



OPEN

Small RNA sequencing provides insights into molecular mechanism of flower development in *Rhododendron pulchrum* Sweet

Bo Fang^{1,2}, Zhiwei Huang^{1,2}, Yirong Sun¹, Wanjing Zhang¹, Jiaojun Yu¹, Jialiang Zhang¹, Hongjin Dong¹ & Shuzhen Wang¹✉

Rhododendron pulchrum sweet, a member of the Ericaceae family possessing valuable horticultural properties, is widely distributed in the temperate regions. Though serving as bioindicator of metal pollution, the molecular mechanism regulating flowering in *R. pulchrum* is very limited. Illumina sequencing was performed to identify critical miRNAs in the synthesis of flavonoids at different developmental stages. Totally, 722 miRNAs belonging to 104 families were screened, and 84 novel mature miRNA sequences were predicted. The miR166, miR156, and miR167-1 families were dominant. In particular, 126 miRNAs were significantly differentially expressed among four different flowering stages. Totally, 593 genes were differentially regulated by miRNAs during the flower development process, which were mostly involved in “metabolic pathways”, “plant hormone signal transduction”, and “mitosis and regulation of biosynthetic processes”. In pigment biosynthesis and signal transduction processes, gra-miR750 significantly regulated the expression of flavonoid 3',5'-hydroxylase; aof-miR171a, aof-miR171b, aof-miR171c, cas-miR171a-3p, and cas-miR171c-3p could regulate the expression of DELLA protein; aof-miR390, aof-miR396b, ath-miR3932b-5p, cas-miR171a-3p, aof-miR171a, and aof-miR171b regulated BAK1 expression. This research showed great potentials for genetic improvement of flower color traits for *R. pulchrum* and other *Rhododendron* species.

Rhododendron genus contains approximately 1000 evergreen and deciduous species, as well as thousands of commercial hybrids, which is widely distributed in Europe, Asia, and North America^{1,2}. As a typical large vascular plant genus within the Ericaceae family, *Rhododendron* possess valuable medicinal and horticultural properties^{3,4}. Besides bioindicator of metal pollution (Pb, Zn, and Cd), the evergreen *R. pulchrum* is also famous due to the beautiful vegetative forms, bright-colored flowers, and long flowering period ranging from late March to early May⁵. Therefore, *R. pulchrum* is widely served as roadside trees in urban and rural areas⁵.

During the past three decades, genetic analyses and transcriptomics studies have identified several hundred genes involved in flower development, and the majority encode transcription factors and transcriptional co-regulators, proteins involved in epigenetic control of gene expression, and microRNAs (miRNAs)^{6,7}. In particular, transcription factor-encoding genes AINTEGUMENTA (ANT), auxin efflux carrier Pin-Formed 1 (PIN1), Aintegumenta-like 6 (AIL6), regulatory gene LEAFY (LFY), MADS-domain transcription factors Agamous-like 24 (AGL24), Short Vegetative Phase (SVP), and shoot identity gene Terminal Flower1 (TFL1) were all deeply studied^{8,9}. Based on major advances in technology, genome-wide studies of transcription factor-binding sites, proteomic analyses, and imaging techniques have all been adopted to clarify molecular mechanisms underlying flowering¹⁰.

MiRNAs (approximately 22 nt), a class of non-coding single strand molecules negatively regulating gene expressions both at transcriptional and posttranscriptional levels, play critical roles in biological processes, including embryogenesis and organ development^{11–14}, and metabolism^{15–18}. In particular, functional mature miRNAs could be incorporated into Argonaute 1 (AGO1)-containing RNA-induced silencing complex (RISC), and causing degradation or translational repression in sequence-specific manner^{19,20}. The miRNAs often regulate

¹College of Biology and Agricultural Resources, Huanggang Normal University, Huanggang 438000, Hubei, People's Republic of China. ²These authors contributed equally: Bo Fang and Zhiwei Huang. ✉email: wangshuzhen710@whu.edu.cn

corresponding mRNA targets through guiding cleavages between 10 and 11th nucleotides in complementary region²¹.

Though molecular methods have been utilized to study the growth, development, and metabolism regulation, the basic molecular research of *R. pulchrum* is still limited, especially for the molecular mechanisms underlying anthocyanin synthesis. In order to explore the potential roles of miRNAs during flowering process, miRNA expression profiles of *R. pulchrum* flowers at four different developmental stages were investigated with high throughput small RNA sequencing technology. Differentially expressed miRNAs among different flower developmental stages and corresponding targets were also clarified. This research will be benefit for genetic improvement of *R. pulchrum* and other *Rhododendron* species.

Results

Sequence analysis of small RNA in *R. pulchrum* flowers

R. pulchrum flowers at four different developmental stages were separately collected, including stage I (floral bud stage, late-March, dormant flower buds, floral organs had been well formed containing style, petal primordium, stamen primordia, sepal primordia, pistil primordia, and ovary), stage II (early flowering stage, early-April, 1–2 days before full bloom), stage III (full-flowering stage, middle-April, completely open petals, observable pistils and stamens), and stage IV (flower withering stage, late-April, petals began to fall) (Fig. 1). Twelve small RNA libraries (three libraries for each developmental stage) were constructed and sequenced. Totally, 28,256,644–33,991,292, 32,259,303–39,111,962, 29,838,331–36,968,844, and 29,218,896–41,838,094 raw reads were generated; while 25,977,837–31,816,069, 28,096,248–32,886,303, 22,332,972–27,916,820, and 23,035,469–33,163,852 clean reads (≥ 18 nt) were obtained after removing low quality reads and adaptors for stage I, stage II, stage III, and stage IV respectively (Table 1). The Q20 values were all above 99.68%. GC contents varied from 51.52% (stage I) to 55.04% (stage III). In particular, similar length distributions of small RNA were observed for four set samples of different developmental stages. Among 18–40 nt small RNAs, the majority ranged from 21 to 25 nt in length, and the percentages ranged within 48.24–84.15%. In particular, the most abundant sRNAs were 24nt, and the percentages of 24 nt sRNAs were 59.46%, 37.06%, 20.03%, and 19.12% for stage I, stage II, stage III, and stage IV respectively. However, the second abundant sRNAs were 25 nt in stage I samples, but 21 nt for the rest samples. The third abundant sRNAs were 22 nt in samples at stage II, stage III, and stage IV; while sRNAs with length of 21 nt was the third abundant type for stage I samples (Fig. S1).

Unique reads were obtained through removing redundant sequence. In total, 75.34%, 78.24%, 90.39%, and 87.04% clean reads could be mapped to the reference genome (*R. simsii* genome) for stage I, stage II, stage III, and stage IV, respectively. Particularly, the most abundant sRNA were mapped to chromosome 5 (9.54–19.2%), chromosome 2 (5.56%–6.71%), and chromosome 6 (4.44–5.12%) of *R. simsii* genome (Fig. 2). After annotation, 19,940,384–24,302,325, 22,623,935–26,572,199, 21,446,057–26,735,509, and 21,089,468–30,160,793 small RNA

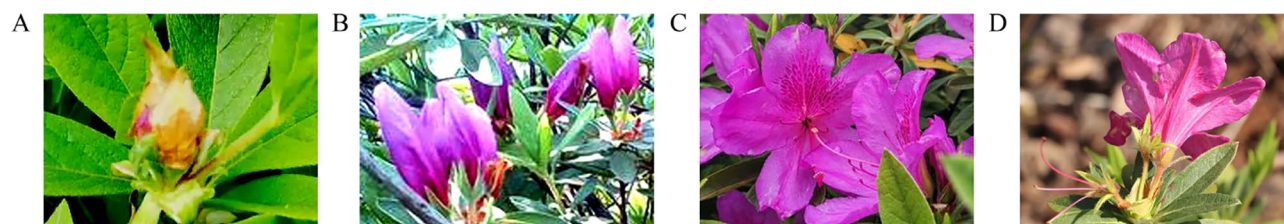


Figure 1. Flower tissues of four different stages: (A–D) representing stage I (floral bud stage), stage II (early flowering stage), stage III (full-flowering stage), and stage IV (flower withering stage), respectively.

RNA type	Stage I	Stage II	Stage III	Stage IV
Raw reads	28,256,644–33,991,292	32,259,303–39,111,962	29,838,331–36,968,844	29,218,896–41,838,094
Clean reads	25,977,837–31,816,069	28,096,248–32,886,303	22,332,972–27,916,820	23,035,469–33,163,852
Q20	99.69–99.8%	99.72–99.79%	99.68–99.71%	99.7–99.71%
GC%	51.52–51.89%	52.71–52.75%	54.88–55.04%	54.06–54.15%
Mapped percentage	75.34%	78.24%	90.39%	87.04%
Annotation reads	19,940,384–24,302,325	22,623,935–26,572,199	21,446,057–26,735,509	21,089,468–30,160,793
Annotation percentage	75.3–75.38%	78.2–78.29%	90.25–90.49%	87.04–87.19%
rRNA	209,874–239,374	328,326–355,140	275,098–335,100	370,317–459,196
tRNA	27,892–31,441	33,640–36,550	20,185–23,230	34,532–42,451
snRNA	12,830–14,212	13,143–14,526	7218–8099	11,913–13,908
snoRNA	14,290–15,991	12,979–14,493	7499–8312	10,740–12,315
Known_miRNA	1500–1622	1678–1843	1231–1435	1.454–1533

Table 1. Small RNA categorization in *R. pulchrum* flowers at different developmental stages.

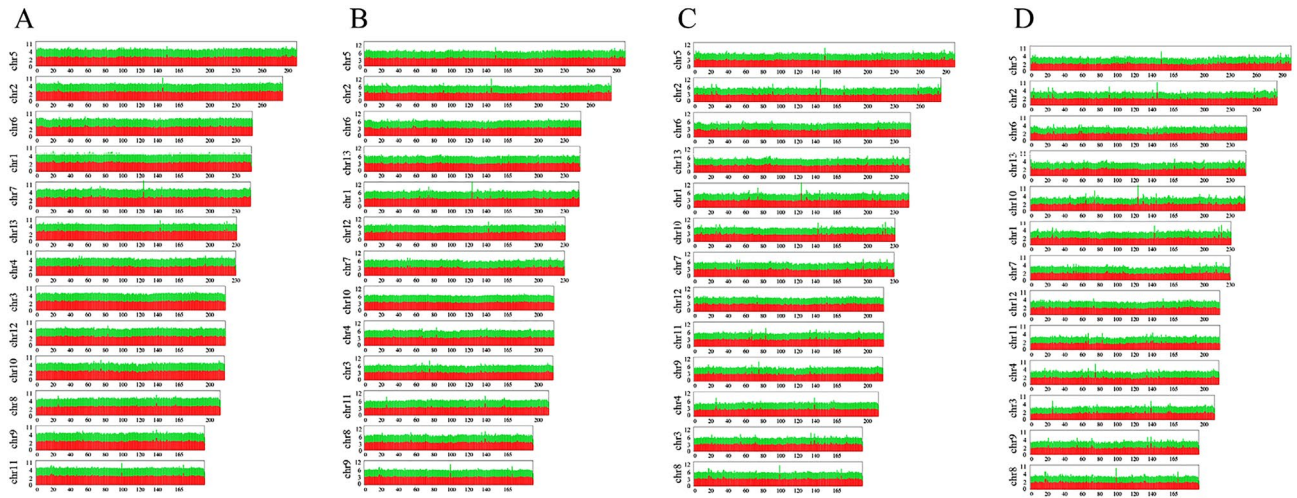


Figure 2. The genome distribution of sRNAs in four flower samples at stage I (A), stage II (B), stage III (C), and stage IV (D), respectively. The red parts in Y-axis represents log₁₀ count, and the green part represents log₁₀ category. “Chr” was short for chromosome.

tags were obtained for stage I-stage IV samples of *R. pulchrum* flowers, which could be further classified into miRNAs, tRNAs, rRNAs, snRNA, and snoRNA.

Clarification of miRNAs based on deep sequencing

Totally, 104 miRNA families were obtained, among which miR166 family was the largest represented family with 24 members (Tables S1, S2). Moreover, miR156 (17 members), miR167-1 (17 members), miR159 (16 members), and miR171_1 (16 members) were also the dominant families. Based on alignment results, homologous sequences of these identified miRNA were mainly in *Arabidopsis lyrata*, *Malus domestica*, *Populus trichocarpa*, *Glycine max*, and *Oryza sativa*. The first position of known miRNAs had obvious uracil (U) base preference (Fig. S2A). Besides bases at the first position, the 3'-end bases also possessed U base preference. Moreover, the 8th and 9th bases had slight G base preference, while C base preference existed at 23th position (Fig. S2B).

In related to the 722 known miRNAs, 437, 348, 454, 367 miRNAs were present in stage I, stage II, stage III, and stage IV samples, respectively. Totally, 211 miRNAs regulated expression of corresponding genes across all four developmental stages of *R. pulchrum* flowers (Fig. 3 and Table S3). Particularly, 90, 70, 98, and 70 miRNA were unique to stage I, stage II, stage III, and stage IV, respectively. Furthermore, 5, 11, 8, 15, 64, and 12 miRNAs were expressed in “stage I” vs “stage II”, “stage II” vs “stage III”, “stage III” vs “stage IV”, “stage I” vs “stage IV”, “stage I” vs “stage III”, and “stage II” vs “stage IV”, respectively (Fig. 3 and Table S3). Moreover, there were 68 miRNAs were commonly expressed in samples at three developmental stages.

The miRNA sequences, which met the threshold of miRDeep2 analysis but possessing no known homologous miRNA gene families in miRBase, were used to predict novel miRNAs (total score > 0). A total of 84 novel mature miRNA sequences were identified in *R. pulchrum* flowers (Table S4). Totally, four length types of miRNAs were

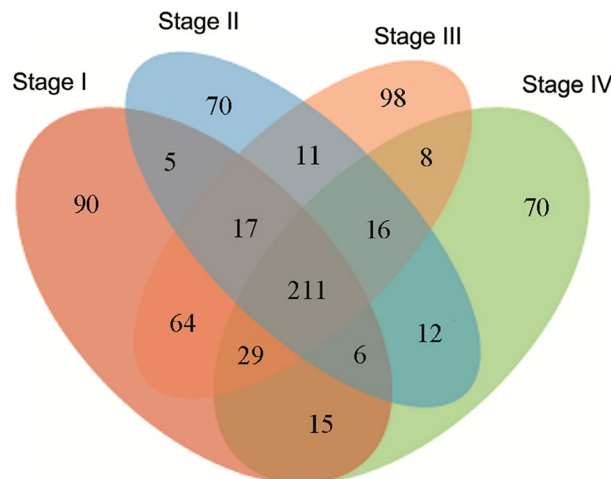


Figure 3. Venn diagram of identified miRNA in four stages.

predicted, including 21 nt (55, 65.476%), 22nt (24, 28.571%), 20 nt (4, 4.762%), and 24 nt (1, 1.190%) (Fig. S3A). The first nucleotide showed G and U bias for 20 nt and 21 nt miRNAs, respectively. For these 24 nt miRNAs, the first nucleotide was all U base. Among these 84 novel miRNAs, the 1th, 2th, 7th, and 12th position all had U base preference (Fig. S3B). Bases at 8th and 23th position had obvious C base preference. Moreover, the 24th bases in novel miRNAs were all G base.

Identification of differentially expressed miRNAs

Based on fold change (≥ 1 or ≤ -1) and P-value (< 0.05) criteria, a total of 126 miRNAs were significantly differentially expressed. Compared with samples at stage I, 31 miRNAs were up-regulated and 34 miRNAs were down-regulated in samples at stage II (Fig. 4A and Table S5). Compared with samples at stage II, 34 miRNAs were up-regulated and 37 miRNAs were down-regulated in samples at stage III (Fig. 4B and Table S5). Compared with samples at stage III, numbers of the up-regulated and down-regulated miRNAs in samples at stage IV were 33 and 30, respectively (Fig. 4C and Table S5).

Statistical analysis was carried out through calculating Pearson's correlation coefficient (r) value based on DEGs, and these biological replicates in each group showed high correlation (Fig. 5). Furthermore, miRNA profiles of flower tissues at stage II and stage IV were clustered first, and then grouped with that of flowers sampled at stage I. However, the miRNA profiles of flowers at full-flowering stage (Stage III) showed less association with the other three flower samples (Fig. 5A). Furthermore, gene cluster showed that these DEGs could be mainly clustered into two clades: clades 1 containing miRNA profiles of "Stage I" and "Stage II"; clade 2 consisting of miRNA profiles of "Stage III" and "Stage IV" (Fig. 5B).

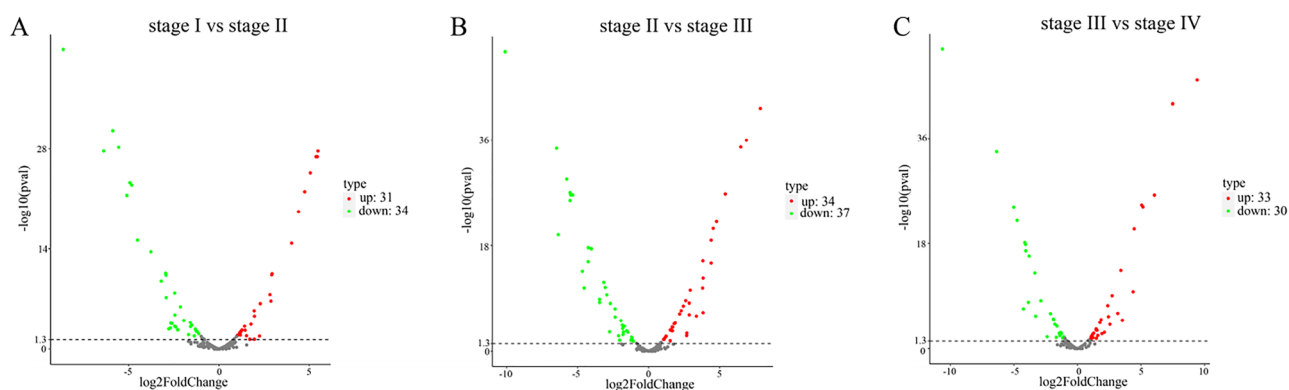


Figure 4. Volcano plots of miRNA data between stage I and stage II (A), stage II and stage III (B), as well as stage III and stage IV (C).

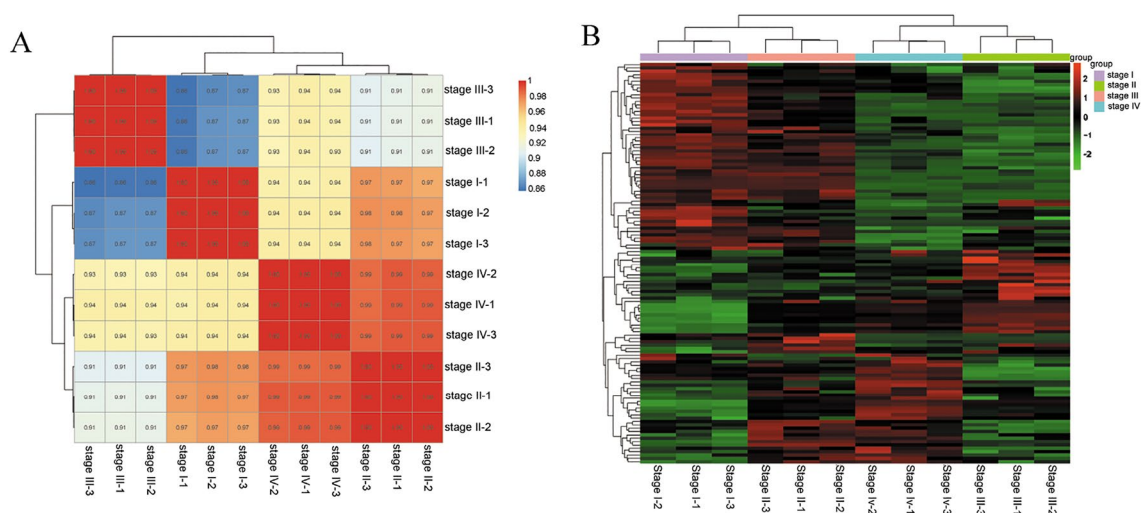


Figure 5. Distance analysis of *R. pulchrum* flower tissues collected from four developmental stages by calculating Pearson's correlation coefficient: (A) sample correlation matrix ($p < 0.001$, $FDR < 0.001$); (B) gene cluster analysis ($p < 0.001$, $FDR < 0.001$).

Target prediction and GO enrichment analysis

With transcriptome data of *R. pulchrum* flower obtained in our previous study serving as target database, 1170 miRNA target site (593 target genes) were obtained. Among these miRNA-target pairs, 84.57% of miRNAs potentially targeted multiple unigenes, ranging from 2 to 157. During *R. pulchrum* flower development from stage I to stage II, a total of 1081 potential miRNA targets were categorized into molecular functions (227), cellular components (109), and biological processes (745) according to GO-based enrichment analysis (Table S6). “organic substance catabolic process” (GO:1901575), “cellular catabolic process” (GO:0044248), and “extracellular region” (GO:0005576) were the main enriched terms (Fig. 6A). Among GO terms associated with biologic process, the target genes were mainly involved in “monosaccharide transporter” and “xenobiotic detoxification” pathways (Fig. S4A). In the cellular component category, target genes were tightly associated with “nucleus and extracellular space” (Fig. S4B). In molecular function, most target genes were related to “isomerase activity”, “electron transfer activity”, and “acid phosphatase activity” (Fig. S4C).

From stage II to stage III, a total of 1328 potential miRNA targets were categorized into molecular functions (258), cellular components (170), and biological processes (900) (Table S6). In particular, “drug transporter activity” (GO:0090484), “drug transmembrane transporter activity” (GO:0015238), and “signaling receptor activity” (GO:0038023) were the most enriched terms (Fig. 6B). Among GO terms related to biologic process, the majority target genes were involved in “histone lysine methylation”, “mRNA splicing”, and “transcription initiation” (Fig. S5A). In cellular component category, target genes were mainly associated with “small-subunit processome assembly”, “1,3- β -D-glucan synthetase complex”, and “nucleolus” (Fig. S5B). For molecular function category, target genes were mainly related to “ADP binding”, “NAD⁺ binding”, and “clathrin binding” (Fig. S5C).

From stage III to stage IV, a total of 951 potential miRNA targets were categorized into molecular functions (169), cellular components (87), and biological processes (695) (Table S6). Particularly, “cellular aromatic compound metabolic process” (GO:0006725), “cellular catabolic process” (GO:0044248), and “organic substance catabolic process” (GO:1901575) were the abundant enriched terms (Fig. 6C). For GO terms associated with biologic process, the target genes were mainly involved in “histone lysine methylation”, “transcription initiation”, and “mRNA splicing” (Fig. S6A). Among cellular component category, target genes were tightly associated with “small subunit processome”, “nucleus”, and “apoplast” (Fig. S6B). In molecular function category, most target genes were related to “NAD⁺ binding”, “copper ion binding”, and “4 iron,4 sulfur cluster binding” (Fig. S6C).

KEGG pathway analysis of potential targets of known miRNAs

During *R. pulchrum* flower development process from stage I to stage II, 52 pathways have been identified, mainly including “Plant hormone signal transduction” (ko04075), “Ubiquitin mediated proteolysis” (ko04120), “mRNA surveillance pathway” (ko03015), and “Amino sugar and nucleotide sugar metabolism” (ko00520) (Fig. 7A and Table S7). For the “Flavone and flavonol biosynthesis” pathway (ko00944), *gra*-miR8750 significantly regulated the expression of flavonoid 3',5'-hydroxylase, which was involved in the biosynthesis of secondary metabolites including 3'-O-Methylfluteolin, quercetin, and myricetin (Fig. 8).

During *R. pulchrum* flower development process from stage II to stage III, 65 pathways have been identified, and the most abundant were “Amino sugar and nucleotide sugar metabolism” (ko00520), “Galactose metabolism” (ko00052), “Glyoxylate and dicarboxylate metabolism” (ko00630), and “Plant hormone signal transduction” pathway (ko04075) (Fig. 7B and Table S7). In the “Plant hormone signal transduction” pathway, a list of miRNAs showed great potentials for the signal transduction. During the environmental information processing, *aof*-miR171a, *aof*-miR171b, *aof*-miR171c, *cas*-miR171a-3p, and *cas*-miR171c-3p could regulate the expression of DELLA protein, which further affecting the diterpenoid biosynthesis (Fig. 9). Moreover, *aof*-miR390, *aof*-miR396b, *ath*-miR3932b-5p, *cas*-miR171a-3p, *aof*-miR171a, and *aof*-miR171b regulated brassinosteroid biosynthesis through changing the expression level of BAK1 protein; while *cas*-miR171c-3p and *aof*-miR390 could alter the expression of BR11 protein, a leucine-rich repeat receptor involved in the brassinosteroid (BR) perception (Fig. 9).

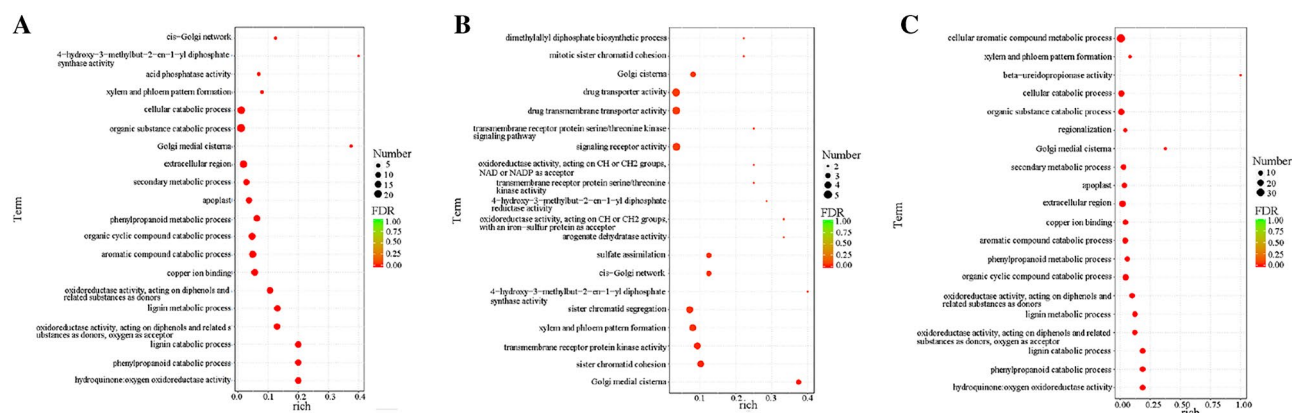


Figure 6. GO annotations of potential miRNA targets during *R. pulchrum* flower development from stage I to stage II (A), stage II to stage III (B), and stage III to stage IV (C).

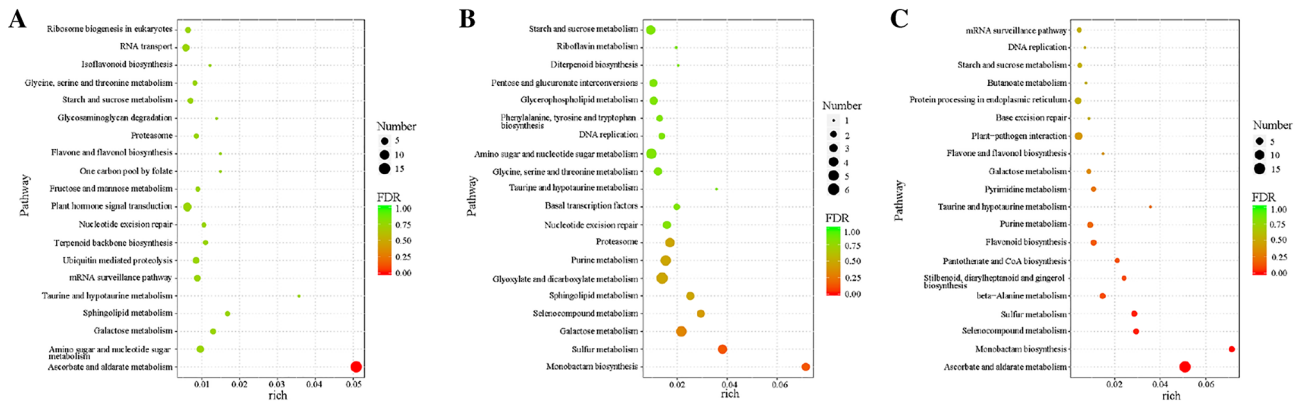


Figure 7. KEGG pathway analysis of potential miRNA targets during *R. pulchrum* flower development from stage I to stage II (A), stage II to stage III (B), and stage III to stage IV (C).

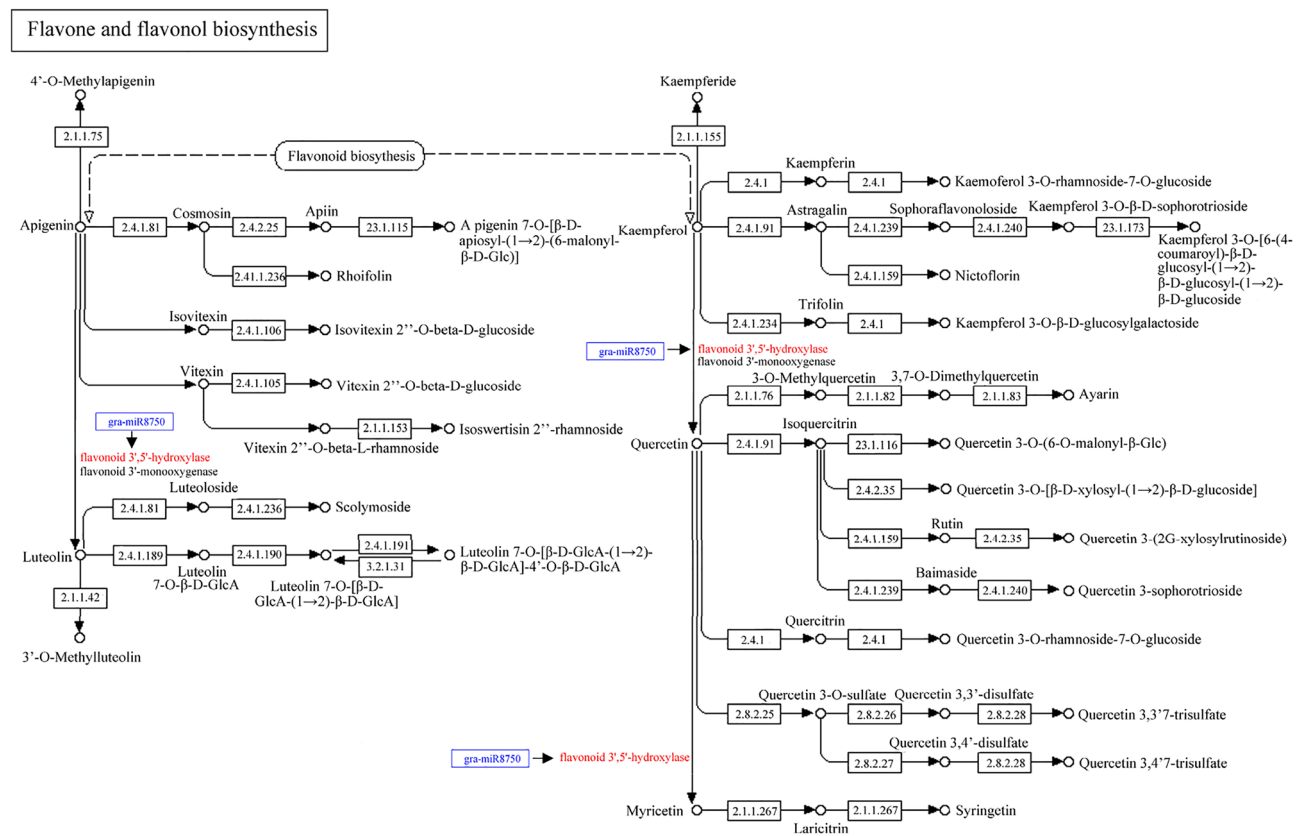


Figure 8. Metabolic regulation of miRNA in the process of flavone and flavonol biosynthesis.

During *R. pulchrum* flower development process from stage III to stage IV, 34 pathways were well screened out, mainly involving “Ascorbate and aldarate metabolism” (ko00053), “Plant-pathogen interaction” (ko04626), “Plant hormone signal transduction” (ko04075), and “Protein processing in endoplasmic reticulum” (ko04141) (Fig. 7C and Table S7). For the “Plant hormone signal transduction” (ko04075), auxin response factor (ARF), DELLA, and MYC2 were significantly regulated by miRNAs (Fig. 9). In particular, the ata-miR408-3p could affect the expression of ARF in tryptophan metabolism. Moreover, these miRNAs consisting of aof-miR171a, aof-miR171b, aof-miR171c, cas-miR171a-3p, and cas-miR171c-3p, all accounted for the diterpenoid biosynthesis through affecting the expression of DELLA protein. Furthermore, aof-miR160a, aof-miR160c, and csi-miR160c-5p regulated the expression of transcription factor MYC2, which was involved in α -Linolenic acid metabolism.

Validation of differentially expressed miRNAs by real-time qPCR assay

To verify the miRNA sequencing data, nine genes exerting diverse expression profiles at four developmental stages were randomly selected for qRT-PCR validation, including rsi-MIR 158-3, rsi-MIR 159-8, rsi-MIR 166-2,

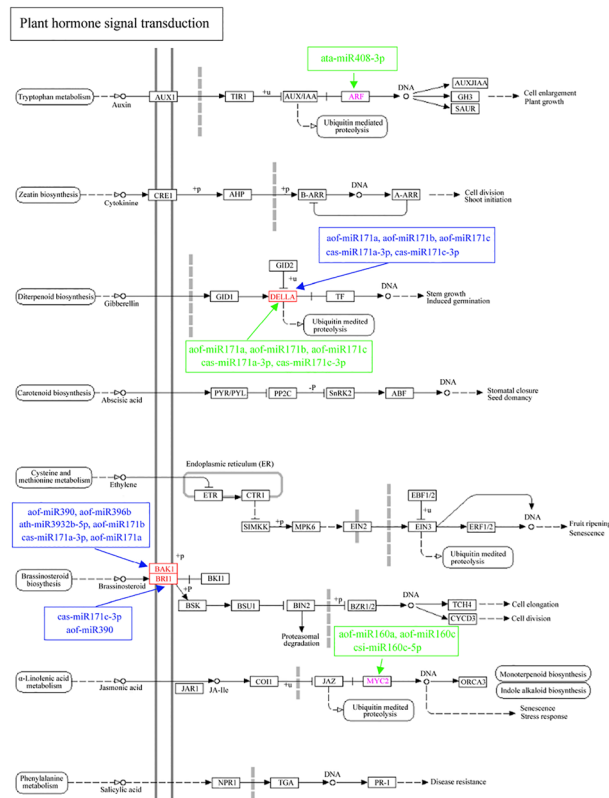


Figure 9. Metabolic regulation of miRNA in the process of plant hormone signal transduction.

rsi-MIR 167_1-3, rsi-MIR 171_1-11, rsi-MIR 396-1, rsi-MIR 398-6, rsi-MIR 535-2, and rsi-undef-1. The qPCR analysis showed that up-regulated and down-regulated results of miRNAs were similar with deep sequencing results (Fig. 10). Good correlation ($R^2 = 0.8919$) was obtained between sequencing data and qRT-PCR results, further confirmed the high reliability of miRNA sequencing data. Expression pattern of corresponding target genes was also consistent with prediction, which was verified by qRT-PCR analysis (data not shown). The amount of rsi-MIR 159-8, rsi-MIR 167_1-3, and rsi-undef-11 transcript decreased gradually during flowering process, but slightly increased at the withering stage. Moreover, the expression levels of rsi-MIR158-3, rsi-MIR166-2, rsi-MIR171_1-11, rsi-MIR396-1, and rsi-MIR398-6 were all highest at early flowering stage, while transcript of rsi-MIR535-2 was peaked at full-flowering stage.

Discussion

The miRNAs have been evidenced as key regulators in gene expression, developmental processes, and even stress tolerance²². Recently, the bioinformatics, as well as sequencing approaches and technologies have facilitated the rapid and accurate miRNA detection and analysis at single-base pair resolution²³. In related to the Ericaceae family with great ornamental properties, understanding flowering mechanism will benefit corresponding genetic improvement of flowers traits. However, there is currently limited miRNA information on regulating mechanism underlying flowering of *R. pulchrum*, which impedes further genetic improvement.

In this study, 722 conserved miRNAs representing 104 miRNA families associated with different flower developmental processes were identified in *R. pulchrum* at four different stages. According to Choudhary et al.²⁴, the non-coding miRNA might be crucial entity in remodeling genetic architecture of *R. pulchrum*, and the fluctuations in miRNA expression could be induced by developmental factors. The number of miRNAs regulating flower development of *R. pulchrum* (722 known miRNAs and 84 novel miRNAs) is higher than that of Korla fragrant pear at anthesis (73 miRNAs)²⁵, that of *Carya cathayensis* genome (51 conserved miRNAs belong to 16 families and 195 novel miRNAs)²⁶, as well as that of ornamental *Paeonia rockii* and *Paeonia ostii* petals (22 conserved miRNAs and 27 novel miRNAs)²⁷.

Totally, 126 miRNAs were significantly differently expressed, inferring that these miRNAs were inducible and responded to certain developmental stage. The variable miRNA expression among tissues of different developmental stages can be associated with the process of growth and development of *R. pulchrum*. Moreover, exogenous environmental cues also mediate phase transition through networking of flowering pathways as well as corresponding components. In particular, miR166, miR156, miR159, and miR171_1 families were dominant for *R. pulchrum* flower development, which was partly similar with that of maize²⁸, *Camellia oleifera*²⁹, and *Vigna mungo*³⁰. The miR166/165 group and corresponding target genes regulate shoot apical meristem and floral development through WUSCHEL-CLAVATA) pathway³¹. In *Paeonia rockii* petals, variegation formation was

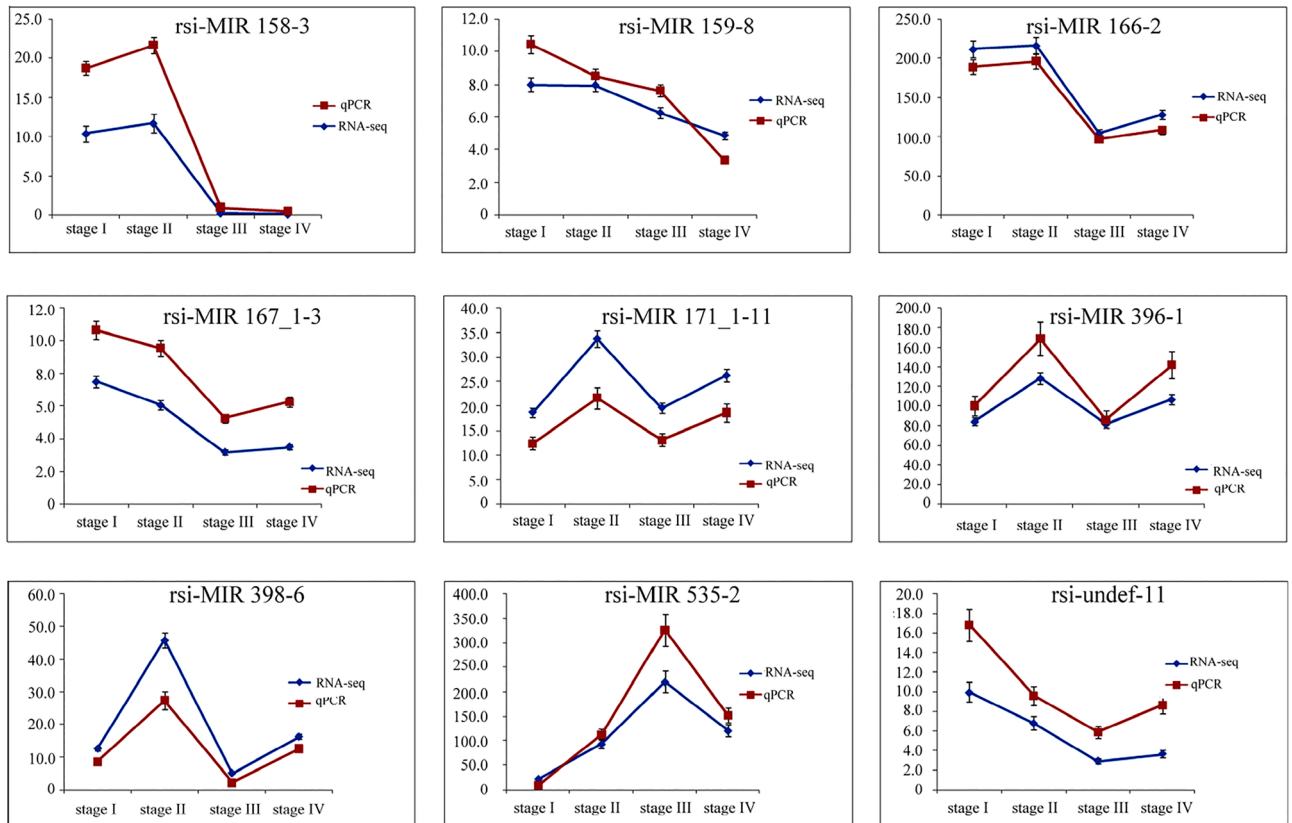


Figure 10. Correlation analyses of nine miRNAs between Illumina sequencing and qRT-PCR data.

regulated by miR168, miR396a, miR159c, and novel_miR_05, which might target MYB transcription factors, CHS, and ABC transporter²⁷.

During process of *R. pulchrum* flower development, 1170 miRNA target site (593 target genes) have been screened, much less than that of *R. arboretum* leaves (27,139 predicted targets)²⁴. Similarly, these predicted target genes were all involved in metabolism, reproduction, and response to abiotic stimuli. In particular, amino sugar, nucleotide sugar, galactose, glyoxylate, dicarboxylate, ascorbate, and aldarate were all involved in the metabolism processes. Moreover, a negative correlation was observed between the expression of miRNAs and corresponding targets. It further highlighted the critical role of miRNA-target pairs in perceiving environmental variability and monitoring flowering growth. The phylogenetic clustering further supported the lineage-specific evolution and function of putative miRNA sequence in *R. pulchrum*.

Flavonoids are ubiquitous plant secondary metabolites, and their hydroxylation pattern could greatly affect color, stability, and even antioxidant capacity. In particular, hydroxylation pattern of the B-ring of flavonoids is mainly determined by flavonoid 3'-hydroxylase (F3'H) and flavonoid 3',5'-hydroxylase (F3',5'H)³². As cytochrome P450-dependent monooxygenases, flavonoid 3',5'-hydroxylase could hydroxylate the B-ring of flavonoids at both 3'- and 5'-position, respectively³³. In particular, the F3'5'H activity facilitates the synthesis of delphinidin-based anthocyanins, providing the basis for lilac to blue flower colours. During *R. pulchrum* flower development, gra-miR8750 might exert some effects towards the expression of flavonoid 3',5'-hydroxylase, which further greatly affect flower color.

Gibberellic acid controls several aspects of plant development and growth through GA-GID-DELLA signaling module³⁴. With regard to the chlorophyll biosynthetic pathway, DELLA stabilization could lead to increased accumulation of Pchl_{ide} and PORs in etiolated seedlings³⁵. The miR171 family can destabilize mRNAs encoding SCARECROW-Like transcription factors (SCL6/SCL6-IV, SCL22/SCL6-III and SCL27/SCL6-II), which could cause developmental defects in leaves and flowers³⁶. However, SCARECROW-like transcription factors are important for development of roots, where no miR171 was produced³⁶. During flower developmental process in *R. pulchrum*, miR171 might also significantly affected the expression of DELLA and BAK proteins, which also deeply validated the importance of miR171 for plant development.

This study contributes a complete profile of miRNAs in *R. pulchrum*. This documentation of genome-wide profiling of miRNAs, their targets, and expression will enhance the understanding of developmental strategies in *Rhododendron* species. Unraveling the complex developmental mechanisms of *R. pulchrum* will benefit the genetic improvement of flower traits in *Rhododendron* species.

Materials and methods

Materials

R. pulchrum plants used in this research were over 10 years, which were cultivated in Huanggang Botanical garden (114.535° E, 30.278° N, 25 m). We have the permission to collect samples. The samples of *R. pulchrum* flowers at four different developmental stages were separately collected, including stage I (floral bud stage, late-March, dormant flower buds, floral organs had been well formed containing style, petal primordium, stamen primordia, sepal primordia, pistil primordia, and ovary), stage II (early flowering stage, early-April, 1–2 days before full bloom), stage III (full-flowering stage, middle-April, completely open petals, observable pistils and stamens), and stage IV (flower withering stage, late-April, petals began to fall). Flowers collected from three individual plants were pooled as one sample. For each stage, three biological replicates were collected. After being frozen in liquid nitrogen immediately, these samples were stored at -80°C until further extraction. The Study complies with local and national regulations and guidelines.

Small RNA library construction and sequencing

Total RNAs were isolated using TRIzol Reagent (Invitrogen) according to the manufacturer's instructions. The purity and concentration of RNAs were checked with a Nanodrop 2000 system (Thermo, MA, USA) and a Qubit® RNA Assay Kit in Qubit® 2.0 Fluorometer (Life Technologies, CA, USA), respectively. Moreover, RNA integrity was verified with the RNA Nano 6000 Assay Kit and an Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, USA). Small RNA libraries were constructed with TruSeq Small RNASample prep Kit. After PCR amplification, the enriched libraries were added with sequencing adaptors. Then, the libraries were purified, and checked with Agilent High Sensitivity DNA Kit on the Agilent Bioanalyzer 2100 system. Quantitative analysis of the libraries were performed with Quant-iT PicoGreen dsDNA Assay Kit. Sequences were carried out on HiSeq 2500 PE125 platform (single-end).

Identification of miRNAs through deep sequencing

Raw reads (fastq format) were processed through both custom Perl and Python scripts to obtain clean reads. In particular, certain range of lengths were chosen for downstream analysis. Small RNA tags were mapped to the reference genome of *R. simsii* (<https://www.ncbi.nlm.nih.gov/genome/94195>). The mapped small RNA tags were further aligned to Repeat Masker, Rfam database, and Nr (NCBI nonredundant protein sequences) to remove tags originating from repeat sequences, protein-coding genes, tRNA, rRNA, snRNA, and snoRNA. Afterwards, the remaining tags were aligned with known miRNAs from other plant species in miRBase21.0 and modified by mirdeep2 software^{37,38}. Based on hairpin structure of miRNA precursor, novel miRNA were predicted with miREvo and mirdeep2 softwares through exploring secondary structure, dicer cleavage site, as well as minimum free energy of small RNA tags unannotated in the former steps^{38,39}. Particularly, the miFam.dat (<http://www.mirbase.org/ftp.shtml>) was used to screen families of known miRNAs²⁹. The miRNA target genes were predicted with psRNATarget server (<http://plantgrn.noble.org/psRNATarget/>). To expand the utility of sequencing data, sRNAs obtained from the same developmental stage were pooled and were further assembled into a non-redundant set.

Differential expression analysis of miRNAs

The miRNA expression levels were calculated by TPM (transcript per million) using a normalization formula $\text{Normalized expression} = \text{Mapped read count} / \text{Total reads} \times 1,000,000$ ⁴⁰. Differential expression analysis between two samples was carried out with the DESeq R package (1.8.3). Benjamini & Hochberg method was used to adjust the P-values, and 0.05 was set as the threshold for significantly differential expression by default²⁹.

Target prediction, GO enrichment and KEGG pathway analysis

Novel miRNA, significantly expressed and conserved, were selected and used for target prediction using the psRNA Target server by psRobot_tar in psRobot⁴¹. Target gene candidates of differentially expressed miRNAs were performed through Gene Ontology (GO) enrichment analysis implemented with Goseq based Wallenius non-central hyper-geometric distribution⁴². Moreover, GO terms were also submitted to enrichment analysis by agriGO tool⁴³. Functional annotation of miRNA was also performed through KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analysis. KOBAS software was used to test the statistical enrichment of target gene candidates involved in KEGG pathways⁴⁴.

Validation of miRNA expression with qRT-PCR analysis

Total RNA was isolated from *R. pulchrum* flowers with TRIzol reagent (Invitrogen) according to the manufacturer's instructions. The quality and quantity of RNA were tested with agarose gel electrophoresis and ultraviolet spectrophotometer. Mir-X miRNA First-Strand Synthesis Kit and Mir-X miRNA qRT-PCR TB Green Kit were selected (Takara) in this research. Nine miRNAs were randomly selected for real-time qPCR analysis on a Step OnePlus Real-Time PCR system (Applied Biosystems, USA). The 5' primer were miRNA-specific (Table S8), and the 3' primer is the mRQ 3' primer supplied with the kit. Relative expression level of each miRNA was normalized against EF1 expression levels, and fold-change was calculated according to the $2^{-\Delta\Delta\text{CT}}$ method. In particular, three technical replicates were carried out. Real-time PCR reaction was conducted at 95°C for 5 min, followed by 40 cycles (incubations at 95°C for 15 s, and 60°C for 31 s). The dissociation curve was analyzed at 55°C – 95°C .

Data availability

The datasets generated and analysed during current study are available in the NCBI database under the Bio-Project, Biosample, and SRA numbers of PRJNA485857, SAMN09829198-SAMN09829201, and SRR7698534-SRR7698537, respectively.

Received: 4 July 2023; Accepted: 12 October 2023

Published online: 20 October 2023

References

- Meijón, M., Cañal, M. J., Valledor, L., Rodríguez, R. & Feito, I. Epigenetic and physiological effects of gibberellin inhibitors and chemical pruners on the floral transition of azalea. *Physiol. Plant.* **141**(3), 276–288 (2011).
- Tymon, L. S. *et al.* Phylogeny and taxonomy of *Erysiphe* spp. on *Rhododendron*, with a special emphasis on North American species. *Mycologia* **114**(5), 887–899 (2022).
- Bhattacharyya, D. *Rhododendron* species and their uses with special reference to Himalayas—A review. *Assam Univ. J. Sci. Technol.* **7**, 161–167 (2011).
- Wang, S. *et al.* Comparative analysis of microsatellite, SNP, and InDel markers in four *Rhododendron* species based on RNA-seq. *Breed. Sci.* **68**, 536–544 (2018).
- Suzuki, K., Yabuki, T. & Ono, Y. Roadside *Rhododendron pulchrum* leaves as bioindicators of heavy metal pollution in traffic areas of Okayama, Japan. *Environ. Monit. Assess.* **149**(1–4), 133–141 (2009).
- Takeda, Y. *et al.* Molecular and epigenetic control of aldosterone synthase, CYP11B2 and 11-Hydroxylase, CYP11B1. *Int. J. Mol. Sci.* **24**(6), 5782 (2023).
- Allahverdy, J. & Rashidi, N. MicroRNAs induced by *Listeria monocytogenes* and their role in cells. *Microb. Pathog.* **175**, 105997 (2023).
- Vachon, G., Engelhorn, J. & Carles, C. C. Interactions between transcription factors and chromatin regulators in the control of flower development. *J. Exp. Bot.* **69**, 2461–2471 (2018).
- Thomson, B. & Wellmer, F. Molecular regulation of flower development. *Curr. Top. Dev. Biol.* **131**, 185–210 (2019).
- Prunet, N., Yang, W., Das, P., Meyerowitz, E. M. & Jack, T. P. SUPERMAN prevents class B gene expression and promotes stem cell termination in the fourth whorl of *Arabidopsis thaliana* flowers. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 7166–7171 (2017).
- Wójcik, A. M. & Gaj, M. MiR393 contributes to the embryonic transition induced in vitro in *Arabidopsis* via the modification of the tissue sensitivity to auxin treatment. *Planta* **244**, 231–243 (2016).
- Zhong, T. *et al.* Identification and characterization of microRNAs in the goat (*Capra hircus*) rumen during embryonic development. *Front. Genet.* **8**, 163 (2017).
- Wang, Y. *et al.* Improvement in fruit quality by overexpressing miR399a in woodland strawberry. *J. Agric. Food Chem.* **65**, 7361–7370 (2017).
- Xing, H. *et al.* Genome-wide investigation of microRNAs and expression profiles during rhizome development in ginger (*Zingiber officinale* roscoe). *BMC Genom.* **23**(1), 49 (2022).
- Liang, R. *et al.* Using RNA sequencing to identify putative competing endogenous RNAs (ceRNAs) potentially regulating fat metabolism in bovine liver. *Sci. Rep.* **7**, 6396 (2017).
- Yang, W. C., Guo, W., Zan, L. S., Wang, Y. N. & Tang, K. Q. Bta-miR-130a regulates the biosynthesis of bovine milk fat by targeting peroxisome proliferator-activated receptor gamma. *J. Anim. Sci.* **95**, 2898–2906 (2017).
- Sulas, P. *et al.* A large set of miRNAs is dysregulated since the earliest steps of human hepatocellular carcinoma development. *Am. J. Pathol.* **188**, 785–794 (2017).
- Khezri, M. R., Yousefi, K., Zolbanin, N. M. & Ghasemnejad-Berenji, M. MicroRNAs in the pathophysiology of Alzheimer's disease and Parkinson's disease: An overview. *Mol. Neurobiol.* **59**(3), 1589–1603 (2022).
- Zhang, M., Xian, H. C., Dai, L., Tang, Y. L. & Liang, X. H. MicroRNAs: Emerging driver of cancer perineural invasion. *Cell Biosci.* **11**(1), 117 (2021).
- Abel, Y. *et al.* The interaction between RPA3 and TRBP reveals a possible involvement of the HSP90/R2TP chaperone complex in the regulation of miRNA activity. *Nucleic Acids Res.* **50**(4), 2172–2189 (2022).
- Meyers, B. C. *et al.* Criteria for annotation of plant MicroRNAs. *Plant Cell.* **20**, 3186–3190 (2008).
- Li, H. *et al.* Characterization of the stress associated microRNAs in *Glycine max* by deep sequencing. *BMC Plant Biol.* **11**, 170 (2011).
- McCormick, K. P., Willmann, M. R. & Meyers, B. C. Experimental design, preprocessing, normalization and differential expression analysis of small RNA sequencing experiments. *Silence* **2**(1), 2 (2011).
- Choudhary, S., Thakur, S., Majeed, A. & Bhardwaj, P. Exploring microRNA profiles for circadian clock and flowering development regulation in Himalayan *Rhododendron*. *Genomics* **111**(6), 1456–1463 (2019).
- Ma, L. *et al.* Cloning and identification of novel mirnas in the flower organs of Korla fragrant pear at anthesis. *J. Horticult. Sci. Biotechnol.* **94**(3), 1–12 (2019).
- Sun, Z. C., Zhang, L. S. & Wang, Z. J. Genome-wide analysis of mirnas in *Carya cathayensis*. *BMC Plant Biol.* **17**(1), 228 (2017).
- Shi, Q. Q. *et al.* Identification of microRNAs and their targets involved in *Paeonia rockii* petal variegation using high-throughput sequencing. *J. Am. Soc. Horticult. Sci.* **144**(2), 1–22 (2019).
- Li, D. *et al.* Deep sequencing of maize small RNAs reveals a diverse set of microRNA in dry and imbibed seeds. *PLoS One* **8**(1), e55107 (2013).
- Liu, X. X., Luo, X. F., Luo, K. X., Liu, Y. L. & Pan, T. Small RNA sequencing reveals dynamic microRNA expression of important nutrient metabolism during development of *Camellia oleifera* fruit. *Int. J. Biol. Sci.* **15**(2), 416–429 (2019).
- Kundu, A., Paul, S., Dey, A. & Pal, A. High throughput sequencing reveals modulation of microRNAs in *Vigna mungo* upon Mungbean Yellow Mosaic India virus inoculation highlighting stress regulation. *Plant Sci.* **257**, 96–105 (2017).
- Jung, J. H. & Park, C. M. MIR166/165 genes exhibit dynamic expression patterns in regulating shoot apical meristem and floral development in *Arabidopsis*. *Planta* **225**(6), 1327–1338 (2007).
- Liu, S., Ju, J. & Xia, G. Identification of the flavonoid 3'-hydroxylase and flavonoid 3',5'-hydroxylase genes from *Antarctic moss* and their regulation during abiotic stress. *Gene* **543**(1), 145–152 (2014).
- Seitz, C., Ameres, S., Schlangen, K., Forkmann, G. & Halbwirth, H. Multiple evolution of flavonoid 3',5'-hydroxylase. *Planta* **242**(3), 561–573 (2015).
- Sun, X. *et al.* A functionally required unfoldome from the plant kingdom: Intrinsically disordered N-terminal domains of GRAS proteins are involved in molecular recognition during plant development. *Plant Mol. Biol.* **77**, 205–223 (2011).
- Ma, Z. *et al.* *Arabidopsis* miR171-targeted scarecrow-like proteins bind to GT cis-elements and mediate gibberellin-regulated chlorophyll biosynthesis under light conditions. *PLoS Genet.* **10**(8), e1004519 (2014).
- Cho, J. & Paszkowski, J. Regulation of rice root development by a retrotransposon acting as a microRNA sponge. *Elife.* **6**, e30038 (2017).

37. Kozomara, A. & Sam, G. J. MiRBase: Annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.* **42**, D68–D73 (2014).
38. Friedländer, M. R., Mackowiak, S., Li, N., Chen, W. & Rajewsky, N. MiRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res.* **40**, 37–52 (2012).
39. Wen, M., Shen, Y., Shi, S. & Tang, T. MiREvo: An integrative microRNA evolutionary analysis platform for next-generation sequencing experiments. *BMC Bioinform.* **13**, 140 (2012).
40. Zhou, L. *et al.* Integrated profiling of microRNAs and mRNAs: microRNAs located on Xq27.3 associate with clear cell renal cell carcinoma. *PLoS One* **5**(12), e15224 (2010).
41. Wu, H. J., Ma, Y., Chen, T., Wang, M. & Wang, X. PsRobot: A web-based plant small RNA meta-analysis toolbox. *Nucleic Acids Res.* **40**, W22–W28 (2012).
42. Young, M. D., Wakeeld, M. J., Smyth, G. K. & Oshlack, A. Goseq: Gene ontology testing for RNA-seq datasets. *ResearchGate* **8**, 1–26 (2012).
43. Candar-Cakir, B., Arican, E. & Zhang, B. Small RNA and degradome deep sequencing reveals drought- and tissue-specific microRNAs and their important roles in drought-sensitive and drought-tolerant tomato genotypes. *Plant Biotechnol. J.* **14**, 1727–1746 (2016).
44. Mao, X., Cai, T., Olyarchuk, J. G. & Wei, L. Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics* **21**, 3787–3793 (2005).

Acknowledgements

This work was supported by research grants from National Natural Science Foundation of China (NSFC 31500995), Natural Science Foundation of Hubei Provincial Department of Science and Technology (2021CFB320), and Scientific and Technological Research Project of Hubei Provincial Department of Education (D20222902 and B2022204).

Author contributions

F.B. and H.Z. designed the experiments and prepared the draft manuscript. Y.J., Z.J. and D.H. collected plant materials. S.Y. and Z.W. performed the experiments and data analysis. W.S.Z. took charge of the whole research.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-44779-z>.

Correspondence and requests for materials should be addressed to S.W.

Reprints and permissions information is available at www.nature.com/reprints.

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023