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

Bo Fang
Huanggang Normal University
Zhiwei Huang
Huanggang Normal University
Yirong Sun
Huanggang Normal University
Wanjing Zhang
Huanggang Normal University
Jiaojun Yu
Huanggang Normal University
Jialiang Zhang
Huanggang Normal University
Hongjin Dong
Huanggang Normal University
Shuzhen Wang (wangshuzhen710@whu.edu.cn)
Huanggang Normal University

Article

Keywords: Rhododendron pulchrum Sweet, MicroRNA, Illumina sequencing, anthocyan synthesis, genetic improvement

Posted Date: July 26th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3139653/v1

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Scientific Reports on October 20th, 2023. See the published version at https://doi.org/10.1038/s41598-023-44779-z.

Title page: #

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

Bo Fang[#], Zhiwei Huang[#], Yirong Sun, Wanjing Zhang, Jiaojun Yu, Jialiang Zhang, Hongjin Dong, Shuzhen Wang*

College of Biology and Agricultural Resources, Huanggang Normal University, Huanggang, 438000, Hubei Province, P.R. China

Running title: Small RNA-seq revealling flower development in *R. pulchrum*

* Author to whom correspondence should be addressed;

contributing equally to this work;

E-Mail: wangshuzhen710@whu.edu.cn; Tel. and Fax: +86-0713-8833606.

Abstract: *Rhododendron pulchrum* Sweet, a member of the Ericaceae family

possessing valuable medicinal and 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 clarify the function of miRNAs in the synthesis of anthocyanin 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 were dominant families. In particular, 126 miRNAs were significantly differentially expressed among four different flowering stages. Based on integrated GO and KEGG function annotations, the miRNA target genes were mostly involved in metabolic pathways, plant hormone signal transduction, mitosis and regulation of biosynthetic processes. Totally, 593 genes were differentially regulated by miRNAs during the flower development process. 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.

Key words: *Rhododendron pulchrum* Sweet; MicroRNA; Illumina sequencing; anthocyan synthesis; genetic improvement

Introduction

Rhododendron genus contains approximately 1000 evergreen and deciduous species, as well as thousands of commercial hybrids, is widely distributed in Europe, Asia, and North America (Meijón et al. 2011; Tymon et al. 2022). As a typical large vascular plant genus within the Ericaceae family, species belonging to *Rhododendron* genus possess valuable medicinal and horticultural properties, and are mainly distributed in Northern Hemisphere (Bhattacharyya 2011; Wang et al. 2018). Besides bioindicator of metal pollution (Pb, Zn, and Cd), the evergreen *R. pulchrum* is more famous due to their

beautiful vegetative forms, bright-colored flowers, and long flowering period ranging from late March to early May (Suzuki et al. 2009). Therefore, *R. pulchrum* is widely served as roadside trees in urban and rural areas (Suzuki et al. 2009).

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, or microRNAs (miRNAs) (Takeda et al. 2023; Allahverdy and Rashidi, 2023). In particular, transcription factor-encoding genes AINTEGUMENTA (ANT), auxin efflux carrier Pin-Formed 1 (PIN1) and Aintegumenta-like 6 (AIL6), regulatory gene LEAFY (LFY), MADS-domain transcription factors Agamous-like 24 (AGL24) and Short Vegetative Phase (SVP), as well as shoot identity gene Terminal Flower1 (TFL1) were also deeply studied (Vachon et al. 2018; Thomson and Wellmer, 2019). 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 underlining flowering (Prunet et al. 2017). However, the systems-wide understanding of gene regulatory networks underlying flower developmental process needs more efforts.

MiRNAs (approximately 22 nt), a class of non-coding single strand molecules negatively regulating gene expressions both at transcriptional and posttranscriptional levels via cleaving targets or repressing protein translation, play critical roles in biological processes, including embryogenesis and organ development (Wójcik and Gaj, 2016; Zhong et al. 2017; Wang et al. 2017; Xing et al. 2022), metabolism (Liang et al. 2017; Yang et al. 2017), as well as tumorigenesis and chronic disease in animals (Sulas et al. 2017; Khezri et al. 2022). In particular, functional mature miRNAs could be incorporated into Argonaute 1 (AGO1)-containing RNA-induced silencing complex (RISC) target endogenous transcripts, and causing degradation or translational repression in sequence-specific manner (Zhang et al. 2021; Abel et al. 2022). The specific position of target AGO1 protein is often found between the 10^{th} and 11^{th}

nucleotides of target mRNA (Meyers et al. 2008).

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 synthesis of anthocyanin. In order to explore the potential role of miRNAs during the flowering process, miRNA expression profiles of flowers at four different developmental stages were investigated with high throughput next generation 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.

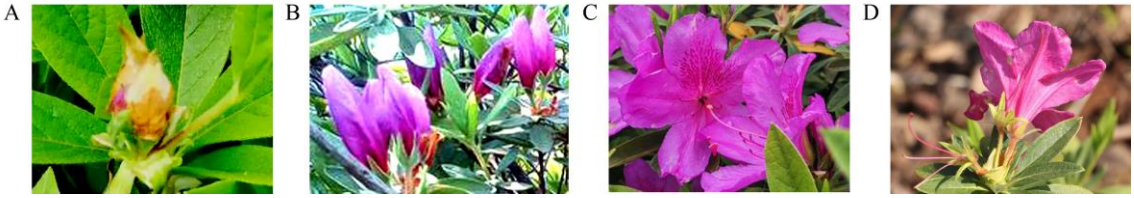

Figure 1 Flower tissues of four different stages: A-D representing stage I (floral bud stage), stage II (early flowering stage), stage III (full-flowering stage), and stage IV (flower withering stage), respectively.

Results

Sequence analysis of small RNA in R. pulchrum flowers

Twelve small RNA libraries (three libraries for each developmental stage) were constructed and sequenced for tissue samples (Figure 1). 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 (≧18nt) 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 were observed in these different sample types. Among 18-40 nt small RNAs, the majority ranged from 21 to 25 nt in length, and the percentages were 82.79%-84.15%,

66.97%-67.84%, 54.29%-56.00%, and 48.24%-52.75%, respectively. 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 samples at stage I, but 21 nt for the rest samples. The third abundant sRNAs were 22nt in samples at stage II, stage III, and stage IV; while 21 nt in samples at stage I (Figure 2).

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

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	7,218-8,099	11,913-13,908
snoRNA	14,290-15,991	12,979-14,493	7,499-8,312	10,740-12,315
Known _miRNA	1,500-1,622	1,678-1,843	1,231-1,435	1.454-1,533

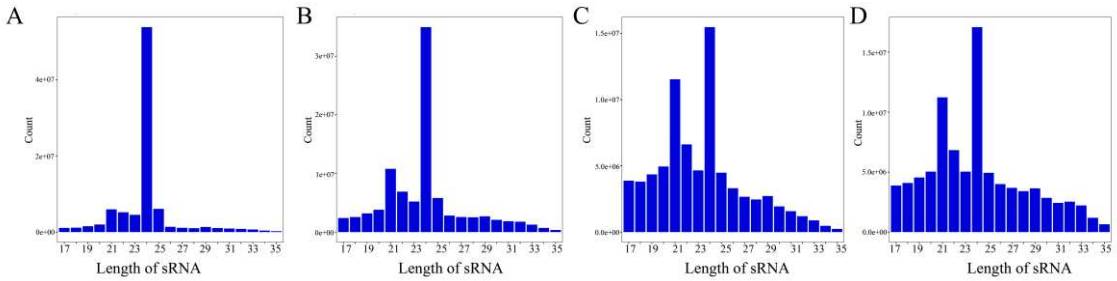

Figure 2 Length distribution of sRNA in four flower samples at stage I (A), stage II (B), stage III (C), and stage IV (D), respectively.

Unique reads were obtained through removing redundant sequence. In relation to *R. simsii* genome, 75.34%, 78.24%, 90.39%, and 87.04% clean reads were mapped to the reference genome for stage I, stage II, stage III, and stage IV, respectively. Particularly, the most abundant siRNA 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 (Figure 3). After annotation, 19,940,384-24,302,325, 22,623,935-26,572,199, 21,446,057-26,735509, and 21,089,468-30,160,793 small RNA tags were obtained in *R. pulchrum* flower at different developmental stages, which were classified into miRNAs, tRNAs, rRNAs. The percentage of mapped reads ranged from 75.30%-75.38% (stage I) to 90.25%-90.49% (stage III).

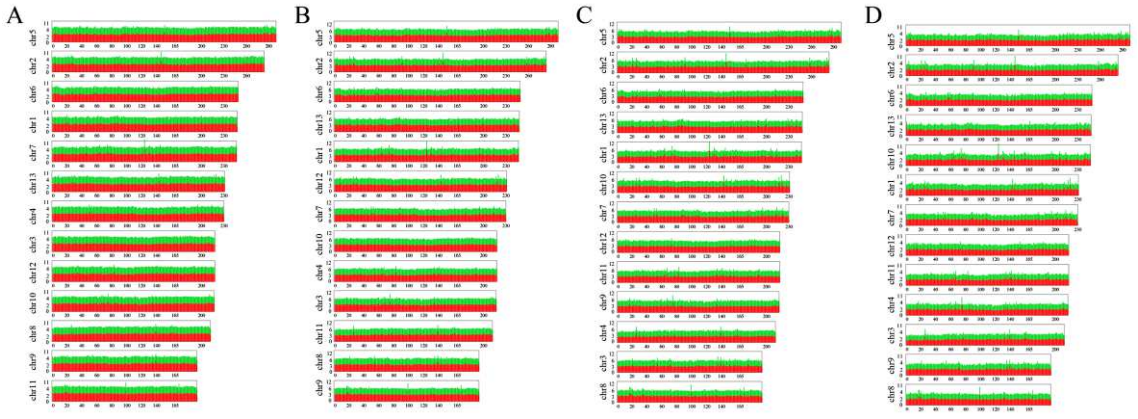

Figure 3 The genome distribution of sRNAs in four flower samples at stage I (A), stage II (B), stage III (C), and stage IV (D), respectively. The red parts in Y-axis represents log10 count, and the green part represents log10 category. “Chr” was short for chromosome.

Clarification of miRNAs based on deep sequencing

A set of 722 miRNAs belonging to 104 families were obtained, among which miR166 family was the largest represented family with 24 members (Table 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 miRNA had obvious uracil (U) base preference (Figure 4A). Besides base at the first position, the 3'-end base also possessed U base preference. Moreover, the 8th and 9th bases had slight G base preference, while C base preference existed at 23th base position (Figure 4B). Totally, 211 miRNAs were commonly expression in all four stages (Figure 5 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, and 15 miRNAs were expressed in stage I and stage II, stage II and stage III, stage III and stage IV, as well as stage I and stage IV, respectively (Figure 5 and Table S3).

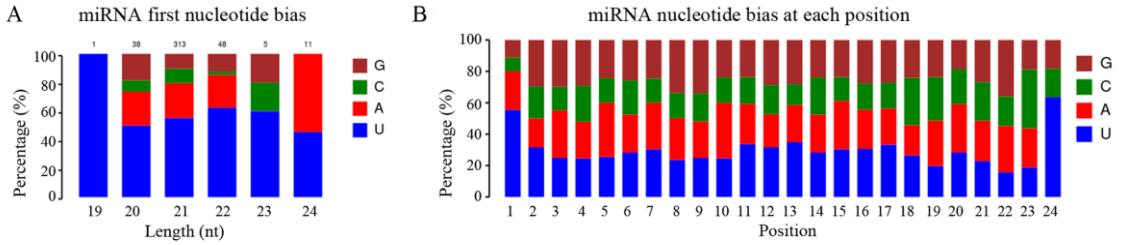

Figure 4 First nucleotide bias (A) and nucleotide bias at each position (B) of identified miRNA in *R. pulchrum*.

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 set of 84 novel mature miRNA sequences were identified in *R. pulchrum* flowers (Table S4). Totally, four length types of miRNAs were predicted, including 21 nt (55, 65.476%), 22nt (24, 28.571%), 20 nt (4, 4.762%), and 24 nt (1, 1.190%) (Figure 6A). 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 1^{th}, 2^{th}, 7^{th}, and 12^{th} position all had U base

preference (Figure 6B). Bases at 8th and 23th position had obvious C base preference. Moreover, the 24th base was complete G.

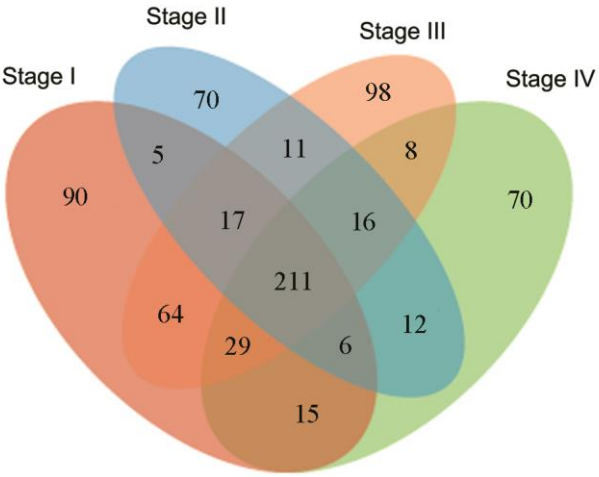

Figure 5 Venn diagram of identified miRNA in four stages.

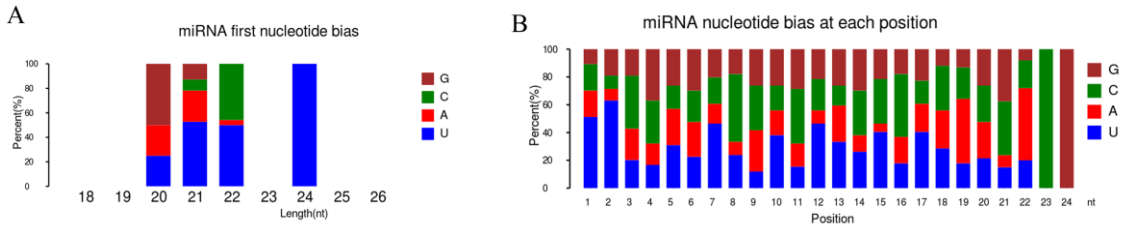

Figure 6 First nucleotide bias (A) and nucleotide bias at each position (B) of novel predicted miRNA in *R. pulchrum*.

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 differently expressed. In particular, 87 miRNAs were commonly expressed in flowers at all four different stages. Compared with stage I, a total of 31 miRNAs were up-regulated and 34 miRNAs were down-regulated in stage II (Figure 7A and Table S5). Compared with stage II, a set of 34 miRNAs were up-regulated and 37 miRNAs were down-regulated in stage III (Figure 7B and Table S5). Compared with stage III, numbers of the up-regulated and down-regulated miRNAs at stage IV were 33 and 30, respectively (Figure 7C and Table S5).

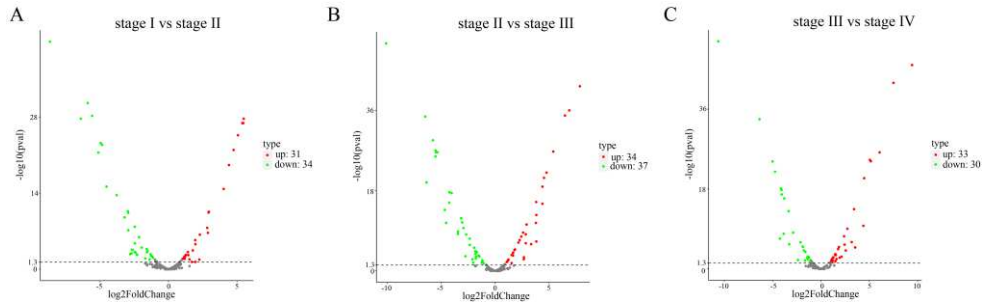

Figure 7 Volcano plots of miRNA data between stage I and stage II (A), stage II and stage III (B), as well as stage III and stage IV (C).

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 (Figure 8). Furthermore, flower tissues at stage II and stage IV were clustered first, and then grouped with flowers sampled at stage I. However, flowering at full-flowering stage (Stage III) showed less association with these other three flower samples (Figure 8A). Furthermore, gene cluster showed that these DEGs could be mainly clustered into two clades: clades 1 containing “Stage I” and “Stage II”, clade 2 consisting of “Stage III” and “Stage IV” (Figure 8B).

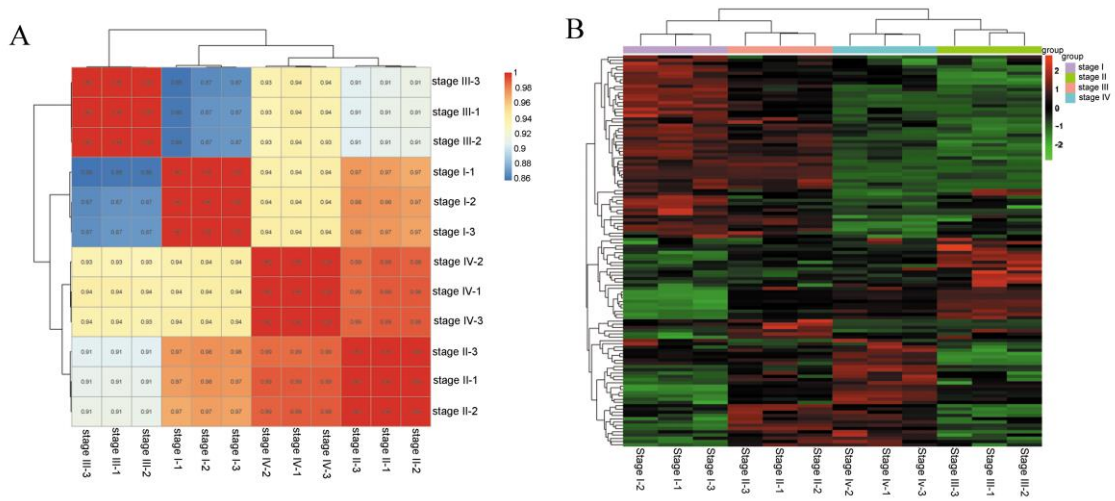

Figure 8 Distance analysis of *R. pulchrum* flower tissues collected from four developmental stages by calculating Pearson's correlation coefficient: (A) sample correlation matrix ($p < 0.001$, $FDR < 0.001$), (B) gene cluster analysis ($p < 0.001$, $FDR < 0.001$).

Target prediction and GO enrichment analysis

With transcriptome data of *R. pulchrum* flower obtained in our previous study serving as target database, 1,170 miRNA target site (593 target genes) were obtained. Among these miRNA-target pairs, 84.57% of these miRNAs potentially targeted mutiple unigenes, ranging from 2 to 157. During *R. pulchrum* flower development from stage I to stage II, a total of 1,081 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), “extracellular region” (GO:0005576), “organic cyclic compound catabolic process” (GO:1901361), and “aromatic compound catabolic process” (GO:0019439) were the main enriched terms (Figure 9A). Among GO terms associated with biologic process, the target genes were mainly involved in “monosaccharide transporter”, “xenobiotic detoxification”, and “defense response” pathways (Figure 10A). In the cellular component category, target genes were tightly associated with “nucleus and extracellular space” (Figure 10B). In molecular function, most target genes were related to “isomerase activity”, “electron transfer activity”, “acid phosphatase activity”, “structural constituents”, “monosaccharide transmembrane transporter activity”, “lipid binding, ADP binding”, and “NAD+ binding” (Figure 10C).

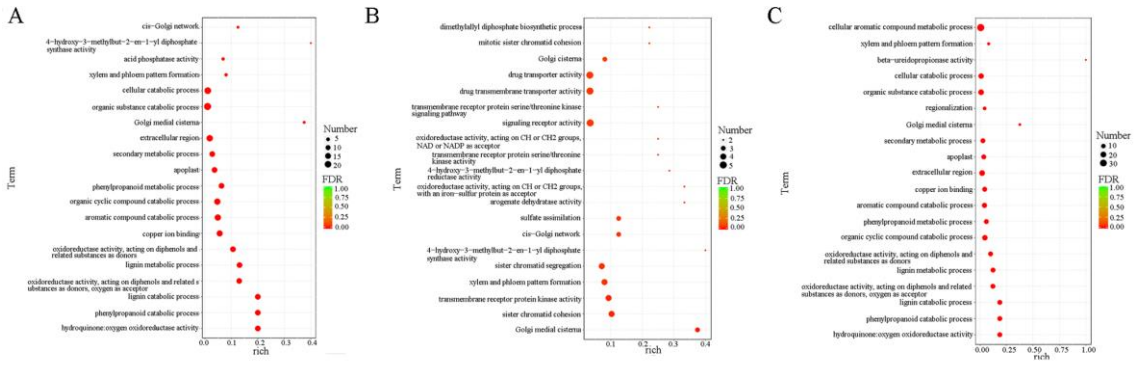

Figure 9 GO annotations of potential miRNA targets during *R. pulchrum* flower development from stage I to stage II (A), stage II to stage III (B), and stage III to stage IV (C).

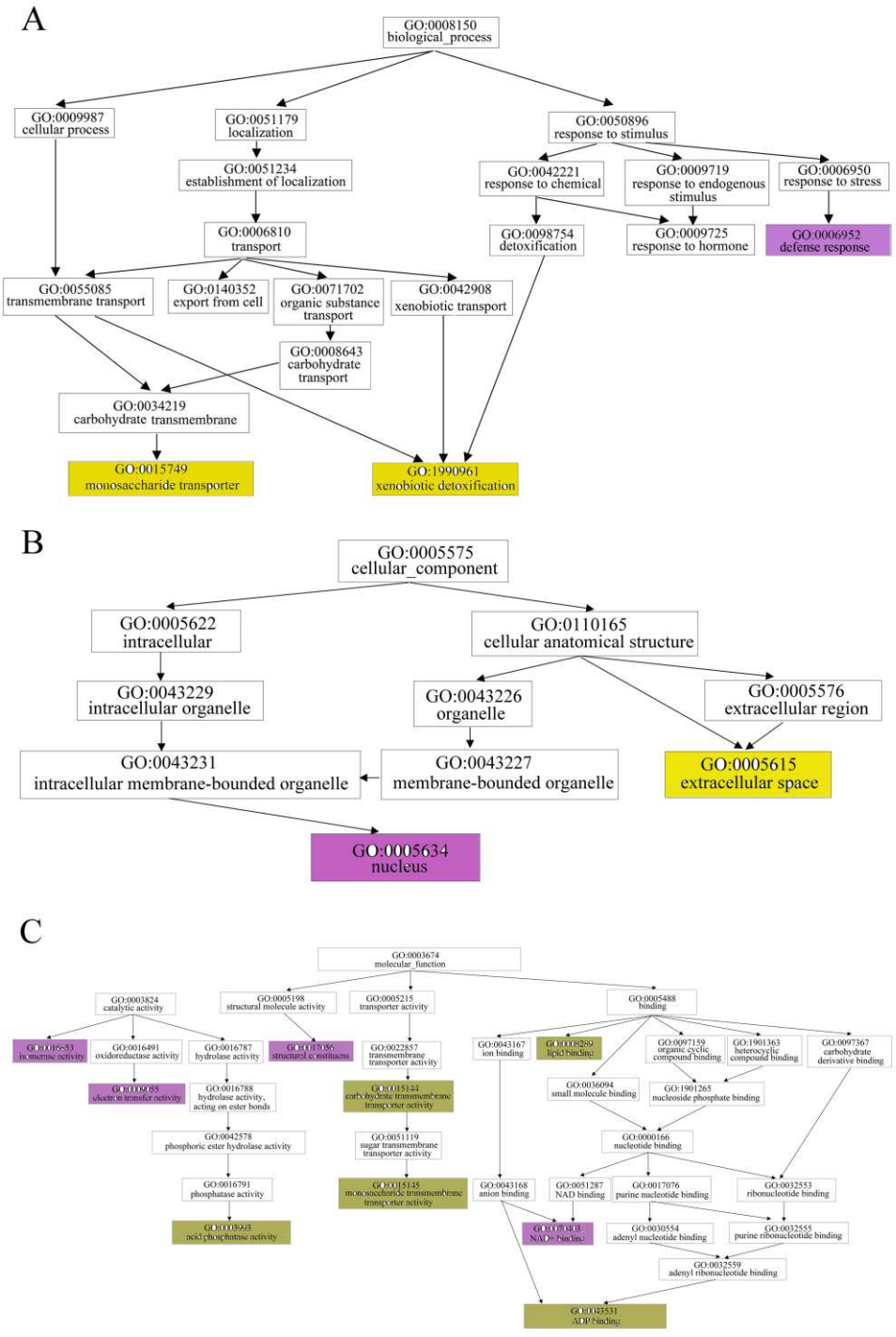

Figure 10 GO enrichment analysis of potential miRNA targets of differently expressed miRNAs during *R. pulchrum* flower development from stage I to stage II: biologic process (A), cellular component (B), molecular function (C).

From stage II to stage III, a total of 1,328 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),

224 “drug transmembrane transporter activity” (GO:0015238), “signaling receptor activity”
225 (GO:0038023), “sister chromatid segregation” (GO:0000819), “xylem and phloem
226 pattern formation” (GO:0010051), “transmembrane receptor protein kinase activity”
227 (GO:0019199), and “sister chromatid cohesion” (GO:0007062) were the most enriched
228 terms (Figure 9B). Among GO terms related to biologic process, the majority target
229 genes were involved in “histone lysine methylation”, “mRNA splicing”, “transcription
230 initiation”, “base-excision repair”, “ribosome biogenesis”, “defence response”, and
231 “response to hormone pathways” (Figure 11A). In cellular component category, target
232 genes were mainly associated with “small-subunit processome assembly”,
233 “1,3-β-D-glucan synthetase complex”, and “nucleolus” (Figure 11B). For molecular
234 function category, target genes were mainly related to “ADP binding”, “NAD^+ binding”,
235 “clathrin binding”, “4 iron,4 sulfur cluster binding”, “sulfotransferase activity”,
236 “histone-lysine N-methyltransferase activity”, “1,3-beta-β-glucan synthase activity”,
237 and “DNA-directed DNA polymerase activity” (Figure 11C).
238 From stage III to stage IV, a total of 951 potential miRNA targets were categorized
239 into molecular functions (169), cellular components (87), and biological processes (695)
240 (Table S6). Particularly, “cellular aromatic compound metabolic process”
241 (GO:0006725), “cellular catabolic process” (GO:0044248), “organic substance
242 catabolic process” (GO:1901575), “extracellular region” (GO:0005576), “aromatic
243 compound catabolic process” (GO:0019439), and “organic cyclic compound catabolic
244 process” (GO:1901361) were the abundant enriched terms (Figure 9C). For GO terms
245 associated with biologic process, the target genes were mainly involved in “histone
246 lysine methylation”, “transcription initiation”, “mRNA splicing”, “covalent chromatin
247 modification”, “lignin catabolic process”, “ribosome biogenesis”, “base-excision repair”,
248 “response to hormone pathways” (Figure 12A). Among cellular component category,
249 target genes were tightly associated with “small subunit processome”, “nucleus”, and
250 “apoplast” (Figure 12B). In molecular function category, most target genes were related
251 to “NAD+ binding”, “copper ion binding”, “4 iron,4 sulfur cluster binding”,

"hydroquinone:oxygen oxidoreductase activity", "histone-lysine N-methyltransferase activity", "acid phosphatase activity", "phosphotransferases", "DNA-directed DNA polymerase activity" (Figure 12C).

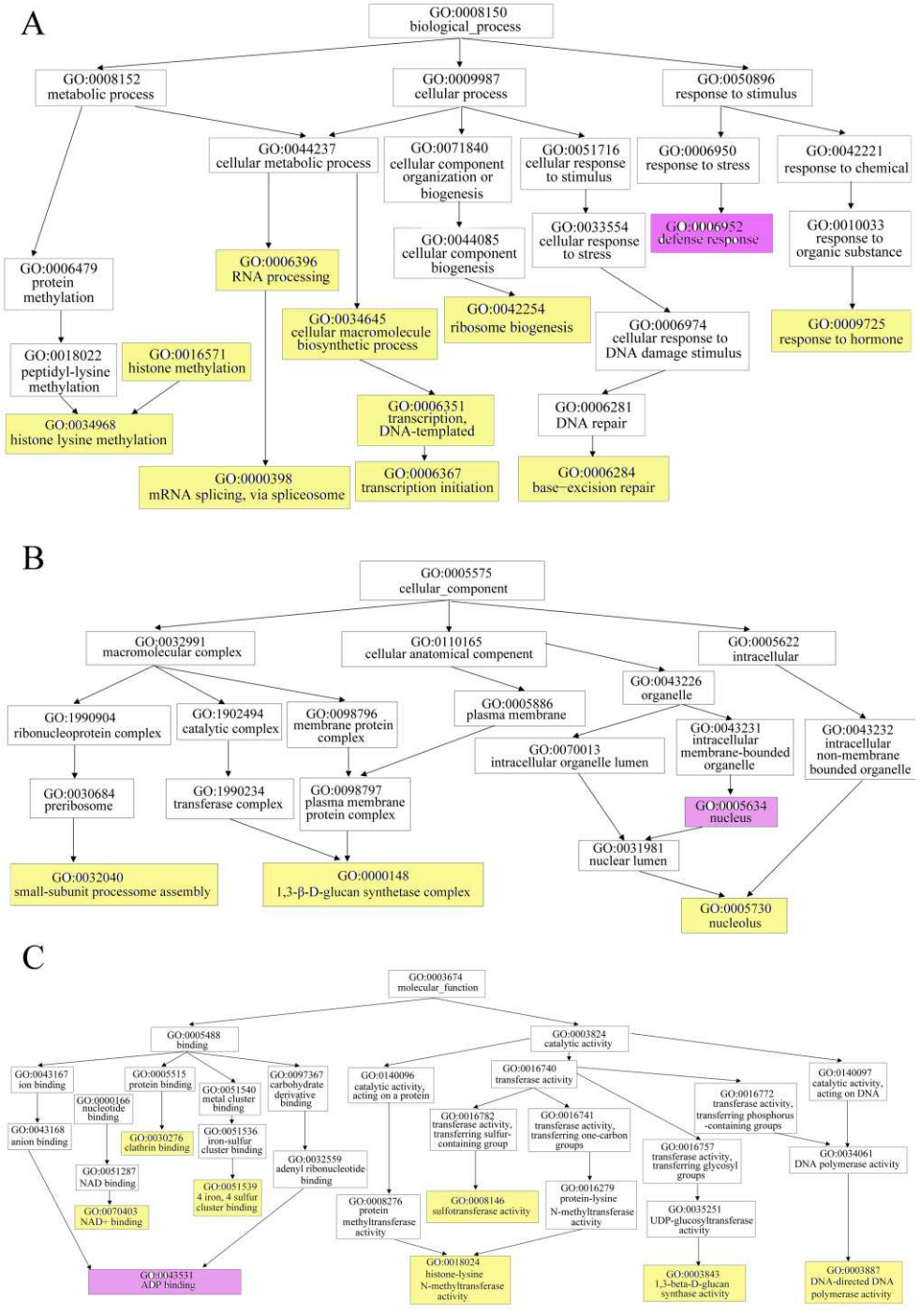

Figure 11 GO enrichment analysis of potential miRNA targets of differently expressed miRNAs during *R. pulchrum* flower development from stage II to stage III: biologic process (A), cellular component (B), molecular function (C).

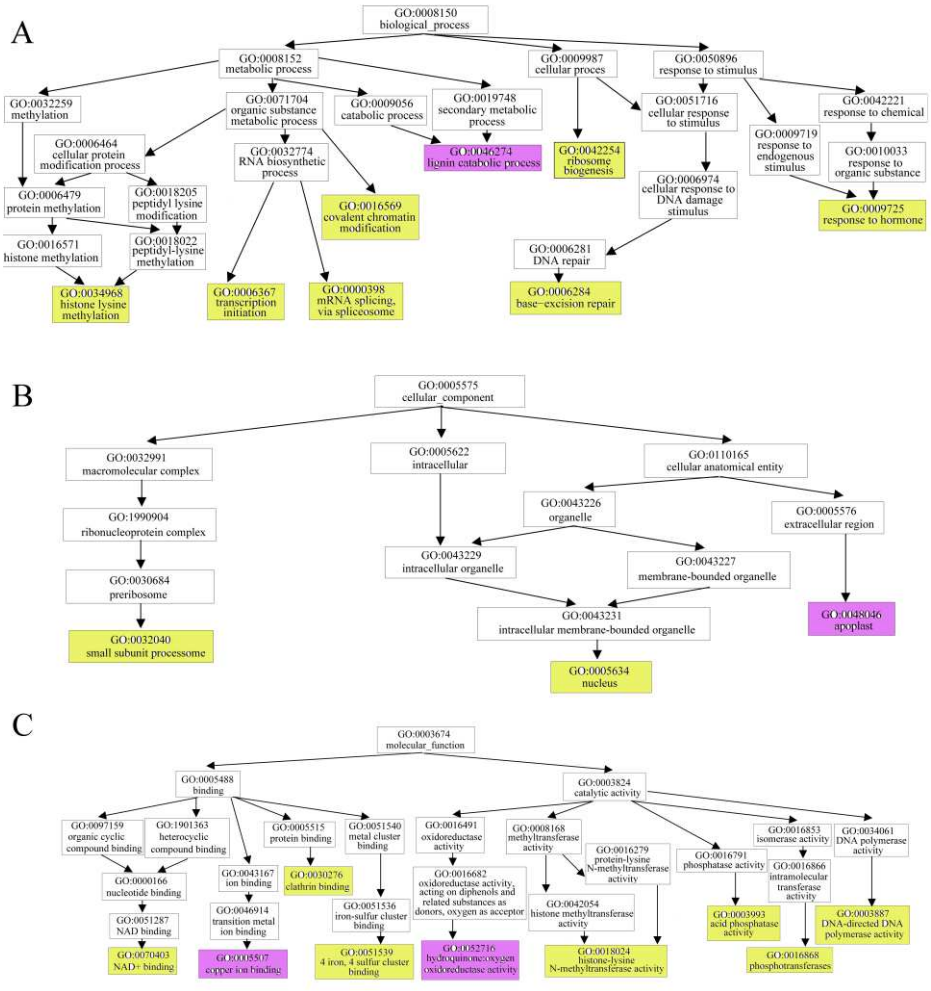

Figure 12 GO enrichment analysis of potential miRNA targets of differently expressed miRNAs during *R. pulchrum* flower development from stage III to stage IV: biologic process (A), cellular component (B), molecular function (C).

KEGG pathway analysis of potential targets of known miRNAs

During *R. pulchrum* flower development 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), “Amino sugar and nucleotide sugar metabolism” (ko00520) (Figure 13A and Table S7). For the “Flavone and flavonol biosynthesis” pathway (ko00944), gra-miR750 significantly regulated the expression of flavonoid 3’,5’-hydroxylase, which was involved in the biosynthesis of secondary metabolites including 3’-O-Methylluteolin, quercetin, and myricetin (Figure 14).

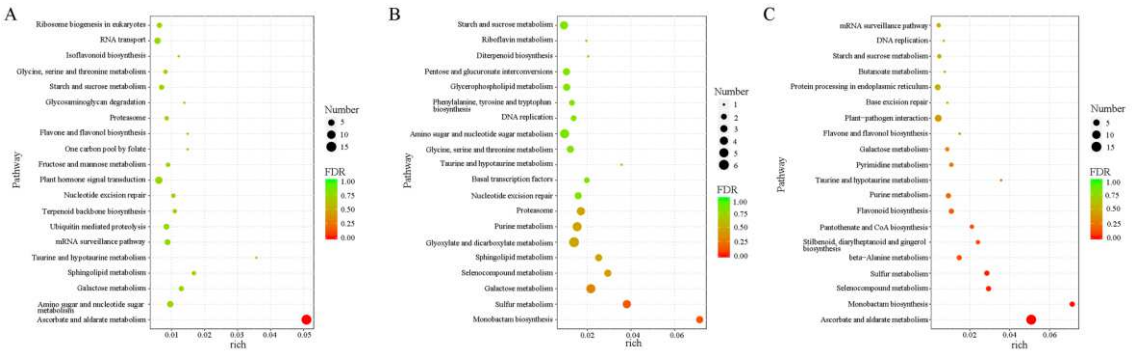

Figure 13 KEGG pathway analysis of potential miRNA targets during *R. pulchrum* flower development from stage I to stage II (A), stage II to stage III (B), and stage III to stage IV (C).

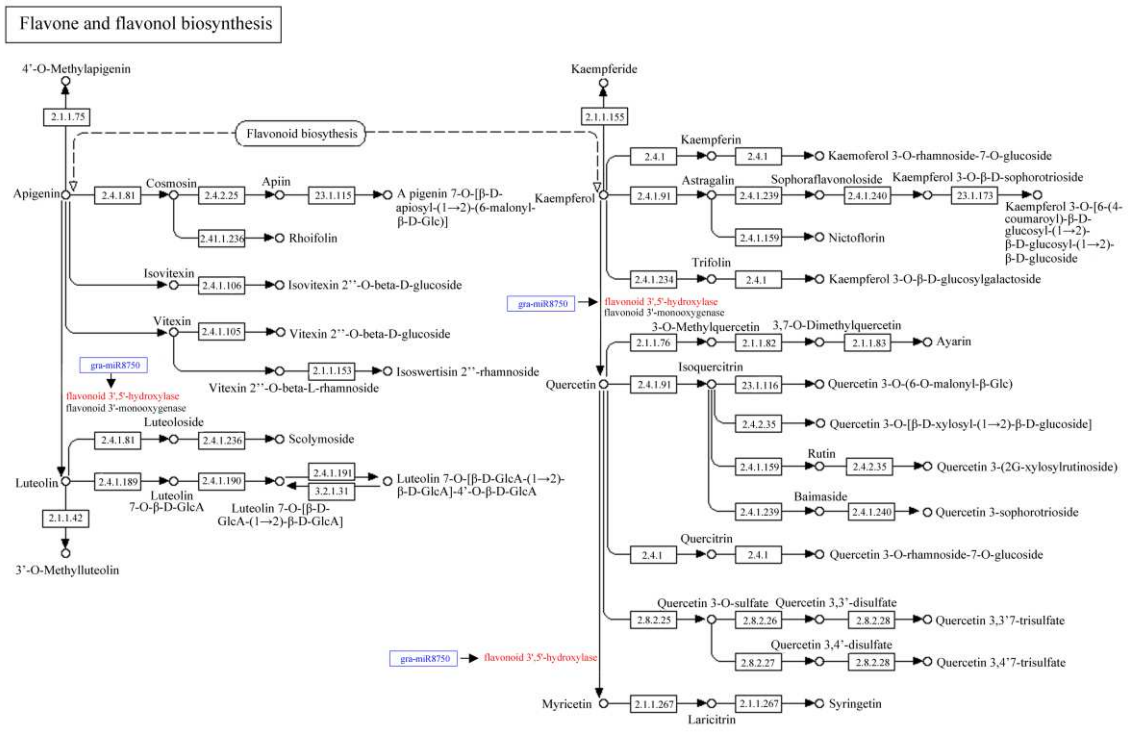

Figure 14 Metabolic regulation of miRNA in the process of flavone and flavonol biosynthesis.

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) (Figure 13B 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 (Figure 15). 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 (Figure 15).

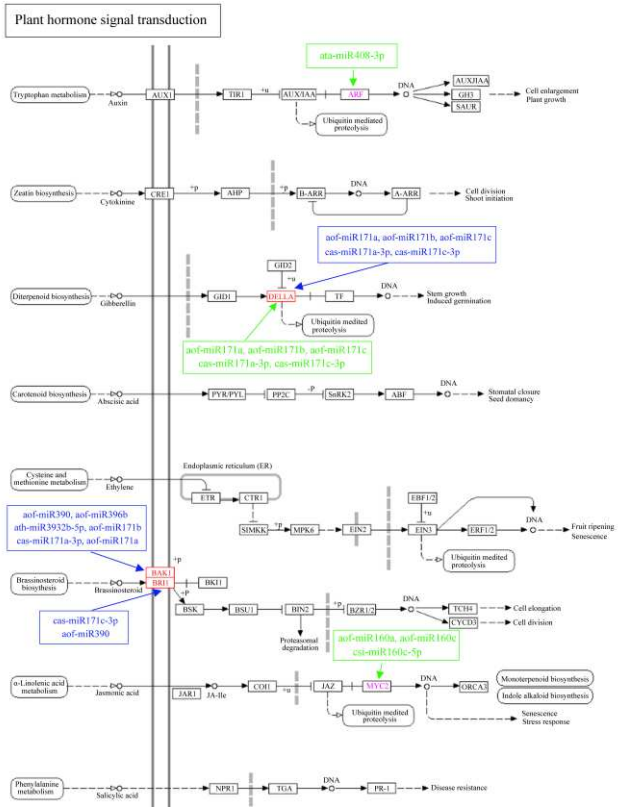

Figure 15 Metabolic regulation of miRNA in the process of plant hormone signal transduction.

From stage II to stage III, 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) (Figure 13C and Table S7). For the "Plant hormone signal transduction" (ko04075), auxin response factor (ARF), DELLLA, and MYC2 were significantly regulated by miRNAs (Figure 15). 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, the aof-miR160a, aof-miR160c, and csi-miR160c-5p regulated the expression of transcription factor MYC2 involved in the 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, 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-11 (Table 2). The qPCR analysis showed that up-regulated and down-regulated results of miRNAs were similar to deep sequencing results (Figure 16). Good correlation (R^2 =0.8919) was obtained between sequencing data and qRT-PCR results, further confirmed the high reliability of miRNA sequencing data. Epression pattern of corresponding target genes was also consistent with prediction. 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.

Table 2 Information on primers of miRNA for real-time qPCR amplification.

No.	miRNA ID	5’ Primers (5’-3’)
1	rsi-MIR158-3	TCCCAAATGTAGACAAAGC
2	rsi-MIR159-8	TTTGGATTGAAGGGAGCTCTC
3	rsi-MIR162-2	TGGACGCAGCGGTTCATCGATC
4	rsi-MIR167_1-3	TGAAGCTGCCAGCATGATCTAA
5	rsi-MIR171_1-11	TTGAGCCGTGCCAATATCACT
6	rsi-MIR396-1	TTCAAGAAAGCTGTGGGAAG
7	rsi-MIR398-6	TGTGTTCTCAGGTCACCCCTT

8	rsi-MIR535-2	TGACGACGAGAGAGAGCACGC
9	rsi-undef-11	TGCCTGGCTCCCTGTATGCCA

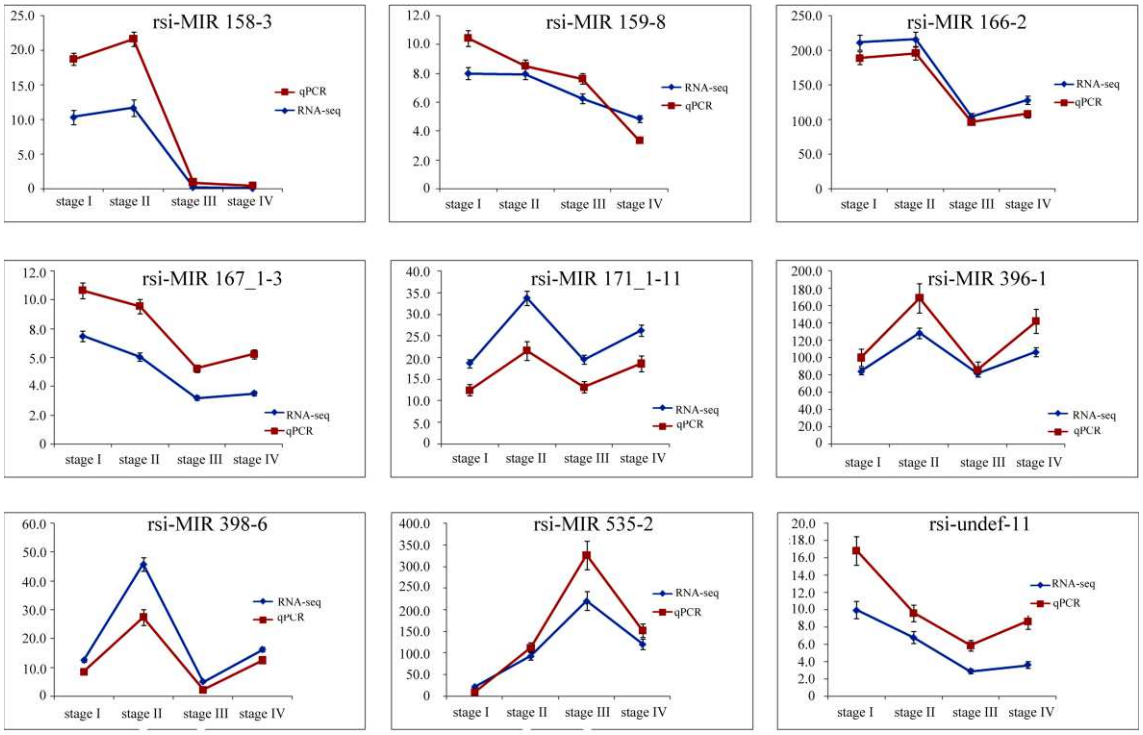

Figure 16 Correlation analyses of nine miRNAs between Illumina sequencing and qRT-PCR data.

Discussion

During evolution, plants have evolved considerable degree of developmental plasticity in response to environmental signals (Li et al. 2011). The miRNAs have been evidenced as key regulators in gene expression, developmental processes, and even stress tolerence (Li et al. 2011). Recently, the bioinformatics, as well as sequencing approaches and technologies have facilitated the rapid and accurate miRNA detection and analysis at single-base pair resolution (McCormick et al. 2011). 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. However, only 256 conserved miRNAs have been detected in leaves of *R. arboretum* located on Himalayas (Choudhary et al. 2019). According to Choudhary et al. (2019), 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. In particular, miR166, miR156, miR159, and miR171_1 families were dominant for *R. pulchrum* flower development, which was partly similar with that of maize (Li et al. 2013), *Camellia oleifera* (Liu et al. 2019), and *Vigna mungo* (Kundu et al. 2017). According to Jung and Park (2007), miR166/165 group and corresponding target genes regulate shoot apical meristem and floral development through WUSCHEL-CLAVATA) pathway. A set of 84 novel mature miRNA sequences (21nt-24nt), showing G and U bias, were identified in *R. pulchrum* flowers, which were less than that of *R. arboretum* leaves (210 novel miRNAs) (Choudhary et al. 2019).

For four different flowering stages, 90, 70, 98, and 70 miRNA were expressed only in floral bud stage, early flowering stage, full-flowering stage, and flower withering stage, respectively. 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 adaptability of *R. pulchrum*. Moreover, exogenous environmental cues also mediate phase transition through networking of flowering pathways as well as corresponding components.

During process of *R. pulchrum* flower development, 1,170 miRNA target site (593 target genes) have been screened, much less than that of *R. arboretum* leaves (27,139 predicted targets) (Choudhary et al. 2019). 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) (Liu et al. 2014). As cytochrome P450-dependent monooxygenases, flavonoid 3',5'-hydroxylase could hydroxylate the B-ring of flavonoids at both 3'- and 5'-position, respectively (Seitz et al. 2015). 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-miR750 significantly regulated 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 module (Sun et al. 2011). With regard to the chlorophyll biosynthetic pathway, DELLA stabilization could lead to increased accumulation of Pchlide and PORs in etiolated seedlings (Ma et al. 2014). 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 flowers, (Cho et al. 2017). However, SCARECROW-like transcription factors are important for development of roots, where no miR171 was produced (Cho et al. 2017). During flower developmental process in *R. pulchrum*, miR171 significantly regulated 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 miRNA, 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 was 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. After being frozen in liquid nitrogen immediately, these samples were stored at -80 °C until further extraction.

Small RNA library construction and sequencing

Total RNAs of each stage were isolated from flower tissues using TRlzol 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). Three biological replicates were performed for each sample. 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 Aglient High Sensitivitty 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 was 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 (Kozomara and Sam, 2014; Friedländer et al. 2012). 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 (Wen et al. 2012; Friedländer et al. 2012). Particularly, the miFam.dat (http://www.mirbase.org/ftp.shtml) was used to screen families of known miRNAs (Liu et al. 2019). 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 Normalized expression = Mapped read count/Total reads ×1000000) (Zhou et al. 2010). 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 default (Liu et al. 2019).

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 psRobot (Wu et al. 2012). Target gene candidates of differentially expressed miRNAs were performed through Gene Ontology (GO) enrichment analysis implemented with GOseq based Wallenius non-central hyper-geometric distribution (Young et al. 2012). Moreover, GO terms were also submitted to enrichment analysis by agriGO tool (Candar-Cakir et al. 2016). 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 pathways (Mao et al. 2005).

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, 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 $2^{-\Delta\Delta CT}$ method. In particular, three technical and three biological replicates were carried out. The real-time PCR reaction was conducted at 95˚C for 5 min, followed by 40 cycles (incubations at 95˚C for 15 sec, and 60˚C for 31 sec). The dissociation curve was analyzed at 55˚C

476 -95˚C.

477 Data Availability

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

481 Conflict of interest

482 There are no competing interests in relation to papers accepted for publication. There
483 are also no actual or potential conflicts of interest including any financial, personal or
484 other relationships with other people or organizations.

485 Acknowledgment

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

491 Author contributions

492 Fang B, and Huang Z designed the expriments and prepared the draft manuscript. Yu J,
493 Zhang J, and Dong H collected plant materials. Sun Y and Zhang W performed the
494 experiments and data analysis. Wang SZ took charge of the whole research.

495 References

496 Abel, Y., Charron, C., Virciglio, C., Bourguignon-Igel, V., Quinternet, M., Chagot, M.E.,
497 Robert, M.C., Verheggen, C., Branlant, C., Bertrand, E., Manival, X., Charpentier,
498 B. & Rederstorff, M. The interaction between RPAP3 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).

Allahverdy, J. & Rashidi, N. MicroRNAs induced by Listeria monocytogenes and their role in cells. *Microb Pathog.* **175**, 105997 (2023).

Bhattacharyya, D. *Rhododendron* species and their uses with special reference to Himalayas—a review. *Assam University J. Sci. Tech.* **7**, 161–167 (2011).

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).

Cho, J. & Paszkowski, J. Regulation of rice root development by a retrotransposon acting as a microRNA sponge. *Elife*. **6**, e30038 (2017).

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).

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).

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).

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).

Kozomara, A. & Sam, G.J. MiRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.* **42**, D68–D73 (2014).

525 Li, H., Dong, Y., Yin, H., Wang, N., Yang, L., Liu, X., Wang, Y., Wu, J. & Li, X.
526 Characterizaiton of the stress associated microRNAs in *Glycine max* by deep
527 sequencing. *BMC Plant Biol.* **11**, 170 (2011).
528 Liang, R., Han, B., Li, Q., Yuan, Y., Li, J. & Sun D. Using RNA sequencing to identify
529 putative competing endogenous RNAs (ceRNAs) potentially regulating fat
530 metabolism in bovine liver. *Sci. Rep.* **7**, 6396 (2017).
531 Liu, S., Ju, J. & Xia, G. Identification of the flavonoid 3'-hydroxylase and flavonoid
532 3',5'-hydroxylase genes from *Antarctic moss* and their regulation during abiotic
533 stress. *Gene* **543(1)**, 145–152 (2014).
534 Liu, X.X., Luo, X.F., Luo, K.X., Liu, Y.L. & Pan, T. Small RNA sequencing reveals
535 dynamic microrna expression of important nutrient metabolism during
536 development of *Camellia oleifera* fruit. *Int. J. Biol. Sci.* **15(2)**, 416–429 (2019).
537 Ma, Z., Hu, X., Cai, W., Huang, W., Zhou, X., Luo, Q., Yang, H., Wang, J. & Huang, J.
538 Arabidopsis miR171-targeted scarecrow-like proteins bind to GT cis-elements and
539 mediate gibberellin-regulated chlorophyll biosynthesis under light conditions. *PLoS*
540 *Genet.* **10(8)**, e1004519 (2014).
541 Mao, X., Cai, T., Olyarchuk, J.G. & Wei, L. Automated genome annotation and pathway
542 identification using the KEGG Orthology (KO) as a controlled vocabulary.
543 *Bioinformatics.* **21**, 3787–3793 (2005).
544 McCormick, K.P., Willmann, M.R. & Meyers, B.C. Experimental design, preprocessing,
545 normalization and differential expression analysis of small RNA sequencing
546 experiments. *Silence* **2(1)**, 2 (2011).
547 Meyers, B.C., Axtell, M.J., Bartel, B., Bartel, D.P., Baulcombe, D., Bowman, J.L., Cao,
548 X., Carrington, J.C., Chen, X., Green, P.J., Griffiths-Jones, S., Jacobsen, S.E.,
549 Mallory, A.C., Martienssen, R.A., Poethig, R.S., Qi, Y., Vaucheret, H., Voinnet, O.,
550 Watanabe, Y., Weigel, D. & Zhu, J.K. Criteria for annotation of plant MicroRNAs.
551 *Plant Cell*. **20**, 3186–3190 (2008).
552 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).

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. Nati. Acad. Sci. USA* **114**, 7166–7171 (2017).

Seitz, C., Ameres, S., Schlangen, K., Forkmann, G. & Halbwirth, H. Multiple evolution of flavonoid 3',5'-hydroxylase. *Planta.* **242(3)**, 561–573 (2015).

Sulas, P., Di Tommaso, L., Novello, C., Rizzo, F., Rinaldi, A., Weisz, A., Columbano, A., & Roncalli, M. A large set of miRNAs is dysregulated since the earliest steps of human hepatocellular carcinoma development. *Am. J. Pathol.* **188**, 785–794 (2017).

Sun, X., Xue, B., Jones, W.T., Rikkerink, E., Dunker, A.K. & Uversky, V.N. 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).

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., Demura, M., Kometani, M., Karashima, S., Yoneda, T. & Takeda, Y. Molecular and epigenetic control of aldosterone synthase, CYP11B2 and 11-Hydroxylase, CYP11B1. *Int. J. Mol. Sci.* **24(6)**, 5782 (2023).

Thomson, B. & Wellmer, F. Molecular regulation of flower development. *Curr. Top. Dev. Biol.* **131**, 185–210 (2019).

Tymon, L.S., Bradshaw, M., Götz, M., Braun, U., Peever, T.L. & Edmonds, R.L. Phylogeny and taxonomy of Erysiphe spp. on Rhododendron, with a special emphasis on North American species. *Mycologia*. **114(5)**, 887–899 (2022).

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).

Wang, S., Li, Z., Guo, X., Fang, Y., Xiang, J. & Jin, W. Comparative analysis of microsatellite, SNP, and InDel markers in four *Rhododendron* species based on RNA-seq. *Breed. Sci.* **68**, 536–544 (2018).

Wang, Y., Zhang, J., Cui, W., Guan, C., Mao, W. & Zhang, Z. Improvement in fruit quality by overexpressing miR399a in woodland strawberry. *J. Agric. Food Chem.* **65**, 7361–7170 (2017).

Wen, M., Shen, Y., Shi, S. & Tang, T. MiREvo: an integrative microRNA evolutionary analysis platform for next-generation sequencing experiments. *BMC Bioinformatics.* **13**, 140 (2012).

Wójcik, A.M. & Gaj, M. MiR393 contributes to the embryogenic transition induced in vitro in Arabidopsis via the modification of the tissue sensitivity to auxin treatment. *Planta*. **244**, 231–243 (2016).

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).

Xing, H., Li, Y., Ren, Y., Zhao, Y., Wu, X. & Li, H.L. Genome-wide investigation of micrornas and expression profiles during rhizome development in ginger (Zingiber officinale roscoe). *BMC Genomics*. **23(1)**, 49 (2022).

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).

Young, M.D., Wakeeld, M.J., Smyth, G.K. & Oshlack, A. Goseq: Gene Ontology testing for RNA-seq datasets. Alicia Oshlack (2012).

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).

Zhong, T., Hu, J., Xiao, P., Zhan, S., Wang, L., Guo, J., Li, L., Zhang, H. & Niu, L. Identification and characterization of microRNAs in the goat (*Capra hircus*) rumen during embryonic development. *Front Genet.* **8**, 163 (2017).

609 Zhou, L., Chen, J., Li, Z., Li, X., Hu, X., Huang, Y., Zhao, X., Liang, C., Wang, Y., Sun,
610 L., Shi, M., Xu, X., Shen, F., Chen, M., Han, Z., Peng, Z., Zhai, Q., Chen, J.,
611 Zhang, Z., Yang, R., Ye, J., Guan, Z., Yang, H., Gui, Y., Wang, J., Cai, Z. & Zhang,
612 X. Integrated profiling of microRNAs and mRNAs: microRNAs located on Xq27.3
613 associate with clear cell renal cell carcinoma. *PLoS One*. **5(12)**, e15224 (2010).
614
615