

University of Kentucky  
UKnowledge

Plant and Soil Sciences Faculty Publications

Plant and Soil Sciences

2-22-2017

# The miRNAome of *Catharanthus roseus*: Identification, Expression Analysis, and Potential Roles of MicroRNAs in Regulation of Terpenoid Indole Alkaloid Biosynthesis

Ethan M. Shen  
University of Kentucky

Sanjay Kumar Singh  
University of Kentucky, [sanjaysingh@uky.edu](mailto:sanjaysingh@uky.edu)

Jayadri S. Ghosh  
University of Kentucky, [jayadri.ghosh@uky.edu](mailto:jayadri.ghosh@uky.edu)

Barunava Patra  
University of Kentucky, [barunava.patra@uky.edu](mailto:barunava.patra@uky.edu)

Priyanka Paul  
University of Kentucky, [priyanka.kasabiyanka@gmail.com](mailto:priyanka.kasabiyanka@gmail.com)

Follow this and additional works at: [https://uknowledge.uky.edu/pss\\_facpub](https://uknowledge.uky.edu/pss_facpub)

 Part of the [Cell and Developmental Biology Commons](#), [Genomics Commons](#), [Molecular Biology Commons](#), and the [Plant Sciences Commons](#).  
Right click to open a feedback form in a new tab to let us know how this document benefits you.

## Repository Citation

Shen, Ethan M.; Singh, Sanjay Kumar; Ghosh, Jayadri S.; Patra, Barunava; Paul, Priyanka; Yuan, Ling; and Pattanaik, Sitakanta, "The miRNAome of *Catharanthus roseus*: Identification, Expression Analysis, and Potential Roles of MicroRNAs in Regulation of Terpenoid Indole Alkaloid Biosynthesis" (2017). *Plant and Soil Sciences Faculty Publications*. 92.  
[https://uknowledge.uky.edu/pss\\_facpub/92](https://uknowledge.uky.edu/pss_facpub/92)

This Article is brought to you for free and open access by the Plant and Soil Sciences at UKnowledge. It has been accepted for inclusion in Plant and Soil Sciences Faculty Publications by an authorized administrator of UKnowledge. For more information, please contact [UKnowledge@sv.uky.edu](mailto:UKnowledge@sv.uky.edu).

---

**Authors**

Ethan M. Shen, Sanjay Kumar Singh, Jayadri S. Ghosh, Barunava Patra, Priyanka Paul, Ling Yuan, and Sitakanta Pattanaik

**The miRNAome of *Catharanthus roseus*: Identification, Expression Analysis, and Potential Roles of MicroRNAs in Regulation of Terpenoid Indole Alkaloid Biosynthesis****Notes/Citation Information**

Published in *Scientific Reports*, v. 7, article no. 43027, p. 1-13.

© The Author(s) 2017

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

**Digital Object Identifier (DOI)**

<https://doi.org/10.1038/srep43027>

# SCIENTIFIC REPORTS



OPEN

## The miRNAome of *Catharanthus roseus*: identification, expression analysis, and potential roles of microRNAs in regulation of terpenoid indole alkaloid biosynthesis

Received: 01 September 2016

Accepted: 18 January 2017

Published: 22 February 2017

Ethan M. Shen<sup>1,2,\*</sup>, Sanjay K. Singh<sup>1,\*</sup>, Jayadri S. Ghosh<sup>1</sup>, Barunava Patra<sup>1</sup>, Priyanka Paul<sup>1</sup>, LingYuan<sup>1</sup> & Sitakanta Pattanaik<sup>1</sup>

MicroRNAs (miRNAs) regulate numerous crucial biological processes in plants. However, information is limited on their involvement in the biosynthesis of specialized metabolites in plants, including *Catharanthus roseus* that produces a number of pharmaceutically valuable, bioactive terpenoid indole alkaloids (TIAs). Using small RNA-sequencing, we identified 181 conserved and 173 novel miRNAs (cro-miRNAs) in *C. roseus* seedlings. Genome-wide expression analysis revealed that a set of cro-miRNAs are differentially regulated in response to methyl jasmonate (MeJA). *In silico* target prediction identified 519 potential cro-miRNA targets that include several auxin response factors (ARFs). The presence of cleaved transcripts of miRNA-targeted ARFs in *C. roseus* cells was confirmed by Poly(A) Polymerase-Mediated Rapid Amplification of cDNA Ends (PPM-RACE). We showed that auxin (indole acetic acid, IAA) repressed the expression of key TIA pathway genes in *C. roseus* seedlings. Moreover, we demonstrated that a miRNA-regulated ARF, CrARF16, binds to the promoters of key TIA pathway genes and repress their expression. The *C. roseus* miRNAome reported here provides a comprehensive account of the cro-miRNA populations, as well as their abundance and expression profiles in response to MeJA. In addition, our findings underscore the importance of miRNAs in posttranscriptional control of the biosynthesis of specialized metabolites.

As sessile organisms, plants have evolved unique mechanisms to defend themselves in adverse environmental conditions. Plants synthesize thousands of specialized metabolites that play unique roles in plant growth, development, and defense. Many of these metabolites are beneficial for humans. *Catharanthus roseus* (L.) G. Don, commonly known as Madagascar periwinkle, synthesizes over 130 terpenoid indole alkaloids (TIA), including the pharmaceutically important anti-neoplastic compounds, vinblastine and vincristine<sup>1</sup>. Biosynthesis of TIAs is highly complex, involving multiple sub-cellular compartments. Tryptamine, derived from the indole branch, and secologanin, from the seco-iridoid branch, are condensed to form the first TIA, strictosidine, catalyzed by STRICTOSIDINE SYNTHASE (STR). The biosynthesis of TIAs is induced by a number of factors, including fungal elicitors<sup>2</sup>, UV light<sup>3</sup>, wounding<sup>4</sup>, cold<sup>5</sup>, and drought stress<sup>6</sup>. The phytohormone jasmonic acid (JA) and its methyl ester, methyl jasmonate (MeJA), are major elicitors of TIA biosynthesis in *C. roseus*. JA-responsive expression of TIA pathway genes are positively regulated by transcription factors (TFs), such as the JA-responsive AP2/ERFs, ORCA2<sup>7</sup>, ORCA3<sup>8</sup>, ORCA4 and ORCA5<sup>9</sup>, as well as the bHLH factors, CrMYC2<sup>10</sup>, BIS1<sup>11</sup>, and BIS2<sup>12</sup>, and the WRKY TF, CrWRKY1<sup>13</sup>. Several repressors, including the G-box binding basic leucine-zipper (bZIP)

<sup>1</sup>Department of Plant and Soil Sciences, University of Kentucky, 1401 University Drive, Lexington, KY 40546, USA.

<sup>2</sup>Math, Science, and Technology Center, Paul Laurence Dunbar High School, 1600 Man o' War Boulevard, Lexington, KY 40513, USA. \*These authors contributed equally to this work. Correspondence and requests for materials should be addressed to L.Y. (email: lyuan3@uky.edu) or S.P. (email: spatt2@uky.edu)

factors, GBF1 and GBF2<sup>14</sup>, and zinc finger proteins, ZCT1, ZCT2, and ZCT3<sup>15,16</sup>, negatively regulate the expression of TIA pathway genes in *C. roseus*. Other phytohormones, such as cytokinin (CK) and auxin, also affect TIA biosynthesis. In *C. roseus* cell suspension culture, CK enhances the accumulation of alkaloids<sup>17</sup>, whereas auxin negatively regulates the expression of key TIA biosynthesis genes, including *TRYPTOPHAN DECARBOXYLASE (TDC)* and *STR*<sup>18,19</sup>. However, the molecular mechanism of auxin-mediated regulation of the TIA pathway is unknown. Auxin is an essential phytohormone that plays pivotal roles in plant growth and development. Auxin regulates gene expression through a signal transduction pathway which includes the F-box protein TRANSPORT INHIBITOR RESPONSE 1/AUXIN SIGNALING F-BOX PROTEINS (TIR1/AFB), AUXIN-RESPONSE FACTOR (ARF), and Auxin/INDOLE-3-ACETIC ACID (Aux/IAA) proteins. ARFs are DNA binding proteins that recognize auxin-responsive (AuxRE) elements in target promoters to activate or repress expression<sup>20</sup>. In Arabidopsis, expression of ARFs are regulated posttranscriptionally by microRNAs (miRNAs)<sup>21,22</sup>.

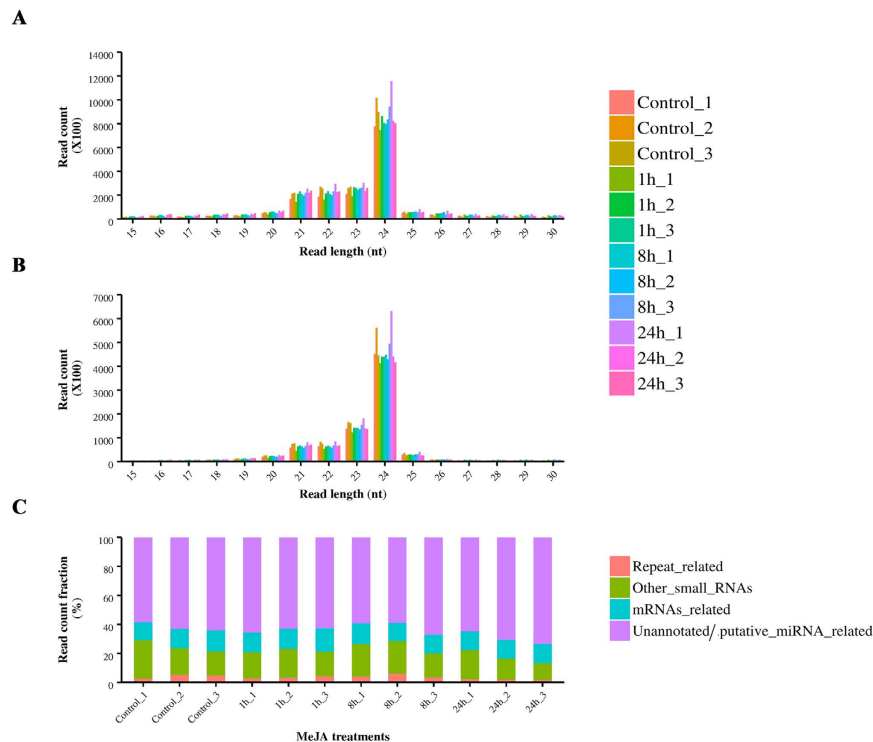
miRNAs comprise a major class of endogenous non-coding small regulatory RNAs approximately 21 to 24 nucleotides in length. They are present in a variety of organisms from algae to plants<sup>23</sup>. In plants, mature miRNAs are processed from primary transcripts by DICER-LIKE 1 RNase (DCL1), and subsequently loaded onto the ARGONAUTE (AGO) protein(s) to form RNA-induced silencing complexes (RISCs)<sup>24</sup>. The miRNA-loaded RISC binds to the target mRNA in a sequence-specific manner to either degrade the mRNA<sup>25</sup> or prevent it from translation<sup>26</sup>. Many plant miRNAs are evolutionarily conserved among plant species<sup>27,28</sup>. miRNAs are involved in numerous biological processes, including plant hormone homeostasis<sup>29</sup>, root development<sup>30,31</sup>, leaf morphogenesis<sup>32,33</sup>, flower development<sup>34</sup>, and embryogenesis<sup>35,36</sup>. However, information on the involvement of miRNA in regulation of specialized metabolite biosynthesis is limited<sup>37–39</sup>. Whether miRNAs are involved in regulating TIA biosynthesis in *C. roseus* has not been extensively investigated.

A previous study has identified 81 conserved and 7 novel miRNAs in *C. roseus* seedlings using deep sequencing<sup>40</sup>. The number of identified miRNAs in the report is considerably lower than the average number of miRNAs in plants studied thus far. Moreover, as the *C. roseus* genome sequence was unavailable during the previous study, the tomato genome was used as the reference to predict novel miRNAs. Here, we systematically identified miRNAs in *C. roseus* by sequencing four small RNA libraries from the control and JA-treated seedlings. We identified a total of 354 *C. roseus* miRNAs, including 181 conserved and 173 novel miRNAs. All identified miRNAs were mapped to the recently available *C. roseus* reference genome<sup>41</sup>. Expression of selected miRNAs and their potential targets were validated using quantitative reverse transcription PCR (qRT-PCR). In addition, we predicted a set of MeJA-responsive miRNAs that target a group of ARFs in *C. roseus*. Cleaved ARF transcripts in *C. roseus* cells were identified by Poly(A) Polymerase-Mediated Rapid Amplification of cDNA Ends (PPM-RACE). We hypothesized that ARFs bind to TIA pathway gene promoters to repress their activity, and at least some of these ARFs are regulated by miRNAs. We demonstrated that a JA-responsive, miRNA regulated ARF, CrARF16, binds to key TIA pathway gene promoters to repress their expression. Our findings provide a comprehensive account of the *C. roseus* miRNAome, and suggest that miRNAs and ARFs are involved in the regulation of TIA biosynthesis.

## Results

**Small RNA populations in *C. roseus* seedlings.** A total of 12 small RNA (sRNA) libraries were constructed from the control and MeJA-treated *C. roseus* seedlings (Supplementary Fig. S1). Sequencing of small RNA libraries, from control and seedlings treated with MeJA for 1 h, 8 h, and 24 h, produced approximately 76.5 million (M), 76.8 M, 79 M, and 84.5 M raw reads, respectively (Supplementary Table S1). Among the raw reads, 88–91% were found to contain adapter sequences and thus chosen for further analysis. After removing low-quality reads and trimming the adaptor sequences, approximately 64.1 M, 65.5 M, 69.8 M, and 73 M clean reads were obtained for the control, 1 h-, 8 h-, and 24 h-MeJA-treated libraries, respectively (Supplementary Table S1). A total of 60.7 M (control), 62.2 M (1 h), 66.6 M (8 h), and 69.3 M (24 h) sequences were successfully mapped to the *C. roseus* reference genome. Sequences mapped to the *C. roseus* reference genome were further subjected to removal of transfer and ribosomal RNAs (t/rRNAs). Size distribution analysis of the sRNA sequences after filtering showed that all libraries exhibit similar distribution in length, with the most abundant class being 24 nt (Fig. 1A). After removal of redundant reads, the 24 nt reads remained most abundant, followed by the 23 nt reads (Fig. 1B). Overall, putative miRNAs and un-annotated sRNAs constitute 60–75% of the total cleaned reads in all twelve libraries (Fig. 1C).

**Conserved miRNAs in *C. roseus*.** To identify conserved miRNAs in *C. roseus*, the unique sequences were aligned to the mature sequences of known miRNAs deposited in the miRBase database<sup>42</sup>. A total of 181 conserved miRNAs, representing 33 miRNA families, were identified in *C. roseus* (Supplementary Table S2). Conserved miRNA varied from 18 nt to 24 nt in length. Most of the conserved miRNA families reported by Prakash *et al.*<sup>40</sup> are also present in our dataset. However, we could not detect miR169, miR530, miR828, miR2199, miR5139, miR5368, and miR6173 in our dataset; on the other hand, we detected miR2275, miR5532, miR5538, miR6300, miR6478, miR8016, miR858, miR397, miR390, miR393, miR394, and miR395, that were not reported by Prakash *et al.*<sup>40</sup>. In addition, 65 precursor sequences, representing 28 conserved miRNAs, were also predicted from the *C. roseus* genome (Supplementary Table S3). Several cro-miRNA families were found to be encoded by multiple loci in the *C. roseus* genome. For instance, cro-miR159 is encoded by two different loci (cro\_scaffold\_3041204 and cro\_scaffold\_3039079). Of the 65 identified precursor sequences, cro-miR159, -miR396, and -miR166 showed high abundance, having an average normalized count (NC) >1,000 (Supplementary Table S4), similar to those observed in other plant species, including radish and banana<sup>43,44</sup>. Cro-miR2275, -miR477, and -miR2111 were found to be least abundant, with an average NC less than 10. A considerable variation in expression was also noticed among loci coding for the same miRNA family (Supplementary Table S4). For instance, the average NC for the cro-miR319 family varied from 15–542. Similar variations in read abundance were also observed among members of the miRNA families such as cro-miR396 (1–1440), -miR395 (1–94), -miR171 (0.3–44), and -miR156

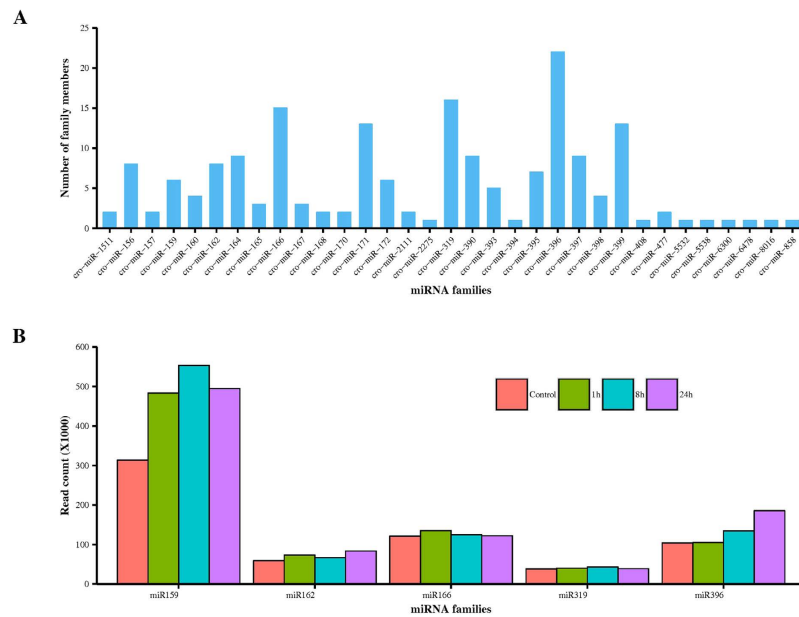


**Figure 1. Length distribution and different classes of sequence reads identified in the small RNA libraries of *C. roseus*.** (A) Size distribution of the total reads. (B) Size distribution of unique reads. (C) Distribution of different classes of sequence reads derived from the small RNA libraries. Putative miRNAs and un-annotated sRNAs constitute 60–75% of the total cleaned reads in all twelve libraries. Each small RNA library (control, 1 h-, 8 h-, or 24 h-MeJA treated) is represented by three biological replicates. *C. roseus* seedlings treated with 100  $\mu$ M MeJA for 1 h, 8 h, or 24 h were used for RNA isolation and library preparation. Mock-treated seedlings served as control.

(1–194) (Supplementary Table S4). The average minimal folding free energy (MFE) value of the miRNA precursors was  $-49.2 \text{ kcal mol}^{-1}$ ; and the length of the precursors ranged from 75 nt to 377 nt, with an average length of 166 nt (Supplementary Table S3). The predicted secondary structures of the known miRNA precursors and locations of mature miRNAs in precursors were shown in Supplementary Fig. S2. A significant difference in the number of members was recorded for each conserved miRNA family (Fig. 2A). The largest family identified was *cro-miR396* with 22 members, followed by *cro-miR319* and *-miR166*, with 16 and 15 members, respectively. Of the remaining 30 families, 21 families were represented by 2–13 members while 9 were represented by a single member (Fig. 2A). Considerable variation in expression was observed among the individual families (Fig. 2B), and also among members of the same miRNA family. For instance, of the 6 members of the *cro-miR159* family, *cro-miR159a* showed highest abundance (average NC > 1,000) while *cro-miR159c* was found to be least abundant (average NC < 1.0) (Supplementary Table S5).

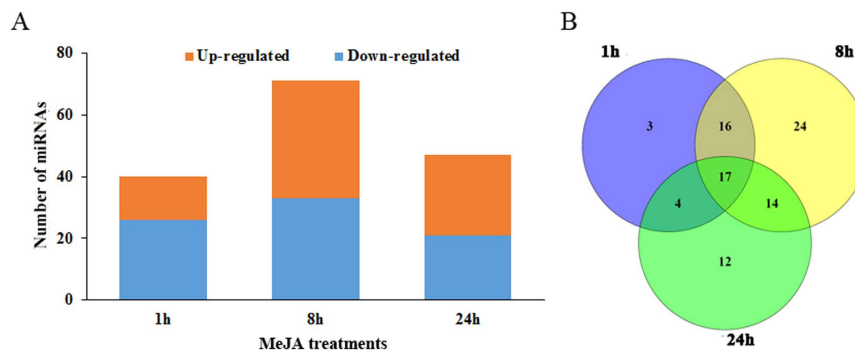
**Novel *C. roseus* miRNAs.** A total of 173 novel *C. roseus* miRNAs were predicted based on established prediction criteria<sup>45,46</sup> (Supplementary Tables S6 and S7). The length of these novel miRNA precursors, ranging from 55–491 nt and averaging in 137 nt, agreed with the commonly observed length of miRNA precursors in plants<sup>47</sup>. The average minimal folding free energy (MFE) value of these miRNA precursors was  $-43.2 \text{ kcal mol}^{-1}$ . The predicted secondary structures of these novel miRNA precursors and locations of mature miRNAs in precursors are shown in Supplementary Fig. S3. The sequence read abundance of the majority of *C. roseus* novel miRNAs were low compared to the conserved *cro-miRNAs* (Supplementary Table S8). A considerable variation in expression was also observed in *C. roseus*-specific miRNAs. Of the 173 *C. roseus*-specific miRNAs, the NC for 22 miRNAs were less than 10, whereas that for 134 miRNAs ranged from 10 to 200. Seventeen abundant novel *C. roseus*-specific miRNAs, including *cro-novel-35*, *cro-novel-51*, and *cro-novel-56*, have NC of more than 200 (Supplementary Table S8). The authenticity of predicted novel miRNAs is usually supported by the existence of complementary sequences<sup>46</sup>. In *C. roseus*, 162 of the 173 novel miRNAs have complementary miRNAs.

**MeJA-responsive miRNAs in *C. roseus* seedlings.** We analyzed the small RNA libraries to identify differentially expressed miRNAs (DEMs) in *C. roseus* seedlings with or without MeJA treatment. We identified 40 (14 upregulated and 26 downregulated), 71 (38 upregulated and 33 downregulated), and 47 (26 upregulated and 21 downregulated) JA-responsive miRNAs after 1 h, 8 h, and 24 h of MeJA treatment, respectively (Fig. 3A and Supplementary Table S9). Thirty-six conserved miRNAs, belonging to 14 miRNA families, such as *cro-miR160* and *-miR164*, were differentially expressed in at least one time-point. Additionally, expression



**Figure 2. Number of members and abundance of known miRNA families identified in *C. roseus*.**

(A) Number of members in each conserved miRNA family in *C. roseus*. (B) Sequence read counts of top five conserved miRNA families in the small RNA libraries. The transcript abundance of each family in the control, as well as 1 h-, 8 h-, and 24 h-MeJA treated library is shown.

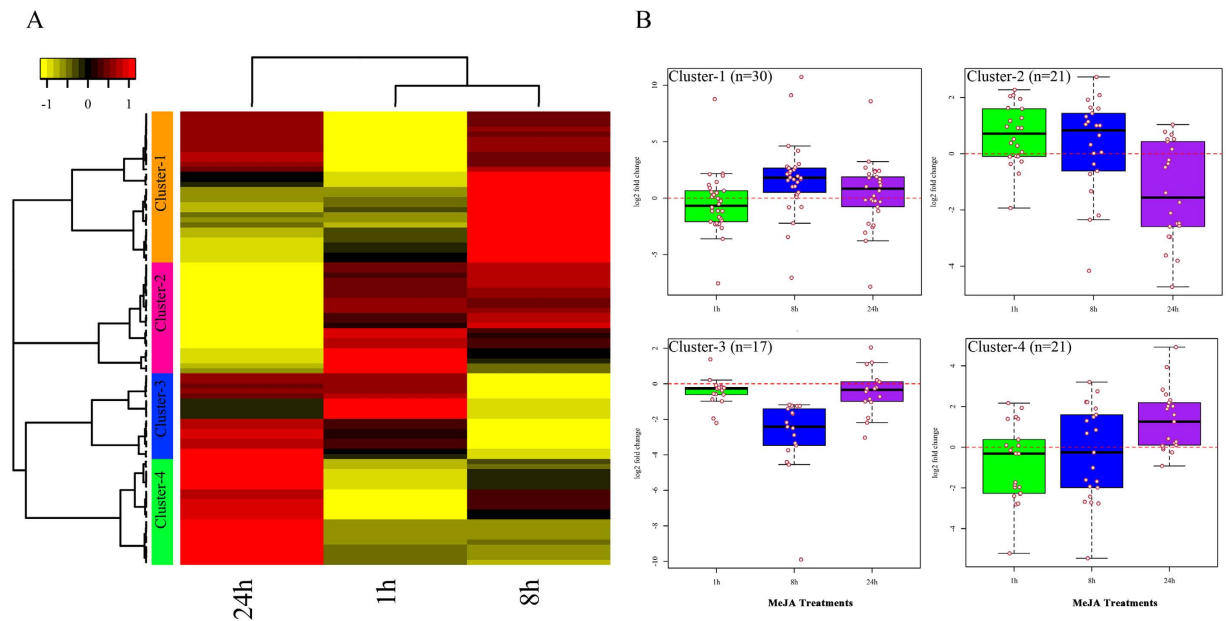


**Figure 3. Differential expression analysis of miRNAs in *C. roseus* seedlings. (A)** Numbers of up- and downregulated miRNAs in *C. roseus* seedlings after MeJA-treatment for 1 h, 8 h, or 24 h (B) Venn diagram showing the overlap of differentially expressed miRNAs (DEMs) between the three time points.

of 11 miRNAs, cro-miR319a-5p, cro-miR164b, cro-miR390b-3p-1, cro-miR164a, cro-miR393d, cro-novel-62, cro-novel-92, cro-novel-73, cro-novel-65, cro-novel-10, and cro-novel-63, were increased after MeJA treatment in all time-points. In contrast, 6 miRNAs, cro-novel-64, cro-novel-64, cro-novel-54, cro-novel-130, cro-miR398f, cro-miR319a-3p.2-3p, and cro-miR319a, were downregulated at all time-points.

Venn-diagram analysis revealed that 17 DEMs were common in all three time-points, whereas 16, 14, and 4 DEMs were common in 1 h–8 h, 8 h–24 h, and 1 h–24 h comparisons, respectively (Fig. 3B and Supplementary Table S9). Among the 17 common DEMs, 7 miRNAs (cro-novel-73, cro-novel-92, cro-novel-62, cro-novel-10, cro-miR393d, cro-miR164a, and cro-miR319a-5p) were upregulated and 5 miRNAs (cro-miR319a-3p.2-3p, cro-miR319a, cro-miR398f, cro-novel-64, and cro-novel-130) were downregulated at all three time points (Fig. 3B). In addition, the expression of five miRNAs (cro-novel-71, cro-novel-38, cro-novel-98, cro-miR168a, and cro-miR160h) were down 1 h following MeJA treatment, but were up after 8 h and 24 h. The DEMs were divided into four groups according to their expression patterns (Fig. 4A). The mean expression of DEMs in group-I ( $n = 30$ ) increased after MeJA-treatment, peaking around 8 h. Those in group-II ( $n = 21$ ) were initially upregulated by MeJA but gradually downregulated from 8 h to 24 h. The members of group-III ( $n = 17$ ) were downregulated in response to MeJA; however, some members, including cro-novel-2, cro-miR396, and cro-miR396a-3p-4, were upregulated significantly between 8 h and 24 h. The mean expression of DEMs in group-IV ( $n = 21$ ) increased continuously for 24 h (Fig. 4B).



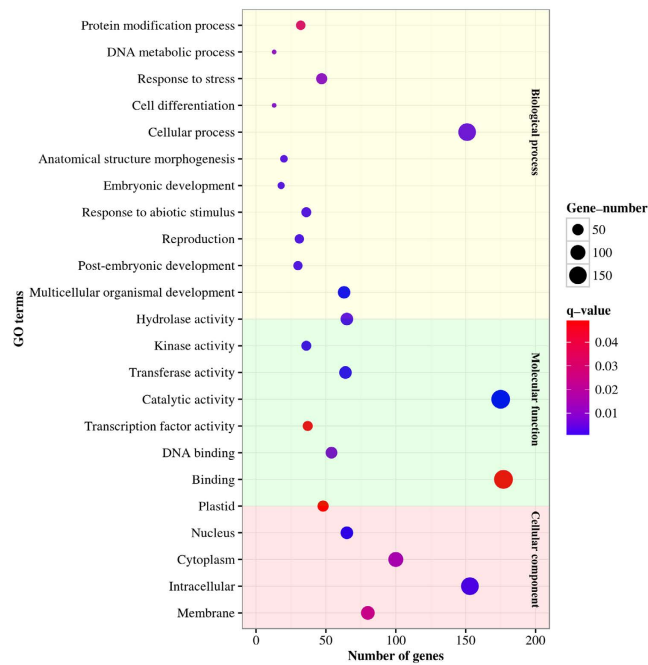


**Figure 4. Hierarchical cluster analysis of DEMs in *C. roseus* seedlings.** (A) Heat-map of the log<sub>2</sub>-fold change values of DEMs after 1 h, 8 h, and 24 h of MeJA treatment compared to the control. DEMs were divided into four different clusters (indicated by color-coded bars with numbers) based on their expression patterns. Columns denote the time points (1 h, 8 h, and 24 h) of MeJA treatment. The scale of heat-map is given as log<sub>2</sub> fold change with a range from  $-1.0$  (yellow) to  $+1.0$  (red). (B) Box plots represent the log<sub>2</sub> fold change (MeJA-treatment/control) in gene expression after MeJA treatment in four different clusters. Each single dot represents a miRNA.

**Identification and classification of cro-miRNA targets.** In order to explore the roles of DEMs in diverse biological processes, their putative targets were predicted using the open source web server, psRNATarget<sup>48</sup>, with default parameters. The transcript sequences of the *C. roseus* genome were used as a reference set. Detailed information related to the predicted targets for conserved and *C. roseus*-specific miRNAs were included in Supplementary Table S10. A total of 519 candidate target genes were identified for 80 cro-miRNAs (149 target genes for 33 conserved miRNA and 374 target genes for 47 novel miRNAs, respectively), with an average of 6 targets per miRNA. The majority of the target genes (65.3%) were predicted to be regulated by transcript degradation, whereas the remaining targets are regulated by translational repression. Furthermore, 55 genes were targeted by more than one cro-miRNA. Among the conserved miRNA targets, many encode TFs, such as MYB TFs (targeted by cro-miR159), NAC domain-containing proteins (targeted by cro-miR164), and auxin response factors (regulated by cro-miR160). In addition, genes encoding receptor-like protein kinase, pentatricopeptide repeat (PPR) and tetratricopeptide repeat (TPR)-like superfamily proteins, and major facilitator superfamily proteins, were also identified as potential targets of cro-miR396 and cro-miR398. Potential targets were also identified for 47 *C. roseus*-specific miRNAs. These targets included mRNA encoding receptor-like protein kinase, ubiquitin protein ligase 6, blue-copper-binding protein, sterol methyltransferase 2, and F-box family protein. In addition, some *C. roseus*-specific miRNAs were also found to target TFs, including basic bHLH, bZIP domain proteins, and ARF family TFs (Supplementary Table S10).

The potential targets of the DEMs were annotated to better understand the biological functions of cro-miRNAs. Gene ontology (GO) analysis revealed that these target genes could be classified into 11 biological processes, 8 molecular functions, and 5 cellular components (Fig. 5). For biological processes, “cellular process”, “multicellular organism development”, and “response to stress”, were the three most dominant GO categories. With regard to cellular components, “intracellular” and “membrane” were the two most abundant GO terms. The GO term “intracellular” was significantly enriched for two children GO terms, “nucleus” and “cytoplasm”. The three most dominant GO terms in molecular functions were “binding”, “catalytic activity”, and “transferase activity”. Analysis of the GO term ‘transcription factor activity’ revealed that several families of TFs, such as GRAS, MYB, ARF, NAC, and GRE, were more dominant in this category. A significant number of genes associated with ‘kinase activity’ (mostly receptor-like kinases) were also found to be targeted by the cro-miRNAs. The GO term ‘catalytic activity’ was enriched with children GO terms, “transferase activity” and “hydrolase activity”.

Previous studies suggest that auxin negatively affects the expression of key TIA pathway genes in *Catharanthus*<sup>18,19</sup>. However, the underlying molecular mechanism of IAA-mediated repression is not well studied. ARFs are key components in auxin signaling pathway. Our target prediction analyses revealed *CrARF10*, *CrARF16*, and *CrARF17* as potential targets of MeJA-induced miR160. *HAM3*, a GRAS domain TE, was also found to have a high target prediction score in our analysis. We performed PPM-RACE on three selected targets, *ARF10*, *ARF16*, and *HAM3* to demonstrate that they are indeed targeted by the respective miRNAs. Cloning of the PCR products, followed by sequencing of the PPM-RACE products, confirmed that *ARF10* and *ARF16* were



**Figure 5. Gene ontology (GO) analysis of target genes of DEMs in *C. roseus* seedlings.** Each GO term is represented by a single circle, that the color indicates the q-values and the significance of the GO term, and the size is proportional to the gene numbers. The Y-axis represents the names of enrichment GO terms. The X-axis represents the number of genes.

cleaved by the conserved *cro-miR160*, whereas the GRAS domain TF *HAM3* was cleaved by *cro-miR171b-3p-2*, in *C. roseus* cells (Fig. 6A).

**Expression profiles of DEMs and their targets.** We used qRT-PCR to validate the small RNA sequencing (sRNA-seq) results of the DEMs and their potential targets. Expression of eleven (5 conserved and 6 novel) DEMs (Fig. 6B) and thirteen corresponding targets were measured using qRT-PCR (Fig. 6C). Expression of *cro-miR393d*, *cro-miR164a*, and *cro-miR164b* were upregulated, and that of *cro-miR168a* was downregulated, following 1 h MeJA treatment. There was no significant change in the expression of *miR160a* in response to 1 h MeJA treatment. The expression of all the selected conserved miRNAs were induced following 8 h and 24 h MeJA treatments. All analyzed *C. roseus*-specific miRNA were upregulated after 8 h and 24 h JA treatments. These results are in good agreement with the sRNA-seq data.

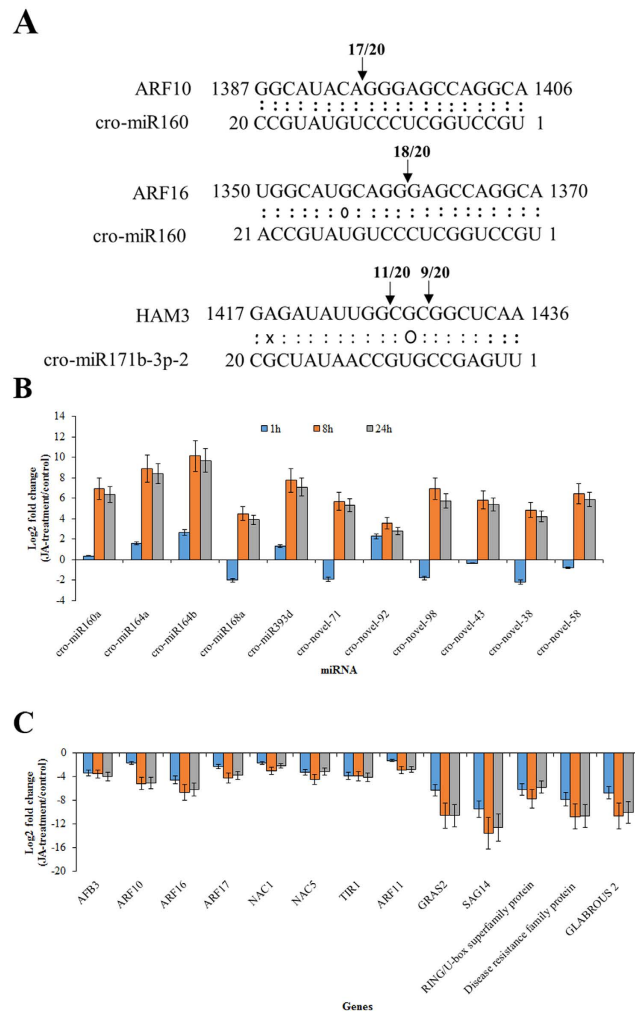
Usually, miRNAs and their target genes are expected to have contrasting expression patterns. We thus analyzed the expression of a number of TF genes (*ARF10*, *ARF16*, *ARF17*, *TIR1*, *AFB3*, *NAC1*, *NAC5* and *GRAS2*) which were predicted to be targeted by conserved and novel miRNAs, including *cro-miR160*, *cro-miR3953d*, *cro-miR164*, *cro-novel-43*, and *cro-novel-38*. As expected, the expression of the conserved and novel *cro-miRNAs* were inversely correlated to those of their targets (Fig. 6C).

**Negative regulation of TIA structural and regulatory genes by auxin.** TIA biosynthesis is influenced by a number of phytohormones including auxin and MeJA. Our results showed that expression of *CrARF10*, *CrARF16*, and *CrARF17* were repressed in MeJA-treated *C. roseus* seedlings. In addition, these three ARFs were identified as targets of MeJA-induced miRNA *cro-miR160*. The expression of *TIR1* and *AFB3*, crucial components in auxin signaling, were also altered in the MeJA-treated seedlings.

Auxin has been shown to represses the expression of several TIA pathway genes, including *TDC* and *STR*, in *C. roseus* cell lines<sup>18,19</sup>. Previous findings, and our analysis of JA-treated *C. roseus* seedlings, confirmed that most TIA biosynthetic pathway genes are induced by exogenous application of JA (Supplementary Fig. S4). We thus hypothesized that JA and auxin act antagonistically to regulate the TIA pathway in *C. roseus*. To test our hypothesis, we analyzed the expression of key regulatory and structural genes in *C. roseus* seedlings treated with IAA for different time periods. As shown in Fig. 7, auxin repressed the expression of key regulatory (*ORCA3*, *ORCA4*, *ORCA5*, *BIS1*) and structural genes (*STR*, *TDC*, *G10H*, *SLS*) in the TIA pathway. The transcript levels of *BIS1*, *ORCA4*, *SLS*, and *G10H* were significantly decreased at all time-points following IAA treatment. The expression of *ORCA3* was not affected by IAA at 1 h, but reduced after a longer treatment. The expression of *TDC*, *SGD*, and *STR*, were moderately affected by IAA treatment (Supplementary Fig. S5). In addition, we also measured the expression of two ARFs, *CrARF10* and *CrARF16*, in IAA-treated *Catharanthus* seedlings using qPCR. Expression of *CrARF10* and *CrARF16* were induced significantly following IAA treatment (Supplementary Fig. 6).

The repression of TIA pathway genes by IAA led us to hypothesize that ARFs repress TIA pathway genes by binding to key pathway gene promoters. To test our hypothesis, the promoter sequences of *TDC*, *G10H*, and *STR*





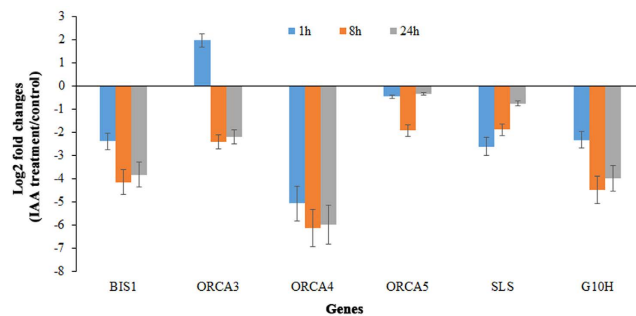
**Figure 6. PPM-RACE and experimental validation of selected conserved and novel cro-miRNAs, as well as their predicted targets.** (A) PPM-RACE confirmed the cleavage of *CrARF10* (CRO\_T007962) and *CrARF16* (CRO\_T023928) by cro-miR160 and that of *HAM3* (CRO\_T018940) by cro-miR171b-3p-2. cro-miR171b-3p-2 cleaves *HAM3* at two sites. The numbers above the sequences indicate the numbers of cleaved fragments detected at the cro-miR160 or cro-novel-43 target sites relative to the total fragments sequenced. For example, of 20 sequenced PCR fragments of *ARF10*, 17 were cleaved at position 1394 by cro-miR160. Watson-Crick pairing (:), G–U wobbles (O), and mismatched base pairing (X) are indicated. (B,C) Quantitative RT-PCR validated the expression of selected conserved and *C. roseus*-specific miRNAs (B) and their targets (C). The expression values of miRNAs and their targets were normalized against the expression of the endogenous controls, U6 snRNA and RPS9, respectively. The log<sub>2</sub>-transformed fold changes in mRNA levels (MeJA-treated/control) of miRNAs and their targets is shown. The data represent the mean values  $\pm$  SD of three biological replicates.

were PCR-amplified, from *Catharanthus* genomic DNA using gene-specific primers, and cloned into a plant expression vector containing a firefly luciferase reporter<sup>9</sup>; the promoter-driven transcriptional activities were measured, alone or in the presence of *CrARF16*, in a tobacco protoplast assay. As shown in Fig. 8, all three promoters individually generated significant background luciferase activity. However, co-electroporation of the promoters-reporter plasmids with *CrARF16* significantly repressed the basal activities of all three promoters.

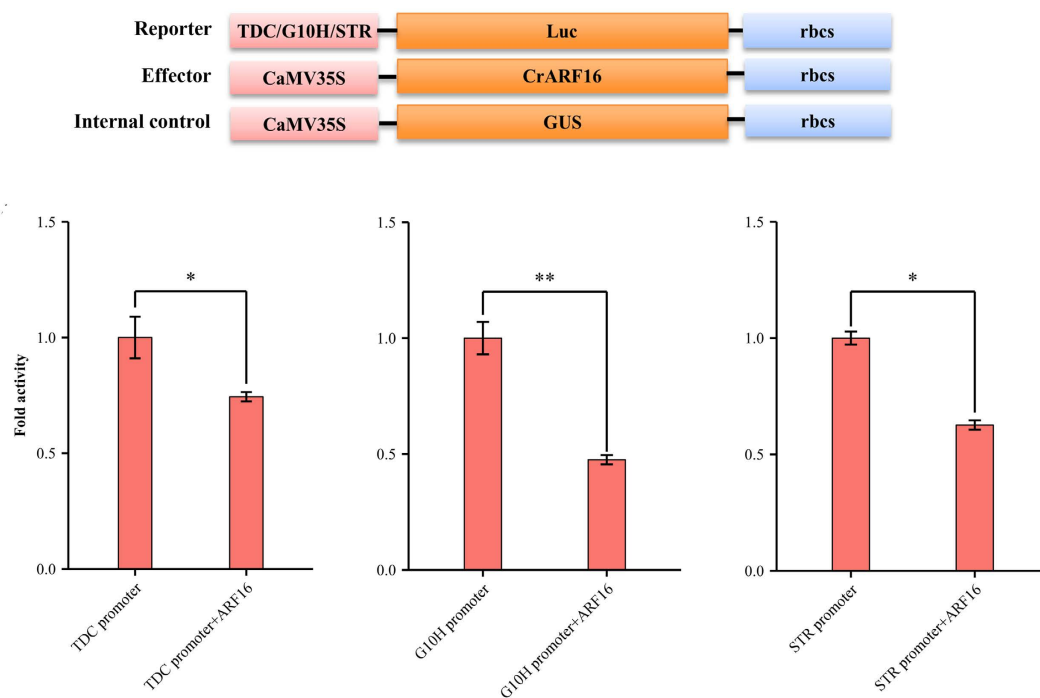
## Discussion

Here, we report the identification, expression analysis, target prediction and validation of conserved and novel miRNAs in *C. roseus* seedlings. Additionally, we identified a set of JA-responsive cro-miRNAs and attempted to explore their potential roles in posttranscriptional control of TIA biosynthesis.

**The miRNAome of *C. roseus*.** Using high-throughput sequencing, we identified 354 cro-miRNAs (Supplementary Tables S2, S3, S6 and S7), a number that is close to what have been reported from *Arabidopsis* (427), *Zea mays* (321), *Sorghum bicolor* (241), and *Populus trichocarpa* (401), but less than those from *Medicago truncatula* (756), *Oryza sativa* (713), and *Glycine max* (639). A previous study has identified 81 conserved miRNAs in *Catharanthus* seedlings<sup>40</sup>. Most of the conserved miRNA family members identified in that study are present



**Figure 7. Quantitative RT-PCR analysis of selected TIA biosynthetic pathway genes in control and IAA-treated *C. roseus* seedlings.** The relative abundance of individual gene is presented as the ratio of the IAA-treated vs. control seedlings. The expression value of each gene was normalized against the expression of endogenous control, RPS9. The log<sub>2</sub>-transformed fold changes in mRNA levels (IAA-treated/control) of the target genes is shown. The data represent the mean values  $\pm$  SD of two biological replicates.



**Figure 8. Repression of TDC, G10H, and STR promoter activities in tobacco cells by CrARF16.** TDC, G10H, and STR promoters fused to the *fire-fly* luciferase reporter were electroporated into tobacco protoplasts either alone or with the effector plasmid expressing CrARF16. A plasmid containing the  $\beta$ -glucuronidase (GUS) reporter, controlled by the CaMV35S promoter and *rbcs* terminator, was used as a normalization control. Schematic diagrams of the reporter, effector, and internal control plasmids used in this assay are shown on the top. Luciferase and GUS activities were measured 20 h after electroporation. Luciferase activity was normalized against the GUS activity. The reporter alone without effectors served as the control. Data presented are means  $\pm$  SDs of three biological replicates. Asterisks indicate a significant difference compared to the control at \* $P < 0.08$  or \*\* $P < 0.005$  (t-test).

in our dataset. The 24 nt long sRNAs dominated the *C. roseus* sRNA transcriptome, indicating the dominance of heterochromatic small interfering RNA (siRNA) (Fig. 1A,B). Similar observations have been made for many other plant species, including peanut, hot pepper, cucumber, rice and trifoliate orange<sup>49–53</sup>. The high percentage of 24 nt sRNAs probably plays a vital role in preserving genome integrity by heterochromatic-histone modification<sup>54</sup>. Our dataset also shows a high ratio of total reads/unique reads for the 21 nt class. The majority of the identified cro-miRNAs were 21 nt in length, which is the canonical size for miRNAs generated from DCL1 processing<sup>55</sup>. Based on the recently published draft *C. roseus* genome, which is not fully assembled and annotated, we were able to predict 65 precursor sequences representing 28 conserved miRNAs in *C. roseus* genome (Supplementary Table S3). Cro-miR159, -miR396, -miR166, -miR162, and -miR319 are highly abundant in *C. roseus* (Supplementary Table S4). MiR159 and miR319 are conserved family of miRNA and important for plant

growth, morphogenesis, and reproduction, by regulating the expression of a number of MYB and TCP family TF genes<sup>56</sup>. The conserved miR396 regulates the GROWTH-REGULATING FACTOR (GRF) family TFs, which control cell proliferation in *Arabidopsis* leaves<sup>57</sup>. MiR162 targets DICER-LIKE1 (DCL1), which catalyzes the formation of miRNA sequences, thus controlling miRNA biosynthesis by negative feedback regulation<sup>58</sup>. MiR166 and miR159 are involved in abiotic stress responses such as cold and salinity<sup>59,60</sup>. A previous study has identified conserved miRNAs from 99 different tissues in 34 different plant species<sup>61</sup>. We show that the *C. roseus* miRNAome harbors most of the ancient miRNA families, including miR160, miR159, and miR164 (Fig. 2A, Supplementary Table S2). In addition to the conserved miRNAs, all sequenced and analyzed plant genomes contain family- or species-specific miRNAs that may have originated and diverged on scales ranging from family to species<sup>62</sup>. The *C. roseus* genome harbors numerous miRNAs that lack close orthologs in other plant species (Supplementary Table S7). Consistent with previous observations<sup>63,64</sup>, we also found that the majority of the *C. roseus*-specific miRNAs show relatively low expression when compared to the conserved miRNAs (Supplementary Table S8). A previous study<sup>40</sup> has identified a limited number of *cro*-miRNAs that is significantly lower than the average numbers of miRNAs reported in other plant species. The miRNAome reported here provides a comprehensive account of the numbers and expression profiles of conserved and novel miRNAs in *C. roseus* seedlings. Unlike the previous report, the majority of the *cro*-miRNAs identified in the study are mapped to the *C. roseus* draft genome. Moreover, expression analysis of the selected conserved and novel *cro*-miRNAs by qRT-PCR complements the sRNA-seq data and validates our miRNA prediction criteria (Fig. 6A–C).

**JA-responsive miRNAs target key regulators in auxin signaling.** JA-responsive expression of regulatory and structural genes is a hallmark of TIA pathway. Through modulating the expression of the TIA pathway genes, exogenous application of JA increases the production of TIAs in *C. roseus* seedlings, hairy roots, and cell cultures<sup>65–67</sup>. A number of JA-responsive regulatory genes have been isolated and characterized for their roles in regulation of the TIA pathway. However, the influence of JA on miRNA expression is not well studied. In *Taxus chinensis*, MeJA downregulates *miR156*, *miR168*, *miR169*, *miR172*, *miR172*, *miR396*, *miR480*, and *miR1310*, but upregulates *miR164* and *miR390*<sup>68</sup>. Our analysis shows that a number of *cro*-miRNAs were also differentially expressed in response to MeJA treatment (Fig. 3A,B; Supplementary Table S9) and the DEMs were divided into four independent clusters based on their expression (Fig. 4A,B). We observed a significant induction of *cro-miR160*, *cro-miR168*, and *cro-miR393* after JA treatment. Target prediction of MeJA-responsive DEMs identified 519 potential targets (Supplementary Table S10). GO analysis led to a better understanding of their biological functions (Fig. 5). Detailed analysis of the GO term, “transcription factor activity” revealed the significance of several auxin signaling genes, including *TIR1*, *CrARF10*, *CrARF16*, and *CrARF17*.

ARFs are a plant-specific TF family that controls auxin-regulated transcription. Most ARFs have a conserved N-terminal DNA binding domain, C-terminal dimerization domain, and a non-conserved middle region that confers transcriptional repression or activation<sup>69,70</sup>. ARFs bind to AuxREs, which are found in the promoters of early auxin response genes, including *Aux/IAA*, *SMALL AUXIN UPREGULATED RNA (SAUR)*, and *GH3*, to either enhance or repress their transcription<sup>71,72</sup>. In *Arabidopsis*, *ARF10*, *ARF16*, and *ARF17* are reported to be targeted by miR160<sup>73</sup>. We observed a significant upregulation in *cro-miR160* after MeJA treatment (Fig. 6B). As expected, expression of *CrARF10*, *CrARF16*, and *CrARF17* were significantly downregulated after MeJA-treatment in *C. roseus* seedlings (Fig. 6C). Other than *cro-miR160*, MeJA also induced the expression of *cro-miR164a*, *cro-miR164b*, *cro-miR393d*, and *cro-novel-92*, which are predicted to target the auxin signaling genes, including *NAC1*, *NAC5*, *TIR1*, *AFB3*, and *ARF11*. The JA-responsive differential expression and targeting of IAA signaling genes by these MeJA-induced *cro*-miRNAs suggest that they can potentially affect TIA biosynthesis through regulating key factors in IAA signaling.

**Cr-miRNA-targeted ARFs are potentially involved in regulation of the TIA pathway.** Previous studies have reported the negative effects of auxin on expression of several structural genes in the TIA pathway<sup>18,19</sup>. However, the molecular mechanism underlying auxin-mediated repression of TIA pathway genes is still elusive. Our results not only confirmed the previous findings, but also showed that auxin significantly downregulates the expression of several key transcriptional regulators, including *ORCA3*, *ORCA4*, and *BIS1*, in TIA biosynthetic pathway (Fig. 7). In addition, we showed that auxin upregulates the expression of two ARFs, *CrARF10* and *CrARF16*, in *Catharanthus* seedlings. We also demonstrated that the miRNA-targeted *CrARF16* represses the activity of key TIA pathway gene promoters, such as *TDC*, *STR*, and *G10H* in plant cells (Fig. 8). Taken together our findings suggest that auxin and JA act antagonistically to regulate the expression of TIA pathway genes, in part, through the posttranscriptional regulation by *cro*-miRNAs. Auxin represses the expression of key TIA pathway genes, possibly mediated by the *CrARF* repressors. MeJA-induced expression of *cro*-miRNAs result in the degradation of *CrARF* repressors, leading to activation of TIA pathway genes.

In conclusion, this study provides a comprehensive account of the *C. roseus* miRNAome and the possible roles of *cro*-miRNAs in regulating the biosynthesis of TIAs. In addition, targeting of *CrARFs* by MeJA-induced *cro*-miRNAs to attenuate the repression of auxin highlights antagonistic functions of two key phytohormones in TIA pathway regulation. Our findings provide a starting point for further investigation of the regulatory roles miRNAs in specialized metabolite biosynthesis, in particular, TIAs in *C. roseus*.

## Materials and Methods

**Plant Materials, treatments, and RNA isolation.** *Catharanthus roseus* (L.) G. Don cv. ‘Cooler Apricot’ seeds were surface-sterilized and germinated on half-strength Murashige and Skoog (MS) basal medium. Two week-old seedlings were treated with 100  $\mu$ M MeJA or 10  $\mu$ M indole acetic acid (IAA) for 1 h, 8 h, or 24 h and frozen immediately in liquid nitrogen. Mock-treated seedlings were used as controls. Total RNA were isolated from 100 mg of control and MeJA-treated seedlings using miRVana miRNA isolation Kit with phenol (ThermoFisher

Scientific, USA). RNA quantity was determined using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). Quality of the RNA samples was determined using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA). RNA samples with RNA integrity number (RIN) above 8 were used for library preparation. For expression analysis of regulatory and structural genes in the TIA pathway, RNA were isolated from the control and MeJA- or IAA-treated seedlings using RNeasy plant mini kit (QIAGEN, USA).

**High-throughput sequencing of small RNAs.** Two micrograms of RNA from each sample was sent to the Sequencing and Genotyping Center at the Delaware Biotechnology Institute at the University of Delaware for small RNA library preparation and sequencing. The libraries were pooled together and sequenced on an Illumina HiSeq 2500. Deep sequencing was performed in triplicates for each treatment for a 50 cycle single end run. The data quality was checked at the Sequencing and Genotyping Center and sequencing reads were provided in the FASTq format.

**Analysis of small RNA sequencing data.** Low quality and contaminated reads were removed as described previously<sup>74</sup>. Sequences smaller than 18 nucleotide (nt) and larger than 24 nt were also removed. Reads ranging from 18 to 24 nt in length, were mapped to the *C. roseus* reference genome sequence<sup>41</sup> with the software, Bowtie2<sup>75</sup>, with no mismatches allowed. The unmapped sequences were removed from analysis. Reads, matched with the *C. roseus* genome, were compared with non-coding sRNAs deposited in NCBI GenBank and RNA family (Rfam) databases. Reads that matched ribosomal RNAs (rRNAs), transfer RNAs (tRNAs), small nucleolar RNAs (snoRNAs), and small nuclear RNAs (snRNAs), were filtered and removed. The remaining reads were used for downstream analyses.

**Identification of conserved and *C. roseus*-specific (novel) miRNAs.** For conserved cro-miRNA identification, sequence reads were aligned with known miRNA sequences from other plant species deposited in miRBase 21 (<http://www.mirbase.org/index.shtml>)<sup>42,76</sup> with a maximum of two mismatches. The unaligned reads were then subjected to miR-PREFeR pipeline to predict novel miRNA candidates. miR-PREFeR<sup>45</sup> utilizes expression patterns of miRNAs and follows the criteria for plant miRNA annotation<sup>46</sup> to accurately predict plant miRNAs from one or more small RNA-seq samples. The pipeline generates candidate regions and candidate mature sequences of each candidate region based on alignment depth of 20 reads. Predicted novel cro-miRNAs were further screened and validated in this study. The secondary structures of precursors for all conserved and novel cro-miRNA candidates were constructed using the RNAfold software<sup>77</sup>.

**Differential expression analysis of JA responsive cro-miRNAs.** Differential expression of cro-miRNAs was calculated from read count data using DESseq2 of the R package as described<sup>78</sup>. Cro-miRNAs with log<sub>2</sub> fold change differences  $\geq 1$ , p-value  $\leq 0.05$ , and false discovery rate (FDR)  $\leq 0.1$ , were considered to be differentially expressed. The R package 'gplots' was used for heat map generation of differentially expressed cro-miRNAs. Normalized count was calculated as described previously<sup>44</sup>.

**Prediction and annotation of cro-miRNA target genes.** The potential target prediction of all identified cro-miRNAs was conducted using the plant small RNA target server (psRNATarget; <http://plantgrn.noble.org/psRNATarget/>), a widely used web-based tool<sup>48</sup>. Candidate targets were analyzed by locally installed BLASTX search against the NCBI Nr database with the default parameters. BiNGO 3.0.3 plug-in<sup>79</sup> of Cytoscape<sup>80</sup> was used for gene ontology (GO) analysis of target genes of JA responsive cro-miRNAs.

**Quantitative RT-PCR validation of selected, differentially expressed cro-miRNAs and their potential targets.** Expression of cro-miRNAs were analyzed using the poly(T) adaptor RT-PCR method<sup>81</sup>. One  $\mu\text{g}$  of RNase-free DNase I treated total RNA was polyadenylated at 37 °C for 60 min in a 10  $\mu\text{l}$  reaction volume containing 0.08 units poly (A) polymerase. The polyadenylated RNA was then reverse transcribed in a 20  $\mu\text{l}$  reaction mix with SuperScript III Reverse Transcriptase (Invitrogen, USA) and oligo (dT) adaptor, following the manufacturer's instructions. For expression analysis of regulatory and structural genes in TIA pathway, synthesis of first strand cDNA from total RNA, and Quantitative RT-PCR (qRT-PCR), were performed as described previously<sup>13</sup>. The comparative cycle threshold (Ct) method (bulletin no. 2; Applied Biosystems, <http://www.applied-biosystems.com>) was used to measure the transcript levels. U6 small nuclear ribonucleic acid (snRNA) and Ribosomal Protein S9 (RPS9), were used as the normalization controls for miRNA and target mRNA, respectively. The primers used in RT-qPCR are listed in Supplementary Table S11. All experiments were performed using two biological replicates with three technical replicates. Specificity of miRNA primers was determined by cloning the PCR products into the pGEM-T Easy vector (Promega, USA), followed by sequencing.

**Poly(A) polymerase-mediated rapid amplification of cDNA ends.** Poly(A) polymerase-Mediated Rapid Amplification of cDNA Ends (PPM-RACE) was conducted using the protocol described previously<sup>82</sup> to map the cleavage sites of target transcripts. Briefly, total RNA (2  $\mu\text{g}$ ) isolated from *C. roseus* seedlings were polyadenylated with poly (A) polymerase and used for cDNA synthesis using oligodT primer with an adapter. The first strand cDNA served as template to amplify the cleaved products using adapter and gene-specific primers. The PCR products were cloned into the pGEM-easy vector and sequenced.

**Transcriptomic Analysis of MeJA-treated *C. roseus* tissues.** To analyze the effect of MeJA on *C. roseus* transcriptome, publically available RNAseq data were obtained from the sequence read archive database (accession number PRJNA185483). Raw reads were cleaned and filtered as described previously<sup>74</sup>. Finally, cleaned reads were mapped to the *C. roseus* reference sequence<sup>41</sup> with Bowtie2<sup>75</sup>, and Fragments Per Kilobase of transcript per Million mapped reads (FPKM) values were calculated by eXpress<sup>83</sup>.



**Plasmid construction, protoplast isolation and electroporation.** For the protoplast assay, reporter plasmids were generated by cloning the *TDC*, *STR*, or *G10H* promoters into a pUC vector containing the fire-fly luciferase and *rbcS* terminator. The effector plasmids were generated by cloning *CrARF16* into a modified pBlue-Script vector containing the cauliflower mosaic virus (CaMV) 35S promoter and *rbcS* terminator. A plasmid containing  $\beta$ -glucuronidase (*GUS*) reporter, controlled by the *CaMV35S* promoter and *rbcS* terminator, was used as an internal control in the protoplast assay. The reporter plasmids were electroporated, alone or with the effector plasmid, into tobacco protoplasts. Protoplast isolation from tobacco cell suspension cultures, electroporation, and measurement of luciferase and GUS activities in protoplasts were performed as described previously<sup>84</sup>.

## References

- Holland, J. F. *et al.* Vincristine treatment of advanced cancer: a cooperative study of 392 cases. *Cancer Res* **33**, 1258–1264 (1973).
- Menke, F. L. H., Parchmann, S., Mueller, M. J., Kijne, J. W. & Memelink, J. Involvement of the Octadecanoid Pathway and Protein Phosphorylation in Fungal Elicitor-Induced Expression of Terpenoid Indole Alkaloid Biosynthetic Genes in *Catharanthus roseus*. *Plant Physiol* **119**, 1289–1296 (1999).
- Ramani, S. & Chelliah, J. UV-B-induced signaling events leading to enhanced-production of catharanthine in *Catharanthus roseus* cell suspension cultures. *BMC Plant Biol* **7**, 61–61, doi: 10.1186/1471-2229-7-61 (2007).
- Vazquez-Flota, F., Carrillo-Pech, M., Minero-García, Y. & De Lourdes Miranda-Ham, M. Alkaloid metabolism in wounded *Catharanthus roseus* seedlings. *Plant Physiol Biochem* **42**, 623–628 (2004).
- Dutta, A., Sen, J. & Deswal, R. Downregulation of terpenoid indole alkaloid biosynthetic pathway by low temperature and cloning of a AP2 type C-repeat binding factor (CBF) from *Catharanthus roseus* (L). G. Don. *Plant Cell Rep* **26**, 1869–1878 (2007).
- Jaleel, C. A. *et al.* Alterations in osmoregulation, antioxidant enzymes and indole alkaloid levels in *Catharanthus roseus* exposed to water deficit. *Colloids and Surfaces. B, Biointerfaces* **59**, 150–157 (2007).
- Menke, F. L. H., Champion, A., Kijne, J. W. & Memelink, J. A novel jasmonate- and elicitor-responsive element in the periwinkle secondary metabolite biosynthetic gene *Str* interacts with a jasmonate- and elicitor-inducible AP2-domain transcription factor, ORCA2. *The EMBO J* **18**, 4455–4463 (1999).
- van der Fits, L. & Memelink, J. ORCA3, a Jasmonate-Responsive Transcriptional Regulator of Plant Primary and Secondary Metabolism. *Science* **289**, 295–297 (2000).
- Paul, P. *et al.* A differentially regulated AP2/ERF transcription factor gene cluster acts downstream of a MAP kinase cascade to modulate terpenoid indole alkaloid biosynthesis in *Catharanthus roseus*. *New Phytol* **213**, 1107–1123 (2017).
- Zhang, H. *et al.* The basic helix-loop-helix transcription factor CrMYC2 controls the jasmonate-responsive expression of the ORCA genes that regulate alkaloid biosynthesis in *Catharanthus roseus*. *Plant J* **67**, 61–71 (2011).
- Van Moerkercke, A. *et al.* The bHLH transcription factor BIS1 controls the iridoid branch of the monoterpenoid indole alkaloid pathway in *Catharanthus roseus*. *Proc Natl Acad Sci USA* **112** (2015).
- Van Moerkercke, A. *et al.* The basic helix-loop-helix transcription factor BIS2 is essential for monoterpenoid indole alkaloid production in the medicinal plant *Catharanthus roseus*. *Plant J* **88**, 3–12 (2016).
- Suttipanta, N. *et al.* The transcription factor CrWRKY1 positively regulates the terpenoid indole alkaloid biosynthesis in *Catharanthus roseus*. *Plant Physiol* **157**, 2081–2093 (2011).
- Sibérlil, Y. *et al.* *Catharanthus roseus* G-box binding factors 1 and 2 act as repressors of strictosidine synthase gene expression in cell cultures. *Plant Mol Biol* **45**, 477–488 (2001).
- Pauw, B. *et al.* Zinc finger proteins act as transcriptional repressors of alkaloid biosynthesis genes in *Catharanthus roseus*. *J Biol Chem* **279**, 52940–52948 (2004).
- Rizvi, N. F., Weaver, J. D., Cram, E. J. & Lee-Parsons, C. W. Silencing the Transcriptional Repressor, ZCT1, Illustrates the Tight Regulation of Terpenoid Indole Alkaloid Biosynthesis in *Catharanthus roseus* Hairy Roots. *PLoS One* **11**, e0159712, doi: 10.1371/journal.pone.0159712 (2016).
- Papon, N. *et al.* Cytokinin and ethylene control indole alkaloid production at the level of the MEP/terpenoid pathway in *Catharanthus roseus* suspension cells. *Planta Med* **71**, 572–574 (2005).
- Goddijn, O. M., de Kam, R., Zanetti, A., Schilperoord, R. & Hoge, J. H. Auxin rapidly down-regulates transcription of the tryptophan decarboxylase gene from *Catharanthus roseus*. *Plant Mol Biol* **18**, 1113–1120 (1992).
- Pasquali, G. *et al.* Coordinated regulation of two indole alkaloid biosynthetic genes from *Catharanthus roseus* by auxin and elicitors. *Plant Mol Biol* **18**, 1121–1131 (1992).
- Wang, R. & Estelle, M. Diversity and specificity: auxin perception and signaling through the TIR1/AFB pathway. *Current Opin Plant Biol* **21**, 51–58 (2014).
- Mallory, A. C., Bartel, D. P. & Bartel, B. MicroRNA-directed regulation of Arabidopsis AUXIN RESPONSE FACTOR17 is essential for proper development and modulates expression of early auxin response genes. *Plant Cell* **17**, 1360–1375 (2005).
- Curaba, J., Singh, M. B. & Bhalla, P. L. miRNAs in the crosstalk between phytohormone signalling pathways. *J Exp Bot* **65**, 1425–1438 (2014).
- Bartel, D. P. MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215–233 (2009).
- Rogers, K. & Chen, X. Biogenesis, turnover, and mode of action of plant microRNAs. *Plant Cell* **25**, 2383–2399 (2013).
- Baumberger, N. & Baulcombe, D. C. Arabidopsis ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs. *Proc Natl Acad Sci USA* **102**, 11928–11933 (2005).
- Brodersen, P. *et al.* Widespread translational inhibition by plant miRNAs and siRNAs. *Science* **320**, 1185–1190 (2008).
- Jones-Rhoades, M. W., Bartel, D. P. & Bartel, B. MicroRNAs and their regulatory roles in plants. *Annu Rev Plant Biol* **57**, 19–53 (2006).
- Sun, G. MicroRNAs and their diverse functions in plants. *Plant Mol Biol* **80**, 17–36 (2012).
- Meng, Y. *et al.* Genome-wide survey of rice microRNAs and microRNA-target pairs in the root of a novel auxin-resistant mutant. *Planta* **230**, 883–898 (2009).
- Khan, G. A. *et al.* MicroRNAs as regulators of root development and architecture. *Plant Mol Biol* **77**, 47–58 (2011).
- Meng, Y., Ma, X., Chen, D., Wu, P. & Chen, M. MicroRNA-mediated signaling involved in plant root development. *Biochem Biophys Res Commun* **393**, 345–349 (2010).
- Song, J. B., Huang, S. Q., Dalmay, T. & Yang, Z. M. Regulation of leaf morphology by microRNA394 and its target LEAF CURLING RESPONSIVENESS. *Plant Cell Physiol* **53**, 1283–1294 (2012).
- Juarez, M. T., Kui, J. S., Thomas, J., Heller, B. A. & Timmermans, M. C. microRNA-mediated repression of rolled leaf1 specifies maize leaf polarity. *Nature* **428**, 84–88 (2004).
- Wang, Z. J., Huang, J. Q., Huang, Y. J., Li, Z. & Zheng, B. S. Discovery and profiling of novel and conserved microRNAs during flower development in *Carya cathayensis* via deep sequencing. *Planta* **236**, 613–621 (2012).
- Lin, Y. & Lai, Z. Comparative Analysis Reveals Dynamic Changes in miRNAs and Their Targets and Expression during Somatic Embryogenesis in Longan (*Dimocarpus longan* Lour.). *PLoS ONE* **8**, e60337, doi: 10.1371/journal.pone.0060337 (2013).
- Zhai, L. *et al.* Genome-Wide Identification of Embryogenesis-Associated microRNAs in Radish (*Raphanus sativus* L.) by High-Throughput Sequencing. *Plant Mol Biol Rep* **32**, 900–915 (2014).

37. Jia, X. *et al.* Small tandem target mimic-mediated blockage of microRNA858 induces anthocyanin accumulation in tomato. *Planta* **242**, 283–293 (2015).
38. Gou, J.-Y., Felippes, F. F., Liu, C.-J., Weigel, D. & Wang, J.-W. Negative Regulation of Anthocyanin Biosynthesis in Arabidopsis by a miR156-Targeted SPL Transcription Factor. *Plant Cell* **23**, 1512–1522 (2011).
39. Sharma, D. *et al.* MicroRNA858 Is a Potential Regulator of Phenylpropanoid Pathway and Plant Development. *Plant Physiol* **171**, 944–959 (2016).
40. Prakash, P., Ghosliya, D. & Gupta, V. Identification of conserved and novel microRNAs in *Catharanthus roseus* by deep sequencing and computational prediction of their potential targets. *Gene* **554**, 181–195 (2015).
41. Kellner, F. *et al.* Genome-guided investigation of plant natural product biosynthesis. *Plant J* **82**, 680–692 (2015).
42. Kozomara, A. & Griffiths-Jones, S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res* **42**, D68–73 (2014).
43. Nie, S. *et al.* Identification of bolting-related microRNAs and their targets reveals complex miRNA-mediated flowering-time regulatory networks in radish (*Raphanus sativus* L.). *Scientific Rep* **5**, 14034, doi: 10.1038/srep14034 (2015).
44. Bi, F., Meng, X., Ma, C. & Yi, G. Identification of miRNAs involved in fruit ripening in Cavendish bananas by deep sequencing. *BMC Genomics* **16**, 1–15, doi: 10.1186/s12864-015-1995-1 (2015).
45. Lei, J. & Sun, Y. miR-PREFeR: an accurate, fast and easy-to-use plant miRNA prediction tool using small RNA-Seq data. *Bioinformatics* **30**, 2837–2839 (2014).
46. Meyers, B. C. *et al.* Criteria for annotation of plant MicroRNAs. *Plant Cell* **20**, 3186–3190 (2008).
47. Thakur, V. *et al.* Characterization of statistical features for plant microRNA prediction. *BMC Genomics* **12**, 108, doi: 10.1186/1471-2164-12-108 (2011).
48. Dai, X. & Zhao, P. X. psRNATarget: a plant small RNA target analysis server. *Nucleic Acids Res*, **39**, W155–W159 (2011).
49. Chi, X. *et al.* Identification and characterization of microRNAs from peanut (*Arachis hypogaea* L.) by high-throughput sequencing. *PLoS One* **6**, e27530, doi: 10.1371/journal.pone.0027530 (2011).
50. Hwang, D. G. *et al.* The hot pepper (*Capsicum annuum*) microRNA transcriptome reveals novel and conserved targets: a foundation for understanding MicroRNA functional roles in hot pepper. *PLoS One* **8**, e64238, doi: 10.1371/journal.pone.0064238 (2013).
51. Martinez, G., Forment, J., Llave, C., Pallas, V. & Gomez, G. High-throughput sequencing, characterization and detection of new and conserved cucumber miRNAs. *PLoS One* **6**, e19523, doi: 10.1371/journal.pone.0019523 (2011).
52. Morin, R. D. *et al.* Comparative analysis of the small RNA transcriptomes of *Pinus contorta* and *Oryza sativa*. *Genome Res* **18**, 571–584 (2008).
53. Song, C. *et al.* Deep sequencing discovery of novel and conserved microRNAs in trifoliolate orange (*Citrus trifoliata*). *BMC Genomics* **11**, doi: 10.1186/1471-2164-11-431 (2010).
54. Axtell, M. J. Classification and comparison of small RNAs from plants. *Annual review of plant biology* **64**, 137–159, doi: 10.1146/annurev-arplant-050312-120043 (2013).
55. Vazquez, F. Arabidopsis endogenous small RNAs: highways and byways. *Trends in Plant Sci* **11**, 460–468 (2006).
56. Palatnik, J. F. *et al.* Sequence and Expression Differences Underlie Functional Specialization of Arabidopsis MicroRNAs miR159 and miR319. *Dev Cell* **13**, 115–125 (2007).
57. Rodriguez, R. E. *et al.* Control of cell proliferation in Arabidopsis thaliana by microRNA miR396. *Development* **137**, 103–112 (2010).
58. Xie, Z., Kasschau, K. D. & Carrington, J. C. Negative Feedback Regulation of Dicer-Like1 in Arabidopsis by microRNA-Guided mRNA Degradation. *Curr Biol* **13**, 784–789 (2003).
59. Valiollahi, E., Farsi, M. & Kakhki, A. M. Sly-miR166 and Sly-miR319 are components of the cold stress response in *Solanum lycopersicum*. *Plant Biotech Rep* **8**, 349–356 (2014).
60. Kitazumi, A., Kawahara, Y., Onda, T. S., De Koeyer, D. & de los Reyes, B. G. Implications of miR166 and miR159 induction to the basal response mechanisms of an andigena potato (*Solanum tuberosum* subsp. andigena) to salinity stress, predicted from network models in Arabidopsis. *Genome* **58**, 13–24 (2015).
61. Chavez Montes, R. A. *et al.* Sample sequencing of vascular plants demonstrates widespread conservation and divergence of microRNAs. *Nature Comm* **5**, 3722 (2014).
62. Cupepus, J. T., Fahlgren, N. & Carrington, J. C. Evolution and functional diversification of MIRNA genes. *Plant Cell* **23**, 431–442 (2011).
63. Axtell, M. J. Evolution of microRNAs and their targets: Are all microRNAs biologically relevant? *Biochim et Biophys Acta (BBA) - Gene Regulatory Mechanisms* **1779**, 725–734 (2008).
64. Fahlgren, N. *et al.* High-throughput sequencing of Arabidopsis microRNAs: evidence for frequent birth and death of MIRNA genes. *PLoS One* **2**, e219, doi: 10.1371/journal.pone.0000219 (2007).
65. Aerts, R. J., Schäfer, A., Hesse, M., Baumann, T. W. & Slusarenko, A. Signalling molecules and the synthesis of alkaloids in *Catharanthus roseus* seedlings. *Phytochem* **42**, 417–422 (1996).
66. Gantet, P., Imbault, N., Thiersault, M. & Doireau, P. Necessity of a Functional Octadecanoic Pathway for Indole Alkaloid Synthesis by *Catharanthus roseus* Cell Suspensions Cultured in an Auxin-Starved Medium. *Plant Cell Physiol* **39**, 220–225 (1998).
67. Rijhwani, S. K. & Shanks, J. V. Effect of elicitor dosage and exposure time on biosynthesis of indole alkaloids by *Catharanthus roseus* hairy root cultures. *Biotech Prog* **14**, 442–449 (1998).
68. Qiu, D. *et al.* High throughput sequencing technology reveals that the taxoid elicitor methyl jasmonate regulates microRNA expression in Chinese yew (*Taxus chinensis*). *Gene* **436**, 37–44 (2009).
69. Guilfoyle, J. T., Ulmasov, T. & Hagen, G. The ARF family of transcription factors and their role in plant hormone-responsive transcription. *Cell Mol Life Sci* **54**, 619–627 (1998).
70. Tiwari, S. B., Hagen, G. & Guilfoyle, T. The Roles of Auxin Response Factor Domains in Auxin-Responsive Transcription. *Plant Cell* **15**, 533–543 (2003).
71. Li, S.-b., Xie, Z.-Z., Hu, C.-G. & Zhang, J.-Z. A review of auxin response factors (ARF) in plants. *Front Plant Sci* **7**, doi: 10.3389/fpls.2016.00047 (2016).
72. Hagen, G. & Guilfoyle, T. Auxin-responsive gene expression: genes, promoters and regulatory factors. *Plant Mol Biol* **49**, 373–385, doi: 10.1023/a:1015207114117 (2002).
73. Wang, J.-W. *et al.* Control of Root Cap Formation by MicroRNA-Targeted Auxin Response Factors in Arabidopsis. *Plant Cell* **17**, 2204–2216 (2005).
74. Singh, S. K. *et al.* RNA-sequencing Reveals Global Transcriptomic Changes in *Nicotiana tabacum* Responding to Topping and Treatment of Axillary-shoot Control Chemicals. *Scientific Rep* **5**, 18148, doi: 10.1038/srep18148 (2015).
75. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Meth* **9**, 357–359 (2012).
76. Kozomara, A. & Griffiths-Jones, S. miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res* **39**, D152–157 (2011).
77. Lorenz, R. *et al.* ViennaRNA Package 2.0. *Algorithms for Molecular Biology* **6**, 1–14, doi: 10.1186/1748-7188-6-26 (2011).
78. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 1–21 (2014).
79. Maere, S., Heymans, K. & Kuiper, M. BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics* **21**, 3448–3449 (2005).



80. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498–2504 (2003).
81. Fiedler, S. D., Carletti, M. Z. & Christenson, L. K. In *RT-PCR Protocols: Second Edition* (ed Nicola, King) 49–64 (Humana Press, 2010).
82. Wang, C. & Fang, J. RLM-RACE, PPM-RACE, and qRT-PCR: an integrated strategy to accurately validate miRNA target genes. *Met Mol Bio* **1296**, 175–186 (2015).
83. Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotech* **28**, 511–515 (2010).
84. Bai, Y. *et al.* Flavonoid-related basic helix-loop-helix regulators, NtAn1a and NtAn1b, of tobacco have originated from two ancestors and are functionally active. *Planta* **234**, 363–375 (2011).

### Acknowledgements

We thank Kathy Shen for critical reading of the manuscript. This work is supported partially by the Harold R. Burton Endowed Professorship to L.Y. and by the National Science Foundation under Cooperative Agreement No. 1355438 to L.Y.

### Author Contributions

L.Y. and S.P. designed research; E.S., S.S., B.P., J.G. and P.P. performed experiments; E.S., S.S. and S.P. analyzed data; S.S., S.P. and L.Y. wrote the paper.

### Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Shen, E. M. *et al.* The miRNAome of *Catharanthus roseus*: identification, expression analysis, and potential roles of microRNAs in regulation of terpenoid indole alkaloid biosynthesis. *Sci. Rep.* **7**, 43027; doi: 10.1038/srep43027 (2017).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017