Genomics xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Exploring microRNA profiles for circadian clock and flowering development regulation in Himalayan *Rhododendron*

Shruti Choudhary, Sapna Thakur, Aasim Majeed, Pankaj Bhardwaj*

Molecular Genetics Laboratory, Department of Plant Sciences, Central University of Punjab, City Campus, Mansa Road, Bathinda 151001, India

ARTICLE INFO

Keywords:
Circadian clock
Development
Gene regulation
MicroRNA profiles
Rhododendron
Small RNA sequencing

ABSTRACT

miRNA is a non-coding, yet crucial entity in remodeling the genetic architecture. *Rhododendron arboreum* of Himalayas grows and even flower under fluctuating climate. sRNA from leaves of vegetative and reproductive periods was sequenced to elucidate its seasonal associations. Conserved (256) and novel (210) miRNAs and their precursors were located based on homology with plant databases and transcriptome of the species. 27,139 predicted targets were involved with metabolism, reproduction, and response to abiotic stimuli. A comparative analysis showed differential expression of 198 miRNAs with season-specific abundance of 103 miRNAs. Specific isoforms of 11 miRNA families exhibited a temporal expression and targeted different genes implying a complex regulation. The variable miRNA expression among the tissues of different conditions can be associated with the adaptability of the species, which will prove essential for further study on miRNAs mediating seasonal response. Moreover, exogenous cues also mediate phase transition via networking of flowering pathways and their components. In this context, 18 known families and 77 novel miRNAs modulating 117 genes crucial in circadian entrainment were filtered. A negative correlation was obtained between the expression of 18 of these miRNAs and their targets when tested through quantitative-PCR. It highlighted the role of miRNA-target pairs in perceiving environmental variabilities and monitoring flowering growth. Furthermore, a phylogenetic clustering was performed, which supported the lineage-specific evolution and function of putative miR156 sequence in the species. This documentation of genome-wide profiling of miRNA, their targets, and expression will enhance the understanding of developmental and climate-tolerance strategies in high-altitude trees.

1. Introduction

Abiotic factors test the survival capacity of living organisms. How these environmental cues trigger and schedule the growth activities is a matter of investigation. A transition in developmental phase is mediated by a gradient in molecular signal and in the cellular ability to respond to these signals. Besides the protein-coding components, the non-coding (nc) regions of the genome affect the environment-sensing ability of an organism and vice versa [1]. The production and movement of a diffusible RNA signal from leaves followed by its action against floral repressors activate a set of downstream genes associated with the developmental phase changes. Henceforth, it can be concluded that ncRNA or small RNA (sRNA) is a connecting link between the ambient environment and meristem transition [2].

MicroRNAs (miRNAs) are 19–25 nucleotide (nt) long, endogenous ncRNAs having stem-loop structures. These regulatory molecules originate from the processing of long precursors known as, pri-miRNAs. miRNAs negatively regulate the gene expression either by cleavage or translational repression at the post-transcriptional level or by methylation at the transcriptional level [3].

A crucial player in the adaptation of an organism is the circadian clock. The biological system recognizes the variabilities in environmental factors and induces physiological and behavioral reprogramming [4]. The machinery comprises of perception and relaying tools, the central oscillator as time-keeper, and the output biochemical and developmental processes. Though the clock is self-sustaining, the internal oscillators can also synchronize themselves with the local environment through the process of entrainment. The circadian feedback including the “core” clock genes not only maintains their own rhythm but also influence a set of downstream clock-controlled genes. Besides handling the morphological and diurnal activities, the clock is involved in the temporal regulation of genes associated with adaptive or phase transitions in plants [4,5].

A two-way coordination can be seen among miRNA accumulation levels and the components of circadian rhythm as well as external factors such as light and temperature. Photoperiod-dependent

* Corresponding author.
E-mail address: pankajbhardwaj@cup.edu.in (P. Bhardwaj).

https://doi.org/10.1016/j.ygeno.2018.09.019
Received 19 July 2018; Received in revised form 23 August 2018; Accepted 25 September 2018

flowering is one of the output processes of clock-mediated pathways that had been recently shown under miRNA regulation [6]. Similarly, silencing via miRNAs connects the temperature and the resulting shifts in the functional gene expression [7]. Several miRNA family members such as mir172, mir156, mir159, mir398 etc., have been reported to control the floral timing and other developmental responses under normal physiological as well as during multiple stress conditions. The miRNA-mediated circadian control can further determine the reproductive success of a plant by ensuring that there is a harmony between blooming time and favorable season.

Rhododendron arboreum of Ericaceae constitute the native flora of the Himalayas and has become an economically important species. The variations in phenology in response to global deviations in climate [8,9] has made it enticing to understand the allied mechanisms. A differential transcriptome profiling in the species earlier demonstrated that the rhythmic processes were induced among the two developmental seasons (manuscript under review). However, genome-scale data to explain the miRNA-directed regulation of the biological time-keeping system in Rhododendrons is still unavailable. In the present study, sRNA was sequenced using the high-throughput technology to demonstrate the temporal control elicited by miRNAs in the tree species. The study provides evidence for the functional involvement of sRNA during season/development transitions and emphasizes to view circadian clock and its components with a new perspective.

2. Materials and methods

2.1. Plant material and sRNA sequencing

Leaf tissues from same plant of *R. arboreum* were sampled from two sites of Himachal Pradesh (31.945910°N; 76.894685°E and 31.726562°N; 76.862830°E). The samples were collected approximately 2½ h after dawn and the duration of daylength was 11 ± 1 h. The tissues were collected first in February–March signifying reproductive growth and later in October for miRNA expression analysis during vegetative growth. The tissues were frozen immediately in liquid nitrogen and stored at −80 °C until further use. For the genome-wide discovery of miRNAs and their role in the regulation of seasonal development, total RNA enriched with sRNA was isolated. The quality and integrity were checked on Bioanalyzer (Agilent Technologies, USA). sRNA fraction was isolated from the total RNA using the mirVana Kit (Ambion, USA) followed by size-separation on denaturing gel. sRNA from leaves of flowering and vegetative season were pooled together as pool I and II, respectively.

The sRNA sequencing libraries were prepared using NEB NextR Multiplex sRNA Library Prep Set for Illumina (NEB, USA.) following the manufacturer's instructions. 5′ and 3′ adapters were ligated to the sRNAs, reverse transcribed, and amplified. The sequencing was performed using Illumina NextSeq500 Single-end sequencing technology.

2.2. sRNA analysis

5′ and 3′- adapter sequences as well as low quality (Q value < 10) or short reads were removed using UEA sRNA workbench v3.0_ALPHA [10]. The pre-processed reads were filtered based on their length (16–35 nt). Each read was counted as a tag and length distribution of these unique tags was recorded. The tags were then searched for other types of ncRNAs available at NCBI (rRNA, tRNA, snRNA, and snoRNA). The resulting unaligned reads were clustered and filtered based on 100% coverage and similarity to generate read count. The tags were also mapped to the *R. arboreum* transcriptome data (NCBI accession No.: SRP092027) generated from an earlier study by our group [11]. A brief of the workflow adopted for the present analyses is presented in Fig. 1.

2.3. miRNA prediction

Clustered reads (16–35 nt long) were used for prediction of known or conserved miRNA sequences. The homology search was done using BLASTN at an e-value cut-off of 0.001 against the known miRNA precursors and mature miRNAs of plant kingdom as retrieved from the miRBase database (release 21.0) [12]. Tags with > 1 read count, with ≤1 mismatch, and at 95–100% identity with the miRBase were considered as conserved counterparts and retained for further analysis.

Tags not matching any of the above databases were considered unclassified and were processed for novel miRNA prediction. The tags with ≥5 read counts were kept for the analysis to reduce false positives. Additionally, the tags perfectly mapping to *R. arboreum* transcriptome were taken as the potential precursors. Putative miRNA candidates were predicted using the miRCat tool [10]. Once the sequences were mapped to the reference, the regions covered with sRNAs were designated as the putative sRNA loci and the most abundant reads were referred to as the likely novel miRNA sequences. The sequences were mapped to the miRBase to ensure their novel nature. Flanking sequences surrounding this sRNA were extracted from the reference and folded using RNAfold to produce a secondary structure. miRCat further trimmed the secondary structure based on certain features: (i) < 3 adjacent mismatches between miRNA and miRNA*; (ii) GC content of 25–70%; (iii) hairpin length = 60 nt, with ≥50% pairing; (iv) sRNA length = 20–22; and (v) adjusted minimum free energy or δG (AMFE; signifies the stability of RNA secondary structure) of the hairpin = −25.

2.4. miRNA target prediction and annotations

miRNAs with copy number ≥5 were considered for target prediction by the miRanda-3.3a tool along with reference transcriptome. A strict alignment in the seed region (offset positions 2–8) was set to prevent the detection of the target sites containing gaps or non-canonical base pairing. miRNA hits having δG ≤ −25 were assumed as the targets of the reported conserved miRNAs. Similarly, the targets of putative novel miRNAs were predicted using the psRNATarget webserver following default parameters [13].

The target genes were searched for homology with the coding sequences of a recently sequenced closely related species, *R. delavayi* and also with the TAIR10 database. The gene ontology (GO) terms were assigned to the predicted targets with the help of PANTHER. Enrichment analysis of each term was also performed using WebGestalt server at the adjusted *p*-value of < 0.05. The genes encoding for specific transcription factors (TFs) and pathways were identified by aligning against the plant TF database and KAAS web-servers, respectively. Further, the genes regulating the circadian system and flowering time were filtered based on KEGG and literary references.

2.5. Differential expression of miRNAs

Read count table for sRNA reads obtained for both the pools were generated and compared. The differential expression analysis was carried based on the normalized read-count using the DESeq tool. The expression levels of miRNA targets were further validated in the leaf tissues of the two seasons by quantitative real-time (q-RT) PCR. cDNA was prepared using the High Capacity cDNA Reverse Transcription kit (Applied Biosystems, USA). PCR was done in the Step OnePlus™ RT-PCR machine (Applied Biosystems, USA) using PowerUp™ SYBR™ Green master mix (Applied Biosystems, USA) and specific primers.

The thermal cycler profile comprised of an initial denaturation of 10 min at 95 °C, followed by 40 cycles each of denaturation for 30 s at 95 °C, annealing of 30 s at 58 °C, and elongation of 1 min at 72 °C, and a final melting curve step. Ct value or the threshold cycle of each target gene was normalized according to β-Actin (taken as endogenous control). The relative quantities of transcripts were calculated using the δCt

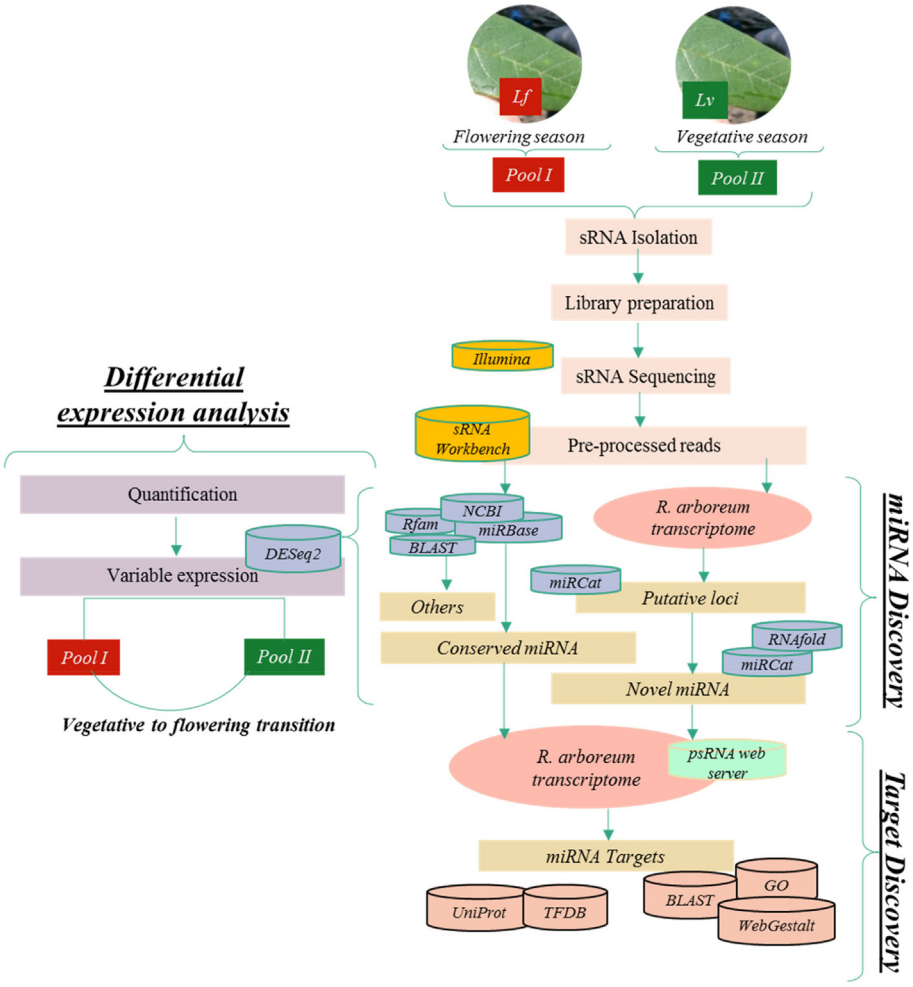


Fig. 1. A scheme for small RNA analysis in *R. arboreum*.

method and the expression fold-change was transformed to $\log_2$ scale.

2.6. Phylogenetic analysis

For the phylogenetic tree of the miR156 family, the validated sequences of mature plant miRNAs were downloaded from miRBase [12] and searched against the miRNA sequences of *R. arboreum* at the e-value cut-off of 1e-10. Then the sequences were aligned using MUSCLE. Finally, the phylogenetic tree was constructed using the Maximum Likelihood (ML) method by IQ-TREE based on ultrafast bootstrap [14] and SH-like approximate likelihood ratio test (SH-alrt) [15] with 1000 replicates. Phylogram was visualized using Archaeopteryx tool [16].

3. Results

3.1. sRNA sequencing

Two sRNA libraries constructed from the mature and healthy tissues of the reproductive and vegetative season were sequenced on the Illumina's platform. > 20 million reads were generated for the analysis. Pre-processing to remove low-quality and duplicate reads, adapter sequences and size selection between 16 and 35 nt reduced the total read count to ~15 million (Table 1). 21, 22, 23, and 24-nt sRNA species accounted for 18.4%, 7.1%, 8.9, and 20.4% respectively, representing 55% of the lot (Fig. S1a). After removal of the reads matching the structural ncRNAs (snoRNA, snRNA, tRNA, and rRNA), the remaining 3.2 million reads were used for miRNA prediction.

Table 1
Statistics for the sRNA library generated for *R. arboreum*.

Feature	Pool I	Pool II
Raw reads (in million) (unique)	11.5 (4.4)	9.9 (2.4)
Trimmed reads (in million) (unique)	9.2 (1.7)	8.9 (1.9)
Filtered reads 16–35 nt long (in million) (unique)	7.9 (1.5)	7.1 (1.7)
rRNA	302	237
snoRNA	68	63
snRNA	56	64
tRNA	16	12
Total structural sRNA	442	376

3.2. Discovery of miRNAs and their loci

Based on sequence homology with the plant-specific miRNAs, conserved miRNAs were identified. A total of 4174 and 4477 unique sRNA reads (16–35 nt) were aligned to the miRBase. As a result, 198 and 197 conserved miRNAs were identified in the pool I and II, respectively (Table 2). Briefly, a total of 240 conserved miRNAs were predicted for the two pools that fell under the category of 51 known miRNA families (Tables S2a and S2b).

Secondly, for the identification of novel miRNAs, the sRNA reads from the two pools were mapped to the *R. arboreum* transcriptome [11]. 169,399 reads with the abundance of ≥5 were retained for further screening. The mapped 91,477 reads were then used for prediction of novel miRNAs using the miRCat tool (Table 2). Precursor sequences were extracted by extending the mapped reads in the flanking regions

Table 2
A summary of conserved miRNA and novel miRNA discovery.

Feature	Pool I	Pool II
Conserved miRNA discovery using miRBase		
Reads aligned to miRBase	4174	4477
Length of the aligned reads (nt)	19–35	
Known unique miRNA	198	197
Length of the subject miRNA (nt)	18–24	
Known miRNA with read count ≥50	47	43
Known miRNA with read count ≥10	89	85
miRNA family	47	47
Total conserved miRNAs	**240**	
Total conserved miRNA families	**51**	
Novel miRNA discovery using miRCat		
Length of the input (nt)	16–35	
Filtered reads with abundance ≥5 (million)	11.7	
Input unique sequences	169,279	
Reads aligned to the *R. arboreum* transcriptome	91,477	
Potential sRNA precursors or loci reported	442	
Potential miRNA candidates	226	
miRNA candidates with similarity to the known miRNA	16	
Total novel miRNAs	210	

and their ability to form a stem-loop structure. The potential precursor sequences (442) were extracted and folded with the RNAfold package followed by a final processing through the miRCat pipeline (Table S2c).

We located 226 non-redundant novel miRNA candidates from the sRNA sequencing data procured from the two pools (Table 2). In addition to the criteria described in the materials and methods section to reduce false positives, the stability of secondary structures of the miRNA precursors was also tested statistically. 16 novel miRNAs exhibited similarity with the conserved miRNAs with a slight difference of sequence or length. The main families among them were: miR156, miR157, miR159, miR160, miR162, miR166, miR167, miR169, miR172, miR396, miR398, miR535, and miR7122 (Table S2c). All the other sequences that were not categorized into any of the conserved families were finally termed as putatively novel miRNAs (Table S2d). Summarizing the miRNA discovery, a total of 466 miRNAs were identified, of which, 256 were conserved and the rest (210) were considered novel.

3.3. Target discovery

To understand the biological role of miRNA, predictions of their targets is necessary. Targets of both conserved and novel identified miRNAs were predicted using the miRanda tool and psRNATarget webserver, respectively. Of the 466 miRNAs, 4517 and 22,874 putative targets were predicted for 117 known and 197 novel miRNAs, respectively. In total, 308 miRNAs targeted 27,139 transcripts (Table S2e). The number of targets for the conserved miRNAs varied from 1 to 81 whereas, for novel ones, it ranged from 95 to 234 (Table 3). 81.2% of the novel miRNA targets were regulated through cleavage and the remaining by translational repression (details not shown).

Table 3
A summary of target discovery.

Feature	Pool I	Pool II
Target discovery for conserved miRNAs		
miRNAs for which targets were predicted	88	84
Target transcripts	1934	1860
Unique targets	1544	1479
Targets per miRNA	1–81	
miRNA per target	1–9	
Target discovery for novel miRNA		
miRNAs for which targets were predicted	197	
Target transcripts	14,329	10,105
Targets per miRNA	95–234	
miRNA per target	1–26	

3.4. Functional diversity of miRNA targets

The GO and pathway analysis revealed that the highest proportion of genes were linked to metabolism, especially biosynthesis of secondary metabolites (Tables S3a and S3b). Following enrichment analysis, distinct processes that were being targeted by miRNAs included mRNA (192), ncRNA (180), amino acid (188), and cofactor (180) metabolic process; seed/embryo development (181); response to cold (137) etc. The significant pathway hits were: carbon, galactose, alanine-aspartate-glutamate, and glyoxylate-dicarboxylate metabolism; mRNA surveillance; and glycolysis/ gluconeogenesis.

Additionally, the conserved PFAM domain analysis was accomplished for 6054 targets. It was observed that majority of these targets had a phosphokinase domain, followed by pentatricopeptide repeat, F-box, NB-ARC, WD40, leucine-rich repeat motif/domain etc. (Table S3a). Furthermore, 3648 targets were categorized under 24 protein classes. The major ones included oxidoreductase, transporter, transferase, hydrolase, and nucleic acid binding in a proportion of 10.2%, 10.4%, 13.7%, 16.6%, and 17.8%, respectively. Similarly, another set of 1254 targets were the TF-encoding genes belonging to 57 families, with bHLH, FAR1, C3H, MYB, C2H2, and ARF in higher proportions (Table S3a). On the other hand, 37% of the potential miRNA targets might have an unknown function as they could not be annotated.

3.5. Expression levels of miRNAs

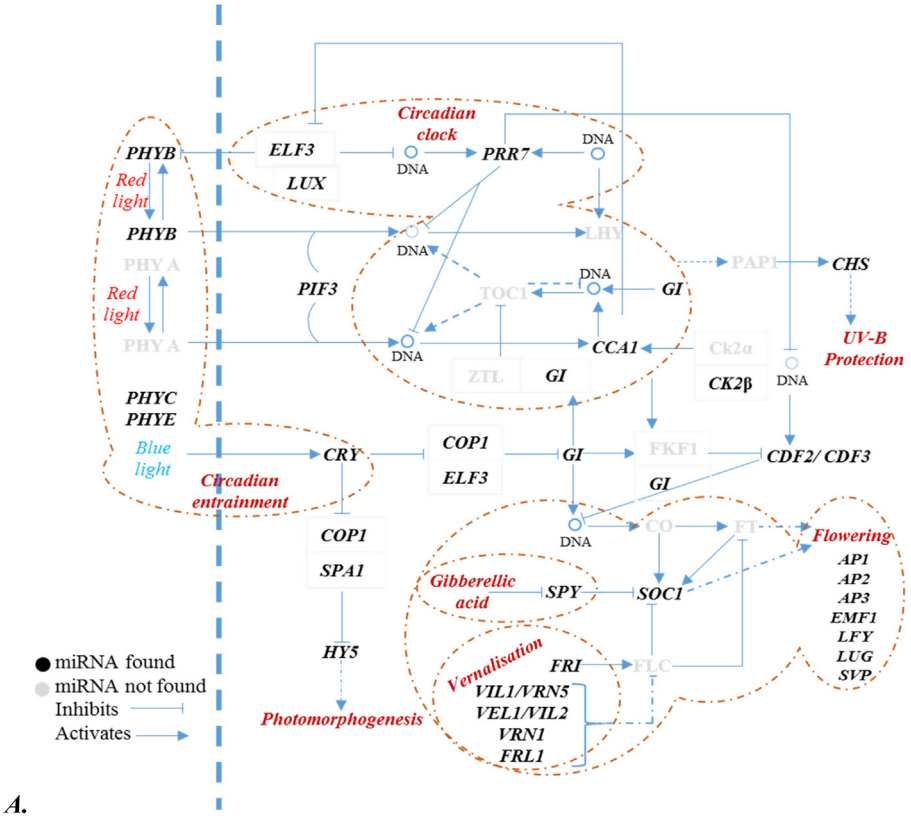
To gain insights into the putative roles of miRNAs, the expression pattern of the novel and known miRNAs were compared across the two pools by DESeq (Tables S4a and S4b). A large fraction of the conserved miRNAs (65% of 240) belonging to 42 families were expressed in both the pools (Table S4a). 41 and 44 miRNAs were up and downregulated in pool I, respectively whereas 43 and 42 miRNAs were exclusive to pool I and II, respectively (Table 4). Similarly, the expression pattern of novel miRNAs also varied between the two pools (Table S4b). Of the 179 novel miRNAs expressed in both the pools, 48 and 65 were found up and downregulated, respectively in pool I. On the other hand, 10 and 8 each were exclusive to pool I and II, respectively (Table 4). Altogether, the temporal expression patterns revealed here will help in elucidating the specific roles of miRNAs in plant development.

3.6. Dissecting the miRNA-mediated regulation of circadian clock and flowering genes

84 gene families represented by 388 targets were involved with the circadian clock and the output flowering time/development pathway. Of all these transcripts, 225 targets expressed in both the pools were differentially regulated. The structural genes related to circadian clock and their putative regulated miRNA families have been enlisted in Fig. 2. We observed that the targets encoded for proteins like COP1, ELF3/4, LUX, Ck2β, CHS, GI, FKF1, HY5, CRY, PHYB/C/E, PRR7, FKF1,

Table 4
A summary of differential expression for miRNA.

Statistics	Total	Up	Down	Neutral
For conserved miRNAs				
Expressed in both samples	155	44	41	70
Expressed only in flowering season	43			
Expressed only in vegetative season	42			
For novel miRNAs				
Expressed in both samples	179	48	65	66
Expressed only in flowering season	10			
Expressed only in vegetative season	8			



A.

Gene	miRNA
AP1 (APETALA 1)	ra_pnov_cluster51
AP2 (APETALA 2)	ra_pnov_cluster186, ra_pnov_cluster23, miR172
AP3 (APETALA 3)	ra_pnov_cluster148
PRR7 (Pseudo-response regulator 7)	gma-miR159a-5p, ra_pnov_cluster182
CCA1 (Circadian clock Associated 1)	ra_pnov_cluster89
CDF2 (Cyclic dof factor 2)	ra_pnov_cluster118
CDF3 (Cyclic dof factor 3)	ra_pnov_cluster31
CHS (Chalcone synthase)	ra_pnov_cluster167
COP1 (Ubiquitin-protein ligase)	ra_pnov_cluster125, ra_pnov_cluster158
CRY1 (Cryptochrome-1)	mdm-miR1511, ra_pnov_cluster54
ELF3 (EARLY FLOWERING 3)	ra_pnov_cluster167, ra_pnov_cluster31
FRI (FRIGIDA)	ath-miR8175, ra_pnov_cluster23
FRL1 (FRIGIDA-like)	ra_pnov_cluster9
GI (GIGANTEA)	ra_pnov_cluster128, vvi-miR845c
HY5 (Transcription factor)	ra_pnov_cluster195, ra_pnov_cluster23
LFY (Leafy)	ra_pnov_cluster17
LUG (Transcriptional corepressor LEUNIG)	ra_pnov_cluster51
LUX (LUX ARRHYTHMO)	ra_pnov_cluster106, ra_pnov_cluster28
PHYB (Phytochrome B)	ra_pnov_cluster98
PHYC (Phytochrome C)	gma-miR398d, ra_pnov_cluster197
PHYE (Phytochrome E)	ra_pnov_cluster169
SOC1 (Suppressor of Constans 1)	ata-miR167e-5p, ra_pnov-cluster77
PIF3 (Phytochrome-interacting factor 3)	ra_pnov_cluster19, ra_pnov_cluster29
SPY (SPINDLY)	bra-miR162-3p, ra_pnov_cluster145
SPA4 (Protein SPA1 related 4)	ra_pnov_cluster84
SVP (Short vegetative protein)	zma-miR159c-5p
VIL1 (VIN3-like protein 1)	ppe-miR211d, ra_pnov_cluster167
VIL2 (VIN3-like protein 2)	bgy-miR156, ra_pnov_cluster144
VRN1 (Vernalization 1)	ra_pnov_cluster182
CK2b (Casein kinase II subunit beta)	ra_pnov_cluster45, ra_pnov_cluster54

B.

neutrally expressed
downregulated during flowering
upregulated during flowering

Fig. 2. The enzyme-coding structural genes involved in the rhythmic processes as constructed from KEGG and literary references, (A) miRNAs for the targets (colored in **black)** were found during sRNA profiling; (B) Various genes associated with their putative miRNAs, whose expression levels in the two seasons are summarized by color coding.

CCA1, CDF1/2/3, and PIF3. Here, 27 novel miRNAs and 10 conserved miRNA families were the prime regulators (Fig. 2). For the flowering time, initiation, and development controllers, 13 conserved miRNA families as well as 64 novel miRNAs were demonstrated (Table S1b). SVP, LUG, LFY, EMF1, AP1/2/3, SOC1, SPY, SPL, FRI, AGL6/16, EFL4, FIE, FLK, FLX, FPA, GAI, GID1B, and VRN1 were the major targets.

To validate the miRNAs involved in the circadian pathway regulation, we performed qRT-PCR on their respective targets. The selected target genes of 18 miRNAs were PHYE, PIF3, EMF1, CCA1, COP1, PHYB, ERK, FLK, GI, AMPK, SPL13b, AP2, FAS2, FRL1, LUX, Cke/s, HEN4, and FPA. The sequence of the primers designed is enlisted in Table S1c. For most of the targets, an inverse relation was obtained between the expression patterns of miRNA and their targets (Fig. 3). However, the same was not true for three of these targets (HEN4, FAS2, and GI), which might be due to the fact that some other mechanism must be regulating their expression.

3.7. Phylogenetic analysis of miR156 family in R. arboreum

To access the phylogenetic features of the highly conserved miRNA in *R. arboreum,* the stem-loop miRNA precursors were aligned to the miRBase [12]. It was observed that 16 hairpins were homologous to that of other plants (Table S2d). Among the highly conserved miRNA sequences, ra_pnov_cluster158 annotated to the miR156 family of the

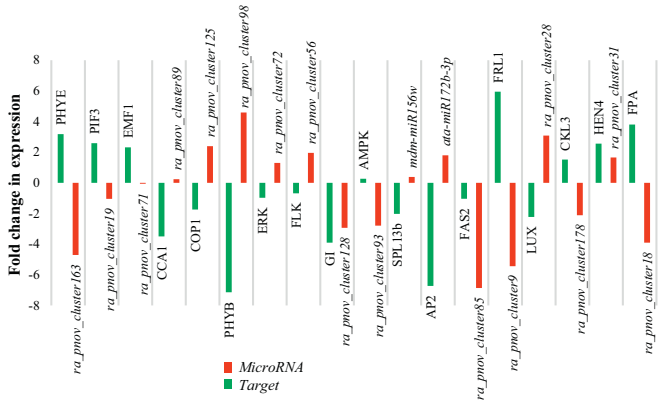


Fig. 3. Change in the expression profiles of selected genes and their microRNAs in the leaves of flowering and vegetative season as obtained via quantitative PCR and small RNA sequencing data, respectively.

plant miRNA database was examined in detail. Since the sequence similarity of conserved miRNA also lie in its precursor, a phylogenetic tree based on the pre-miRNA sequences of the plant miR156 family was constructed (Fig. 4). The mature miRNA sequence (ra_pnov_cluster158) formed a subclade with the corresponding members in *Manihot esculenta* (mes-miR156k), *Citrus sinensis* (csi-miR156f), *Ricinus communis* (rco-miR156e), *Theobroma cacao* (tcc-miR156a), and *Vitis vinifera* (vvi-miR156a). Around 151 reads had a match with the pre-miR156 precursor (Table S2c). It was further confirmed that the reads and precursor followed this rule in the complementary (*) region also.

4. Discussion

The ncRNAs via their silencing modes negatively regulate gene expression. The accumulation pattern of miRNAs is known to make striking differences during different developmental stages in the lifespan of an organism. The high-throughput sequencing technology has provided a massive amount of genomic information, especially in non-model organisms. The genome-wide sRNA profile generated for *R. arboreum* distinguished the biological role of miRNAs during and after the reproductive season, characterized by low and high temperatures, respectively. It was established that various miRNAs were expressed varyingly in the two seasons. As our understanding develops, miRNAs also modulated the developmental processes, especially the circadian clock and its output pathways, which were studied in detail.

29% of the total reads were removed from the sRNA sequencing data to avoid the selection of contaminants. The 24-nt long miRNAs dominated the present dataset similar to other studies [17,18]. The size distribution further indicated the high-quality of the sequences generated (Fig. S1a). Taking into account the transcriptome of *R. arboreum* sequenced [11] and miRBase, both conserved and novel miRNAs were identified from the sRNA libraries prepared from the two seasons. The commonly employed algorithms also led to the prediction of novel miRNAs and their putative loci. ~54% of the unique reads mapping to the transcriptome were used to predict miRNA precursors having an optimum AMFE value.

Sequence analysis classified the conserved miRNAs into 51 families postulating their requirement during a change in season or growing phase. A major proportion of these miRNAs was found conserved in Arabidopsis (34), *Glycine* max (30), rice (26), and so on (Tables S2a and S2b). miR159 was detected as the most abundant miRNA family in the present study and was predicted to target 640 unigenes. It is one of the ancient miRNA family in the plant kingdom and is also the most dominant one in Arabidopsis. The usual targets of miR159 are the genes encoding for the TFs having R2R3 MYB domain [19,20]. In addition to the MYB TFs, we predicted 158 putative targets of miR159. These genes encoded for ABC transporters, ARR−B, B3, G2-like, HSF, MADS, bZIP, and ERF proteins (Table S3a), and those moderating biological clock and development such as SVP, PRR7, BRM, MYB33, and ABF1 (Table S1b).

27,139 miRNA targets were identified by the target-prediction algorithms. Of these, 11% and 59% unigenes were the targets of 309 conserved and novel miRNAs, respectively. 41% of the target transcripts were categorized under the GO and pathway databases. Moreover, the protein domains, class, and TF groupings were defined for 22%, 13%, and 5% of all the targets, respectively (Tables S3a and S3b). Of the various TF families targeted by the conserved miRNAs, a majority were involved in plant development (Table S3a). The known miRNAs targeted TFs with similar functions, as demonstrated in other species. For example, we found NAC TFs as the targets of miR164, Squamosa-binding protein (SBP) and homeobox for miR165, NFY for miR169, and Scarecrow-like proteins (GRAS) for miR171. Besides, novel miRNAs were also found to target some crucial developmental factors. 153 ABC transporters were targeted by 5 conserved miRNAs (miR396a-5p, miR156, miR535d, miR172a, and miR3630-3p) as well as by 70 novel miRNA sequences. Similarly, 3 conserved miRNA families (miR159, miR172, and miR4995) and 55 novel miRNAs targeted 81 heat shock factors/proteins. Target discovery followed by their

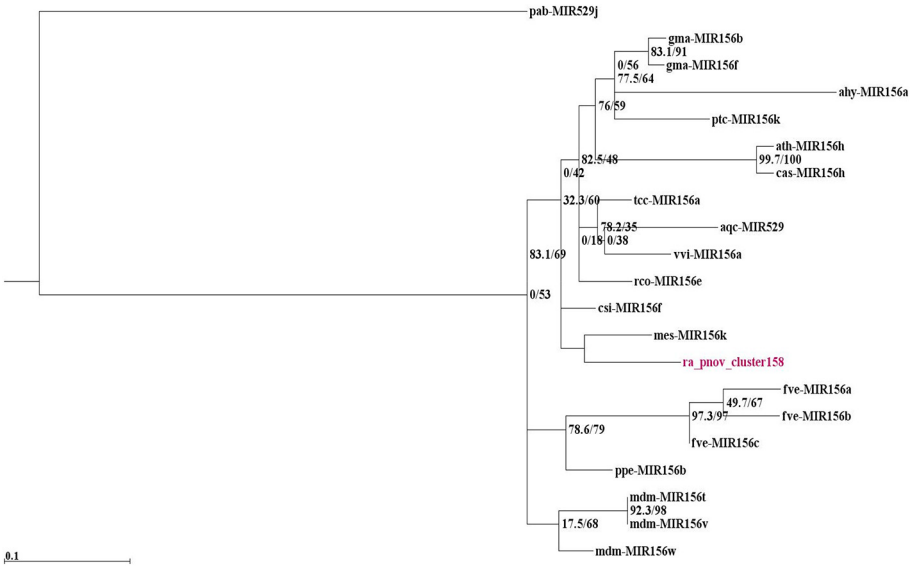


Fig. 4. A phylogenetic tree based on the homologs of miR156 precursors in different plants, whose sequences were retrieved from the miRBase; red color depicts the subclade containing the *R. arboreum* putative miR156 homolog (ra_pnov_cluster158). (Scale represents branch length and internal node data denotes bootstrap values based on SH-alrt and ultrafast bootstrap). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

annotations resolved that the novel miRNAs were functionally more diverse than those of conserved miRNAs. Overall, the existing sRNA profiling elucidated a robust target prediction in silico. However, the actual miRNA-target interactions will need further experimental validations.

Upon comparing the expressions of conserved miRNAs across the two pools, 35% (85 of 240) had variable expression in response to cold (absolute value of $\log_2$ fold-change ≥ 1). ~24% of miRNAs identified were found exclusively expressed in either of the seasons. miR828, miR1507, miR2275, miR4376, miR6263, miR8781 were exclusive to pool I. While miR1023, miR162_1, miR162_2, miR169_2, miR2673, miR947 were exclusive to pool II. It should be further pointed out that different isoforms of a single miRNA can target distinct genes at different time periods to exhibit specific regulations. Our study also showed that the expression of members of 11 conserved miRNA families was variable across the two pools (Table S4a). For instance, miR172a and miR172c-5p were expressed neutrally; miR172c and miR172d were expressed only in pool I; and miR172b, miR172d-3p, and miR172f were exclusive to pool II. The same held true for the members of miR159, miR164, miR166, miR167, miR171, miR396, miR399, miR403, miR482, and miR5067 families. Similarly, 57% of the novel RNAs were differentially regulated and 9% were exclusively expressed in either of the two pools (Fig. S1a). Overall, the diversity in miRNA expression patterns proves their significance during different growing periods.

Circadian clock maintains a tempo of developmental responses as per the environmental conditions [21,22]. Five of the “core” and five photoreceptor genes were located as targets of diverse miRNAs. These genes communicate with the flowering pathways to regulate a set of floral pathway integrators and, in turn, activate the genes needed for reproductive development. These transcripts were targeted by 77 novel and 30 conserved miRNAs. The expression levels of miRNAs and of most of their targets expressed during two seasons submitted a coordination between genes and their regulatory system. The reported target genes required for day-length measurements and their miRNA-directed regulation during flowering agreed with the literature (Fig. 2; Table S1b) [23,24].

Another environment-specific flowering strategy is linked to temperature. A shift in annual temperature patterns affects the expression of functional genes, including the regulation via miRNA silencing. In this relation, a downregulation of miR211 that targeted a transcript encoding for VIN3-like protein (VIL1) was seen (Fig. 2). Such genes tend to express only after an extended cold period [31] and their way of regulation advocate an epigenetic basis for vernalization in *R. arboreum*. A networking among the other targets of the flowering regulation pathway has been created using the String database (Fig. S1d). The results clearly exhibited the complexity of regulation between miRNA (both novel and conserved) and their targets.

The miR156-miR172 duo is the master mediator of vegetative-to-reproductive transition. The expression of miR156 is generally on a higher side in developing seedlings than that of miR172. We observed an inverse relationship between the two miRNAs during flowering that agreed with the development in Arabidopsis [25,26]. miR156, along with eight novel miRNAs, was found suppressing its default target i.e., SBP-like (SPL) transcripts [27], during vegetative growth. Extending the clock's output process, miR172 has been demonstrated as the floral timing regulator that causes translational repression of APEATLA2 (AP2) gene subfamily [28]. We also located AP2 as the target of a miR172 homolog and of two novel miRNAs, which were upregulated in the flowering season. Further, an inverse relation obtained via q-RT PCR between the selected target and miRNA expression supports that the selected miRNAs are regulating their candidate targets (Fig. 3).

The annotation and phylogeny of the miR156 family in *R. arboreum* and its homologs identified in other plant species were studied to infer evolutionary relationships. The analysis further ascertained the function of putative miR156 sequence in the species (Fig. 4). The putative miR156 precursor of *R. arboreum*, ra_pnov_cluster158, grouped with mes-miR156k of *M. esculenta* as well as with its counterparts from *Citrus*, *Ricinus*, *Vitis,* and *Theobroma.* Secondly, this putative sequence was observed to target SPL in the present study and moreover, its expression was upregulated during the reproductive season. These results were consistent with the reports on the accumulation pattern of miR156 and its target in other plants [25–27]. It further indicated that the function of this miRNA family in the regulation of developmental timing is evolutionarily conserved in flowering plants [29]. The annotation for ra_pnov_cluster158 targets further revealed E3 ubiquitin ligase, kinase, and cytochrome P450 genes similar to earlier studies [29,30]. An additional targeting of B3 and C3H domain-containing protein, FAR1-related, Homeobox, hydrolase, methyltransferase, phosphatase, and oxidase indicated an expansion of the functions of the miR156 family. The phylogenetic tree and differential expression of the miR156 family in *R. arboreum* have strengthened the results of similarity searches as well as provided a probable function to the putative miR156 sequence in the species.

5. Conclusion

Silencing the gene expression at transcriptional or post-transcriptional levels is typical of miRNA. The regulators can sense the endogenous cues; modulate their interaction with the crucial genes to govern the timing and progression of an organism through a specific developmental phase. The present sequencing data has led to the generation of miRNA expression profiles in *R. arboreum* and revealed the complexity of sRNAs in controlling diverse cellular processes during the two growing periods. Furthermore, the functional annotations for the predicted targets have implicated the putative roles of both abundant novel and conserved miRNAs in managing the circadian clock as well as the flowering/vegetative switch. Collectively, the study extends the share of miRNAs to organize plant behavior and the associated physiological consequences in higher organisms. The study will offer new perspectives into the regulation of transcription during development and seasonal changes in tree species.

Acknowledgments

The work is supported by the University Grants Commission, New Delhi, India (BSR-UGC 30-13/2014) and RSM (CUPB/CC/14/OO/4507). The study presented here is a part of the project aiming for diversity and genetic structure analysis in *R. arboreum*. ST and SC want to thank the Indian Council for Medical Research for the fellowships towards Ph.D.

The authors sincerely thank anonymous reviewers for their valuable comments that contributed significantly in further improvement of this manuscript.

Conflict of interest

The authors declare that they have no competing interests.

Authors' contribution

PB conceived the study and designed and organized all the experiments. SC and ST collected the samples for RNA isolation and collated the sequence data. AM assisted with the phylogenetic analysis and wet-lab experiments. SC carried out the bioinformatics analysis, compiled and analyzed the results, performed PCR characterization, and wrote the manuscript. PB further improved the work with his valuable suggestions. All the authors have read and approved the final manuscript.

Data archiving statement

The raw small RNA and transcriptome sequence data are available

from the NCBI under the study accession No.: SRP148106 and SRP092027, respectively.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ygeno.2018.09.019.

References

[1] G. Szittya, D. Silhavy, A. Molnár, Z. Havelda, Á. Lovas, L. Lakatos, Z. Bánfalvi, J. Burgyán, Low temperature inhibits RNA silencing-mediated defence by the control of siRNA generation, EMBO J. 22 (2003) 633–640.

[2] D. Jin, Y. Wang, Y. Zhao, M. Chen, MicroRNAs and their cross-talks in plant development, J. Genet. Genomics 40 (2013) 161–170.

[3] L. He, G.J. Hannon, MicroRNAs: small RNAs with a big role in gene regulation, Nat. Rev. Genet. 5 (2004) 522–531.

[4] R.M. Green, S. Tingay, Z.-Y. Wang, E.M. Tobin, Circadian rhythms confer a higher level of fitness to Arabidopsis plants, Plant Physiol. 129 (2002) 576–584.

[5] Y. Xie, Y. Liu, H. Wang, X. Ma, B. Wang, G. Wu, H. Wang, Phytochrome-interacting factors directly suppress MIR156 expression to enhance shade-avoidance syndrome in Arabidopsis, Nat. Commun. 8 (2017) 348.

[6] N. Mehta, H.-Y.M. Cheng, Micro-managing the circadian clock: the role of microRNAs in biological timekeeping, J. Mol. Biol. 425 (2013) 3609–3624.

[7] N. Ohama, H. Sato, K. Shinozaki, K. Yamaguchi-Shinozaki, Transcriptional regulatory network of plant heat stress response, Trends Plant Sci. 22 (2017) 53–65.

[8] R.W. Brooker, Plant-plant interactions and environmental change, New Phytol. 171 (2006) 271–284.

[9] K.S. Gaira, R.S. Rawal, B. Rawat, I.D. Bhatt, Impact of climate change on the flowering of *Rhododendron arboreum* in central Himalaya India, Curr. Sci. 106 (2014) 1735–1738.

[10] M.B. Stocks, S. Moxon, D. Mapleson, H.C. Woolfenden, I. Mohorianu, L. Folkes, F. Schwach, T. Dalmay, V. Moulton, UEA sRNA workbench: a suite of tools for analysing and visualizing next generation sequencing microRNA and small RNA datasets, Bioinformatics 28 (2012) 2059–2061.

[11] S. Choudhary, S. Thakur, R.A. Najar, A. Majeed, A. Singh, P. Bhardwaj, Transcriptome characterization and screening of molecular markers in ecologically important Himalayan species (*Rhododendron arboreum*), Genome 61 (2018) 417–428.

[12] S. Griffiths-Jones, R.J. Grocock, S. van Dongen, A. Bateman, A.J. Enright, miRBase: microRNA sequences targets and gene nomenclature, Nucleic Acids Res. 34 (2006) D140–D144.

[13] X. Dai, P.X. Zhao, psRNATarget: a plant small RNA target analysis server, Nucleic Acids Res. 39 (2011) W155–W159.

[14] D.T. Hoang, O. Chernomor, A. von Haeseler, B.Q. Minh, S.V. Le, UFBoot2: improving the ultrafast bootstrap approximation, Mol. Biol. Evol. 35 (2017) 518–522.

[15] S. Guindon, J.-F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, O. Gascuel, New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of phyML 3.0, Syst. Biol. 59 (2010) 307–321.

[16] M.V. Han, C.M. Zmasek, phyloXML: XML for evolutionary biology and comparative genomics, BMC Bioinformatics 10 (2009) 356–361.

[17] S. Paul, A. Kundu, A. Pal, Identification and expression profiling of *Vigna mungo* microRNAs from leaf small RNA transcriptome by deep sequencing, J. Integr. Plant Biol. 56 (2014) 15–23.

[18] S.B. Ghag, U.K. Shekhawat, T.R. Ganapathi, Small RNA profiling of two important cultivars of banana and overexpression of miRNA156 in transgenic banana plants, PLoS One 10 (2015) e0127179.

[19] R. Rajagopalan, H. Vaucheret, J. Trejo, D.P. Bartel, A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*, Genes Dev. 20 (2006) 3407–3425.

[20] K.D. Kasschau, N. Fahlgren, E.J. Chapman, C.M. Sullivan, J.S. Cumbie, S.A. Givan, J.C. Carrington, Genome-wide profiling and analysis of Arabidopsis siRNAs, PLoS Biol. 5 (2007) e57.

[21] R. Henriques, H. Wang, J. Liu, M. Boix, L.F. Huang, N.H. Chua, Antiphasic regulatory module comprising CDF5 and its antisense RNA FLORE links the circadian clock to photoperiodic flowering, New Phytol. 216 (2017) 854–867.

[22] Y. Joo, V. Fragoso, F. Yon, I.T. Baldwin, S.G. Kim, Circadian clock component LHY tells a plant when to respond photosynthetically to light in nature, J. Integr. Plant Biol. 59 (2017) 572–587.

[23] Y.H. Song, S. Ito, T. Imaizumi, Flowering time regulation: photoperiod-and temperature-sensing in leaves, Trends Plant Sci. 18 (2013) 575–583.

[24] S. Garcia-Mora, M.J. de Leone, M. Yanovsky, Time to grow: circadian regulation of growth and metabolism in photosynthetic organisms, Curr. Opin. Plant Biol. 35 (2017) 84–90.

[25] G. Wu, M.Y. Park, S.R. Conway, J.-W. Wang, D. Weigel, R.S. Poethig, Sequential action of miR156 and miR172 regulates developmental timing in Arabidopsis, Cell 138 (2009) 750–759.

[26] Q.-H. Zhu, C.A. Helliwell, Regulation of flowering time and floral patterning by miR172, J. Exp. Bot. 62 (2010) 487–495.

[27] J.-W. Wang, B. Czech, D. Weigel, miR156-regulated SPL transcription factors define an endogenous flowering pathway in *Arabidopsis thaliana*, Cell 138 (2009) 738–749.

[28] M.J. Aukerman, H. Sakai, Regulation of flowering time and floral organ identity by a microRNA and its APETALA2-like target genes, Plant Cell 15 (2003) 2730–2741.

[29] B. Khatabi, S. Arikit, R. Xia, S. Winter, D. Oumar, K. Mongomake, B.C. Meyers, V.N. Fondong, High-resolution identification and abundance profiling of cassava (*Manihot esculenta* Crantz.) microRNAs, BMC Genomics 17 (2016) 85–103.

[30] J.-W. Wang, R. Schwab, B. Czech, E. Mica, D. Weigel, Dual effects of mir156-targeted SPL genes and cyp78a5/kluh on plastochron length and organ size in *Arabidopsis thaliana*, Plant Cell 20 (2008) 1231–1243.

[31] S. Sung, R.M. Amasino, Vernalization in Arabidopsis thaliana is mediated by the PHD finger protein VIN3, Nature 427 (2004) 159.