A comparative analysis of small RNA sequencing data in tubers of purple potato and its red mutant reveals small RNA regulation in anthocyanin biosynthesis

Fang Liu Equal first author, [1], **Peng Zhao** Equal first author, [1], **Guangxia Chen** [1], **Yongqiang Wang** [1], **Yuanjun Yang** Corresp. [1]

[1] Institute of Vegetables, Shandong Academy of Agricultural Sciences, Jinan, China

Corresponding Author: Yuanjun Yang
Email address: yangyuanjun@263.net.cn

Anthocyanins are a group of natural pigments acting as stress protectants induced by biotic/abiotic stress in plants. Although the metabolic pathway of anthocyanin has been studied in potato, the roles of miRNAs on the metabolic pathway remain unclear. In this study, a purple tetraploid potato of SD92 and its red mutant of SD140 were selected to explore the regulation mechanism of miRNA in anthocyanin biosynthesis. A comparative analysis of small RNAs between SD92 and SD140 revealed that there were 179 differentially expressed miRNAs, including 65 up- and 114 down-regulated miRNAs. Furthermore, 31 differentially expressed miRNAs were predicted to potentially regulate 305 target genes. KEGG pathway enrichment analysis for these target genes showed that plant hormone signal transduction pathway and plant-pathogen interaction pathway were significantly enriched. The correlation analysis of miRNA sequencing data and transcriptome data showed that there were 140 negative regulatory miRNA-mRNA pairs. The miRNAs included miR171 family, miR172 family, miR530b_4 and novel_mir170. The mRNAs encoded transcription factors, hormone response factors and protein kinases. All these results indicated that miRNAs might regulate anthocyanin biosynthesis through transcription factors, hormone response factors and protein kinase.

A comparative analysis of small RNA sequencing data in tubers of purple potato and its red mutant reveals small RNA regulation in anthocyanin biosynthesis

Fang Liu[1*], Peng Zhao[1*], Guangxia Chen[1], Yongqiang Wang[1], Yuanjun Yang[1]

[1]Institute of Vegetables, Shandong Academy of Agricultural Sciences, Shandong Key Laboratory of Greenhouse Vegetable Biology, Shandong Branch of National Vegetable Improvement Center, Jinan, Shandong, China.
[*]The authors contributed equally to the work.

Corresponding Author:
Yuanjun Yang
Gongye North Road 23788, Jinan, Shandong, 250100, China
Email address: yangyuanjun@263.net.cn

Abstract

Anthocyanins are a group of natural pigments acting as stress protectants induced by biotic/abiotic stress in plants. Although the metabolic pathway of anthocyanin has been studied in potato, the roles of miRNAs on the metabolic pathway remain unclear. In this study, a purple tetraploid potato of SD92 and its red mutant of SD140 were selected to explore the regulation mechanism of miRNA in anthocyanin biosynthesis. A comparative analysis of small RNAs between SD92 and SD140 revealed that there were 179 differentially expressed miRNAs, including 65 up- and 114 down-regulated miRNAs. Furthermore, 31 differentially expressed miRNAs were predicted to potentially regulate 305 target genes. KEGG pathway enrichment analysis for these target genes showed that plant hormone signal transduction pathway and plant-pathogen interaction pathway were significantly enriched. The correlation analysis of miRNA sequencing data and transcriptome data showed that there were 140 negative regulatory miRNA-mRNA pairs. The miRNAs included miR171 family, miR172 family, miR530b_4 and novel_mir170. The mRNAs encoded transcription factors, hormone response factors and protein kinases. All these results indicated that miRNAs might regulate anthocyanin biosynthesis through transcription factors, hormone response factors and protein kinase.

Introduction

Anthocyanins are flavonoid compounds, which are secondary metabolites. They are natural food pigments found in edible parts of fruits, vegetables and crops (Chiu et al. 2010). The foods rich in anthocyanin present bright colors and are popular with people (Bimpilas et al. 2016). Moreover, anthocyanins also have antioxidant activity and can protect human beings from disease or reduce the damage of disease. The anthocyanin extracts from purple rice protect cardiac function in STZ-induced diabetes rat hearts by inhibiting cardiac hypertrophy and fibrosis (Chen et al. 2016). Anthocyanins from red potato show anti-hepatotoxity in rats with toxicity of D-galactosamine (Han et al. 2006). Anthocyanin extracts from bilberries and blackcurrants have protective activity on acute acetaminophen-induced hepatotoxicity in rats (Cristani et al. 2016).

In anthocyanin biosynthesis, phenylalanine is a primary precursor. Then under the action of a series of enzymes, the substances of coumaroyl CoA, dihydroflavonols, leucoanthocyanins and anthocyanins are successively produced. Anthocyanin biosynthesis is regulated by structural genes and their transcription factors. Some genes regulating anthocyanin biosynthesis have been isolated and characterized in potato, such as *f3'5'h* (Jung et al. 2005), *dfr* (De Jong et al. 2003), *developer* (*D*) locus (Jung et al. 2009), *AN1* (D'Amelia et al. 2014) and *StMYB44* (Liu et al. 2019).

Small RNAs usually consist of 20-30 nucleotides and widely exist in eukaryotic organisms. According to their biogenesis modes, small RNAs are distinguished into three major types, namely miRNA, siRNA and piRNA (Axtell 2013; Chen 2009). Small RNAs guide biological processes at DNA or RNA level, for example, the cleavage of complementary RNAs. Different types of small RNAs have similar molecular functions. Both miRNAs and siRNAs can inhibit translation of target mRNAs, and both siRNAs and piRNAs can direct chromatin modifications (Chen 2009). miRNAs regulate target mRNAs through transcript cleavage and/or translational inhibition. Conserved miRNAs play vital roles in many plant physiological processes, such as development, stress responses, primary and secondary metabolism (Gou et al. 2011; Jones-Rhoades et al. 2006; Matzke et al. 2009; Xia et al. 2012).

So far, miRNAs have been proved to be involved in the regulation of anthocyanin biosynthesis. miRNA858a and HYPOCOTYL 5 (HY5) can repress the expression of *MYB-LIKE 2* (*MYBL2*), thus leading to the activation of anthocyanin biosynthesis pathway (Wang et al. 2016). Increasing miR156 activity promotes anthocyanin accumulation, while reducing miR156 activity leads to a high level of flavonol (Gou et al. 2011). Both miR828 and miR858 regulate *VvMYB114* to promote anthocyanin biosynthesis in grapes (Tirumalai et al. 2019). The miRNA involved in anthocyanin biosynthesis pathway are also reported in apple (Hu et al. 2021), tomato

(Jia et al. 2015), potato (Bonar et al. 2018) and kiwifruit (Li et al. 2019). However, there are few studies on the post-transcriptional regulation of miRNA in potato anthocyanin biosynthesis. In the study, a comparative miRNA analysis and the expression analysis of miRNA-mRNA were performed between purple flesh potato, SD92, and its red flesh mutant, SD140. These results will shed light on the regulation mechanism of miRNA in potato anthocyanin biosynthesis .

Materials & Methods

Plant materials

SD92, commonly known as Hei Jingang, was a tetraploid potato with purple skin and purple flesh. SD140 is a mutant of SD92. Its skin and flesh colors were red (Liu et al. 2018; Liu et al. 2015). Two materials were planted in a greenhouse for two months at 20 ± 2 °C with a photoperiod of 16 h light/8 h dark.

Sample library construction and sequencing

Fresh tubers (diameter 4-5 cm) from three individual plants were harvested for three biological replicates, cleaned with sterilized water, frozen in liquid nitrogen and finally stored at -80˚C. Total RNA extraction of the samples was performed with a modified Trizol reagent (Liu et al. 2018) for library construction and validation of miRNA sequencing data.

Small RNA was isolated and the library was constructed in accordance with the protocol of Preparing Samples for Analysis of Small RNA (Illumina, USA). The 18-30 nt RNA segments were separated from total RNA by polyacrylamide gel electrophoresis, then ligated with 3’ adaptor (GAACGACATGGCTACGATCCGACTT) and 5’ adaptor (AGTCGGAGGCCAAGCGGTCTTAGGAAGACAA). The resulting segments were employed to synthesize first-strand cDNA. The cDNA was amplified and only cDNA with both 3’ and 5’ adaptors was enriched. Finally, the fragments of 100~120 bp were separated to constructe the library. After library quantification and single-stranded DNA cyclization, the library was sequenced by BGISEQ-500 technology. The raw data was deposited into NCBI BioProject database (PRJNA824931).

miRNA identification and prediction

The impurities of raw data, including 5’ primer contaminants, no-insert tags, oversized insertion tags, low quality tags, poly-A tags and the tags without 3’ primer, were excluded from the raw data to obtain clean tags. Low-quality tags were tags whose base quality values were less than 20, accounting for more than 50% of the total bases. The clean tags were mapped to potato reference genome PGSC_DM v4.03 (http://solanaceae.plantbiology.msu.edu/data) by Bowtie2 (Langmead et al. 2009) and small RNA databases miRBase (Kozomara & Griffiths-Jones 2014), snoRNA (Yoshihama et al. 2013) and Rfam (Nawrocki et al. 2015). If a small RNA could be mapped to more than one database, the small RNA annotation followed the searching priority of

miRBase > snoRNA > Rfam. One small RNA was only mapped to one category. The small RNAs mapped to Rfam database were validated by cmsearch (Nawrocki & Eddy 2013). The novel miRNA was determined by miRA (Evers et al. 2015) according to the characteristic hairpin structure of miRNA precursor. Small interfering RNA (siRNA), a 22-24 nt double-strand RNA, was identified by the characteristic of one strand 2 nt shorter than the other (Jagla et al. 2005).

miRNA expression and screening of differentially expressed miRNAs (DEMs)

The expression level of miRNA was estimated by the transcripts per kilobase million (TPM) ('t Hoen et al. 2008). The differential expression was calculated by DEGseq (Wang et al. 2010) based on MA-plot method (Yang et al. 2002). The P-value calculated for each gene was adjusted to Q-value for multiple testing corrections by two alternative strategies. The miRNAs with expression fold change > 2 and Q-value < 0.001 were defined as differentially expressed miRNAs. The volcano plot and heatmap of differentially expressed miRNAs were obtained by Excel 2016 and MeV (Saeed et al. 2003), respectively.

Target gene prediction, Gene Ontology (GO) and KEGG pathway enrichment analyses

TargetFinder (Fahlgren & Carrington 2010) and psRobot (Wu et al. 2012) were used to predict the target genes of miRNAs. All target genes were mapped to GO-terms in the database (http://www.geneontology.org/) and KEGG Orthology (Kanehisa et al. 2008) pathways. The enrichment analyses of GO terms and KEGG pathways were performed by the hypergeometric test based on GO::TermFinder (Boyle et al. 2004). The P-values were adjusted by Bonferroni method (Abdi 2007). The adjusted P-value was defined as Q-value. The terms with Q-value < 0.05 were defined as significantly enriched terms.

Expression validation of miRNAs

RNAs were digested by DNaseI (Thermo, USA) to remove genome DNA. First-strand cDNA was synthesized by miRNA First Strand cDNA Synthesis Kit (Sangon Biotech, China) using tailing reaction method. Real-time quantitative PCR (RT-qPCR) was performed with UltraSYBR Mixture Kit (CWBIO, China) by using *18S rRNA* (GenBank: X67238.1) as a reference gene. The primers of *18S rRNA* and miRNAs were listed in Table 1. The universal reverse primer for miRNAs was supplied from miRNA First Strand cDNA Synthesis Kit. Three biological replicates were performed. Significant difference of miRNA expression between SD92 and SD140 was identified by student's t-test ($p < 0.05$).

Results

Sequencing and classification of potato small RNAs

To identify the miRNAs regulating potato flesh color, six small RNA libraries were constructed and sequenced. The counts of raw tags of six libraries ranged from 28,058,311 to

30,152,601 (Table 2). Low quality tags, invalid adapter tags, poly-A tags and short valid length tags (shorter than 18 nt) were removed to obtain clean tags. The percentages of clean tags of six libraries ranged from 92.10% to 95.22%, which indicated the sequencing data could be used for subsequent analyses. Of the six libraries, 19-25 nt length tags accounted for 87.9% - 96.4% of the total tags, and the 24 nt tags were the most abundant (Table S1). More than 85.04% of the total clean tags from six libraries were mapped to the reference genome (Table S2). Therefore, the sequencing data should accurately reflect small RNA expression and could be used for differential expression analysis of small RNA. To classify and annotate small RNAs, the clean tags were mapped to small RNA databases miRBase, snoRNA and Rfam. The types and proportion of identified small RNAs were similar within six libraries. The intergenic miRNAs were the most abundant (Table S3).

Identification of known and novel miRNAs

There were about 300 known miRNAs and 160 novel miRNAs identified in every library (Table 3). In total, 356 known miRNAs belonging to 121 miRNA families were identified (Table S4), and miR172 family was the most abundant family where 21 members were identified. The nucleotide bias analyses on these non-redundant known miRNAs (Fig. S1A) showed that the first and 24th nucleotide preferred to be uracil (U), and adenine (A) was the dominant nucleotide in the 10th nucleotide position. Meanwhile, several nucleotide positions were conserved. For example, the proportions of four kinds of nucleotides were nearly equal in the 4th, 9th and 16th nucleotide position (Fig. S1A).

Unmapped tags were further used to predict novel small RNAs. Totally, 171 novel miRNAs were identified for six libraries. The mature sequences, star sequences and precursor sequences of 171 novel miRNAs were listed in Table S5. The length of the novel miRNAs ranged from 19 to 30 nucleotides. Most of the nucleotide positions preferred to be uracil (U) or adenine (A) (Fig. S1B). Two exceptions were the 9th and 11th nucleotide where the dominant nucleotides were guanine (G) and cytosine (C), respectively.

Differentially expressed miRNAs between SD92 and SD140

To further validate the expression changes of miRNAs between SD92 and SD140, 15 miRNAs from 11 different miRNA families were randomly selected to be tested by RT-qPCR (Fig. 1). The results of RT-qPCR showed the same expression regulation pattern with miRNA sequencing data, which suggested that the miRNA sequencing result was reliable. What's more, the results showed 6 miRNAs were differentially expressed between SD92 and SD140 (P<0.05). Different miRNAs from the same miRNA family displayed the same regulation pattern. For example, both miR166a-3p and miR166d-5p_2 were from miR166 family and exhibited higher expression levels in SD140 than in SD92.

A total of 179 differentially expressed miRNAs were identified in this study, including 107 known miRNAs and 72 novel miRNAs (Fig. 2A, Table S6). Among the differentially expressed miRNAs, 65 and 114 were confirmed to be up- and down-regulated in SD140, respectively. These miRNAs belonged to 49 miRNA families. Of the 49 miRNA families, miR399 and miR172 family were the two largest families, which contained 10 and 9 miRNA members, respectively. Interestingly, the members of miR399 and miR172 families were significantly down-regulated in SD140, respectively.

Target gene prediction of miRNAs

To further explore the function of miRNAs, the target genes (mRNAs) of all miRNAs were predicted by psRobot and TargetFinder. Totally, 7,416 target genes were identified for 450 miRNAs where 897 target genes were confirmed as targets of 116 miRNAs by both softwares. Among these 897 target genes, 305 genes were regulated by 31 differentially expressed miRNAs (Table S7).

GO and KEGG pathway enrichment analysis of target genes

GO enrichment analysis of the above 305 target genes showed that the biological process ontology included 47 GO terms. “Cellular macromolecule metabolic process” and “macromolecule metabolic process” were the most abundant GO terms, containing 77 genes, respectively.

The cellular component ontology included 16 GO terms, and the most abundant terms were “cell” and “cell part”, which contained 115 genes, respectively. The molecular function ontology included 10 GO terms. The GO term “binding” contained 126 genes, which was the most abundant term in molecular function (Fig. 3).

To explore the possible function of target genes, KEGG pathway enrichment analysis was performed. The 305 target genes of 31 DEMs were distributed in 6 first-level and 33 second-level KEGG pathways, respectively. The first-level KEGG pathway term “metabolism” was the most abundant, including 10 second-level KEGG pathway terms. Thirty-eight target genes were assigned in the second-level KEGG pathway term “signal transduction”, which was the most abundant second-level KEGG pathway term (Fig. 4).

Among the enriched top 20 pathways, only two pathways, “plant hormone signal transduction” and “plant-pathogen interaction”, were defined as significantly enriched pathways ($p < 0.05$), which comprised 24 target genes, respectively (Fig. 5 and Table S8). This indicated that the DEMs between SD92 and SD140 might be involved in plant-pathogen interaction and hormone signal transduction.

Target genes of miRNAs involved in regulation of anthocyanin biosynthesis

Generally, plant miRNAs regulate target mRNAs through two major mechanisms, transcript cleavage and translational inhibition (Chen 2009), thus there are negative regulation relationship

in the expressions of miRNA and corresponding target genes. In our previous study, a comparative transcriptome analysis was performed between SD92 and SD140 (Liu et al. 2018). By combining transcriptome sequencing data (SRA accession number: SRP125987) and miRNA sequencing data of present study, 31 differentially expressed miRNAs and corresponding target mRNAs were identified and listed in Table S9. Among them, the differentially expressed miRNAs negatively regulating target mRNAs were screened, and 140 miRNA-mRNA pairs were confirmed. In these miRNAs-mRNAs pairs, miRNAs contained 5 known miRNA families and 12 novel miRNAs. These mRNAs corresponded to 71 genes (Table 4). These genes mainly encoded transcription factors, quamosa promoter binding protein, hormone response factors, protein kinases and disease resistance protein.

Transcription factors affect anthocyanin biosynthesis by regulating the expressions of structural genes (D'Amelia et al. 2014; Liu et al. 2016). In this study, we focused on the regulation of miRNA on transcription factors in anthocyanin biosynthesis (Table 4). *PGSC0003DMG400006604*, *PGSC0003DMG400011046* and *PGSC0003DMG400012038*, which were regulated by miR172b, encoded AP2 transcription factor SlAP2e, RAP2-7-like and RAP2-7, respectively. The target gene of miR530b_4, *PGSC0003DMG400025479*, encoded AP2-like transcription factor TOE3. *PGSC0003DMG400011457* encoded WRKY transcription factor 48 and was regulated by miR172e-5p. Both *PGSC0003DMG400004826* and *PGSC0003DMG400018279*, which were regulated by novel_mir170, encoded transcription factor ERF039-like and MYB35-like, respectively.

Hormones improve the biosynthesis of anthocyanins (Zhang et al. 2011; Palma-Silva et al. 2016), so we did research on miRNA regulating hormones in this experiment in order to throw light on miRNA regulation mechanism on anthocyanins biosynthesis. In this study, RAP2-7 and RAP2-7-like, which were regulated by miR172b, were ethylene-responsive transcription factors. TOE3 transcription factor, which was regulated by miR172b and miR530b_4, was also responsive to ethylene (Table 4). The target gene of miR171b-3p, *PGSC0003DMG400012683*, encoded the DELLA protein that was an inhibitor of GA signal transduction.

Protein kinases were involved in anthocyanin biosynthesis (Li et al. 2016). Protein kinases regulated by miRNA were investigated in this study. Both *PGSC0003DMG400018811* and *PGSC0003DMG400024795*, which were regulated by novel_mir170, encoded LRR receptor-like serine/threonine protein kinase ERECTA and RCH1, respectively. *PGSC0003DMG400026383* encoded receptor-like protein kinase and was regulated by novel_mir117.

There were also significant changes in the expression of target genes regulated by other miRNAs, such as *PGSC0003DMG402007414*, which was target gene of novel_mir105 and novel_mir143, but the gene function was unknown.

Discussion

Generally, miRNAs play an important role in some kinds of plant biological processes such as growth, development and stress response (Jones-Rhoades et al. 2006). The functions of miRNAs in plant anthocyanin biosynthesis have been reported in some species, including Arabidopsis (Gou et al. 2011; Wang et al. 2016), apple (Hu et al. 2021), grape (Tirumalai et al. 2019), tomato (Jia et al. 2015), sweet potato (He et al. 2019) and kiwi fruit (Li et al. 2019).

In this study, miR399 and miR172 families were the two largest differentially expressed miRNA families. The expressions of miR399 family (miR399a_6, miR399i, miR399j_2) and miR172 family (miR172e-5p, miR172b) were down-regulated in SD140. miR172 inhibits flavonoid biosynthesis through suppressing the expression of an AP2 transcription factor that positively regulates *MdMYB10* (Ding et al. 2022). In SD140, miR172b was down-regulated, and its target gene encoding AP2-like factor was up-regulated, indicating that miR172b regulated the change in anthocyanin biosynthesis from petunidin to pelargonidin through AP2-like factor. Both miR399 expression and anthocyanin accumulation are increased under Pi-deficiency conditions (Chen et al. 2018; Hsieh et al. 2009). miR399 is related to anthocyanin accumulation. However, the target gene of miR399 was unknown in SD92 and SD140, so the regulation mechanism of miR399 in anthocyanin biosynthesis remains unclear and needs further study.

miR171 family (miR171a-3p, miR171b-3p, miR171b-3p_2) was up-regulated in SD140 (Table S6). miR171 is down-regulated and anthocyanin accumulation is up-regulated under water deficit (Ghorecha et al. 2014). miR171 is related with anthocyanin accumulation. The target gene of miR171b-3p, *PGSC0003DMG400012683*, encoded DELLA protein. DELLA proteins are important repressors of GA signaling (Chai et al. 2022; Sukiran et al. 2022). Plant hormones are involved in anthocyanin biosynthesis, such as auxin (Ji et al. 2015; Liu et al. 2014), abscisic acid (ABA) (Balint & Reynolds 2013; Leão et al. 2014) and gibberellic acid (GA) (Loreti et al. 2008). GA represses the sucrose accumulation in anthocyanin synthesis (Loreti et al. 2008) and decreases anthocyanin accumulation under low temperature or phosphate starvation (Jiang et al. 2007; Zhang et al. 2011). Moreover, the KEGG pathway “plant hormone signal transduction” comprising of 24 target genes was significantly enriched in this study, which suggested that plant hormones were involved in the anthocyanin biosynthesis in SD92 and SD140. Thus, it indicated that miR171b-3p probably regulated the change of anthocyanin biosynthesis in SD92 and SD140 through DELLA protein.

miR828 were frequently reported to be involved in anthocyanin biosynthesis regulation (Bonar et al. 2018; Tirumalai et al. 2019). In potato, miR828 is associated with purple tuber skin and flesh color rich in anthocyanin. One member of miR828 family, miR828a_1, was identified in SD92 and SD140, but was not significantly expressed differentially between SD92 and

SD140. These results indicated that miR828a_1 might not regulate the change of anthocyanin biosynthesis between SD92 and SD140.

The accumulation of anthocyanin was reported to be related with miR156 (Gou et al. 2011). In this study, miRNA156 was differentially expressed between SD92 and SD140. Its target gene mainly encoded squamosa promoter binding protein and cell cycle checkpoint protein RAD17. These target genes regulated by miR156a-5p need further study in anthocyanin biosynthesis.

A novel miRNA, novel_mir170, was down-regulated in SD140 (4.81 vs 0.14). It regulated a number of target genes, which mainly encoded protein kinase, ethylene responsive transcription factor ERF039-like and transcription factor MYB35-like. Protein kinases play an important role in anthocyanin biosynthesis. Plant sucrose-nonfermenting 1 (SNF1)-related protein kinase was involved in anthocyanin accumulation regulated by MdbHLH3 (Liu et al. 2017; Shen et al. 2017). Anthocyanin biosynthesis was regulated by mitogen-activated protein kinase (Luo et al. 2017; Wersch et al. 2018). In this experiment, the two target genes of novel_mir170 encoding LRR receptor-like serine/threonine-protein kinase were up-regulated, which were consistent with the metabolism data (Liu et al. 2022). These results showed that novel_mir170 regulated the change of anthocyanin biosynthesis through LRR receptor-like serine/threonine-protein kinase in SD92 and SD140. MYB transcription factor can regulate the biosynthesis of anthocyanin by regulating the expression of structural genes (D'Amelia et al. 2014). Target gene of novel_mir170, which encoded MYB transcription factor, was up-regulated. These results showed that novel_mir170 regulated the anthocyanin biosynthesis by regulating the expression of MYB. Ethylene is closely related to the biosynthesis of anthocyanin (Chen et al. 2022; Jeong et al. 2010). In this study, the target gene of novel_mir170 encoding ethylene responsive transcription factor ERF039 was up-regulated. These results indicated that novel_mir170 regulated anthocyanin biosynthesis by up-regulating the expression of ethylene responsive transcription factor. In conclusion, novel_mir170 was an important novel miRNA identified in this study and might be an important miRNA for regulation of anthocyanin biosynthesis.

Conclusions

A comparative small RNA sequencing analysis between purple potato and its mutant revealed that there were 179 differentially expressed miRNAs, consisting of 65 up- and 114 down-regulated miRNAs, respectively. miR399 and miR172 families were the two largest differentially expressed miRNA families. 31 differentially expressed miRNAs were predicted to potentially regulate 305 target genes. The miRNA sequencing data and the transcriptome data showed that miR171 family and miR172 family regulated the change in anthocyanin biosynthesis from petunidin to pelargonidin through DELLA protein and AP2-like transcription

factor, respectively. A novel miRNA, novel_mir170, regulated anthocyanin biosynthesis by serine/threonine-protein kinase and MYB transcription factor.

Acknowledgements

We sincerely appreciate Dr. Yumeng Huo for his help in designing miRNA primers for RT-qPCR.

References

Abdi H. 2007. The Bonferonni and Šidák corrections for multiple comparisons in Encyclopedia of Measurement and Statistics. Thousand Oaks, CA: SAGE.

Axtell MJ. 2013. Classification and comparison of small RNAs from plants. *Annu Rev Plant Biol* 64:137-159. 10.1146/annurev-arplant-050312-120043

Balint G, and Reynolds AG. 2013. Impact of exogenous abscisic acid on vine physiology and grape composition of Cabernet Sauvignon. *Am J Enol Viticult* 64:74-87. 10.5344/ajev.2012.12075

Bimpilas A, Panagopoulou M, Tsimogiannis D, and Oreopoulou V. 2016. Anthocyanin copigmentation and color of wine: The effect of naturally obtained hydroxycinnamic acids as cofactors. *Food Chem* 197:39-46. 10.1016/j.foodchem.2015.10.095

Bonar N, Liney M, Zhang R, Austin C, Dessoly J, Davidson D, Stephens J, McDougall G, Taylor M, Bryan GJ, and Hornyik C. 2018. Potato mir828 is associated with purple tuber skin and flesh color. *Front Plant Sci* 9:1742. 10.3389/fpls.2018.01742

Boyle EI, Weng S, Gollub J, Jin H, Botstein D, Cherry JM, and Sherlock G. 2004. GO::TermFinder--open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. *Bioinformatics* 20:3710-3715. 10.1093/bioinformatics/bth456

Chai Z, Fang J, Huang C, Huang R, Tan X, Chen B, Yao W, and Zhang M. 2022. ScAIL1 modulates plant defense responses by targeting DELLA and regulating GA and JA signaling. *J Exp Bot* 73:6727-6743. 10.1093/jxb/erac339

Chen M, Gu H, Wang L, Shao Y, Li R, and Li W. 2022. Exogenous ethylene promotes peel color transformation by regulating the degradation of chlorophyll and synthesis of anthocyanin in postharvest mango fruit. *Front Nutr* 9:911542. 10.3389/fnut.2022.911542

Chen X. 2009. Small RNAs and their roles in plant development. *Annu Rev Cell Dev Biol* 25:21-44. 10.1146/annurev.cellbio.042308.113417

Chen YF, Shibu MA, Fan MJ, Chen MC, Viswanadha VP, Lin YL, Lai CH, Lin KH, Ho TJ, Kuo WW, and Huang CY. 2016. Purple rice anthocyanin extract protects cardiac function in STZ-induced diabetes rat hearts by inhibiting cardiac hypertrophy and fibrosis. *J Nutr*

Biochem 31:98-105. 10.1016/j.jnutbio.2015.12.020

Chen Y, Wu P, Zhao Q, Tang Y, Chen Y, Li M, Jiang H, and Wu G. 2018. Overexpression of a phosphate starvation response AP2/ERF gene from physic nut in Arabidopsis alters root morphological traits and phosphate starvation-induced anthocyanin accumulation. *Front Plant Sci* 9:1186. 10.3389/fpls.2018.01186

Chiu LW, Zhou X, Burke S, Wu X, Prior RL, and Li L. 2010. The purple cauliflower arises from activation of a MYB transcription factor. *Plant Physiol* 154:1470-1480. 10.1104/pp.110.164160

Cristani M, Speciale A, Mancari F, Arcoraci T, Ferrari D, Fratantonio D, Saija A, Cimino F, and Trombetta D. 2016. Protective activity of an anthocyanin-rich extract from bilberries and blackcurrants on acute acetaminophen-induced hepatotoxicity in rats. *Nat Prod Res* 30:2845-2849. 10.1080/14786419.2016.1160235

D'Amelia V, Aversano R, Batelli G, Caruso I, Castellano Moreno M, Castro-Sanz AB, Chiaiese P, Fasano C, Palomba F, and Carputo D. 2014. High AN1 variability and interaction with basic helix-loop-helix co-factors related to anthocyanin biosynthesis in potato leaves. *Plant J* 80:527-540. 10.1111/tpj.12653

De Jong WS, De Jong DM, De Jong H, Kalazich J, and Bodis M. 2003. An allele of dihydroflavonol 4-reductase associated with the ability to produce red anthocyanin pigments in potato (Solanum tuberosum L.). *Theor Appl Genet* 107:1375-1383. 10.1007/s00122-003-1395-9

Ding T, Tomes S, Gleave AP, Zhang H, Dare AP, Plunkett B, Espley RV, Luo Z, Zhang R, Allan AC, Zhou Z, Wang H, Wu M, Dong H, Liu C, Liu J, Yan Z, and Yao JL. 2022. microRNA172 targets APETALA2 to regulate flavonoid biosynthesis in apple (Malus domestica). *Hortic Res* 9: uhab007. 10.1093/hr/uhab007

Evers M, Huttner M, Dueck A, Meister G, and Engelmann JC. 2015. miRA: adaptable novel miRNA identification in plants using small RNA sequencing data. *BMC Bioinformatics* 16:370. 10.1186/s12859-015-0798-3

Fahlgren N, and Carrington JC. 2010. miRNA target prediction in plants. *Methods Mol Biol* 592:51-57. 10.1007/978-1-60327-005-2_4

Ghorecha V, Patel K, Ingle S, Sunkar R, and Krishnayya NS. 2014. Analysis of biochemical variations and microRNA expression in wild (Ipomoea campanulata) and cultivated (Jacquemontia pentantha) species exposed to in vivo water stress. *Physiol Mol Biol Plants* 20:57-67. 10.1007/s12298-013-0207-1

Gou JY, Felippes FF, Liu CJ, Weigel D, and Wang JW. 2011. Negative regulation of anthocyanin biosynthesis in Arabidopsis by a miR156-targeted SPL transcription factor. *Plant Cell* 23:1512-1522. 10.1105/tpc.111.084525

Han KH, Hashimoto N, Hashimoto M, Noda T, Shimada K, Lee CH, Sekikawa M, and Fukushima M. 2006. Red potato extract protects from D-galactosamine-induced liver injury in rats. *Biosci Biotechnol Biochem* 70:2285-2288. 10.1271/bbb.60097

He L, Tang R, Shi X, Wang W, Cao Q, Liu X, Wang T, Sun Y, Zhang H, Li R, and Jia X. 2019. Uncovering anthocyanin biosynthesis related microRNAs and their target genes by small RNA and degradome sequencing in tuberous roots of sweetpotato. *BMC Plant Biol* 19:232. 10.1186/s12870-019-1790-2

Hsieh LC, Lin SI, Shih AC, Chen JW, Lin WY, Tseng CY, Li WH, and Chiou TJ. 2009. Uncovering small RNA-mediated responses to phosphate deficiency in Arabidopsis by deep sequencing. *Plant Physiol* 151:2120-2132. 10.1104/pp.109.147280

Hu Y, Cheng H, Zhang Y, Zhang J, Niu S, Wang X, Li W, Zhang J, and Yao Y. 2021. The MdMYB16/MdMYB1-miR7125-MdCCR module regulates the homeostasis between anthocyanin and lignin biosynthesis during light induction in apple. *New Phytol* 231:1105-1122. 10.1111/nph.17431

Jagla B, Aulner N, Kelly PD, Song D, Volchuk A, Zatorski A, Shum D, Mayer T, De Angelis DA, Ouerfelli O, Rutishauser U, and Rothman JE. 2005. Sequence characteristics of functional siRNAs. *RNA* 11:864-872. 10.1261/rna.7275905

Jeong SW, Das PK, Jeoung SC, Song JY, Lee HK, Kim YK, Kim WJ, Park YI, Yoo SD, Choi SB, Choi G, and Park YI. 2010. Ethylene suppression of sugar-induced anthocyanin pigmentation in Arabidopsis. *Plant Physiol* 154:1514-1531. 10.1104/pp.110.161869

Ji X-H, Wang Y-T, Zhang R, Wu S-J, An M-M, Li M, Wang C-Z, Chen X-L, Zhang Y-M, and Chen X-S. 2015. Effect of auxin, cytokinin and nitrogen on anthocyanin biosynthesis in callus cultures of red-fleshed apple (Malus sieversii f.niedzwetzkyana). *Plant Cell, Tissue and Organ Culture (PCTOC)* 120:325-337. 10.1007/s11240-014-0609-y

Jia X, Shen J, Liu H, Li F, Ding N, Gao C, Pattanaik S, Patra B, Li R, and Yuan L. 2015. Small tandem target mimic-mediated blockage of microRNA858 induces anthocyanin accumulation in tomato. *Planta* 242:283-293. 10.1007/s00425-015-2305-5

Jiang C, Gao X, Liao L, Harberd NP, and Fu X. 2007. Phosphate starvation root architecture and anthocyanin accumulation responses are modulated by the gibberellin-DELLA signaling pathway in Arabidopsis. *Plant Physiol* 145:1460-1470. 10.1104/pp.107.103788

Jones-Rhoades MW, Bartel DP, and Bartel B. 2006. MicroRNAs and their regulatory roles in plants. *Annu Rev Plant Biol* 57:19-53. 10.1146/annurev.arplant.57.032905.105218

Jung CS, Griffiths HM, De Jong DM, Cheng S, Bodis M, and De Jong WS. 2005. The potato P locus codes for flavonoid 3',5'-hydroxylase. *Theor Appl Genet* 110:269-275. 10.1007/s00122-004-1829-z

Jung CS, Griffiths HM, De Jong DM, Cheng S, Bodis M, Kim TS, and De Jong WS. 2009. The

potato developer (D) locus encodes an R2R3 MYB transcription factor that regulates expression of multiple anthocyanin structural genes in tuber skin. *Theor Appl Genet* 120:45-57. 10.1007/s00122-009-1158-3

Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, and Yamanishi Y. 2008. KEGG for linking genomes to life and the environment. *Nucleic Acids Res* 36:D480-484. 10.1093/nar/gkm882

Kozomara A, and Griffiths-Jones S. 2014. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res* 42:D68-73. 10.1093/nar/gkt1181

Langmead B, Trapnell C, Pop M, and Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10:R25. 10.1186/gb-2009-10-3-r25

Leão P, Lima M, Costa J, and Trindade D. 2014. Abscisic acid and ethephon for improving red color and quality of crimson seedless grapes grown in a tropical region. *Am J Enol Viticult* 66:37-45. 10.5344/ajev.2014.14041

Li S, Wang W, Gao J, Yin K, Wang R, Wang C, Petersen M, Mundy J, and Qiu JL. 2016. MYB75 phosphorylation by MPK4 is required for light-induced anthocyanin accumulation in Arabidopsis. *Plant Cell* 28:2866-2883. 10.1105/tpc.16.00130

Liu F, Chen G, Zhang Y, Zhao P, Dong D, Wang Y, Wang S, and Yang Y. 2022. A comparative analysis of metabolome reveals the regulation of the anthocyanin biosynthesis branch in potatoes. *Potato Res* 66:1-20. 10.1007/s11540-022-09586-5

Li Y, Cui W, Wang R, Lin M, Zhong Y, Sun L, Qi X, and Fang J. 2019b. MicroRNA858-mediated regulation of anthocyanin biosynthesis in kiwifruit (Actinidia arguta) based on small RNA sequencing. *PLoS One* 14:e0217480. 10.1371/journal.pone.0217480

Liu Y, Lin-Wang K, Espley RV, Wang L, Yang H, Yu B, Dare A, Varkonyi-Gasic E, Wang J, Zhang J, Wang D, and Allan AC. 2016. Functional diversification of the potato R2R3 MYB anthocyanin activators AN1, MYBA1, and MYB113 and their interaction with basic helix-loop-helix cofactors. *J Exp Bot* 67:2159-2176. 10.1093/jxb/erw014

Liu F, Yang Y, Gao J, Ma C, and Bi Y. 2018. A comparative transcriptome analysis of a wild purple potato and its red mutant provides insight into the mechanism of anthocyanin transformation. *PLoS One* 13:e0191406. 10.1371/journal.pone.0191406

Liu XJ, An XH, Liu X, Hu DG, Wang XF, You CX, and Hao YJ. 2017. MdSnRK1.1 interacts with MdJAZ18 to regulate sucrose-induced anthocyanin and proanthocyanidin accumulation in apple. *J Exp Bot* 68:2977-2990. 10.1093/jxb/erx150

Liu Y, Lin-Wang K, Deng C, Warran B, Wang L, Yu B, Yang H, Wang J, Espley RV, Zhang J, Wang D, and Allan AC. 2015. Comparative transcriptome analysis of white and purple potato to identify genes involved in anthocyanin biosynthesis. *PLoS One* 10:e0129148.

10.1371/journal.pone.0129148

Liu Y, Lin-Wang K, Espley RV, Wang L, Li Y, Liu Z, Zhou P, Zeng L, Zhang X, Zhang J, and Allan AC. 2019. StMYB44 negatively regulates anthocyanin biosynthesis at high temperatures in tuber flesh of potato. *J Exp Bot* 70:3809-3824. 10.1093/jxb/erz194

Liu Z, Shi MZ, and Xie DY. 2014. Regulation of anthocyanin biosynthesis in Arabidopsis thaliana red pap1-D cells metabolically programmed by auxins. *Planta* 239:765-781. 10.1007/s00425-013-2011-0

Loreti E, Povero G, Novi G, Solfanelli C, Alpi A, and Perata P. 2008. Gibberellins, jasmonate and abscisic acid modulate the sucrose-induced expression of anthocyanin biosynthetic genes in Arabidopsis. *New Phytol* 179:1004-1016. 10.1111/j.1469-8137.2008.02511.x

Luo J, Wang X, Feng L, Li Y, and He JX. 2017. The mitogen-activated protein kinase kinase 9 (MKK9) modulates nitrogen acquisition and anthocyanin accumulation under nitrogen-limiting condition in Arabidopsis. *Biochem Biophys Res Commun* 487:539-544. 10.1016/j.bbrc.2017.04.065

Matzke M, Kanno T, Daxinger L, Huettel B, and Matzke AJ. 2009. RNA-mediated chromatin-based silencing in plants. *Curr Opin Cell Biol* 21:367-376. 10.1016/j.ceb.2009.01.025

Nawrocki EP, Burge SW, Bateman A, Daub J, Eberhardt RY, Eddy SR, Floden EW, Gardner PP, Jones TA, Tate J, and Finn RD. 2015. Rfam 12.0: updates to the RNA families database. *Nucleic Acids Res* 43:D130-137. 10.1093/nar/gku1063

Nawrocki EP, and Eddy SR. 2013. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* 29:2933-2935. 10.1093/bioinformatics/btt509

Palma-Silva C, Ferro M, Bacci M, and Turchetto-Zolet AC. 2016. De novo assembly and characterization of leaf and floral transcriptomes of the hybridizing bromeliad species (Pitcairnia spp.) adapted to neotropical inselbergs. *Mol Ecol Resour* 16:1012-1022. 10.1111/1755-0998.12504

Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M, Sturn A, Snuffin M, Rezantsev A, Popov D, Ryltsov A, Kostukovich E, Borisovsky I, Liu Z, Vinsavich A, Trush V, and Quackenbush J. 2003. TM4: a free, open-source system for microarray data management and analysis. *Biotechniques* 34:374-378. 10.2144/03342mt01

Shen X, Guo X, Zhao D, Zhang Q, Jiang Y, Wang Y, Peng X, Wei Y, Zhai Z, Zhao W, and Li T. 2017. Cloning and expression profiling of the PacSnRK2 and PacPP2C gene families during fruit development, ABA treatment, and dehydration stress in sweet cherry. *Plant Physiol Biochem* 119:275-285. 10.1016/j.plaphy.2017.08.025

Sukiran NA, Pollastri S, Steel PG, and Knight MR. 2022. Plant growth promotion by the interaction of a novel synthetic small molecule with GA-DELLA function. *Plant Direct*

6:e398. 10.1002/pld3.398

't Hoen PA, Ariyurek Y, Thygesen HH, Vreugdenhil E, Vossen RH, de Menezes RX, Boer JM, van Ommen GJ, and den Dunnen JT. 2008. Deep sequencing-based expression analysis shows major advances in robustness, resolution and inter-lab portability over five microarray platforms. *Nucleic Acids Res* 36:e141. 10.1093/nar/gkn705

Tirumalai V, Swetha C, Nair A, Pandit A, and Shivaprasad PV. 2019. miR828 and miR858 regulate VvMYB114 to promote anthocyanin and flavonol accumulation in grapes. *J Exp Bot* 70:4775-4792. 10.1093/jxb/erz264

Wang L, Feng Z, Wang X, Wang X, and Zhang X. 2010. DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* 26:136-138. 10.1093/bioinformatics/btp612

Wang Y, Wang Y, Song Z, and Zhang H. 2016. Repression of MYBL2 by both microRNA858a and HY5 leads to the activation of anthocyanin biosynthetic pathway in Arabidopsis. *Mol Plant* 9:1395-1405. 10.1016/j.molp.2016.07.003

Wu HJ, Ma YK, Chen T, Wang M, and Wang XJ. 2012. PsRobot: a web-based plant small RNA meta-analysis toolbox. *Nucleic Acids Res* 40:W22-28. 10.1093/nar/gks554

Xia R, Zhu H, An YQ, Beers EP, and Liu Z. 2012. Apple miRNAs and tasiRNAs with novel regulatory networks. *Genome Biol* 13:R47. 10.1186/gb-2012-13-6-r47

Yang YH, Dudoit S, Luu P, Lin DM, Peng V, Ngai J, and Speed TP. 2002. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res* 30:e15. 10.1093/nar/30.4.e15

Yoshihama M, Nakao A, and Kenmochi N. 2013. snOPY: a small nucleolar RNA orthological gene database. *BMC Res Notes* 6:426. 10.1186/1756-0500-6-426

Zhang Y, Liu Z, Liu R, Hao H, and Bi Y. 2011. Gibberellins negatively regulate low temperature-induced anthocyanin accumulation in a HY5/HYH-dependent manner. *Plant Signal Behav* 6:632-634. 10.4161/psb.6.5.14343

Table 1(on next page)

Primer sequences of miRNAs for real-time quantitative PCR

Table 1 Primer sequences of miRNAs for real-time quantitative PCR

Primer	Direction	Sequence (5'-3')
18S rRNA	Forward	CCTGGTCGGCATCGTTTA
18S rRNA	Reverse	CGAACAACTGCGAAAGCAT
miR156a-5p	Forward	TGACAGAAGAGAGTGAGCAC
miR166a-3p	Forward	TCGGACCAGGCTTCATTCC
miR166d-5p_2	Forward	GGAATGTTGTCTGGCTCGAGG
miR171b-3p	Forward	TTGAGCCGTGCCAATATCAC
miR171b-3p_2	Forward	TTGAGCCGCGTCAATATCTCT
miR172b	Forward	GGAATCTTGATGATGCTGCA
miR172e-5p	Forward	GCAACATCATCAAGATTCACA
miR399a_6	Forward	GCCAAAGGAGAATTGCCCTG
miR399i	Forward	CCAAAGGAGAGCTGCCCTG
miR399j_2	Forward	TGCCAAAGGAGAGTTGCCCTA
miR530a	Forward	TGCATTTGCACCTGCACCTT
miR828a_1	Forward	CGCTGTCTTGCTCAAATGAGTATTC
novel_mir32	Forward	ATTAACTTTGGCCAGCATC
novel_mir105	Forward	GGACCCTTGGCGAAGTCACC
novel_mir143	Forward	CACTGAGTTGGACCCTTGGC
novel_mir170	Forward	GCGAGCGAATTAGATTCATTGTTTGA

Table 2(on next page)

Summary of sequencing data for each sample

Table 2 Summary of sequencing data for each sample

Sample name	Raw tag count	Low quality tag	Invalid adapter tag	Poly A tag	Tag length < 18	Clean tag	Q20 of clean tag(%)	Percentage of clean tag(%)
SD140_1	30,152,601	521,573	1,211,217	765	296,890	28,122,156	99.30	93.27
SD140_2	29,662,224	559,145	642,637	1,307	285,077	28,174,058	99.20	94.98
SD140_3	29,108,569	439,201	1,438,318	979	420,200	26,809,871	99.20	92.10
SD92_1	28,058,311	476,281	601,154	814	262,128	26,717,934	99.00	95.22
SD92_2	28,907,701	462,036	684,333	2,174	265,810	27,493,348	99.30	95.11
SD92_3	29,706,600	544,647	816,486	1,600	341,405	28,002,462	99.20	94.26

Table 3(on next page)

Summary of detected small RNAs for each sample

Table 3 Summary of detected small RNAs for each sample

Sample name	Known miRNA	Novel miRNA	Known siRNA	Novel siRNA
SD140_1	290	151	0	12,518
SD140_2	293	161	0	13,671
SD140_3	284	145	0	12,447
SD92_1	275	166	0	13,373
SD92_2	304	161	0	11,225
SD92_3	311	168	0	13,147

Table 4(on next page)

Differentially expressed miRNAs and negatively regulated target genes

Table 4 Differentially expressed miRNAs and negatively regulated target genes

miRNA	Target Gene	Gene Annotation
miR156a-5p	*PGSC0003DMG400022824*	squamosa promoter-binding protein 1-like
miR156a-5p	*PGSC0003DMG400023962*	uncharacterized protein
miR156a-5p	*PGSC0003DMG400029156*	cell cycle checkpoint protein RAD17
miR156a-5p	*PGSC0003DMG400032817*	squamosa promoter-binding protein 1-like
miR156a-5p	*PGSC0003DMG400034310*	squamosa promoter-binding-like protein 12
miR171b-3p	*PGSC0003DMG400009015*	DEAD-box ATP-dependent RNA helicase 24
miR171b-3p	*PGSC0003DMG400012683*	DELLA protein
miR172b	*PGSC0003DMG400004006*	floral homeotic protein APETALA 2
miR172b	*PGSC0003DMG400006604*	AP2 transcription factor SlAP2e
miR172b	*PGSC0003DMG400011046*	ethylene-responsive transcription factor RAP2-7-like
miR172b	*PGSC0003DMG400012038*	ethylene-responsive transcription factor RAP2-7
miR172b	*PGSC0003DMG400027904*	floral homeotic protein APETALA 2-like
miR172b	*PGSC0003DMG400030080*	phosphatidylinositol/phosphatidylcholine transfer protein SFH4
miR172b miR530b_4	*PGSC0003DMG400025479*	AP2-like ethylene-responsive transcription factor TOE3
miR172e-5p	*PGSC0003DMG400010386*	malate dehydrogenase, glyoxysomal
miR172e-5p	*PGSC0003DMG400011457*	probable WRKY transcription factor 48
miR172e-5p	*PGSC0003DMG400011477*	putative lysine-specific demethylase JMJ16
miR172e-5p	*PGSC0003DMG400021020*	uncharacterized protein
miR172e-5p novel_mir32	*PGSC0003DMG400014214*	uncharacterized protein
miR482e-5p novel_mir117	*PGSC0003DMG400030780*	uncharacterized protein
miR530a	*PGSC0003DMG400010027*	dof zinc finger protein DOF3.5-like
miR530a	*PGSC0003DMG400022193*	pirin-like protein
miR530a	*PGSC0003DMG400030421*	transcription initiation factor IIA large subunit
miR530a	*PGSC0003DMG400038860*	uncharacterized protein
miR530b_4	*PGSC0003DMG400001126*	uncharacterized protein
miR530b_4	*PGSC0003DMG400030587*	non-specific lipid-transfer protein 2-like
novel_mir32	*PGSC0003DMG400003436*	uncharacterized protein
novel_mir32	*PGSC0003DMG400007187*	probable protein S-acyltransferase 1
novel_mir32	*PGSC0003DMG400009055*	uncharacterized protein
novel_mir32	*PGSC0003DMG400011113*	putative disease resistance protein RGA3
novel_mir32	*PGSC0003DMG400012875*	protein disulfide isomerase-like 1-3
novel_mir32	*PGSC0003DMG400016798*	polyadenylate-binding protein 2-like
novel_mir32	*PGSC0003DMG400017569*	protein disulfide-isomerase-like
novel_mir32	*PGSC0003DMG400027301*	caffeic acid 3-O-methyltransferase-like
novel_mir32	*PGSC0003DMG400032155*	linoleate 13S-lipoxygenase 2-1, chloroplastic
novel_mir32	*PGSC0003DMG400043688*	uncharacterized protein

novel_mir42	*PGSC0003DMG400008897*	L-type lectin-domain containing receptor kinase IV.1-like
novel_mir54	*PGSC0003DMG400032120*	UPF0496 protein At3g19330-like
novel_mir61	*PGSC0003DMG400004296*	late blight resistance protein homolog R1B-16
novel_mir61	*PGSC0003DMG400004756*	late blight resistance protein homolog R1A-10
novel_mir61	*PGSC0003DMG400007867*	disease resistance protein RGH3
novel_mir61	*PGSC0003DMG400007870*	late blight resistance protein homolog R1A-3
novel_mir61	*PGSC0003DMG400007872*	late blight resistance protein homolog R1C-3
novel_mir61	*PGSC0003DMG400031244*	THUMP domain-containing protein 1 homolog
novel_mir61	*PGSC0003DMG402007871*	disease resistance protein RGH3
novel_mir67	*PGSC0003DMG400008560*	uncharacterized protein
novel_mir67	*PGSC0003DMG400017053*	uncharacterized protein
novel_mir67	*PGSC0003DMG400030551*	multicopper oxidase LPR2
novel_mir75	*PGSC0003DMG400003887*	uncharacterized protein
novel_mir75	*PGSC0003DMG400009731*	probable S-adenosylmethionine-dependent methyltransferase
novel_mir75	*PGSC0003DMG400017312*	RING finger protein 44
novel_mir75	*PGSC0003DMG400025978*	uncharacterized protein
novel_mir78	*PGSC0003DMG400000774*	RNA-binding protein 2
novel_mir89	*PGSC0003DMG400006945*	senescence-associated carboxylesterase 101-like
novel_mir105 novel_mir143	*PGSC0003DMG402007414*	uncharacterized protein
novel_mir117	*PGSC0003DMG400020645*	ycf54-like protein
novel_mir117	*PGSC0003DMG400026383*	probable receptor-like protein kinase
novel_mir117	*PGSC0003DMG400031180*	uncharacterized protein
novel_mir128	*PGSC0003DMG400034633*	uncharacterized protein
novel_mir128	*PGSC0003DMG400037457*	uncharacterized protein
novel_mir128	*PGSC0003DMG400043850*	uncharacterized protein
novel_mir170	*PGSC0003DMG400000513*	galactinol-sucrose galactosyltransferase 5
novel_mir170	*PGSC0003DMG400002541*	60S ribosomal protein L37-3
novel_mir170	*PGSC0003DMG400004826*	ethylene-responsive transcription factor ERF039-like
novel_mir170	*PGSC0003DMG400007189*	proteasome subunit alpha type-3-like, partial
novel_mir170	*PGSC0003DMG400008432*	uncharacterized protein
novel_mir170	*PGSC0003DMG400012159*	KAT8 regulatory NSL complex subunit 3
novel_mir170	*PGSC0003DMG400018279*	transcription factor MYB35-like
novel_mir170	*PGSC0003DMG400018811*	LRR receptor-like serine/threonine-protein kinase ERECTA
novel_mir170	*PGSC0003DMG400024795*	LRR receptor-like serine/threonine-protein kinase RCH1
novel_mir170	*PGSC0003DMG400033933*	hypothetical protein SDM1_41t00024

2

Figure 1

Expression analysis of miRNAs by RT-qPCR.

The values are represented by mean ± standard deviation (n = 3). Student's t-test, $P < 0.05$.

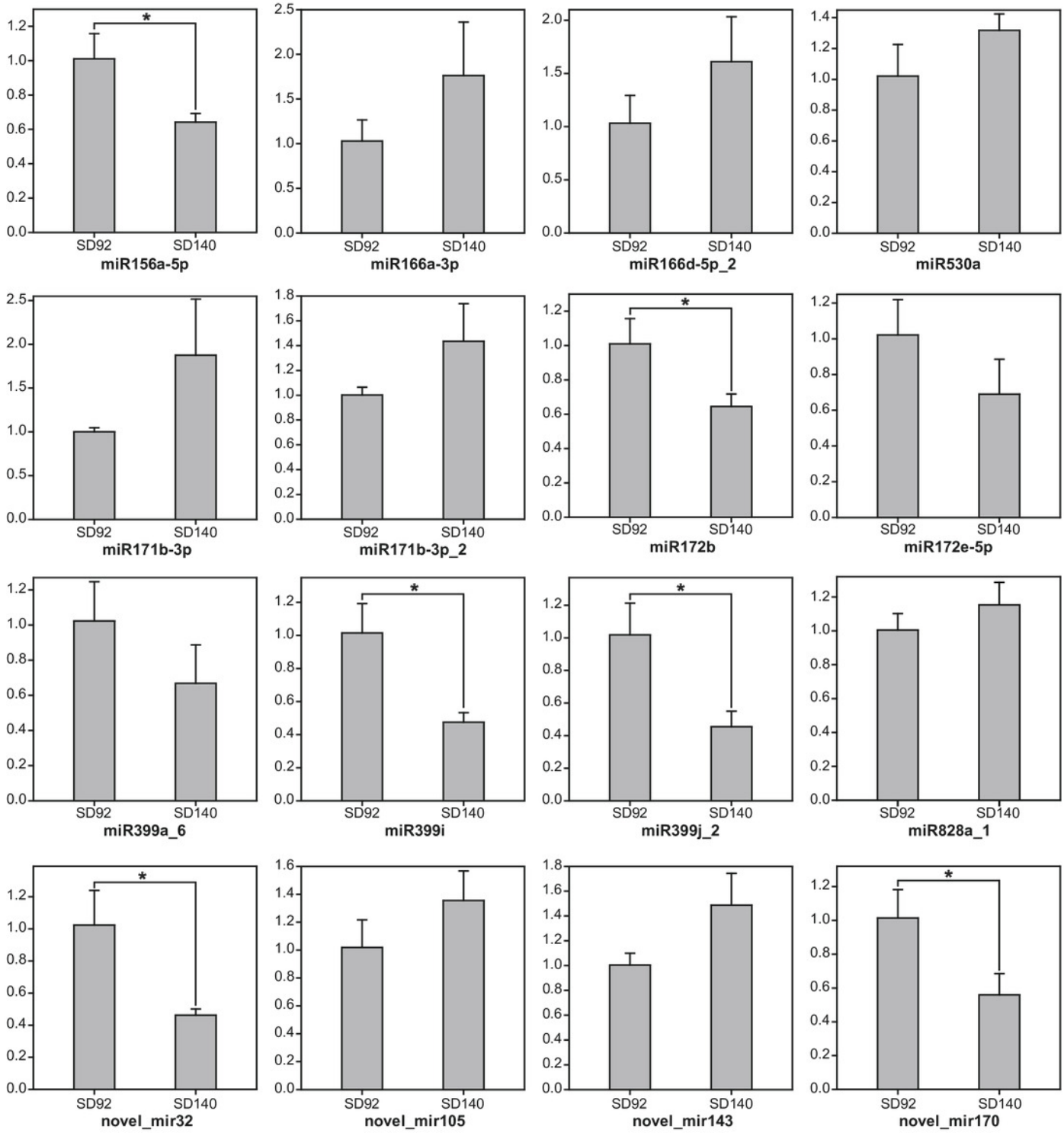

*
SD92
SD140
miR156a-5p
miR166a-3p
miR166d-5p_2
miR530a
miR171b-3p
miR171b-3p_2
miR172b
miR172e-5p
miR399a_6
miR399i
miR399j_2
miR828a_1
novel_mir32
novel_mir105
novel_mir143
novel_mir170

Figure 2

Identification of differentially expressed miRNAs between SD92 and SD140.

(A) Volcano plot of differentially expressed miRNAs between SD92 and SD140. The cutoff values of fold change and Q-value are > 2 and < 0.001, respectively. Up-regulated and down-regulated miRNAs are indicated by red and blue dots. (B) Heatmap of differentially expressed miRNAs in three biological replicates. Hierarchical clustering was performed by complete linkage method and Euclidean distance.

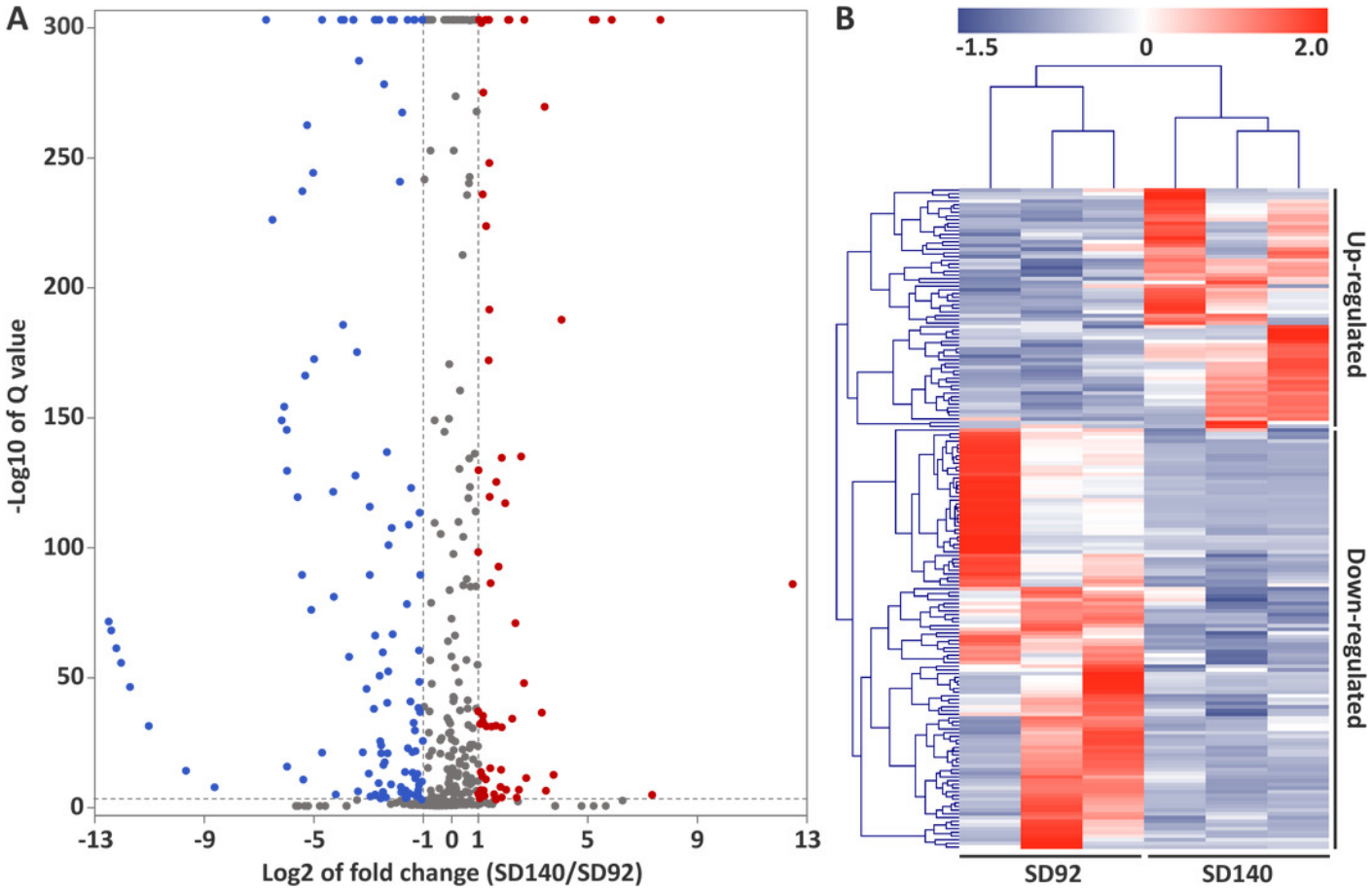


Figure 3

GO classification of predicted target genes of the differentially expressed miRNAs.

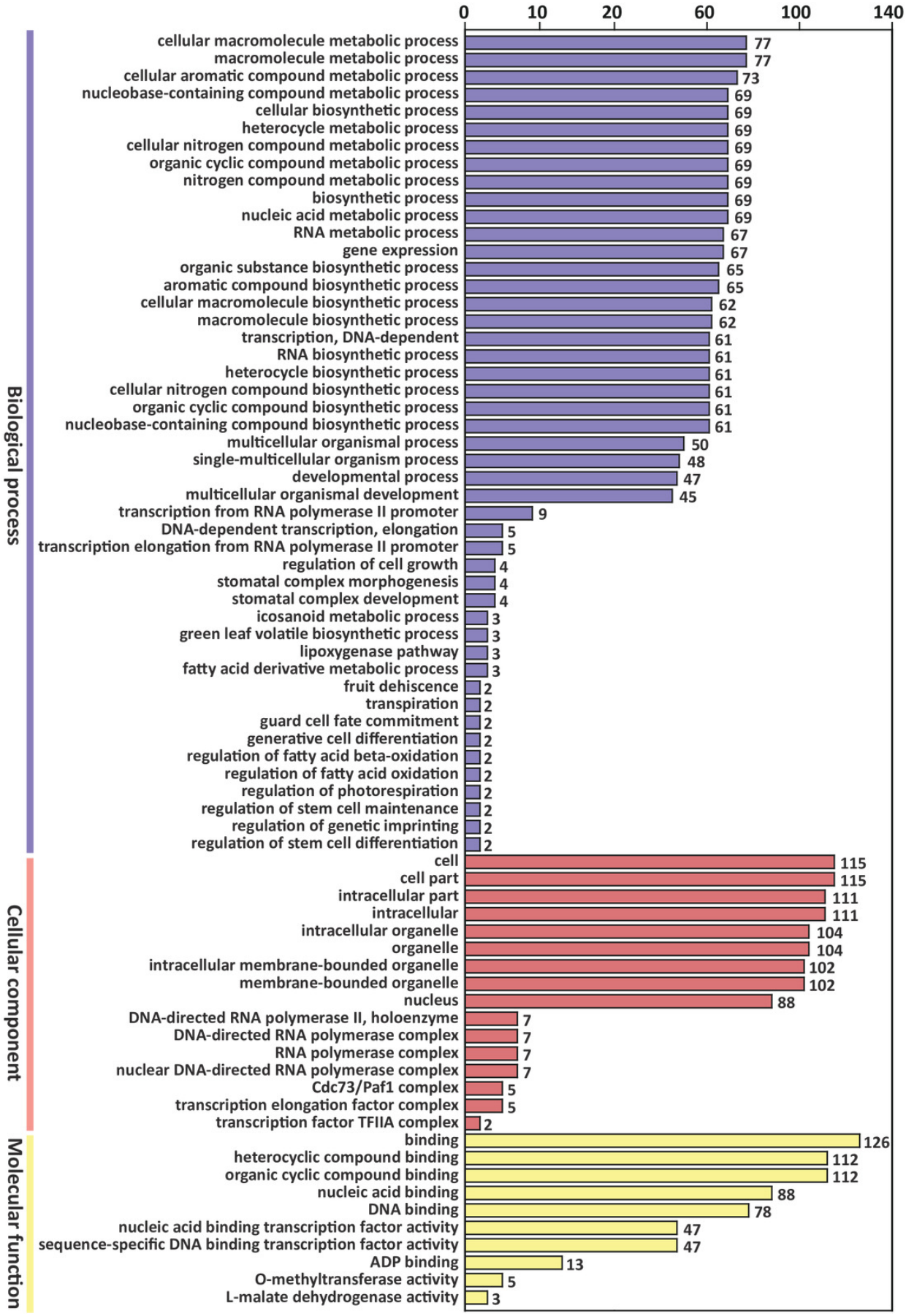

0
10
20
60
100
140
Biological process
cellular macromolecule metabolic process 77
macromolecule metabolic process 77
cellular aromatic compound metabolic process 73
nucleobase-containing compound metabolic process 69
cellular biosynthetic process 69
heterocycle metabolic process 69
cellular nitrogen compound metabolic process 69
organic cyclic compound metabolic process 69
nitrogen compound metabolic process 69
biosynthetic process 69
nucleic acid metabolic process 69
RNA metabolic process 67
gene expression 67
organic substance biosynthetic process 65
aromatic compound biosynthetic process 65
cellular macromolecule biosynthetic process 62
macromolecule biosynthetic process 62
transcription, DNA-dependent 61
RNA biosynthetic process 61
heterocycle biosynthetic process 61
cellular nitrogen compound biosynthetic process 61
organic cyclic compound biosynthetic process 61
nucleobase-containing compound biosynthetic process 61
multicellular organismal process 50
single-multicellular organism process 48
developmental process 47
multicellular organismal development 45
transcription from RNA polymerase II promoter 9
DNA-dependent transcription, elongation 5
transcription elongation from RNA polymerase II promoter 5
regulation of cell growth 4
stomatal complex morphogenesis 4
stomatal complex development 4
icosanoid metabolic process 3
green leaf volatile biosynthetic process 3
lipoxygenase pathway 3
fatty acid derivative metabolic process 3
fruit dehiscence 2
transpiration 2
guard cell fate commitment 2
generative cell differentiation 2
regulation of fatty acid beta-oxidation 2
regulation of fatty acid oxidation 2
regulation of photorespiration 2
regulation of stem cell maintenance 2
regulation of genetic imprinting 2
regulation of stem cell differentiation 2
Cellular component
cell 115
cell part 115
intracellular part 111
intracellular 111
intracellular organelle 104
organelle 104
intracellular membrane-bounded organelle 102
membrane-bounded organelle 102
nucleus 88
DNA-directed RNA polymerase II, holoenzyme 7
DNA-directed RNA polymerase complex 7
RNA polymerase complex 7
nuclear DNA-directed RNA polymerase complex 7
Cdc73/Paf1 complex 5
transcription elongation factor complex 5
transcription factor TFIIA complex 2
Molecular function
binding 126
heterocyclic compound binding 112
organic cyclic compound binding 112
nucleic acid binding 88
DNA binding 78
nucleic acid binding transcription factor activity 47
sequence-specific DNA binding transcription factor activity 47
ADP binding 13
O-methyltransferase activity 5
L-malate dehydrogenase activity 3

Figure 4

First-level and second-level KEGG pathway classification of predicted target genes of the DEMs.

Six different first-level KEGG pathway were distinguished in different colors.

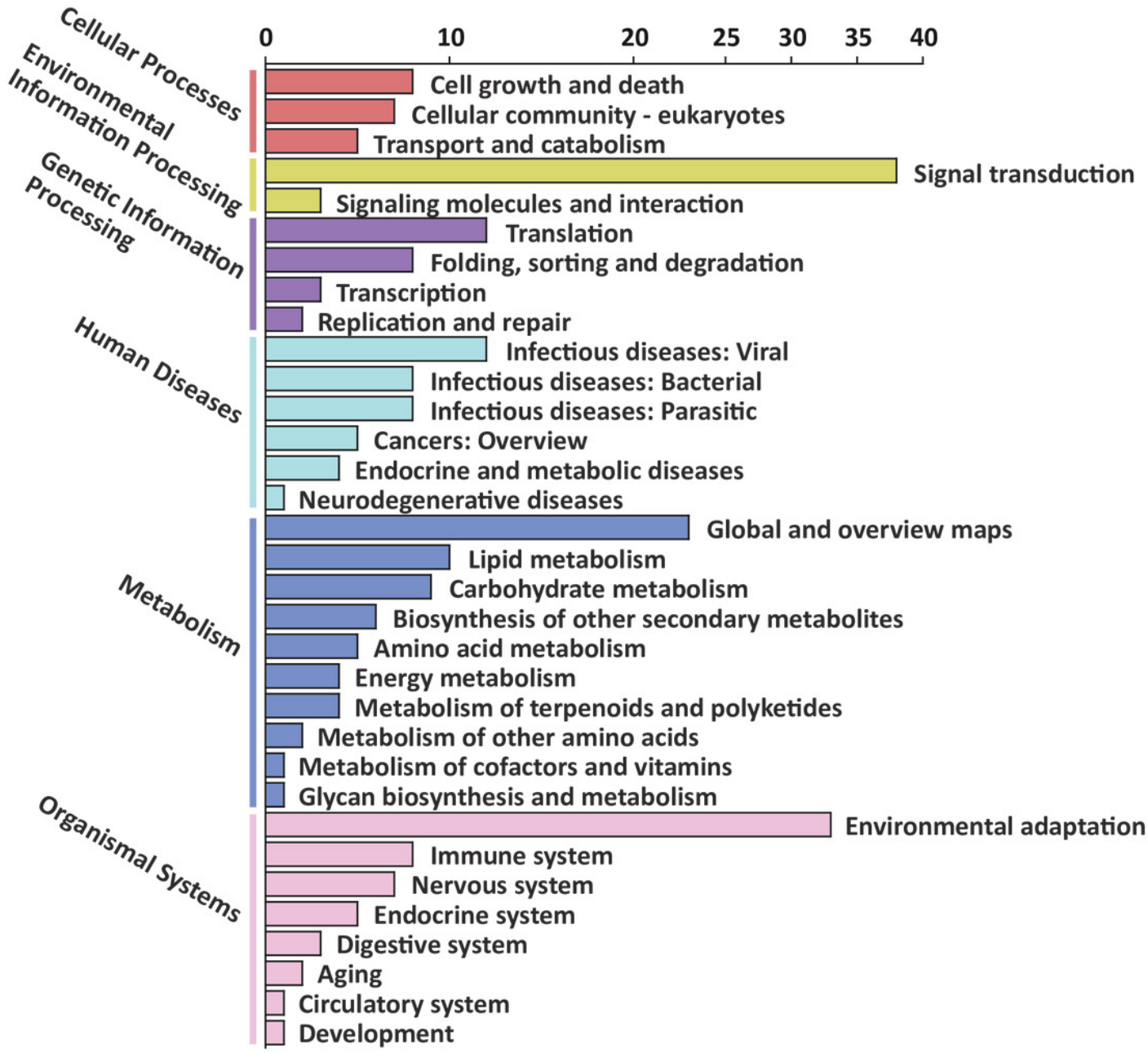
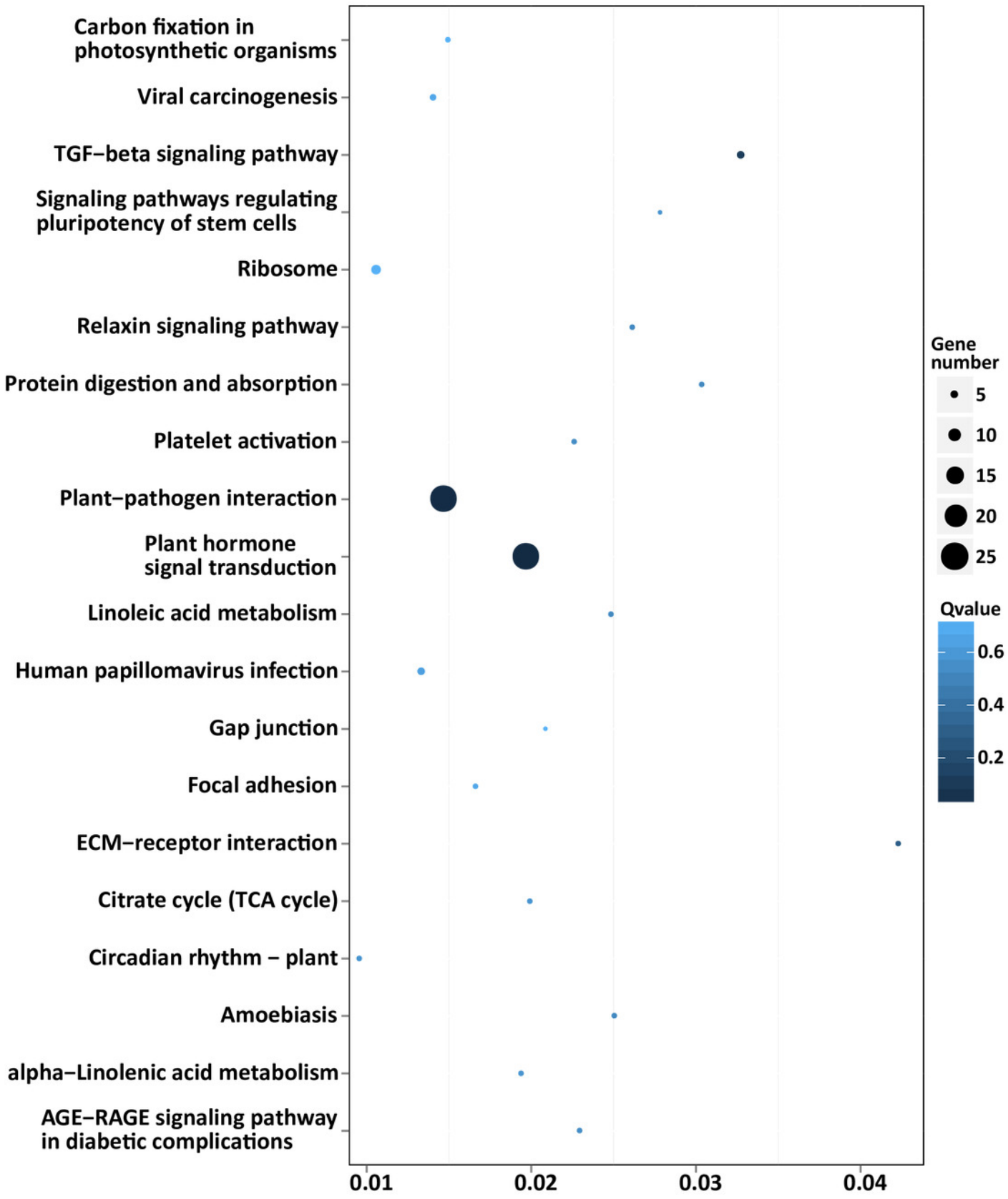


Figure 5

Scatterplot of enriched KEGG pathways of predicted target genes of the DEMs.

X axis indicates the rich factor. The rich factor is the ratio of DEMs target gene numbers annotated in the pathway term to all gene numbers annotated in the pathway. Y axis indicates KEGG pathways.


Carbon fixation in photosynthetic organisms
Viral carcinogenesis
TGF−beta signaling pathway
Signaling pathways regulating pluripotency of stem cells
Ribosome
Relaxin signaling pathway
Protein digestion and absorption
Platelet activation
Plant−pathogen interaction
Plant hormone signal transduction
Linoleic acid metabolism
Human papillomavirus infection
Gap junction
Focal adhesion
ECM−receptor interaction
Citrate cycle (TCA cycle)
Circadian rhythm − plant
Amoebiasis
alpha−Linolenic acid metabolism
AGE−RAGE signaling pathway in diabetic complications
0.01
0.02
0.03
0.04
Gene number
5
10
15
20
25
Qvalue
0.6
0.4
0.2