



OPEN ACCESS

EDITED BY

Guo-Fei Tan,
Guizhou Academy of Agricultural Sciences
(CAAS), China

REVIEWED BY

Guo Jian,
Yangzhou University, China
Zhenhui Wang,
Jilin Agriculture University, China

*CORRESPONDENCE

Lin Ruan

✉ alinche@126.com

Seping Dai

✉ daiseping@126.com

Yixun Yu

✉ yuyixun@scau.edu.cn

RECEIVED 01 May 2023

ACCEPTED 06 July 2023

PUBLISHED 27 July 2023

CITATION

Li H, Wang W, Liu R, Tong B, Dai X, Lu Y,
Yu Y, Dai S and Ruan L (2023) Long non-
coding RNA-mediated competing
endogenous RNA regulatory network
during flower development and color
formation in *Melastoma candidum*.
Front. Plant Sci. 14:1215044.
doi: 10.3389/fpls.2023.1215044

COPYRIGHT

© 2023 Li, Wang, Liu, Tong, Dai, Lu, Yu, Dai
and Ruan. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Long non-coding RNA-mediated competing endogenous RNA regulatory network during flower development and color formation in *Melastoma candidum*

Hui Li^{1,2}, Wei Wang¹, Rui Liu³, Botong Tong^{3,4}, Xinren Dai³,
Yan Lu⁵, Yixun Yu^{2*}, Seping Dai^{1*} and Lin Ruan^{1*}

¹Department of Botany, Guangzhou Institute of Forestry and Landscape Architecture, Guangzhou, China, ²College of Forestry and Landscape Architecture, South China Agricultural University, Guangzhou, China, ³State Key Laboratory of Tree Genetics and Breeding, Chinese Academy of Forestry, Beijing, China, ⁴State Key Laboratory of Tree Genetics and Breeding, Northeast Forestry University, Harbin, China, ⁵Jiangsu Key Laboratory for the Research and Utilization of Plant Resources, Institute of Botany, Chinese Academy of Sciences, Nanjing, Jiangsu, China

M. candidum, an evergreen shrubby flower known for its superior adaptation ability in South China, has gained increased attention in garden applications. However, scant attention has been paid to its flower development and color formation process at the non-coding RNA level. To fill this gap, we conducted a comprehensive analysis based on long non-coding RNA sequencing (lncRNA-seq), RNA-seq, small RNA sequencing (sRNA-seq), and widely targeted metabolome detection of three different flower developmental stages of *M. candidum*. After differentially expressed lncRNAs (DELncRNAs), differentially expressed mRNAs (DEmRNAs), differentially expressed microRNAs (DEmiRNAs), and differentially synthesized metabolites (DSmets) analyses between the different flower developmental stages, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) were conducted to identify some key genes and metabolites in flavonoid, flavone, anthocyanin, carotenoid, and alkaloid-related GO terms and biosynthetic pathways. Three direct-acting models, including antisense-acting, cis-acting, and trans-acting between lncRNAs and mRNAs, were detected to illustrate the direct function of lncRNAs on target genes during flower development and color formation. Based on the competitive endogenous RNA (ceRNA) regulatory theory, we constructed a lncRNA-mediated regulatory network composed of DELncRNAs, DEmiRNAs, DEmRNAs, and DSmets to elucidate the indirect role of lncRNAs in the flower development and color formation of *M. candidum*. By utilizing correlation analyses between DERNAs and DSmets within the ceRNA regulatory network,

alongside verification trials of the ceRNA regulatory mechanism, the study successfully illustrated the significance of lncRNAs in flower development and color formation process. This research provides a foundation for improving and regulating flower color at the lncRNA level in *M. candidum*, and sheds light on the potential applications of non-coding RNA in studies of flower development.

KEYWORDS

lncRNA, ceRNA regulatory mechanism, flower development, flower color formation, metabolites, *Melastoma candidum*

1 Introduction

lncRNAs are an influential class of molecules with a length of more than 200 nt (Wang and Chang, 2011). They originate from exonic, intronic, intragenic, and intergenic promoter regions, as well as 3' and 5' UTR enhancer sequences, and are transcribed in either a sense or antisense direction (Zhang et al., 2013). Traditionally, lncRNAs were thought to be meaningless molecules. In recent years, research has proven that lncRNAs regulate many biological processes in organisms such as gene expression adjustment (Guttman et al., 2009), post-transcription, post-translation (Crick et al., 1961), and chromosome modification (Bertone et al., 2004; Yanofsky, 2007) by combining with corresponding proteins. Unlike protein-coding genes, most lncRNAs are usually expressed at low levels and lack strong sequence conservation between species (Cabili et al., 2011; Necsulea et al., 2014). A number of studies have provided evidence that lncRNAs play a significant role in response to stress (Wunderlich et al., 2014), male sterility (Ding et al., 2012), phosphate homeostasis (Franco-Zorrilla et al., 2007), flowering time regulation (Heo and Sung, 2011), and flower and pollen development (Liu et al., 2010; Kang and Liu, 2015) in plants.

An increasing body of evidence indicating that lncRNAs may affect gene expression by either *cis*-acting on their chromosomes or *trans*-acting protein-encoding genes to carry out their functions (Wu et al., 2019). It is possible that long non-coding RNAs located upstream of a gene may associate with the promoter or other *cis*-acting elements of co-expressed genes in order to regulate gene expression at the transcriptional or post-transcriptional level. A large number of long non-coding RNAs could overlap with transcription factor binding sites, potentially preventing transcription factors from binding to the corresponding sites (de la Fuente, 2010; Chen, 2016). In some cases, lncRNAs are found to enhance the binding rate of transcription factors to nearby binding sites (de la Fuente, 2010; Yang et al., 2019). A study by Li et al. (2022) found that *cis*-acting LNC_002115 regulates hickory female floral development by influencing both *PHO2* and *SVP*. Kang and Liu (2015) identified a significant number of lncRNAs from 35 different flower and fruit tissues of diploid strawberries. They inferred a *cis* or *trans*-acting relationship between lncRNAs and their targets based on correlation analysis of lncRNAs and their

target gene expression trend. In addition, they discovered that lncRNAs are not well conserved between species of plants.

As endogenous target mimics, lncRNAs can bind with miRNAs on mRNA response elements (MREs) and mitigate miRNAs' cleavage effect on target genes by sponge-like actions (Wu et al., 2019). Due to their similar structural characteristics, lncRNAs may also be negatively regulated by miRNAs through a similar mechanism to mRNAs (Chaofeng et al., 2013). In *Arabidopsis*, a ceRNA regulatory relationship showed that lncRNA IPS1 could influence the expression level of *PHO2* by binding to miR399 (Franco-Zorrilla et al., 2007). lncRNAs, *COOLAIR* and *COLDAIR*, could repress the *FLC* gene expression through an epigenetic silencing mechanism to regulate flowering time (Heo and Sung, 2011; Kim and Sung, 2017). Compared with the wild type, overexpression of lncRNA npc48 increased the rosette diameter and leaf serration and delayed flowering time (Ben Amor et al., 2009). Besides, ceRNA regulatory relations has also been demonstrated in other plants, such as cucumber (He et al., 2020), pepper (Zuo et al., 2019), rice (Xu et al., 2016), maize (Zhu et al., 2017), and tomato (Yang et al., 2019). In rice, lncRNA osa-eTM160 attenuated the repression of osa-miR160 on *ARF18* throughout the embryonic anther by target mimicry ways (Wang et al., 2017). Fang et al. (2019) found that a rice lncRNA Ef-dc transcribed from the antisense strand of the flowering activator *SOCI* locus can positively regulate the expression of *SOCI* and balance yield with maturity duration. Wang et al. (2017) found that lncRNA osa-eTM160 could attenuate the repression of osa-miR160 on osa-*ARF18* during the early developmental stage of the anther. A genome-wide association study (GWAS) analysis and ceRNA network by Xu et al. (2021) identified the key regulatory mechanism of the LTCONS_00034157- miRNA167h- PsTPS1 in the later flowering process of *Prunus sibirica*.

Flower color is an important trait that determines the ornamental quality and landscaping application value (Zhu et al., 2019b). It is the result of pigment metabolite accumulation in the vacuoles of flower epidermal cells (Mol et al., 1998; Deng et al., 2013). Flower color is predominantly due to the production of flavonoids carotenoids or betalains. (Holton, 1995). Betalains, one type of alkaloid, are water-soluble nitrogenous pigments derived from the amino acid L-tyrosine, mainly classified as red-violet betacyanins and yellow betaxanthins. In Caryophyllales, betalains

exclusively replace anthocyanin (Tanaka et al., 1998; Deng et al., 2013) to attract pollinators and seed dispersers (Sunnadeniya et al., 2016). Some researchers have demonstrated that the relevant enzymes for the production of anthocyanins are not expressed in betalains-producing plants at the biochemical level (Shimada et al., 2007; Brockington et al., 2011). However, enzymes, genes and biosynthetic pathways involved in betalain production are much less well-studied than those of flavonoids and carotenoids.

The change in color both within flowers and in isolated pigments involves a range of biochemical mechanisms. Some of the factors influencing color are temperature, co-pigments, pH, metals, sugars, anthocyanin stacking, and cell shape (Bowles et al., 2006; Shoji et al., 2007). The earliest research suggested that pH, metal-complex theory, and Metalloanthocyanins could be the key factors to determine flower color. In 1913, Willstätter and Everest proposed the pH theory based on the observation of a pigment from blue cornflowers and rose, cyanin, which could display red color under acidic media and blue color under alkaline solutions (Willstätter and Everest, 1913). The metal complex theory states that anthocyanin can form complexation with metal ions such as Mg^{2+} (Mitsui et al., 1959), Fe^{3+} , Al^{3+} (Bayer et al., 1966), Ga^{3+} , In^{3+} , Co^{3+} , Mn^{2+} , Zn^{2+} and Cd^{2+} to show different colors in plants flower (Kondo et al., 1998). Metalloanthocyanins theory holds that anthocyanins, flavones, and metal ions fix at 6: 6: 2 in blue flowers (Takeda, 2006). With biological development, more and more flower color can be explained by genes or metabolite levels. Nowadays, scientists believe that flower color is predominantly controlled by the production of flavonoids, carotenoids, and alkaloid-related compounds, such as betalains (Holton, 1995). For example, researchers have found that yellow flower petals often contain yellowish xanthophylls, β -carotenoids, and chrysanthemum. All of them belong to the class of carotenoids (Nielsen et al., 2003; Kishimoto et al., 2004). Some species such as roses, and carnations are lacking blue because of the absence of Flavonoid3',5'-hydroxylase, a key enzyme catalyzing the hydroxylation reaction between dihydrokaempferol and dihydromyricetin (Holton and Cornish, 1995; Kondo et al., 1998).

M. candidum belongs to the Melastomataceae family which is mainly centered in Southeast Asia and extends to India, South China, and Northern Australia (Liu et al., 2014). Although many species of Melastomataceae have a relatively high degree of overlap in geographic distributions and flower periods, some members still face reproductive isolation problems. Additionally, many members of the Melastomataceae are dull colors, which limits their use in garden industries. Here, we took *M. candidum* as a research object and conducted lncRNA-seq, sRNA-seq, mRNA-seq, and a widely targeted metabolome for three development stages of the flowers. Based on differentially expressed RNA analyses, we developed three functional acting models, including antisense, *cis*, and *trans* - acting models. Two networks of antisense and *cis* models were constructed in flavonoid, anthocyanin, carotenoid, and alkaloid-related pathways. According to the ceRNA theory, a lncRNA-mediated regulatory network was also built in the aforementioned pathways. A correlation among all kinds of RNAs and metabolites was conducted to illustrate the relationship among them. This research aims to unveil the role of lncRNA in flower development in *M. candidum*.

2 Materials and methods

2.1 RNA extraction, cDNA library construction, and sequencing

Three developmental stages of the flower of *M. candidum* including closed buds with white petals (McI), closed buds with pink petals (McII), and opened buds with pink petals (McIII) were collected and immediately frozen in liquid nitrogen. Three replicates of each stage sample were taken from three seedlings. Total RNA was extracted by using an OminiPlant RNA Kit (DNase I) (CW2598, CWBIO, Taizhou, China) according to the manufacturer's protocol. An Agilent 2100 Bioanalyzer was utilized to assess the quality of the RNA (Agilent Technologies, Palo Alto, CA, USA). To generate a cDNA library, we fragmented the mRNA using a fragmentation buffer and reverse-transcribed the resulting small fragments into cDNA using random primers. The second strand cDNA was synthesized by employing DNA polymerase I, RNaseH, dNTP, and buffer. After purification with the poly(A) and PCR extraction kit (Qiagen), the synthesized products were further purified using the QiaQuick PCR extraction kit (Qiagen, Venlo, The Netherlands) and then ligated to Illumina sequencing adapters. Afterward, the second-strand cDNA was digested using the enzyme UNG (Uracil-N-Glycosylase), and the resulting products were separated by size on an agarose gel before being amplified by PCR. The PCR products were then sequenced on Illumina HiSeq TM 4000 platforms.

2.2 Filter of the raw data and alignment against the genome

After obtaining sequence data, raw reads consisting of adapters or low-quality bases were filtered by the fastp software (version 0.18.0) (Chen et al., 2018b) with steps: 1) removing reads containing adapters; 2) removing reads containing more than 10% of unknown nucleotides (N); 3) removing low quality reads containing more than 50% of low quality (Q-value \leq 20) bases. After filtering, the clean short reads were aligned to the ribosome RNA (rRNA) database using Bowtie2 (version 2.2.8) (Langmead and Salzberg, 2012) to eliminate the rRNA-mapped reads. After building an index of the *M. candidum* genome, the paired-end clean reads were mapped to the reference genome by using HISAT2 (version 2.1.0) (Kim et al., 2015) with a default parameter "-rna-strandedness RF".

2.3 Transcripts reconstruction and annotation

The reconstruction of transcripts was carried out with the software Stringtie (version 1.3.4) (Pertea et al., 2015). To identify the novel transcripts among the reconstructed transcripts, all the reconstructed transcripts were aligned to the reference genome and were then categorized into 12 categories using the Cuffcompare program (Trapnell et al., 2010). A novel transcript was defined if it had one of the following class codes - u, i, j, x, c, e, or o. To further

identify new genes, we used the following parameters: length of transcript > 200 bp, number of exons > 1 (Lu et al., 2019). All the novel transcripts were then aligned to the Nr, KEGG, and GO databases to obtain protein functional annotations.

2.4 lncRNA prediction and classification

Software including Coding-Non-Coding Index (CNCI, version2) (Sun et al., 2013) and Coding Potential Calculator 2 (CPC2, version 0.9-r2) (Kong et al., 2007) (<http://cpc.cbi.pku.edu.cn/>) were used to assess the protein-coding potential of novel transcripts. Only transcripts which meet the protein-coding-score criteria (CNCI sequence-score < 0 and CPC2 coding probability < 0.5) were considered as lncRNAs. Based on the location relative to protein-encoding genes (PCG), lncRNAs are categorized into five classes: intergenic lncRNAs (located between two PCGs), bidirectional lncRNAs (situated on the opposite strand but within 1 kb of the promoter on the sense strand), intronic lncRNAs (located within an intron of a PCG on the sense strand), antisense lncRNAs (transcribed from the opposite strand of a PCG), and sense lncRNA (spanning multiple introns or exons within a PCG) (Ahmad et al., 2021).

2.5 Quantification of transcripts and differentially expressed analysis

The abundance of the transcript was quantified using StringTie (Pertea et al., 2015). An FPKM (fragment per kilobase of transcript per million mapped reads) value was calculated for each transcription region to mitigate the effect of varying transcript lengths and sequencing data amounts on expression. DEGs of coding RNAs and lncRNAs were analyzed separately by using DESeq2 software (Li and Dewey, 2011) between two versus groups by the following steps: 1) normalization of read counts; 2) calculation of *p*-value; 3) correction of value to get a false discovery rate (FDR) value. Differentially expressed lncRNAs were screened with the threshold: Fold change ≥ 2 , FDR ≤ 0.05 .

2.6 Differential analyses of miRNAs, mRNAs, and metabolites

For miRNAs, expression level was calculated and normalized by transcripts per million (TPM) methods. EdgeR (Robinson et al., 2010) were used to conducted differential analysis. DE miRNAs were identified with the absolute threshold \log_2 (fold change) ≥ 0.585 , *p*-value ≤ 0.05 . For mRNAs, HISAT2 (Kim et al., 2015) were used to align to the genome of *M. candidum* (<http://evolution.sysu.edu.cn/Sequences.html>). The FPKM value was used to quantify expression abundance and variations. DE mRNAs were identified by DESeq2 (Love et al., 2014) package with the absolute threshold \log_2 (fold change) ≥ 1 and *p*-value ≤ 0.05 . Metabolites were analyzed by using an LC-ESI-MS/MS system (UPLC, Shim-pack UFLC SHIMADZU CBM30A, <http://www.shimadzu.com.cn/>; and MS/MS (Applied Biosystems 6500 QTRAP, <http://www.appliedbiosystems.com.cn/>) (Chen et al., 2013) and variable importance in projection (VIP) score of (O)PLS model was applied to rank the metabolites and distinguish two versus groups. DSmetS were identified with the absolute threshold \log_2 (fold change) ≥ 0.585 , vip ≤ 1 .

www.appliedbiosystems.com.cn/) (Chen et al., 2013) and variable importance in projection (VIP) score of (O)PLS model was applied to rank the metabolites and distinguish two versus groups. DSmetS were identified with the absolute threshold \log_2 (fold change) ≥ 0.585 , vip ≤ 1 .

2.7 lncRNA-mRNA association analysis

In order to identify three direct-acting models of lncRNAs and mRNAs including antisense-regulation, *cis*-regulation, and trans-regulation, the software RNAplex (version 0.2) (Shen et al., 2014) (<http://www.tbi.univie.ac.at/RNA/RNAplex.1.html>) was used to predict the complementary correlation and detect the interaction between antisense lncRNAs and mRNAs. Antisense lncRNAs can form complementary base pairing with mRNAs. The ViennaRNA package (Lorenz et al., 2011) in R software was used to predict the best base pairing based on the calculation of minimum free energy through thermodynamics structure. *Cis* lncRNAs could regulate neighboring genes on the same allele. lncRNAs with the unknown region were annotated again to identify *cis*-regulators. Trans lncRNAs could regulate co-expressed genes far from them. The correlation of expression between lncRNAs and protein-coding genes was used to identify a trans-regulation relationship with a Pearson correlation of more than 0.999.

2.8 Validation of lncRNA-seq result by qRT-PCR

In order to verify the lncRNA-seq result, we chose 12 lncRNAs to perform qRT-PCR. The primers of lncRNAs used in qRT-PCR were listed in Table S1. 0.5 μ g total RNA of three stages flowers were reverse-transcribed into first-strand cDNA using the PrimeScript RT reagent Kit gDNA Eraser (Takara, Dalian, China). Then, the SYBR @Premix Ex Taq TMII (Takara, Dalian, China) was used according to the manufacturer's instruction for qRT-PCR of lncRNAs on the Illumina Eco real-time PCR system (Illumina, USA). The α -tubulin gene of *M. candidum* was used as the internal reference gene. Ct values were then calculated by the $2^{-\Delta\Delta Ct}$ algorithm. Primers sequences, length of PCR products and PCR amplification efficiency for each pair of primers are listed in Table S1. Relative expression levels of selected lncRNAs at different stages of flowers were analyzed using One-way ANOVA with multiple comparison by GraphPad Prism 9. The graphs were visualized by GraphPad Prism 9 and Adobe Illustrator 2020.

2.9 Construction and validation of lncRNA-mediated ceRNA network

The ceRNA network was constructed based on the following rules: 1) negative correlation between miRNA and lncRNAs, as well as miRNA and mRNAs with calculation of Spearman correlation coefficient (SCC); 2) positive correlation between lncRNAs and mRNAs with calculation of Spearman correlation coefficient (SCC);

3) enrich degree of lncRNA binding to miRNAs and miRNAs binding to mRNAs. Relationships among ceRNA members and metabolites were visualized by the ggalluvial package in R3.6.1 software.

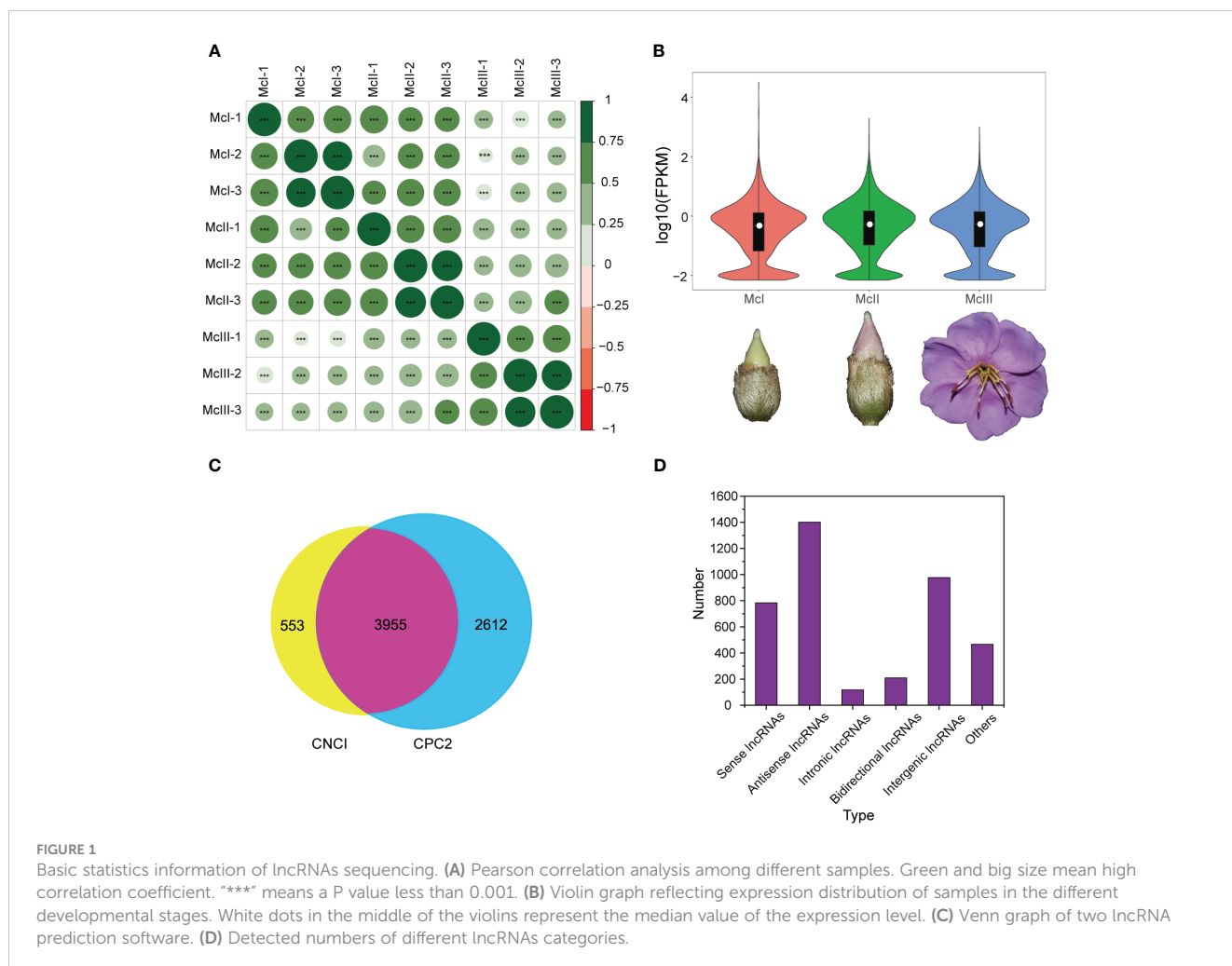
To verify the lncRNA-miRNA-mRNA regulatory chains, the transient coexpression experiments were performed in leaves of *Nicotiana benthamiana* according to Lu et al. (Lu et al., 2019). Precursors of miRNA (pre-miRNA), lncRNAs, and mRNAs were cloned from mixture of cDNA of three-stages flowers. For miRNAs that could not obtain pre-miRNAs in the cDNA, we performed overlap PCR to ligate these mature miRNAs into the vector. Pre-miRNAs and lncRNAs were ligated into pCAMBIA2300 at *Sma*I and *Xba*I sites. mRNAs were ligated into pCAMBIA1300 at *Xba*I and *Kpn*I sites. The vector harboring pre-miRNAs, lncRNAs, and mRNAs were then transformed into *Agrobacterium tumefaciens* GV3101 strain. Equal amounts of agrobacterial cell cultures containing miRNAs, lncRNAs, and mRNAs were mixed respectively. Then, the mixture was infiltrated into leaves of *N. benthamiana*. After being incubated in dark for two days, the infiltrated tobacco leaves were observed under laser scanning confocal microscopy (LSCM) with EGFP = 780, and collected for total RNA extractions (CW2598, CWBIO, Taizhou, China) and qRT-PCR (Takara, Dalian, China). Tobacco L23 was used as an

internal reference (Liu et al., 2012). The primers were listed in Table S1.

3 Results

3.1 Statistics information of sequenced lncRNAs

Prior to conducting differentially expressed analyses, we performed basic statistical analyses on the sequencing data, such as examining correlations between samples, evaluating expression distributions, and identifying types of lncRNAs. Sample correlation was determined based on the expression level of lncRNAs (Figure 1A). Correlations among samples of each stage had a relatively high coefficient, indicating these samples meet our analysis requirements. A relatively high average expression level was observed at the late stage (Figure 1B). In order to increase the reliability of the results, we used CNCI and CPC2 to identify lncRNAs in sequenced samples. A total of 4,508 lncRNAs were identified by CNCI and a total of 6,567 lncRNAs were identified by CPC2. To ensure greater accuracy, we took the intersection of the results from both software and obtained 3,955 lncRNAs



(Figure 1C). Five types of lncRNA were identified, including sense lncRNAs, antisense lncRNAs, intronic lncRNAs, bidirectional lncRNAs, and intergenic lncRNAs. It should be noted that some lncRNAs in the graph cannot be assigned a precise category. Among the identified lncRNA, the proportion of antisense lncRNAs was the highest, while that of intronic lncRNAs was the lowest (Figure 1D). Overall, the lncRNAs were suitable for further analysis.

3.2 Differentially expressed analyses of lncRNAs

For all lncRNAs identified, we performed differential expression analyses among different stages (Figure 2). There were 190 up-regulated and 228 down-regulated lncRNAs in the McII vs MCI group (Figure 2A; Table S2). In the comparison between the McIII and MCI group, we identified 552 up-regulated and 602 down-regulated lncRNAs (Figure 2B; Table S3). 426 lncRNAs were up-regulated and 454 lncRNAs were down-regulated in the McIII vs McII group (Figure 2C; Table S4). Differential expression analyses results

indicated huge differences of lncRNA expression levels between McIII and the other two stages. Then, we gave a statistics among the three comparison groups using Venn method which yielded 67 common lncRNAs (Figure 2D). As lncRNAs play a variety of roles during the different stages of development, we used the union of three comparison groups to conduct further analysis and identify the roles of lncRNAs during the different stages of flower development.

3.3 Antisense, *cis*, and trans-acting model of lncRNAs and functional annotation of their targeted mRNAs

It has been reported that lncRNAs interact with mRNAs in three distinct ways including *cis*, trans, and antisense interactions (Nie et al., 2012). *Cis* model means that lncRNAs could induce reconstruction and histone modification to influence the combination of transcript factors with promoter and enhancer or combine with transcription elements directly to affect the expression of the protein-encoding genes. In the antisense model, it is assumed that lncRNAs could directly bind to mRNAs to

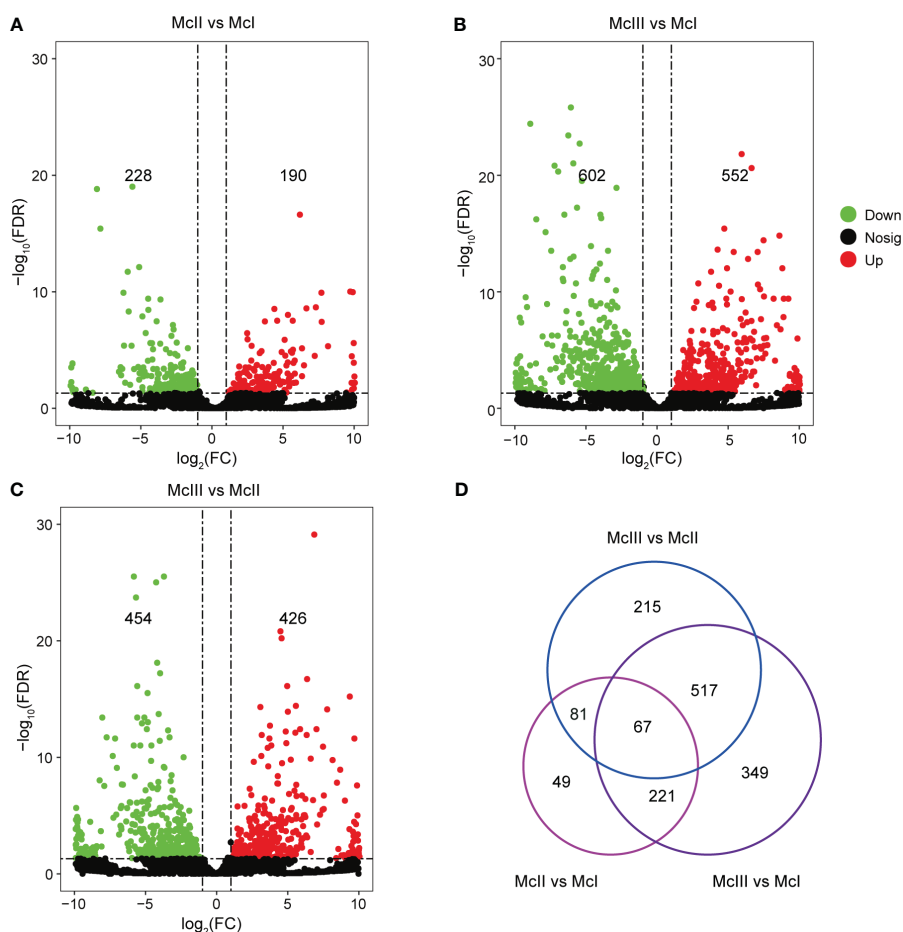


FIGURE 2

Differentially expressed lncRNAs (DElncRNAs) in different comparing groups. (A) DElncRNAs in stageII (McII) versus stageI (MCI) group. (B) DElncRNAs in stageIII (McIII) versus stageI (MCI) group. (C) DElncRNAs in stageIII (McIII) versus stageII (McII) group. Red dots mean up-regulated lncRNAs, and green dots mean down-regulated lncRNAs. (D) Venn graphs of different versus groups.

influence alternative splicing of mRNAs or to decay mRNAs into siRNAs, which is closely related to the stability of mRNAs. The trans model is often able to be judged by the opposite expression trends between lncRNAs and their mRNA counterparts. LncRNAs could sometimes bind to proteins translated by mRNAs, altering their activity, structure, or position to influence their corresponding mRNAs. Here we also confirmed three functional models for differentially expressed lncRNAs and then GO and KEGG analyses were performed for their corresponding mRNAs (Figures 3A, B; Tables S5-S10). In Figures 3A, B, some GO terms associated with flower development and color formation were chosen for visualization, including alkaloids, pigments, flavonoids,

hormones, and flower development-related terms or path. As a result of the three acting models, the trans model enriched more genes than the other two functional models. However, we could not find the antisense models of these DELncRNAs in the flavone biosynthetic process of the result of GO analysis as well as the isoflavonoid biosynthetic process, the folate biosynthesis, and the brassinosteroid biosynthetic process of the results of KEGG analysis (Figures 3A, B). Similarly, there was also lack of *cis* models of these DELncRNAs in the Tropane Alkaloid biosynthetic process and the Alkaloid metabolic process of the results of GO analysis, and the isoflavonoid biosynthesis and Anthocyanin biosynthesis of the results of KEGG analysis (Figures 3A, B).

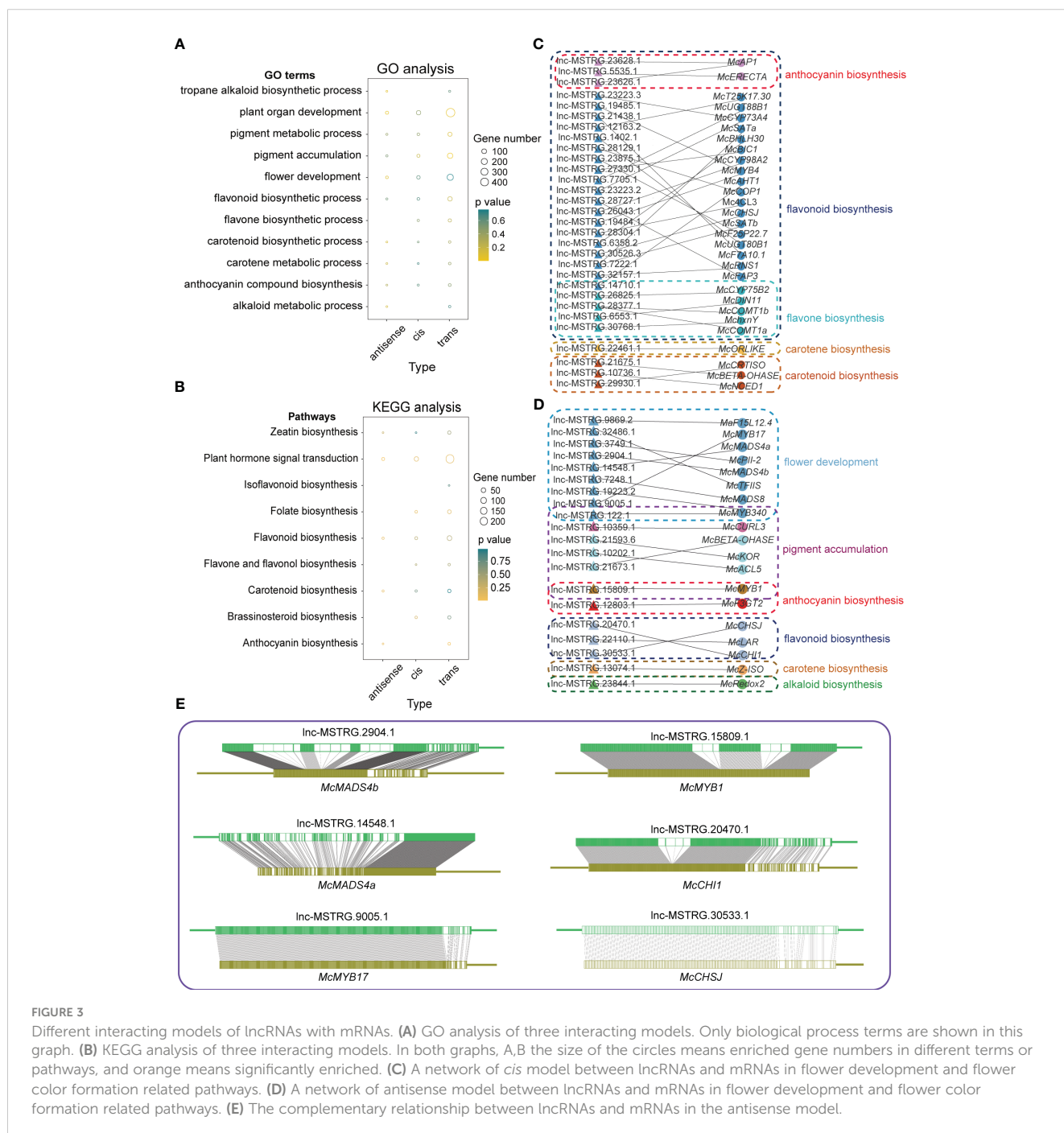


FIGURE 3 Different interacting models of lncRNAs with mRNAs. (A) GO analysis of three interacting models. Only biological process terms are shown in this graph. (B) KEGG analysis of three interacting models. In both graphs, A,B the size of the circles means enriched gene numbers in different terms or pathways, and orange means significantly enriched. (C) A network of *cis* model between lncRNAs and mRNAs in flower development and flower color formation related pathways. (D) A network of antisense model between lncRNAs and mRNAs in flower development and flower color formation related pathways. (E) The complementary relationship between lncRNAs and mRNAs in the antisense model.

To further investigate the potential roles of lncRNAs in regulating gene expression by antisense and *cis* models, we built lncRNAs-target genes networks (Figures 3C, D). In the *cis* model, we identified 26 lncRNAs that function on genes involved in flavonoid biosynthesis. Of these, three lncRNAs, lnc-MSTRG.23628.1, lnc-MSTRG.5535.1, and lnc-MSTRG.23626.1, were related to anthocyanin biosynthesis, while five lncRNAs, lnc-MSTRG.14710.1, lnc-MSTRG.26825.1, lnc-MSTRG.28377.1, lnc-MSTRG.6553.1, and lnc-MSTRG.30768.1 were related to flavone biosynthesis. In addition to flavonoid biosynthesis, we found one lncRNA in carotene biosynthesis and three lncRNAs in carotenoid biosynthesis (Figure 3C). In contrast to the *cis* model, we observed fewer lncRNAs involved in flavonoid biosynthesis in the antisense model, with only three lncRNAs, lnc-MSTRG.15809.1, lnc-MSTRG.12803.1, and lnc-MSTRG.30533.1, were detected. No lncRNAs involved in flavone biosynthesis process were identified (Figure 3D). Overall, our findings suggest that there are more *cis* models than antisense models involved in regulating flower development and color formation in *M. candidum*. The antisense model involves lncRNAs forming complementary relationships with genes. To investigate this further, we examined several genes, *McMADS4a*, *4b*, *McMYB1,17*, *McCHS1*, and *McCH11*, along with their corresponding lncRNAs to assess their complementary relationship. Our results showed a high degree of complementarity between the lncRNAs and their target genes (Figure 3E).

3.4 Validation and expression model of lncRNAs in three development stages of *M. candidum*

In order to confirm the accuracy of our lncRNA-seq data, we performed qRT-PCR assays on 12 selected lncRNAs. The expression patterns of most of these lncRNAs were found to be consistent with the results obtained from the sequencing analysis. Specifically, lnc-MSTRG.17619.1, lnc-MSTRG.29477.1, and lnc-MSTRG.34245.1 were highly expressed during the McI stage, while lnc-MSTRG.10215.1, lnc-MSTRG.11402.2, lnc-MSTRG.223785.1, lnc-MSTRG.24719.1, lnc-MSTRG.29558.1, and lnc-MSTRG.29930.1 were highly expressed during the McII stage (Figure 4A). Additionally, lnc-MSTRG.28008.2 and lnc-MSTRG.11415.1 were highly expressed during the McIII stage, while lnc-MSTRG.24133.1 was highly expressed in both the McII and McIII stages (Figure 4A). We found that the results obtained from the lncRNA-seq analysis and the qRT-PCR experiments were positively correlated, with a slope of 1.147 and an R value of 0.727 for the McII vs McI group, and a slope of 0.968 and an R value of 0.645 for the McIII vs McII group (as shown in Figure 4B). Based on these findings, we concluded that the lncRNA-seq analysis was valid and accurate.

We investigated the expression pattern of these lncRNAs during different stages of the flower development process. In the McII vs McI group, a total of six DElncRNAs were confirmed, of which four were down-regulated and two were up-regulated (Figures 5A, D). Ten DElncRNAs were differentially expressed between the McIII and McI groups, of which three were downregulated and seven were upregulated (Figures 5B, E). Five DElncRNAs related to flower color formation were found between McIII and McII, including one that was down-regulated and four that were up-regulated (Figures 5C, F).

During the transition from the second to the third stage of flower development, lnc-MSTRG.10215.1 and lnc-MSTRG.24133.1 were up-regulated and exhibited an upward trend in expression, indicating their corresponding competing endogenous mRNAs were key regulator in regulating this period of development. Conversely, lnc-MSTRG.34245.1, lnc-MSTRG.29477.1, and lnc-MSTRG.17619.1 showed a decline in expression from the second stage onwards (Figures 5A, B), indicating their corresponding competing endogenous mRNAs were less involved in this process. It was observed that lnc-MSTRG.11415.1 and lnc-MSTRG.29558.1 began to be highly expressed in the third period (McIII), suggesting that these two lncRNAs will play a critical role in flower development after color formation (Figures 5B, C, E, F). There is only a low expression of lnc-MSTRG.11402.2 in the second stage, showing that its target gene plays a negative role in flower coloration process.

3.5 lncRNAs-mediated ceRNA regulatory network build and verification

To further clarify the regulatory mechanism of the aforementioned lncRNAs, we first examined the supplementary relation between lncRNAs and miRNAs, miRNAs and target genes (Figure S1). Subsequently, we constructed a ceRNA regulatory network, which included lncRNAs, miRNAs, mRNAs, and metabolites for flower formation-related pathways (Figure 6A). As shown in the graph, lnc-MSTRG.10215.1 and lnc-MSTRG.11402.2 were found to regulate *McF3H* through interacting with miR5207, ultimately influencing dihydromyricetin (DHM), dihydroquercetin (DHQ), and pinobanksin (PBA). Dihydroquercetin (DHQ), also known as 3, 5, 7, 3, 4 - pentahydroxy flavanone or taxifolin, is a bioactive flavonoid that is considered one of the rarest and most effective natural antioxidants (Yu et al., 2021). Both dihydroquercetin and dihydromyricetin are colorless dihydroflavones that can be reduced to leucoanthocyanidins under the action of the dihydroflavonol-4-reductase (DFR) enzyme. Subsequently, under the catalysis of downstream enzymes, the leucoanthocyanidins undergo a process of the transformation into orange pelargonidin, reddish-purple cyanidin, and violet-blue delphinidin (Tanaka and Brugliera, 2013; Liu et al., 2019). Pinobanksin, one of the most common phenolic constituents of pine heartwood, was first isolated from *Pinus Banksiana* (Erdtman, 1944). When combined with zinc ion, this compound is specific to 3-hydroxyflavanones to produce much deeper colors during the reducing process. Pinobanksin usually produces orange-red when reduced with magnesium or zinc and hydrochloric acid (Lindstedt, 1950). *p*-coumaroyl Shikimic acid is a by-product of the anthocyanin pathway, derived from 4-coumaroyl-CoA by the formation of an ester bond with shikimic acid catalyzed by Shikimate O-hydroxycinnamoyl transferase (HCT) (Paliyath et al., 2009; Li et al., 2021). In the presence of the enzymes C4H1 and C3H3, this compound can be catalyzed again to transform into caffeoyl shikimic acid (Chen et al., 2011). Both *p*-coumaroyl Shikimic acid and caffeoyl shikimic acid are key compounds in Phenylpropanoid biosynthesis during lignin formation. The Sankey graph showed that miR3704, miR480, and miR1220 had cleavage effects on *McHST*, *McAHT1*, and *McPHT3*, respectively. Additionally, we discovered that lnc-MSTRG.24133.1

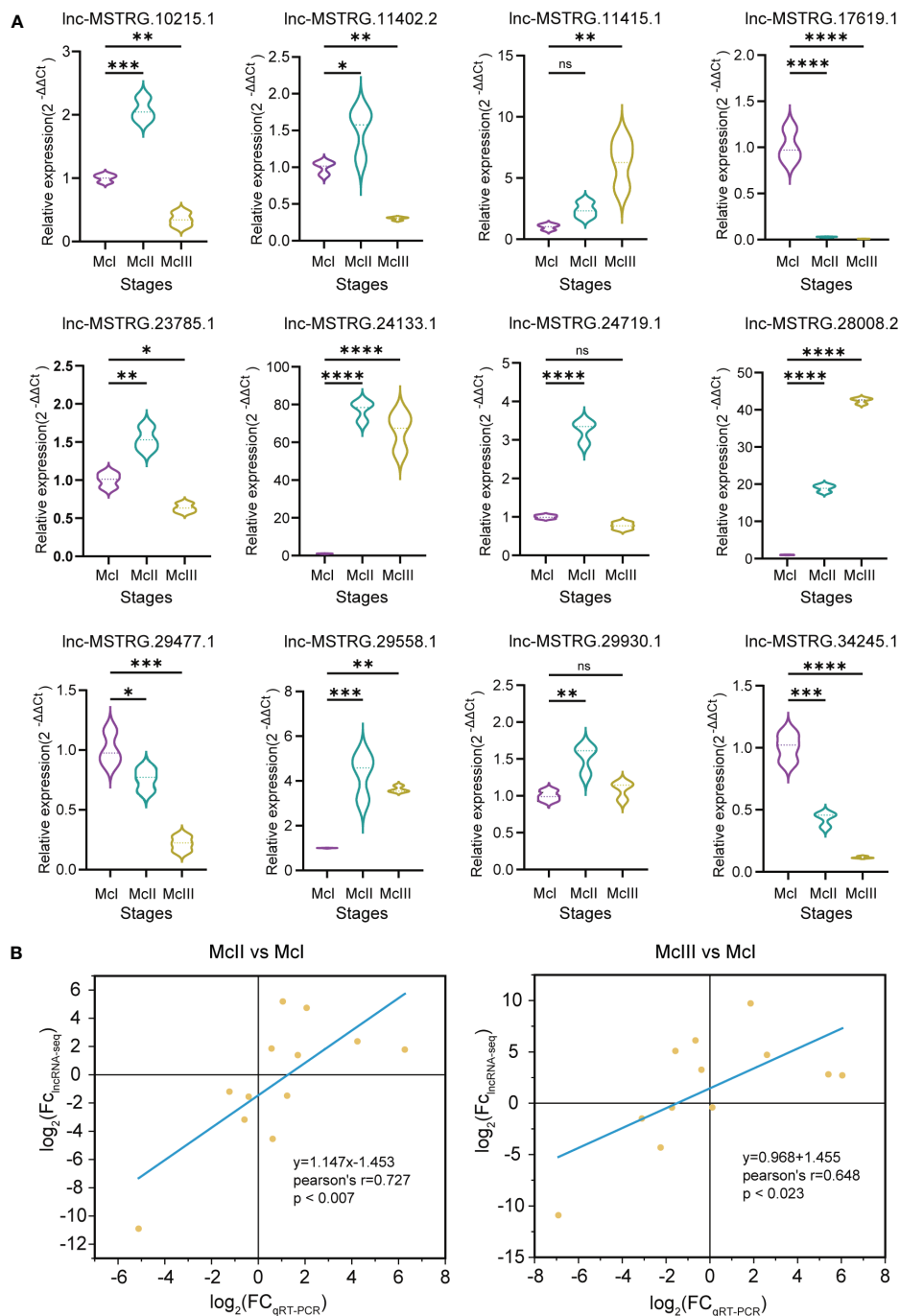


FIGURE 4 Validation of lncRNA results. **(A)** qRT-PCR results of 12 selected lncRNAs. Three colors represent three stages. Dot lines within the violin graphs mean median values. "ns" means $p > 0.05$; "**" means $p \leq 0.05$; "***" means $p \leq 0.01$; "****" means $p \leq 0.001$; "*****" means $p \leq 0.0001$. **(B)** Fitting result between qRT-PCR and lncRNA-seq. In both qRT-PCR and miRNA-seq results, fold change (FC) was normalized by \log_2 algorithm.

could adsorb miR3704 to maintain a balance between miRNA and target gene. Three lncRNAs, lnc-MSTRG.17619.1, lnc-MSTRG.29477.5, and lnc-MSTRG.34245.1, were found to bind to miR480 while another set of three lncRNAs, lnc-MSTRG.23785.1, lnc-MSTRG.29558.1, and lnc-MSTRG.29930.1 were found to bind to

miR1220. These results indicate that the three lncRNAs that ceRNA regulatory mechanism play an important role in the formation of *p*-coumaroyl Shikimic acid, which ultimately determine flower color. Three specific genes, namely *McUGT88B1*, *McUGT88F3*, and *McCYP98A2*, were found to have no associated with metabolites.

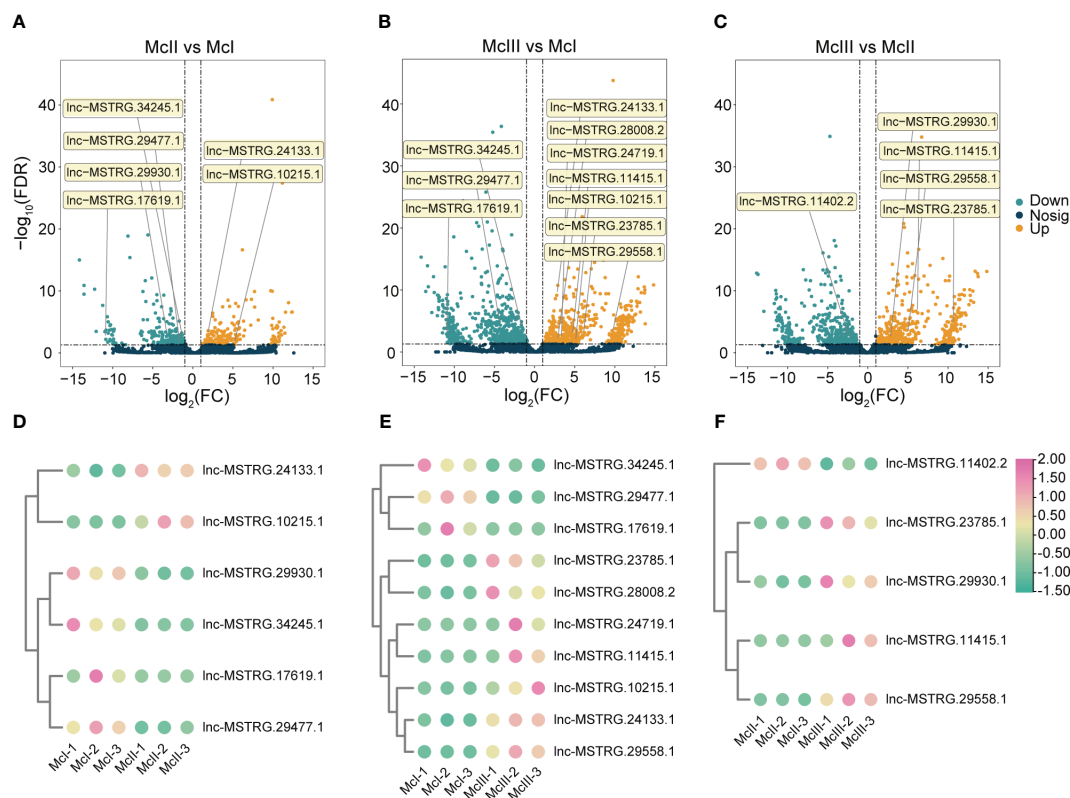


FIGURE 5

Distribution and expression heatmaps of selected lncRNAs. (A) Selected lncRNAs in McII vs Mcl group. (B) Selected lncRNAs in McIII vs Mcl group. (C) Selected lncRNAs in McIII vs McII group. In (A–C), orange means up-regulated lncRNAs, and green means down-regulated lncRNAs. (D) Heatmap of selected lncRNAs in McII vs Mcl group. (E) Heatmap of selected lncRNAs in McIII vs Mcl group. (F) Heatmap of selected lncRNAs in McIII vs McII group. In graphs (D–F), pink dots mean high expression level, and the light green dot means low expression level.

This may indicate that the three genes related to ceRNAs do not significantly affect metabolism. It is also possible that annotation of the metabolite database is incomplete.

We conducted a correlation analysis to determine the relationship between ceRNA members and metabolites in specific pathways (Figure 6B). Each ceRNA regulatory routine showed a negative correlation between miRNAs and both lncRNAs and mRNAs, whereas a positive correlation was observed between lncRNAs and mRNAs, indicating the presence of a true competitive mechanism between lncRNAs and mRNAs. lnc-MSTRG.28008.2 and *McASP2* were negatively correlated with L-phenylalanine, however, miR8131-3p was positively correlated with L-phenylalanine, indicating that this ceRNA regulatory chain is responsible for L-phenylalanine synthesis (Figure 6B). It is known that both the flavonoid and anthocyanin pathways originate from the amino acid L-phenylalanine, which is deaminated by the action of phenylalanine ammonia-lyase (PAL) to produce trans-cinnamic acid and ammonia (Fasoula et al., 1995). lnc-MSTRG.29477.1, lnc-MSTRG.29558.1, *McAHT1*, and *McPHT3* were positively correlated with *p*-coumaroyl Shikimic acid, whereas, miR480 and miR1220 were negatively correlated with *p*-coumaroyl Shikimic acid, indicating that these two ceRNAs regulatory chains were key regulators in the synthesis of *p*-coumaroyl Shikimic acid. However,

the rest of ceRNA regulatory networks did not display obvious putative correlations, which could be attributed to more complex regulatory mechanisms like multiple ceRNA chains, histone modification, RNA methylation modification process, etc. Further studies are needed to gain a better understanding of these complex regulatory mechanisms.

To further validate the identified lncRNA-mediated ceRNA regulatory network, we employed pCAMBIA1300 harboring target genes, and pCAMBIA2300 harboring lncRNAs and miRNAs to perform tobacco injection and qRT-PCR experiment. Ordinarily, miRNAs have cleavage effects on the target genes, when lncRNAs bind with miRNAs, the cleavage effects will be weakened. We selected 7 lncRNA-mediated ceRNA regulatory routines which included 6 genes, 5 miRNAs, and 6 lncRNAs to validate the ceRNA regulatory mechanism in tobacco (Figure 7). In accordance with our anticipation, when only injecting the miRNAs and target genes into leaves of tobacco, the fluorescence signal became weak (Figure 7), and the expression level of target genes began to decrease. When we added the lncRNAs in corresponding ceRNA chains, the fluorescence signal was regained, and the expression level of the target gene also began to increase (Figure 7). These results provided evidence for the involvement of lncRNA-mediated ceRNA regulation in *M. candidum*.

3.6 Functional model of lncRNA-mediated ceRNA during flower development of *M. candidum*

Based on the aforementioned results, we summarized some key lncRNA-mediated ceRNA regulatory relationship during flower development (Figure 8; Figure S2). lncRNAs could bind with a post-transcriptional product of genes directly to degrade them into siRNAs (small interference RNAs). Another common path of lncRNAs is to compete with mRNAs to bind to miRNAs, thereby easing the cleavage effect of miRNA and facilitating more efficient translation of mRNAs (Figure 8A). We observed some ceRNA regulatory mechanisms operating in the flavonoid pathway.

Specifically, lnc-MSTRG.11402.2-miR5207-*McF3H* formed a regulatory chain that acts as a catalyst for the transformation from pinocembrin to pinobanksin (Figure 8B). lnc-MSTRG.28008.2-miR8131-3p-*McASP2* catalyzes the conversion of phenylpyruvate into L-phenylalanine. lnc-MSTRG.29930.1, lnc-MSTRG.23785.1, and lnc-MSTRG.29558.1 could bind with miR1220 to regulate *McPHT3* genes; lnc-MSTRG.24133.1 could affect *McHST* by interacting with miR3704; lnc-MSTRG.34245.1, lnc-MSTRG.29477.1, and, lnc-MSTRG.17619.1 completed with miR480 to bind to *McAHT1*. All these three combinations together influenced the metabolic process from *p*-Coumaroyl-CoA to *p*-Coumaroyl shikimic acid (Figure 8B). Out of all the observed lncRNA-miRNA-mRNA-metabolites regulatory chains, lnc-MSTRG.29930.1,

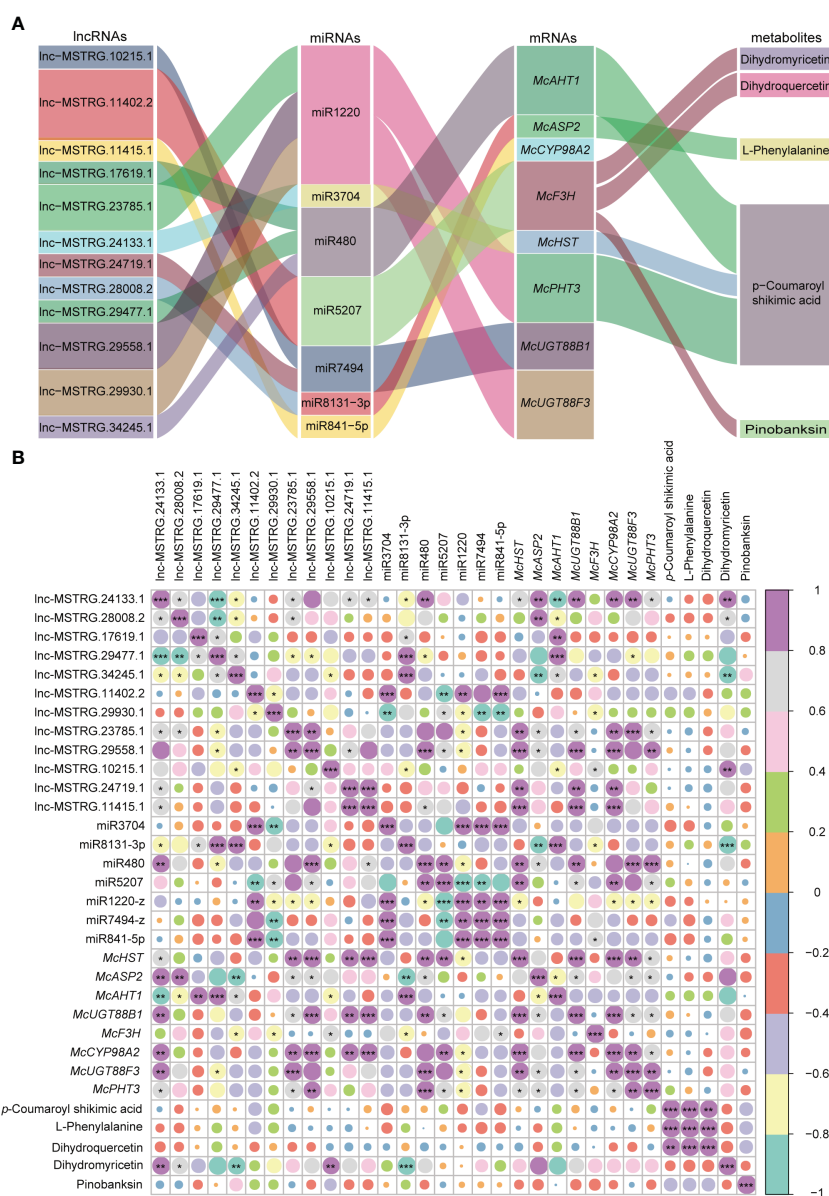


FIGURE 6 Sankey and correlation relationship map of selected lncRNAs, miRNAs, mRNAs, and metabolites. **(A)** Corresponding relationship among lncRNAs, miRNAs, mRNAs, and metabolites. **(B)** Correlation relationship among lncRNAs, miRNAs, mRNAs, and metabolites. Purple means a high correlation coefficient, and green means a low correlation coefficient. **** means *P*-value is significant at 0.001 level, *** means the *P*-value is significant at 0.01 level, and ** means *P*-value is significant at 0.05 level.

McPHT3, and *p*-Coumaroyl shikimic acid were lowly expressed in the second stage, whereas miR1220 was highly expressed in the second stage, indicating the presence of a ceRNA regulatory mechanism. It was found that *lnc*-MSTRG.24133.1, *McHST*, and *p*-Coumaroyl shikimic acid were highly expressed in the third stage, but miR3704 was lowly expressed, suggesting that this ceRNA regulatory chain also plays an important role in the biosynthesis of *p*-Coumaroyl shikimic acid (Figure S4). We also found some irregular ceRNA regulatory chains, for example, *lnc*-MSTRG.11402.2-miR5207-

McF3H, which was thought to be involved in the synthesis of dihydroquercetin, dihydromyricetin, and pinobanksin. However, we did not observe any corresponding expression trends for these metabolites; *lnc*-MSTRG.17619.1, *lnc*-MSTRG.29477.1, *lnc*-MSTRG.34245.1-miR480-*McAHT1*-*p*-Coumaroyl shikimic acid displayed putative trends in the first stage, whereas, in the third stage, they were inconsistent with the putative results (Figures S3, S4). These irregular phenomena may be caused by other post-transcription regulating or modifications mechanisms.

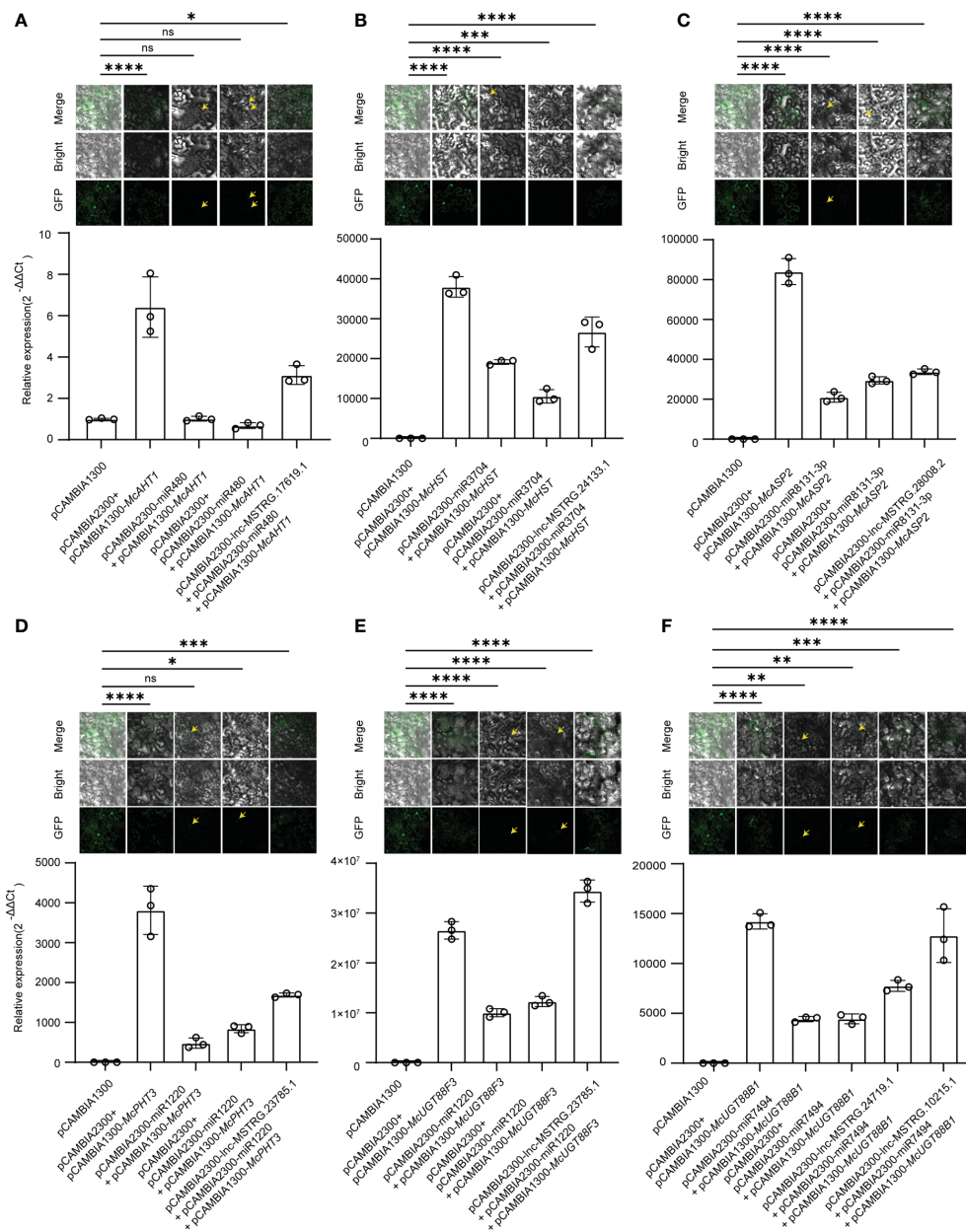


FIGURE 7 Validation of lncRNA-Mediated ceRNA regulatory network. **(A)** Validation of *lnc*-MSTRG.17619.1-miR480-*McAHT1* regulatory chain. **(B)** Validation of *lnc*-MSTRG.24133.1-miR3704-*McHST* regulatory chain. **(C)** Validation of *lnc*-MSTRG.28008.2-miR8131-3p-*McASP2* regulatory chain. **(D)** Validation of *lnc*-MSTRG.23785.1-miR1220-*McPHT3* regulatory chain. **(E)** Validation of *lnc*-MSTRG.23785.1-miR1220-*McUGT88F3* regulatory chain. **(F)** Validation of *lnc*-MSTRG.24719.1-miR7494-*McUGT88B1* and *lnc*-MSTRG.10215.1-miR7494-*McUGT88B1* regulatory chains. The pictures above column are GFP fluorescence graphs obtained using a laser scanning confocal microscope (LSCM). "*" means $p \leq 0.05$; "***" means $p \leq 0.01$; "****" means $p \leq 0.001$; "****" means $p \leq 0.0001$; "ns" means not significant. Yellow arrows point to weak fluorescence in the images.

4 Discussion

Early studies classified lncRNAs as transcriptional noise because of their low transcription levels and conservation. To date, lncRNAs are increasingly recognized as an important regulatory molecule in animals and plants. In plants, Numerous studies have proven lncRNAs' key roles in many biological processes, including flower time (Heo and Sung, 2011), nutrition metabolism (Borah et al., 2018), root development (Chen et al., 2018a), abiotic stress (Gai et al., 2018), etc. In this study, we conducted lncRNAs sequencing for the flower of *M. candidum*. In total, 1,499 DELncRNAs were identified. Several lncRNAs with cis, trans, and antisense functions on mRNA were identified, and the lncRNAs-mediated ceRNA regulatory mechanism was verified. Our research proved that lncRNAs play a pivotal role in flower development and color control in *M. candidum* via direct or indirect ways.

The statistics databases of lncRNAs include CPC, CNCI, PFAM protein structure domain (PFAM), and coding potential assessment tool (CPAT), which will lead to significant discrepancies in the number of lncRNAs detected. For example, by using CPC and CNCI databases, Zhu et al. (2019a) identified 32,036 lncRNAs in the

leaf and shoot of *Camellia sinensis*. Ye et al. (2019) identified 1,860, 3,342, 6,102, and 5,543 lncRNAs by CNCI, CPC, PFAM, and CPAT databases, respectively, in eight tissues of *Ginkgo biloba*, finally, they obtained 1,270 common lncRNAs. Our sequencing results showed that CPC2 and CNC, respectively, produced 4,508 and 6,567 lncRNAs. Finally, 3,955 common lncRNAs based on Venn analysis were identified. Our lncRNA number is more than Ye et al.'s research but far less than Zhu et al.'s research. Since lncRNAs lack codon regions, they are less conserved than protein-coding genes. They may possess a conserved motif but are not easily detected by the BLAST method (Kang and Liu, 2015). Numerous studies had reported that lncRNAs also possessed tissue-specific expression profiles (Ding et al., 2019; Zong et al., 2021). Pairwise comparisons between tissues and stages revealed different proportions of lncRNA isoforms and loci among different tissues or stages in *Fragaria vesca* (Kang and Liu, 2015). So, we concluded that these differences in detected lncRNA numbers in our research are normal and they may be caused by the different detected tissues and species. Further analysis showed that our lncRNAs data was moderate and sufficient to investigate their regulatory relationships involved.

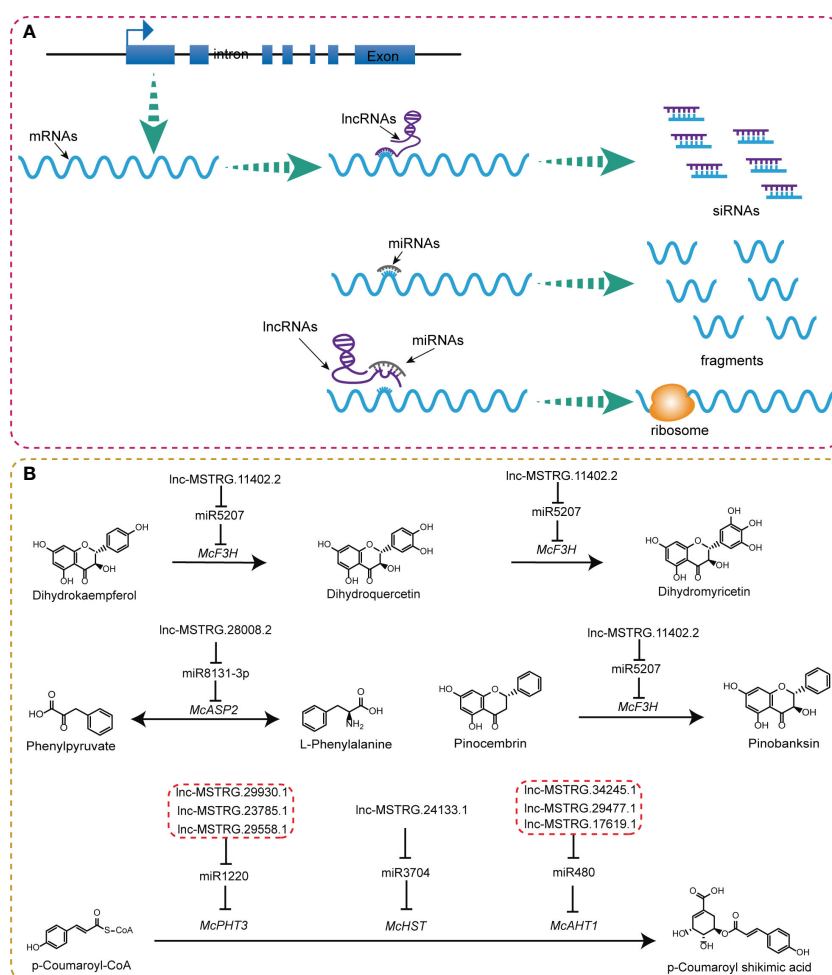


FIGURE 8 LncRNAs-mediated ceRNAs regulatory mechanism in flavonoids pathway. (A) Functional model of lncRNA. (B) Location of lncRNA-mediated ceRNA regulatory chains in corresponding compound biosynthesis process of flavonoids pathway.

The importance of lncRNAs in coloring and pigment formation has been demonstrated by several studies. For example, TCONS_01039552 and PONTK.3920.2 could regulate the expression of *F3H* in sea buckthorn fruit and *Solanum tuberosum*, respectively (Zhang et al., 2018; Bao et al., 2022). Our research identified two lncRNAs, lnc-MSTRG.10215.1 and lnc-MSTRG.11402.2, that can regulate *F3H* via ceRNA regulatory mechanism. (Figure 6A). In strawberry fruit, TRINITY_DN48515_c0_g3_i1 and TRINITY_DN1328_c0_g1_i1 could positively and negatively correlate with *CHI* and *CHS*, respectively (Lin et al., 2018); in *Solanum tuberosum* L, PONTK.2668.1 and PONTK.2668.15 could regulate *CHS* (Bao et al., 2022). In our research, we observed that lnc-MSTRG.30526.3 and lnc-MSTRG.30533.1 could regulate *CHS* by both *cis* and antisense models (Figures 3C, D). lnc-MSTRG.20470.1 could regulate *CHI1* through the antisense model (Figure 3D). Zhu et al. (2019a) found that LCONS00054003 targets *4CL* by *cis* model in fresh leaf and shoot of *Camellia sinensis*. Similarly, our research found that lnc-MSTRG.28304.1 targeted *4CL3*, one member of *4CL*. In addition, lnc-MSTRG.28377.1 targeted on two *COMT1* members, *McCOMT1a* and *McCOMT1b*, and lnc-MSTRG.19485.1 targeted *McMYB4* (Figure 3C). Previous studies demonstrated that *4CL3* displays a strong preference for 4-coumaric acid as substrate and is expressed at high levels in flowers but not in lignified organs suggesting that the primary function of *4CL3* is to provide activated 4-coumaric acid for the chalcone synthase (*CHS*) reaction that feeds the flavonoid-specific branch pathways in *Arabidopsis* (Ehlting et al., 1999; Kumar and Ellis, 2003). *AtMYB4* may affect pollen development by altering the flux of the phenylpropanoid pathway and pollen wall composition (Preston et al., 2004). *COMT1* could catalyze the conversion of caffeic acid to ferulic acid and of 5-hydroxyferulic acid to sinapic acid (Gowri et al., 1991). *COMT1* also methylates 5-hydroxyferuloyl CoA derivatives and flavonols with vicinal aromatic dihydroxy groups, such as quercetin (Fellenberg et al., 2012). In the antisense model, lncRNAs can pair with target genes, such as *McMYB1*, *McMYB17*, *McMADS4*, and *McMADS8* to regulate gene expression and control anthocyanin biosynthesis and pigment accumulation in floral organs. It was reported that a decrease of *MYB1* expression inhibits anthocyanin biosynthesis in bagged Chinese bayberry fruit, suggesting that *MYB1* may be involved in anthocyanin biosynthesis (Niu et al., 2010). In apples, the transcript level of *MYB17* was highly correlated with anthocyanin level, suggesting its role in pigment accumulation. The MADS-box motif has been identified in three different classes of genes in floral organs: A, B, and C. The genes of class A and C are responsible for the development of sepals and carpels, respectively. And the genes of classes B and C together control the formation of stamens (Kang et al., 1998). Additionally, we found some genes exist in both the ceRNA regulatory mechanism and the *cis* model, including *AHT1*, *UGT88B1*, and *CYP98A1*. These results indicated that lncRNAs are key factors to control flower color formation and development through diverse mechanisms in *M. candidum*.

In the trans model of lncRNA regulation, lncRNAs are able to control gene expression from a distance by interacting with target genes at a different chromosomal locus. (Fukuda et al., 2019; Fu et al., 2020). Previous study demonstrated that DEGs regulated by lncRNAs in the trans model were significantly higher than DEGs regulated by

cis and miRNA-mediated models, in which 413 lncRNAs were found to be capable of regulating 6060 genes in the trans model (Zong et al., 2021). Our GO and KEGG analysis results also showed a similar phenomenon. According to our results, a lot of lncRNAs have been shown to regulate protein-coding genes. In GO analysis, we found that the antisense model of lncRNAs could not be detected in the flavone biosynthesis process, or the alkaloid metabolic process; *cis* model of lncRNAs could not be detected in the tropane alkaloid biosynthetic process, and the alkaloid metabolic process. KEGG analysis results revealed that the antisense model of lncRNAs was missing in isoflavonoid biosynthesis pathways, folate biosynthesis pathways, flavone and flavonol biosynthesis pathways, as well as brassinosteroid biosynthesis pathways. The *cis* model of lncRNAs was not observed in the pathway for the synthesis of isoflavonoids and anthocyanins. In addition, the enriched gene numbers in the trans model were also significantly higher than in other acting models (Figures 3A, B). While numerous studies have demonstrated that these three acting models exist, researchers are unable to confirm their role directly due to the absence of related techniques. It is urgent to find a method to affirm the exact function of lncRNAs.

Tobacco injection assay is usually influenced by many factors such as ratio of GV3101 harboring different vectors, activity of the bacteria, injection volume of bacteria, and state of the tobacco leaves. These factors could influence experiment result directly. In our qRT-PCR assay of tobacco, we observed that the expression levels of *McUGT88F3* were higher in pCAMBIA2300-lnc-MSTRG.23785.1 + pCAMBIA2300-miR1220 + pCAMBIA1300-*McUGT88F3* combination than in pCAMBIA2300 + pCAMBIA1300-*McUGT88F3* combination. This phenomenon was not consistent with our expectation. In one hand, although we have controlled OD of bacteria to 0.7 - 1.0 and standardized the final OD to ~ 0.8 before injection, the exact states of the bacteria, such as viability of the bacteria and DNA quality in the bacteria, are still hard to keep consistent, which could affect the experiment results. Furthermore, the quality of tobacco leaves, such as the thickness, age, and texture, could affect the efficiency and volume of the injection, eventually causing different expression levels of target genes. Therefore, strict control of the experimental conditions is crucial for the accurate interpretation of the results. Another possible reason is that there may be highly conserved miRNAs in tobacco and *M. candidum*. When injecting the exogenous target genes, the endogenous miRNAs of tobacco could also degrade these exogenous target genes, leading to the decrease of the expression level of them. The presence of lncRNA may contribute to the enhancement of *McUGT88F3* expression by protecting it from degradation by both endogenous and exogenous miRNAs, eventually leading to higher expression level of target genes in experiment group than in control group. Further studies are needed to confirm these hypotheses.

5 Conclusion

Collectively, the sole flower color of *M. candidum* greatly limits its application in the gardening industry, which is also a common problem for many garden plants. In our research, we identified specific lncRNAs

that regulate flower development or color formation by comparing different stages of flower development of *M. candidum*. Plant development involves a complex set of physiological and biochemical reactions, such as splicing, methylation modification, histone modification at post-transcription or translation level, which could indirectly influence the expression patterns of functional genes. Since most current research focuses mainly on the direct regulation of functional genes or enzymes, some other key factors controlling flower color may be neglected. Despite lncRNAs being a relatively new classification of non-coding RNAs, they play various roles in plants and should be more widely studied in the future. Our research has proven the pivotal role of lncRNAs during flower development and color formation, offering a new avenue for regulating flower color at non-coding RNA levels in *M. candidum*.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: NCBI BioProject accession number: PRJNA884434.

Author contributions

LR, SD and YY designed this research; HL and WW collected the samples for this research; HL analysed, visualized the data, and wrote the draft; HL, BT, RL, XD and YL conducted the verification experiments; LR, SD and YY revised the paper. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by Funding of scientific research projects for postdoc (Grant No. 2022BSHKYZZ); project of

Guangzhou Ecological Garden Science & Technology Collaborative Innovation Center (202206010058).

Acknowledgments

We greatly appreciate prof Renchao Zhou of Sun Yat-sen University for providing the genome of *M. candidum*; Guangzhou Genedenovo Biotechnology for sequencing and bioinformatics analysis; Prof Quanzi Li of Chinese Academy of Forestry for providing experimental equipment and filed for this research.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2023.1215044/full#supplementary-material>

References

- Ahmad, P., Bensaoud, C., Mekki, I., Rehman, M. U., and Kotsyfakis, M. (2021). Long Non-Coding RNAs and their potential roles in the vector-host-pathogen triad. *Life* 11, 56. doi: 10.3390/life11010056
- Bao, Y., Nie, T., Wang, D., and Chen, Q. (2022). Anthocyanin regulatory networks in *Solanum tuberosum* L. leaves elucidated via integrated metabolomics, transcriptomics, and StAN1 overexpression. *BMC Plant Biol.* 22, 1–17. doi: 10.1186/s12870-022-03557-1
- Bayer, E., Egeter, H., Fink, A., Nether, K., and Wegmann, K. (1966). Complex formation and flower colors. *Angew. Chem. Int. Ed. Engl.* 5, 791–798. doi: 10.1002/anie.196607911
- Ben Amor, B., Wirth, S., Merchan, F., Laporte, P., d'Aubenton-Carafa, Y., Hirsch, J., et al. (2009). Novel long non-protein coding RNAs involved in *Arabidopsis* differentiation and stress responses. *Genome Res.* 19, 57–69. doi: 10.1101/gr.080275.108
- Bertone, P., Stolc, V., Royce, T. E., Rozowsky, J. S., Urban, A. E., Zhu, X., et al. (2004). Global identification of human transcribed sequences with genome tiling arrays. *Science* 306, 2242–2246. doi: 10.1126/science.1103388
- Borah, P., Das, A., Milner, M. J., Ali, A., Bentley, A. R., and Pandey, R. (2018). Long non-coding RNAs as endogenous target mimics and exploration of their role in low nutrient stress tolerance in plants. *Genes* 9, 459. doi: 10.3390/genes9090459
- Bowles, D., Lim, E.-K., Poppenberger, B., and Vaistij, F. E. (2006). Glycosyltransferases of lipophilic small molecules. *Annu. Rev. Plant Biol.* 57, 567–597. doi: 10.1146/annurev.arplant.57.032905.105429
- Brockington, S. F., Walker, R. H., Glover, B. J., Soltis, P. S., and Soltis, D. E. (2011). Complex pigment evolution in the *Caryophyllales*. *New Phytol.* 190, 854–864. doi: 10.2307/20869113
- Cabili, M. N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A., et al. (2011). Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* 25, 1915–1927. doi: 10.1101/gad.174466.11
- Chaofeng, T., Minghua, W., and Guiyuan, L. (2013). The interaction between lncRNA and microRNA contributes to tumor. *Chin. J. Biochem. Mol. Biol.* 29, 1029–1034.
- Chen, L.-L. (2016). Linking long noncoding RNA localization and function. *Trends Biochem. Sci.* 41, 761–772. doi: 10.1016/j.tibs.2016.07.003
- Chen, W., Gong, L., Guo, Z., Wang, W., Zhang, H., Liu, X., et al. (2013). A novel integrated method for large-scale detection, identification, and quantification of widely targeted metabolites: application in the study of rice metabolomics. *Mol. Plant* 6, 1769–1780. doi: 10.1093/mp/ss080

- Chen, H.-C., Li, Q., Shuford, C. M., Liu, J., Muddiman, D. C., Sederoff, R. R., et al. (2011). Membrane protein complexes catalyze both 4- and 3-hydroxylation of cinnamic acid derivatives in monolignol biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* 108, 21253–21258. doi: 10.1073/pnas.1116416109
- Chen, L., Shi, S., Jiang, N., Khanzada, H., Wassan, G. M., Zhu, C., et al. (2018a). Genome-wide analysis of long non-coding RNAs affecting roots development at an early stage in the rice response to cadmium stress. *BMC Genomics* 19, 1–10. doi: 10.1186/s12864-018-4807-6
- Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018b). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. doi: 10.1101/274100
- Crick, F., Barnett, L., Brenner, S., and Watts-Tobin, R. J. (1961). General nature of the genetic code for proteins. *Nature* 192, 1227–1232. doi: 10.1038/1921227a0
- de la Fuente, A. (2010). From ‘differential expression’ to ‘differential networking’—identification of dysfunctional regulatory networks in diseases. *Trends Genet.* 26, 326–333. doi: 10.1016/j.tig.2010.05.001
- Deng, J., Chen, S., Yin, X., Wang, K., Liu, Y., Li, S., et al. (2013). Systematic qualitative and quantitative assessment of anthocyanins, flavones and flavonols in the petals of 108 lotus (*Nelumbo nucifera*) cultivars. *Food Chem.* 139, 307–312. doi: 10.1016/j.foodchem.2013.02.010
- Ding, J., Lu, Q., Ouyang, Y., Mao, H., Zhang, P., Yao, J., et al. (2012). A long noncoding RNA regulates photoperiod-sensitive male sterility, an essential component of hybrid rice. *Proc. Natl. Acad. Sci. U.S.A.* 109, 2654–2659. doi: 10.1073/pnas.1121374109
- Ding, Z., Tie, W., Fu, L., Yan, Y., Liu, G., Yan, W., et al. (2019). Strand-specific RNA-seq based identification and functional prediction of drought-responsive lncRNAs in cassava. *BMC Genomics* 20, 1–13. doi: 10.1186/s12864-019-5585-5
- Ehltling, J., Büttner, D., Wang, Q., Douglas, C. J., Somssich, I. E., and Kombrink, E. (1999). Three 4-coumarate: coenzyme A ligases in *Arabidopsis thaliana* represent two evolutionarily divergent classes in angiosperms. *Plant J.* 19, 9–20. doi: 10.1046/j.1365-313X.1999.00491.x
- Erdtman, H. (1944). The phenolic constituent of pineheartwood. *Sven. Kem. Tidskr.* 56, 26.
- Fang, J., Zhang, F., Wang, H., Wang, W., Zhao, F., Li, Z., et al. (2019). Ef-cd locus shortens rice maturity duration without yield penalty. *Proc. Natl. Acad. Sci. U.S.A.* 116, 18717–18722. doi: 10.1073/pnas.1815030116
- Fasoula, D. A., Stephens, P. A., Nickell, C. D., and Vodkin, L. O. (1995). Cosegregation of purple-throat flower color with dihydroflavonol reductase polymorphism in soybean. *Crop Sci.* 35, 1028–1031. doi: 10.2135/cropsci1995.0011183X003500040017x
- Fellenberg, C., van Ohlen, M., Handrick, V., and Vogt, T. (2012). The role of CCoAMT1 and COMT1 in *Arabidopsis* anthers. *Planta* 236, 51–61. doi: 10.1007/s00425-011-1586-6
- Franco-Zorrilla, J. M., Valli, A., Todesco, M., Mateos, I., Puga, M. I., Rubio-Somoza, I., et al. (2007). Target mimicry provides a new mechanism for regulation of microRNA activity. *Nat. Genet.* 39, 1033–1037. doi: 10.1038/ng2079
- Fu, L., Ding, Z., Tan, D., Han, B., Sun, X., and Zhang, J. (2020). Genome-wide discovery and functional prediction of salt-responsive lncRNAs in duckweed. *BMC Genomics* 21, 1–14. doi: 10.1186/s12864-020-6633-x
- Fukuda, M., Nishida, S., Kakei, Y., Shimada, Y., and Fujiwara, T. (2019). Genome-wide analysis of long intergenic noncoding RNAs responding to low-nutrient conditions in *Arabidopsis thaliana*: possible involvement of trans-acting siRNA3 in response to low nitrogen. *Plant Cell Physiol.* 60, 1961–1973. doi: 10.1093/pcp/pcz048
- Gai, Y.-P., Yuan, S.-S., Zhao, Y.-N., Zhao, H.-N., Zhang, H.-L., and Ji, X.-L. (2018). A novel lncRNA, MuLnc1, associated with environmental stress in Mulberry (*Morus multicaulis*). *Front. Plant Sci.* 9, doi: 10.3389/fpls.2018.00669
- Gowri, G., Bugos, R. C., Campbell, W. H., and Dixon, M. R. A. (1991). Stress Responses in Alfalfa (*Medicago sativa* L.) X. Molecular cloning and expression of S-Adenosyl-L-Methionine: caffeic acid3-O-methyltransferase, a key enzyme of lignin biosynthesis. *Plant Physiol.* 97, 7–14. doi: 10.1104/pp.97.1.7
- Guttman, M., Amit, I., Garber, M., French, C., Lin, M. F., Feldser, D., et al. (2009). Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* 458, 223. doi: 10.1038/nature07672
- He, X., Guo, S., Wang, Y., Wang, L., and Sun, J. (2020). Systematic identification and analysis of heat stress-responsive lncRNAs, circRNAs and miRNAs with associated coexpression and ceRNA networks in cucumber (*Cucumis sativus* L.). *Physiol. Plant* 168, 736–754. doi: 10.1111/pp.12997
- Heo, J. B., and Sung, S. (2011). Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. *Science* 331, 76–79. doi: 10.1126/science.1197349
- Holton, T. A. (1995). Modification of flower colour via manipulation of P450 gene expression in transgenic plants. *Drug Metab. Drug Interact.* 12, 359–368. doi: 10.1515/DMDI.1995.12.3-4.359
- Holton, T., and Cornish, E. (1995). Genetics and biochemistry of anthocyanin biosynthesis. *Plant Cell* 7, 1071–1083. doi: 10.26877/asset.v4i1.11659
- Kang, C., and Liu, Z. (2015). Global identification and analysis of long non-coding RNAs in diploid strawberry *Fragaria vesca* during flower and fruit development. *BMC Genomics* 16, 1–15. doi: 10.1186/s12864-015-2014-2
- Kang, H. G., Jeon, J. S., and Lee, S. (1998). “Identification of class B and class C floral organ identity genes from rice plants. *Plant Mol. Biol.* 38, 1021–1029. doi: 10.1023/A:1006051911291
- Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi: 10.1038/nmeth.3317
- Kim, D.-H., and Sung, S. (2017). Vernalization-triggered intragenic chromatin loop formation by long noncoding RNAs. *Dev. Cell* 40, 302–312. e304. doi: 10.1016/j.devcel.2016.12.021
- Kishimoto, S., Maoka, T., Nakayama, M., and Ohmiya, A. (2004). Carotenoid composition in petals of chrysanthemum (*Dendranthema grandiflorum* (Ramat.) Kitamura). *Phytochemistry* 65, 2781–2787. doi: 10.1016/j.phytochem.2004.08.038
- Kondo, T., Ueda, M., Isobe, M., and Goto, T. (1998). A new molecular mechanism of blue color development with protocyanin, a supramolecular pigment from cornflower, *Centaurea cyanus*. *Tetrahedron Lett.* 39, 8307–8310. doi: 10.1016/S0040-4039(98)01858-9
- Kong, L., Zhang, Y., Ye, Z.-Q., Liu, X.-Q., Zhao, S.-Q., Wei, L., et al. (2007). CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res.* 35, 345–349. doi: 10.1093/nar/gkm391
- Kumar, A., and Ellis, B. E. (2003). 4-Coumarate: CoA ligase gene family in *Rubus idaeus*: cDNA structures, evolution, and expression. *Plant Mol. Biol.* 51, 327–340. doi: 10.1023/A:1022004923982
- Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923
- Li, B., and Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinf.* 12, 1–16. doi: 10.1186/1471-2105-12-323
- Li, X., Fan, J., Luo, S., Yin, L., Liao, H., Cui, X., et al. (2021). Comparative transcriptome analysis identified important genes and regulatory pathways for flower color variation in *Paphiopedilum hirsutissimum*. *BMC Plant Biol.* 21, 1–17. doi: 10.1186/s12870-021-03256-3
- Li, C., Jin, H., Zhang, W., Qin, T., Zhang, X., Pu, Z., et al. (2022). Whole-transcriptome analysis reveals long noncoding RNAs involved in female floral development of hickory (*Carya cathayensis* Sarg.). *Front. Genet.* 13, doi: 10.3389/fgene.2022.910488
- Lin, Y., Jiang, L., Chen, Q., Li, Y., Zhang, Y., Luo, Y., et al. (2018). Comparative transcriptome profiling analysis of red- and white-fleshed strawberry (*Fragaria × ananassa*) provides new insight into the regulation of the anthocyanin pathway. *Plant Cell Physiol.* 59, 1844–1859. doi: 10.1093/pcp/pcy098
- Lindstedt, G. (1950). Constituents of Pine heartwood. XXI. The structure of pinobanksin. *Acta Chem. Scand.* 4, 772–781. doi: 10.3891/acta.chem.scand.04-0772
- Liu, T., Chen, Y., Chao, L., Wang, S., Wu, W., Dai, S., et al. (2014). Extensive hybridization and introgression between *Melastoma candidum* and *M. sanguineum*. *PLoS One* 9, e96680. doi: 10.1371/journal.pone.0096680
- Liu, H., Lou, Q., Ma, J., Su, B., Gao, Z., and Liu, Y. (2019). Cloning and functional characterization of dihydroflavonol 4-reductase gene involved in anthocyanidin biosynthesis of grape hyacinth. *Int. J. Mol. Sci.* 20, 4743. doi: 10.3390/ijms20194743
- Liu, F., Marquardt, S., Lister, C., Swiezewski, S., and Dean, C. (2010). Targeted 3′ processing of antisense transcripts triggers *Arabidopsis* FLC chromatin silencing. *Science* 327, 94–97. doi: 10.1126/science.1180278
- Liu, D., Shi, L., Han, C., Yu, J., Li, D., and Zhang, Y. (2012). Validation of reference genes for gene expression studies in virus-infected *Nicotiana benthamiana* using quantitative real-time PCR. *PLoS One* 7, e46451. doi: 10.1371/journal.pone.0046451
- Lorenz, R., Bernhart, S. H., Höner zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P. F., et al. (2011). ViennaRNA package 2.0. *Algorithms Mol. Biol.* 6, 1–14. doi: 10.1186/1748-7188-6-26
- Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 1–21. doi: 10.1186/PREACCEPT-8897612761307401
- Lu, Y., Deng, S., Li, Z., Wu, J., Liu, Q., Liu, W., et al. (2019). Competing endogenous RNA networks underlying anatomical and physiological characteristics of poplar wood in acclimation to low nitrogen availability. *Plant Cell Physiol.* 60, 2478–2495. doi: 10.1093/pcp/pcz146
- Mitsui, S., Hayashi, K., and Hattori, S. (1959). Further studies on commelinin, a crystalline blue metallo-anthocyanin from the flowers of commelina studies on anthocyanins. *Proc. Jpn. Acad.* 35, 169–174. doi: 10.2183/pjab1945.35.169
- Mol, J., Grotewold, E., and Koes, R. (1998). How genes paint flowers and seeds. *Trends Plant Sci.* 3, 212–217. doi: 10.1016/S1360-1385(98)01242-4
- Necsulea, A., Soumillon, M., Warnefors, M., Liechti, A., Daish, T., Zeller, U., et al. (2014). The evolution of lncRNA repertoires and expression patterns in tetrapods. *Nature* 505, 635–640. doi: 10.1038/nature12943
- Nie, L., Wu, H. J., Hsu, J. M., Chang, S. S., and Hung, M. C. (2012). Long non-coding RNAs: versatile master regulators of gene expression and crucial players in cancer. *Am. J. Transl. Res.* 4, 127–150. doi: 10.1155/2012/952452
- Nielsen, K. M., Lewis, D. H., and Morgan, E. R. (2003). Characterization of carotenoid pigments and their biosynthesis in two yellow flowered lines of *Sandersonia aurantiaca* (Hook). *Euphytica* 130, 25–34. doi: 10.1023/A:1022328828688
- Niu, S.-S., Xu, C.-J., Zhang, W.-S., Zhang, B., Li, X., Lin-Wang, K., et al. (2010). Coordinated regulation of anthocyanin biosynthesis in Chinese bayberry (*Myrica rubra*) fruit by a R2R3 MYB transcription factor. *Planta* 231, 887–899. doi: 10.1007/s00425-009-1095-z
- Paliyath, G., Murr, D. P., Handa, A. K., and Lurie, S. (2009). *Postharvest biology and technology of fruits, vegetables, and flowers* John Wiley & Sons, (2th Ed), USA.

- Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290–295. doi: 10.1038/nbt.3122
- Preston, J., Wheeler, J., Heazlewood, J., Li, S. F., and Parish, R. W. (2004). *AtMYB32* is required for normal pollen development in *Arabidopsis thaliana*. *Plant J.* 40, 979–995. doi: 10.1111/j.1365-313X.2004.02280.x
- Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616
- Shen, S., Park, J. W., Lu, Z.-x., Lin, L., Henry, M. D., Wu, Y. N., et al. (2014). rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proc. Natl. Acad. Sci. U.S.A.* 111, E5593–E5601. doi: 10.1073/pnas.1419161111
- Shimada, S., Otsuki, H., and Sakuta, M. (2007). Transcriptional control of anthocyanin biosynthetic genes in the *Caryophyllales*. *J. Exp. Bot.* 58, 957–967. doi: 10.1093/jxb/erl256
- Shoji, K., Miki, N., Nakajima, N., Momonoi, K., Kato, C., and Yoshida, K. (2007). Perianth bottom-specific blue color development in tulip cv. Murasakizuisho requires ferric ions. *Plant Cell Physiol.* 48, 243–251. doi: 10.1093/pcp/pcl060
- Sun, L., Luo, H., Bu, D., Zhao, G., Yu, K., Zhang, C., et al. (2013). Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. *Nucleic Acids Res.* 41, e166. doi: 10.1093/nar/gkt646
- Sunnadeniya, R., Bean, A., Brown, M., Akhavan, N., Hatlestad, G., Gonzalez, A., et al. (2016). Tyrosine hydroxylation in betalain pigment biosynthesis is performed by cytochrome P450 enzymes in beets (*Beta vulgaris*). *PLoS One* 11, e0149417. doi: 10.1371/journal.pone.0149417
- Takeda, K. (2006). Blue metal complex pigments involved in blue flower color. *Proc. Jpn. Acad.* 82, 142–154. doi: 10.2183/pjab.82.142
- Tanaka, Y., and Brugliera, F. (2013). Flower colour and cytochromes P450. *Phytochem. Rev.* 368, 20120432. doi: 10.1007/s1101-006-9003-7
- Tanaka, Y., Tsuda, S., and Kusumi, T. (1998). Metabolic engineering to modify flower color. *Plant Cell Physiol.* 39, 1119–1126. doi: 10.1093/oxfordjournals.pcp.a029312
- Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., Van Baren, M. J., et al. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28, 511–515. doi: 10.1038/nbt.1621
- Wang, K. C., and Chang, H. Y. (2011). Molecular mechanisms of long noncoding RNAs. *Mol. Cell* 43, 904–914. doi: 10.1016/j.molcel.2011.08.018
- Wang, M., Wu, H.-J., Fang, J., Chu, C., and Wang, X.-J. (2017). A long noncoding RNA involved in rice reproductive development by negatively regulating osa-miR160. *Sci. Bull.* 62, 470–475. doi: 10.1016/j.scib.2017.03.013
- Willstätter, R., and Everest, A. E. (1913). Untersuchungen über die Anthocyane. I. Über den Farbstoff der Kornblume. *Justus Liebigs Ann. Chem.* 401, 189–232. doi: 10.1002/jlac.19134010205
- Wu, X., Shi, T., Iqbal, S., Zhang, Y., Liu, L., and Gao, Z. (2019). Genome-wide discovery and characterization of flower development related long non-coding RNAs in *Prunus mume*. *BMC Plant Biol.* 19, 1–17. doi: 10.1186/s12870-019-1672-7
- Wunderlich, M., Groß-Hardt, R., and Schöffl, F. (2014). Heat shock factor HSF2a involved in gametophyte development of *Arabidopsis thaliana* and its expression is controlled by a heat-inducible long non-coding antisense RNA. *Plant Mol. Biol.* 85, 541–550. doi: 10.1007/s11103-014-0202-0
- Xu, W., Bao, W., Liu, H., Chen, C., Bai, H., Huang, M., et al. (2021). Insights into the molecular mechanisms of late flowering in *Prunus sibirica* using whole-genome and transcriptome analyses. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.802827
- Xu, X.-W., Zhou, X.-H., Wang, R.-R., Peng, W.-L., An, Y., and Chen, L.-L. (2016). Functional analysis of long intergenic non-coding RNAs in phosphate-starved rice using competing endogenous RNA network. *Sci. Rep.* 6, 1–12. doi: 10.1038/srep20715
- Yang, Z., Yang, C., Wang, Z., Yang, Z., Chen, D., Wu, Y., et al. (2019). LncRNA expression profile and ceRNA analysis in tomato during flowering. *PLoS One* 14, e0210650. doi: 10.1371/journal.pone.0210650
- Yanofsky, C. (2007). Establishing the triplet nature of the genetic code. *Cell* 128, 815–818. doi: 10.1016/j.cell.2007.02.029
- Ye, J., Cheng, S., Zhou, X., Chen, Z., Kim, S. U., Tan, J., et al. (2019). A global survey of full-length transcriptome of *Ginkgo biloba* reveals transcript variants involved in flavonoid biosynthesis. *Ind. Crops Prod.* 139, 111547. doi: 10.1016/j.indcrop.2019.111547
- Yu, Z.-W., Zhang, N., Jiang, C.-Y., Wu, S.-X., Feng, X.-Y., and Feng, X.-Y. (2021). Exploring the genes involved in biosynthesis of dihydroquercetin and dihydromyricetin in *Ampelopsis grossedentata*. *Sci. Rep.* 11, 1–14. doi: 10.1038/s41598-021-95071-x
- Zhang, G., Chen, D., Zhang, T., Duan, A., Zhang, J., and He, C. (2018). Transcriptomic and functional analyses unveil the role of long non-coding RNAs in anthocyanin biosynthesis during sea buckthorn fruit ripening. *DNA Res.* 25, 465–476. doi: 10.1093/dnares/dsy017
- Zhang, J., Mujahid, H., Hou, Y., Nallamilli, B. R., and Peng, Z. (2013). Plant long ncRNAs: a new frontier for gene regulatory control. *Am. J. Plant Sci.* 04, 1038–1045. doi: 10.4236/ajps.2013.45128
- Zhu, H.-h., Yang, J.-x., Xiao, C.-h., Mao, T.-y., Zhang, J., and Zhang, H.-y. (2019b). Differences in flavonoid pathway metabolites and transcripts affect yellow petal colouration in the aquatic plant *Nelumbo nucifera*. *BMC Plant Biol.* 19, 1–18. doi: 10.1186/s12870-019-1886-8
- Zhu, C., Zhang, S., Fu, H., Zhou, C., Chen, L., Chang, X., et al. (2019a). Transcriptome and phytochemical analyses provide new insights into long non-coding RNAs and characteristic secondary metabolites of oolong tea (*Camellia sinensis*) in solar-withering. *Front. Plant Sci.* 10, 1683. doi: 10.21203/rs.2.11183/v1
- Zhu, M., Zhang, M., Xing, L., Li, W., Jiang, H., Wang, L., et al. (2017). Transcriptomic analysis of long non-coding RNAs and coding genes uncovers a complex regulatory network that is involved in maize seed development. *Genes* 8, 274. doi: 10.3390/genes8100274
- Zong, X., Wang, S., Han, Y., Zhao, Q., Xu, P., Yan, Q., et al. (2021). Genome-wide profiling of the potential regulatory network of lncRNA and mRNA in *Melilotus albus* under salt stress. *Environ. Exp. Bot.* 189, 104548. doi: 10.1016/j.envexpbot.2021.104548
- Zuo, J., Wang, Y., Zhu, B., Luo, Y., Wang, Q., and Gao, L. (2019). Network analysis of noncoding RNAs in pepper provides insights into fruit ripening control. *Sci. Rep.* 9, 1–11. doi: 10.1038/s41598-019-45427-1