

# 桔梗白花和紫花发育过程基因差异表达研究

齐书娅 严一字 金卓 吴松权

(延边大学农学院, 133000, 吉林延吉)

**摘要** 为挖掘影响桔梗紫花花色形成基因和途径, 利用百迈客云平台 BMKCloud 对桔梗白花和紫花 3 个时期 (花蕾期、转色期、开放期) 进行了转录组测序, 筛选桔梗花青素代谢相关差异表达基因。筛选到 5 个 KEGG 通路, 与花青素代谢相关。本试验共筛选出 11 个参与花青素生物合成的候选基因, 包括 9 个关键酶基因 (HCT、CHI、DFR、F3'5'H、ANS、3'GT、CCR、NAR、HPPR) 和 2 个糖基酶基因 ( $\beta$ -葡萄糖苷酶 BoGH3B 和 UGT79B6), 持续表达影响桔梗紫花花色形成。本研究构建了桔梗紫花形成的表达调控通路, 筛选出调控桔梗紫花形成的候选基因, 为研究桔梗紫花的遗传和分子机制研究提供参考。

**关键词** 转录组; 花青素; 桔梗; 花色; 差异表达基因

桔梗 [*Platycodon grandiflorus* (Jacq.) A.DC.] 为桔梗科桔梗属多年生草本植物, 是传统地道药材, 别名包袱花、铃铛花、僧帽花等<sup>[1]</sup>。除食用和药用价值外, 桔梗因其花开似铃铛, 有蓝、紫、白等花色, 有单瓣、重瓣等花型, 植株挺拔优美, 花朵色彩清新淡雅, 花期在 7-9 月, 也是一种很好的园林观赏植物<sup>[2-3]</sup>。然而, 桔梗花作为桔梗的副产品并不被重视, 过去多为药用, 现已开发作切花, 花药并用, 且少有病虫害, 较易管理, 很有发展前景。

基因表达分析是阐明植物生命周期中遗传、信号传导及代谢途径等复杂调控过程的重要工具<sup>[4]</sup>。目前, 转录组测序技术已成为分析生物体基因转录表达水平的普遍方法之一<sup>[5]</sup>, 且已广泛应用于忍冬<sup>[6]</sup>、木兰<sup>[7]</sup>和款冬花<sup>[8]</sup>等植物花香物质代谢通路的分子机理研究。近年来, 桔梗的研究主要集中在栽培应用、药用成分检测及功能分析等方面<sup>[9]</sup>。因此, 对桔梗花发育过程中的转录组学研究尤为重要。本研究以桔梗白花为对照, 研究不同发育时期的桔梗紫花呈色差异表达基因。采用高通量测序平台对紫花桔梗花瓣进行转录组测序分析, 比较研究紫花花瓣不同时期的基因差异表达情况, 经基因功能注释和代谢通路富集, 筛选出差异表达基因, 研究结果可为后期开展影响桔梗花色的关键基因深入挖掘和功能研究提供科学参考。

## 1 材料与方法

### 1.1 试验材料

试验材料为多年生桔梗白花 (B) 和紫花 (Z), 由延边大学作物育种实验室保存。于 2023 年 8 月晴朗天气选取无病虫害、生长旺盛的鲜花, 擦除花上灰尘和水珠后, 按花苞期 (A)、转色期 (B)、开放期 (C) 进行采样, 各设 3 次重复 (图 1)。采取的样品经液氮速冻固定, 研磨混匀, 分装入 10 mL 离心管并编号, 随后转移至 -80 °C 冰箱保存, 用于后续转录组测序及 RT-qPCR 验证试验。



图 1 不同发育时期的桔梗紫花和白花

Fig.1 Purple and white flowers of *P. grandiflorus* at different development stages

### 1.2 RNA 的提取与质检

根据 Genstar 公司提供的操作方案分离和纯化总样品的 RNA。用 NanoDrop ND-2000 对总 RNA 质检。再通过 Agilent2100/LabChip GX 检测 RNA

作者简介: 齐书娅, 研究方向为长白山药食同源植物资源与利用, E-mail: 1152258498@qq.com

吴松权为通信作者, 研究方向为长白山药食同源植物资源与利用, E-mail: arswsq@ybu.edu.cn

基金项目: 吉林省科技厅重点研发项目“高品质地道药材桔梗的生态种植技术集成”(20220204048YY)

收稿日期: 2024-02-05; 修回日期: 2024-03-05; 网络出版日期: 2024-06-19

的完整性，同时通过琼脂糖电泳的方法进行验证。

1.3 转录组文库的构建及测序

样品检测合格后，进行文库构建，主要流程如下：

（1）用带有 Oligo（dT）的磁珠富集真核生物 mRNA；（2）加入 Fragmentation Buffer 将 mRNA 进行随机打断；（3）以 mRNA 为模板，合成第一条 cDNA 链及二链，并进行 cDNA 纯化；（4）纯化的双链 cDNA 再进行末端修复、加 A 尾并连接测序接头，然后用 AMPure XP beads 进行片段大小选择；（5）最后通过 PCR 富集得到 cDNA 文库。

文库质检合格后，使用 Illumina NovaSeq6000 测序平台进行 PE150 模式测序。

1.4 分析流程

测序数据下机后，使用百迈客云平台 BMKCloud（www.biocloud.net）提供的生物信息学分析流程进行数据分析。将 Clean Data 进行序列组装，获得该物种的 Unigene 库。基于此，进行文库质量评估、结构水平分析、差异表达分析、GO、

KEGG 富集分析等。

1.5 测序样品的相关性分析

采用 Spearman 相关性分析来评估各个样品的 3 次重复之间结果的可靠性。

1.6 实时定量聚合酶链式反应（quantitative real-time polymerase chain reaction, qRT-PCR）验证

选取 11 个明显富集到苯丙酸代谢、类黄酮生物合成、苯丙氨酸代谢、花青素生物合成途径的基因，进行 qRT-PCR 分析，验证转录组数据。利用 Primer 3 plus 设计引物（表 1），参照 GenStar 公司的 2×RealStar Green Fast Mixture 试剂盒说明书进行 qRT-PCR 检测。扩增体系为 20 μL（1 μL cDNA，上、下游引物各 0.5 μL，10 μL 2×RealStar Green Fast Mixture，8.0 μL ddH<sub>2</sub>O）。反应程序：95℃ 60 s；95℃ 10s，60℃ 30 s，40 个循环。以 *Actin* 为内参基因，每个取样时间点设 3 次生物学重复，每个样品设置 3 次技术重复。使用 2<sup>-ΔΔCt</sup> 方法进行表达量的计算，用 SPSS 20.0 进行统计分析，用 Graph

表 1 qRT-PCR 引物序列  
Table 1 Primer sequences of qRT-PCR

基因 ID Gene ID	正向引物 Forward primer (5'-3')	反向引物 Reverse primer (5'-3')
<i>Actin</i>	CATCCACGAGACCACCTACAA	CGAGCCACCAATCCAAACA
TRINITY_DN23720_c0_g2	AATGCTCAATCAGTCCCTG	CTAACCGTATTCCCGTCTT
TRINITY_DN2567_c0_g1	TAGTAGCGGCAGCAGAAGC	TTCATCAACAACGGCATCA
TRINITY_DN967_c1_g1	AGTATAGGTTATGCCTGTGG	TTATGGGATTGATGCTCTT
TRINITY_DN2406_c1_g1	GTTTGCCAATTATGGACTC	AACCCAATTCGTTCTGTTA
TRINITY_DN3035_c0_g1	CCAAGAGTGAGGTCTGGCTGAG	GGGCGGTGACGGAGAAGTA
TRINITY_DN37472c0_g1	CATTGGCTCGTGGTTAGTC	AAGTTGGTGTGAGCTTTTCG
TRINITY_DN5772_c0_g1	GGAATCGGAATGGTTGAGT	TTGGCTAGGCAGTGTAAGC
TRINITY_DN6749_c0_g1	ATGTGGAGACGGAACGACG	TGACAGCCCTGGGAAAGAA
TRINITY_DN8143_c2_g1	TGCCCAATCCAATTATACCA	ACCGACGAAGTTGCTGACTA
TRINITY_DN8697_c0_g1	TGAGCCTTCAGTTTCTTCC	GGGTGCCTTTGATTGTTAG
TRINITY_DN3637_c0_g1	TTCGGCAACAAGTATGAAA	TCCAGTCCTCTCTTAGCA

Pad Prism 8 软件进行图像绘制。

2 结果与分析

2.1 转录组测序质量分析

使用百迈客云平台对桔梗白花和紫花的花苞期、转色期、开放期进行高通量转录组测序。结果见表 2，经过测序质量控制，共得到 187.22 Gb 质量控制后的数据，每个 cDNA 文库有效序列占比均在 90%以上；每组样本碱基质量值≥20（Q20）均

在 99%以上；碱基质量值≥30（Q30）均在 97%以上；GC 碱基含量在 43%~45%，说明样品测序质量较好，碱基组成平衡。以上结果表明测序数据质量和数量满足后续分析的要求。

2.2 转录组测序数据组装

如表 3 所示，N50 长度的非冗余序列为 2542，组装完整性较高，1000 bp 以下的基因数占比 49.12%，1000~2000 bp 的基因数占比 23.75%，2000 bp 以上的基因数占比 27.13%。通过 Trinity

表 2 测序数据分析汇总表  
Table 2 Summary table of sequencing data analysis

样本 ID Sample ID	读长总数 Read sum	碱基总数 Base sum	GC 碱基含量 GC base content (%)	Q20 (%)	Q30 (%)
BA1	22673594	6791955690	43.69	99.64	97.86
BA2	23522001	7046338862	44.34	99.62	97.75
BA3	20665494	6190548036	43.94	99.68	98.12
BB1	19419289	5817024270	43.96	99.64	97.86
BB2	21363691	6399577508	44.15	99.62	97.73
BB3	20202020	6051810874	44.07	99.66	98.03
BC1	21679804	6492693228	42.67	99.60	97.60
BC2	22115453	6624717198	43.30	99.63	97.80
BC3	25825530	7734409312	42.96	99.65	97.95
ZA1	23120475	6924725334	43.28	99.68	98.17
ZA2	24396786	7306725610	43.76	99.69	98.22
ZA3	24180417	7241221932	43.22	99.66	98.03
ZB1	24876074	7451447804	43.91	99.66	98.02
ZB2	21070788	6312223716	44.02	99.65	97.93
ZB3	22153286	6636874790	44.23	99.61	97.71
ZC1	22804663	6831635192	43.71	99.66	97.97
ZC2	22750568	6815684020	43.69	99.61	97.64
ZC3	25896801	7757837316	43.87	99.69	98.20

表 3 组装结果统计  
Table 3 Statistical of assembly results

所有组合转录本长度 All combination transcripts length (bp)	总数 Total number	百分比 Percentage (%)
100~300	25 306	12.54
300~500	31 668	15.69
500~1000	42 152	20.89
1000~2000	47 928	23.75
>2000	54 755	27.13
总数 Total number	201 809	
总长度 Total length	303 409 855	
N50 长度 N50 length	2452	
平均长度 Mean length	1503	

测序，共拼接得到 201 809 条转录本，平均长度为 1503 bp，上述结果表明，组装完整性与测序数据质量均较高，可充分满足后续分析要求。

2.3 转录组测序样品的相关性分析

对样本进行 Spearman 相关性分析，相关系数 ( $R^2$ ) 越接近 1，表明样品之间表达模式的相似度越高， $R^2 \geq 0.80$  说明高度相关， $0.60 < R^2 < 0.80$  说明中度相关，结果见图 2。每个样品间的  $R^2$  基本都接近 0.80，表明样品间表达模式相似度高，可用于后续试验。

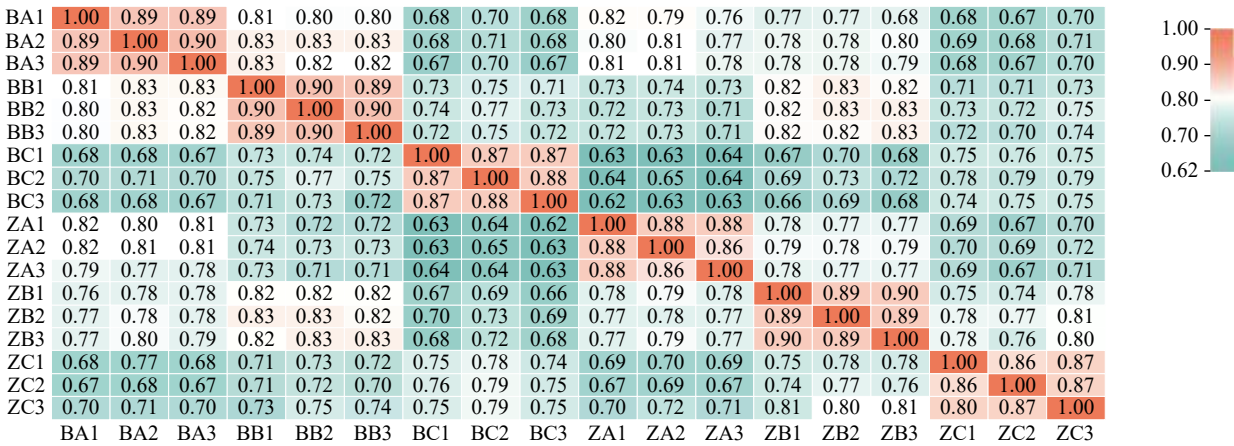


图 2 样本的相关性分析  
Fig.2 Correlation analysis of samples

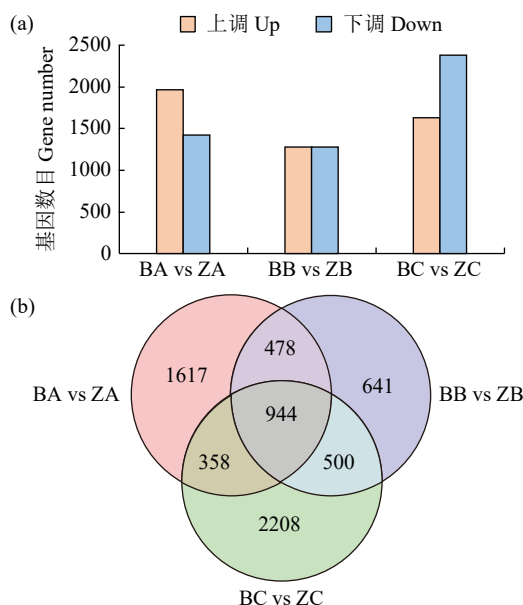
2.4 差异表达基因数目统计

基于基因在各样本中的 Count 值，使用差异分

析软件进行差异表达基因筛选，对于有生物学重复的差异分组，用 DESeq2 软件进行差异分析；对于

没有生物学重复的差异分组,使用 EBSeq 软件进行差异分析。在差异表达基因检测过程中,将差异倍数 (fold change)  $\geq 2$  且发现错误率 (FDR)  $< 0.01$  作为筛选标准。并用柱状图直观地展示整体的上下调数量,结果如图 3a 所示,BA vs ZA 中,有 3397 个差异表达基因 (DEGs) 响应差异表达,其中,上调 1970 个,下调 1427 个;BB vs ZB 中,有 2563 个 DEGs 响应差异表达,其中,上调 1282 个,下调 1281 个;BC vs ZC 中,有 4010 个 DEGs 响应差异表达,其中,上调 1630 个,下调 2380 个。

通过图 3b 可以看出,获得至少在一次比较中为差异表达基因共 9970 个,其中 944 个 DEGs 为 3 个时期共有,在花朵开放过程中持续表达,花苞期 (A) 特有 DEGs 为 1617 个、转色期 (B) 特有 DEGs 为 641 个、开放期 (C) 特有 DEGs 为 2208 个。



(a) 各比较组差异表达基因数目统计柱状图; (b) 各比较组差异表达基因韦恩图。  
(a) Histogram of the number of differentially expressed genes in each comparison group; (b) Venn diagram of differentially expressed genes for each comparison group.

图 3 差异表达基因数目统计

Fig.3 Statistics on the number of differentially expressed genes

## 2.5 差异基因表达水平火山图

利用火山图可以了解不同表达基因的总体分布,如图 4 所示,横轴表示各样品间基因的差异表达倍数,纵轴表示各基因表达量的变化统计学显著性。3 个比较组中红点代表显著上调表达的基因,蓝点代表显著下调表达的基因,由此可见,本研究选取的转录组数据包含大量具备统计学意义的显

著差异基因,且上调和下调数量差异较小。同时花朵生长发育过程中 DEGs 数量先降低再升高,在开放期 (C) DEGs 数量最多。

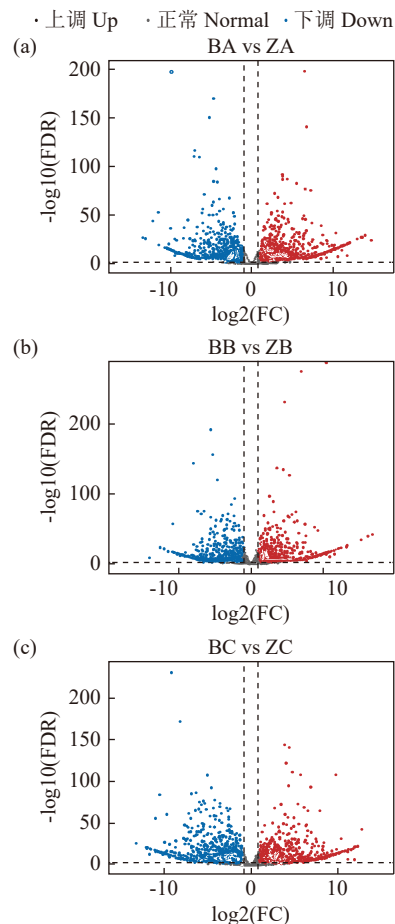


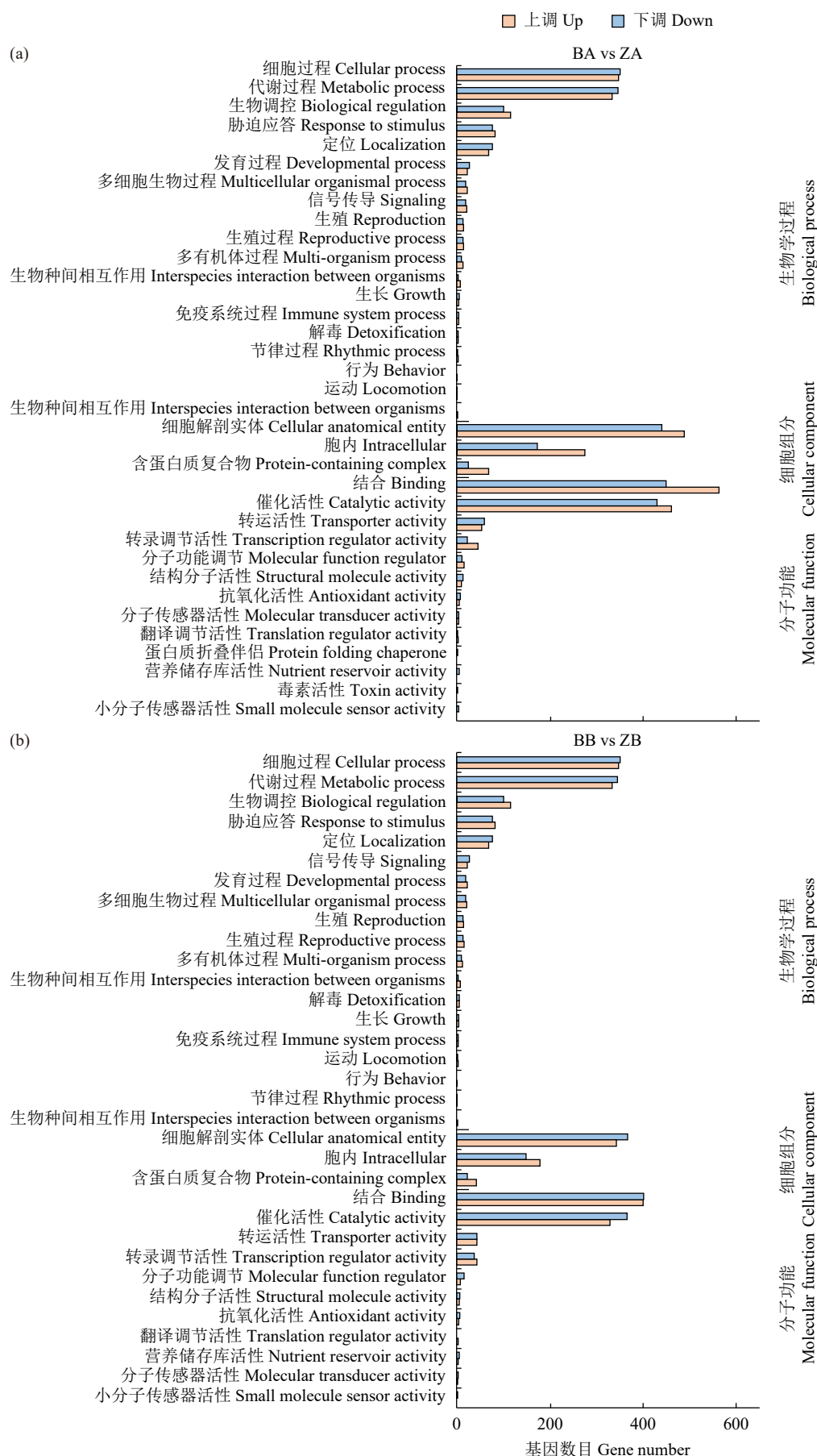
图 4 差异表达基因火山图

Fig.4 Volcano diagram of differentially expressed genes

## 2.6 DEGs 的 GO 富集分析

为了探究桔梗生长发育过程中花色形成相关差异表达基因的生物学功能,对 3 个时期鉴定到的特有 DEGs 进行 GO 功能注释。GO 数据库是 GO 组织 (gene ontology consortium) 在 2000 年构建的一个结构化的标准生物学注释系统,旨在建立基因及其产物知识的标准词汇体系,对于各物种都适用。GO 注释系统是一个有向无环图,包含 3 个主要分支,即生物学过程 (biological process)、分子功能 (molecular function) 和细胞组分 (cellular component)。

如图 5 所示,进一步对 BA vs ZA、BB vs ZB、BC vs ZC 特有 DEGs 进行 GO 富集程度分析,以  $P < 0.05$  为条件对 GO 条目进行筛选,共获得 45 个显著富集的 GO Term,其中生物学过程功能 23 个、





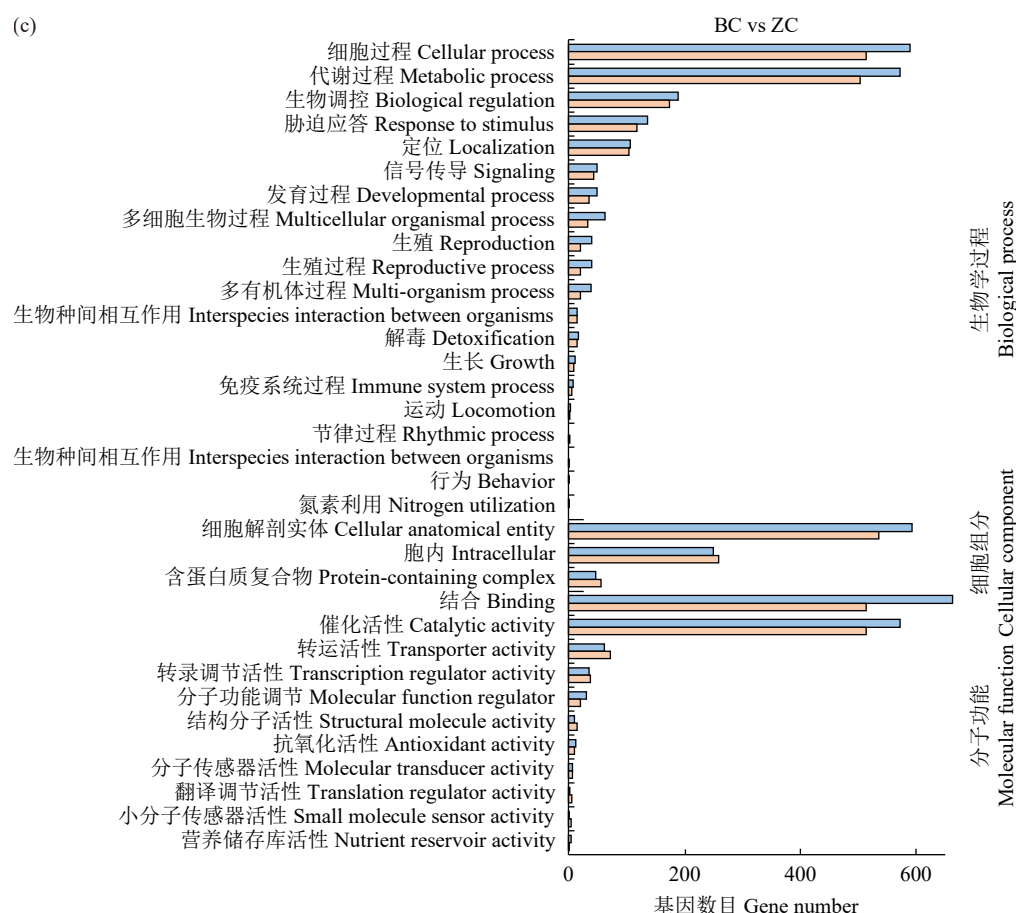


图 5 各时期 GO 富集柱状图

Fig.5 Histogram of GO enrichment in each period

细胞组分 6 个、分子功能 16 个。在生物学过程 (biological process) 中细胞过程途径的差异基因最多, 3 个时期上调基因分别为 507、347、516 个; 下调基因分别为 416、351、593 个。分子功能中结合途径的差异基因最多, 3 个时期上调基因分别为 562、399、515 个; 下调基因分别为 448、401、664 个。细胞组分中细胞途径的差异基因最多, 3 个时期上调基因分别为 489、343、536 个, 下调基因分别为 438、367、597 个。

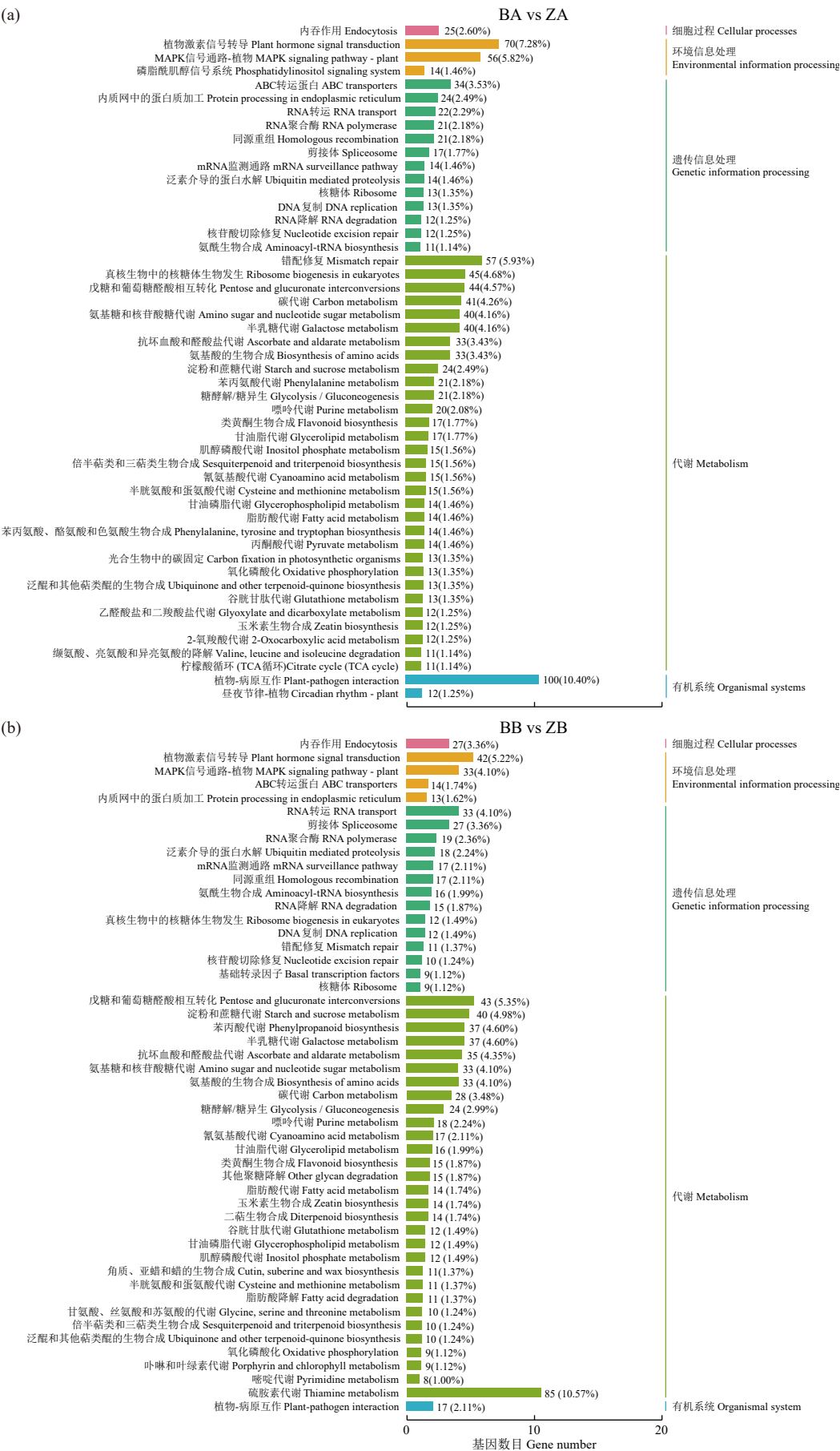
特有的 DEGs 的 GO 富集数据表明, 在细胞组分类别中, 主要富集在细胞、含蛋白质复合物、胞内等 GO 注释中; 在生物学过程类别中, 主要富集在细胞过程、代谢过程、生物调控等 GO 注释中, 表明花色形成起始于相关基因在细胞内开始催化、转运的功能, 为生物学功能提供物质基础; 在分子功能类别中, 主要富集在结合、催化活性、转运活性等 GO 注释中, 细胞代谢旺盛, 并在细胞内形成花色形成所需物质, 保证花朵呈色。在花色相关细胞的形成过程中, 花青素是植物次生代谢产物, 糖

苷衍生物。因此推测, 归类于代谢过程和催化活性的 DEGs, 可能为参与桔梗花青素生物合成差异相关的良好候选基因。

## 2.7 DEGs 的 KEGG 富集分析

如图 6 所示, 在 KEGG 数据库中 BA vs ZA 特有 DEGs 共有 804 个注释到 125 个通路上, BB vs ZB 特有 DEGs 共有 628 个注释到 120 个通路上。BC vs ZC 特有 DEGs 共有 1580 个注释到 127 个通路上。DEGs 被注释到以下 5 大类通路中, 分别是细胞过程、环境信息处理、遗传信息处理、代谢过程、有机系统。

以  $P < 0.05$  为条件对 KEGG 通路进行筛选, 特有 DEGs 显著富集的通路包括植物-病原互作 (ko04626)、植物激素信号转导 (ko04075)、戊糖和葡萄糖醛酸相互转化 (ko00040)、碳代谢 (ko01200)、淀粉和蔗糖代谢 (ko00500) 等。是植物体内最重要的基础代谢, 为类黄酮代谢途径和花青素合成提供大量能量, 同时也为桔梗花朵各生长发育时期生理生化进程做准备。



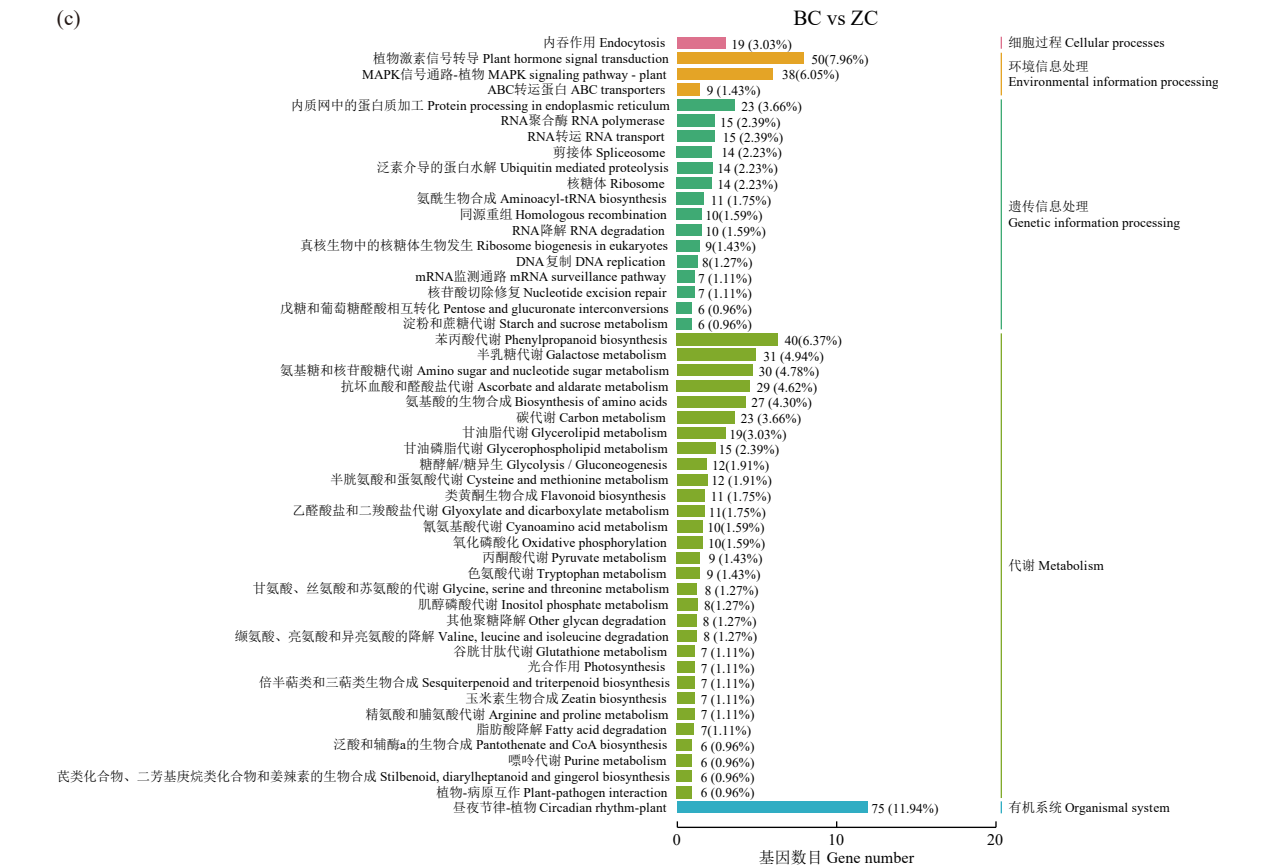


图 6 各时期 DEGs 的 KEGG 浓缩图

Fig.6 KEGG enrichment of DEGs in different periods

通过 KEGG 富集分析,筛选出各时期比较组与花色有关通路及持续表达基因,推测以下基因与桔梗紫花花色形成有关(表 4)。

如图 7,通过筛选出的各时期关键基因的表达量,可以看出 TRINITY\_DN5772\_c0\_g1 (*F3'5'H* 基因)在桔梗紫花转色期表达量最高,为紫花呈色提

表 4 各时期比较组花色关键结构基因差异表达

Table 4 Differential expression of key structural genes of flower color in the comparison groups in different periods					
代谢通路 Pathway ID	通路名称 Name of pathway	基因 ID Gene ID	基因名称 Gene name	详细描述 Detailed description	表达 Expression
ko00940	苯丙酸代谢	TRINITY_DN23720_c0_g2	HCT	羟基肉桂酰辅酶 A 莽草酸酯	上调
		TRINITY_DN2567_c0_g1	CCR	肉桂酰辅酶 A 还原酶	上调
ko00941	类黄酮生物合成	TRINITY_DN967_c1_g1	β-葡萄糖苷酶 BoGH3B	β-葡萄糖苷酶 BoGH3B	上调
		TRINITY_DN2406_c1_g1	CHI	查尔酮-黄酮异构酶	上调
		TRINITY_DN3035_c0_g1	NAR	柚皮素, 2-氧代戊二酸 3-双加氧酶	上调
		TRINITY_DN37472_c0_g1	DFR	二氢黄酮醇 4-还原酶	上调
		TRINITY_DN5772_c0_g1	F3'5'H	类黄酮-3',5'-羟化酶	上调
		TRINITY_DN6749_c0_g1	ANS	花青素合酶	上调
ko00360	苯丙氨酸代谢	TRINITY_DN8143_c2_g1	HPPR	羟苯基丙酮酸还原酶	下调
ko00942	花青素生物合成	TRINITY_DN8697_c0_g1	3'GT	花青素 3-O-葡萄糖基转移酶	上调
		TRINITY_DN3637_c0_g1	UGT79B6	UDP-糖基转移酶 79B6	下调

供大量前体物质。

2.8 qRT-PCR 结果分析

为了确认测序分析中获得的转录物,进一步验证转录组测序数据的准确性,选取了上述总结的 11 个差异基因进行 qRT-PCR 验证(图 8)。将 11 组数据的实时荧光定量鉴定结果与转录组数据(RNA-seq)进行比较,结果表明,这些基因的表达趋势与转录组测序结果基本一致,说明 qRT-PCR 数据和转录组测序数据之间具有良好的一致性,本研究中用于花青素合成相关基因分析的转录组数



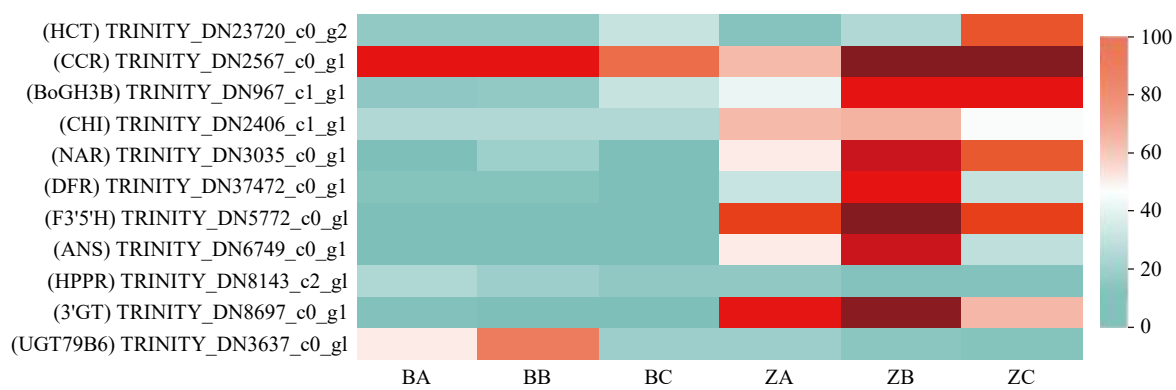
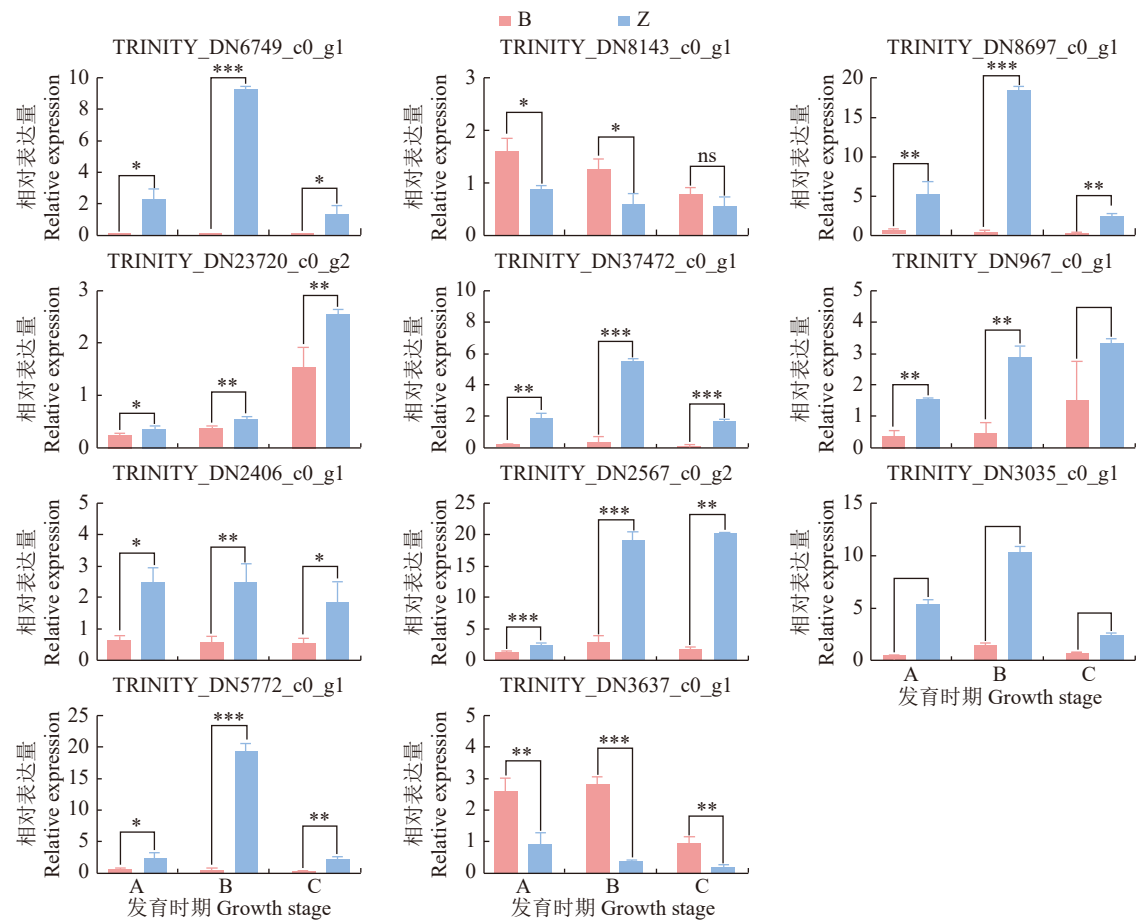


图 7 关键基因表达量热图  
Fig.7 Heat map of key gene expression



“\*”、“\*\*”和“\*\*\*” 分别代表同一比较组中白花和紫花在  $P < 0.05$ 、 $P < 0.01$  和  $P < 0.001$  水平上差异显著。  
“\*”, “\*\*” and “\*\*\*” indicate that white flowers and purple flowers in the same comparison group were significantly different at  $P < 0.05$ ,  $P < 0.01$ , and  $P < 0.001$  levels, respectively.

图 8 差异基因实时荧光定量结果  
Fig.8 qRT-PCR results of differential genes

据是可靠的。

### 3 讨论

花色是观赏植物最重要的观赏与经济性状之一<sup>[10]</sup>，了解决定植物花色形成的分子基础及其遗传调控过程是加速观赏植物花色育种的重要前提<sup>[11]</sup>。

桔梗作为药、食、赏兼用的经济植物，目前对桔梗的研究主要集中在药用和食用方面，针对桔梗花色研究和提高观赏效果文章鲜有报道，因而了解桔梗花色形成分子机理显得尤为重要。

#### 3.1 花青素生物合成途径

花青素生物合成途径是一般类黄酮代谢途径

的延伸<sup>[12]</sup>，第一阶段是苯丙氨酸在苯丙氨酸裂解酶（phenylalanine ammonia-lyase, PAL）、肉桂酸羟化酶（cinnamate 4-hydroxylase, C4H）、香豆酰辅酶 A 连接酶（4-coumarate CoA ligase, 4CL）的作用下生成香豆酰辅酶 A（4-coumarate, CoA）<sup>[13]</sup>。

第二阶段由查尔酮合成酶（chalcone synthase, CHS）催化丙二酰辅酶 A 和香豆酰辅酶 A 生成查尔酮，随后，查尔酮异构酶（chalcone isomerase, CHI）将查尔酮异构成三羟基黄烷酮，黄烷酮-3-羟化酶（flavanone-3-hydroxylase, F3H）将三羟基黄烷酮进一步羟基化形成二氢黄酮醇。二氢黄酮醇在不同酶的催化作用下生成不同的物质，其中一部分通过类黄酮-3'-羟化酶（flavonoid-3'-hydroxylase, F3'H）将 B 环 3'位羟基化生成矢车菊素的前体物质二氢槲皮素，还有一部分通过类黄酮-3',5'-羟化酶（flavonoid-3',5'-hydroxylase, F3'5'H）将 B 环 3',5'位置羟基化，生成飞燕草素的前体物质二氢杨梅黄酮<sup>[14-15]</sup>。

第三阶段二氢黄酮醇、二氢槲皮素和二氢杨梅黄酮在二氢黄酮醇还原酶（dihydroflavonol-4-reductase, DFR）和花青素合成酶（anthocyanidin synthase, ANS）作用下形成有色花青素，在植物细胞中花青素通常不能以单体的形式存在，因此他们往往会与糖苷结合转化成稳定的花青素苷<sup>[13]</sup>。

本试验也探究出与花色形成有关的通路，如苯丙酸代谢、类黄酮生物合成、苯丙氨酸、酪氨酸和色氨酸的生物合成、苯丙氨酸代谢及花青素生物合成等。还有甜菜色素生物合成、类胡萝卜素生物合成、异黄酮生物合成等通路被发现，但只有少量差异表达基因富集。

### 3.2 花青素生物合成关键酶

类黄酮类、类胡萝卜素类和生物碱类是自然界中主要的三大色素类群，类黄酮物质中的花青素苷是绝大多数观赏植物重要的呈色物质之一<sup>[16]</sup>。而 PAL、CHS、CHI、DFR、F3H 和 4CL 均是植物花青素生物合成路径中的关键酶，在花青素生物合成方面发挥重要作用，影响着植物花色的变化<sup>[17]</sup>。本研究发现 CHI、DFR、F3'5'H、ANS 等关键酶基因影响桔梗紫花呈色。王冲等<sup>[18]</sup>在解析大花君子兰花瓣花青素代谢差异中发现，主要富集与花青素合成相关的通路，包括苯丙素生物合成途径、类黄酮生物合成途径、异黄酮生物合成途径、花青素生物合

成途径和苯丙氨酸代谢生物合成途径。陈家龙等<sup>[19]</sup>研究表明，在花青苷生物合成途径中，与白花相比，粉花中 C4H、CHS、CHI 和 DFR 的表达量显著增加，F3H 和 LDOX 的表达量显著降低，而 HCT 和 F3'H 的同源基因有的表达上调，有的下调。这一结果表明 C4H、CHS、CHI 以及 DFR 可能是导致粉花品种花色呈色的关键基因，而 DFR 基因专门用于合成花青苷<sup>[20]</sup>。

### 3.3 蓝紫色花形成的关键酶与代谢产物

F3'5'H 均属于细胞色素 P450 家族的单加氧酶，在多种植物中都被报道是调控蓝色花形成的主要结构基因<sup>[21]</sup>。F3'5'H 是合成飞燕草素的主效基因，将其导入缺乏此 2 种基因的植物中，可长出非传统的蓝紫色花卉<sup>[22]</sup>。Huai 等<sup>[23]</sup>将三色堇 F3'5'H 基因导入月季中，会使翠雀素积累，从而得到了蓝紫色的月季。本试验也发现在类黄酮生物合成中 F3'5'H 基因在紫花转色期中显著表达。

糖基转移酶是生成花色苷的关键酶<sup>[24]</sup>，Fukuchi-Mizutani 等<sup>[25]</sup>和 Nakatsuka 等<sup>[26]</sup>在三花龙胆呈色机理研究中分别从花瓣中纯化到花青素 3'-O-葡萄糖基转移酶 3'GT 和花青素 5-O-葡萄糖基转移酶（Gt5GT7），并通过底物特异性分析和转基因功能验证，证实 Gt5GT7 和 3'GT 能相继催化合成飞燕草素 3,5,3'-O-三葡萄糖苷，为最终产物深蓝色龙胆飞燕草素的合成提供底物。本试验发现在花青素生物合成通路中，3'GT 在紫花转色期表达量显著高于白花。

飞燕草素及其衍生物的合成和积累是绝大多数蓝紫色花瓣着色的基础<sup>[27]</sup>。前人<sup>[28-29]</sup>通过多基因联合修饰工程，降低花青素和天竺葵素的比例，增加飞燕草素的积累，在玫瑰和菊花中培育出了蓝色的花朵将桔梗 *PgF3'5'H* 基因通过抑瘤拟南杆菌转化到烟草中。*PgF3'5'H* 在转基因植株中的表达水平相对高于对照。异源表达使转基因烟草花产生新的以飞燕草苷为基础的花青素<sup>[30]</sup>。

### 3.4 内参基因选择

实时荧光定量 PCR（qRT-PCR）是研究基因表达的一项重要技术，通过对内参基因的检测从而达到对目的基因表达情况的实时监测<sup>[31]</sup>。理想的内参基因在植物不同生长发育阶段和不同组织中均能稳定表达，不受环境影响<sup>[32]</sup>。作为传统内参基因，*Actin* 基因已被广泛应用。*Actin* 基因不论

在核苷酸还是氨基酸水平上都具有高度的保守性和同源性,在各种组织中表达相对稳定,因而常被作为研究其他基因的表达模式及调控机制的分子内标,在半定量和定量 PCR 研究中用作内参,以确定目标基因的相对表达量<sup>[33]</sup>。

作为花瓣在不同发育时期的内参基因选择上,*Actin* 基因已在荷花不同花色品种及不同花发育期间表达稳定<sup>[34]</sup>;在芍药花瓣不同发育时期和不同组织中,*Actin* 均是表达比较稳定的内参基因<sup>[35]</sup>;Vekemans 等<sup>[36]</sup>比较珙桐的苞片与花中一些基因的表达量时,选用 *Actin* 作为内参基因,定量结果准确地反映出了目标基因表达量在两者中的显著变化;齐香玉等<sup>[37]</sup>在茉莉花不同发育阶段的内参基因选择上也验证了 *Actin* 基因为最优。

## 4 结论

综上,通过比较分析并挖掘桔梗紫花的差异基因表达,为后续研究桔梗紫花花色特异基因表达调控及分析相关基因功能提供一定的参考价值。通过筛选紫花特异且高表达的花青素生物合成基因。在桔梗白花与紫花对比中,KEGG 分析结果表明,苯丙酸代谢、类黄酮生物合成、花青素生物合成途径中差异基因高度表达与桔梗紫花花色形成有关;共筛选出 9 条关键酶基因(HCT、CHI、DFR、F3'5'H、ANS、3'GT、CCR、NAR、HPPR)以及 2 个糖基酶基因( $\beta$ -葡萄糖苷酶 BoGH3B 和 UGT79B6),持续表达影响桔梗紫花花色形成,为桔梗花青素生物合成途径关键酶基因进一步研究提供了依据,也为将来花青素生物合成途径的解析及调控机制的研究奠定了基础。

## 参考文献

- [1] 中国科学院中国植物志编辑委员会. 中国植物志. 北京: 科学出版社, 2016.
- [2] Wu S J, Meng Y. Comprehensive development and utilization of wild *Platycodon*. Northern Horticulture, 2010, 24: 219-221.
- [3] Yang F F, Guo X, Hu B X, et al. Reserach progress in the use diversity and directive breeding of *Platycodon grandiflorum*. Progress in Modern Biomedicine, 2015, 15(14): 2751-2756.
- [4] Chen C B, Wu J Y, Hua Q Z, et al. Identification of reliable reference genes for quantitative real-time PCR normalization in pitaya. Plant Methods, 2019, 15(1): 1-12.
- [5] Chen L, Li L N, Dai Y P, et al. *De novo* transcriptome analysis of *Osmanthus serrulatus* Rehd. flowers and leaves by Illumina sequencing. Biochemical Systematics and Ecology, 2015, 61: 531-540.
- [6] Xia Y, Chen W W, Xiang W B, et al. Integrated metabolic profiling and transcriptome analysis of pigment accumulation in *Lonicera japonica* flower petal during colour-transition. BMC Plant Biology, 2021, 21: 98.
- [7] Dhandapani S, Jin J J, Sridhar V, et al. Integrated metabolome and transcriptome analysis of *Magnolia champaca* identifies biosynthetic pathways for floral volatile organic compounds. BMC Genomics, 2017, 18(1): 463.
- [8] 贾岩, 张福生, 肖淑贤, 等. 款冬花不同发育阶段的代谢组学 and 比较转录组学分析. 中国生物化学与分子生物学报, 2017, 33(6): 615-623.
- [9] 李国清, 毕研文, 陈宝芳, 等. 中草药桔梗人工栽培研究进展. 农学学报, 2016, 6(7): 55-59.
- [10] 戴思兰, 洪艳. 基于花青素苷合成和呈色机理的观赏植物花色改良分子育种. 中国农业科学, 2016, 49(3): 529-542.
- [11] 李淑娟, 侯静, 冒燕, 等. 决定植物花色的分子机制与遗传调控. 西南林业大学学报, 2012, 32(6): 92-97.
- [12] 王江昱, 田淑婷, 张涵, 等. 单子叶观赏植物花青素苷呈色机制研究进展. 扬州大学学报(农业与生命科学版), 2022, 43(5): 101-114.
- [13] Hichri I, Barrieu F, Bogs J, et al. Recent advances in the transcriptional regulation of the flavonoid biosynthetic pathway. Journal of Experimental Botany, 2011, 62(8): 465-2483.
- [14] Timothy A, Edwina C. Genetics and biochemistry of anthocyanin biosynthesis. The Plant Cell, 1995, 7(7): 1071-1083.
- [15] 刘淑华, 臧丹丹, 孙燕, 等. 青素生物合成途径及关键酶研究进展. 土壤与作物, 2022, 11(3): 336-346.
- [16] Saigo T, Wang T, Watanabe M, et al. Diversity of anthocyanin and proanthocyanin biosynthesis in land plants. Current Opinion in Plant Biology, 2020, 55: 93-99.
- [17] 赵历强, 单春苗, 张声祥, 等. 基于转录组测序的细风轮花青素合成途径及关键酶基因分析. 植物研究, 2020, 40(6): 886-896.
- [18] 王冲, 宋阳. 基于转录组数据解析大花君子兰花瓣花青素代谢差异. 分子植物育种, 2023, 21(4): 1093-1102.
- [19] 陈家龙, 侯蓉苗, 朱建军, 等. 木槿两个花色品种的花瓣转录组测序分析. 分子植物育种, 2022, 20(8): 2507-2516.
- [20] Wang W J, Liu G S, Niu H X, et al. The F-box protein COI1 functions up stream of MYB305 to regulate primary carbohydrate metabolism in tobacco (*Nicotiana tabacum* L. cv. TN90). Journal of Experimental Botany, 2014, 65(8): 2147-2160.
- [21] 韩科厅, 胡可, 戴思兰. 观赏植物花色的分子设计. 分子植物育种, 2008, 6(1): 16-24.
- [22] 彭玉帅, 王如峰, 张陆军. 花青素生物合成的关键酶及其调控因子. 中草药, 2014, 5(1): 131-136.
- [23] Huai G L, Qi P, Yang H J, et al. Characteristics of  $\alpha$ -Gal epitope, anti-Gal antibody,  $\alpha$ -1, 3 galactosyltransferase and its clinical exploitation (Review). International Journal of Molecular Medicine, 2016, 37(1): 11-20.
- [24] 周琳, 邹红竹, 韩璐璐, 等. 糖基转移酶在花瓣色泽形成中的作用研究进展. 园艺学报, 2022, 49(3): 687-700.
- [25] Fukuchi-Mizutani M, Okuhara H, Fukui Y, et al. Biochemical and molecular characterization of a novel UDP-glucose: anthocyanin 3'-O-glucosyltransferase, a key enzyme for blue anthocyanin biosynthesis from gentian. Plant Physiology, 2003, 32: 1652-1663.
- [26] Nakatsuka T, Sato K, Takahashi H, et al. Cloning and characterization of the UDP-glucose: anthocyanin 5-O-glucosyltransferase gene from blue-flowered gentian. Journal of Experimental Botany, 2008, 59(6): 1241-1252.

- [27] Tanaka Y, Brugliera F. Flower colour and cytochromes P450. Philosophical Transactions of the Royal Society B, 2013, 368 (1612): 20120432.
- [28] Noda N, Yoshioka S, Kishimoto S, et al. Generation of blue chrysanthemums by anthocyanin B-ring hydroxylation and glucosylation and its coloration mechanism. Science Advances, 2017, 3: e1602785.
- [29] Noda N. Recent advances in the research and development of blue flowers. Breeding Science, 2018, 68(1): 79-87.
- [30] Lulin Ma, Wenjie Jia, Qing Duan, et al. Heterologous expression of *Platycodon grandiflorus* PgF3'5'H modifies flower color pigmentation in tobacco. Genes, 2023, 14(10): 1920.
- [31] Giulietti A, Overbergh L, Valckx D, et al. An overview of Real-time quantitative PCR: applications to quantify cytokine gene expression. Methods, 2001, 25(4): 386-401.
- [32] Czechowski T, Stitt M, Altmann T, et al. Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis*. Plant Physiology, 2005, 139(1): 5-17.
- [33] Thellin O, Zorzi W, Lakaye B, et al. Housekeeping genes as internal standards: use and limits. Journal of Biotechnology, 1999, 75: 291-295.
- [34] 王彦杰, 陈叶清, 薛泽云, 等. 荷花花瓣着色过程实时荧光定量 PCR 内参基因的筛选及验证. 南京农业大学学报, 2017, 40(3): 408-415.
- [35] 李健. 芍药实时定量 PCR 内参基因的筛选和验证. 分子植物育种, 2017, 15(7): 2544-2549.
- [36] Vekemans D, Viaene T, Caris P, et al. Transference of function shapes organ identity in the dove tree inflorescence. New Phytologist, 2012, 193(1): 216-228.
- [37] 齐香玉, 陈双双, 冯景, 等. 茉莉花实时荧光定量 PCR 内参基因的筛选与验证. 华北农学报, 2020, 35(6): 22-30.

## Study on Differential Gene Expression during the Development of White and Purple Flowers of *Platycodon grandiflorus*

Qi Shuya, Yan Yizi, Jin Zhuo, Wu Songquan

(Agricultural College, Yanbian University, Yanji 133000, Jilin, China)

**Abstract** To explore the genes and pathways influencing the color formation of *Platycodon grandiflorus* purple flowers, the BMKCloud platform from Biomarker Technologies was used to conduct transcriptome sequencing on both white and purple *P. grandiflorus* flowers at three stages (bud, color changing, and blooming). The differentially expressed genes related to anthocyanin metabolism in *P. grandiflorus* were screened. Five KEGG pathways related to anthocyanin metabolism were screened out. A total of 11 candidate genes involved in anthocyanin biosynthesis were screened out, including nine key enzyme genes (HCT, CHI, DFR, F3'5'H, ANS, 3'GT, CCR, NAR, HPPR) and two glycosylase genes ( $\beta$ -glucosidase BoGH3B and UGT79B6), and their continuous expression affects the formation of purple flowers of *P. grandiflorus*. This research constructed an expression regulation pathway for the formation of *P. grandiflorus* purple flowers and identified the candidate genes involved in this process, thus providing new insights for the genetic and molecular mechanism research of *P. grandiflorus* purple flowers.

**Key words** Transcriptome; Anthocyanins; *Platycodon grandiflorus*; Flower color; Differently expressed genes