Journal of Genetics (2020)99:31
https://doi.org/10.1007/s12041-020-01194-x

RESEARCH ARTICLE

Identification of anthocyanin biosynthesis related microRNAs and total microRNAs in *Lonicera edulis* by high-throughput sequencing

JIE CUI[1]*, ZENGLE GAO[1], BINSHENG LI[2], JUNLIANG LI[1], XINYAN LI[1], CONGYU WANG[1], DAYOU CHENG[1] and CUIHONG DAI[1]

[1]*Department of Chemistry and Chemical Engineering, Harbin Institute of Technology, Harbin 150001, Heilongjiang People's Republic of China*
[2]*Plant Introduction and Domestication Laboratory, Forest Botanical Garden in Heilongjiang, Harbin 150040, People's Republic of China*
*For correspondence. E-mail: cuijie@hit.edu.cn.

Received 16 May 2019; revised 31 December 2019; accepted 14 February 2020

Abstract. miRNAs are important regulators of plant gene expression. There are few studies on the regulation of miRNAs in *Lonicera edulis*. We used high-throughput sequencing technology to analyse miRNAs in *L. edulis*, aiming to identify miRNAs and elucidate their function in *L. edulis*. In the present study, we employed the high-throughput sequencing technology to profile miRNAs in *L. edulis*. A total of 51,819,072 small RNA tags with sizes ranging from 18 to 30 nt were obtained, indicating that *L. edulis* have a large and diverse small RNA population. Bioinformatic analysis identified 507 mature miRNAs, and 16 predicted novel miRNAs that are likely to be unique to *L. edulis*. Three miRNAs related to anthocyanin biosynthesis were locked by gene ontology (GO) analysis and target gene analysis. The selected three miRNAs are relatively high in the expression of *L. edulis*. Some of the previous studies have studied these types of miRNAs involved in the anthocyanin metabolism pathway in fruits. Among them, expression profiles of three conserved miRNAs were validated by stem loop qRT-PCR. Further, the potential target genes of conserved and novel miRNAs were predicted and subjected to GO annotation. Enrichment analysis of the GO-represented biological processes and molecular functions revealed that these target genes were potentially involved in a wide range of metabolic pathways and developmental processes. In particular, different families of miRNAs can directly or indirectly regulate anthocyanin biosynthesis. In recent years, the research on miRNAs has become more and more clear, but the research on miRNAs involved in the regulation of anthocyanin synthesis of *L. edulis* is still lagging. This study provides a useful resource for further elucidation of the functional roles of miRNAs during fruit development and ripening.

Keywords. anthocyanins; micro RNAs; high-throughput sequencing; *Lonicera edulis.*

Introduction

Anthocyanidin is a water-soluble pigment that imparts red, blue and purple flavonoid secondary metabolites to plant tissues. The presence of anthocyanins is beneficial to the adaptation of plants to adverse conditions such as low temperature, UV stress and pathogen attack (Chiu *et al.* 2010). Anthocyanins are synthesized by the phenylpropane pathway and stored in vacuoles (Li *et al.* 2016), but their cellular transport is not known. Anthocyanins have health-care functions and are resistant to certain ageing-related diseases, cardiovascular diseases, and cancers (Yoshimoto *et al.* 1999; Wang *et al.* 2000; Kong *et al.* 2003). Yuan *et al.* (2014) showed that cyanidin is the main anthocyanin component of *Lonicera japonica* L. Liu *et al.* (2016) used high performance liquid chromatography (HPLC) detection technology and UV detector to qualitatively determine the fruit of *L. edulis*. The main anthocyanins in the fruit of *L. edulis* are mainly cyanidin. The biosynthesis process of anthocyanins is controlled by two types of genes: structural genes and regulatory genes. Among them, the structural

JC, experimental guidance and analysis; BL, provided the resource, experimental sample planting, cultivation, management, picking; JL and XL, analysed the data; CW, collect samples; DC & CD, revised the manuscript.

Published online: 19 April 2020

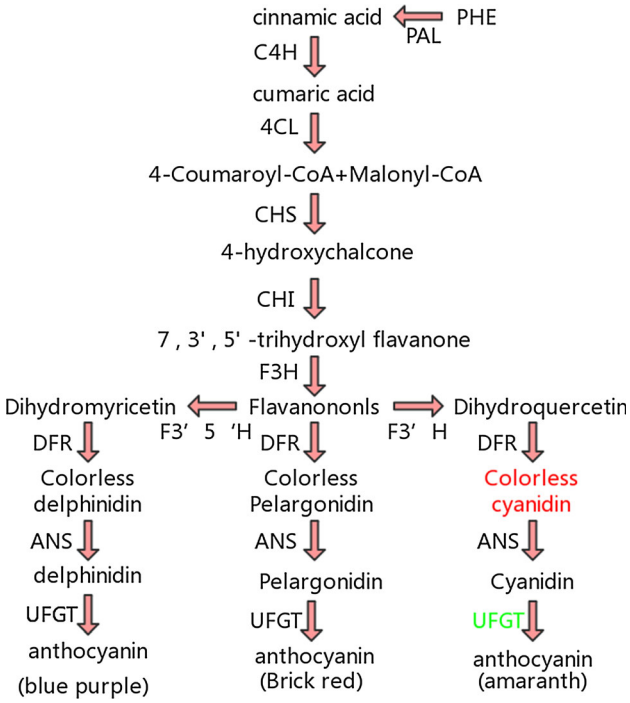

Figure 1. Anthocyanin biosynthesis pathway of *L. edulis*.

gene encodes a series of biosynthetic enzymes, which synthesize anthocyanins capable of colouring fruits under the catalysis of enzymes. Regulatory genes regulate the biosynthesis of anthocyanins by regulating structural genes. The anthocyanin biosynthesis pathway is regulated at the transcriptional level by MYB, bHLH and WD40 transcriptional regulators (Xie *et al.* 2012). At present, the anthocyanin biosynthesis pathway of *L. edulis* has been clarified, as shown in figure 1.

L. edulis is distributed in China's Jilin province, eastern Heilongjiang, Inner Mongolia, Sichuan, and Northwest China; Russia's Far East; North Korea and Japan. It is rare and unique in flavour and rich in natural pigments. It is very popular with consumers at home and abroad. It is known as the 'king of the third generation fruit'. *L. edulis* is rich in anthocyanins and has functions such as delaying ageing and scavenging free radicals (Yoshimoto *et al.* 1999). Studies have reported that fruits contain a variety of microRNAs (miRNAs), such as grapes, apples and lychees. We have detected a variety of miRNAs in *L. edulis* by high-throughput sequencing technology, which has a variety of biological functions. Mature miRNAs are a group of 20–24 nt noncoding small RNAs that play an important role in plant growth, development, metabolism and stress responses by inhibiting their targets at the post-transcriptional level. Although the regulation of miRNAs in plants is well known, there are a few studies on the regulatory mechanisms of miRNAs in the anthocyanin synthesis of *L. edulis*. It has been reported that some miRNAs can target transcription factors related to anthocyanin biosynthesis. In grapes, studies have shown that miRNAs play a regulatory role in the expression of WD40 protein and play an important role in the biosynthesis process of anthocyanins (An *et al.* 2012). Rock (2013) showed that miR828 targets two MYBs genes in grapes, namely VvMYBA6 and VvMYBA7 genes, which play an important role in the synthesis of flavonoids in grapes (Christopher 2013). Studies have shown that the gene encoding MdMYB9 in apple is a predicted target of mdm-miR858, and the expression levels of MdMYB9 and miR858 are reported to have opposite trends (Dong *et al.* 2016). In apples, miR858, miR828 and miR159 form a complex and large regulatory network by interacting with multiple MYBs, and participate in the regulation of apple growth and anthocyanin metabolism. Studies have shown that MYB113 and MYB114 in the MYB transcription factor redundantly regulate the production of anthocyanins by regulating the anthocyanin biosynthesis gene UFGT (Gonzalez *et al.* 2008). The target UFGT of miRNAs associated with anthocyanin synthesis identified in this study was identified, and UFGT is a key structural gene downstream of the anthocyanin synthesis pathway. Li *et al.* (2010) showed that *UFGT* gene is closely related to anthocyanin synthesis in apple and pear. Zhang *et al.* (2018) found that UFGT in strawberry fruit is expressed during fruit colouration.

Sun *et al.* (2016) found that miR157 and miR8181 were first identified in radish roots to regulate anthocyanin biosynthesis by regulating target gene SPL. Yue *et al.* (2017) analysed miRNAs in blueberries by high-throughput sequencing and degradation groups, which identified their targets. Studies have found that miRNAs involved in the anthocyanin biosynthesis pathway in blueberries include vas-miR-21, whose target is the gene encoding chalcone isomerase (CHI). It was confirmed by degradation group analysis that vas-miR167a in the young fruit could cleave the transcript of flavanone 3-hydroxylase (F3H) and the transcript of flavanone 3-hydroxylase (F3H) as vas-miR167a Target. The vas-miR858b in blueberry was identified, and its target is a transcription factor encoding TT2 type MYB. Studies have shown that MIR156 can positively regulate the accumulation of anthocyanins by directed cleavage of SPL transcripts in Arabidopsis and rice (Xie *et al.* 2012). miR828 and miR858 are capable of negatively regulating anthocyanin biosynthesis by cleaving MYB transcription factors (Jia *et al.* 2015). In addition, studies have shown that miR858 has been identified in Arabidopsis, apple, cotton and tomato, and that miR858 is predicted to target up to 66 MYB factors in apples (Xia *et al.* 2012; Guan *et al.* 2014). Blocking the expression of miR858 can increase anthocyanin accumulation in tomato fruit (Jia *et al.* 2015). These miRNAs play an important role in fruit development and maturation. These findings provide new insights into the further identification of miRNAs in fruits, deepening our understanding of the functional properties of miRNAs, and their understanding of the regulation of anthocyanin biosynthesis in *L. edulis*.

Figure 2. Fruits of the five periods from development to maturity (a: No. 6 'Blue Sky'; b: No. 9 'Blue Forest' for Canada. a-1/a-2. young fruit; b-1/b-2. green ripe stage; c-1/c-2. half colour change; d-1/d-2. colour changed; e-1/e-2. mature).

Materials and methods

Plant material and sample collection

From 9 May to 23 June, *L. edulis* was sampled in the plantation area of *L. edulis* plant of Heilongjiang Forest Botanical Garden. Each material was picked at least three times, each time in 10 plants. The collection was carried out from five periods of fruit development to maturity. Fruits with the same growth state were collected and immediately wrapped in tin foil and pre-cooled with liquid nitrogen, and stored at −80°C. The *L. edulis* variety used in the experiment was the introduction of variety no. 6 'Blue Sky' and No. 9 'Blue Forest' for Canada (figure 2).

Anthocyanin analysis

The anthocyanin content was determined by solvent extraction method combined with pH differential method. The anthocyanin content of sassafras fruit was measured and the accumulation trend of anthocyanin content in different varieties and different periods was compared. Anthocyanin content (mg/100 g) = (A/εL) × MW × DF × V/Wt × 100(A, absorbance; ε, cornflower-3-glucoside extinction coefficient, 26,900; L, cuvette optical path, 1 cm; MW, cyanidin-3-glucoside molecular weight, 449.2; DF, dilution factor; V, final volume, mL; Wt, product weight, mg; A = (Aλ vis-max-A700) pH 1.0 − (Aλ vis-max-A700) pH 4.5.).

Bioinformatics analysis of sRNA sequencing data

We removed the annotated noncoding RNA readings including rRNA, tRNA, snRNA and snoRNA. After annotating all the small RNA fragments, we used the unannotated fragments to make novel miRNA predictions. For the identification of miRNAs, the target gene prediction, GO function annotation and KEGG pathway of target genes were performed. The detailed data processing steps are as follows: (i) removed tags with lower sequencing quality. (ii) Removed tags contaminated with 5' adapter. (iii) Removed tags without 3' adapter sequences. (iv) Removed tags without inserted fragments. (v) Removed tags containing polyA. (vi) Removed tags with length <18 nt. (vii) Calculated the length distribution of small RNA fragments.

The clean reads are then compared to the reference set and other small RNA databases for comparison using the alignment software AASRA. miRNA-based precursors are capable of forming hairpin secondary structures, and we used miRDeep2 and miRA for new miRNA predictions. The expression level of small RNAs is normalized using TPM. the standardized data can be directly used for subsequent differential comparison analysis. The formula for calculating TPM is as follows: TPM = C*10^6/N. C represents the copy number of a certain miRNA in the sample, and N represents the total number of copies of the sample to the genome.

Differential expression analysis of miRNAs

The miRNA was extracted using ABI's mirVanaTM miRNA Isolation kit and the protocol was followed according to the instructions. The reverse transcription of miRNA was performed using the Taq Man MicroRNA Reverse Transcription kit from ABI, USA, and the procedure was followed according to the instructions. Quantification of miRNA and its target genes using the qRT-PCR reaction instructions ABI's Taq Man Gene Expression Master Mix, the procedure was followed according to the instructions. The relative abundance of gene expression was determined using the

calculated relative expression change $2^{-\Delta\Delta Ct}$ method and miR159 as an internal control gene. All gene expression data were derived from three biological replicates with three technical replicates per biological sample.

In summary, based on high-throughput sequencing analysis, we identified 507 known and 16 potential novel miRNAs and predicted their target genes. Our findings also indicate that miRNAs target genes are involved in the regulation of cellular metabolic processes, metabolic processes, structural molecule activit, auxin mediated signalling pathways, and anthocyanin biosynthesis. These findings provide new insights into the fact that miR396a-3p and miR156h-3p regulate structural genes downstream of the anthocyanin synthesis pathway and are further involved in the synthesis of anthocyanins in the fruit ripening process. The data presented in this paper provide a basis for the regulation of miRNAs in *L. edulis*, and provide new insights into the effects of miRNAs on the development and maturation of *L. edulis*.

Results

Analysis of anthocyanin content and UFGT enzyme activity in L. edulis

***Analysis of anthocyanin content in L. edulis*:** The content of anthocyanins of the two varieties 'No. 6' and 'No. 9' was measured by solvent immersion pH difference method (figure 3). The content of anthocyanin in 'No. 6' and 'No. 9' gradually accumulated and increased from fruit development to maturity. The content of anthocyanins in the young fruit stage, green ripening stage and half-colouring stage of each variety was almost zero. The colour period begins to accumulate rapidly and reaches its highest level at maturity. The anthocyanin content of No. 6 variety was 253.73 mg/100 g, and the anthocyanin content of No. 9 variety was 75.24 mg/100 g; The content of anthocyanins in the ripening stage of the fruit: 'No. 6 > 'No. 9'. The content of anthocyanin in the fruit was in the final stage of accumulation. The content of anthocyanin in 'No. 6' was the highest, and the content of anthocyanin in 'No. 9' was the lowest, which was consistent with the result of direct observation, i.e. the fruit ratio of 'No. 6' variety was 'No. 9'. The colour of the fruit of the variety is shallow.

***Analysis of UFGT enzyme activity in L. edulis*:** The colour of the *L. edulis* is mainly determined by the amount of anthocyanins. UFGT controls the synthesis of anthocyanins. UFGT is a key enzyme encoding the last step in the anthocyanin synthesis pathway. Its main function is to transfer the glucosyl group on UDP-glucose to the C3 hydroxyl group of the anthocyanin molecule. The focus of this study was on the No. 6 'Blue Sky' in *L. edulis*, so the enzyme activity of the No. 6 'Blue Sky' fruit was further determined. The activity of UFGT enzyme shown in figure 4 gradually increased with the accumulation of fruit to mature pigment, which was consistent with the accumulation trend of anthocyanins. It was further shown that enhancing the activity of UFGT has an important effect on the accumulation of anthocyanins in *L. edulis*. The enhancement of UFGT activity promotes the accumulation of anthocyanins.

***High-throughput sequencing of sRNA in L. edulis*:** The data obtained by sequencing the fruit of *L. edulis* is compared with the noncoding database of the 22nd edition of the miRBase database or piRNA database (partial animal), snoRNA (human or partial plant). Noncoding small RNAs of each sample were detected using siRNA or piRNA prediction. The low quality data is first removed and then analysed downstream to obtain the amount of sequencing data for each sample. The valid data after initial processing of sRNA group sequencing data includes sRNA sequence information and its expression information (RPM). Known

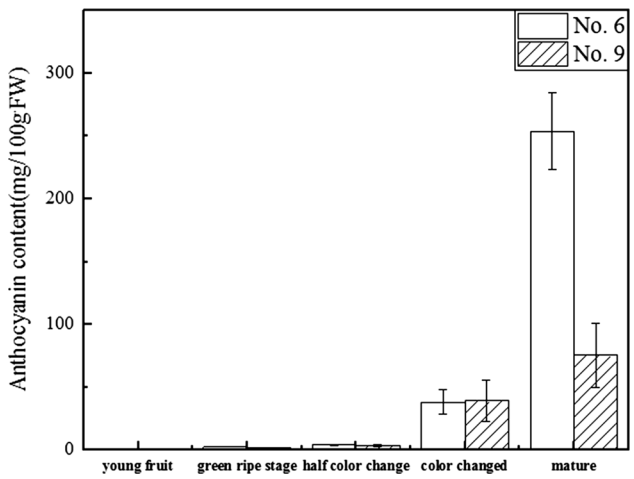

Figure 3. The content of anthocyanins in different periods of 'No. 6' and 'No.9' varieties of *L. edulis*.

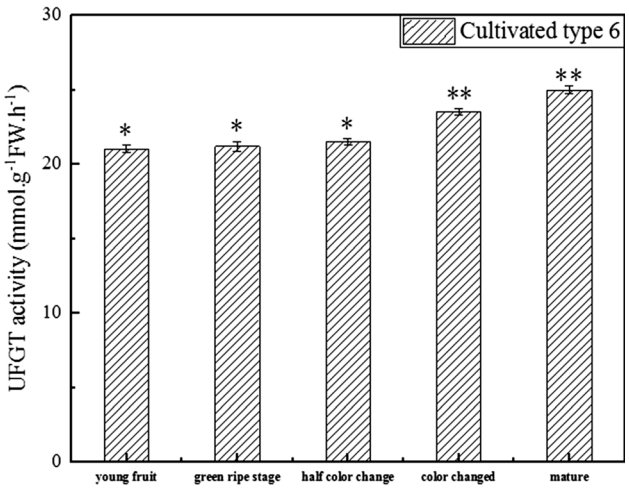

Figure 4. UFGT enzyme activity in different periods of *L. edulis* (No. 6 'Blue Sky').

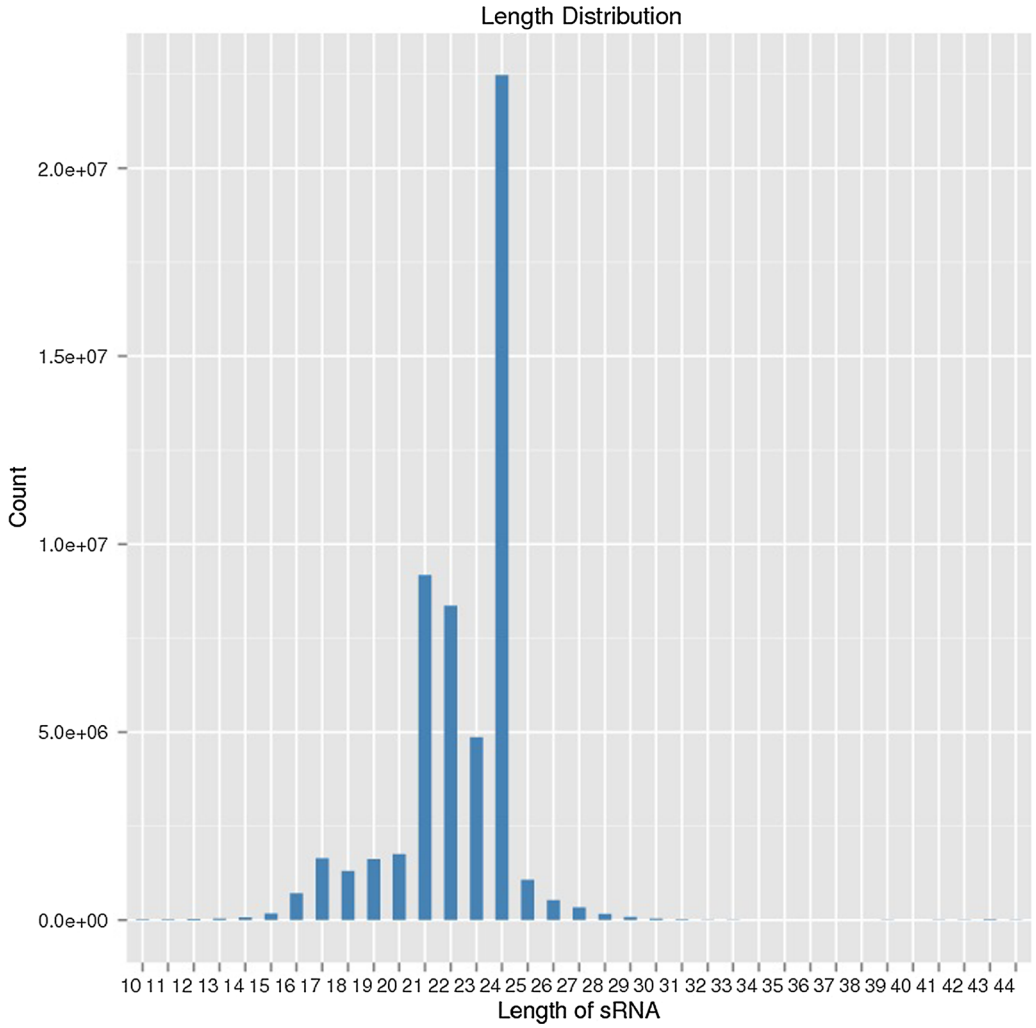

Figure 5. sRNA length range distribution.

miRNAs and novel miRNAs were analysed from valid data obtained from the sRNA group. The distribution of small RNA lengths in the sequencing data was statistically found. The distribution of small RNA lengths in *L. edulis* was concentrated at 18–30 nt (figure 5), which was consistent with the distribution of miRNA length in plants. Further classification of single sequence lengths indicates that 24-nt miRNAs are the predominant type, followed by 21-nt miRNAs. This is consistent with the length distribution of sRNA reported in apples and lychees (Xia *et al.* 2012; Guan *et al.* 2014; Yao *et al.* 2015).

Further classification of single sequence lengths indicates that the 24-nt miRNA is the major type of small RNA. After filtering the data in the classification annotation, the clean tag is compared to a known small RNA database, including miRBase, Rfam, siRNA, piRNA, snoRNA. The genomic alignment of all samples was counted, and 51,819,072 total labels were obtained, with 10,581,923 mapping labels, and the ratio reached 20.42%. Small RNAs are classified and annotated by comparison with known sRNA databases. The proportions of various small RNA classification annotations are shown in figure 6. The results of miRNA classification annotation showed that the proportion of miRNA was 3.4%. It was found that most of the small RNA classification annotations in *L. edulis* were unknown and needed to be further explored.

Base preference analysis of miRNAs in L. edulis

The mature and precursor information of the miRNAs on the alignment is selected by alignment with mature matrices of known miRNAs in the miRbase database. Figure 7a shows the base distribution of small RNA nucleic acids. It can be observed from the figure that G occupies the base position of most sRNAs, followed by C, again A, a small part of U, and G at 24 nt sRNA. The G base distribution ratio is the largest. Subsequent identification of the top base distribution of conserved miRNAs in *L. edulis* is shown in figure 7b. Position-specific base analysis revealed that G is the preferred starting base for these conserved miRNA sequences. Base C was found to occupy the base position of the second

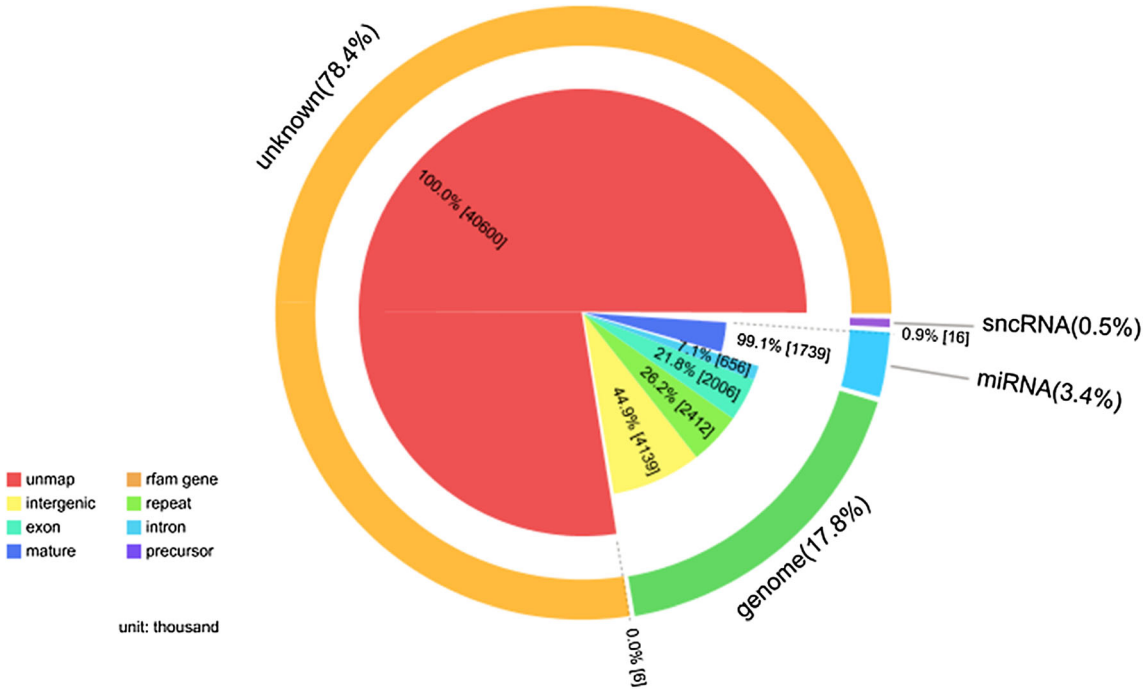

Figure 6. Small RNA classification annotation.

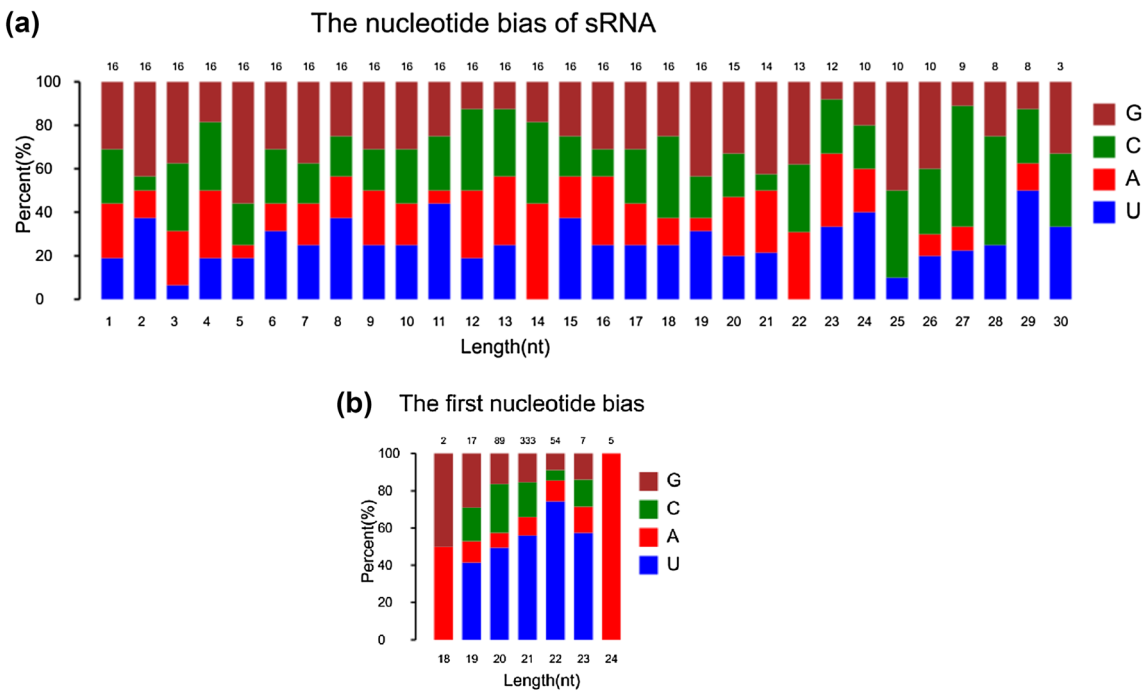

Figure 7. Nucleotide composition of the conserved miRNAs and their bias positions. (a) Nucleotide composition and bias for individual positions from the 5′ end of miRNA sequences. (b) Nucleotide composition and bias for the first position of miRNAs with different length.

of 19 to 23 nt, while A was still abundant at 24 nt. Further analysis of the starting bases of miRNAs with different nucleotide lengths revealed that the 'G' bias was predominantly present in the 18–21 nt sequence, with U predominantly occupying the first basal position in 19–23 nt.

GO analysis of miRNA target genes

Through the analysis of GO, the function of miRNA in *L. edulis* can be divided into three categories: biological process, cell composition and molecular function, and

divided into 50 functional groups. The biological process involves a large proportion of functions such as biological processes, defense responses, transcriptional regulation, and transcriptional regulation of DNA. The cellular components are mainly concentrated in the composition of the nuclear, cytoplasm and plasma membrane. Molecular functions mainly focus on molecular function, protein binding, ATP binding, catalytic functions of DNA, ADP binding, and DNA sequence-specific binding (figure 8). In addition, it is also involved in a variety of biological pathways, such as flavonoid biosynthesis pathways, flower development, mRNA processing, signal transduction, and kinase activity.

Anthocyanins in the *L. edulis* are a kind of water-soluble natural pigment widely present in plants and belong to flavonoids. Flavonoids are secondary metabolites of plants. Enrichment analysis of biological processes and molecular functions represented by GO suggests that these target genes may be involved in a wide range of metabolic pathways and developmental processes. miR159a_1, miR156h-3p and miR396a-3p were screened for anthocyanin-related genes, and their target genes were GAMYB-like gene and flavonoid-3-O-glycosyltransferase. The target genes predicted by miR156h-3p and miR396a-3p are identical. The MYB protein has been identified as a major regulator of the anthocyanin biosynthetic pathway. Flavonoid-3-O-glycosyltransferase is a key gene downstream of the anthocyanin synthesis pathway, which converts anthocyanidin into anthocyanins. High-throughput sequencing of miRNAs in *L. edulis* provides new insights into the role of miRNAs in the synthesis of anthocyanins in *L. edulis*.

Identification of mature miRNAs and prediction of novel miRNAs in L. edulis

The sequences annotated as other small RNAs were removed, and the remaining sequences were homologously aligned with the plant miRNA sequences in the miRBase database. A total of 507 mature miRNAs and 16 novel miRNAs were identified. The mature miRNAs family was found to have the largest number of members of the miR156 family, including 36; followed by the miRNA159 family, containing 25; There are also parts that only identify one family member, such as miR5139 and miR1128. We predicted nine miRNAs directly involved in the regulation of anthocyanin synthesis, such as miR156h-3p, miR159a_1, miR167d-p3, miR394b-p3, miR394a-5p, miR396a-3p, MIR828a-p3, miR5139 and miR1128. Because some of the miRNAs are expressed in *L. edulis* and the accumulation of anthocyanins in the fruit is not high, only some miRNAs with high expression levels are quantified in this study (miR156h-3p and miR396a-3p). The stem–loop structure diagram of the precursor of some novel miRNAs is shown in figure 9 and table 1, and the red colour is indicated as the mature sequence.

Expression analysis of anthocyanin content and UFGT enzyme activity in L. edulis

***Identification of miRNA targets related to anthocyanin synthesis in L. edulis*:** This study identified nine miRNAs associated with anthocyanin synthesis, based on miRNA identification results and target prediction results. qRT-PCR was performed on miRNAs with high expression in *L. edulis*,

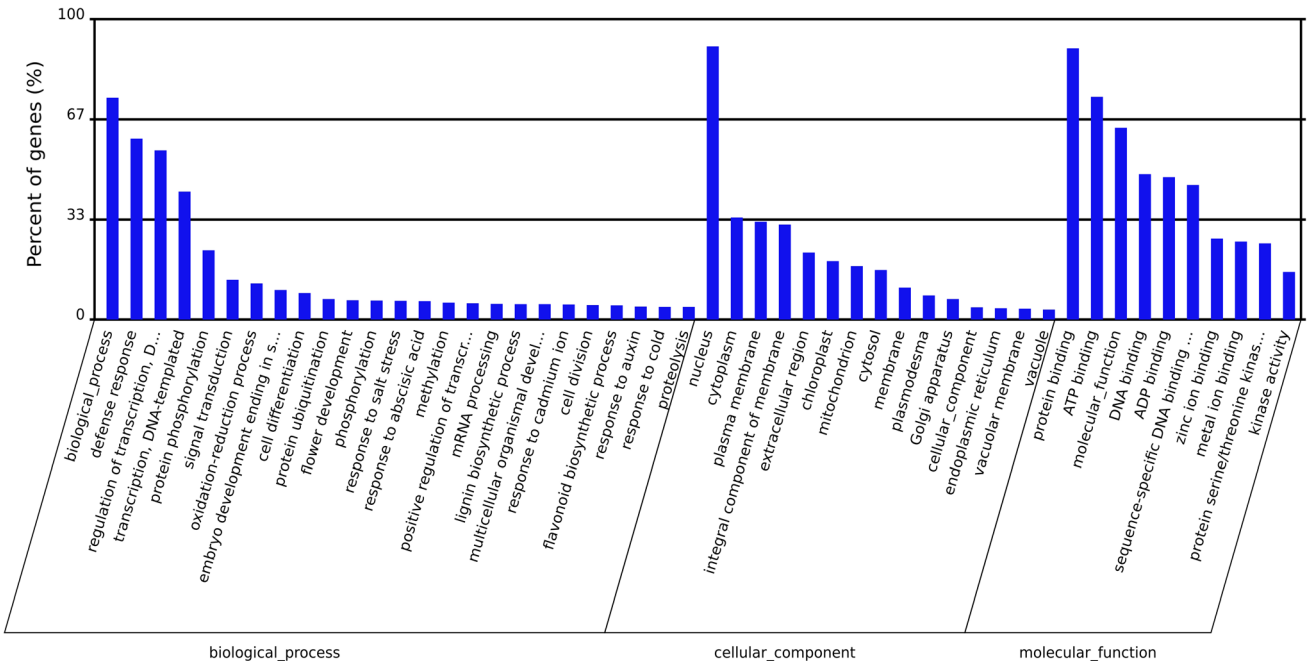

Figure 8. The GO functional annotation of target gene.

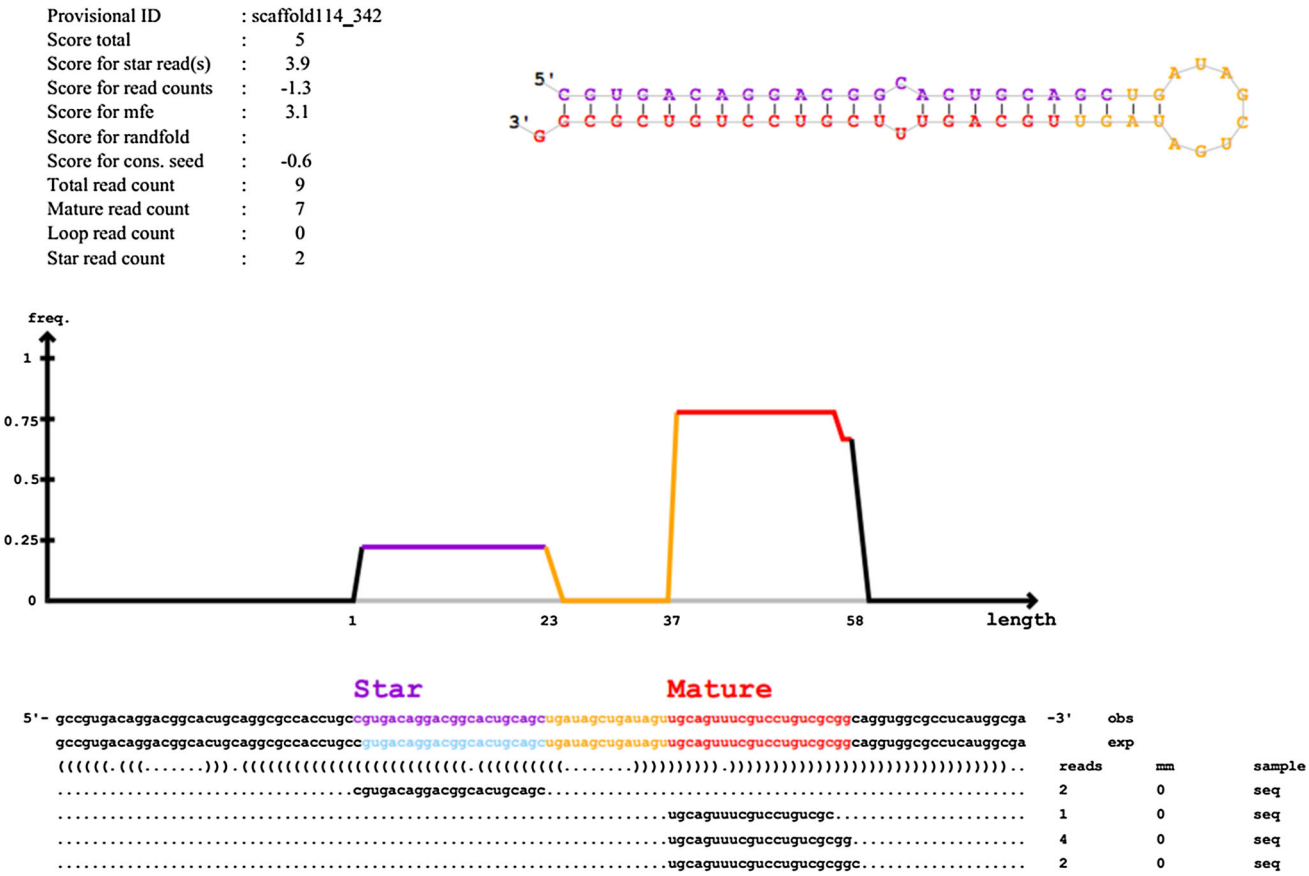

Figure 9. Prediction of the stem–loop structure of some new miRNAs precursors.

namely miR159a_1, miR156h-3p and miR396a-3p. It was found that miR159a_1 targets GAMYB-like and predicts that miR156h-3p and miR396a-3p cotarget flavonol 3-O-glucosyltransferase. This study uses miR159a_1 as an internal reference. Primer sequences can be found in table 2.

***qRT-PCR to verify anthocyanin biosynthesis-related miRNAs and their target genes*:** Our study found that the *UFGT* gene in *L. edulis* can be targeted by miR156h-3p and miR396a-3p, suggesting that miR156h-3p and miR396a-3p may regulate *UFGT* gene expression during anthocyanin synthesis in *L. edulis*. miR156h-3p and miR396a-3p were significantly negatively correlated with target genes during anthocyanin biosynthesis. Through analysis and target gene prediction, it was found that miR156h-3p and miR396a-3p may be negative regulators of anthocyanin biosynthesis. Figure 10 shows the results of analysis of the expression levels of miRNAs involved in anthocyanin synthesis in *L. edulis*. In the figure 10a clearly shows that miR396a-3p has a significant negative regulatory relationship with its target gene UFGT. The expression of miR396a-3p increased gradually from the young fruit stage to the green ripening stage, and gradually decreased from the green ripening stage to the colour-changing stage, and the expression level gradually increased from the colour-changing stage to the mature stage. The *L. edulis* showed partial colour in the half-colouring period, and the anthocyanin began to accumulate, which was consistent with the initial increase of the expression of *UFGT* gene in the half-colouring period. Further, the *UFGT* gene is involved in the biosynthesis of anthocyanins in *L. edulis*, and miR396a-3p plays a negative regulatory role in the anthocyanin synthesis pathway. In the figure 10b clearly shows that miR156h-3p has a significant negative regulatory relationship with its target gene UFGT. The expression of miRNA156h-3p decreased with the accumulation of anthocyanins, which further proved that miRNA156h-3p negatively regulated the accumulation of anthocyanins in *L.* edulis. UFGT is the last key gene downstream of the anthocyanin synthesis pathway, which can convert anthocyanidin into anthocyanins. The target genes predicted by our high-throughput sequencing have an important role in understanding the colouration mechanism and the anthocyanin metabolism of *L. edulis*.

Discussion

High-throughput sequencing of L. edulis

The response of miRNAs to plant development, biotic and abiotic stress is important. The development of high-throughput sequencing technology has identified miRNAs

Table 1. Statistics of sequencing reads.

miRNA id	Chromosome	Strand	Sequence (mature)	Sequence (star)
novel_mir1	NW_015805964.1	+	AGAAGCGTCCTCAGCGGCGGAC	GTTGTTTGGGAATGCAGCC
novel_mir2	NW_015807502.1	+	CTAGGTTGGTGACCTCCATTG	ACTTTGGAGGGGTGCTCGCCTAAGGTC
novel_mir3	NW_015811216.1	+	GCCACGGTAATTCCAGCTCCAATAGCGTA	GTTGTTGCAGTTAAAAAGCTCGTAGTTGG
novel_mir4	NW_015814417.1	+	TTGTTTTAATTAAACAATCGGATTCCCCT	GGAATCCGACTGTTTAATTAAAACAAAGCA
novel_mir5	NW_015814943.1	-	AGAAGGGACGAGACGACCGGTGCACAC	GGACCGGCCGGCCCATCCCAAA
novel_mir6	NW_015841767.1	+	CTCCGAAGTTTCCCTCAGGATAG	CAATGATTAGAGGCATCGGGG
novel_mir7	NW_015851805.1	+	GGTCAGCAATGCTCATGCCGACGCTGGCGTC	CCTGCCAGTAGTCATATGCTTGTCTCAA
novel_mir8	NW_015852265.1	+	AGGCTCGCGCCCAAGGTTTTGCAGCGACC	CGCGCCCTCCTACTCATCGGGGCCTGGC
novel_mir9	NW_015882221.1	+	GAGACCCGGGTTCAAGTCCC	AGTTGGTTAGGATACTCGGCTCTCACCC
novel_mir10	NW_015885490.1	+	TAAGGATTGACAGACTGAGAGCTCTTTCTT	GGCAATAATAAGTCTGTGATGCCCTTAGAT
novel_mir11	NW_015916166.1	+	TTCATGGTTCGATATTCTG	AAAAATAGCTCGACGCCATGATGA
novel_mir12	NW_015921343.1	+	GGGTGTTTGGTCTAGTGGTATGA	ACCACTAGACCACTGGTGCCTT
novel_mir13	NW_015922802.1	+	ATGCGCCCCGGTCGGATGTGGAACGG	ATCCGCTGATCGACTCGGGGCGTGGA
novel_mir14	NW_015923133.1	+	GGCCGTAGTCCCTCTAAGGTACATGGTCG	CCAGCTTCTTAGAGGGACTATGGCCTTT
novel_mir15	NW_015929661.1	+	CTCTCGGCAACGGATATCTCGGCTCTCGC	TGAATTGCAGAATCCCGTGAACCATCGAGT
novel_mir16	NW_015941351.1	+	CGGGGTATTGTAAGTGGCAGAGTGGCCTTG	CATGATCCACTGCGATTCAGCCCTTTGTCG

and their targets in many plant species. However, studies on miRNAs and their targets in *L. edulis* are very limited. Therefore, we completed high-throughput sequencing to fully characterize the regulation of miRNAs during the development of *L. edulis*. However, it is unclear whether miRNA is involved in the biosynthesis of anthocyanin in *L. edulis*, and has not been studied so far. Therefore, this study used high-throughput deep sequencing technology to identify miRNAs and their target genes related to anthocyanin biosynthesis. In this study, experiments have shown that the length of small RNA is most abundant at 24 nt, followed by 21 nt, which is consistent with the length distribution of sRNA reported in apples and litchi (Xia *et al.* 2012; Guan *et al.* 2014; Yao *et al.* 2015). Finally, a total of 507 known miRNAs and 16 new miRNAs were identified in *L. edulis*. In addition, we predicted targets for miRNAs and found that most target genes are conserved among plant species. These miRNAs play an important role in plant development to maturity.

miRNAs and their targets related to L. edulis of Anthocyanins biosynthesis

There is increasing evidence that miRNAs are involved in the biosynthesis of anthocyanins, such as miR858 and miR159 (Jia *et al.* 2015; Wang *et al.* 2017). By analysing the expression profiles of miRNAs at different developmental stages, Zhao (2018) found that mul-miR159a was significantly different in the process of mulberry from green to red, and it is speculated that this miRNA may have a regulatory role in the synthesis of mulberry anthocyanins. To verify the biological function of miR159a in mulberry, the miRNA gene was cloned and its regulation in anthocyanin synthesis was analysed. It was found that mul-miR159a may participate in the regulation of anthocyanin synthesis through the targeting of *Mul-MYB33* gene, and negatively regulate the synthesis and accumulation of anthocyanins. miR858 (Xia *et al.* 2012; Guan *et al.* 2014; Jia *et al.* 2015) has been identified in Arabidopsis, apple, cotton and tomato, and studies by Jia *et al.* (2015) indicate that miRNA858 is a negative regulator of anthocyanin synthesis. miRNA858 regulates the biosynthesis of anthocyanin in tomato by regulating two R2R3-MYB transcription factors. miRNA858 was also present in the sequenced of *L. edulis*, but its quantitative analysis found that its expression was extremely low, so miRNA858 was not the focus of this study.

Wang *et al.* (2017) cloned the MYB transcription factor gene VvGAMYB in a grape gibberellin signaling pathway. Studies have shown that VvGAMYB is a target gene of miR159, its role is regulated by miR159, and the expression of miR159 and VvGAMYB is negatively correlated. VvGAMYB participates in the flowering process of grapes through the gibberellin flowering pathway. In this experiment, the high-throughput sequencing study found that the

Table 2. Primers used in qRT-PCR.

Primer name	Sequence (5′→3′)
miR395a_5-1F	CGCTGAAGTGTTTGGGGG
miR395a_5-1R	CAGTGCAGGGTCCGAGGTAT
miR6300-1F	AGCGGGCGTCGTTGTAGTAT
miR6300-1R	CGCAGGGTCCGAGGTATTC
miR396a-3p_5-1F	AGCGGTTCAATAAAGCTGTGG
miR396a-3p_5-1R	CGCAGGGTCCGAGGTATTC
miR397a-1F	ACGGGCTCATTGAGTGCAG
miR397a-1R	CGCAGGGTCCGAGGTATTC
miR397b-1F	TCAGCCCATTGAGTGCAGC
miR397b-1R	CGCAGGGTCCGAGGTATTC
miR160a-5p-1F	TTCTGCCTGGCTCCCTGTA
miR160a-5p-1R	CAGTGCAGGGTCCGAGGTAT
miR156h-3p-1F	GTTTGTGTCGTATCTCGGGCA
miR156h-3p-1R	CGCAGGGTCCGAGGTATTC
miR159a_1-1F	GTGCCATTTGGATTGAAGGG
miR159a_1-1R	CAGTGCAGGGTCCGAGGTAT
UFGT-1F	GCCCCAGGAGGATATTGAGC
UFGT-1R	CTGCCATATCTGCGGCAAAC

target gene of miRNA159a_1 in *L. edulis* is GAMYB-like gene. Based on the above studies, it is predicted that miRNA159a_1 may be involved in the growth and development of *L. edulis*. The German research group (Gou *et al.* 2011) found that the expression of anthocyanin in plants expressing miR156 increased, and the expression of genes required for anthocyanin synthesis (DFR, F3'H, ANS, UGT75C1) was upregulated, further confirming that miR156 plays an important role in plant of anthocyanin biosynthesis.

qRT-PCR verification of growth and development-related miRNAs in L. edulis

To reveal the expression pattern of anthocyanin biosynthesis-related miRNAs in *L. edulis*, eight miRNAs, namely miR160a-5p, miR6300, miR395a_5, miR396a-3p, miR397a, miR397b, miR159a_1 and miR156h-3p were selected for RT-qPCR verification. Primer sequences are shown in table 2. The target of miRNA in plants is mostly an important transcription factor in plant life activities, so miRNA plays an important role in plant development and stress response. The target of miR160 is the ARF family transcription factor, and miR160 can regulate bud differentiation and root development by inhibiting ARF10, ARF16, ARF17 (Wang *et al.* 2005). Gao *et al.* (2017) found that miR6300 was present in lily by high-throughput sequencing and degradation group analysis. miR6300 has rarely been reported in other species. This study found that miR6300 is present in *L. edulis*, and its regulation needs to be further explored. miR395 was found in *A. thaliana*, miR395 is thought to target mRNA encoding ATP thiourease. miR395 is upregulated in *Arabidopsis* during sulphate-limited

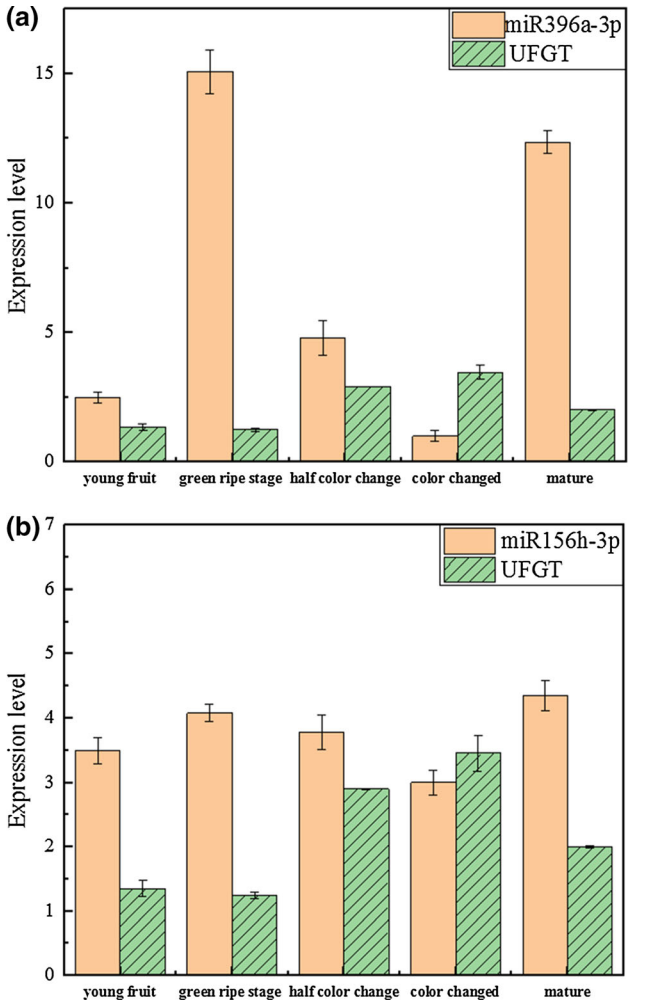

Figure 10. Relative expression of miRNA and its target genes related to anthocyanin synthesis in *L. edulis*: (a) Expression analysis of miR396a-3p and its target gene UFGT in *L. edulis*. (b) Expression analysis of miR156h-3p and its target gene UFGT in *L. edulis*.

conditions, when the mature miRNA then regulates sulphur transporters and ATP sulphurylases, and participate in the pathway of sulphate metabolism (Liang *et al.* 2010). Tang *et al.* (2018) conducted a number of studies on *Arabidopsis* and rice, and found that the target of miR396 is growth regulator GRFs. OsmiR396 and its target OsGRF in rice can regulate leaf development, plant height, meristem function, flowering time, inflorescence structure and seed size (Kuijt *et al.* 2014; Liu *et al.* 2014; Che *et al.* 2015; Duan *et al.* 2015; Gao *et al.* 2015; Luo *et al.* 2005). Zhao *et al.* (2015) found that miR397 is involved in the specific expression of tomato root, flower, mature fruit and callus. These results suggested that Le LACmiR397 was related to resistance to Syringic acid and Sinapic acid in tomato. Zhao (2018) used mulberry in different developmental stages as experimental materials to analyse the expression changes of miRNAs in mulberry at different developmental stages by using small

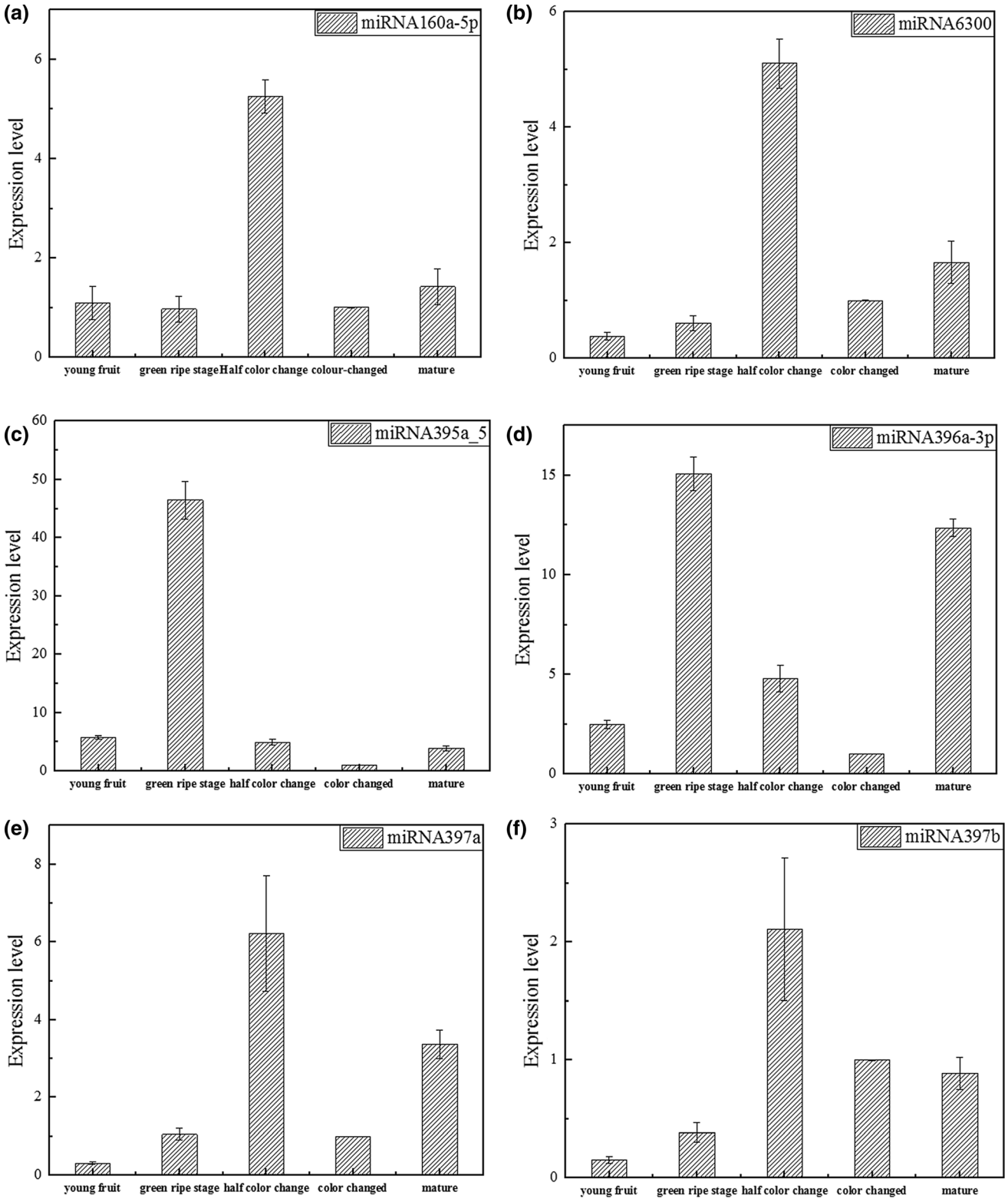

Figure 11. Expression of miRNAs related to growth and development in *L. edulis*.

RNA digital analysis technology based on Illumina high-throughput sequencing. The miR159 involved in the synthesis of mulberry anthocyanins was identified, and it was found that mul-miR159a can play a negative regulatory role in the biosynthesis of anthocyanins by targeting the Mul-MYB33 gene.

Gou *et al.* (2011) found that the negative regulation of anthocyanin biosynthesis in *Arabidopsis* by a miR156-

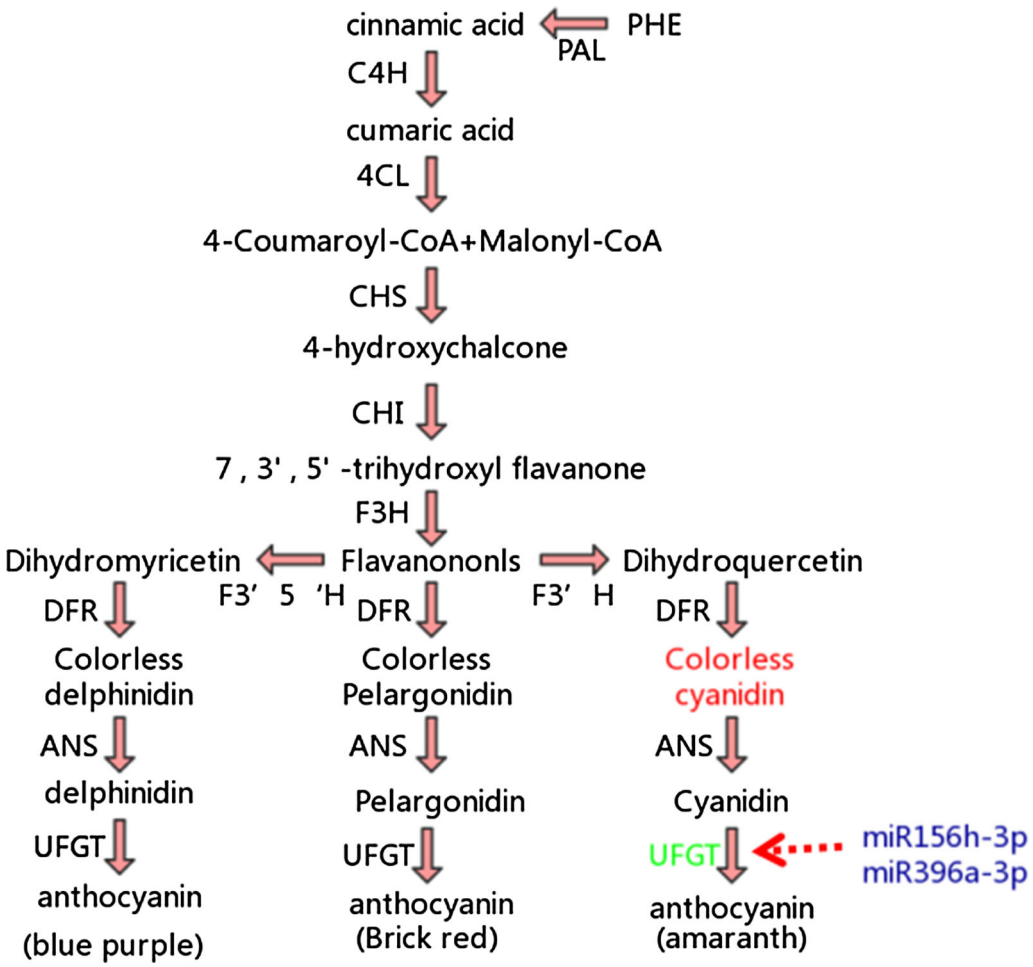

Figure 12. The enzyme-coding structural genes involved in anthocyanin biosynthetic pathway associated with the putative regulated miRNAs and their expression levels. PAL, phenylalanine ammonia-lysae; C4H, cinnamate 4-hydoxylase; 4CL, 4-coumar-ate CoA ligase; CHS, chalcone synthase; CHI, chalcone isomerase; F3H, flavanone 3-hydroxylase; F3′H, flavonoid 3′-hydroxylase; F3′5′H, flavonoid-3′, 5′-hydroxylase; DFR, dihydroflavonol 4-reductase; ANS, anthocyanidin synthase; 3GT, UDP-glucose-flavonoid 3-O-glucosyltransferase.

targeted SPL transcription factor, the figure below shows the miRNA expression map. miR160-5p was gradually upregulated from the young fruit stage to the half-colour change stage, and was the highest in the half-turn colour stage, and then upregulated after the half-turn colour stage to the mature stage (figure 11a). The expression level of miR6300 increased significantly from the young fruit stage to the half-colour change stage, and it was the highest in the half-transfection stage. The *L. edulis* began to colour in the half-colouring stage, and the expression of miR6300 gradually decreased from the half-colouring stage to the colour-changing stage, and the expression level of miR6300 gradually increased from the colour-changing stage to the mature stage (figure 11b). We speculate that miR160-5p and miR6300 play important roles in the development of *L. edulis*. The development of *L. edulis* plays an important role in the accumulation of anthocyanins. The anthocyanins of *L. edulis* began to accumulate in the half-colouring stage, and the fruit began to colour. The expression trends of miR395a_5 and miR396a-3p were consistent in *L. edulis*. The expression trends of miR395a_5 and miR396a-3p were consistent. The expression of sorghum fruit gradually increased from the young fruit stage to the green ripe stage, and the expression level of miRNA gradually decreased from the green ripening stage to the colour change stage. The expression of miR395a_5 was significantly higher than that of miR396a-3p (figure 11, c&d). The miR397 family members miR397a and miR397b gradually increased from the young fruit stage to the half colour stage and reached the highest in the half colour period of the *L. edulis*. The expression of miR397a gradually decreased from the half colour change period to the colour changed period of *L. edulis*, and gradually increased from the colour-changing stage to the maturity stage. The expression level of miR397b gradually decreased from the half colour stage to the maturity stage of *L. edulis* (figure 11, e&f). The anthocyanin of the *L. edulis* began to accumulate in the half-transfection phase and miR397a and miR397b accumulated to the highest in the half-transfection phase. It is further speculated that miR397a and miR397b may act indirectly on some

biologically active molecules to regulate the accumulation of anthocyanins. We speculate that some growth regulators that control plant growth may promote fruit colouration during the growth and development of sassafras fruit, which affects the accumulation of anthocyanins.

In general, genes involved in anthocyanin biosynthesis have been extensively characterized (Gandía-Herrero and García-Carmona 2013). Anthocyanin synthesis related genes can be divided into two types: structural genes encoding anthocyanin biosynthesis enzymes and regulatory genes controlling the expression of structural genes. In addition to regulating anthocyanin biosynthesis by cutting transcription factors to affect the expression of structural genes, miRNAs can also directly inhibit the expression of structural genes. As shown in figure 12, structural genes involved in the anthocyanin biosynthesis pathway include phenylalanine ammonia-lysae (PAL), cinnamate 4-hydoxylase (C4H), 4-coumar-ate CoA ligase (4CL), chalcone synthase (CHS), chalcone isomerase (CHI), flavanone 3-hydroxylase (F3H), flavonoid 3′-hydroxylase (F3′H), flavonoid-3′, 5′-hydroxylase (F3′ 5′H), flavonol synthase (FLS), dihydroflavonol 4-reductase (DFR), anthocyanin synthase (ANS) and UDP-glucose-flavonoid 3-O-glucosyltransferase (3GT).

In the present study, one structural gene was found to be targeted by two miRNAs with different expression levels (figure 12), namely miR396a-3p and miR156h-3p. The relationship between the two miRNAs and the target gene was found to be negatively regulated by the expression of the gene. The target gene UFGT was expressed at a low level from young fruit stage to green ripening stage. At this time, the *L. edulis* was green and did not accumulate anthocyanins.

Pigment accumulation began from the *L. edulis* in the half color change stage, and the expression of UFGT gene was upregulated, which promoted the accumulation of anthocyanins. It was further speculated that UFGT plays an important role in the accumulation of anthocyanins in *L. edulis*, while miR396a-3p and miR156h-3p affect the accumulation of anthocyanins by negatively regulating UFGT.

Acknowledgements

This research was funded by National Natural Science Foundation Project and Heilongjiang Provincial Financial Forestry Science and Technology Promotion Fund, grant number 31571731 and 2015ST-11. Thanks to Prof. Li Binsheng of Heilongjiang Forest Botanical Garden for providing the resource, experimental sample planting, cultivation, management, picking; and thanks to the sequencing support of Huada Gene.

References

An X. H., Tian Y., Chen K. Q., Wang X. F. and Hao Y. J. 2012 The apple WD40 protein MdTTG1 interacts with bHLH but not MYB proteins to regulate anthocyanin accumulation. *J. Plant Physiol.* **169**, 710–717.

Che R., Tong H., Shi B., Liu Y., Fang S., Liu D. *et al.* 2015 Control of grain size and rice yield by GL2-mediated brassino steroid responses. *Nat. Plants* **2**, 15195.

Chiu L. W., Zhou X., Burke S., Wu X., Prior R. L. and Li L. 2010 The purple cauliflower arises from activation of a MYB transcription factor. *Plant Physiol.* **154**, 1470–1480.

Christopher D. 2013 Trans-acting small interfering RNA4: key to nutraceutical synthesis in grape development? *Trends Plant Sci.* **18**, 601–610.

Dong Q., Fei Y., Rui M., Xiaobing J., Huijuan Y., Ziyi G. *et al.* 2016 Identification of microRNAs and their targets associated with fruit-bagging and subsequent sunlight re-exposure in the 'Granny Smith' apple exocarp using high-throughput sequencing. *Front Plant Sci.* **7**, 27.

Duan P., Ni S., Wang J., Zhang B., Xu R., Wang Y. *et al.* 2015 Regulation of OsGRF4 by OsmiR396 controls grain size and yield in rice. *Nat. Plants* **2**, 15203.

Gandía-Herrero F and García-Carmona F. 2013 Biosynthesis of betalains: yellow and violet plant pigments. *Trends Plant Sci*. **18**, 334–343.

Gao F., Wang K., Liu Y., Chen Y., Chen P., Shi Z. *et al.* 2015 Blocking miR396 increases rice yield by shaping inflorescence architecture. *Nat. Plants* **2**, 15196.

Gao X, Cui Q, Cao Q. Z, Liu Q, He H. B, Zhang D. M *et al.* 2017 Transcriptome-wide analysis of *Botrytis elliptica* responsive microRNAs and their targets in Lilium Regale Wilson by high-throughput sequencing and degradome analysis. *Front. Plant Sci.* **8**, 753.

Gonzalez A., Zhao M., Leavitt J. M. and Lloyd A. M. 2008 Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in *Arabidopsis* seedlings. *Plant J.* **53**, 814–827.

Gou J. Y., Felippes F. F., Liu C. J., Weigel D. and Wang J. W. 2011 Negative regulation of anthocyanin biosynthesis in *Arabidopsis* by a miR156-targeted SPL transcription factor. *Plant Cell* **23**, 1512–1522.

Guan X., Pang M., Nah G., Shi X., Ye W., Stelly D. M. *et al.* 2014 miR828 and miR858 regulate homoeologous MYB2 gene functions in *Arabidopsis trichome* and cotton fibre development. *Nat. Commun.* **5**, 3050.

Jia X., Shen J., Liu H., Li F., Ding N., Gao C. *et al.* 2015 Small tandem target mimic-mediated blockage of microRNA858 induces anthocyanin accumulation in tomato. *Planta* **242**, 283–293.

Kong J. M., Chia L. S., Goh N. K., Chia T. F. and Brouillard R. 2003 Analysis and biological activities of anthocyanins. *Phytochemistry (Amsterdam)* **64**, 923–933.

Kuijt S. J. H., Greco R., Agalou A., Shao J., "t Hoen C. C. J., Overnas E. *et al.* 2014 Interaction between the growth-regulating factor and knotted1-like homeobox families of transcription factors [W]. *Plant Physiol.* **164**, 1952–1966.

Li J. C., Li T. Z., Wang Z. G. and Li B. J. 2010 Cloning and expression of UFGT gene in skin of Max Red Bartlett. *Acta Botanica Boreali-Occidentalia Sinica* **30**, 30–34.

Li Z., Yu J. J., Peng Y. and Huang B. R. 2016 Metabolic pathways regulated by abscisic acid, salicylic acid, and γ-aminobutyric acid in association with improved drought tolerance in creeping bentgrass (agrostis stolonifera). *Physiol. Plantarum* **159**, 42.

Liang G., Yang F. and Yu D. 2010 Microrna395 mediates regulation of sulfate accumulation and allocation in *Arabidopsis thaliana*. *Plant J.* **62**, 1046–1057.

Liu H., Guo S., Xu Y., Li C., Zhang Z., Zhang D. *et al.* 2014 OsmiR396d-regulated OsGRFs function in floral organogenesis in rice through binding to their targets OsJMJ706 and OsCR4. *Plant Physiol.* **165**, 160–174.

Liu P., Huo J. W., Wang Y., Qin D. and Zhao L. J. 2016 De novo transcriptome sequencing of blue honeysuckle fruit (*Lonicera caerulea* L.) and analysis of major genes involved in anthocyanin biosynthesis. *Acta Physiol. Plant.* **38**, 180.

Luo A. D., Liu L., Tang Z. S., Bai X. Q., Cao S. Y. and Chu C. C. 2005 Down-regulation of OsGRF1 gene in rice rhd1 mutant results in reduced heading date. *J. Integr. Plant Biol.* **47**, 8.

Rock C. D. 2013 Trans-acting small interfering rna4: key to nutraceutical synthesis in grape development? *Trends Plant Sci.* **18**, 601–610.

Sun Y., Qiu Y., Duan M., Wang J., Zhang X., Wang H. *et al.* 2016 Identification of anthocyanin biosynthesis related microRNAs in a distinctive chinese radish (*Raphanus sativus* L.) by high-throughput sequencing. *Mol. Genet. Genomics* **292**, 1–15.

Tang Y., Liu H., Guo S., Wang B., Li Z., Chong K. *et al.* 2018 OsmiR396d affects gibberellin and brassino steroid signaling to regulate plant architecture in rice. *Plant Physiol.* **176**, 946.

Wang C. J., Wang J. M., Lin W. L., Chu C. Y., Chou F. P. and Tseng T. H. 2000 Protective effect of hibiscus anthocyanins against tert-butyl hydroperoxide-induced hepatic toxicity in rats. *Food Chem. Toxicol*. **38**, 411–416.

Wang J. W., Wang L. J., Mao Y. B., Cai W. J., Xue H. W. and Chen X. Y. 2005 Control of root cap formation by microRNA-targeted auxin response factors in *Arabidopsis. Plant Cell.* **17**, 2204–2216.

Wang M. Q., Xie Z. P., Sun X., Li X. P., Zhu X. D., Wang C. et al. 2017 Function analysis of miR159 and its target gene VvGAMYB in grape flower development. *Acta Horticult. Sinica*. https://doi.org/10.16420/j.issn.0513-353x.2016-0816.

Xia R., Zhu H., An Y. Q., Beers E. P. and Liu Z. 2012 Apple miRNAs and tasiRNAs with novel regulatory networks. *Genome Biol.* **13**, Article R47.

Xie K., Shen J., Hou X., Yao J., Li X., Xiao J. *et al.* 2012 Gradual increase of miR156 regulates temporal expression changes of numerous genes during leaf development in rice. *Plant Physiol.* **158**, 1382–1394.

Yao F., Zhu H., Yi C., Qu H. and Jiang Y. 2015 MicroRNAs and targets in senescent litchi fruit during ambient storage and post-cold storage shelf life. *BMC Plant Biol.* **15**, 181.

Yoshimoto M., Okuno S., Yoshinaga M., Yamakawa O. and Yamada J. 1999 Antimutagenicity of sweetpotato (*Ipomoea batatas*) roots. *Biosci. Biotech. Bioch*. **63**, 537–541.

Yuan Y., Wang Z., Jiang C., Wang X. and Huang L. 2014 Exploiting genes and functional diversity of chlorogenic acid and luteolin biosyntheses in *Lonicera japonica* and their substitutes. *Gene* **534**, 408–416.

Yue J. Y., Lu X. H., Zhang H., Ge J., Gao X. L. and Liu Y. S. 2017 Identification of conserved and novel microRNAs in blueberry. *Front. Plant Sci.* **8**, 1155.

Zhang Y. T., Hu W. J., Peng X. R., Sun B., Wang X. R. and Tang H. R. 2018 Characterization of anthocyanin and proanthocyanidin biosynthesis in two strawberry genotypes during fruit development in response to different light qualities. *J. Photoch. Photobio. B.* **186**, 225–231.

Zhao H. N. 2018 *Identification and regulation functions of miRNAs involved in anthocyanin synthesis in mulberry.* Shandong Agricultural University, Shandong.

Zhao X. Y., Pang M. L., Zhao Q., Ren Y. R., Hao Y. J. and You C. X. 2015 Cloning and expression analysis of tomato LeLAC (mir397) gene. *Acta Horticul. Sinica* **42**, 1285–1298.

Corresponding editor: Manoj Prasad