

# 基于转录组测序的紫芽茶树花青素合成 相关基因分析

蒋会兵<sup>1</sup>, 夏丽飞<sup>1</sup>, 田易萍<sup>1</sup>, 戴伟东<sup>2</sup>, 孙云南<sup>1</sup>, 陈林波<sup>1</sup>

(<sup>1</sup> 云南省农业科学院茶叶研究所/云南省茶树种质资源创新与配套栽培技术工程研究中心/云南省茶学重点实验室, 勐海 666201;

<sup>2</sup> 中国农业科学院茶叶研究所, 杭州 310008)

**摘要:**为筛选调控茶树芽叶紫色性状相关候选基因,以紫芽茶树品种紫娟、绿芽茶树品种云抗10号和福鼎大白茶为材料,利用液质联用法(HPLC-MS)、转录组测序(RNA-Seq)和实时荧光定量PCR技术(qRT-PCR),分析了紫芽和绿芽茶树品种芽叶花青素、儿茶素含量及合成相关基因的表达水平。结果表明,紫娟、云抗10号和福鼎大白茶芽叶花青素含量分别为5.05 mg/g、0.08 mg/g和0.15 mg/g,紫芽与绿芽茶树间花青素含量差异显著。3个茶树品种儿茶素含量分别为15.12 mg/g、19.52 mg/g和15.09 mg/g,差异不显著。转录组测序共获得255079条转录本序列(Transcripts),组装出166118条单基因簇(Unigenes)。所得单基因簇注释为GO分类中54446条,KOG分类中30666条,KEGG分类中20336条。DEG-Seq分析得到紫芽与绿芽茶树品种间共有434条差异表达基因,涉及到38个代谢通路。进一步筛选到类黄酮生物合成途径相关基因112条,有15条基因在紫芽与绿芽茶树品种间差异表达,其中PAL(Cluster-1850.70337)、CHS(Cluster-21971.0)、ANS(Cluster-1850.80848)和UGT(Cluster-1850.77163)基因在紫芽茶树品种中上调表达,在绿芽茶树品种间表达差异不显著,表达趋势与花青素积累情况相似。FLS(Cluster-1850.81068)基因在绿芽茶树品种中均为上调表达,可能与儿茶素合成相关。qRT-PCR随机检测9个差异表达基因,表明不同茶树品种间差异表达基因的相对表达水平与转录组测序结果基本一致。因此,从基因表达模式推测,PAL、CHS、ANS和UGT基因可能在紫芽茶树芽叶花青素生物合成途径中起关键作用。本研究探索了茶树芽叶紫色性状的转录组信息,筛选到一些差异表达基因,为进一步研究茶树芽叶呈色机制提供参考。

**关键词:**茶树品种;紫芽;花青素;转录组;差异基因

## Transcriptome Analysis of Anthocyanin Synthesis Related Genes in Purple Bud Tea Plant

JIANG Hui-bing<sup>1</sup>, XIA Li-fei<sup>1</sup>, TIAN Yi-ping<sup>1</sup>, DAI Wei-dong<sup>2</sup>, SUN Yun-nan<sup>1</sup>, CHEN Lin-bo<sup>1</sup>

(<sup>1</sup> Tea Research Institute, Yunnan Academy of Agricultural Sciences/Yunnan Engineering Research Center of Tea Germplasm Innovation and Matching Cultivation/Yunnan Provincial Key Laboratory of Tea Science, Menghai 666201;

<sup>2</sup> Tea Research Institute, Chinese Academy of Agricultural Sciences, Hangzhou 310008)

**Abstract:** In order to identify the candidate genes that are involved in purple color of tea plants, three tea genotypes consisting of purple bud 'Zijuan', green bud 'Yunkang No. 10' and 'Fuding Dabaicha' were used. We determined the contents and compositions of anthocyanin and catechin by high performance liquid chromatography-mass (HPLC-MS), and also analyzed gene expression of synthesis-related genes by RNA-sequencing (RNA-Seq) and quantitative real-time (qRT-PCR). HPLC-MS analysis results showed that the contents of anthocyanin in 'Zijuan', 'Yunkang No. 10' and 'Fuding Dabaicha' were 5.05 mg/g, 0.08 mg/g and 0.15 mg/g, respectively. The anthocyanin

收稿日期:2018-01-26 修回日期:2018-03-16 网络出版日期:2018-07-18

URL: <http://kns.cnki.net/kcms/detail/11.4996.S.20180717.1211.002.html>

基金项目:国家自然科学基金项目(31560220);茶树生物学与资源利用国家重点实验室开放基金项目(SKLT0F20150105);云南省人才培养计划项目(2015HB105)

第一作者研究方向为茶树种质资源与遗传育种研究, E-mail: jianghb119@126.com。夏丽飞为共同第一作者, E-mail: xialif@163.com

通信作者: 陈林波, 研究方向为茶树育种与生物技术研究, E-mail: chenlinbo2002@sina.com

content was significantly different in purple bud vs. green bud tea cultivars ( $P < 0.05$ ). Whereas the catechin contents of three tea cultivars were 15.12 mg/g, 19.52 mg/g and 15.09 mg/g, respectively, no significant difference was detected. By RNA-Seq, a total of 255079 reads was obtained and assembled into 166118 unigenes. Of that, 54446 were assigned to GO terms, 30666 were assigned to KOG categories, and 20336 were assigned to KEGG pathways. DEG-Seq analysis revealed 434 differentially expressed genes between purple bud vs. green bud tea cultivars, belonging to 38 metabolic pathways. 15 of 112 genes in anthocyanin biosynthesis pathway were differentially expressed between purple bud vs. green bud tea cultivars. Out of that, *PAL* (cluster-1850.70337), *CHS* (cluster-21971.0), *ANS* (cluster-1850.80848) and *UFGT* (cluster-1850.77163) were up-regulated in purple bud tea cultivar, while no significant difference between green bud tea cultivars was detected. In addition, *FLS* (Cluster-1850.81068) genes were up-regulated in green bud tea cultivars, possibly associated with the catechin accumulation. Therefore, this study revealed a set of differentially expressed genes that associated with the anthocyanin synthesis of purple bud tea plant, and particularly, *PAL*, *CHS*, *ANS* and *UFGT* genes may play central role.

**Key words:** Tea plant (*Camellia sinensis* (L.) Kuntze); purple shoot; anthocyanin; transcriptome; differentially expressed genes

茶树 [*Camellia sinensis* (L.) Kuntze] 为山茶属植物, 是我国重要的木本经济作物。茶叶中富含多酚类物质, 是形成茶叶品质的主要成分<sup>[1]</sup>。紫娟、苔香紫和红芽佛手等紫芽茶不仅含有丰富的儿茶素和黄酮类化合物, 而且花青素含量显著高于一般茶类, 具有更为明显的抗氧化、抗肿瘤等药理保健作用<sup>[2-3]</sup>, 越来越受人们关注。茶树芽叶紫色性状作为特色茶树品种改良的重要经济性状之一, 探明茶树芽叶花青素合成机制在理论和应用中均具有重要意义。

茶树芽叶呈紫色与其叶片内花青素的合成与累积密切相关, 紫芽茶中主要含有飞燕草素、矢车菊素和天竺葵素等, 当含量较高时能使茶树嫩叶呈现紫色、红色或紫红色<sup>[4]</sup>。近年来, 已对茶树类黄酮物质合成途径中的主要结构基因苯丙氨酸解氨酶 (*PAL*, phenylalanine ammonialyase)、查耳酮合成酶 (*CHS*, chalcone synthase)、查耳酮异构酶 (*CHI*, chalcone isomerase)、二氢黄酮醇 4-还原酶 (*DFR*, dihydroflavonol 4-reductase)、花青素合成酶 (*ANS*, anthocyanidin synthase)、无色花色苷还原酶 (*LAR*, leucoanthocyanidin reductase)、黄烷酮 3-羟基化酶 (*F3H*, flavanone 3-hydroxylase)、类黄酮 3'-羟基化酶 (*F3'H*, flavonoid 3'-hydroxylase)、类黄酮-3', 5'-羟基化酶 (*F3'5'H*, flavonoid 3', 5'-hydroxylase)、花色苷还原酶 (*ANR*, anthocyanidin reductase) 等<sup>[5-7]</sup>, 以及少数相关调节基因 *MYB* (vmyb avian myeloblastosis viral oncogene homolog)、*bHLH* (basic helix-loop-helix) 和 *WDR* (WD-repeat) 等<sup>[8]</sup> 进行了分离、克隆和表达研究, 为探索茶树芽叶呈色的遗传机制奠定了基础。然而花青素的合成与累积过程复杂, 涉及众

多酶基因的共同作用, 受到转录因子的调控, 以及内外因子的影响等<sup>[9]</sup>。目前茶树基因组信息还比较匮乏, 与茶树芽叶紫色性状形成相关的花青素合成调控机理还不完全明确, 紫芽和绿芽茶树芽叶中花青素合成相关酶基因的表达差异性是怎样的, 哪些是在紫芽茶树芽叶中促进花青素合成与积累的关键酶基因还有待进一步研究证实。

随着转录组测序技术日益成熟, 高通量测序得以应用于茶树研究领域, 成为茶树基因组研究和品种选育的重要工具<sup>[10]</sup>。本研究利用高通量测序技术对紫芽和绿芽茶树品种进行了转录组测序和生物信息学分析, 得到一些与花青素合成相关酶基因, 并应用数字基因表达谱和实时荧光定量 PCR 分析了紫芽与绿芽茶树品种间的差异表达基因, 筛选出影响茶树芽叶呈色的候选基因, 为研究茶树芽叶呈色的分子机理提供参考。

## 1 材料与方法

### 1.1 试验材料

选用云南省农业科学院茶叶研究所品种园内的紫娟 (紫芽)、云抗 10 号 (绿芽) 和福鼎大白茶 (绿芽) 为材料构建单基因簇 (Unigenes) 数据库。2017 年 10 月, 分别采集各材料同一时期 (1 芽 1 叶) 样品, 一部分微波固样, 做花青素、儿茶素含量测定; 一部分液氮固样, -80 °C 冰箱保存, 做转录组分析。

### 1.2 花青素和儿茶素含量测定

采用液质联用 (HPLC-MS, High Performance Liquid Chromatography-Mass Spectrometry) 法<sup>[11]</sup> 测定样品花青素含量。取 0.2 g 样品, 加入 10 mL, 20%

酸性乙醇(含 0.1% 盐酸),超声波提取(40 kHz, 20 min, 室温),离心(8000 r/min, 10 min)取上清液,沉淀物重复 2 次,合并上清液(30 mL),旋转蒸发(45 °C),3% 甲醇定容至 5 mL, SPE 柱净化,复溶(2 mL 0.1% 盐酸水溶液),过滤膜(0.45 μm)。Agilent 1100 液相色谱仪串联离子阱质谱联用,流动相 A 相(0.1% 甲酸-水溶液)、B 相(100% 乙腈)。进样量 5 μL,流速 0.3 mL/min,柱温 40 °C。电喷雾离子源(ESI),正离子模式扫描( $m/z$  200-650),干燥气体(N<sub>2</sub>, 12 L/min, 350 °C),喷雾气压(30 psi),毛细管电压(3500 V)。

采用高效液相色谱法(HPLC, High Performance Liquid Chromatography)法<sup>[12]</sup>测定样品儿茶素含量。取 0.2 g 样品,加入 5 mL 70% 甲醇,水浴提取(70 °C, 10 min),离心(8000 r/min, 10 min)取上清液,沉淀物重复 2 次,合并上清液定容至 10 mL,过滤膜(0.45 μm)。色谱柱(ZORBAX SB-C18 ODS 4.6 mm × 150 mm, 5 μm),流动相 A 相(2% 冰乙酸)、B 相(100% 乙腈)。进样量 10 μL,流速 1 mL/min,柱温 30 °C,波长 280 nm。每个样品 3 次重复。儿茶素总量为各单体儿茶素含量之和。

### 1.3 RNA 提取与文库构建

用 CTAB-LiCl<sub>1</sub> 法<sup>[13]</sup>提取各样品的总 RNA, 2% 琼脂糖凝胶电泳分析 RNA 降解程度及有无污染, Nanodrop 检测 RNA 纯度, Qubit 定量 RNA 浓度, Agilent 2100 检测 RNA 完整性。cDNA 文库构建与测序由北京诺禾致源科技股份有限公司完成,通过 Illumina HiSeq 2500 测序平台,得到原始读序(Raw reads),经过滤得到干净读序(Clean reads),利用

表 1 实时荧光定量 PCR 引物

Table 1 The primers used for real-time quantitative PCR

基因 Gene	基因 ID Accession No.	上游引物 Forward primer(5'-3')	下游引物 Reverse primer(5'-3')	产物长度(bp) Product length
<i>PAL</i>	Cluster-1850.70337	AGAAAAACCGCTGAAGCTATTG	TTTGGCGACTTGCTAACA	141
<i>C4H</i>	Cluster-1850.63303	GTTCATTCCGACCATCCTG	CGTCCAATCTTCTCTCCCA	74
<i>4CL</i>	Cluster-1850.118690	TTCTCCTCCAGATGGGTGTTT	TGGTAGACCCGTGTCTCTG	130
<i>CHS</i>	Cluster-21971.0	GTGCCATCACATTACAACCTTCG	TGAAATGCCTCCACCAAGAT	109
<i>FLS</i>	Cluster-1850.81068	AGTGTTGTTCATTAGGGATTGG	GGTCTAGGACATGGTGGGTAAT	112
<i>DFR</i>	Cluster-1850.18936	GGACCGATGTTGACTTCTGC	GGATGAATGGCGACGAGAT	130
<i>ANR</i>	Cluster-1850.82756	CGCTGTCAATACCAGTGTCC	CCCATACTTGAAACTGAATCCC	157
<i>ANS</i>	Cluster-1850.80848	CGTCCCCAACTCCATTATCA	GCCCACGAAATCCTCACTT	120
<i>UFGT</i>	Cluster-1850.77163	ATTGGGAACITAGAATCAGCC	CGTGGTGGTATGTGAGGTT	183
<i>GAPDH</i>	GE651107.1	GATAGTGTTCACGGTCAATGGA	GCAGCAGCCTTATCCTTATCAG	188

Trinity 软件通过序列之间的重叠信息组装得到重叠群(Contigs),局部组装得到转录本(Transcripts),用 Tgicl 和 Phrap 软件对转录本(Transcripts)进行同源聚类 and 拼接得到单基因簇(Unigenes)。

### 1.4 基因注释、分类和代谢通路分析

通过 Blast 比对工具,单基因簇(Unigenes)与蛋白数据库(NR、NT、KO、Swissprot、PFAM、GO 和 KOG)进行比对,在相似度 > 30%、E-value < 1e-5 条件下得到预测蛋白的 GO 信息。合并归纳 Unigenes 注释信息,进行 KOG 功能分类预测,根据 GO 信息和 KOG 注释信息,对 Unigenes 进行分类。在 E-value < 1e-3 条件下与 KEGG 数据库进行比对,获得 KEGG 注释的代谢通路信息。采用 TMM 对 Read count 数据进行标准化处理,利用 DEG-Seq 在 Qvalue < 0.005 且  $|\log_2(\text{fold change})| > 1$  条件下对差异基因进行差异表达分析。

### 1.5 实时荧光定量 PCR 检测

从测序获得的差异表达基因中,随机挑选 9 个进行实时荧光定量 PCR (Quantitative Real-time, qRT-PCR)测定。采用 CTAB-LiCl<sub>1</sub> 法提取总 RNA,按照 Fermentas 公司的 RevertAid First Strand cDNA Synthesis Kit 说明书合成 cDNA。依据测序所得序列库对应的代表序列,使用 Primer5.0 软件设计实时荧光定量 PCR 引物,以 *GAPDH* (GE651107.1) 基因作为内参基因(表 1),采用比较阈值法,用  $2^{-\Delta\Delta CT}$  法计算基因相对表达量,对照(云抗 10 号)中各目标基因的相对表达量为 1。实时荧光定量 PCR 的反应体系及程序参见 Fermentas 公司的 SYBR Green master mix 说明书。实验设 3 个重复。

## 2 结果与分析

### 2.1 花青素和儿茶素含量比较

从紫娟、云抗 10 号和福鼎大白茶中共检测到 8 种花青素组分,分别为飞燕草色素(Delp, delphinidin)、飞燕草-3-O-半乳糖苷(D3Ga, delphinidin-3-O-Galactoside)、飞燕草-3-O-(6-香豆酰)-半乳糖苷(DCZGa, delphinium-3-O-(6-p-coumaricacyl)-galactosidase)、矢车菊素-3-O-半乳糖苷(C3Ga, cyanidin-3-O-glucoside)、矢车菊素-3-O-(6-香豆酰)-半乳糖苷(CCZDGa, cyanidin-3-O-(6-p-coumaricacyl)-galactosidase)、天竺葵色素(Pelar, pelargonidin)、天竺葵-3-O-葡萄糖苷(P3G, pelargonidin-3,5-diglucoside)和芍药素(Peon, peonin)。紫娟茶树品种中含有除 Delp 外的 7 种成分,其中 DCZGa 含量最高,占花青素总含量的 43.56%。其次是 CCZDGa,占花青素总含量的

29.70%。云抗 10 号和福鼎大白茶中含有 Delp、Pelar、DCZGa 和 CCZDGa 等 4 种花青素组分,各组分含量均较低。紫娟的花青素总含量(5.05 mg/g)分别是云抗 10 号(0.08 mg/g)的 63 倍和福鼎大白茶(0.15 mg/g)的 34 倍,差异显著(表 2)。儿茶素组分检测表明,3 个茶树品种均以表没食子儿茶素没食子酸酯(EGCG, epigallocatechin gallate)和表儿茶素没食子酸酯(ECG, epicatechin gallate)含量较高,儿茶素(C, catechin)、表儿茶素(EC, epicatechin)和表没食子儿茶素(EGC, epigallocatechin)含量均较低,总儿茶素含量分别为 15.12 mg/g、19.52 mg/g 和 15.09 mg/g,差异不显著(表 3)。各品种内总儿茶素含量分别是花青素含量的 3 倍、244 倍和 101 倍。表明在紫芽茶树芽叶类黄酮合成途径中积累了大量花青素和儿茶素,在绿芽茶树芽叶中有大量儿茶素合成,几乎无花青素积累。

表 2 不同茶树品种花青素组分及含量

Table 2 The compositions and contents of anthocyanin in different tea cultivars

序号 No.	化合物 Compounds	含量(mg/g) Contents		
		紫娟 Zijuan	云抗 10 号 Yunkang No. 10	福鼎大白茶 Fuding Dabaicha
1	飞燕草色素 Delp	—	0.02	0.03
2	飞燕草-3-O-(6-香豆酰)-半乳糖苷 DCZGa	2.20	0.03	0.06
3	矢车菊素-3-O-(6-香豆酰)-半乳糖苷 CCZDGa	1.50	0.02	0.04
4	天竺葵色素 Pelar	0.02	0.01	0.02
5	飞燕草-3-O-半乳糖苷 D3Ga	0.52	—	—
6	矢车菊素-3-O-半乳糖苷 C3Ga	0.73	—	—
7	天竺葵-3-O-葡萄糖苷 P3G	0.06	—	—
8	芍药素 Peon	0.02	—	—
总计 Total	花青素含量 Anthocyanin total	5.05a	0.08b	0.15b

“—”表示未检测到;不同字母表示在 0.05 水平上差异显著。下同

“—”indicates not detected. By a different letter are significantly different at 5% probability level. The same as below

表 3 不同茶树品种儿茶素组分及含量

Table 3 The compositions and contents of catechin in different tea cultivars

序号 No.	化合物 Compounds	含量(mg/g) Contents		
		紫娟 Zijuan	云抗 10 号 Yunkang No. 10	福鼎大白茶 Fuding Dabaicha
1	儿茶素 C	1.51	3.74	2.48
2	表儿茶素 EC	1.22	0.56	0.53
3	表没食子儿茶素 EGC	0.69	0.52	0.34
4	表儿茶素没食子酸酯 ECG	4.31	4.03	3.13
5	表没食子儿茶素没食子酸酯 EGCG	7.39	10.67	8.61
总计 Total	总儿茶素含量 Catechin total	15.12a	19.52a	15.09a

相同字母表示在 0.05 水平上差异不显著

By a same letter are no significantly different at 5% probability level

## 2.2 转录组测序与序列拼接

对紫娟、云抗 10 号和福鼎大白茶 1 芽 1 叶进行转录组测序。经去除接头、N 和低质量读序 (Reads), 分别获得 72009220 条、75068850 条和 53594152 条干净读序 (Clean reads), 各样品的转录组碱基数分别为 10.80 Gb、11.26 Gb 和 8.04 Gb。获得序列的 Q20 值均大于 97%, Q30 值均大于 93%, GC 含量均大于 44%。利用 Trinity 对干净读序 (Clean reads) 进行 *de novo* 拼接, 去重复后在 3 个样品中共获得 255079 条转录本 (Transcripts) 序列。取每条基因中最长的转录本作为单基因簇, 共筛选出 166118 条单基因簇, 其最大长度为 29300 bp, 最小长度为 201 bp, 平均长度为 914 bp, N50 为 1344 bp。对转录本及单基因簇的长度分别统计, 其中单基因簇小于 301 bp 的有 23895 条, 占总数的 14.38%。301~500 bp 间的有 43412 条, 占总数的 26.13%。501~1000 bp 间的有 50617 条, 占总数的 30.47%。1001~2000 bp 间的有 31673 条, 占总数的 19.07%。大于 2000 bp 的有 16521 条, 占总数的 9.95% (图 1)。

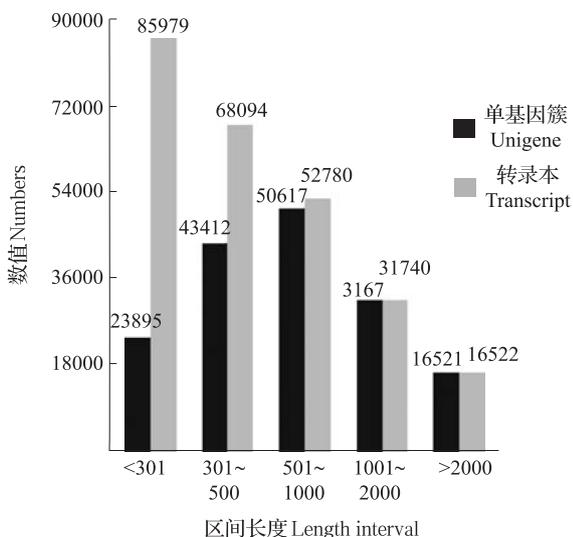


图 1 拼接转录本与 Unigene 基因序列长度分布图

Fig. 1 Statistics between transcripts and unigenes

## 2.3 转录物功能注释与分类

为明确单基因簇序列信息, 使用 Blast 程序将组装得到的单基因簇与 NR、NT、KO、Swissprot、PFAM、GO 和 KOG 等 7 种数据库进行比对, 所有单基因簇均获得注释。其中在 GO 功能分类体系中, 有 54446 条 (32.77%) 单基因簇得到了注释, 被分配到 283476 条 GO 条目下, 可分为生物过程 (135829)、细胞组分 (83742) 和分子功能 (63905) 3 个大类和

56 个功能类别。在生物过程类别下, 细胞过程 (30672 条, 22.58%)、代谢过程 (28540 条, 21.01%) 和单有机体过程 (22355 条, 16.46%) 3 个组分涉及的单基因簇较多。在细胞组分类别下, 细胞 (16467 条, 19.66%)、细胞部分 (16462 条, 19.66%) 和细胞器 (11059 条, 13.21%) 3 个组分涉及的单基因簇较多。在分子功能类别下, 结合 (30049 条, 47.02%) 和催化活性 (23983 条, 37.53%) 两个分组包含了 80% 以上的单基因簇 (图 2)。

在 KOG 数据库中, 有 30666 条单基因簇得到注释, 被分配到 26 个 KOG 功能类别。其中, 一般功能基因 (5135 条, 16.74%) 包含的单基因簇最多, 其次是翻译后修饰、蛋白质折叠和分子伴侣 (3556 条, 11.60%)。其他类别的基因表达丰度都各不相同, 其单基因簇数量都在 3000 条以下。最少的为细胞运动和未命名蛋白, 单基因簇数量分别为 15 条和 1 条 (图 3)。

在 KEGG 数据库中, 有 20336 条单基因簇获得注释, 可归为 5 个层次、19 个亚层次及 130 个代谢通路。5 大层次包括了细胞过程 (1265 条)、环境信息处理 (952 条)、遗传信息处理 (5579 条)、新陈代谢 (11505 条) 和有机系统 (1035 条), 其中最大的分类是新陈代谢。在 19 个亚层次代谢通路中, 翻译 (2346 条)、碳水化合物代谢 (2253 条)、分选与降解 (1811 条)、整理 (1692 条) 及折叠氨基酸代谢 (1401 条) 占据该分类系统的前 5 位。在 130 个代谢通路中, 排前 10 位的有核糖体 (ko03010)、碳代谢 (ko01200)、植物-病原互作 (ko04626)、氨基酸合成 (ko01230)、内质网蛋白处理 (ko04141)、嘌呤代谢 (ko00230)、剪接体 (ko03040)、RNA 运输 (ko03013)、植物激素信号转导 (ko04075)、淀粉和蔗糖代谢 (ko00500) 及内吞作用 (ko04144), 其单基因簇数分别为 1041 条、1033 条、902 条、869 条、681 条、613 条、596 条、587 条、583 条和 543 条 (图 4)。

## 2.4 不同茶树品种基因差异表达分析

将 3 个不同茶树品种两两比对, 即紫娟/云抗 10 号、紫娟/福鼎大白茶、云抗 10 号/福鼎大白茶筛选差异表达基因, 对 3 个组合筛选的差异表达基因绘制维恩图 (图 5)。结果表明, 在紫娟/云抗 10 号中获得 1248 条差异表达基因, 上调基因数 630 条, 下调基因数 618 条。在紫娟/福鼎大白茶中获得 1985 条差异表达基因, 上调基因数 1208 条, 下调基因数 777 条。在云抗 10 号/福鼎大白茶中获得 2182 条差异表达基因, 上调基因数 1276 条, 下调基

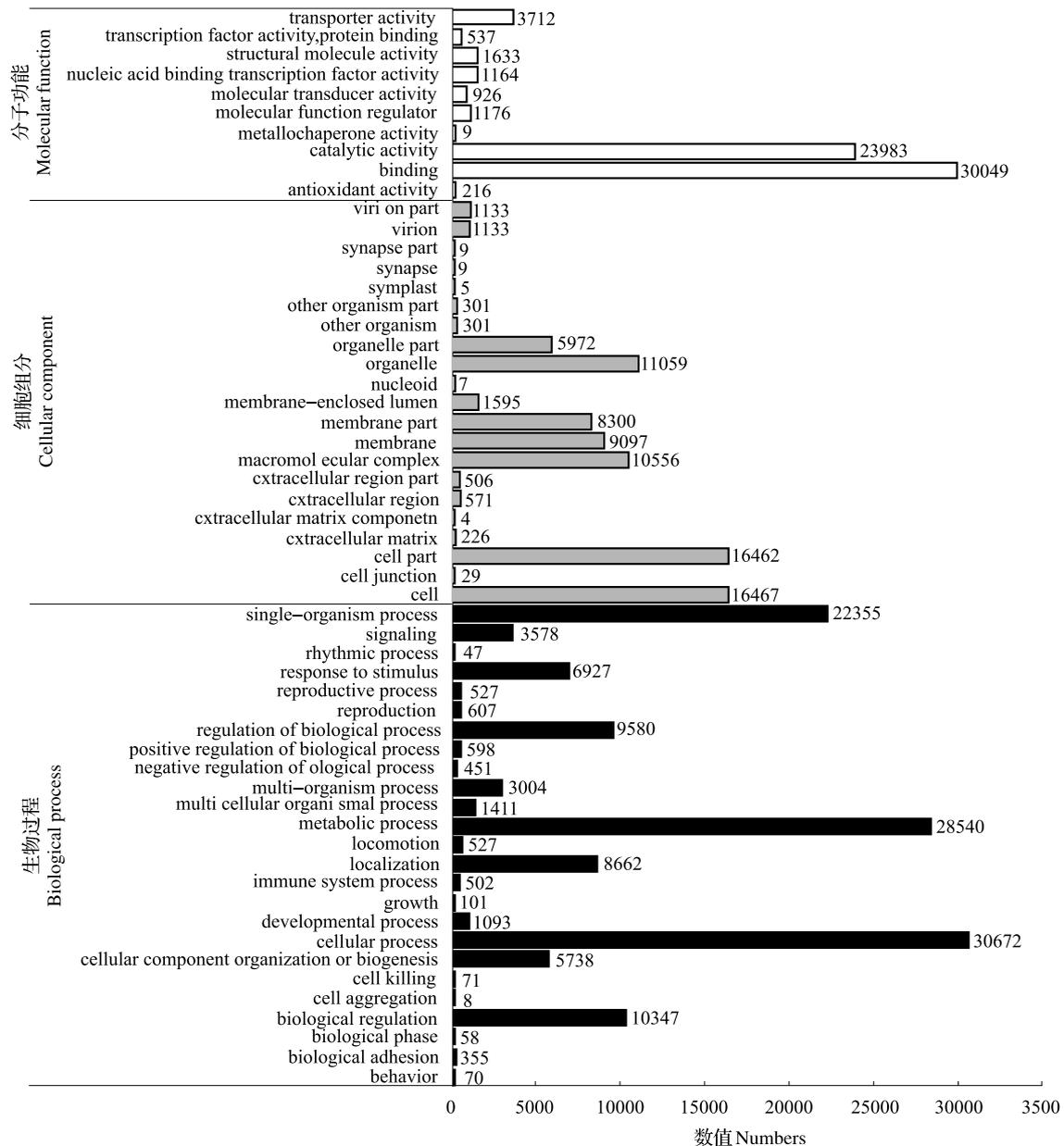


图2 功能基因的 GO 注释情况

Fig. 2 Gene ontology annotation of differentially expressed genes

因数 906 条。KEGG 富集分析显示, 在处理组 (紫芽品种) 与两个对照组 (绿芽品种) 间共有 434 条差异表达基因, 这些差异基因主要涉及到氨基糖和核苷酸糖代谢, 淀粉和蔗糖代谢, 植物激素信号转导, 核糖体, 缬氨酸、亮氨酸和异亮氨酸降解, 内质网中蛋白处理, 苯丙氨酸、酪氨酸和色氨酸生物合成, 剪接体, 半胱氨酸和蛋氨酸代谢, 植物-病原互作, 糖酵解和糖质新生, 丙酮酸代谢, RNA 降解, 甘氨酸、丝氨酸和苏氨酸代谢, 过氧化物酶体, 黄酮和黄酮醇生物合成, 内吞作用等 38 个代谢通路。

## 2.5 花青素合成途径关键基因筛选和验证

通过 NCBI 数据库进行 Blast 比对, 共筛选到与类黄酮生物合成途径相关表达基因 112 条, 包括苯丙氨酸解氨酶 (PAL, phenylalanine ammonia-lyase) 30 条、肉桂酸羟化酶 (C4H, cinnamate acid 4-hydroxylase) 13 条、4-香豆酰 CoA 连接酶 (4CL, 4-coumarate: CoA ligase) 15 条、查耳酮合成酶 (CHS, chalcone synthase) 15 条、查耳酮异构酶 (CHI, chalcone isomerase) 2 条、黄酮 3-羟基化酶 (F3H, flavanone 3-hydroxylase) 2 条、类黄酮 3'-羟基化酶 (F3'H, flavonoid 3'-hydroxylase) 1 条、类黄酮 3', 5'-羟基化酶

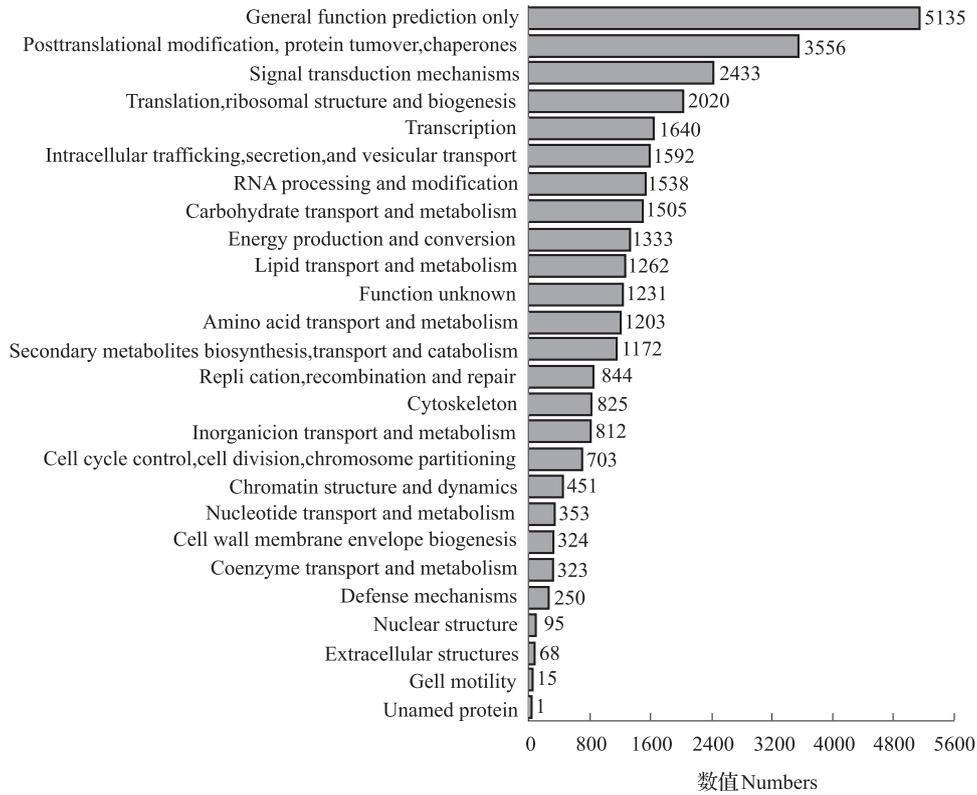


图 3 功能基因的 KOG 注释情况

Fig. 3 KOG annotation of differentially expressed genes

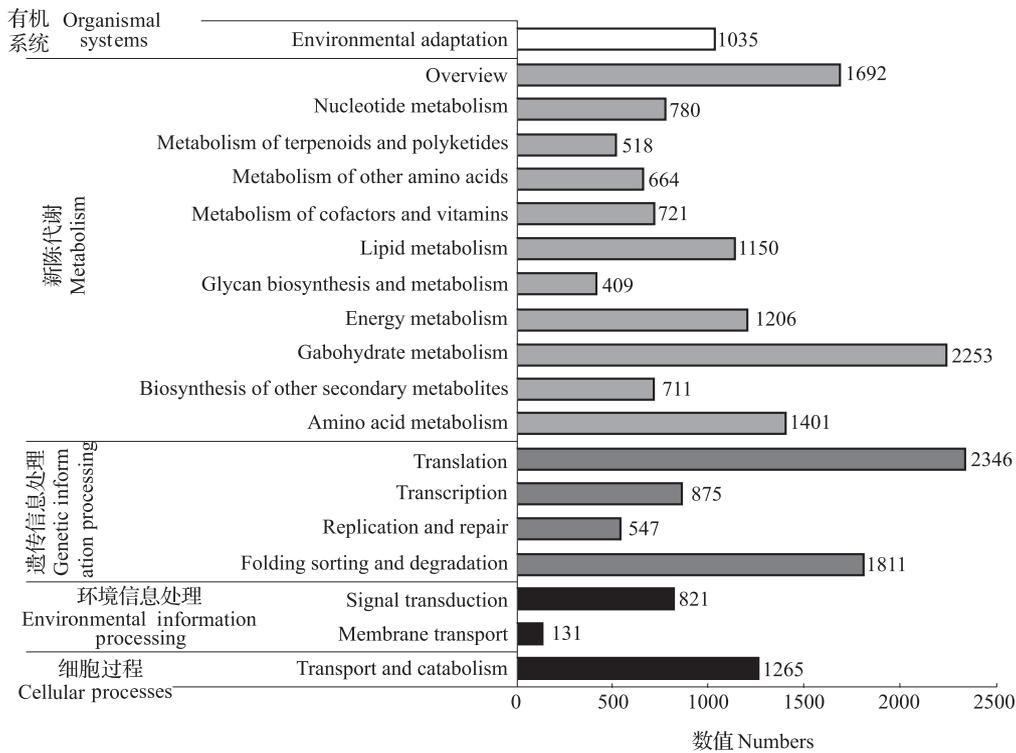
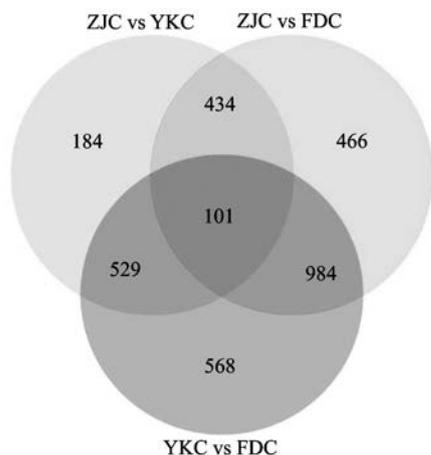


图 4 功能基因的 KEGG 代谢通路分析

Fig. 4 KEGG metabolism pathway categories



ZJC vs YKC: Zijuan vs Yunkang No. 10; ZJC vs FDC: Zijuan vs Fuding Dabaicha; YKC vs FDC: Yunkang No. 10 vs Fuding Dabaicha

图5 差异基因维恩图

Fig. 5 The Venn of differentially expressed genes

(F3'5'H, flavonoid 3',5'-hydroxylase) 6条、二氢黄酮醇4-还原酶(DFR, dihydroflavonol 4-reductase) 11条、花色素合成酶(ANS, anthocyanidin synthase) 4条、黄酮醇合成酶(FLS, flavonol synthase) 4条、无色花色素还原酶(LAR, leucoanthocyanidin reductase) 6条、花色素还原酶(ANR, anthocyanidin reductase) 2条和类黄酮糖基转移酶(UFGT, UDP-glycose: flavonoid glucosyltransferase) 1条。使用RPKM法计算基因表达量(图6),结果显示,有15条基因在紫芽与绿芽茶树品种间差异表达,其中PAL(Cluster-1850.70337, Cluster-18341.0)、CHS(Cluster-21971.0, Cluster-20219.0)、ANS(Cluster-1850.80848)和UFGT(Cluster-1850.77163)等6条差异基因在紫芽茶树品种紫娟中上调表达,在绿芽茶树品种云抗10号和福鼎大白茶中表达差异不显著。CAH(Cluster-2495.0, Cluster-5457.0)、4CL(Cluster-1850.85078, Cluster-1850.84740)和DFR(Cluster-1850.18936, Cluster-1850.16996)等6条差异基因在紫娟和云抗10号中表达差异不显著,在福鼎大白茶中下调表达。FLS(Cluster-1850.81151, Cluster-1850.81068)基因在绿芽茶树品种云抗10号和福鼎大白茶中均为上调表达,在2个绿芽品种间差异不显著。ANR(Cluster-1850.82756)基因在云抗10号中上调表达,但在紫娟和福鼎大白茶间表达差异不显著(图6)。说明紫芽与绿芽茶树品种间花青素含量差异与合成相关基因的表达水平有关。

以GAPDH(GE651107.1)为内参基因,利用qRT-PCR方法对PAL、CAH、4CL、CHS、FLS、DFR、

ANR、ANS和UFGT等9个差异表达基因进行了标准化验证。比较3个茶树品种发现,PAL基因在紫娟中的表达量最高,分别是云抗10号和福鼎大白茶的2.38倍和2.42倍。CAH、4CL和DFR在紫娟和云抗10号中的表达量较高,在福鼎大白茶中表达量较低。CHS在紫娟中的表达量最高,分别是云抗10号和福鼎大白茶的2.77倍和28.06倍。FLS在紫娟中的表达量较低,分别是云抗10号的0.50倍和福鼎大白茶的0.33倍。ANR在云抗10号中的表达量较高,在紫娟和福鼎大白茶中的表达量相当。ANS在紫娟中的表达量最高,分别是云抗10号和福鼎大白茶的5.73倍和8.23倍。UFGT在紫娟中的表达量最高,分别是云抗10号和福鼎大白茶的1.68倍和2.25倍。qRT-PCR检测结果显示,PAL、CAH、4CL、CHS、ANS和UFGT等基因在紫芽茶树品种紫娟中的表达量均高于绿芽茶树品种云抗10号和福鼎大白茶,而FLS基因在紫芽茶树品种中的表达量低于2个绿芽茶树品种。总体而言,所检测的9个差异表达基因在不同茶树品种中的表达水平与转录组测序结果基本一致(图6),说明表达谱测序结果可靠。

### 3 讨论

本研究从紫娟、云抗10号和福鼎大白茶芽叶中共检测到8种花青素组分,其中DCZGa和CCZDGa两种组分含量较高,与花青素积累量密切相关。紫芽茶树品种紫娟花青素含量分别是绿芽茶树品种云抗10号的63倍和福鼎大白茶的34倍,差异极显著。3个品种间总儿茶素含量差异不显著,各品种内总儿茶素含量均明显大于花青素含量,差异极显著。表明在紫芽茶树芽叶类黄酮合成过程中积累了大量花青素和儿茶素,而绿芽茶树芽叶几乎无花青素的合成,这同以往的研究结果一致<sup>[4,11]</sup>。

通过高通量测序,从3个茶树品种样本中共获得了255079条转录本序列,筛选出166118条单基因族。其中有54446条单基因族注释到GO分类的56个功能类别中,有30666条单基因族注释到KOG分类的26个功能类别中,有20336条单基因族注释到KEGG分类的130个代谢通路中。数字基因表达谱分析显示,紫芽与绿芽茶树品种间共有434条差异表达基因,KEGG富集分析涉及到38个代谢通路。这些结果增加了紫娟叶片花青素形成过程的转录组数据信息。本研究筛选到与花青素、儿茶素代谢途径相关基因112条,其中15条基因在紫芽与绿芽

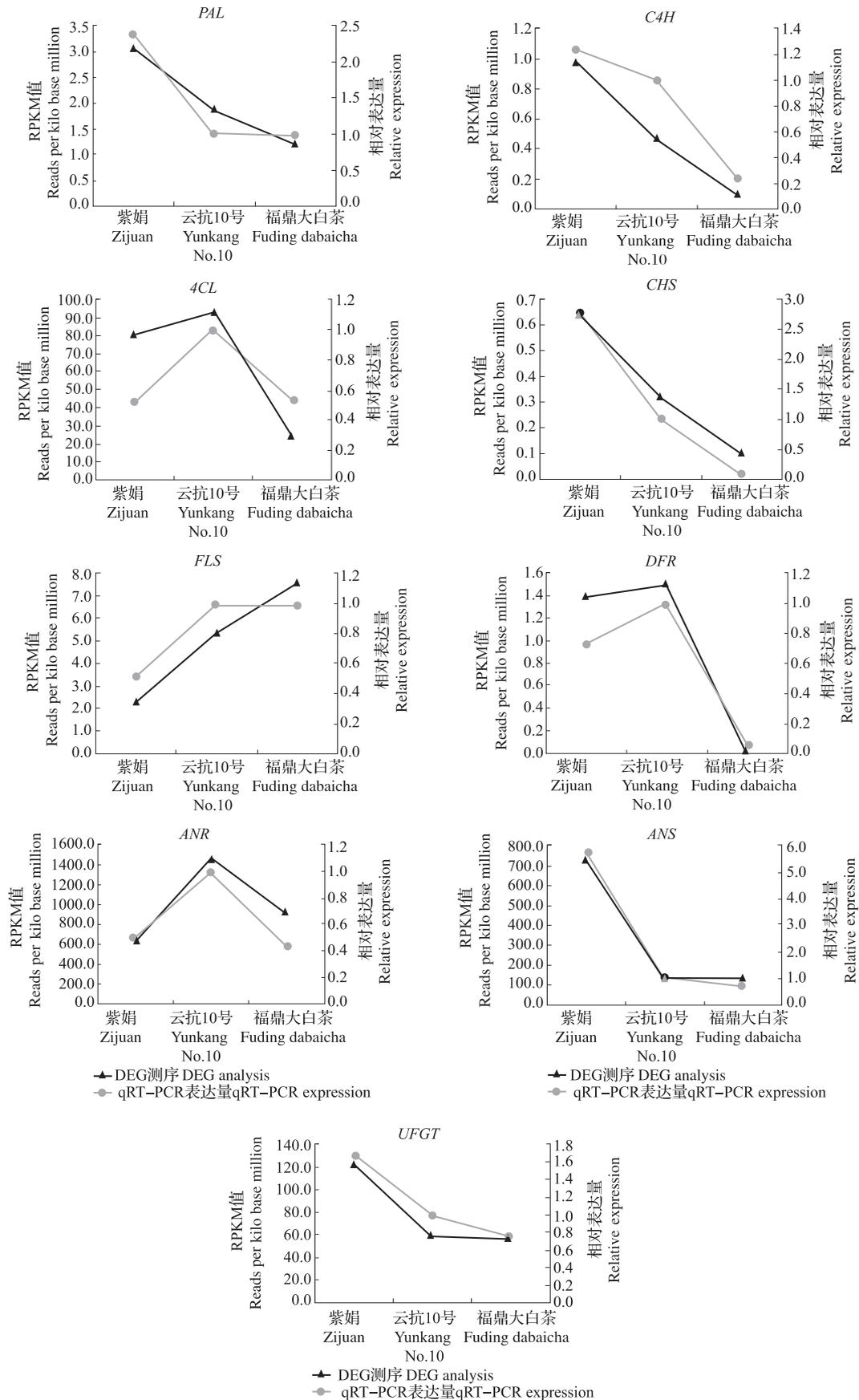


图6 DEG 测序结果与 qRT-PCR 表达量比对

Fig. 6 The transcriptional analysis of key genes by DEG sequencing and qRT-PCR

茶树品种间差异表达,包含了 *PAL*、*C4H*、*4CL*、*CHS*、*FLS*、*DFR*、*ANS*、*ANR* 和 *UFGT* 等 9 个酶基因。除 *FLS* 和 *ANR* 在紫娟中下调表达外,其余差异基因在紫娟中均为上调表达。进一步利用实时荧光定量 PCR 方法对测序结果中差异表达基因进行了验证,结果显示,9 个差异表达基因在不同茶树品种间的相对表达水平与高通量测序结果基本一致,说明测序结果准确反映了紫娟与云抗 10 号和福鼎大白茶基因表达的差异性。

茶树中花青素和儿茶素均是由类黄酮代谢途径合成,共享了早期一部分的生物合成途径<sup>[14]</sup>(图 7)。研究表明,茶树中 *PAL*、*CHS* 和 *DFR* 等基因的表达与儿茶素类化合物的累积呈正相关<sup>[15]</sup>。本研究中 *PAL*、*CHS* 基因在紫芽茶树品种紫娟中上调表达,在两个绿芽茶树品种云抗 10 号和福鼎大白

茶间表达差异不显著,与花青素积累情况相吻合,说明 *PAL*、*CHS* 基因的表达还与花青素的累积密切相关。这与周天山等<sup>[16]</sup>和周琼琼等<sup>[17]</sup>对茶树芽叶紫化关键酶基因的表达分析结果(*PAL*、*CHS* 基因在紫芽茶树资源中上调表达)一致,而与马春雷等<sup>[5]</sup>研究结果(*PAL* 在紫芽茶树资源中下调表达)不同。也验证了蓝花喜盐鸢尾中 *CHS* 上调表达对花青素合成起关键作用<sup>[18]</sup>。夏涛等<sup>[14]</sup>认为 *DFR* 基因在不同儿茶素含量的细胞系之间表达水平差异较大。本研究检测到 *C4H*、*4CL* 和 *DFR* 在紫娟和云抗 10 号中的表达量较高,在福鼎大白茶中下调表达,这些基因在紫芽和绿芽茶树品种中均有表达,可能是由于不同茶树品种间遗传因素的影响,使得儿茶素和多酚类物质的积累不同<sup>[19]</sup>,其作用机制有待进一步研究。

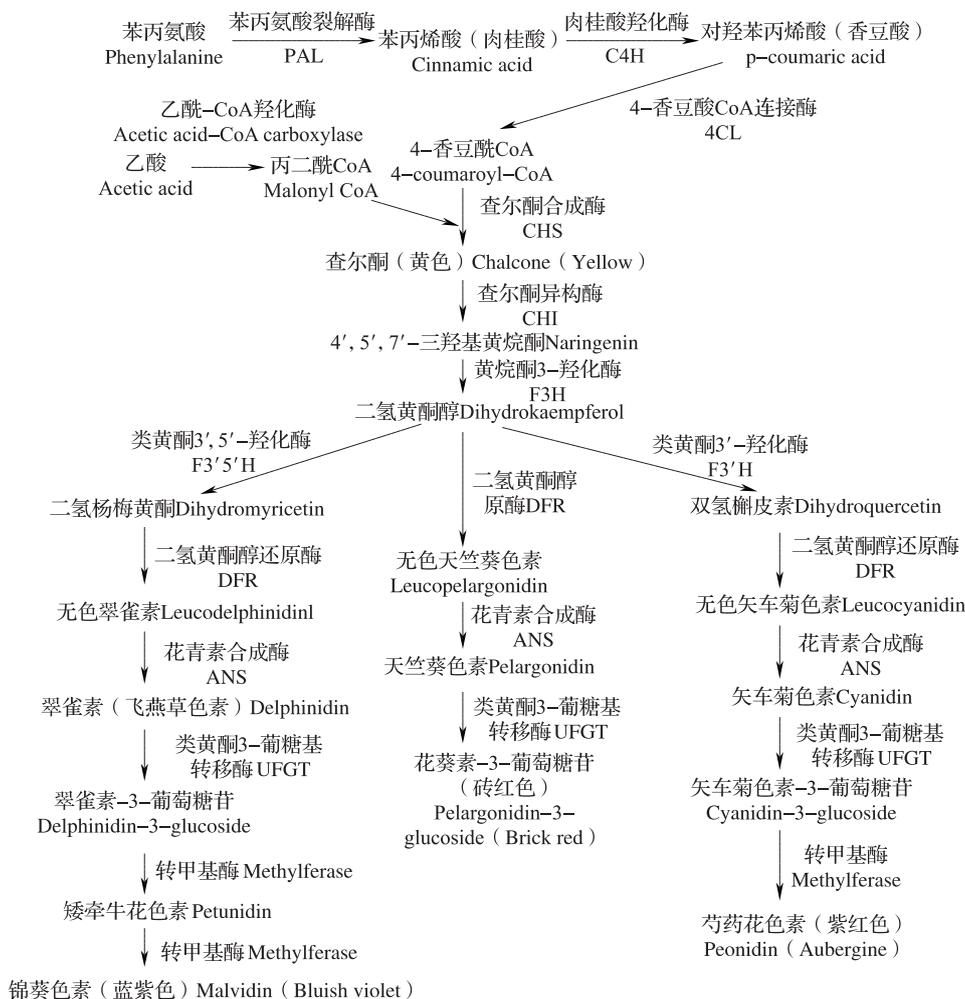


图 7 花青素生物合成途径<sup>[30]</sup>

Fig. 7 Biosynthesis pathway of anthocyanin<sup>[30]</sup>

*ANS* 基因的高表达使紫芽果皮、种皮、糊粉层皆呈紫色<sup>[20]</sup>, *UFGT* 是苹果果皮<sup>[21]</sup>、桃果肉<sup>[22-23]</sup> 呈红色的关键基因。本研究中 *ANS* 和 *UFGT* 基因在紫芽茶树品种紫娟中上调表达, 在两个绿芽茶树品种云抗 10 号和福鼎大白茶间表达差异不显著, 与花青素积累情况一致, 可能是导致紫娟花青素含量高而叶片呈紫色的关键酶基因。其原因可能与 *ANS* 能催化无色花青素形成有色花青素<sup>[24]</sup>, *UFGT* 催化不稳定的花青素经糖基化修饰转化为稳定的花青素等有关<sup>[25]</sup>。 *FLS* 基因在烟草中过量表达, 可增加黄酮醇类化合物积累, 而减少花青素合成<sup>[26]</sup>。本研究检测到 *FLS* 基因在紫芽茶树品种紫娟中下调表达, 在绿芽茶树品种云抗 10 号和福鼎大白茶中上调表达, 与花青素积累情况相反, 推测在绿芽茶树芽叶类黄酮合成途径中, 可能存在 *FLS* 和 *DFR* 相互竞争底物<sup>[27]</sup>, 使黄烷醇含量增加, 无色花色素减少, 从而影响花青素合成, 导致绿芽茶树叶片花青素含量极低。Pang 等<sup>[28]</sup> 研究表明, 茶树中的 *ANRI* 和 *ANR2* 基因能将花青素和飞燕草色素转化成表儿茶素类 (EC 和 EGC)。周天山等<sup>[16]</sup> 研究表明, *ANRI* 和 *ANR2* 在紫芽茶中上调表达, 与其总儿茶素含量情况不一致, 认为紫芽茶总儿茶素含量偏低与花青素还原酶 (*ANR*) 无关。本研究中, *ANR* (Cluster-1850. 82756) 基因在云抗 10 号中上调表达, 在紫娟和福鼎大白茶间表达差异不显著, 表达趋势与总儿茶素含量情况相吻合, 说明 *ANR* 基因在紫芽和绿芽茶树品种中均有表达, 其作用机制可能与茶树芽叶儿茶素合成有关<sup>[29]</sup>。

本研究分析了紫芽和绿芽茶树品种芽叶花青素、儿茶素的含量和组分, 以转录组测序为基础, 探讨了茶树芽叶花青素合成相关基因的差异表达情况。表明紫芽茶树品种芽叶花青素含量与其生物合成途径相关基因的表达水平有关, 紫芽与绿芽茶树品种间芽叶花青素合成相关基因表达存在较大的差异性。从基因表达模式推测, *PAL* (Cluster-1850. 70337)、*CHS* (Cluster-21971. 0)、*ANS* (Cluster-1850. 80848) 和 *UFGT* (Cluster-1850. 77163) 基因可能在紫芽茶树品种芽叶花青素生物合成途径中起关键作用, 但茶树花青素代谢调控机制复杂, 受遗传因素及多种环境因子影响, 相关转录调控机理有待进一步研究证实。

#### 参考文献

[1] 宛晓春. 茶叶生物化学, 3 版. 北京: 中国农业出版社, 2003: 9-30

- [2] Jiang L H, Shen X J, Shoji T, Kanda T, Zhou J C, Zhao L M. Characterization and activity of anthocyanins in Zijuan tea (*Camellia sinensis* var. *kitamura*). *Journal of Agricultural and Food Chemistry*, 2013, 61 (13): 3306-3310
- [3] Wang Q P, Peng C X, Gao B, Gong J S. Influence of large molecular polymeric pigments isolated from fermented Zijuan tea on the activity of key enzymes involved in lipid metabolism in rat. *Experimental Gerontology*, 2012, 47 (9): 672-679
- [4] 吴华玲, 何玉媚, 李家贤, 陈栋, 黄华林, 乔小燕, 刘军. 11 个红紫芽茶树新品系的芽叶特性和生化成分研究. *植物遗传资源学报*, 2012, 13 (1): 42-47
- [5] 马春雷, 姚明哲, 王新超, 金基强, 陈亮. 利用基因芯片筛选茶树芽叶紫化相关基因. *茶叶科学*, 2011, 31 (1): 59-65
- [6] 陈林波, 夏丽飞, 孙云南, 梁名志, 张正竹, 李叶云, 宛晓春. 特异茶树品种“紫娟”叶色转变的基因表达差异分析. *茶叶科学*, 2012, 32 (1): 59-65
- [7] Wei K, Wang L Y, Zhang C C, Wu L Y, Li H L, Zhang F, Cheng H. Transcriptome analysis reveals key flavonoid 3'-hydroxylase and flavonoid 3',5'-hydroxylase genes in affecting the ratio of dihydroxylated to trihydroxylated catechins in *Camellia sinensis*. *PLoS One*, 2015, 10 (9): e0137925
- [8] 贡年娣, 郭丽丽, 王弘雪, 赵磊, 王婕, 王文钊, 刘亚军, 王云生, 高丽萍, 夏涛. 茶树两个 MYB 转录因子基因的克隆及功能验证. *茶叶科学*, 2014, 34 (1): 36-44
- [9] 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
- [10] Xia E H, Zhang H B, Sheng J, Li K, Zhang Q J, Kim C, Zhang Y, Liu Y, Zhu T, Li W, Huang H, Tong Y, Nan H, Shi C, Shi C, Jiang J J, Mao S Y, Jiao J Y, Zhang D, Zhao Y, Zhao Y J, Zhang L P, Liu Y L, Liu B Y, Yu Y, Shao S F, Ni D J, Eichler E E, Gao L Z. The tea tree genome provides insights into tea flavor and independent evolution of caffeine biosynthesis. *Molecular Plant*, 2017, 10 (6): 866-877
- [11] 李智, 王日为, 张丽霞, 韩晓阳. 茶树紫色芽叶花青素组分分析及结构推测. *茶叶科学*, 2014, 34 (3): 279-287
- [12] 郭颖, 黄峻榕, 陈琦, 吴雪原, 程浩, 吴琼. 茶叶中儿茶素类测定方法的优化. *食品科学*, 2016, 37 (6): 137-141
- [13] 史成颖, 宛晓春, 江昌俊, 孙俊. 提取高质量茶树总 RNA 的方法研究. *安徽农业大学学报*, 2007, 34 (3): 360-363
- [14] 夏涛, 高丽萍. 类黄酮及茶儿茶素生物合成途径及其调控研究进展. *中国农业科学*, 2009, 42 (8): 2899-2908
- [15] Mamati G E, Liang Y R, Lu J L. Expression of basic genes involved in tea polyphenol synthesis in relation to accumulation of catechins and total tea polyphenols. *Journal of the Science of Food and Agriculture*, 2006, 86 (3): 459-464
- [16] 周天山, 王新超, 余有本, 肖瑶, 钱文俊, 肖斌, 杨亚军. 紫芽茶树类黄酮生物合成关键酶基因表达与总儿茶素、花青素含量相关性分析. *作物学报*, 2016, 42 (4): 525-531
- [17] 周琼琼, 孙威江. 茶树芽叶紫化的生理生化分析及其关键酶基因的表达. *生物技术通报*, 2015, 31 (1): 86-91
- [18] 徐凌云, 王俊丽, 周宜君. 喜盐鸢尾花色形成关键基因的克隆及表达分析. *植物遗传资源学报*, 2017, 18 (2): 340-348
- [19] 梅菊芬, 徐德良, 汤茶琴, 卢媛, 徐琪. 茶树花青素及其种质资源的研究和利用进展. *热带农业工程*, 2013, 37 (1): 42-46
- [20] 徐熙, 任明见, 李鲁华, 杨喜翠, 徐如宏. 贵紫芽 1 号籽粒色素形成相关基因的差异表达. *中国农业科学*, 2018, 51 (2): 203-216
- [21] Liao L, Vimolmangkang S, Wei G C, Zhou H, Korban S S, Han Y P. Molecular characterization of genes encoding leucoanthocyanidin reductase involved in proanthocyanidin biosynthesis in apple. *Frontiers in Plant Science*, 2015, 10 (6): 243
- [22] Jiao Y, Ma R J, Shen Z J, Yan J, Yu M L. Gene regulation of anthocyanin biosynthesis in two blood-flesh peach (*Prunus persica* (L.) Batsch) cultivars during fruit development. *Journal of Zhejiang University Science B*, 2014, 15 (9): 809-819

- [23] 丁体玉,曹珂,方伟超,朱更瑞,陈昌文,王新卫,王力荣. 红肉桃两类花色苷积累模式与相关基因表达差异. 中国农业科学, 2017, 50(13): 2553-2563
- [24] 孙玉燕,段蒙蒙,邱杨,张晓辉,沈镛,王海平,李锡香. 心里美萝卜花青素合成酶基因 *RsANS* 克隆及花青素生物合成相关基因表达分析. 植物遗传资源学报, 2016, 17(5): 889-896
- [25] Kobayashi S, Ishimaru M, Ding C K, Yakushiji H, Goto N. Comparison of UDP-glucose: flavonoid 3-O-glucosyltransferases (UFGT) gene sequences between white grapes (*Vitis vinifera*) and their sports with red skin. Plant Science an International Journal of Experimental Plant Biology, 2001, 160(3): 543-550
- [26] Zhou X W, Fan Z Q, Chen Y, Zhu Y L, Li J Y, Yin H F. Functional analysis of a flavonol synthase-like gene from *Camellia nitidissima* reveals its roles in flavonoid metabolism during floral pigmentation. Journal of Biosciences, 2013, 38(3): 593-604
- [27] Rani A, Singh K, Ahuja P S, Kumar S. Molecular regulation of catechins biosynthesis in tea [*Camellia sinensis* (L.) O. Kuntze]. Gene, 2012, 495(2): 205-210
- [28] Pang Y Z, Abeyasinghe I S B, He J, He X Z, Huhman D, Mewan K M, Sumner L W, Yun J F, Dixon R A. Functional characterization of proanthocyanidin pathway enzymes from tea and their application for metabolic engineering. Plant Physiology, 2013, 161(3): 1103-1116
- [29] Wu Z J, Li X H, Liu Z W, Xu Z S, Zhuang J. *De novo* assembly and transcriptome characterization: novel insights into catechins biosynthesis in *Camellia sinensis*. BMC Plant Biology, 2014, 14(1): 277
- [30] 贾赵东,马佩勇,边小峰,杨清,郭小丁,谢一芝. 植物花青素合成代谢途径及其分子调控. 西北植物学报, 2014, 34(7): 1496-1506

## 欢迎订阅 2019 年《茶叶科学》征订启事

《茶叶科学》为中文核心期刊,中国科学引文及中国学术期刊综合评价数据库来源期刊,CA、CAB 等收录。主要报道最新涉茶科技成果,内容包括茶树栽培、育种、病虫害防治、茶叶加工、生化、机械、技术经济、茶饮料、茶食品和保健品、茶的医用保健等。

双月刊,每期定价 20 元,全年订价 120 元。征订方式:

(1) 邮汇地址:浙江省杭州市梅灵南路 9 号中国农业科学院茶叶研究所,收款人:《茶叶科学》编辑部,邮编:310008。

(2) 银行汇款:开户银行:中国农业银行杭州市西湖支行,帐号:19000101040005296,开户名:中国农业科学院茶叶研究所。银行汇款时请在附言中注明“茶叶科学”。

电话:0571-86651482, 86651902

E-mail:cykx@vip.163.com(订刊)

传真:0571-86651482

网址:www.tea-science.com

## 欢迎订阅 2019 年《河南农业科学》

《河南农业科学》是河南省农业科学院主办的综合性农业科技期刊。本刊连续被评为全国中文核心期刊、中国科技核心期刊、中国农业核心期刊,被遴选为中国科学引文数据库(CSCD)来源期刊、RCCSE 中国核心学术期刊。栏目设置有:综述、作物栽培·遗传育种、农业资源与环境、植物保护、园艺、畜牧·兽医、农产品加工与安全·农业工程·农业信息技术。

月刊,每期定价 18.00 元,全年 216 元。各地邮局均可订阅,邮发代号:36-32。如错过订期,可直接与本刊编辑部联系订阅。

地址:郑州市花园路 116 号

邮编:450002

电话:0371-65739041

E-mail:hnykx@163.com

网址:http://www.hnykx.org.cn