Identification of microRNAs in the green and red sectors of *Amaranthus tricolor* L. leaves based on Illumina sequencing data

Shengcai Liu [1], **Liyun Peng** [1], **Junfei Pan** [1], **Xiao Wang** [1], **Chunli Zhao** [1], **Chunzhen Cheng** [1], **Zihao Zhang** [1], **Yuling Lin** [1], **Xu XuHan** [Corresp., 1, 2], **Zhongxiong Lai** [Corresp. 1]

[1] Institute of Horticultural Biotechnology, Fujian Agriculture and Forestry University, Fuzhou, China

[2] Institut de la Recherche Interdisciplinaire de Toulouse, Toulouse, France

Corresponding Authors: Xu XuHan, Zhongxiong Lai
Email address: xxuhan@163.com, laizx01@163.com

Betalains are abundant in amaranth plants. Additionally, the betalain molecular structure and metabolic pathway differ from those of betanin in beet plants. To date, only a few studies have examined the regulatory roles of miRNAs in betalain biosynthesis in plants. Thus, we constructed small RNA libraries for the red and green sectors of amaranth leaves to identify miRNAs associated with betalain biosynthesis. We identified 198 known and 41 novel miRNAs. Moreover, 216 miRNAs were distributed in 44 miRNA families, including miR156, miR159, miR160, miR166, miR172, miR319, miR167, miR396, and miR398. An analysis of all unigene sequences in an amaranth transcriptome database resulted in the detection of 493 target genes for the 239 screened miRNAs. The targets included *SPL2*, *ARF18*, *ARF6*, and *NAC*. A quantitative real-time polymerase chain reaction validation of 20 miRNAs and nine target genes revealed expression-level differences between the red and green sectors of amaranth leaves. This study involved the application of an Illumina sequencing platform to identify miRNAs regulating betalain metabolism in amaranth plants. The data presented herein may provide insights into the molecular mechanisms underlying the regulation of betalain biosynthesis in amaranth and other plant species.

Identification of microRNAs in the green and red sectors of *Amaranthus tricolor* L. leaves based on Illumina sequencing data

Shengcai Liu [1], Liyun Peng [1], Junfei Pan [1], Xiao Wang [1], Chunli Zhao [1], Chunzhen Cheng [1], Zihao Zhang [1], Yuling Lin [1], Xu XuHan [1,2,*], Zhongxiong Lai [1,*]

1. Institute of Horticultural Biotechnology, Fujian Agriculture and Forestry University, Fuzhou 350002, China
2. Institut de la Recherche Interdisciplinaire de Toulouse, Toulouse, 31300, France

Corresponding authors:
Zhongxiong Lai
laizx01@163.com

Xu XuHan
xxuhan@163.com

Abstract: Betalains are abundant in amaranth plants. Additionally, the betalain molecular structure and metabolic pathway differ from those of betanin in beet plants. To date, only a few studies have examined the regulatory roles of miRNAs in betalain biosynthesis in plants. Thus, we constructed small RNA libraries for the red and green sectors of amaranth leaves to identify miRNAs associated with betalain biosynthesis. We identified 198 known and 41 novel miRNAs. Moreover, 216 miRNAs were distributed in 44 miRNA families, including miR156, miR159, miR160, miR166, miR172, miR319, miR167, miR396, and miR398. An analysis of all unigene sequences in an amaranth transcriptome database resulted in the detection of 493 target genes for the 239 screened miRNAs. The targets included *SPL2*, *ARF18*, *ARF6*, and *NAC*. A quantitative real-time polymerase chain reaction validation of 20 miRNAs and nine target genes revealed expression-level differences between the red and green sectors of amaranth leaves. This study involved the application of an Illumina sequencing platform to identify miRNAs regulating betalain metabolism in amaranth plants. The data presented herein may provide insights into the molecular mechanisms underlying the regulation of betalain biosynthesis in amaranth and other plant species.

Key words: *Amaranthus tricolor*; betalains; miRNA; target gene; Illumina sequencing platform

Introduction

A microRNA (miRNA) is a small (20–30 nucleotides) non-coding RNA molecule in eukaryotic organisms (Bartel 2004; Ha & Kim 2014; Sun et al. 2014; Voinnet 2009). Its primary function involves the down-regulation of gene expression at the post-transcriptional level (e.g., translation inhibition, target gene cleavage, and epigenetic modification) (Baulcombe 2004). MicroRNAs are among the most important gene regulatory elements, and help mediate physiological and metabolic processes (Bologna & Voinnet 2014; Ren & Yu 2012), including growth and development (Curaba et al. 2012; Shamimuzzaman & Vodkin 2012; Xie et al. 2015a; Zhang et al. 2017), hormone signaling (Mallory et al. 2005; Wang & Guo 2015; Yan et al. 2016), and

responses to external stresses (Budak et al. 2015; Ding et al. 2013; Ferdous et al. 2015; Ma et al. 2015) .

The regulation of secondary metabolism has been a topic of considerable interest among plant scientists. Plant secondary metabolism enables plants to adapt to environmental conditions during long-term evolution, and substantially affects plant defense mechanisms (Aninbon et al. 2016; Bartwal et al. 2013). Additionally, the practical applications of secondary metabolism include the production of industrial raw materials, such as spices (Schweiggert et al. 2007), pigments (Jiménez-Aguilar et al. 2015), and condiments. Exposure to biotic and/or abiotic stresses may increase secondary metabolite contents (Kong et al. 2017), while also inducing changes to miRNAs and the related gene expression levels (Sosa-Valencia et al. 2017; Sunkar 2010; Xie et al. 2015b). The synthesis of secondary metabolites is mediated by the production, interactions, and regulation of signaling molecules (e.g., reactive oxygen species, Ca^{2+}, and hormones) (Akula & Ravishankar 2011). The expression of the related genes is regulated by the binding of a *trans*-acting factor by a *cis*-regulatory element. Moreover, the synthesis of secondary metabolites is regulated by cross-talk stimulated by environmental conditions, which can activate the secondary metabolism pathway *via* signal transduction pathways (Patra et al. 2013).

Previous studies revealed that miRNA can influence plant secondary metabolism by regulating hormone synthesis and signal transduction (Gupta et al. 2017a; Liu et al. 2016). When plants are exposed to biotic and abiotic stresses, the coordinated responses of various hormone signal transduction pathways or the effects of hormones on other signal transduction pathways regulate plant stress tolerance and secondary metabolism (Shriram et al. 2016). Moreover, hormone-regulated miRNAs mediate the synthesis of secondary metabolites by targeting transcription factors, synthetases, and signal transduction elements related to hormones (Li & Zachgo 2013; Liu et al. 2009).

Mahajan et al. (2011) applied bioinformatics techniques to predict that many genes involved in plant secondary metabolism are regulated by miRNAs. These miRNAs target genes

encoding synthetases and transcription factors. For example, miR408 targets a gene encoding a laccase, which is a type of polyphenol oxidase. Meanwhile, miR482 and miR1448 reportedly down-regulate the expression of *PPO* in potato (Chi et al. 2015), and Vv-miR058 regulates the expression of *Vv-PPO* (Ren et al. 2014). Moreover, miRNAs can regulate flower color, with miR171, miR166i, miR159e, miR845, and miR396e more abundant in white roses than red roses (Kim et al. 2012; Roy et al. 2016). Additionally, miR396e targets the gene encoding cytochrome P450, whose function is related to flower color. Earlier studies concluded that cytochrome P450 is involved in betalain metabolism (Hatlestad et al. 2012; Sunnadeniya et al. 2016). Furthermore, miR156, miR166, miR414, and miR2102 of *Panicum virgatum* target the MYB transcription factor genes (Xie et al. 2014). The R2R3-MYB transcription factor is involved in phenylpropane metabolism and betalain metabolism (Hatlestad et al. 2015; Lloyd et al. 2017; Polturak et al. 2016; Xie et al. 2016). Some miRNAs have been observed to regulate anthocyanin biosynthesis (Bai et al. 2017; Gou et al. 2011; Gupta et al. 2017b; Jia et al. 2015; Liu et al. 2016; Qu et al. 2016; Sun et al. 2017; Wang et al. 2016). Thus, enhancing miR156 activity may promote anthocyanin accumulation, while inhibiting miR156 activity may lead to increased flavonol contents. This is because miR156 targets the gene encoding SPL9, which is a protein that can destabilize the MYB-bHLH-WD40 transcriptional activation complex to negatively regulate the anthocyanin biosynthetic pathway. However, it remains unclear which miRNAs regulate betalain metabolism.

There are considerable differences in the molecular structures of betalains and anthocyanins, which have never been observed in the same plant (Tanaka et al. 2008). Betalains are water-soluble, nitrogen-containing pigments that exist in some species of the order Caryophyllales and higher fungi (Hatlestad & Lloyd 2015; Khan & Giridhar 2015). Additionally, betalains exhibit stronger antioxidant activities than anthocyanins. The biosynthesis of betalains in plants is initiated by tyrosinase, with a series of enzymatic reactions as well as spontaneous and non-enzymatic additions ultimately resulting in the production of betalains (Hatlestad & Lloyd 2015). The molecular mechanism underlying betalain biosynthesis is similar to that

regulating anthocyanin biosynthesis, with both involving the same regulatory factors and similar synthesis conditions, effect factors, and functional components (Celli & Brooks 2016; Davies 2015; Sakuta 2014).

There have been several relatively recent investigations regarding betalain metabolism (Brockington et al. 2015; Harris et al. 2012; Hatlestad et al. 2015; Hatlestad et al. 2012; Sunnadeniya et al. 2016). However, these studies focused on key enzymes and transcription factors, such as polyphenol oxidase (PPO), 3, 4-dihydroxyphenylalanine (DOPA), DOPA 4,5-Dioxygenase (DODA), and MYB. However, there have been no reports describing miRNA-mediated regulation of betalains.

Amaranthus tricolor L. is an annual herb that is rich in betalains, which are the characteristic pigments of amaranth plants (Shukla et al. 2003; Strack et al. 2003). Amaranth species grow relatively quickly and produce a large biomass. Additionally, a convenient pigment-extraction method has been developed for these plants. Consequently, *Amaranthus* species are a valuable alternative to beets as a source of natural betalains (Cai et al. 2005). Therefore, we used the Illumina HiSeq 2500 sequencing platform and bioinformatics techniques to analyze *A. tricolor* to identify miRNAs and their target genes related to betalain metabolism. We herein describe the identified miRNAs and their target genes in terms of their roles in amaranth betalain metabolism, which may provide the basis for future studies aimed at enhancing betalain production.

Materials and Methods

Materials

Leaves were collected from ten 30-day-old *A. tricolor* cv. Dahong seedlings, after which RNA was extracted from the red sectors (RS) and green sectors (GS). The RNA samples were subsequently used for a transcriptome analysis.

Methods

Total RNA extraction and quality assessment

Total RNA was extracted from the leaf RS and GS samples using Trizol reagent (Invitrogen, USA). The extracted RNA was treated with DNase I to remove genomic DNA and then quantified with a NanoDrop 2000 spectrophotometer (Thermo, USA). The RNA integrity was assessed using a 2100 Bioanalyzer (Agilent Technologies) as well as by denaturing agarose gel electrophoresis with ethidium bromide staining. Subsequent analyses were conducted with RNA samples that satisfied the following criteria: A260/A280 ratio = 1.9–2.1; 28S/18S RNA ratio > 1.5; and RNA integrity value ≥ 8.0. At least 1.5 μg total RNA (≥ 250 ng μl^{-1}) was used per analysis.

Construction, detection, and sequencing of small RNA libraries

After analyzing each RNA sample, we constructed two cDNA libraries according to the Small RNA Sample Prep Kit (Illumina). The cDNA library quality and insert fragment length were checked using the 2100 Bioanalyzer. The cDNA libraries were sequenced with the Illumina HiSeq 2500 single-end 50-nucleotide system.

Bioinformatics analysis of small RNAs from *Amaranthus tricolor* leaves

The raw sequencing data for each cDNA library were transformed by base calling into raw data or raw reads, which were stored in a FASTQ file format. To ensure the data were accurately analyzed, adapter fragments, sequences with a poly-A tail, low-quality reads with ambiguous bases ("N"), and reads with < 18 or > 30 bases were removed from the raw reads to yield the clean reads. Sequence length distribution and quantity were also calculated. The clean reads were then aligned with the sequences available in the Silva, GtRNAdb, Rfam, and Repbase databases. We obtained unannotated reads for the miRNA, but not the ribosomal RNA (rRNA), transfer RNA (tRNA), small nuclear RNA (snRNA), small nucleolar RNA (snoRNA), non-coding RNA, and repetitive sequences. We identified the putative amaranth miRNAs by searching the miRBase v21.0 database. Meanwhile, the sRNAs that were not mapped to any pre-miRNAs in

the miRBase or classified into any categories in the Rfam (11.0) database were predicted with the miRDeep2 program and the amaranth transcriptome data.

Bioinformatics-based prediction of amaranth miRNA target genes

With the known and novel amaranth miRNA sequences as probes, the psRNA Target (http://bioinfo3.noble.org/ps RNATarget/) program was used to search all unigene sequences in the amaranth transcriptome database to identify candidate target genes (Zhang 2005). The miRNA target genes were predicted based on the following criteria: (1) there were no more than four mismatches between the sRNA and target gene (G–U pair = 0.5 mismatch); (2) the miRNA/target gene complex should not consist of more than two adjacent sites with mismatches; (3) in the miRNA/target gene complex, the adjacent sites at positions 2–12 from the 5′ end of the miRNA cannot be mismatched; (4) in the miRNA/target gene complex, the adjacent sites at positions 11 and 12 cannot be mismatched; (5) in the miRNA/target gene complex, positions 1–12 from the 5′ end of the miRNA should not include more than 2.5 mismatches; and (6) the minimum free energy (MFE) of the miRNA/target gene complex should not be less than 75% of the MFE when the miRNA is combined with the optimal complement. Genes were identified as candidate miRNA targets when the final sequence score was ≤ 5. The putative miRNA target genes were functionally characterized by using their sequences as queries for BLASTX searches of nonredundant databases, such as the GO, KEGG, nr, Pfam, KOG, and COG databases.

Quantitative real-time polymerase chain reaction analysis of miRNAs and their target genes in the red and green sectors of amaranth leaves

A quantitative real-time polymerase chain reaction (qRT-PCR) assay was used to validate the accuracy of the miRNAs in the sRNA database as well as their target genes. The analyzed RNA samples were the same as those used for the sRNA library construction. The qRT-PCR was completed as previously described (Liu et al., 2011). We examined 21 miRNAs and nine target genes using the LightCycler 480 qPCR instrument (Roche Applied Science, Switzerland), with all reactions completed in triplicate. Of the 21 miRNAs, miR390 was the most stably expressed

in the GS (17.69 ± 0.01 cycles) and RS (17.69 ± 0.13 cycles) of amaranth leaves. Thus, it was used as the reference for miRNA quantification. Additionally, *EF1a* was used as the reference for target gene quantification. Gene expression levels in the GS were set to 1.0.

Data analysis

Data analyses were completed using the SPSS software package (version 19) (SPSS, Chicago, IL, USA). All figures were prepared with GraphPad Prism 6.01 (GraphPad Software Inc., USA).

Results

Distribution and quantity of small RNAs in the red and green sectors of amaranth leaves

The cDNA libraries representing the sRNAs in the RS and GS of amaranth leaves were sequenced using high-throughput Illumina sequencing technology. We obtained 39,521,807 clean reads after removing adapters, low-quality reads, reads with < 18 or > 30 nucleotides, and reads with poly-A sequences (Table 1). Among these reads, 24,581,600 (62.20%) were common to the RS and GS, while 9,618,206 (24.34%) and 5,322,001 (13.47%) reads were specific to the GS and RS, respectively. A total of 15,038,207 unique reads were obtained, of which 1,963,804 (13.06%), 8,198,440 (54.52%), and 4,875,963 (32.42%) were common to both sectors, specific to the GS, and specific to the RS, respectively.

An analysis of the sRNA data revealed 22,196,710 and 17,325,097 clean reads as well as 10,162,244 and 6,839,767 unique sRNAs for the GS and RS, respectively. The distribution and quantity of sRNAs based on the clean reads for the GS and RS were calculated. There were considerably fewer sRNAs than other known sRNA types, including rRNA, siRNA, snRNA, snoRNA, and tRNA. Most of the unique sRNAs, including miRNAs, for the GS (85.99%) and RS (73.46%) were unannotated. Therefore, we speculated that many novel sRNAs are involved in the regulation of betalain metabolism. Alternatively, in-depth studies of plant sRNA libraries may be needed to clarify the betalain metabolic process in amaranth plants. The distribution and quantity of the different types of sRNA are presented in Table 2

Length distribution of clean and unique reads for small RNAs in the red and green sectors of amaranth leaves

We observed substantial variability in the length of amaranth sRNA sequences (Supplementary Fig. 1 and Table 3). The clean and unique reads were mainly 21–24 nucleotides long, which is consistent with the read lengths in most angiosperms. The different sRNA lengths in plants may be related to the corresponding function and how the target genes are regulated (Wu et al. 2010). Using Bowtie software, the clean reads were mapped to the *A. tricolor* transcriptome (SRA: SRR5930345). Additionally, similar length distributions were observed for the clean and unique reads.

Identification of known miRNAs in the red and green sectors of amaranth leaves

To identify known miRNAs related to betalain metabolism in *A. tricolor*, the two libraries were compared with the whole mature plant miRNA sequences in the miRBase 21.0 database (http://www.mirbase.org/ftp.shtml). Of the 198 known mature miRNAs that were identified (including miR156, miR159, miR166, miR168, miR408, and miR319), 191 were detected in the GS and RS libraries. Most of the miRNAs were 21 nucleotides long. Unexpectedly, none of the known miRNAs were longer than 22 nucleotides. The length distributions (18–22 nucleotides) of the known miRNAs are provided in Supplementary Fig. 2A.

Prediction of novel miRNAs in the red and green sectors of amaranth leaves

Novel miRNAs can be predicted based on the important characteristics of the inner stem-loop structure. The sequences of miRNAs from the same family are highly conserved in most plant species. An analysis of an amaranth transcriptome sequence database using the miRDeep2 program enabled the prediction of novel amaranth miRNAs. We ultimately obtained 41 candidate miRNAs, with all but one present in the GS and RS. These novel miRNAs were 21–24 nucleotides long (Supplementary Fig. 2B), with 20 of them (48.78%) comprising 24 nucleotides. Moreover, we observed a bias in the first base of miRNAs depending on the sequence length. For

example, the first base tended to be C, U, U/A, and A in miRNAs with 21, 22, 23, and 24 nucleotides, respectively (Supplementary Fig. 3).

Analysis of predicted miRNA families in the red and green sectors of amaranth leaves

Using the SSEARCH program, which is useful for finding a short sequence within a miRNA library, we compared the obtained miRNAs with the mature miRNAs in the miRBase 21.0 database (comparison parameter set to E−10). The miRNAs were then classified into known miRNA families according to the lowest E-value for the comparisons. Of the 239 miRNAs, 216 were distributed in 44 miRNA families. The number of miRNAs differed considerably among the various families, with half containing only one member. The other 22 families comprised multiple members, including 16 families with 2–9 members, and six families with ≥ 10 members. The family with the most identified members was miR166 (31 members), followed by miR396 (20 members), miR159/319 (30 members), miR156 (14 members), miR167 (13 members), and miR172 (13 members). Additionally, some novel miRNAs were classified into known miRNA families. The possibility these miRNAs represent new members of specific *A. tricolor* miRNA families will need to be verified (Supplementary Table S1).

Prediction of miRNA target genes in the red and green sectors of amaranth leaves

To investigate the regulatory effects of miRNAs on gene expression, we used the TargetFinder program to identify the genes targeted by miRNAs. With the amaranth miRNA sequences as probes, we detected 491 targets for 218 of the 232 screened miRNAs from the GS. Additionally, 480 targets were identified for 217 of the 231 screened miRNAs from the RS. Meanwhile, 249 targets were revealed for 183 of the 198 known miRNAs from both sectors. Moreover, 255 targets were identified for the 41 novel miRNAs in both sectors. A total of 493 targets were detected for 224 of the 239 screened miRNAs in both sectors (Table 4). Several miRNAs targeted multiple genes, suggesting these miRNAs have diverse functions. The potential targets

of the known miRNAs among different plant species were either known or newly identified.

To clarify the target gene functions, we conducted BLAST searches of the nr, Swiss-Prot, GO, COG, KEGG, KOG, and Pfam databases, using the predicted target gene sequences as queries. Of the 493 target genes, 267 were functionally annotated. Thus, many of the genes may have unknown regulatory functions related to betalain metabolism in amaranth plants. A total of 96, 146, 102, 146, 212, 191, and 267 target genes were annotated based on searches of the COG, GO, KEGG, KOG, Pfam, Swiss-Prot, and nr databases, respectively. The annotated target genes included those encoding squamosa promoter-binding-like protein 2 (SPL2), auxin response factor 18 (ARF18), copper/zinc superoxide dismutase (CSD), auxin response factor 6 (ARF6), AP2-like ethylene-responsive transcription factor, bZIP transcription factor 60, and NAC. An individual miRNA may simultaneously regulate multiple targets, while a target gene may be regulated by multiple miRNAs. The accuracy of the miRNA target gene predictions will need to be verified in future studies.

Of the 146 target genes functionally annotated based on the GO database, 424, 284, and 173 genes were associated with the biological process, cellular component, and molecular function categories, respectively. Additionally, 19 target genes were annotated according to the COG database. The 25 COG functional annotations involved 14 COG functional categories, including ‘General function prediction only’, ‘Translation, ribosomal structure, and biogenesis’, ‘Replication, recombination, and repair’, ‘Transcription’, ‘Post-translational modification, protein turnover, and chaperones’, and ‘Secondary metabolite biosynthesis, transport, and catabolism’ (Supplementary Fig. 4). Similar functional annotations were observed when the KOG database was searched (Supplementary Fig. 5).

Analysis of miRNAs based on the read counts for the red and green sectors of amaranth leaves

The large number of sequences generated by high-throughput sequencing enabled the use of read counts in libraries to estimate miRNA abundance. Relative to the corresponding expression levels in the GS, the most up-regulated miRNAs in the RS were miR159, miR160, miR166,

289 miR172, and miR319 family members, while the most down-regulated miRNAs were miR156,
290 miR167, and miR398 family members. Additionally, miRNAs from the same family tended to
291 exhibit similar expression levels (Fig. 1).

292 Expression analyses completed based on the normalized read counts for each miRNA
293 revealed differences in the abundances of the 239 miRNAs. We observed that 162 miRNAs in
294 the GS and 158 miRNAs in the RS had more than 50 transcripts per million (TPM). Additionally,
295 miR159, miR160, miR166, miR167, and miR396 were abundant in the two libraries. The most
296 abundant miRNA was miR159, with 108,760 and 148,925 TPM in the GS and RS libraries,
297 respectively.

298 Analysis of differential miRNA expression based on transcripts per million in
299 the red and green sectors of amaranth leaves

300 Differences in miRNA expression levels between the RS and GS were analyzed using the online
301 IDEG6 program, with $|\log_2 (FC)| \geq 1$ and FDR ≤ 0.05 as the criteria for identifying differentially
302 expressed miRNAs. Of the 61 miRNAs differentially expressed between the RS and GS, 31 were
303 up-regulated and 30 were down-regulated in the RS relative to the expression level in the GS.
304 Additionally, 23 and 38 differentially expressed miRNAs were novel and known miRNAs,
305 respectively. Among the 23 novel miRNAs, only two were up-regulated and 21 were down-
306 regulated in the RS relative to the expression level in the GS. Of the 38 known miRNAs, 29 were
307 up-regulated and nine were down-regulated in the RS relative to the expression level in the GS.
308 Of the known miRNAs, the expression levels of the miR159, miR164, miR319, miR166,
309 miR171, and miR395 family members were up-regulated, whereas the expression levels of the
310 miR156, miR167, miR390, miR393, and miR398 family members were all down-regulated.

311 Functional analysis of the target genes of the differentially expressed miRNAs

312 The 267 annotated miRNA target genes included 95 that were the targets of differentially
313 expressed miRNAs between the RS and GS. We observed that 40, 56, 39, 49, 73, 64, and 95
314 target genes were annotated based on the COG, GO, KEGG, KOG, Pfam, Swiss-Prot, and nr

databases, respectively.

The 56 target genes annotated according to the GO database were distributed in the three main functional categories (i.e., biological process, cellular component, and molecular function). In the biological process category, most genes were annotated with the metabolic process (37) and cellular process (36) terms. In the cellular component category, most genes were annotated with the cell (23), cell part (23), and organelle (15) terms. Among the genes in the molecular function category, most were annotated with the catalytic activity (26) and binding (25) terms (Supplementary Table S2)

The 39 target genes functionally annotated according to the KEGG database were associated with diverse pathways, including Biosynthesis of amino acids (ko01230), Spliceosome (ko03040), Peroxisome (ko04146), Pyrimidine metabolism (ko00240), Protein export (ko03060), Oxidative phosphorylation (ko00190), and other metabolic pathways.

Analysis of miRNA and target gene expression levels in the red and green sectors of amaranth leaves

To clarify the role of miRNAs in betalain metabolism in *A. tricolor*, the expression levels of 20 miRNAs in the RS and GS of leaves were analyzed with a qRT-PCR assay (Fig. 2). We observed that miR156a/b/c, miR164b, miR166a/b/e-3p/u, miR172d-3p, miR319, and miR396b expression levels were higher in the RS than in the GS. In contrast, miR156a/e-3p, miR162a-5p, miR167, and miR396a-3p expression levels were lower in the RS than in the GS.

Among the novel miRNAs, amt-miR1 was expressed more highly in the RS than in the GS, and was classified in the miR159 family according to SSEARCH. Thus, amt-miR1 appears to be a new member of the miR159 family. Meanwhile, amt-miR3, amt-miR6, and amt-miR26 were expressed at lower levels in the RS than in the GS. According to SSEARCH, amt-miR3 and amt-miR6 belonged to the miR169_1 and miR408 families, respectively. Moreover, amt-miR26 was considered to be a new miRNA.

The regulation of diverse biological processes by miRNAs is mediated by mechanisms

such as the repression of translation and cleavage of target mRNAs. For example, miR156 targeted the genes encoding SPL2, SPL6, and SPL9. Our data revealed that *SPL2* expression was lower in the RS than in the GS, while the *SPL6* and *SPL9* expression levels exhibited the opposite pattern. Additionally, miR164 and miR160 targeted the *NAC* and *TCP4* genes, respectively. The expression levels of these two genes were lower in the RS than in the GS. Meanwhile, the miR166, miR164, and miR396 target genes were *HD-Zip*, *GRF2/9*, and *CSD*, respectively. These genes were more highly expressed in the RS than in the GS (Fig. 3).

Discussion

Identification of miRNAs in amaranth leaves may provide valuable information regarding leaf development and the regulation of secondary metabolism

MicroRNAs are important regulators of many plant biological processes, including leaf development (Mecchia et al. 2013), anther development (Yang et al. 2016; Zhang et al. 2016), cell differentiation (Rodriguez et al. 2016), flowering (Huo et al. 2016; Wang 2014), floral organ morphogenesis (Blein & Laufs 2016; Liu et al. 2017; Reinhart et al., 2002; Rhoades et al., 2002), and responses to environmental stresses (Ohama et al. 2017; Stief et al. 2014). Recent studies have confirmed that miRNAs directly regulate plant secondary metabolism (Liu Juan et al. 2017; Shen et al. 2017; Sun et al. 2017; Yue et al. 2017), with consequences for the production of flavonoids, anthocyanins, and alkaloids. However, there is less information available for the miRNAs involved in betalain metabolism than for the miRNAs related to other secondary metabolic activities. Only some of the miRNAs associated with PPO and P450 production have been identified in a few plant species, including potato (Chi et al. 2015) and grapevine (Ren et al. 2014). Moreover, it is unclear whether these miRNAs influence betalain metabolism. Additionally, although the miRNAs expressed during the amaranth flowering stage have been identified based on sRNA sequencing data (Liu et al. 2011), the miRNAs regulating betalain

368 contents have yet to be characterized. Therefore, identifying and characterizing the miRNAs
369 involved in betalain metabolism in amaranth plants will provide valuable information for
370 elucidating the relevant regulatory mechanism.

371 In this study, we applied an Illumina sequencing platform to further characterize the role
372 of miRNAs affecting the betalain metabolism of amaranth plants. We generated 39,521,807
373 clean reads from the two sequenced libraries. Additionally, 22,196,710 and 17,325,097 clean
374 reads were obtained for the GS and RS, respectively. The abundant sRNA data were analyzed to
375 clarify betalain metabolism in amaranth plants. However, 85.99 and 73.46% of the clean reads
376 for the GS and RS were unannotated, respectively. The results implied that many novel sRNAs
377 help regulate betalain metabolism or more information regarding the constructed sRNA libraries
378 will be needed to further clarify the mechanism mediating betalain metabolism in amaranth
379 plants.

380 The clean and unique reads were mainly 21–24 nucleotides long, which is similar to the
381 findings regarding the corresponding reads in most angiosperms. Additionally, the 24-nt sRNAs
382 were the most common sRNAs. This is consistent with the observations of previous studies
383 involving tomato (Zuo et al. 2012) and persimmon (Luo et al. 2015), but differs from the
384 findings of another study involving trifoliate orange (Zhang et al. 2014). Therefore, the length
385 distribution of sRNAs appears to differ among plant species.

386 There is considerable interest in the regulatory mechanisms underlying plant secondary
387 metabolism. The production of secondary metabolites contributed to the adaptation of plant
388 species to the environment during long-term evolution, while also positively affecting plant
389 defenses (Aninbon et al. 2016; Bartwal et al. 2013). Secondary metabolism has also been
390 important for the production of industrial raw materials, such as spices (Schweiggert et al. 2007),
391 pigments (Jiménez-Aguilar et al. 2015), and condiments. A previous bioinformatics-based study
392 predicted that many genes (e.g., synthetase and transcription factor genes) involved in plant
393 secondary metabolism are regulated by miRNAs (Mahajan et al. 2011). In the current study, we
394 obtained 239 miRNAs associated with amaranth betalain metabolism. These miRNAs may

provide valuable information regarding the relevant regulatory mechanism. The functions of the other sRNAs will need to be explored further.

Members of the miRl66 family may be important positive regulators of betalain biosynthesis in amaranth plants

The miRl66 family is large and highly conserved among diverse plant species. This miRNA family is critical for plant growth and stress responses. Earlier investigations revealed that miR166 down-regulates the expression of the genes encoding HD-ZIP-type transcription factors, including Phabulosa, in *Arabidopsis thaliana* and maize (Chuck & O'Connor 2010; Juarez et al. 2004). There are relatively recent reports confirming that miR166 affects anthocyanin metabolism. For example, miR165/166 negatively regulate anthocyanin biosynthesis (JIA et al. 2013). Another study applied an Illumina sequencing platform to identify candidate target genes of the miR166 family members in blueberry. The identified genes encoded various proteins, including class III homeodomain leucine zipper (HD-ZIP) family members, ethylene response factor (ERF), squamosa promoter-binding-like protein (SPB/SPL), MADS, basic helix-loop-helix (bHLH), R2R3 MYB, and auxin response factor (ARF) (Yue et al. 2017). These transcription factors reportedly affect anthocyanin biosynthesis regulatory pathways in some plant species (Gou et al. 2011; Jia et al. 2015; Yang et al. 2013). In our study, we predicted that miR166 targets *HD-ZIP* genes. Additional research will be needed to verify the genes targeted by miR166. Furthermore, all miR166 family members were more highly expressed in the RS than in the GS, while the expression levels of the corresponding target genes exhibited the opposite pattern. The observed expression trends for miR166 and *HD-ZIP* were consistent with the results of a previous rice study (Liu et al. 2009). Meanwhile, gibberellins can strongly inhibit miR166 expression and the accumulation of anthocyanins (Ahmad & Iqbal 2012) or betalains (Zheng et al. 2016). Therefore, we speculate that miR166 may positively regulate betalain biosynthesis, while having the opposite effect on anthocyanin biosynthesis. This implies that different metabolic pathways mediate anthocyanin and betalain production, thereby enabling both metabolites to co-exist in plants. Alternatively, miR166 may regulate the expression of

other target genes related to betalain metabolism. These possibilities will need to be experimentally validated.

Members of the miRl56 family may negatively regulate betalain biosynthesis in amaranth

The SPL and R2R3-MYB transcription factors negatively regulate flavonoid biosynthesis (Gou et al. 2011; Li & Zachgo 2013; Xu et al. 2015). In *A. thaliana*, miR156 targets *SPL*, which encodes a protein that negatively regulates anthocyanin production by destabilizing the MYB-bHLH-WD transcriptional activation complex, which affects the anthocyanin biosynthetic pathway (Gou et al. 2011). Improving miR156 activity may promote anthocyanin accumulation, while inhibiting miR156 activity may induce flavonol accumulation. The R2R3-MYB transcription factor helps control phenylpropane metabolism (Craven-Bartle et al. 2013) and betalain metabolism (Hatlestad et al. 2015). However, unlike MYBs regulating anthocyanin contents, the MYBs related to betalain metabolism cannot interact with the bHLH of the heterologous anthocyanin MYB-bHLH-WD complexes (Hatlestad et al. 2015). We observed that miR156 was expressed at lower levels in the RS than in the GS, resulting in up-regulated *SPL* expression in the RS, which ultimately prevented MYB from interacting with bHLH. Finally, betalain was biosynthesized in the RS.

Many miRNAs help regulate betalain biosynthesis in amaranth plants

We detected many miRNAs in our sRNA libraries, including miR159, miR319, miR408, miR172, miR482, and miR858. These miRNAs were previously revealed to influence the anthocyanin biosynthetic pathway or other pathways (Boke et al. 2015; Pilon 2011; Wang et al. 2016; Yang et al. 2013). The expression trends of miR159 and miR319 as well as their corresponding targets were similar to those of miR166 and its target. Additionally, miR159 was confirmed to regulate gibberellin signaling and direct the cleavage of mRNA encoding GAMYB-related proteins. Thus, miR159 and miR166 appear to similarly contribute to the regulation of betalain metabolism. Meanwhile, miR408 targets a gene encoding a laccase, which is a type of

polyphenol oxidase (Ravet et al. 2011). Additionally, miR482 and miR1448 down-regulate *PPO* expression levels in potato (Chi et al. 2015), while Vv-miR058 regulates *Vv-PPO* expression (Ren et al. 2014). The PPO enzyme is believed to catalyze the first step of betalain biosynthesis. There are many miRNAs involved in the regulation of betalain metabolism in amaranth plants (Fig. 4).

Conclusions

We constructed sRNA libraries for the RS and GS of amaranth leaves to identify the miRNAs associated with betalain biosynthesis. We identified 198 known and 41 novel miRNAs. Of the 239 screened miRNAs, 224 were observed to target 493 genes in the RS and GS. These targets included *SPL2*, *ARF18*, *ARF6*, and *NAC*. Moreover, miR156a/b/c, miR164b, miR166a/b/e-3p/u, miR172d-3p, miR319, and miR396b expression levels were higher in the RS than in the GS. In contrast, miR156a/e-3p, miR162a-5p, miR167, and miR396a-3p expression levels were lower in the RS than in the GS. Furthermore, a novel miRNA, amt-miR1, was more highly expressed in the RS than in the GS, while amt-miR3, amt-miR6, and amt-miR26 expression levels were lower in the RS than in the GS.

Acknowledgments

This study was supported by the Natural Science Foundation of Fujian Province (2013J05045), the Research Fund of the Education Department of Fujian Province (JA12098), the Youth Academic Backbone of the Horticultural College at Fujian Agriculture and Forestry University (61201400722), and the Major Science and Technology Program of Fujian Province (2015NZ0002-1). We thank Liwen Bianji, Edanz Editing China (www.liwenbianji.cn/ac) for editing the English text of a draft of this manuscript.

Author Contributions Statement

Liu conceived the study, conducted experiments, and wrote the manuscript. Peng, Pan, and Cheng conducted the qRT-PCR experiment and analyzed the data. Wang and Zhao prepared the amaranth and RNA. Zhang and Lin prepared the RNA extracting reagents. Lai and XuHan conceived the study and helped revise the manuscript. All authors read and approved the final version of the manuscript.

References

Ahmad A, and Iqbal M. 2012. Small regulatory RNAs in the regulation of Nitrogen Metabolism of rice under Nitrogen limitation.

Akula R, and Ravishankar GA. 2011. Influence of abiotic stress signals on secondary metabolites in plants. *Plant Signaling & Behavior* 6:1720-1731.

Aninbon C, Jogloy S, Vorasoot N, Patanothai A, Nuchadomrong S, and Senawong T. 2016. Effect of end of season water deficit on phenolic compounds in peanut genotypes with different levels of resistance to drought. *Food Chem* 196:123-129. http://dx.doi.org/10.1016/j.foodchem.2015.09.022

Bai S, Sun Y, Qian M, Yang F, Ni J, Tao R, Li L, Shu Q, Zhang D, and Teng Y. 2017. Transcriptome analysis of bagging-treated red Chinese sand pear peels reveals light-responsive pathway functions in anthocyanin accumulation. *Sci Rep* 7:63. 10.1038/s41598-017-00069-z

Bartel DP. 2004. MicroRNAs. *Cell* 116:281-297. http://dx.doi.org/10.1016/S0092-8674(04)00045-5

Bartwal A, Mall R, Lohani P, Guru SK, and Arora S. 2013. Role of Secondary Metabolites and Brassinosteroids in Plant Defense Against Environmental Stresses. *Journal of Plant Growth Regulation* 32:216-232. 10.1007/s00344-012-9272-x

Baulcombe D. 2004. RNA silencing in plants. *Nature* 431:356-363.

Blein T, and Laufs P. 2016. MicroRNAs (miRNAs) and plant development. *eLS*.

Boke H, Ozhuner E, Turktas M, Parmaksiz I, Ozcan S, and Unver T. 2015. Regulation of the alkaloid biosynthesis by miRNA in opium poppy. *Plant Biotechnol J* 13:409-420. 10.1111/pbi.12346

Bologna NG, and Voinnet O. 2014. The Diversity, Biogenesis, and Activities of Endogenous Silencing Small RNAs in Arabidopsis. *Annual Review of Plant Biology* 65:473-503. 10.1146/annurev-arplant-050213-035728

Brockington SF, Yang Y, Gandia‐Herrero F, Covshoff S, Hibberd JM, Sage RF, Wong GK, Moore MJ, and Smith SA. 2015. Lineage‐specific gene radiations underlie the evolution of novel betalain pigmentation in Caryophyllales. *New Phytologist* 207:1170-1180.

Budak H, Kantar M, Bulut R, and Akpinar BA. 2015. Stress responsive miRNAs and isomiRs in cereals. *Plant Science* 235:1-13.

Cai Y, Sun M, and Corke H. 2005. HPLC characterization of betalains from plants in the amaranthaceae. *Journal of chromatographic science* 43:454-460.

Celli GB, and Brooks MS-L. 2016. Impact of extraction and processing conditions on betalains and comparison of properties with anthocyanins—A current review. *Food Research International*.

Chi M, Liu C, Su Y, Tong Y, and Liu H. 2015. Bioinformatic prediction of upstream microRNAs of PPO and novel microRNAs in potato. *Canadian Journal of Plant Science* 95:871-877. 10.1139/CJPS-2014-308

Chuck G, and O'Connor D. 2010. Small RNAs going the distance during plant development. *Curr Opin Plant Biol* 13:40-45. 10.1016/j.pbi.2009.08.006

Craven - Bartle B, Pascual M, Cánovas FM, and Ávila C. 2013. A Myb transcription factor regulates genes of the phenylalanine pathway in maritime pine. *The Plant Journal* 74:755-766.

Curaba J, Spriggs A, Taylor J, Li Z, and Helliwell C. 2012. miRNA regulation in the early development of barley seed. *BMC Plant Biol* 12. 10.1186/1471-2229-12-120

Davies KM. 2015. S wapping one red pigment for another. *Nat Genet* 47:5.

Ding Y, Tao Y, and Zhu C. 2013. Emerging roles of microRNAs in the mediation of drought stress response in plants. *Journal of experimental botany* 64:3077-3086.

Ferdous J, Hussain SS, and Shi BJ. 2015. Role of microRNAs in plant drought tolerance. *Plant Biotechnol J* 13:293-305.

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

Gupta OP, Karkute SG, Banerjee S, Meena NL, and Dahuja A. 2017a. Contemporary Understanding of miRNA-Based Regulation of Secondary Metabolites Biosynthesis in Plants. *Front Plant Sci* 8:374. 10.3389/fpls.2017.00374

Gupta OP, Nigam D, Dahuja A, Kumar S, Vinutha T, Sachdev A, and Praveen S. 2017b. Regulation of Isoflavone Biosynthesis by miRNAs in Two Contrasting Soybean Genotypes at Different Seed Developmental Stages. *Front Plant Sci* 8:567. 10.3389/fpls.2017.00567

Ha M, and Kim VN. 2014. Regulation of microRNA biogenesis. *Nat Rev Mol Cell Biol* 15:509-524. 10.1038/nrm3838

Harris NN, Javellana J, Davies KM, Lewis DH, Jameson PE, Deroles SC, Calcott KE, Gould KS, and Schwinn KE. 2012. Betalain production is possible in anthocyanin-producing plant species given the presence of DOPA-dioxygenase and L-DOPA. *BMC Plant Biol* 12. Artn 34 10.1186/1471-2229-12-34

Hatlestad GJ, Akhavan NA, Sunnadeniya RM, Elam L, Cargile S, Hembd A, Gonzalez A, McGrath JM, and Lloyd AM. 2015. The beet Y locus encodes an anthocyanin MYB-like protein that activates the betalain red pigment pathway. *Nature Genetics* 47:92-96. 10.1038/ng.3163

Hatlestad GJ, and Lloyd A. 2015. The betalain secondary metabolic network. *Pigments in Fruits and Vegetables*: Springer, 127-140.

Hatlestad GJ, Sunnadeniya RM, Akhavan NA, Gonzalez A, Goldman IL, McGrath JM, and Lloyd AM. 2012. The beet R locus encodes a new cytochrome P450 required for red betalain production. *Nat Genet* 44:816-820. 10.1038/ng.2297

Huo H, Wei S, and Bradford KJ. 2016. DELAY OF GERMINATION1 (DOG1) regulates both seed dormancy and flowering time through microRNA pathways. *Proceedings of the National Academy of Sciences* 113:E2199-E2206.

Jia X, Liang J, Liu Y, Li S, Qiao M, Li R, and Tang G. 2013. Novel functions of MiR165/166 in regulation of anthocyanin biosynthesis and salt stress responses revealed by short tandem target mimic in Arabidopsis thaliana. *2013 Proceedings of the plant biology in China*.

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

Jiménez-Aguilar DM, López-Martínez JM, Hernández-Brenes C, Gutiérrez-Uribe JA, and Welti-Chanes J. 2015. Dietary fiber, phytochemical composition and antioxidant activity of Mexican commercial varieties of cactus pear. *Journal of Food Composition and Analysis* 41:66-73. 10.1016/j.jfca.2015.01.017

Juarez MT, Kui JS, Thomas J, Heller BA, and Timmermans MC. 2004. microRNA-mediated repression of rolled leaf1 specifies maize leaf polarity. *Nature* 428:84-88. 10.1038/nature02363

Khan MI, and Giridhar P. 2015. Plant betalains: Chemistry and biochemistry. *Phytochemistry* 117:267-295. 10.1016/j.phytochem.2015.06.008

Kim J, Park JH, Lim CJ, Lim JY, Ryu JY, Lee BW, Choi JP, Kim WB, Lee HY, Choi Y, Kim D, Hur CG, Kim S, Noh YS, Shin C, and Kwon SY. 2012. Small RNA and transcriptome deep sequencing proffers insight into floral gene regulation in Rosa cultivars. *BMC genomics* 13. Artn 65710.1186/1471-2164-13-657

Kong L, Xie Y, Hu L, Si J, and Wang Z. 2017. Excessive nitrogen application dampens antioxidant capacity and grain filling in wheat as revealed by metabolic and physiological analyses. *Sci Rep* 7:43363. 10.1038/srep43363

Li S, and Zachgo S. 2013. TCP3 interacts with R2R3-MYB proteins, promotes flavonoid biosynthesis and negatively regulates the auxin response in Arabidopsis thaliana. *Plant J* 76:901-913. 10.1111/tpj.12348

Liu Juan , Yuan Yuan, Wang Yaolong, Jiang Chao, Chen Tiying , Zhu Fengjie, Zhao Yuyang , Zhou Junhui , and Luqi H. 2017. Regulation of fatty acid and flavonoid biosynthesis by miRNAs in Lonicera japonica. *Royal Society of Chemistry*:12. 10.1039/C7RA05800D10.1039/c7ra05800d

Liu N, Tu L, Wang L, Hu H, Xu J, and Zhang X. 2017. MicroRNA 157-targeted SPL genes regulate floral organ size and ovule production in cotton. *BMC Plant Biol* 17:7.

Liu Q, Zhang Y-C, Wang C-Y, Luo Y-C, Huang Q-J, Chen S-Y, Zhou H, Qu L-H, and Chen Y-Q. 2009. Expression analysis of phytohormone-regulated microRNAs in rice, implying their regulation roles in plant hormone signaling. *FEBS Lett* 583:723-728. 10.1016/j.febslet.2009.01.020

Liu R, Lai B, Hu B, Qin Y, Hu G, and Zhao J. 2016. Identification of MicroRNAs and Their Target Genes Related to the Accumulation of Anthocyanins in Litchi chinensis by High-Throughput Sequencing and Degradome Analysis. *Front Plant Sci* 7:2059. 10.3389/fpls.2016.02059

Liu S, Lin Y, Chen X, and Lai Z. 2011. Novel and Conserved miRNAs Identification during in vitro Flowering of Amaranthus by Solexa Sequencing . *Chinese Journal of Tropical Crops* 32:1296-1303.

Lloyd A, Brockman A, Aguirre L, Campbell A, Bean A, Cantero A, and Gonzalez A. 2017. Advances in the MYB–bHLH–WD Repeat (MBW) Pigment Regulatory Model: Addition of a WRKY Factor and Co-option of an Anthocyanin MYB for Betalain Regulation. *Plant and Cell Physiology*:pcx075.

Luo Y, Zhang X, Luo Z, Zhang Q, and Liu J. 2015. Identification and characterization of microRNAs from Chinese pollination constant non-astringent persimmon using high-throughput sequencing. *BMC Plant Biol* 15:11. 10.1186/s12870-014-0400-6

Ma C, Burd S, and Lers A. 2015. miR408 is involved in abiotic stress responses in Arabidopsis. *The Plant Journal* 84:169-187.

Mahajan V, Mahajan A, Pagoch SS, Bedi YS, and Gandhi SG. 2011. microRNA mediated regulation of plant secondary metabolism: an in silico analysis. *Journal of Natural Science Biology & Medicine* 2:44.

Mallory AC, Bartel DP, and Bartel B. 2005. MicroRNA-Directed Regulation of Arabidopsis AUXIN RESPONSE FACTOR17 Is Essential for Proper Development and Modulates Expression of Early Auxin Response Genes. *Plant Cell* 17:1360-1375. 10.1105/tpc.105.031716

Mecchia MA, Debernardi JM, Rodriguez RE, Schommer C, and Palatnik JF. 2013. MicroRNA miR396 and RDR6 synergistically regulate leaf development. *Mechanisms of Development* 130:2-13. http://dx.doi.org/10.1016/j.mod.2012.07.005

Ohama N, Sato H, Shinozaki K, and Yamaguchi-Shinozaki K. 2017. Transcriptional regulatory network of plant heat stress response. *Trends Plant Sci* 22:53-65.

Patra B, Schluttenhofer C, Wu Y, Pattanaik S, and Yuan L. 2013. Transcriptional regulation of secondary metabolite biosynthesis in plants. *Biochim Biophys Acta* 1829:1236-1247. 10.1016/j.bbagrm.2013.09.006

Pilon M. 2011. Moving copper in plants. *New Phytologist* 192:305-307.

Polturak G, Breitel D, Grossman N, Sarrion‐Perdigones A, Weithorn E, Pliner M, Orzaez D, Granell A, Rogachev I, and Aharoni A. 2016. Elucidation of the first committed step in betalain biosynthesis enables the heterologous engineering of betalain pigments in plants. *New Phytologist* 210:269-283.

Qu D, Yan F, Meng R, Jiang X, Yang H, Gao Z, Dong Y, Yang Y, and Zhao Z. 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. 10.3389/fpls.2016.00027

Ravet K, Danford FL, Dihle A, Pittarello M, and Pilon M. 2011. Spatiotemporal Analysis of Copper Homeostasis in Populus trichocarpa Reveals an Integrated Molecular Remodeling for a Preferential Allocation of Copper to Plastocyanin in the

Chloroplasts of Developing Leaves. *Plant Physiology* 157:1300-1312. 10.1104/pp.111.183350

Ren G, and Yu B. 2012. Critical roles of RNA-binding proteins in miRNA biogenesis in Arabidopsis. *RNA Biology* 9:1424-1428. 10.4161/rna.22740

Ren GH, Wang BJ, Zhu XD, Mu Q, Wang C, Tao R, and Fang JG. 2014. Cloning, expression, and characterization of miR058 and its target PPO during the development of grapevine berry stone. *Gene* 548:166-173. 10.1016/j.gene.2014.07.021

Rodriguez RE, Schommer C, and Palatnik JF. 2016. Control of cell proliferation by microRNAs in plants. *Current Opinion in Plant Biology* 34:68-76. http://dx.doi.org/10.1016/j.pbi.2016.10.003

Roy S, Tripathi AM, Yadav A, Mishra P, and Nautiyal CS. 2016. Identification and Expression Analyses of miRNAs from Two Contrasting Flower Color Cultivars of Canna by Deep Sequencing. *PloS one* 11:e0147499. 10.1371/journal.pone.0147499

Sakuta M. 2014. Diversity in plant red pigments: anthocyanins and betacyanins. *Plant biotechnology reports* 8:37-48.

Schweiggert U, Carle R, and Schieber A. 2007. Conventional and alternative processes for spice production – a review. *Trends in Food Science & Technology* 18:260-268. http://dx.doi.org/10.1016/j.tifs.2007.01.005

Shamimuzzaman M, and Vodkin L. 2012. Identification of soybean seed developmental stage-specific and tissue-specific miRNA targets by degradome sequencing. *BMC genomics* 13. 10.1186/1471-2164-13-310

Shen EM, Singh SK, Ghosh JS, Patra B, Paul P, Yuan L, and Pattanaik S. 2017. The miRNAome of Catharanthus roseus: identification, expression analysis, and potential roles of microRNAs in regulation of terpenoid indole alkaloid biosynthesis. *Sci Rep* 7:43027. 10.1038/srep43027

Shriram V, Kumar V, Devarumath RM, Khare TS, and Wani SH. 2016. MicroRNAs As Potential Targets for Abiotic Stress Tolerance in Plants. *Front Plant Sci* 7:817. 10.3389/fpls.2016.00817

Shukla S, Pandey V, Pachauri G, Dixit BS, Banerji R, and Singh SP. 2003. Nutritional contents of different foliage cuttings of vegetable amaranth. *Plant Foods for Human Nutrition* 58:1-8. 10.1023/B:QUAL.0000040338.33755.b5

Sosa-Valencia G, Palomar M, Covarrubias AA, and Reyes JL. 2017. The legume miR1514a modulates a NAC transcription factor transcript to trigger phasiRNA formation in response to drought. *Journal of experimental botany* 68:2013-2026. 10.1093/jxb/erw380

Stief A, Altmann S, Hoffmann K, Pant BD, Scheible W-R, and Bäurle I. 2014. Arabidopsis miR156 regulates tolerance to recurring environmental stress through SPL transcription factors. *Plant Cell* 26:1792-1807.

Strack D, Vogt T, and Schliemann W. 2003. Recent advances in betalain research. *Phytochemistry* 62:247-269. Pii S0031-9422(02)00564-2 Doi 10.1016/S0031-9422(02)00564-2

Sun F, Guo G, Du J, Guo W, Peng H, Ni Z, Sun Q, and Yao Y. 2014. Whole-genome discovery of miRNAs and their targets in wheat (Triticum aestivum L.). *BMC Plant Biol* 14:142. 10.1186/1471-2229-14-142

Sun Y, Qiu Y, Duan M, Wang J, Zhang X, Wang H, Song J, and Li X. 2017. Identification of anthocyanin biosynthesis related microRNAs in a distinctive Chinese radish (Raphanus sativus L.) by high-throughput sequencing. *Mol Genet Genomics* 292:215-229. 10.1007/s00438-016-1268-y

Sunkar R. 2010. MicroRNAs with macro-effects on plant stress responses. *Semin Cell Dev Biol* 21:805-811. http://dx.doi.org/10.1016/j.semcdb.2010.04.001

Sunnadeniya R, Bean A, Brown M, Akhavan N, Hatlestad G, Gonzalez A, Symonds VV, and Lloyd A. 2016. Tyrosine hydroxylation in betalain pigment biosynthesis is performed by cytochrome P450 enzymes in beets (Beta vulgaris). *PloS one* 11:e0149417.

Tanaka Y, Sasaki N, and Ohmiya A. 2008. Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. *Plant Journal* 54:733-749. 10.1111/j.1365-313X.2008.03447.x

Voinnet O. 2009. Origin, Biogenesis, and Activity of Plant MicroRNAs. *Cell* 136:669-687. http://dx.doi.org/10.1016/j.cell.2009.01.046

Wang J-J, and Guo H-S. 2015. Cleavage of INDOLE-3-ACETIC ACID INDUCIBLE28 mRNA by microRNA847 upregulates auxin signaling to modulate cell proliferation and lateral organ growth in Arabidopsis. *Plant Cell* 27:574-590.

636 Wang J-W. 2014. Regulation of flowering time by the miR156-mediated age pathway. *Journal of experimental botany* 65:4723-4730.
637 Wang Y, Wang Y, Song Z, and Zhang H. 2016. Repression of MYBL2 by Both microRNA858a and HY5 Leads to the Activation of
638 Anthocyanin Biosynthetic Pathway in Arabidopsis. *Mol Plant* 9:1395-1405. 10.1016/j.molp.2016.07.003
639 Wu L, Zhou H, Zhang Q, Zhang J, Ni F, Liu C, and Qi Y. 2010. MicroRNA Mediates DNA Methylation. *Molecular Cell, ,* 38:465-475.
640 Xie F, Jones DC, Wang Q, Sun R, and Zhang B. 2015a. Small RNA sequencing identifies miRNA roles in ovule and fibre development. *Plant*
641 *Biotechnol J* 13:355-369. 10.1111/pbi.12296
642 Xie F, Stewart CN, Taki FA, He Q, Liu H, and Zhang B. 2014. High-throughput deep sequencing shows that microRNAs play important roles in
643 switchgrass responses to drought and salinity stress. *Plant Biotechnol J* 12:354-366. 10.1111/pbi.12142
644 Xie F, Wang Q, Sun R, and Zhang B. 2015b. Deep sequencing reveals important roles of microRNAs in response to drought and salinity stress in
645 cotton. *Journal of experimental botany* 66:789-804. 10.1093/jxb/eru437
646 Xie L, Liu S, Bai Y, Liu Y, Lin M, Cai S, Zheng X, Xie X, Feng X, Cheng C, Chen Y, and Lai Z. 2016. Cloning and Expression Analysis of
647 Betalain-related Transcription Factor Gene AmMYB1 in Amaranthus tricolor L. . *Acta Botanica Boreali-Occidentalia Sinica* 36:1080-
648 1090.
649 Xu W, Dubos C, and Lepiniec L. 2015. Transcriptional control of flavonoid biosynthesis by MYB-bHLH-WDR complexes. *Trends Plant Sci*
650 20:176-185. 10.1016/j.tplants.2014.12.001
651 Yan J, Zhao C, Zhou J, Yang Y, Wang P, Zhu X, Tang G, Bressan RA, and Zhu J-K. 2016. The miR165/166 Mediated Regulatory Module Plays
652 Critical Roles in ABA Homeostasis and Response in Arabidopsis thaliana. *PLoS Genet* 12:e1006416.
653 Yang F, Cai J, Yang Y, and Liu Z. 2013. Overexpression of microRNA828 reduces anthocyanin accumulation in Arabidopsis. *Plant Cell, Tissue*
654 *and Organ Culture (PCTOC)* 115:159-167. 10.1007/s11240-013-0349-4
655 Yang X, Zhao Y, Xie D, Sun Y, Zhu X, Esmaeili N, Yang Z, Wang Y, Yin G, and Lv S. 2016. Identification and Functional Analysis of
656 microRNAs Involved in the Anther Development in Cotton Genic Male Sterile Line Yu98-8A. *International Journal Of Molecular*
657 *Sciences* 17:1677.
658 Yue J, Lu X, Zhang H, Ge J, Gao X, and Liu Y. 2017. Identification of Conserved and Novel MicroRNAs in Blueberry. *Front Plant Sci* 8.
659 10.3389/fpls.2017.01155
660 Zhang H, Yin L, Wang H, Wang G, Ma X, Li M, Wu H, Fu Q, Zhang Y, and Yi H. 2017. Genome-wide identification of Hami melon miRNAs
661 with putative roles during fruit development. *PloS one* 12:e0180600. 10.1371/journal.pone.0180600
662 Zhang W, Xie Y, Xu L, Wang Y, Zhu X, Wang R, Zhang Y, Muleke EM, and Liu L. 2016. Identification of microRNAs and Their Target Genes
663 Explores miRNA-Mediated Regulatory Network of Cytoplasmic Male Sterility Occurrence during Anther Development in Radish
664 (Raphanus sativus L.). *Front Plant Sci* 7:1054. 10.3389/fpls.2016.01054
665 Zhang X-N, Li X, and Liu J-H. 2014. Identification of Conserved and Novel Cold-Responsive MicroRNAs in Trifoliate Orange (Poncirus
666 trifoliata (L.) Raf.) Using High-Throughput Sequencing. *Plant Molecular Biology Reporter* 32:328-341. 10.1007/s11105-013-0649-1
667 Zhang Y. 2005. miRU: an automated plant miRNA target prediction server. *Nucleic Acids Res* 33:W701-704. 33/suppl_2/W701 [pii]
668 10.1093/nar/gki383
669 Zheng X, Liu S, Cheng C, Guo R, Chen Y, Xie L, Mao Y, Lin Y, Zhang Z, and Lai Z. 2016. Cloning and expression analysis of betalain
670 biosynthesis genes in Amaranthus tricolor. *Biotechnology letters* 38:723.
671 Zuo J, Zhu B, Fu D, Zhu Y, Ma Y, Chi L, Ju Z, Wang Y, Zhai B, and Luo Y. 2012. Sculpting the maturation, softening and ethylene pathway:
672 The influences of microRNAs on tomato fruits. *BMC genomics* 13:7-7. 10.1186/1471-2164-13-7
673
674

Figure Legends

Fig. 1 Read counts of known miRNAs

Fig. 2 Relative expression of miRNAs in the red and green sectors of amaranth leaves
A: Up-regulated miRNAs; B: Down-regulated miRNAs; C: Novel miRNAs

Fig. 3 Relative expression of target genes in the red and green sectors of amaranth leaves

Fig. 4 Schematic diagram of the betalain metabolic pathway network in *Amaranthus tricolor* leaves

Conflict of Interest

The authors declare that they have no conflicts of interest.

Table 1(on next page)

Table1 The summary of high-throughput sequencing data of A. tricolor

Table1 The summary of high-throughput sequencing data of *Am. tricolor*

Types	Total clean reads	Percentage (%)	Unique reads	Percentage (%)
Total_sRNA	39521807	100	15038207	100
GS & RS	24581600	62.2	1963804	13.06
GS_specific	9618206	24.34	8198440	54.52
RS_specific	5322001	13.47	4875963	32.42

Table 2(on next page)

Table 2. Distribution of small RNAs among different categories

Table 2. Distribution of small RNAs among different categories

Types	GS		RS	
	Number	Percentage (%)	Number	Percentage (%)
Total	22196710	100	17325097	100
rRNA	2798820	12.61	4166223	24.05
scRNA	0	0	0	0
snRNA	0	0	0	0
snoRNA	1510	0.01	3060	0.02
tRNA	303729	1.37	420136	2.43
Repbase	6268	0.03	8739	0.05
Unannotated(including miRNA)	19086383	85.99	12726939	73.46

Table 3(on next page)

Table_3_Length_distribution_of_clean_and_unique_reads

Table 3A Length distribution of clean reads

Length(nt)	GS			RS		
	Clean Reads	Clean Mapped I	Percentage (%)	Clean Reads	Clean Mapped Reads	Percentage (%)
18	234635	147547	62.88	345238	236758	68.58
19	283344	154829	54.64	356864	214528	60.11
20	392951	182653	46.48	484936	269133	55.5
21	831971	343378	41.27	1036332	511332	49.34
22	1106581	382859	34.6	1248627	561844	45
23	4863509	723877	14.88	2960096	944252	31.9
24	11262488	754673	6.7	7165283	631763	8.82
25	804876	314761	39.11	875740	434551	49.62
26	567834	319457	56.26	717802	447401	62.33
27	570544	346126	60.67	626677	382527	61.04
28	400688	245199	61.19	439194	271411	61.8
29	509188	324340	63.7	677454	425196	62.76
30	368101	209407	56.89	390854	202407	51.79
Total	22196710	4449106	20.04	17325097	5533103	31.94

Table 3B Length distribution of unique reads

Length(nt)	GS			RS		
	Unique Reads	Unique Mappe	Percentage(%)	Unique Reads	Unique Mapped Reads	Percentage(%)
18	65643	12524	19.08	60534	11737	19.39
19	96656	14745	15.26	87440	14213	16.25
20	156536	19371	12.37	136434	18278	13.4
21	316613	37836	11.95	283952	35281	12.42
22	549129	46293	8.43	429929	42149	9.8
23	2666205	98229	3.68	1400586	62164	4.44
24	5427876	183658	3.38	3725789	134719	3.62
25	375796	21617	5.75	280373	19438	6.93
26	151880	15378	10.13	120233	14752	12.27
27	108504	14511	13.37	93188	14051	15.08
28	88346	14146	16.01	78560	13547	17.24
29	81599	13600	16.67	73906	12826	17.35
30	77461	12855	16.6	68843	11906	17.29
Total	10162244	504763	4.97	6839767	405061	5.92

Table 4(on next page)

Table_4_Statistics_of_miRNAs_and_corresponding_targets

Table 4 Statistics of miRNAs and corresponding targets

Types	All_miRNA	miRNA_with_Target	Target_gene
GS	232	218	491
RS	231	217	480
Known_miRNA	198	183	249
Novel_miRNA	41	41	255
Total	239	224	493

Figure 1(on next page)

Fig.1 Read counts of represented known miRNAs

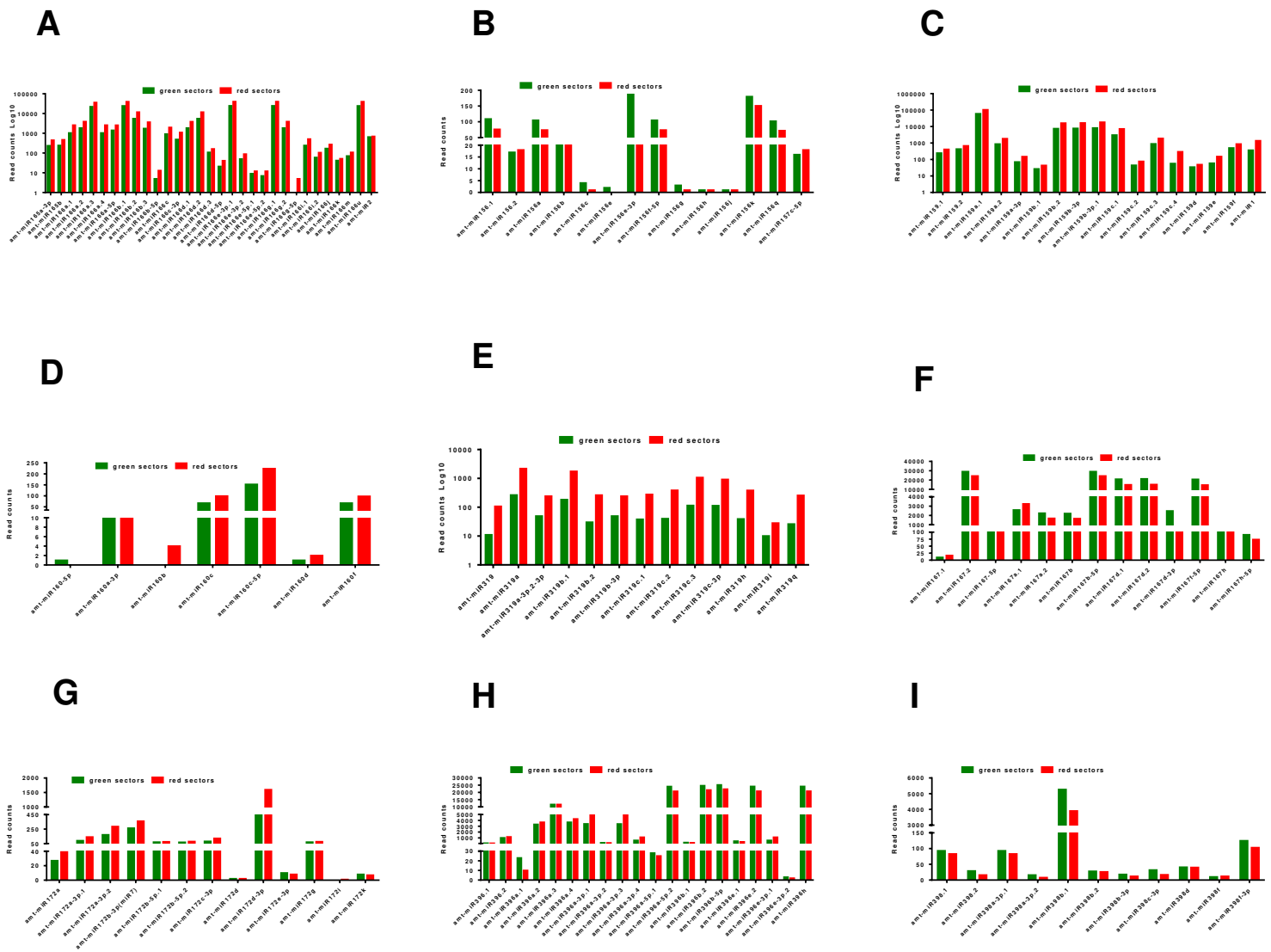

Fig.1 Read counts of represented known miRNAs

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.26723v1 | CC BY 4.0 Open Access | rec: 17 Mar 2018, publ: 17 Mar 2018

Figure 2(on next page)

Fig. 2 Relative expression of miRNAs in the red sectors and green sectors of amaranth

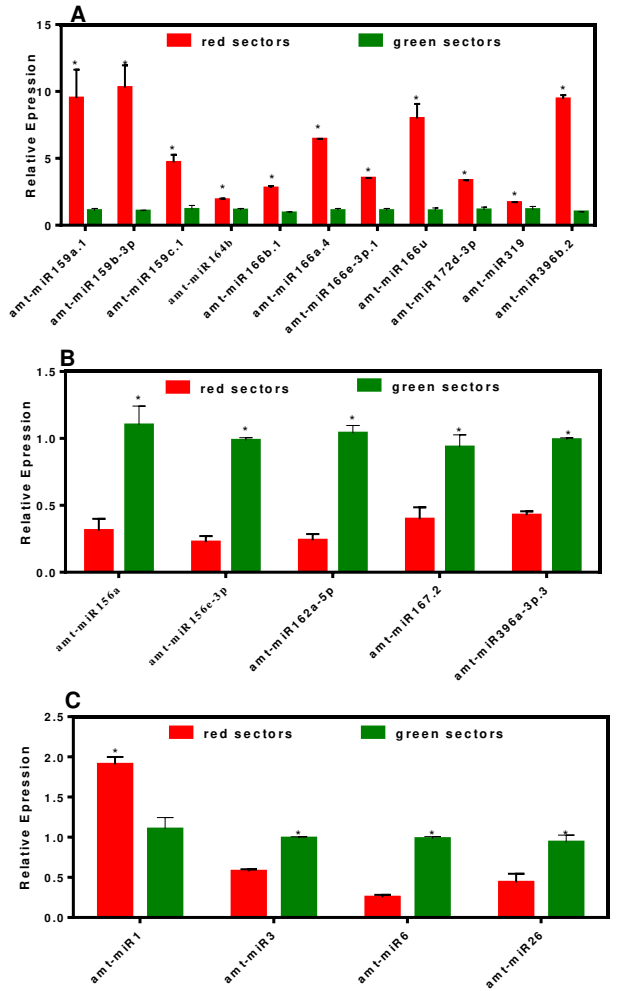

Fig. 2 Relative expression of miRNAs in the red sectors and green sectors of amaranth leaves

Figure 3(on next page)

Fig. 3 Relative expression of target genes in the red sectors and green sectors of amaranth leaves

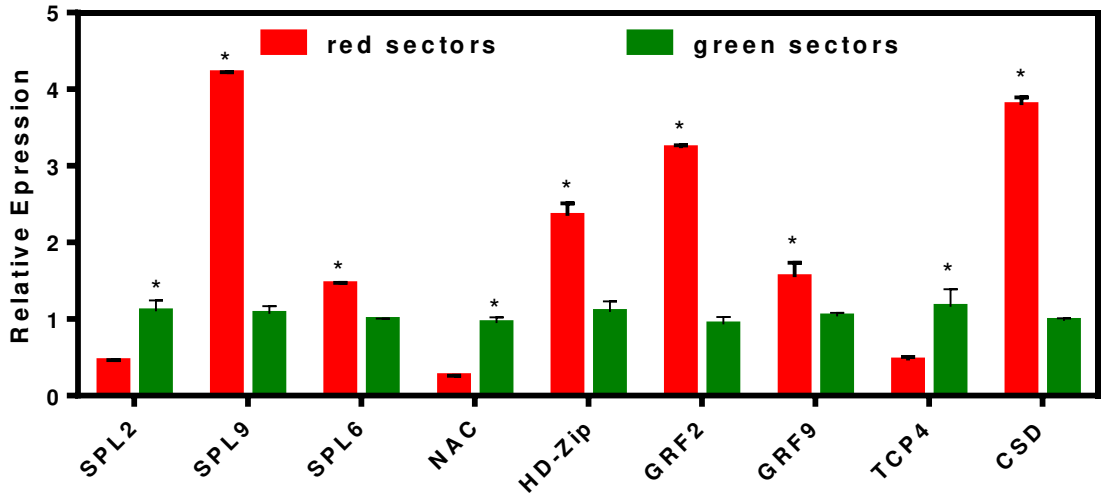

Fig. 3 Relative expression of target genes in the red sectors and green sectors of amaranth leaves

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.26723v1 | CC BY 4.0 Open Access | rec: 17 Mar 2018, publ: 17 Mar 2018

Figure 4(on next page)

Fig.4 Schematic graph showing betalains metabolic pathway networking in A. tricolor leaves

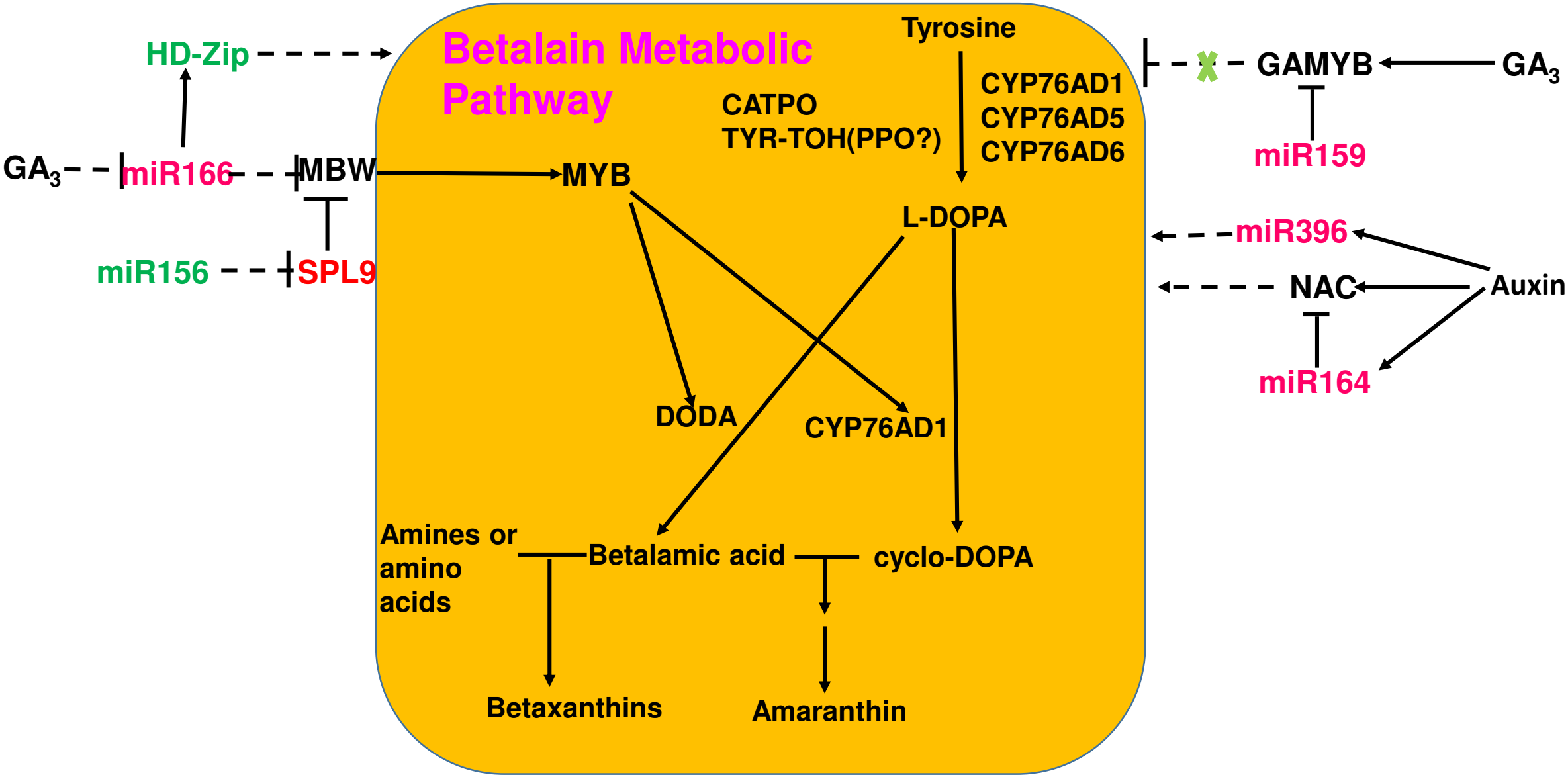

Fig.4 Schematic graph showing betalains metabolic pathway networking in *A. tricolor* leaves

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.26723v1 | CC BY 4.0 Open Access | rec: 17 Mar 2018, publ: 17 Mar 2018