植物研究 2017,37(2):249 ~ 258
Bulletin of Botanical Research

三色(*Bougainvillea peruviana* 'Thimma')在转录水平的甜菜色素和类黄酮积累比较

徐夙侠[1] 黄青云[1] 林春松[1] 黄一锦[2] 胡 欧[1]

(1. 福建省亚热带植物研究所植物生理与生化重点实验室,厦门 361006; 2. 厦门大学附属医院,厦门 361003)

摘 要 *Bougainvillea peruviana* 'Thimma'属于三角梅属,该属植物积累甜菜色素而不是像绝大多数高等植物一样积累花青素。该材料特征同株出现3种颜色:白色、洋红色和白/洋红相间。本研究首次使用3种花色特征的花序(红色Yp、混合色的Ym、白色Yw)作为研究材料进行高通量测序。并通过real-time PCR方法对探测到的花色代谢基因进行验证。共获得平均长度为616 bp的73 325条基因。3种材料的差异显示基因(DEGs)中有327个被注释到甜菜色素合成基因,308个被注释到类黄酮合成基因,466个被注释到花青素合成基因。我们选出8个基因:4个甜菜色素合成基因(PPO,CYP76AD1,cDOPA-5-GT,DODA)和4个花青素合成基因(FLS,DFR,LDOX,3-GT)进行验证。其中,4个甜菜色素合成基因在3种花色材料中的表达较好的正相关于甜菜色素含量。花青素合成途径末端的3个基因(DFR,LDOX,3-GT)在*B. peruviana* 中首次被验证。real-time PCR的验证结果很好的吻合转录组测序的结果。同时,*B. peruviana* 也提供了一个很好的三角梅属植物的生理、生化和分子生物学研究的工具,有效的摒除其他生物学干扰。

关键词 *Bougainvillea peruviana*;转录组;Real-Time PCR;甜菜色素;类黄酮
中图分类号:S68 **文献标志码**:A **doi**:10.7525/j.issn.1673-5102.2017.02.013

Comparison between Betalain and Flavonol Accumulation in Tricolor *Bougainvillea peruviana* 'Thimma' Based on Transcriptome

XU Su-Xia[1] HUANG Qing-Yun[1] LIN Chun-Song[1] HUANG Yi-Jin[2] HU Ou[1]

(1. Key Laboratory of Plant Physiology & Biochemistry, Fujian Institute of Subtropical Botany, Xiamen 361006; 2. The First Affiliated Hospital of Xiamen University, Xiamen 361003)

Abstract *Bougainvillea peruviana* 'Thimma', accumulating betalains, not anthocyanins occurring in majority of plants. The cultivar was characterized with either white, magenta, and white-magenta(variegated) bracts in the same branches. We first used magenta(Yp), bicolor(Ym) and white(Yw) open inflorescences for transcriptome analysis by sequencing. Candidate genes involving in pigment metablisms were validated by real-time PCR. We obtained 73,325 genes with average length of 616 bp. Out of DEGs, 327 candidate genes of betalain-biosynthesis, 308 flavonoid-biosynthesis genes and 466 anthocyanin- accumulation genes were detected. Eight candidate pigment-related genes were verified, four were responsible for betalain production (PPO, CYP76AD1, cDOPA-5-GT, DODA), and four responsible for flavonoids production (FLS, DFR, LDOX, 3-GT). Expression strength of four betalain production genes was well consistent with betalain-accumulation in Yp, Ym and Yw. Full later three anthocyanidin-producing genes(DFR, LDOX, 3-GT) were first verified and the expression varied in three samples. Expression of all eight candidate genes was well consistent with RNA-sequencing data. *B. peruviana* could provide a good tool to study physiological, phytochemical and molecular mechanisms in *Bougainvillea*, effectively eliminating the interference of some factors.

Key words *Bougainvillea peruviana*; transcriptome; real-time PCR; betalain; flavonol

Foundation item: The National Natural Science Foundation of China (Grant No. 31071814; 31372093); Xiamen Creation Plantform of Science&Technology(3502Z20131004)

First author introduction: XU Su-Xia(1973—), female, Ph. D., Associate Professor, Major in Plant secondary motablisms.

Received date: 2016-07-11

Bougainvillea, belonging to Nyctaginaceae family, one member in Caryophyllales, of which nine families are unique for accumulating betalains instead of anthocyanins [1]. So far, no plant has been found that produces both betalain and anthocyanin pigments [1~2]. The mutually exclusive nature of the betalain and anthocyanin production in the plant kingdom is a curious phenomenon [2~3]. *Bougainvilleas* show the specific characteristic in pigment metabolisms with rich flavonoids and betalains together. It is well known that a long common biosynthesis pathway (from CHS to DFR) was shared between anthocyanin- and flavonol-synthesis. Recently, transcriptome analysis of *B. spectabilis* was performed and genes involved in Betalains-/flavonoid-production were identified [4]. It was the first report of full genes data in *Bougainvillea*. Xu et al. [5] validated expression of betalain-/anthocyanidins-biosynthesis genes detected (CYP76AD1, DODA, cyclo-DOPA-5GT, DFR, LDOX) in four *Bougainvilleas*. The correlation between betalain-accumulation and expression of betalain-production genes was not well observed. Betalain-accumulation was complex [5~6].

Given the complexity of pigments metabolisms in bougainvilleas, choosing idea materials will help better unclose the function of these genes in *Bougainvillea*. *B. peruviana*, one of species in *Bougainvillea*, is characterized with tricolor (white, magenta or white-magenta) inflorescence in the same plant contrasting with usual monocolor inflorescence, which is intriguing in horticulture. The variegation is not fixed-pattern, random in white zones or magenta zones as major color. In our study, we used *B. peruviana* to validate the function of betalain-/anthocyanin-biosynthesis genes. Three materials of *B. peruviana* in the same plant including magenta (Yp), white-magenta (Ym) and white (Yw), are characterized with more similar genetic background. The species is a good system to explore pigment metabolisms.

We aimed at further exploring the correlation between flavonoid-production and betalains-production. The results will be help for better understanding the function of these genes and profile of flavonoid-/anthocyanins-production pathway in betalain-production plants.

1 Materials and Methods

1.1 Plant Materials and Growth Conditions

In the study, the cultivar, *B. peruviana* 'Thimma', was planted in *Bougainvilea* Garden in Fujian Institute of Subtropical Botany, China (altitude 24°26′46″, latitude 118° 04′ 04″, annual average temperature 22℃). White-, magenta- and bicolor-bracted samples of *B. peruviana* 'Thimma' were named as Yp, Ym and Yw, respectively (see Fig. 1). White-, magenta-, or bicolor inflorescences from stage Ⅱ & Ⅲ [5] were chosen for transcriptome sequencing, measurement of betalain and flavonol, and expression validation of candidate genes.

Fig. 1 Materials in the study Y. Whole plant; Yp. Magenta inflorescence; Ym. Variegated inflorescence; Yw. White inflorescence

1.2 cDNA library construction and sequencing

Total RNA of each sample was isolated using a Quick RNA isolation kit (Bioteke Corporation, Beijing, China) and then characterized on a 1% agarose gel and examined with a NanoDrop ND1000 spectrophotometer (NanoDrop Technologies, Wilmington,

DE, USA). The RIN(RNA integrity number) values (>8.0) of these samples were assessed using an Agilent 2100 Bioanalyzer(Santa Clara, CA, USA). The construction of the libraries and the RNA-Seq were performed by the Biomarker Biotechnology Corporation(Beijing, China). The purified mRNA was concentrated using Reasy RNA cleaning kit(QIAGEN, Germany) according to the manufacturer's protocol, and then broken into short fragments. RNA Fragment reagent Kit(Autolab Co., China) and Reasy RNA cleaning kit(QIAGEN, Germany) were used to fragment and retrieve the mRNA that has been condensed for 30 s. Taking these cleaved mRNA fragments as templates, a random primer(N_6) and AMV were used to synthesize the first chain, and Secondary Strand Synthesis Enzyme Mix(Autolab Co., China) was used to synthesize the second chain. The resulting cDNAs were then subjected to end-repair and phosphorylation using T4 DNA polymerase and Klenow DNA polymerase. After that, an 'A' base was inserted as an overhang at the 3′ ends of the repaired cDNA fragments and adaptors were subsequently ligated to these cDNA fragments to distinguish the different sequencing samples. To select a size range of templates for downstream enrichment, the products of the ligation reaction were purified and selected on a 2% agarose gel. Next, PCR amplification was performed to enrich the purified cDNA template. MinElute DNA Cleaning Kit extracted the synthesized products(QIAGEN, Germany) according to the protocol. Ampure beads(AgencourtAMPure PCR purification Kit, USA) was used for retrieving adapted products. The quality of retrieved cDNA was checked using the Agilent 2100 Bioanalyzer(RNA Nano Chip, Agilent). Finally, the two libraries were sequenced using an Illumina HiSeq™ 2500.

1.3 De novo assembly and unigene functional annotation

Raw reads were extracted from standard flowgram format(SFF) files using base calling and stored as fastq file. Prior to assembly, the raw reads were subjected for the following processing into clean reads, namely, the adaptor sequence was trimmed using In-house developed program, the low quality, contaminated and vector sequence was trimmed using Bcltofastq 1.8.4(http://www.illumina.com.cn/support/sequencing/downloads.aspx), and sequences shorter than 50 bp were also removed from raw reads. The remaining sequences were then assembled using Trinity (http://TrinityRNASeq.sourceforge.net Released on 2013-11-10) and recognized as unigenes(contigs and singletons) with default parameter.

For each library, short reads were first assembled into longer contigs based on their overlap regions. Then, different contigs from another transcript and their distance were further recognized by mapping clean reads back to the corresponding contigs based on their paired-end information, and thus the sequence of the transcripts was produced. Finally, the potential transcript sequences were clustered using the TGI Clustering tool to obtain uni-transcripts [7].

For the functional annotation of unigenes, the remaining sequences that could putatively encode proteins were searched against the following protein databases: NCBI non-redundant protein database (NR), Swiss-Prot protein database, Kyoto 6 Encyclopedia of Genes and Genomes(KEGG) database, Gene ontology(GO) database using BLASTX search, and nucleotide database(NT) by BLASTN with e-value less than 0.00001. Based on the NR annotation, the Blast2GO program was used to obtain GO annotation of unigenes [8]. The GO functional classification was analyzed using WEGO software.

1.4 Quantification of betalains

According to Kugler et al. [9-10], the aqueous pigment extracts of inflorescence from three blooming stages were diluted with McIlvaine buffer(pH 6.0, citrate-phosphate). The betalain content(BC) was calculated by applying the equation proposed by Cai and Corke [11]: $BC(mg \cdot L^{-1}) = [A \times DF \times MW \times 1\,000/\varepsilon \times L)]$, where, A is the absorption value at the absorption maximum, DF the dilution factor, and l the path length(1 cm) of the cuvette. For quantification of betacyanins, the molecular weight(MW) and molar extinction coefficient(ε) of betanin [MW) 550 $g \cdot mol^{-1}$; ε = 60 000 L/(mol cm) in H_2O;

λ = 538 nm; [12] were applied. Quantitative equivalents of the major betaxanthins (bx) were determined by applying the mean molar extinction coefficient {ε = 48 000 L/(mol cm) in H_2O; λ = 480 nm; [13]} and the molecular weight of Dopa-bx (dopaxanthin; MW = 390 g · mol^{-1}) for extracts from inflorescences of bougainvilleas [10]. All measurements were performed in triplicates using a UV-vis spectromer (Perkin-Elmer, Überlingen, Germany) equipped with UVWinLab V 2.85.04 software (Perkin-Elmer Instruments, Norwalk, CT).

1.5 Quantification of flavonols

Fresh bracts with early blooming flowers were collected in May 2015 and rinsed in ddH_2O. Hydrolysis of the samples followed the method as described by Hertog et al. [14]. Briefly, each sample was repeated three times. First, 0.5 g of ground freeze-dried sample in a flask was mixed with ascorbic acid solution (80 mg in 7.5 mL water). Then the sample was imbibed with the liquid, and 12.5 mL of methanol and 5 mL of 6 mol · L^{-1} HCl were added. Samples were boiled in a water bath under reflux for 2 h at 90℃. Finally, the samples were cooled at room temperature, and then filtered through a 0.22 μm RC filter (SRI, Roth, Germany) prior to injection of 10 μL to HPLC. Triplicates were performed for each sample.

All chemicals were of chromatographic grade in the study. Quercetin and kaempferol (Sigma-Aldrich, Buchs, Germany); isorhamnetin (Sigma-Aldrich, Buchs, Germany). All solution was prepared using ddH_2O (18.2 MΩ). HPLC-grade water was obtained from a Milli-Q System (Millipore, Billerica, MA, USA). All standards were dissolved into methanol to a final concentration of 1.0 mg · mL^{-1} and stored at 4℃.

All measurements were performed on a Waters Acquity™ Ultra Performance LC Spectrometry (Quattro Premier XE) equipped with an autosampler, and a Masslynx4.1 data handling system. The analytical column was BEH C_{18} (50 mm × 2.1 mm I.D., 1.7 μm, Waters, USA). An aliquot of 10 μL solution was injected for HPLC analysis. Chromatograms were acquired at 350 nm for flavonols. Formic acid-acetonitrile (0.1∶99.9, v/v/v) was selected as phase A and TFA-formic acid-acetonitrile-water (0.1∶2∶35∶62.9, v/v/v/v) as phase B [15]. Additional chromatographic conditions: 0 min, B 30%; 12 min, B 48%; 22 min, B 55%; 35 min, B 60%; 40 min, B 30%; the flow rate, 0.3 mL · min^{-1}; the injection volume, 10 μL; temperature, 35℃.

1.6 Quantification by real-time RT-PCR

Eight candidate genes, responsible for betalains, flavonols and anthocyanins were tested by Real-time PCR (Table 1). These candidate genes were polyphenol oxidase (PPO), cytochrome P 450 76AD1 (CYP76AD1), cyclo DOPA-glycosyltransferase (cyclo DOPA-5-GT), Dihydroxy-phenylalanine (DOPA)-4,5-dioxygenase (DODA), flavonol synthase (FLS), dihydroflavonol 4-reductase (DFR), Leucoanthocyanidin dioxygenase (LDOX) and anthocyanidin 3-O-glycosyltransferase (3-O-GT) in the study. Actin 1, which showed no significant change in expression throughout the inflorescence development, was used as control gene to normalize the RT-PCR results. The primers were designed using Primer Premier 5.0 with default parameters (see Table 1). Inflorescences were harvested and immediately frozen in liquid nitrogen and stored at −80℃ for RNA extraction.

First strand cDNA was synthesized from 2 μg of total RNA using PrimerScript™ RT reagent Kit (TaKaRa), according to the manufacturer's instructions. The PCR reaction was performed in 25 μL, containing 12.5 μL 26SYBR Green Master mix (TaKaRa), 300 nmol · L^{-1} each primer and 2 μL 10-fold diluted cDNA template. The PCR reactions were run in Chromo 4 (BioRad, USA) using the following program: 95℃ for 30 s and 45 cycles of 95℃ 5 s, 57℃ 30 s and 72℃ 30 s. The real time RT-PCR reaction was performed in three biological replicates, and three technical repetitions were performed for each biological replicate.

2 Results and Analysis

2.1 Betalain & Flavonol quantification

In this study, we performed betalain quantification of extracts from freeze-dried inflorescences of samples, Yp, Ym and Yw. Among three samples, content of total betalain ranged from 3 085.6 (Yp),

841.7(Ym) to 192.2 mg·kg^{-1} fresh weight(Yw)(Table 2). Betaxanthins are major betalains in Yw whereas betacyanins are major betalains in Yp(the ratio of betacyanins to betaxanthins >1). Content of betaxanthins was similar to that of betacyanins in Ym(the ratio of betacyanins to betaxanthins <1)(Table 2).

Table 1　Candidate genes chosen from redundant ESTs and primers for RT PCR analysis

Protein name	Unigene ID	Direction	Primer sequence	Product Size(bp)
PPO	CL19997Contig1	Forward Reverse	ATAACAACATCATTGCGAGA ACCAAGGAAGTCAAGGAGTA	312
CYP76AD1	CL1Contig6646	Forward Reverse	ATAATGGAAGATATTGGAAAGCC CTAGCTCTGCCATTGCCCATTC	316
DODA	Group7_Unigene_BMK.97176	Forward Reverse	CATGTAGCCATGGGTGCTG GTGAACTTGTAGGAGGCATAGG	98
cyclo-DOPA 5-GT	Group7_Unigene_BMK.94108	Forward Reverse	TTGCTCTGCTGCCAATGG GGTTTAGAATCAAGTGAGAAGC	280
FLS/F3H	CL537Contig1	Forward Reverse	AGAAGGATATGGTGGTCAA ACGGTAGTCAACAACAGA	106
DFR	CL12567Contig2	Forward Reverse	AGTCGATTATTGCAACAAAGGATT TGAATCAAGTTGCGCTGCCAT	298
LDOX	CL1Contig8676	Forward Reverse	TACCCAAAATGCCCTCAACCAG GGTGGCTCACAAAAAACTGCCC	275
Anthocyanidin 3-GT	Group8_Unigene_BMK.25557	Forward Reverse	TAACCAACACGCTAAGTCGG ATCAATGCGAGGTTTATGAC	332
Actin	Unigene29292*	Forward	GAGCATTACAGGAATTGG	252

Table 2　Betalain contents of three samples(Yp,Ym and Yw)

Materials	Betacyanins	Betaxanthins	∑-betalain	Ratio of betacyanins to betaxanthins
Yp	2056.5±1.0	1029.1±0.9	3085.6	2:1
Ym	383.8±0.0	457.9±0.1	841.7	0.84:1
Yw	82.5±0.1	109.7±0.1	192.2	0.75:1

Note:Contents(mg·kg^{-1} fresh weight is means of triplicate determinations(±standard deviation);Means of triplicate determinations with standard deviations at a constant absorbance of 1.00±0.05 at λ_{max}. The same as below.

Table 3　Flavonols contents of three samples(Yp, Ym and Yw)

Materials	Quercetin	Kaempferol	Isorhamnetin	∑-three flavonols	Ratio of betalains to flavonols
Yp	398.6	73.9	400.72	873.22	3.53:1
Ym	726.94	121.06	335.24	1183.24	0.71:1
Yw	436.16	47.68	177.92	661.76	0.29:1
Retention time	5.43/8.21	6.41/9.32	6.60/9.53	—	—

Also, contents of three flavonols (quercetin, kaempferol and Isorhamnetin) were analyzed in the study. The results showed that the highest contents of total flavonols was detected in Ym(1 183.24 mg·kg^{-1}) and the lowest was detected in Yw(661.76 mg·kg^{-1})(Table 3). In Ym and Yw, the ratio of betalains and flavonols was less than one(0.29:1 and 0.71:1) whereas the ratio was more than one(3.53:1). Of three falvonlos, quercetin is the richest flavonol detected and kaempferol is lest detected among three samples.

2.2　Overview of sequencing, assembling and functional annotation

To obtain a global view of the inflorescence transcriptome from three samples of Yp, Ym and Yw, we used the open inflorescence for sequencing. For paired-end reads with a single read length of 100 bp and GC percentages of 44.09%(Yp), 43.31(Ym) and 43.81(Yw), respectively, we obtained 19 530 221(Yp), 26 235 946(Ym) and 22 161 608(Yw) mapped reads, respectively. 127 306 genes with average length of 616 bp and N50 length 953 bp(Table 4). 3 199, 3 168 and 6 091 significantly differentially-expressed genes(DEGs) from the groups of Yp vs Yw, Yw vs Ym and Yp vs Ym, respectively. Out of them, 2 425, 2 641 and 4 490 DEGs matched in the public database, respectively(Table 5).

To identify the putative functions of the assembled unigenes in Yp, Ym and Yw, we searched for sequence similarities in five public databases: NR, Swiss-Prot, KEGG, GO database using the BLASTX search and NT by BLASTN. After Blasting, most of the DEGs(75.3% in the group of Yp vs Yw and

83.4% in the group of Yw vs Ym, 73.7% in the group of Yp vs Ym) were annotated by the NR database; the annotated genes are listed in supplementary Table 1 & Fig. 2. In three groups, Yp-specific, Yw-specific and Ym-specific genes were involved in secondary metabolisms.

Table 4 Overview of the reads from the inflorescence cDNA library of Yw, Ym and Yp

	Yp	Ym	Yw
Reads number	25 209 085	33 259 383	27 948 031
Mapped reads	19 530 221	26 235 946	22 161 608
G + C content(%)	44.09	43.31	43.81
Contig number(in total)	9 676 661		
singleton number(in total)	257 856		
ORF(in total)	126 442		
Mean length(bp)	616		
Unigene N50 length(bp)	953		
DEGs number	5 759		
COG of DEGs	870		
GO of DEGs	1 773		
KEGG of DEGs	524		
SWISS of DEGs	1 890		
nr of DEGs	2 419		

Table 5 Differentially expressed genes (DEGs) among comparison groups based on the three samples (Yp, Ym and Yw)

Comparison group	Yp vs Yw	Yw vs Ym	Yp vs Ym
DEGs	3 219	3 165	6 090
Annotated DEGs	2 425	2 640	4 489
Yp-specific genes	92	—	735
Ym-specific genes	—	190	184
Yw-specific genes	95	303	—

For the differences in the pigment metabolisms between the three samples, we compared their transcriptome data. By the GO analysis, the distribution of gene functions for the cDNA sequences among three samples were similar in three main categories: biological processes, cellular components, and molecular functions(Fig. 3). Among these categories, the cell, cell junction, binding, catalytic activity, organelle, metabolic process and cell process were the groups with the highest representation of genes from three samples. Functional classification and pathway assignments for the three samples were performed through KEGG analyses. Out of 5759 DEGs, 524 DEGs were assigned to KEGG pathways(Table 3). Out of them, 870 DEGs(15%) could be matched in the public database using COG(31%), GO(31%), Kegg(10%), Swiss analysis(33%) and Nr annotation(42%)(Table 4).

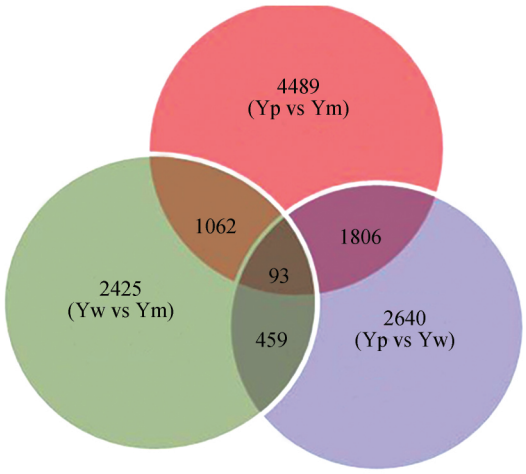

Fig. 2 Venn of DEGs among Yp, Ym and Yw

Gene Ontology(GO) was used to gain a global overview of putative gene function. Most of the identified transcripts appeared to be genes involved in biological process(57%, 58%, 59%), cellular component (8%, 12%, 11%) and molecular function (35%, 30%, 31%, in Yp vs Yw, Yw vs Ym, Yp vs Ym, respectively)(Supplementary Table 1). Biological processes involved in secondary metabolisms were further analyzed and the results were shown in Table 6. Pigments metabolites play vital roles in bougainvilleas growth, development and resistance to various biotic and abiotic stresses. Based on identified transcripts, 3106 DEGs involved in betalain biosynthesis (327, GO: 0005737, GO: 0006725, GO: 0016705, GO:0019825, GO:0008152), flavonoid biosynthetic process (308, GO: 0045431, GO: 0009812, GO: 0051555, GO: 0009963, GO: 0009813), anthocyanin accumulation process (466, GO: 0009718 GO: 0031540, GO: 0043481, GO: 0046283), glycosylation (742, GO:0006493, GO: 0006487, GO: 0006486), pigment biosynthetic process (17, GO:0046148) and pigmentation (17, GO:0043473)(supplementary Table 1 and Table 6).

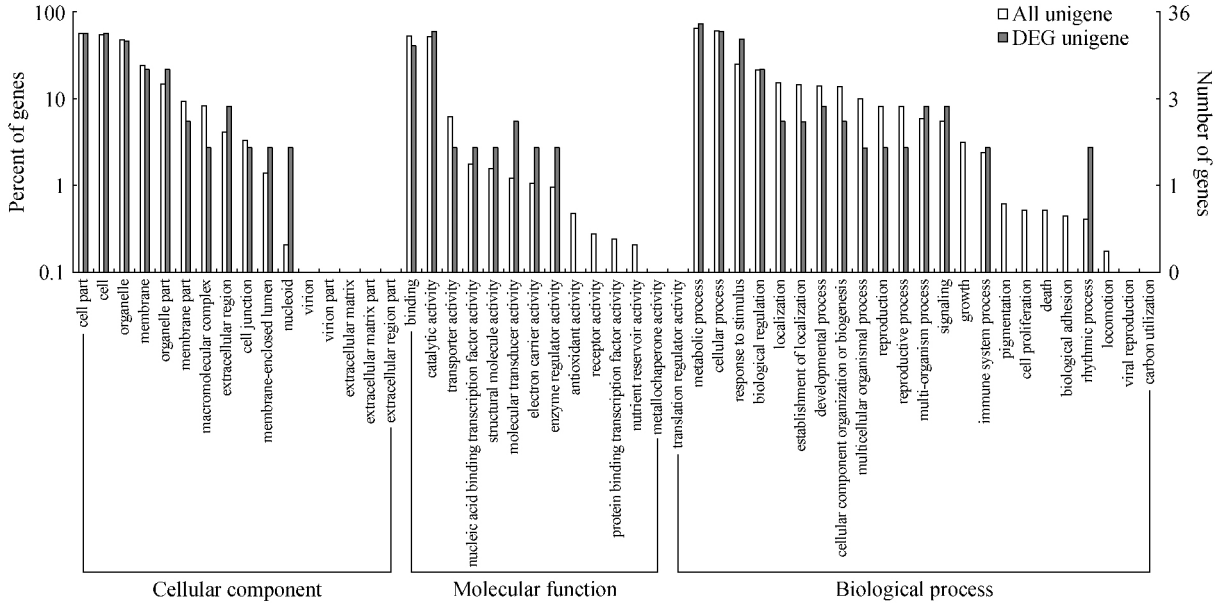

Fig. 3　Functional category distribution of transcriptional changed genes using Gene Ontology(GO) (DFR < 0.01 & fold change ≥ 2)

Table 6　Differentially expressed genes(DEG) involved in pigment metablisms pathway in inflorescences of Yp,Ym and Yw

Pathway	GO ID	DEGs
Betalain	GO:0006725,0016705,0019825,0008152	327
flavonoid biosynthetic process	GO:0045431, 0009813, 0009812,0051555,0009963	308
pigment biosynthetic process	GO:0046148	17
glycosylation	GO:0006493, 0006487, 0006486	742
anthocyanin accumulation process	GO:0031540, 0043481,0046283,0009718	466
pigmentation	GO:0043473	16

2.3　Expression analysis of candidate genes in Yp,Ym and Yw

To verify the accurateness of the RNA-Seq data, we validated the candidate genes involving in pigment biosynthesis through real time RT-PCR. Eight candidate genes were verified. Of all, four genes were responsible for betalain production(PPO, CYP76AD1, cDOPA-5-GT, DODA), four genes responsible for flavonoids production(FLS, DFR, LDOX, 3-GT). The leaves was chosen as the control standardizing other three samples.

The expression of four genes (PPO, cDOPA 5GT, DODA, CYP76AD1) was higher in abundant betalain-accumulation samples from Yp, Ym to Yw. Expression strength of four genes in Ym was between that of Yp and Yw. Similar expression strength was discovered in Yw and leaves(Fig. 4).

The performance of four genes involving in flavonoids production varied much more. Expression of FLS in Yw was higher than that in Ym, which was higher than that in Yp(Fig. 5). Expression of FLS in leaves was significantly less than that in inflorescence. Expression of DFR was higher in Ym than in Yp, which was higher than that in Yw(Fig. 5). Similar to FLS, Expression of DFR in leaves was significantly less than that in inflorescence. Expression of LDOX in Yw was higher than that in Yp, which was higher than that in Ym. Expression of LDOX in leaves was similar to that in Ym(Fig. 5). Expression of 3-GT was significantly higher in Ym than that in Yp, Yw and leaves and similar among Yp, Yw and leaves (Fig. 5).

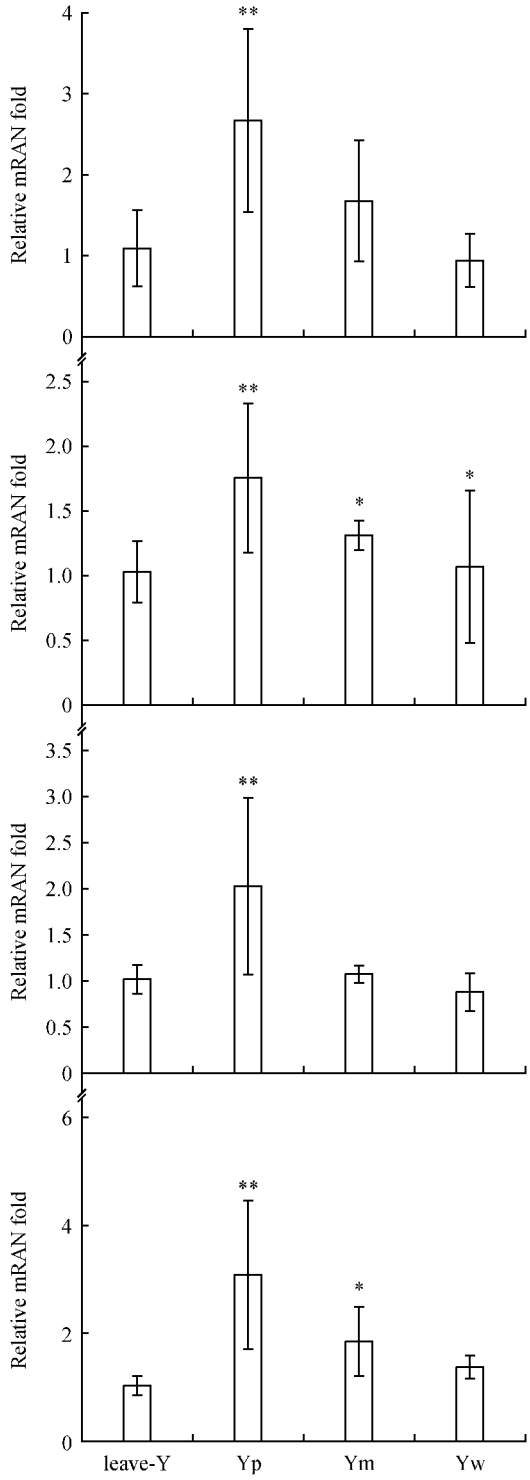

Fig. 4 Expression patterns of candidate genes involving in betalain-production in Yp, Ym and Yw A. PPO; B. CYP76AD1; C. DODA; D. cycloDOPA-5-GT

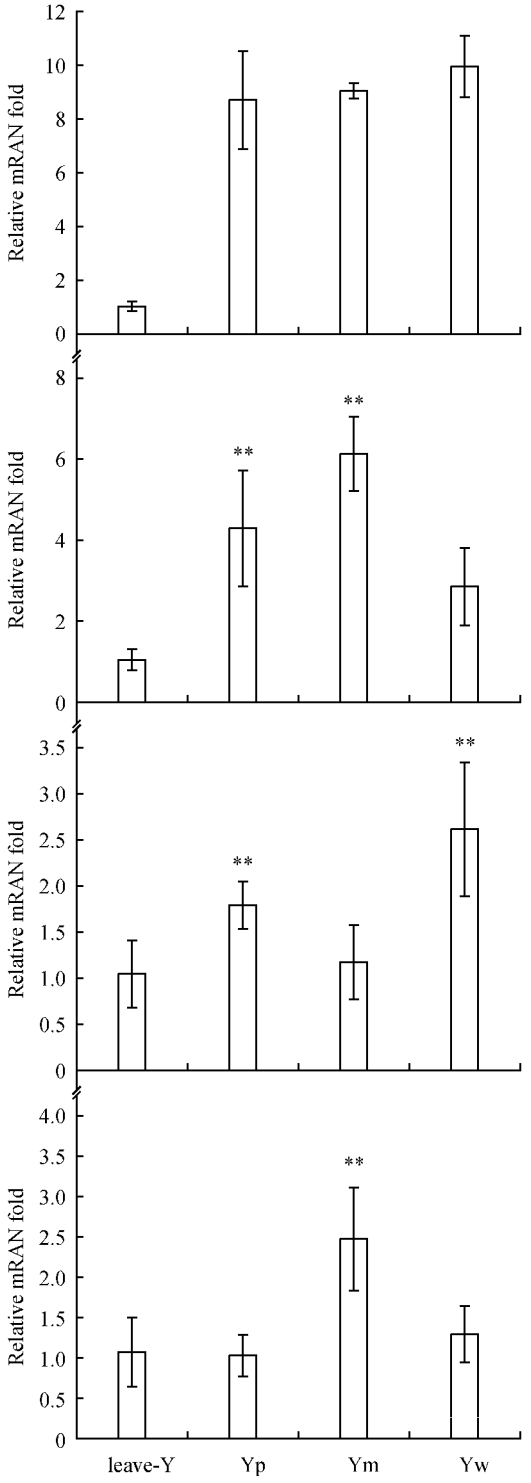

Fig. 5 Expression patterns of candidate genes involving in flavonol /anthocyanindin-production in Yp, Ym and Yw A. FLS/F3H; B. DFR; C. LDOX; D. 3-O-GT

3 Discussion

3.1 Quantification of betalains and flavonols

Based on betalains quantification, higher betalains accumulation was detected in Yp than that in Ym. Richer betalain accumulation was in Ym than in Yw. The quantification clearly suggests that the different color nuances of inflorescences from bougain-

villeas are the result of variations in betaxanthin and betacyanin contents. The results were in accordance with the perceived visual appearance. Betacyanins are major betalains in Yp(ratio of betacyanins to betaxanthins, 2:1) and betaxanthins are major betalains in Ym and Yw(ratio of betacyanins to betaxanthins, 0.75:1 and 0.84:1, respectively)(Table 2). These results were similar to other color cultivars [5].

The three kinds of flavonols are richest in Betalain-production plants [6,16]. Thus, we chose quercetin, kaempferol and isorhamnetin to evaluate contents of flavonols. The quantification results of three flavonols showed that the highest contents of total flavonols was detected in Ym and the lowest was detected in Yw(Table 3). It was observed that the lighter the color was and the smaller the ratio of betalains to flavonols was(Table 3). It seemed that the less betalain accumulated the more flavonols accumulated.

3.2 Expression profile of pigment genes

In the study, all four betalain-producing genes (PPO, CYP76AD1, cyclo-DOPA-5GT and DODA) were verified. The expression level of four genes was well consistent with trends of betalain accumulation (Yp > Ym > Yw). Expression trends of four genes were well consistent with the result of transcriptome sequencing in *B. peruviana*(supplementary table 1). Our previous study showed no well consistence between betalain accumulation and the expression level of three genes (CYP76AD1, cyclo-DOPA-5GT and DODA) in white, purple, orange and red cultivars [5]. It is possible that not only the expression of these genes, but also the supplementation of DOPA (the substrate of DOD [17]), ascorbic acid(the presence of ascorbic acid reducing agent is necessary to transform o-DOPA-quinone back to L-DOPA [18]) might be a regulatory step for betalain biosynthesis in these plants. Thus, in the study, use of *B. peruviana*, Yp, Ym and Yw, eliminate the interference of these factors and verified function of these candidate genes (PPO, CYP76AD1, cyclo-DOPA-5GT and DODA).

Four flavonoids-producing genes(FLS(flavonol-producing), DFR, LDOX, 3-GT(anthocyanin-producing)) were also verified in the study. Flavonol-producing genes(FLS) was first verified in *Bougainvillea*. The expression trends of FLS was well consistent with the result of transcriptome sequencing in *B. peruviana*(supplementary Table 1). Accumulation of flavonols was Yw > Ym > Yp and betalain accumulation was Yw < Ym < Yp in *B. peruviana*. The results suggested that the less betalain accumulation and the more flavonols accumulation. Our results was similar to the research of Tsuda et al. [19], who thought white petunia might accumulate more flavonoids than other color ones with competing substrate with anthocyanins [19]. Similar results occurred in betalain-production plants with different reason from anthocyanin-production plants because betalain synthesis doesnot share common substrate with flavonoids synthesis.

In our previous study, only two anthocyanin-producing candidate genes (DFR, LDOX) were tested and first validated in floral organs and leaves by Real-time PCR [5]. In the study, we verified full later three anthocyanidin-producing genes(DFR, LDOX, 3-GT) in *B. peruviana*. Expression of three genes varied in three samples and no obvious correlation was observed. It was speculated that inhibition of anthocyanin production is not caused by anthocyanidin 3-O-GT in *Bougainvillea*. These findings suggest a fully functioning anthocyanin pathway at least to the stage of anthocyanidin-3-O-GT in bougainvilleas.

In conclusion, although accumulation of betalains and anthocyanins was excluded, both metabolism pathways were not excluded in the betalain-production plants [2~5,20]. Also even regulating factors could be common, i. e. MYB, in regulating betalain-/anthocyanin-production [21]. Further study maybe report betalain-synthesis pathway exists in anthocyanin-production plants.

Acknowledgements We appreciated the technical help by Dr. Dahai Wang from Autolab. LTD. Co., (Beijing, China). SX Xu contributed to the design of the study, data analysis and manuscript writing. QY Huang participated in the cDNA sample preparation, the design of the study and the data analysis. CS Lin helped with RNA Extraction and the data analysis. Yijin Huang contributed to the bioinformatic analysis and helped with the RNA extraction. Ou Hu helped

with RNA extraction and discussed the results.

References

1. Mabry T J. The betacyanins, a new class of red-violet pigments and their phylogenetic significance [M]. //Leone C A. Taxonomic biochemistry, physiology, and serology. New York: Ronald Press, 1964: 239 – 254.
2. Brockington S F, Walker R H, Glover B J, et al. Complex pigment evolution in the *Caryophyllales* [J]. New Phytologist, 2011, 190(4): 854 – 864.
3. Harris N N, Javellana J, Davies K M, et al. Betalain production is possible in anthocyanin-producing plant species given the presence of DOPA-dioxygenase and L-DOPA [J]. BMC Plant Biology, 2012, 12: 34.
4. Xu S X, Huang Q Y, Lin C S, et al. Rapid generation and analysis of expressed sequence tags to uncovering inflorescence secondary metabolism of *Bougainvillea spectabilis* 'Speciosas' by pyrosequencing [J]. Euphytica, 2015, 205(3): 747 – 759.
5. Xu S X, Huang Q Y, Lin C S, et al. Transcriptome comparison reveals candidate genes responsible for the betalain-/anthocyanidin-production in bougainvilleas [J]. Functional Plant Biology, 2016, 43(3): 278 – 286.
6. Heuer S, Richter S, Metzger J W, et al. Betacyanins from bracts of *Bougainvillea glabra* [J]. Phytochemistry, 1994, 37(3): 761 – 767.
7. Pertea G, Huang X Q, Liang F, et al. TIGR gene indices clustering tools (TGICL): a software system for fast clustering of large EST datasets [J]. Bioinformatics, 2003, 19(5): 651 – 652.
8. Conesa A, Götz S. Blast2GO: a comprehensive suite for functional analysis in plant genomics [J]. International Journal of Plant Genomics, 2008, 2008: 619832.
9. Kugler F, Stintzing F C, Carle R. Identification of betalains from petioles of differently colored Swiss chard (*Beta vulgaris* L. ssp. *cicla* (L.) Alef. Cv. Bright Lights) by high-performance liquid chromatography-electrospray ionization mass spectrometry [J]. Journal of Agricultural and Food Chemistry, 2004, 52(10): 2975 – 2981.
10. Kugler F, Stintzing F C, Carle R. Characterisation of betalain patterns of differently coloured inflorescences from *Gomphrena globosa* L. and *Bougainvillea* sp. by HPLC-DAD-ESI-MS^n [J]. Analytical and Bioanalytical Chemistry, 2007, 387(2): 637 – 648.
11. Cai Y, Corke H. Amaranthus betacyanin pigments applied in model food systems [J]. Journal of Food Science, 1999, 64(5): 869 – 873.
12. Wyler H, Meuer U. Zur biogenese der betacyane: versuche mit [2-^{14}C]-dopaxanthin [J]. Helvetica Chimica Acta, 1979, 62(4): 1330 – 1339.
13. Girod P A, Zryd J P. Biogenesis of betalains: purification and partial characterization of DOPA 4,5-dioxygenase from *Amanita muscaria* [J]. Phytochemistry, 1991, 30(1): 169 – 174.
14. Hertog M G L, Hollman P C H, Venema D P. Optimization of a quantitative HPLC determination of potentially anticarcinogenic flavonoids in vegetables and fruits [J]. Journal of Agricultural and Food Chemistry, 1992, 40(9): 1591 – 1598.
15. Yang R Z, Wei X L, Gao F F, et al. Simultaneous analysis of anthocyanins and flavonols in petals of lotus (*Nelumbo*) cultivars by high-performance liquid chromatography-photodiode array detection/electrospray ionization mass spectrometry [J]. Journal of Chromatography A, 2009, 1216(1): 106 – 112.
16. Xu S X, Lin C S, Huang Q Y, et al. Analysis and comparison of flavonoids in three bougainvilleas by micellar electrokinetic chromatography [J]. Bulletin of Botanical Research, 2010, 30(6): 718 – 724.
17. Takahashi K, Takamura E, Sakuta M. Isolation and expression analysis of two DOPA dioxygenases in *Phytolacca americana* [J]. Zeitschrift fur Naturforschung. C, Journal of Biosciences, 2009, 64(7/8): 564 – 573.
18. Gandía-herrero F, García-carmona F. Biosynthesis of betalains: yellow and violet plant pigments [J]. Trends in Plant Science, 2013, 18(6): 334 – 343.
19. Tsuda S, Fukui Y, Nakamura N, et al. Flower color modification of *Petunia hybrida* commercial varieties by metabolic engineering [J]. Plant Biotechnology, 2004, 21(5): 377 – 386.
20. Shimada S, Inoue Y T, Sakuta M. Anthocyanidin synthase in non-anthocyanin-producing *Caryophyllales* species [J]. The Plant Journal, 2005, 44(6): 950 – 959.
21. Hatlestad G J, Akhavan N A, Sunnadeniya R M, et al. The beet Y locus encodes an anthocyanin MYB-like protein that activates the betalain red pigment pathway [J]. Nature Genetics, 2015, 47(1): 92 – 96.