, 2002) yielded the expected 472 bp cDNA fragment (Fig. 1c, lane 2). The ACA10-specific reaction gave rise to a 166 bp fragment seen as a distinct sharp band in an agarose gel, along with some more diffused, larger-sized products likely resulted from nonspecific amplifications (Fig. 1c, lane 1). Sequencing of the 5' RACE clones mapped the cleavage site to 14-nt upstream of the AUG codon (Fig. 1b), as predicted for a miR391-directed cleavage.

Transgenic overexpression of miR391 recapitulated the compact inflorescence phenotypes characteristic of the AtACA10 loss-of-function mutants

To demonstrate a specific causal relationship for miR391-PTGS of ACA10 in Arabidopsis, we examined the outcomes of perturbed miR391 expression through transgenic manipulation. Wild-type Arabidopsis of No-0 ecotype was chosen for this purpose, due to the well characterized ACA10 loss-of-function mutant phenotypes known as *compact inflorescence* (*cif*) observed specifically in this genetic background (George *et al.*, 2008). A set of three binary vectors, each containing a cauliflower mosaic virus (CaMV) 35S promoter-driven *MIR391* expression cassette that differs slightly from each other in the transcription unit, were

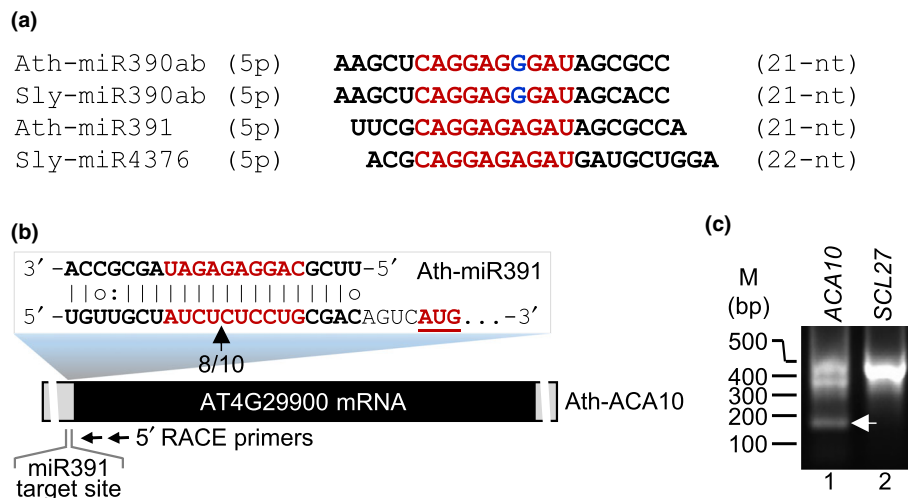


Fig. 1 miR391 directs cleavage of ACA10 mRNA at the 5' UTR in Arabidopsis. (a) Aligned mature miRNA sequences of Arabidopsis (Ath) miR391, tomato (Sly) miR4376, and the deeply conserved miR390 from both species. A 10-nt segment shared by Ath-miR391 and Sly-miR4376, which differs only by 1-nt from that of miR390, is shown in red. (b) A schematic diagram of AtACA10 mRNA showing the predicted miR391-interacting site in its 5' UTR. Shown on top is the sequence alignment between miR391 and its predicted target site. The miR391-directed cleavage, which mapped to 14-nt upstream of the AUG codon (shown in red and underlined) using RLM-5' RACE, is indicated by a vertical arrowhead. The number of sequenced 5' RACE clones confirming the cleavage site is shown underneath the arrowhead. The two horizontal arrowheads represent a set of two ACA10-specific primers used in the 5' RACE. (c) A section of agarose gel image showing detection of the ACA10-specific 5' RACE fragment (lane 1, indicated by a horizontal arrowhead). A reaction specific to miR171-directed cleavage of *AtSCL27* mRNA was included in the assay to serve as a positive control (lane 2). The positions of DNA size markers are shown on the left.

constructed for *Arabidopsis* transformation (Fig. 2a). For control purposes, transgenic plants for each of the three constructs were generated in both No-0 and Ler genetic backgrounds, as disruption of *ACA10* function is not expected to cause noticeable phenotypic changes in Ler or Col-0 as it would in No-0 (Goosey & Sharrock, 2001; George *et al.*, 2008).

While primary transformants (T1) in Ler background exhibited no noticeable phenotypic differences from the wild-type

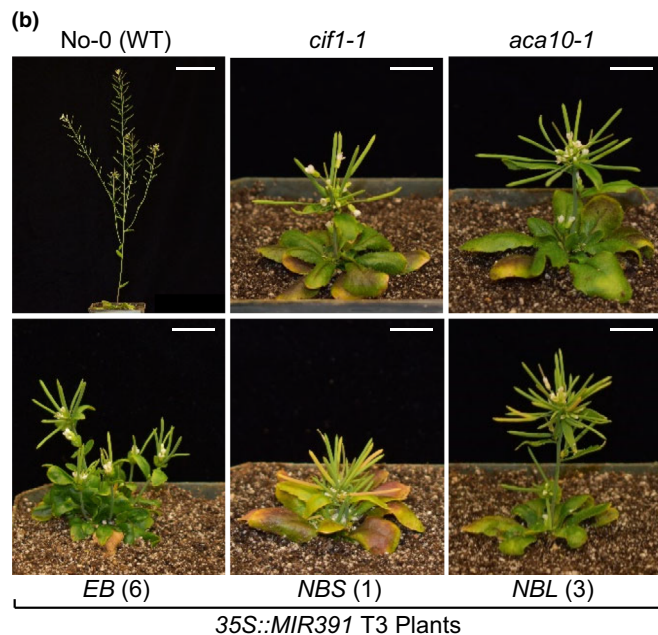
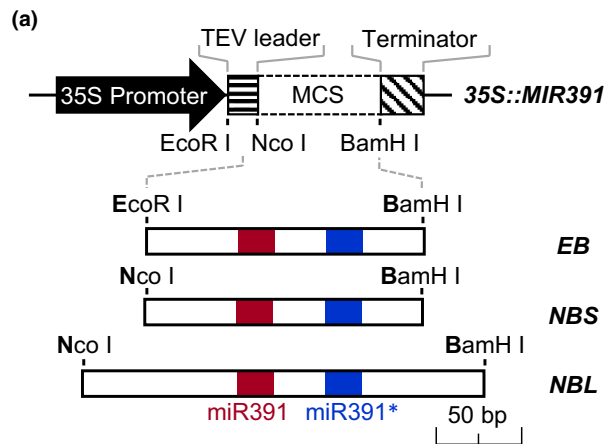


Fig. 2 Transgenic overexpression of miR391 in *Arabidopsis* Nossen-0 (No-0) genetic background recapitulated the phenotypes of the loss-of-function *ACA10* mutant known as *compact inflorescence 1* (*cif1*). (a) Schematic diagrams showing a set of three binary constructs for transgenic overexpression of miR391 driven by cauliflower mosaic virus (CaMV) 35S promoter. These constructs will give rise to transcripts containing either a 149 bp (*EB*, *NBS*) or 225 bp (*NBL*) *Ath-MIR391* precursor sequence, with (*NBS*, *NBL*) or without (*EB*) a 133 bp vector-derived TEV-leader. Designation of the constructs incorporated the restriction sites (*EcoR* I; *Nco* I; and *BamH* I) used for cloning and the length (Short or Long) of the insert. MCS, multiple cloning site. (b) Photographs of representative transgenic lines (T3) along with the wild-type and mutant controls showing the compact inflorescence phenotypes. The scale bar in the No-0 image is 5 cm, and 1 cm in all others.

control plants (data not shown), most of those in No-0 background displayed a range of interesting phenotypes reminiscent of the *cif1* mutants, which is consistent with an effective knockdown of *ACA10* expression due to the transgenic overexpression of miR391. The *cif* phenotypes were consistently observed in the T2 and T3 plants propagated from the selected T1 lines (Figs 2b, S1). Of note, we did not observe substantial phenotypic differences among the transgenic lines of *35S::MIR391_{EB}*, *35S::MIR391_{NBS}*, and *35S::MIR391_{NBL}*, both in terms of the percentage of plants that displayed the *cif*-like phenotypes, or the spectrum of *cif*-like phenotypes observed in each line (Figs 2b, S1). A morphometric survey on mature T2 and T3 plants also provided a more quantitative measurement in indicating the phenotypic similarities between the transgenic lines and the *cif1-1* and *aca10-1* mutant plants grown under same conditions (Fig. S1b,c).

To confirm that the observed *cif*-like phenotypes were indeed attributable to specific knockdown of *ACA10* due to miR391-PTGS, the expression levels of miR391 and *ACA10* in transgenic plants were examined. In sRNA northern blot assays, comparable low levels of miR391 signal, seen as a faint band corresponded to a 21-nt marker position, was detected from each of the two wild-type controls, Col-0 and No-0, respectively (Fig. 3a, lanes 1 and 2 of the two upper panels). Signals of comparable intensity were also detected for the *cif1-1* and *aca10-1* mutants (Fig. 3a, lanes 3 and 4). In each of the selected transgenic lines, however, miR391 accumulated to substantially elevated levels when compared with the wild-type controls (Fig. 3a, lanes 5–10 in the upper panels). These data indicate that effective transgenic overexpression of miR391 was achieved with each of the three transgene constructs. Overexpression of miR391 did not appear to perturb the expression of unrelated endogenous miRNAs, as the levels of miR167 accumulation in transgenic lines were comparable to those observed in the nontransgenic wild-type controls or the mutant lines (Fig. 3a, middle panel).

Specific miR391-PTGS of *ACA10* in transgenic *Arabidopsis*

Expression of *ACA10* was assessed by RT-semi-qPCR or RT-qPCR. In RT-semi-qPCR assays, the *ACA10*-specific band with substantially reduced intensity was seen from each of the transgenic samples (Fig. S2, lanes 4–9 of the top panel). RT-qPCR assays also showed consistent results. Compared with the wild-type control, the relative expression levels of *ACA10* in the *cif1-1* and *aca10-1* mutants were found to be 75.3% and 121.3%, respectively (Fig. 3b), which was in general agreement with previous reports (George *et al.*, 2008). Significantly reduced levels of *ACA10* expression were observed in the transgenic lines, which ranged from 23.1% to 31.3% of the wild-type control (Fig. 3b), confirming an effective miR391-mediated knockdown of *ACA10* expression, as manifested by the phenotypic observations.

As an indirect assay for potential miR391-induced perturbation of miR390 targets, the expression of *Auxin Response Factor 3* (*ARF3*) and *ARF4*, two known targets of miR390-dependent *TAS3*-derived *trans*-acting small interfering RNAs (ta-siRNAs) (Allen *et al.*, 2005) was analyzed. Should miR391 be able to act

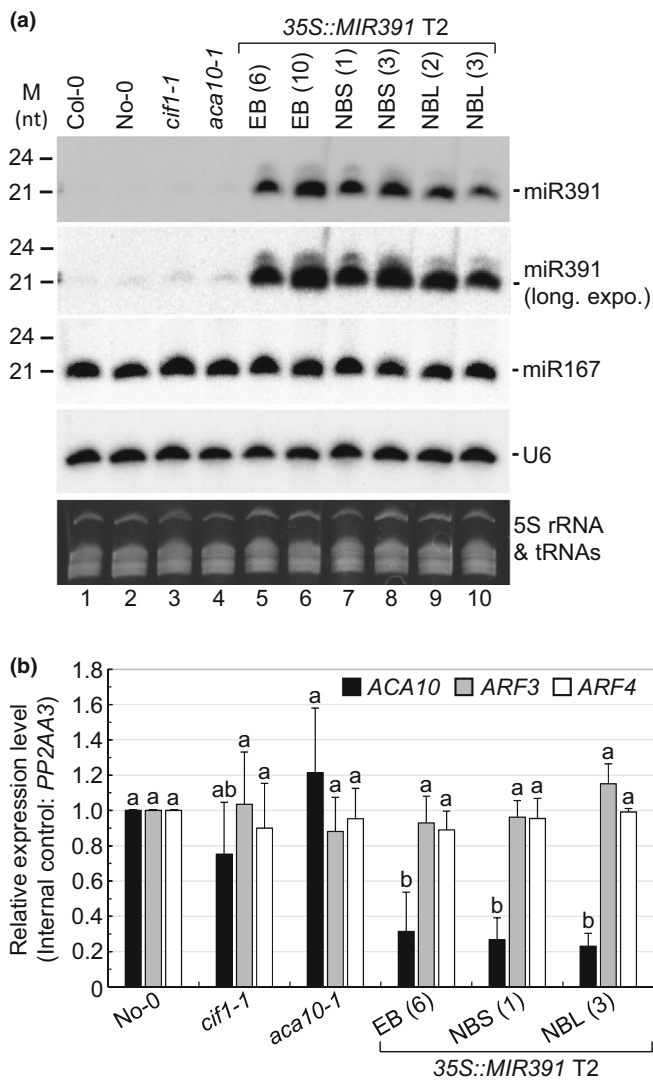


Fig. 3 Expression analysis for miR391 and *ACA10* in 35S::MIR391 transgenic Arabidopsis. (a) sRNA northern blot assays for miR391. Images shown (from top to the bottom) are sRNA blot probed for miR391, miR167, and U6 snRNA, respectively, along with an ethidium bromide-stained gel section corresponding to the zone of 5S rRNA and tRNAs. (b) RT-qPCR assays for *ACA10*, *ARF3* and *ARF4* with *PROTEIN PHOSPHATASE 2A SUBUNIT A3 (PP2AA3; At1g13320)* as an internal control. Each plotted data point represents the mean value calculated from three biological replicates, each assayed with three technical replicates, with an error bar representing the standard deviation. Data points labeled with different alphabets denote statistically significant differences ($P < 0.05$) upon pairwise comparison.

on *TAS3* transcripts, the direct targets of miR390, and promote the biogenesis of *TAS3*-derived ta-siRNAs (*TAS3*-siRNAs), an enhanced knockdown of *ARF3* and *ARF4* expression would have occurred upon transgenic overexpression of miR391. However, results from both RT-semi-qPCR and RT-qPCR assays showed that neither *ARF3* nor *ARF4* exhibited significantly altered expression in transgenic lines compared with the nontransgenic wild-type control or the *cif1-1* and *aca10-1* mutants (Figs 3b, S2, middle panels). These data argue against a regulatory interaction between miR391 and *TAS3* transcripts.

Evidence for widespread miRNA-directed PTGS of ACA family members in seed plants

Our data presented so far strongly indicated that Ath-miR391 is an ortholog of Sly-miR4376, although the two mature miRNAs differ by 1-nt in size and share only a 14-nt identity in alignment (Fig. 1a). These observations, in combination with those made by others (Xia *et al.*, 2013; Chavez Montes *et al.*, 2014), raised the interesting possibility that Ath-miR391 and Sly-miR4376 may represent members of an ACA-targeting miRNA superfamily that has undergone rapid sequence diversification. A closer inspection of several Sly-miR4376-related miRNAs including miR3627 and miR5225 from other dicots, and miR1432 from monocots, revealed a highly conserved 10-nt core (5'-CAGGAGAGAU-3') in mature miRNA sequences. Interestingly, a homologous 10-nt segment is also found in the deeply conserved miR390 family, but differs consistently by 1-nt, with the A in the seventh position replaced by a G (5'-CAGGAGGGAU-3'; Fig. 1a). A search in the public domains using the 10-nt core as query recovered putative members of the ACA-targeting miRNA superfamily from 105 species representing 40 families (Table S2). This collection included many entries that are either new or previously reported miRNAs with diverse designations. Significantly, putative members of the ACA-targeting miRNA superfamily were also found in *Amborella trichopoda*, a basal angiosperm species, and gymnosperm species including ginkgo (*Ginkgo biloba*) and the Norway spruce (*Picea abies*) (Table S2).

If ACA-targeting miRNAs have undergone rapid sequence diversification in association with functional innovation, compensatory sequence changes would be expected in the miRNA-interacting site of cognate ACA targets in diverse lineages, if the miR-PTGS of ACA family members has been maintained. We therefore examined the presence of putative miRNA target site in cDNAs encoding ACA family members homologous to the Ath- and Sly-*ACA10*. To avoid potential pitfalls of recovering false positives due to annotation-associated artifacts, which could occur in short sequence reads-based genomic or transcriptomic assembly, we focused our analysis on the publicly accessible ESTs. This effort collectively identified 79 ACA-specific EST accessions (52 distinct sequences) as putative targets of miR391 superfamily members, which represented 60 species from 23 families (Table S3). The cross-species sequence conservation in the 5' UTR of the putative miRNA-targeted ACA family members was analyzed by MSA. In addition to a nonredundant set of ESTs representing distinct ACA family members, alternative cDNA entries (e.g. 5' RACE clones from (Wang *et al.*, 2011), or cDNAs from annotated reference genomes) were included in some cases for an enhanced phylogenetic representation of the analysis. Strikingly, the sequence corresponding to the putative miRNA-interacting site stood out as the only highly conserved motif other than the ATG start codon in the MSA for ACA sequences in a 129 bp region (120 bp 5' UTR sequence proximal to the start codon, plus the first 9 bp of the coding sequence) analyzed. These features were easily identifiable in an MSA-based heatmap and were further illustrated by the consensus residues which are indicated with asterisks at the bottom of the aligned sequences

(Fig. 4a). Of note, 10 of the 12 most conserved residues identified within the putative miRNA target site corresponded (in reverse complement) to the 10-nt core found in putative ACA-targeting miRNAs (Fig. 4a).

The conserved nature of the putative miRNA-target site within the 5' UTR of mRNAs encoding ACA family members in diverse plant species is indicative of biological relevance as well as potential functional significance of the sequence motif. However, cross-species variations in target site sequences (notably toward the 5' end, upstream of the consensus 10-nt core), as well as their positioning relative to the start codon were readily appreciable from the MSA-based heat map (Fig. 4a).

Compensatory sequence changes at the miRNA-interacting site of the cognate ACA targets were evident, as a near-perfect sequence complementarity characteristic of canonical miRNA:target interaction in plants was observed for all represented species in which such an analysis is possible (see the TF scores in Fig. 4a). Specifically, sequence alignment between a member of the proposed miR391 superfamily and the predicted ACA target is shown for rice (Fig. 4b) and *Amborella trichopoda* (Fig. 4d), the two representative species that occupied the bottom portion of the heatmap (Fig. 4a). Remarkably, the predicted miR1432-directed cleavage of the OsACA mRNA was supported by independent degradome datasets (Fig. 4c), although our attempt of 5' RACE validation with total RNA extracts prepared from young seedling was unsuccessful (data not shown). It remains possible that the use of total RNA instead of poly (A)⁺-enriched fraction may have limited the sensitivity for RLM-5' RACE detection of the predicted cleavage at the 5' UTR of the OsACA mRNA.

Evidence for lineage-specific non-ACA targets of the ACA-targeting miRNA superfamily

Several EF-h/CMLs were previously reported as the predicted targets for miR1432 in rice (*O. sativa* L.) (Lu *et al.*, 2008; Sunkar *et al.*, 2008) and maize (*Z. mays* L.) (Zhang *et al.*, 2009). We wondered if miR1432, a member of the ACA-targeting miRNA superfamily acts as a dual-targeting miRNA in these and perhaps other grass species. In our EST-focused analysis for ACA targets, EST accessions encoding EF-h/CML family members were also recovered from multiple Poaceae species as putative targets of miR1432. The cross-species conservation of putative miR1432 target site in the EF-h/CML-encoding cDNAs was analyzed using MSA with a total of 13 ESTs and three annotated cDNAs representing seven species including rice, maize, barley (*Hordeum vulgare*), wheat (*Triticum aestivum*), sorghum (*Sorghum bicolor*), wild oats (*Avena barbata*), and stiff brome (*Brachypodium distachyon*) (Table S4). MSA using TRANSLATORX (Abascal *et al.*, 2010), a tool for amino acid (aa) sequence-guided DNA sequence alignment, revealed the putative miR1432 target site as a block of sequence with outstanding conservation, as shown in a pictogram for a 30-nt window (Fig. 5a). The target site cDNA sequence is associated with codons for 8 aa residues near the N-terminus of the encoded EF-h/CML proteins (Fig. 5a). Protein-based MSA also identified up to two putative Ca²⁺-binding EF-h/CML

motifs as conserved blocks in the translated cDNA sequences (Fig. S3).

Focusing on rice, RLM-5' RACE was used to validate the predicted miR1432-directed cleavage in an EF-h/CML mRNA (Fig. 5b). A control 5' RACE reaction for miR160-directed cleavage of *OsARF18* (Os06g47150) mRNA (Zhou *et al.*, 2010) yielded the expected 255 bp cDNA fragment (Fig. 5c, lane 1). The *OsEF-h/CML*-specific reaction gave rise to a 248 bp fragment seen as a bright band in an agarose gel (Fig. 5c, lane 2). Sequencing of the 5' RACE clones mapped the cleavage to the predicted site which is 52-nt downstream from the AUG codon (Fig. 5a). This result represents the very first experimental validation for miR1432-directed cleavage of an EF-h/CML-encoding mRNA in grass species. Significantly, it also serves as a solid piece of experimental evidence for lineage-specific acquisition of novel target by members of the proposed miR391 superfamily in Poaceae.

As an independent validation for functional interaction between miR1432 and predicted targets in grass species, we analyzed numerous publicly available degradome sequencing datasets for sequence tags with a perfect match to the predicted EF-h/CML targets. Indeed, miR1432-directed cleavage of EF-h/CML mRNAs was generally supported by degradome data in species for which high quality data are available (Table S4). Notably, available degradome datasets provided strong support for miR1432-directed cleavage of EF-h/CML mRNAs in rice (Fig. 5d), which is consistent with the 5' RACE data, as well as in barley and wheat, as manifested by a high signal-to-noise ratio in the region of interest (Figs 5d, S4). Significantly, the same degradome datasets also provided robust support for miR1432-directed cleavage of the predicted ACA-targets in these species (Fig. S4), indicating that miR1432 family members are truly dual-targeting miRNAs in grasses.

Discussion

We found that Arabidopsis miR391 directs PTGS of ACA10, a central component of Ca²⁺ signaling important for plant development, immunity, and responses to the environment. We then gathered evidence showing that miR-PTGS of ACA family members represents conserved regulatory modules likely common to seed plants. We further showed that the ACA-targeting miR1432 family in grasses also directs PTGS of mRNAs encoding members of the conserved EF-h/CML family proteins thought to function as cellular Ca²⁺ sensors. Collectively, these findings give rise to an interesting emerging picture for conserved miR-PTGS of calcium signaling components in plants directed by members of a rapidly evolving miRNA superfamily.

The miR391-directed PTGS of ACA10 in Arabidopsis

We showed that the AtACA10 mRNA undergoes an internal cleavage *in vivo*, precisely at the predicted miR391 target site within the 5' UTR proximal to the AUG codon. This observation is indicative of active operation of miR391-PTGS of ACA10 in Arabidopsis. Transgenic overexpression of miR391 resulted in

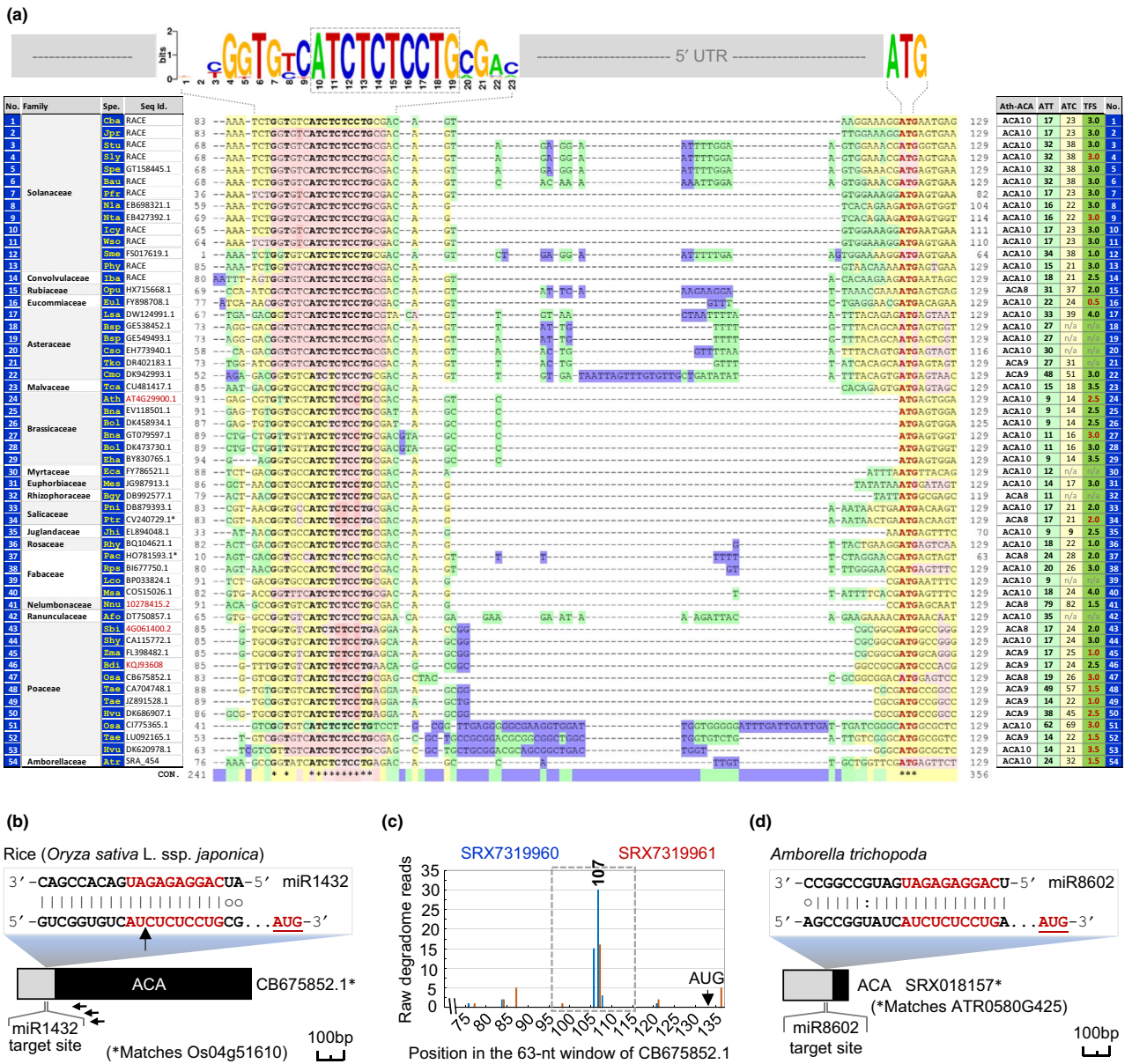


Fig. 4 Evidence for widespread conservation of miRNA-directed PTGS of ACA family members in plants. (a) Conservation of putative miRNA-interacting site in the 5' UTR of ACA mRNAs in diverse species. Shown at the center is a section of MSA-based heatmap for cDNAs encoding ACA family members identified as putative targets of miR391 superfamily members. A segment of up to 129 bp (120 bp of 5' UTR sequence proximal to the ATG, plus the first 9 bp of the ORF) from each of the 54 distinct sequence entries was analyzed. Some of the entries have less than 129 bp 5' UTR sequence available. The residue numbers for each of the aligned sequences are shown on both sides of the heatmap. Consensus residues are indicated by asterisks at the bottom. The conserved nature of the putative miRNA-interacting site spanning a 23 bp window within the aligned 5' UTR is also presented as a pictogram on the top, with the conserved 10-nt core boxed. The family names, three-letter abbreviation of species names, along with accession numbers are shown on the far-left, whereas additional ACA target information is shown on the far-right. Nonclassic EST entries include RACE (the 5' RACE clones from Wang *et al.* (2011)), annotated cDNAs (shown in red), and one SRA read. Accessions with an asterisk (*) are reverse complementary to the coding strand. Ath-ACA: top hit in BLASTx for Arabidopsis ACA homologs; ATT (distance from ATG to the 10-nt core): the number of nt between the ATG and the 3' terminal base of the 10-nt core. ATC (distance from ATG to the cleavage site): the number of nt between the ATG and the predicted miRNA-directed cleavage site. A TargetFinder score (TFS) shown in red indicates degradome data support. Additional information for redundant (with overlapping sequences) EST accessions, ESTs of putative ACA targets with a good TFS but having a 1- or 2-nt mismatch in the conserved 10-nt core, as well as full species names are found in the Supporting Information Table S3. (b) A schematic diagram of a representative ACA-encoding rice EST with the predicted miR1432-interacting site at 5' UTR. Sequence alignment between miR1432 and the predicted target site is shown on the top, with a vertical arrowhead indicating the miRNA-directed cleavage site supported by degradome data. The three horizontal arrowheads represent a set of gene-specific primers used in the 5' RACE. (c) A bar graph constructed by plotting degradome sequencing data against the rice EST sequence. Raw reads of degradome sequence tags from each of two selected datasets (color-coded in red and blue, respectively, with the SRA accession numbers shown on the top) were plotted against a 63-nt window of the EST sequence centered by the 21-nt miRNA target site (boxed with dotted lines). A numerical figure above the peak reads denotes the mapped position for 5' end of the degradome tags. The position of the start codon is also indicated. (d) A schematic diagram of a representative ACA-encoding *Amborella trichopoda* EST (454 platform; see Table S3) with the predicted miRNA-interacting site at 5' UTR.

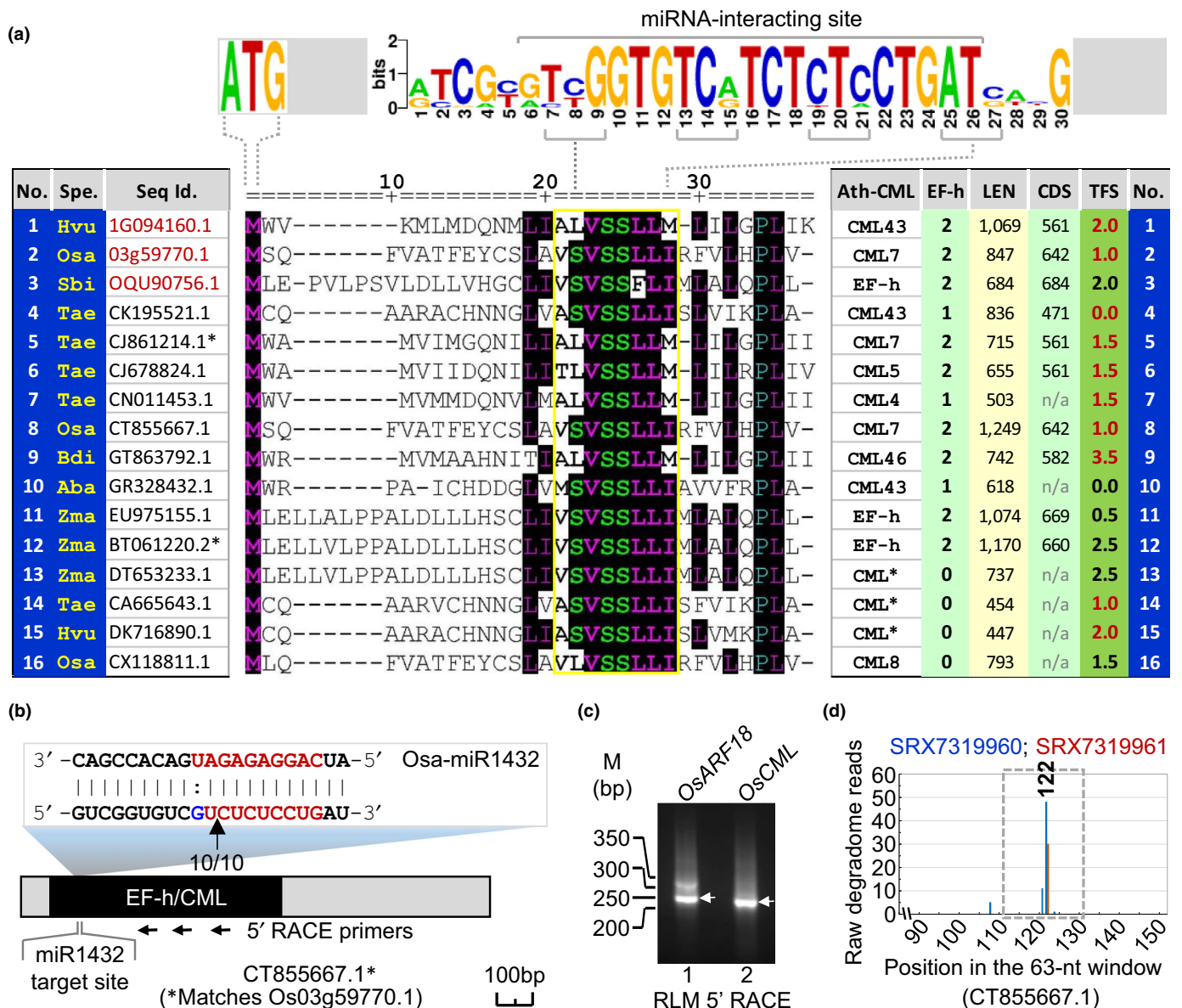


Fig. 5 The miRNA-directed PTGS of EF-h/CML family members in Poaceae. (a) Amino acid (aa) sequence-based MSA analysis identified the segment of peptide encoded by the putative miRNA target site sequence as a highly conserved N-terminus proximal block in the EF-h/CML family members. The residue numbers are indicated on top of the aligned sequence blocks. A pictogram generated based on aa-sequence-guided MSA of cDNA sequences is shown on the top, which illustrates an outstanding sequence conservation at the putative miR1432 target site within a 30 bp window of the EF-h/CML-encoding ESTs from multiple grass species. The three-letter species names along with accession numbers are shown on the far-left, whereas additional CML target information is shown on the far-right. Annotated cDNA accessions are shown in red. Accessions with an asterisk (*) are reverse complementary to the coding strand. Ath-CML: top hit in BLASTX for Arabidopsis homologs. CML* denotes homology identified in BLASTX at NCBI. EF-h: number of conserved EF-h motifs detected (see Supporting Information Fig. S3); LEN and CDS denote the length (bp) of the EST and CDS, respectively. A TargetFinder score (TFS) shown in red indicates degradome data support. Additional information including redundant (with overlapping sequences) ESTs, and full species names are found in Table S4. (b) A schematic diagram of an EF-h/CML-encoding rice EST as a predicted miR1432 target. Sequence alignment between miR1432 and the predicted target site is shown on the top, with a vertical arrowhead indicating the miRNA-directed cleavage sites validated by RLM-5' RACE. The number of sequenced 5' RACE clones confirming the cleavage site is shown underneath the arrowhead. The three horizontal arrowheads represent a set of gene-specific primers used in the 5' RACE. (c) A section of agarose gel image showing detection of the EF-h/CML-specific 5' RACE fragment (lane 2, indicated by a horizontal arrowhead). A reaction specific to miR160-directed cleavage of *OsARF18* (*Os06g47150*) mRNA was included in the assay to serve as a positive control (lane 1). The positions of DNA size markers are shown on the left. (d) A bar graph constructed by plotting degradome reads against the rice EST sequence. Raw reads of degradome tags from each of two selected datasets (color-coded in red and blue, respectively) were plotted against a 63-nt window of the EST sequence centered by the 21-nt miRNA target site (boxed with dotted lines). A numerical figure above the peak reads denotes the mapped position for 5' end of the degradome tags.

profound phenotypic alterations in *Arabidopsis*, most notably in the inflorescence architecture, which essentially recapitulated the *cif* phenotypes characteristic of *AtACA10* loss-of-function mutants in No-0 genetic background (George *et al.*, 2008). Expression analysis confirmed substantially elevated levels of miR391 accumulation in the transgenic plants, concomitant with substantially reduced (up to a 77% reduction) levels of *ACA10* mRNA accumulation. Transgenic overexpression of miR391 did not seem to interfere with the expression of unrelated endogenous miRNAs, as evidenced by miR167 for which no altered expression was observed in any of the transgenic lines examined (Fig. 3a). Such a nonspecific global interference would have occurred if introduction of the transgene (*MIR391* in this case) had resulted in a saturation or an overload of the miRNA pathway machinery, which appears to be unlikely and rarely reported. Importantly, our data did not show any significant change in the expression levels of *ARF3* and *ARF4*, two known targets of *TAS3*-siRNAs, in response to miR391 overexpression (Figs 3b, S2). This observation provided indirect evidence against possible involvement of miR391 in biogenesis of *TAS3*-siRNAs, a process known to be initiated by miR390-directed cleavage of *TAS3* precursor transcripts (Allen *et al.*, 2005). Taken together, these data unambiguously established specific miR391-directed PTGS of *ACA10* in *Arabidopsis* that is distinct from the miR390-dependent, *TAS3*-siRNAs-mediated PTGS of *ARF3* and *ARF4*.

Conservation of miR-PTGS of ACA family members in seed plants

Confirmation of miR391-directed PTGS of *ACA10* in *Arabidopsis* immediately brought up questions on the origin and evolution of the ACA-targeting miRNAs in plants. First and foremost, are *Arabidopsis* miR391, Sly-miR4376, and the homologous miR1432 in grass species evolutionarily related? In a common origin or divergent evolution scenario, how deeply conserved are the ACA-targeting miRNAs in plants? Or, alternatively, do they represent miRNAs that arose through independent, lineage-specific events? Our sequence homology-based search for both putative miRNAs that resemble Ath-miR391, Sly-miR4376, or monocot miR1432 and cognate ACA targets revealed a previously unappreciated level of conservation for miR-PTGS of ACA family members in plants. The common 5' UTR location of target site proximal to the AUG codon, and the highly conserved 10-nt core shared among the putative ACA-targeting miRNAs strongly indicate a common evolutionary origin of the miRNA-ACA regulatory modules identified across a wide spectrum of plant species. Of note, the depth of EST-based analysis on putative ACA targets suffered from the generally limited availability of EST accessions that contain adequate 5' UTR sequence information. In fact, it is interesting to notice that identification of candidate ESTs encoding putative miRNA-targeted ACA family members was biased to species for which a dedicated effort focusing on full-length cDNA identification has been made. Nonetheless, the EST-based ACA target data gathered in this study is consistent with the existence of conserved miRNA-ACA regulatory modules across angiosperm lineages. It is obvious that there was a substantially unbalanced

availability between sRNA and EST datasets for many of the sampled species. This was particularly true for ferns and gymnosperm species where conventional EST data were scarce. Based on the available sRNA data from *Picea abies* and *G. biloba*, it seems safe to conclude that the miRNA-ACA regulatory module also operates in gymnosperms, although the phylogenetic placement of *G. biloba* remains a matter of debate. Our data therefore collectively indicate that the ACA-targeting miRNAs are most likely conserved across seed plants. We propose to consider the related ACA-targeting miRNAs with the conserved 10-nt core as members of the miR391 superfamily.

Sequence and functional diversification in ACA-targeting miRNAs

Having proposed a common origin for ACA-targeting miRNAs in seed plants, several outstanding questions relevant to plant miRNA evolution remain unanswered. In addition to the cross-species variation in mature miRNA sequences, another notable variation among the ACA-targeting miRNAs is the size of mature miRNAs. Specifically, while a subgroup of the proposed superfamily members (which include miR391 in Brassicaceae and miR1432 in Poaceae) represent mature miRNAs of the most common 21-nt species, others exemplified by the Sly-miR4376 represent the 22-nt species known to be functionally associated with biogenesis of target-derived phasiRNAs (Wang *et al.*, 2011) (Fig. 6). Although an asymmetric bulge within the miRNA:miRNA* duplex of a stem-loop precursor has been shown to be critical to give rise to a 22-nt miRNA capable of triggering phasiRNA biogenesis (Chen *et al.*, 2010; Cuperus *et al.*, 2010), what drove the emergence of, or the selection for such an asymmetric bulge in *MIRNA* loci remains poorly understood (Fig. 6). An obvious potential advantage of a 22-nt miR-PTGS system would be a rapid, phasiRNA-mediated destruction of related (e.g. paralogous) target transcripts that otherwise could not be directly targeted by the 22-nt miRNA *per se*. Consistent with the target gene duplication model for *MIRNA* gene evolution (Allen *et al.*, 2004), a recent study has revealed a good correlation between species expressing members of the miR482 superfamily, which are predominantly 22-nt species, and those contain large numbers of *NBS-LRR* (for nucleotide-binding site leucine-rich repeat) genes in their genomes (Gonzalez *et al.*, 2015). In this regard, the proposed miR391 superfamily with an enriched cross-species diversity in mature miRNA sequence, as well as functionally associated variations in mature miRNA size, could be a fertile ground for future study focusing on the evolutionary drive underlying the emergence of 22-nt miRNAs and associated phasiRNAs (Fig. 6).

The initial designation of ACA-targeting miRNAs from diverse plant species as distinct families reflects the unusual extent of sequence variations found in the mature miRNAs. Several factors may have contributed to the observed extensive sequence diversification. First, whereas most plant miRNAs target the coding region or ORF of mRNAs (Jones-Rhoades *et al.*, 2006), the ACA-targeting miRNAs which share a common 5' UTR location of target sites represent a minority of plant miRNAs that target

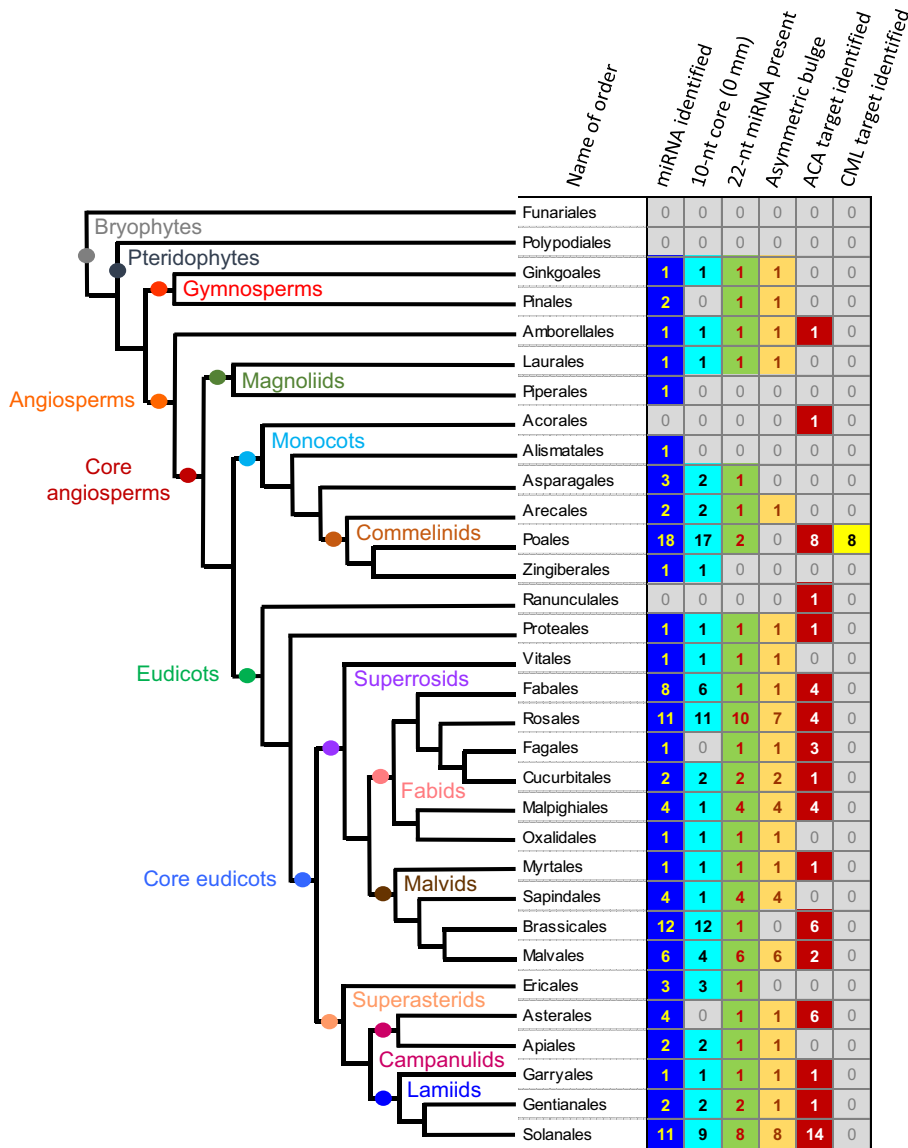


Fig. 6 A phylogenetic map for miRNA-directed PTGS (miR-PTGS) of calcium signaling components in seed plants. An order-based phylogenetic tree showing the number of species in each order for which components of miR-PTGS for either ACA or EF-h/CML family members have been identified. Data presented here (columns from left to right) include: (1) number of species with members of the ACA-targeting miRNA superfamily identified; (2) number of species with ACA-targeting miRNAs containing the conserved 10-nt core sequence without a mismatch; (3) number of species with 22-nt ACA-targeting miRNA present in sRNA-seq data; (4) number of species with 22-nt ACA-targeting miRNA supported by an asymmetric bulge presented in the folded pre-miRNA; (5) number of species with putative ACA target identified; (6) number of species with putative EF-h/CML target identified. Construction of phylogenetic tree followed the *Angiosperm Phylogeny Group* classification (Byng *et al.*, 2016).

noncoding sequences. The noncoding nature of the target site sequence involved in miRNA-target interaction may have permitted sequence diversification through compensatory mutations in both the miRNA and the target sequences due to the absence of coding constraints. Such a 'compensatory mutation' hypothesis could also explain the cross-species offsets in mature miRNA sequences observed among members of the proposed ACA-targeting miRNA superfamily. Second, the lineage-specific acquisition of EF-h/CML targets may have limited the sequence diversification of mature miR1432 in Poaceae, as the ORF location of miRNA target site in CML mRNAs could have imposed coding constraints on the target site sequence. The dual-targeting Poaceae miR1432 and the ACA-targeting miRNAs in other seed plant lineages could therefore have been evolving under different sequence constraints.

Acquisition of EF-h/CML targets in Poaceae represents a regulatory innovation for miR1432 (Fig. 6), a member of the proposed miR391 superfamily. Although the functional

significance for the emergence of dual-targeting miR1432 in Poaceae remains to be examined, two recent reports collectively showed an important regulatory role for rice miR1432 in both grain development and defense response to the fungal pathogen *Magnaporthe oryzae* (Zhao *et al.*, 2019; Li *et al.*, 2021). It is worth noting that although both studies involved transgenic manipulation of miR1432, neither has considered the putative ACA targets, most likely due to a target search effort focusing exclusively on coding sequences. A role for the miR1432-CML regulatory module in rice immunity was further supported by directly manipulating the expression of the miR1432-targeted CML, with higher levels of CML expression associated with an enhanced disease resistance (Li *et al.*, 2021). We noticed that evidence for *OsACOT* (acyl-CoA thioesterase) being a miR1432 target, as presented in Zhao *et al.* (2019), is generally weak. Furthermore, our attempt of RLM-5' RACE in rice yielded a negative result for *OsACOT* (Table S1, and data not shown).

Of note, some plant miRNAs are known to have dual targets of genes involved in the same metabolic pathway. Members of the miR395 family in Arabidopsis, for instance, are known to target mRNAs of both AST68 (also known as 'SULTR2;1', a low-affinity sulfate transporter) (Allen *et al.*, 2005) and APS (isoforms of ATP sulfurylases) (Jones-Rhoades & Bartel, 2004), proteins involved in two distinct steps of sulfur assimilation. Intriguingly, current data including those from transgenic manipulation of miR395 expression support a role for miR395 in coordinated regulation of sulfur translocation and assimilation in response to changing conditions of sulfur supply (Liang *et al.*, 2010; Kawashima *et al.*, 2011). The EF-h/CML proteins are thought to function as cellular Ca²⁺ sensors which undergo conformational changes upon binding to Ca²⁺ ions (Zhu *et al.*, 2015). A few EF-h/CML proteins in plants have been subjected to functional analysis. The Ca²⁺-binding activity for Arabidopsis CML43, as well as its tomato homolog, have been shown biochemically using recombinant proteins (Chiasson *et al.*, 2005; Bender *et al.*, 2014). Interestingly, data from expression analysis and transgenic manipulation have collectively indicated a role for Arabidopsis CML43 and its tomato homolog as mediators of Ca²⁺-dependent signaling in plant immune response to bacterial pathogens (Chiasson *et al.*, 2005). Biochemical evidence also supports Arabidopsis CML7 as a functional Ca²⁺ sensor (Trande *et al.*, 2019). Considering our current understanding of cellular signaling involving CaM or CML proteins as Ca²⁺ sensors and the Ca²⁺-dependent activation of ACA activities, it will be interesting to see if the miRNA-regulated CML and ACA operate in the same signaling pathway in Poaceae.

In conclusion, this work unambiguously establishes miR391-directed PTGS of ACA10 in Arabidopsis and uncovers previously unrecognized conservation for miR-PTGS of ACA family members in seed plants. The data presented here collectively indicate conserved miR-PTGS of core calcium signaling components, directed by members of a miRNA superfamily that is of ancient origin and has undergone rapid sequence diversification associated with functional innovation. The PTGS of calcium signaling components by the miR391 superfamily therefore represents an interesting case of miR-PTGS in plants with a remarkable evolutionarily dynamic feature. Considering the central role of calcium signaling in cellular life, this work is anticipated to stimulate future investigations on regulatory innovation and functional adaptation of ACA-targeting miRNAs associated with sequence diversification in diverse plant lineages.






Acknowledgements

The authors thank the Center for Biotechnology and Genomics at Texas Tech University (TTU) for access to the real-time PCR instrument, and Mohamed Fokar for excellent technical assistance. The authors also thank Benildo de los Reyes for providing rice seeds. This work was supported in part by TTU through an Office of Research & Innovation program.

Author contributions

KA performed experiments, analyzed data, and contributed to preparation of figures and tables. ZZ and AS performed experiments and analyzed data. RAS provided mutants and analyzed data. ZX conceived the research plan, supervised experiments, analyzed data, and wrote the manuscript with feedback from all authors.

ORCID

Komal Attri  <https://orcid.org/0000-0001-5645-8414>
Robert A. Sharrock  <https://orcid.org/0000-0002-0194-0534>
Atinder Singh  <https://orcid.org/0000-0002-5543-9139>
Zhixin Xie  <https://orcid.org/0000-0002-6191-5170>
Zijie Zhang  <https://orcid.org/0000-0003-4051-7189>

Data availability

The data that support the findings of this study are openly available at NCBI.

References

- Abascal F, Zardoya R, Telford MJ. 2010. TRANSLATORX: multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Research* 38: W7–13.
- Addo-Quaye C, Eshoo TW, Bartel DP, Axtell MJ. 2008. Endogenous siRNA and miRNA targets identified by sequencing of the Arabidopsis degradome. *Current Biology* 18: 758–762.
- Allen E, Xie Z, Gustafson AM, Sung GH, Spatafora JW, Carrington JC. 2004. Evolution of microRNA genes by inverted duplication of target gene sequences in *Arabidopsis thaliana*. *Nature Genetics* 36: 1282–1290.
- Allen E, Xie Z, Gustafson AM, Carrington JC. 2005. microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell* 121: 207–221.
- Araki R, Hasumi A, Nishizawa OI, Sasaki K, Kuwahara A, Sawada Y, Totoki Y, Toyoda A, Sakaki Y, Li Y *et al.* 2013. Novel bioresources for studies of *Brassica oleracea*: identification of a kale MYB transcription factor responsible for glucosinolate production. *Plant Biotechnology Journal* 11: 1017–1027.
- Aziz N, Paiva NL, May GD, Dixon RA. 2005. Transcriptome analysis of alfalfa glandular trichomes. *Planta* 221: 28–38.
- Bender KW, Dobney S, Ogunrinde A, Chiasson D, Mullen RT, Teresinski HJ, Singh P, Munro K, Smith SP, Snedden WA. 2014. The calmodulin-like protein CML43 functions as a salicylic-acid-inducible root-specific Ca²⁺ sensor in Arabidopsis. *Biochemical Journal* 457: 127–136.
- Bologna NG, Voinnet O. 2014. The diversity, biogenesis, and activities of endogenous silencing small RNAs in Arabidopsis. *Annual Review of Plant Biology* 65: 473–503.
- Burgan J, Havelda Z. 2011. Viral suppressors of RNA silencing. *Trends in Plant Science* 16: 265–272.
- Byng JW, Chase MW, Christenhusz MJM, Fay MF, Judd WS, Mabberley DJ, Sennikov AN, Soltis DE, Soltis PS, Stevens PF *et al.* 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society* 181: 1–20.
- Chavez Montes RA, de Fatima R-C, De Paoli E, Accerbi M, Rymarquis LA, Mahalingam G, Marsch-Martinez N, Meyers BC, Green PJ, de Folter S. 2014. Sample sequencing of vascular plants demonstrates widespread conservation and divergence of microRNAs. *Nature Communications* 5: 3722.
- Chen C, Li J, Feng J, Liu B, Feng L, Yu X, Li G, Zhai J, Meyers BC, Xia R. 2021. sRNAanno—a database repository of uniformly annotated small RNAs in plants. *Horticulture Research* 8: 45.

- Chen HM, Chen LT, Patel K, Li YH, Baulcombe DC, Wu SH. 2010. 22-Nucleotide RNAs trigger secondary siRNA biogenesis in plants. *Proceedings of the National Academy of Sciences, USA* 107: 15269–15274.
- Chen XM. 2009. Small RNAs and Their Roles in Plant Development. *Annual Review of Cell and Developmental Biology* 25: 21–44.
- Chiasson D, Ekengren SK, Martin GB, Dobney SL, Snedden WA. 2005. Calmodulin-like proteins from Arabidopsis and tomato are involved in host defense against *Pseudomonas syringae* pv. *tomato*. *Plant Molecular Biology* 58: 887–897.
- Clapham DE. 2007. Calcium signaling. *Cell* 131: 1047–1058.
- Clepet C, Joobeur T, Zheng YI, Jublot D, Huang M, Truniger V, Boualem A, Hernandez-Gonzalez ME, Dolcet-Sanjuan R, Portnoy V *et al.* 2011. Analysis of expressed sequence tags generated from full-length enriched cDNA libraries of melon. *BMC Genomics* 12: 252.
- Clough SJ, Bent AF. 1998. Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal* 16: 735–743.
- Crooks GE, Hon G, Chandonia JM, Brenner SE. 2004. WebLogo: a sequence logo generator. *Genome Research* 14: 1188–1190.
- Cuperus JT, Carbonell A, Fahlgren N, Garcia-Ruiz H, Burke RT, Takeda A, Sullivan CM, Gilbert SD, Montgomery TA, Carrington JC. 2010. Unique functionality of 22-nt miRNAs in triggering RDR6-dependent siRNA biogenesis from target transcripts in Arabidopsis. *Nature Structural & Molecular Biology* 17: 997–1003.
- Czechowski T, Stitt M, Altmann T, Udvardi MK, Scheible WR. 2005. Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant Physiology* 139: 5–17.
- Dodd AN, Kudla J, Sanders D. 2010. The language of calcium signaling. *Annual Review of Plant Biology* 61: 593–620.
- Fahlgren N, Carrington JC. 2010. miRNA target prediction in plants. *Methods in Molecular Biology* 592: 51–57.
- Fei Q, Xia R, Meyers BC. 2013. Phased, secondary, small interfering RNAs in posttranscriptional regulatory networks. *Plant Cell* 25: 2400–2415.
- George L, Romanowsky SM, Harper JF, Sharrock RA. 2008. The ACA10 Ca²⁺-ATPase regulates adult vegetative development and inflorescence architecture in Arabidopsis. *Plant Physiology* 146: 716–728.
- German MA, Pillay M, Jeong D-H, Hetawal A, Luo S, Janardhanan P, Kannan V, Rymarquis LA, Nobuta K, German R *et al.* 2008. Global identification of microRNA-target RNA pairs by parallel analysis of RNA ends. *Nature Biotechnology* 26: 941–946.
- Gifford JL, Walsh MP, Vogel HJ. 2007. Structures and metal-ion-binding properties of the Ca²⁺-binding helix-loop-helix EF-hand motifs. *Biochemical Journal* 405: 199–221.
- Gonzalez VM, Muller S, Baulcombe D, Puigdomenech P. 2015. Evolution of *NBS-LRR* gene copies among Dicot plants and its regulation by members of the miR482/2118 superfamily of miRNAs. *Molecular Plant* 8: 329–331.
- Goosey L, Sharrock R. 2001. The Arabidopsis compact inflorescence genes: phase-specific growth regulation and the determination of inflorescence architecture. *The Plant Journal* 26: 549–559.
- Hirakawa H, Nakamura Y, Kaneko T, Isobe S, Sakai H, Kato T, Hibino T, Sasamoto S, Watanabe A, Yamada M *et al.* 2011. Survey of the genetic information carried in the genome of *Eucalyptus camaldulensis*. *Plant Biotechnology* 28: 471–480.
- Hu J, Jin J, Qian Q, Huang K, Ding Y. 2016. Small RNA and degradome profiling reveals miRNA regulation in the seed germination of ancient eudicot *Nelumbo nucifera*. *BMC Genomics* 17: 684.
- Jones-Rhoades MW, Bartel DP. 2004. Computational identification of plant MicroRNAs and their targets, including a stress-induced miRNA. *Molecular Cell* 14: 787–799.
- Jones-Rhoades MW, Bartel DP, Bartel B. 2006. MicroRNAs and their regulatory roles in plants. *Annual Review of Plant Biology* 57: 19–53.
- Kawashima CG, Matthewman CA, Huang S, Lee B-R, Yoshimoto N, Koprivova A, Rubio-Somoza I, Todesco M, Rathjen T, Saito K *et al.* 2011. Interplay of SLIM1 and miR395 in the regulation of sulfate assimilation in Arabidopsis. *The Plant Journal* 66: 863–876.
- Kozomara A, Birgaonu M, Griffiths-Jones S. 2019. miRBase: from microRNA sequences to function. *Nucleic Acids Research* 47: D155–D162.
- Li H, Wang Y, Wu M, Li L, Jin C, Zhang Q, Chen C, Song W, Wang C. 2017. Small RNA sequencing reveals differential miRNA expression in the early development of Broccoli (*Brassica oleracea* var. *italica*) Pollen. *Frontiers in Plant Science* 8: 404.
- Li Y, Zheng Y-P, Zhou X-H, Yang X-M, He X-R, Feng Q, Zhu Y, Li G-B, Wang HE, Zhao J-H *et al.* 2021. Rice miR1432 fine-tunes the balance of yield and blast disease resistance via different modules. *Rice* 14. doi: 10.1186/s12284-021-00529-1.
- Liang G, Yang F, Yu D. 2010. MicroRNA395 mediates regulation of sulfate accumulation and allocation in *Arabidopsis thaliana*. *The Plant Journal* 62: 1046–1057.
- Liu Y, El-Kassaby YA. 2017. Global analysis of small RNA dynamics during seed development of *Picea glauca* and *Arabidopsis thaliana* populations reveals insights on their evolutionary trajectories. *Frontiers in Plant Science* 8. doi: 10.3389/fpls.2017.01719.
- Llave C, Xie Z, Kasschau KD, Carrington JC. 2002. Cleavage of *Scarecrow-like* mRNA targets directed by a class of *Arabidopsis* miRNA. *Science* 297: 2053–2056.
- Lu C, Jeong D-H, Kulkarni K, Pillay M, Nobuta K, German R, Thatcher SR, Maher C, Zhang L, Ware D *et al.* 2008. Genome-wide analysis for discovery of rice microRNAs reveals natural antisense microRNAs (nat-miRNAs). *Proceedings of the National Academy of Sciences, USA* 105: 4951–4956.
- Ma X, Xin Z, Wang Z, Yang Q, Guo S, Guo X, Cao L, Lin T. 2015. Identification and comparative analysis of differentially expressed miRNAs in leaves of two wheat (*Triticum aestivum* L.) genotypes during dehydration stress. *BMC Plant Biology* 15: 21.
- McDowell ET, Kapteyn J, Schmidt A, Li C, Kang JH, Descour A, Shi F, Larson M, Schillmiller A, An L *et al.* 2011. Comparative functional genomic analysis of *Solanum* glandular trichome types. *Plant Physiology* 155: 524–539.
- 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.
- Mochida K, Uehara-Yamaguchi Y, Takahashi F, Yoshida T, Sakurai T, Shinozaki K. 2013. Large-scale collection and analysis of full-length cDNAs from *Brachypodium distachyon* and integration with Pooidae sequence resources. *PLoS ONE* 8: e75265.
- Montgomery TA, Howell MD, Cuperus JT, Li DW, Hansen JE, Alexander AL, Chapman EJ, Fahlgren N, Allen E, Carrington JC. 2008. Specificity of ARGONAUTE7-miR390 interaction and dual functionality in *TAS3* transacting siRNA formation. *Cell* 133: 128–141.
- Nanjo T, Sakurai T, Totoki Y, Toyoda A, Nishiguchi M, Kado T, Igasaki T, Futamura N, Seki M, Sakaki Y *et al.* 2007. Functional annotation of 19,841 Populus nigra full-length enriched cDNA clones. *BMC Genomics* 8: 448.
- Pandey R, Joshi G, Bhardwaj AR, Agarwal M, Katiyar-Agarwal S. 2014. A comprehensive genome-wide study on tissue-specific and abiotic stress-specific miRNAs in *Triticum aestivum*. *PLoS ONE* 9: e95800.
- Restrepo MA, Freed DD, Carrington JC. 1990. Nuclear transport of plant potyviral proteins. *Plant Cell* 2: 987–998.
- Riechmann JL, Heard J, Martin G, Reuber L, Jiang CZ, Keddie J, Adam L, Pineda O, Ratcliffe OJ, Samaha RR *et al.* 2000. Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. *Science* 290: 2105–2110.
- Schauer SE, Jacobsen SE, Meinke DW, Ray A. 2002. *DICER-LIKE1*: blind men and elephants in *Arabidopsis* development. *Trends in Plant Science* 7: 487–491.
- Schreiber AW, Shi BJ, Huang CY, Langridge P, Baumann U. 2011. Discovery of barley miRNAs through deep sequencing of short reads. *BMC Genomics* 12. doi: 10.1186/1471-2164-12-129.
- Shchennikova AV, Beletsky AV, Shulga OA, Mazur AM, Prokhortchouk EB, Kochieva EZ, Ravin NV, Skryabin KG. 2016. Deep-sequence profiling of miRNAs and their target prediction in *Monotropa hypopitys*. *Plant Molecular Biology* 91: 441–458.
- Song XW, Li Y, Cao XF, Qi YJ. 2019. MicroRNAs and their regulatory roles in plant-environment interactions. *Annual Review of Plant Biology* 70: 489–525.
- Sunkar R, Zhou X, Zheng Y, Zhang W, Zhu JK. 2008. Identification of novel and candidate miRNAs in rice by high throughput sequencing. *BMC Plant Biology* 8: 25.

- Suzuki N, Uefuji H, Nishikawa T, Mukai Y, Yamashita A, Hattori M, Ogasawara N, Bamba T, Fukusaki E-I, Kobayashi A *et al.* 2012. Construction and analysis of EST libraries of the trans-polyisoprene producing plant, *Eucommia ulmoides* Oliver. *Planta* 236: 1405–1417.
- Taji T, Sakurai T, Mochida K, Ishiwata A, Kurotani A, Totoki Y, Toyoda A, Sakaki Y, Seki M, Ono H *et al.* 2008. Large-scale collection and annotation of full-length enriched cDNAs from a model halophyte, *Thellungiella halophila*. *BMC Plant Biology* 8: 115.
- Tanaka T, Sakai H, Fujii N, Kobayashi F, Nakamura S, Itoh T, Matsumoto T, Wu J. 2013. bex-db: bioinformatics workbench for comprehensive analysis of barley-expressed genes. *Breeding Science* 63: 430–434.
- Tommaso P, Moretti S, Xenarios I, Orobitch M, Montanyola A, Chang JM, Taly JF, Notredame C. 2011. T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic Acids Research* 39: W13–W17.
- Trande M, Pedretti M, Bonza MC, Di Matteo A, D'Onofrio M, Dominici P, Astegno A. 2019. Cation and peptide binding properties of CML7, a calmodulin-like protein from *Arabidopsis thaliana*. *Journal of Inorganic Biochemistry* 199: 110796.
- Vaucheret H. 2008. Plant ARGONAUTS. *Trends in Plant Science* 13: 350–358.
- Wang L, Du HY, Wuyun TN. 2016. Genome-wide identification of MicroRNAs and their targets in the leaves and fruits of *Eucommia ulmoides* using high-throughput sequencing. *Frontiers in Plant Science* 7. doi: 10.3389/fpls.2016.01632
- Wang Y, Itaya A, Zhong X, Wu Y, Zhang J, van der Knaap E, Olmstead R, Qi Y, Ding B. 2011. Function and evolution of a MicroRNA that regulates a Ca²⁺-ATPase and triggers the formation of phased small interfering RNAs in tomato reproductive growth. *Plant Cell* 23: 3185–3203.
- Xia R, Meyers BC, Liu Z, Beers EP, Ye S, Liu Z. 2013. MicroRNA superfamilies descended from miR390 and their roles in secondary small interfering RNA Biogenesis in Eudicots. *Plant Cell* 25: 1555–1572.
- Xia X, Shao Y, Jiang J, Du X, Sheng L, Chen F, Fang W, Guan Z, Chen S. 2015. MicroRNA expression profile during aphid feeding in chrysanthemum (*Chrysanthemum morifolium*). *PLoS ONE* 10: e0143720.
- Xiang C, Han P, Lutziger I, Wang K, Oliver DJ. 1999. A mini binary vector series for plant transformation. *Plant Molecular Biology* 40: 711–717.
- Xie Z. 2010. Piecing the puzzle together: genetic requirements for miRNA biogenesis in *Arabidopsis thaliana*. *Methods in Molecular Biology* 592: 1–17.
- Xie Z, Allen E, Fahlgren N, Calamar A, Givan SA, Carrington JC. 2005. Expression of *Arabidopsis* MIRNA genes. *Plant Physiology* 138: 2145–2154.
- Yang D-L, Shi Z, Bao Y, Yan J, Yang Z, Yu H, Li Y, Gou M, Wang S, Zou B *et al.* 2017. Calcium pumps and interacting BON1 protein modulate calcium signature, stomatal closure, and plant immunity. *Plant Physiology* 175: 424–437.
- Yang L, Jue DW, Li W, Zhang RJ, Chen M, Yang Q. 2013. Identification of MiRNA from eggplant (*Solanum melongena* L.) by small RNA deep sequencing and their response to *Verticillium dahliae* infection. *PLoS ONE* 8: e72840.
- Zhang L, Chia JM, Kumari S, Stein JC, Liu Z, Narechania A, Maher CA, Guill K, McMullen MD, Ware D. 2009. A genome-wide characterization of microRNA genes in maize. *PLoS Genetics* 5: e1000716.
- Zhang Q, Zhao C, Li M, Sun W, Liu Y, Xia H, Sun M, Li A, Li C, Zhao S *et al.* 2013. Genome-wide identification of *Thellungiella salsuginea* microRNAs with putative roles in the salt stress response. *BMC Plant Biology* 13: 180.
- Zhao YF, Peng T, Sun HZ, Teotia S, Wen HL, Du YX, Zhang J, Li JZ, Tang GL, Xue HW *et al.* 2019. miR1432-OsACOT (Acyl-CoA thioesterase) module determines grain yield via enhancing grain filling rate in rice. *Plant Biotechnology Journal* 17: 712–723.
- Zhou M, Gu L, Li P, Song X, Wei L, Chen Z, Cao X. 2010. Degradome sequencing reveals endogenous small RNA targets in rice (*Oryza sativa* L. ssp. *indica*). *Frontiers in Biology* 5: 67–90.
- Zhu X, Dunand C, Snedden W, Galaud JP. 2015. CaM and CML emergence in the green lineage. *Trends in Plant Science* 20: 483–489.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Phenotypic analysis of transgenic plants overexpressing Ath-miR391 in *Arabidopsis* Nossen-0 (No-0) genetic background.

Fig. S2 Expression analysis for *ACA10*, *ARF3*, and *ARF4* using RT-semi-qPCR assays.

Fig. S3 Putative Ca²⁺-binding EF-hand (EF-h) motifs identified from miR1432-targeted CML family members in Poaceae.

Fig. S4 Evidence for dual targeting of ACA and EF-h/CML family members by miR1432-directed PTGS in grass species.

Table S1 Sequence information for DNA oligonucleotides used as primers or probes.

Table S2 Sequence information for the proposed miR391 superfamily members in diverse plant species.

Table S3 Sequence information for miRNA-targeted ACA family members analyzed in this work.

Table S4 Sequence information for miR1432-targeted EF-h/CML family members analyzed in this work.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.