

# 基于转录组测序的滇山茶花叶呈色机理分析

王 炜<sup>1</sup>, 郑 伟<sup>1\*</sup>, 徐晓丹<sup>2</sup>, 陈 菁<sup>3</sup>, 王谈笑<sup>3</sup>

(1 昆明理工大学 建筑与城市规划学院, 昆明 650500; 2 昆明理工大学 艺术与传媒学院, 昆明 650500; 3 昆明理工大学 现代农业工程学院, 昆明 650500)

**摘 要:** 该研究采用 Illumina Hi-Seq2500 高通量测序技术, 对滇山茶的花叶(绿叶区和金叶区)的 cDNA 进行了转录组测序, 分析滇山茶的花叶呈色机理。测序结果得到了 228 862 条单基因序列, 筛选得到了 4 851 条差异表达基因。2 291 条差异表达基因在 GO 数据库中有功能注释, 1 826 条差异表达基因被注释到 COG 数据库, 这些差异表达基因被分为 23 个主要功能类别, 其中大部分基因与滇山茶的生长和代谢有关。1 363 条差异表达基因注释到了 KEGG 数据库, 进一步筛选出了差异表达基因注释序列最多的 15 条代谢通路, 其中卟啉和叶绿素代谢途径中的 GluRS、 $\delta$ -ALAD、ALAS 基因和类黄酮合成途径中 ANS、LAR、CHS 基因与滇山茶的花叶形成密切相关。研究表明, 滇山茶的花叶呈色是由于叶色相关代谢通路中大量的基因表达量发生变化引起的, 并不是由单基因突变所致, 病毒诱导的多个代谢途径基因沉默可能是滇山茶花叶呈色的主要原因。该研究丰富了滇山茶的转录组信息, 并为滇山茶的彩叶品种培育提供了一定的参考。

**关键词:** 滇山茶; 叶色; 转录组; 叶绿素代谢; 类黄酮生物合成

**中图分类号:** Q789 **文献标志码:** A

## Coloring Mechanism Analysis of Mosaic Leaves in *Camellia reticulata* Lindl. Based on Sequencing of Transcriptome

WANG Wei<sup>1</sup>, ZHENG Wei<sup>1\*</sup>, XU Xiaodan<sup>2</sup>, CHEN Jing<sup>3</sup>, WANG Tanxiao<sup>3</sup>

(1 Faculty of Architecture and City Planning, Kunming University of Science and Technology, Kunming 650500, China; 2 Faculty of Art and Communication, Kunming University of Science and Technology, Kunming 650500, China; 3 Faculty of Modern Agricultural Engineering, Kunming University of Science and Technology, Kunming 650500, China)

**Abstract:** In the present study, the cDNA of *C. reticulata* (floral leaf and green leaf) were sequenced based on Illumina Hi-Seq2500 to analyze the leaf color mechanism. In total, 228 862 unigenes were obtained and 4 851 differential expressed genes (DEGs) were selected. 2 291 DEGs were divided into 31 GO terms, 1 826 DEGs were divided into 23 COG major functional categories. Most of these genes are related to its growth and metabolism. 1 363 DEGs were annotated into KEGG database, 15 pathways of DEGs most annotated were screened out, GluRS,  $\delta$ -ALAD, ALAS, in the pathway of porphyrin and chlorophyll metabolism; ANS, LAR, CHS in the pathway of flavonoid biosynthesis, are the key genes in its floral leaf formation. The coloration of the leaves was caused by a large amount of gene expression changes related to leaf color metabolic pathways, not caused by single gene mutations, the silencing of homologous genes induced by *C. reticulata* mosaic virus in Yunnan is probably the main reason for the formation of *C. reticu-*

收稿日期: 2017-05-18; 修改稿收到日期: 2017-08-20

基金项目: 国家自然科学基金(31101565, 31360488)

作者简介: 王 炜(1993—), 女, 在读硕士研究生, 主要从事园林植物资源及利用研究。E-mail: 1071090370@qq.com

\* 通信作者: 郑 伟, 博士, 副教授, 主要从事园林植物资源及利用研究。E-mail: 57280234@qq.com

lata leaf coloration. These results greatly enriched genetic information of *C. reticulata*, and provide some reference for the cultivation of the colorful leaf varieties.

**Key words:** *Camellia reticulata*; leaf color; transcriptome; chlorophyll metabolism; flavonoid biosynthesis

彩叶植物因其具有色彩艳丽、观赏期长、抗逆性强、易于造景等特点在城市园林绿化中发挥着越来越重要的作用<sup>[1]</sup>,而花叶植物是其重要的组成部分。近年来,园林应用中的花叶植物日益增多,但对花叶呈色的机理研究还十分薄弱,不利于花叶园林植物的科学栽培与繁育。

滇山茶(*Camellia reticulata* Lindl.)是山茶科(Theaceae)常绿乔木<sup>[2]</sup>,是中国特有种、国家二级重点保护植物、云南省八大名花之首、昆明市市花,主要分布于滇中高原和滇西山地<sup>[3]</sup>。滇山茶至今已有1 500多年的栽培历史,花大色艳,是良好的风景园林植物,深受人们的喜爱。滇山茶的栽培品种约有120多个<sup>[4]</sup>,但彩叶品种仍十分少见。

调查发现,滇山茶的花叶植株近年来日益增多。通过形态学、组织细胞学的观测,初步认为滇山茶花叶形成,可能是一种球形的植物病毒诱导所致<sup>[5]</sup>,该病毒主要寄主为滇山茶,有别于近年来侵染山茶(*C. japonica*)和茶梅(*C. sasanqua*)并造成较大危害的杆状病毒——山茶花叶黄斑病毒(*Camellia leaf yellow mottle virus*, CLYMV)<sup>[6]</sup>。然而,这种花叶病毒暂时还没有得到明确的分子鉴定。因此,滇山茶在分子水平的花叶呈色机理研究就显得尤为重要,一方面可对花叶形成的原因进行探索,另一方面可对花叶性状改良提供分子基础。

高通量转录组测序是衔接基因组遗传信息与生物功能蛋白质组的重要纽带<sup>[7]</sup>,该技术可对一组相对性状测序后获得基因表达的差异情况,进行差异基因筛选分析,进而从基因表达量、代谢途径等层面解释相对性状的分子机理。为此,本研究将通过高通量测序技术获得滇山茶的花叶及绿叶转录组信息,对得到的 Unigene 进行功能注释和功能分类,进而分析与滇山茶叶片呈色相关的功能基因、代谢途径及关键基因,探明滇山茶的花叶呈色分子机理,探索花叶形成的原因,为滇山茶的叶色改良研究奠定一定基础。

## 1 材料和方法

### 1.1 试验材料

供试验的滇山茶花叶与绿叶(图 1)均采集于云南省昆明市金殿风景区滇山茶园,来自同一植株,

采集后立即用液氮速冻,后置于固体干冰中带回实验室,保存在-70℃超低温冰箱中备用。

### 1.2 方法

**1.2.1 高通量测序及数据组装** 采用 Trizol 法提取绿叶和花叶滇山茶样品的 RNA。分别采用 Nanodrop、Qubit 2.0、Aglient 2100 方法检测 RNA 样品的纯度、浓度及完整性后,构建 cDNA 文库,文库检验合格后,用 Hi-Sep2500 进行高通量测序,测序读长为 PE125。获得初始数据后,进行数据过滤,截除 Reads 中的测序接头及引物序列并过滤低质量数据后得到高质量 Reads(Clean Data),用 Trinity 对 Clean Data 进行合并组装。首先将测序 Reads 打断为较短的片段(K-mer),然后将这些小片段延伸成较长的片段(Contig),并利用这些片段之间的重叠,得到片段集合(Component),最后利用 De Bruijn 图的方法和测序 Read 信息,在各个片段集合中分别识别转录本序列。

**1.2.2 基因表达量分析及差异表达基因筛选** 采用 Bowtie<sup>[8]</sup>将绿叶和花叶的测序数据对比到 Unigene 库,得到 Mapped reads,根据比对结果,结合 RSEM 估算表达量的分析,利用 FPKM 值表示对应的 Unigene 表达丰度。使用 EBSep<sup>[9]</sup>进行差异表达分析,采用 Benjamini-Hochberg 方法对原有假设检验得到的限制性 P 值进行校正,最终采用校正后的 P 值,即 FDR(False discovery rate)作为差异表达基因筛选的关键指标。在筛选过程中,将  $FDR < 0.01$  且差异倍数  $FC(\text{Fold change}) \geq 2$  作为筛选标准,其中,FC 表示两样品(组)间 FPKM 值的比值。

**1.2.3 差异表达基因的功能注释和富集分析** 使



图 1 滇山茶转录组检测样品

Fig.1 Transcriptome detection samples

用 Blast 软件将差异表达基因的 Unigene 序列与 NR、Swiss-prot、Pfam、KEGG、KOG 和 COG、GO 数据库进行功能注释,用 topGO 软件对注释到 GO 数据库的样品组间差异表达基因进行富集分析。为进一步系统分析基因产物在细胞中的代谢途径以及这些基因产物的功能,将差异表达基因对比到 KEGG 数据库,得到差异表达基因 Pathway,结合富集通路筛选出滇山茶叶色合成的关键差异表达基因。

## 2 结果与分析

### 2.1 测序数据组装及统计

通过高通量测序及质量控制后,共得到 13.12 Gb 有效数据,花叶与绿叶的碱基数据总数分别为 6 965 128 328 和 6 155 169 906。各样品 Q30 碱基百分比不低于 86.84%(表 1)。表明测序数据可靠,可用于后续分析。

对组装结果进行统计,得到序列重叠群(Contig)、转录本序列(Transcript)和单基因序列(Unigene)长度分布表(表 2),组装结果中除了滇山茶的转录本/单基因序列外,还包括了病毒等其他物种的转录本/单基因序列,其中,疑似病毒的序列有 1 985

条,组装得到 276 个重叠群。在序列重叠群中,所占比例最大的是 0~300 bp,有 9 710 940 条,占 99.1%;长度大于 2 000 bp 的有 3 846 条,占 0.04%,平均长度 56.06 bp,N50 长 50 bp。由序列重叠群组装出的 331 311 条转录本序列,平均长度 561.12 bp,N50 长度为 888。由转录本序列组装出的 228 862 条单基因序列,平均长 400.61 bp,N50 长 412 bp。单基因长度主要分布在 200~1 000 bp,占 93.78%。

### 2.2 功能注释

将单基因序列比对到 NR、COG、GO、KEGG 等数据库,得到其注释信息(表 3),结果显示,在 228 862 条单基因序列中,共有 137 455 条单基因序列得到注释,占单基因序列总数的 60.06%,另外有 91 407 条未被注释,认为可能是新基因。单基因序列与 COG 数据库比对,得到 52 312 条同源序列,占单基因序列总数的 22.86%;与 KOG 数据库比对,得到 74 110 条同源序列,占单基因序列总数的 32.38%;与 Swiss-Prot 数据库比对,得到 62 408 条同源序列,占单基因序列总数的 27.27%;与 NR 数据库比对,得到 134 288 条同源序列,占单基因序列总数的 58.68%。

表 1 滇山茶有效数据评估统计  
Table 1 Statistics of clean data

| 样品 Sample        | 编号 Number | 读数 Read    | 碱基 Base       | GC 含量 GC content/% | Q30/% |
|------------------|-----------|------------|---------------|--------------------|-------|
| 花叶 Mosaic leaves | T01       | 27 659 427 | 6 965 128 328 | 47.03              | 87.67 |
| 绿叶 Green leaves  | T02       | 24 439 148 | 6 155 169 906 | 47.27              | 86.84 |

表 2 组装结果统计表  
Table 2 Statistics of assembled results

| 长度区间<br>Length range/bp | 序列重叠群数量<br>Contig number | 转录本序列数量<br>Transcript number | 单基因序列数量<br>Unigene number |
|-------------------------|--------------------------|------------------------------|---------------------------|
| 200~300                 | 9 710 940(99.1%)*        | 162 518(49.05%)              | 142 087(62.08%)           |
| 300~500                 | 56 446(0.58%)            | 70 764(21.36%)               | 49 897(21.80%)            |
| 500~1 000               | 25 709(0.26%)            | 50 039(15.10%)               | 22 648(9.90%)             |
| 1 000~2 000             | 11 409(0.12%)            | 34 535(10.42%)               | 10 667(4.66%)             |
| 2 000+                  | 3 846(0.04%)             | 13 455(4.06%)                | 3 563(1.56%)              |
| 总数 Total number         | 9 808 350                | 331 311                      | 228 862                   |
| 总长度 Total length/bp     | 549 866 640              | 185 905 676                  | 91 685 505                |
| N50 长度 N50 length/bp    | 50                       | 888                          | 412                       |
| 平均长度 Mean length        | 56.06                    | 561.12                       | 400.61                    |

注:\*,表中统计的是长度在区间 0~300 内 Contig 的数量和百分比  
Note: With \* in the table, the number and percentage of Contig in the range 0 to 300 are counted

表 3 单基因序列注释统计

Table 3 Unigene functions annotated

| 数据库<br>Database                     | 注释单基因库<br>Unigene annotated | 百分比<br>Percentage/% |
|-------------------------------------|-----------------------------|---------------------|
| COG                                 | 52 312                      | 22. 86              |
| GO                                  | 82 864                      | 36. 21              |
| KEGG                                | 47 326                      | 20. 68              |
| KOG                                 | 74 110                      | 32. 38              |
| Pfam                                | 73 226                      | 34. 00              |
| Swiss-Prot                          | 62 408                      | 27. 27              |
| NR                                  | 134 288                     | 58. 68              |
| 总注释单基因序列数<br>All unigenes annotated | 137 445                     | 60. 06              |

2.3 差异表达基因的筛选与功能注释

**2.3.1 差异表达基因的筛选** 通过对差异表达基因的筛选,最终得到花叶相对绿叶的差异表达基因为 4 851 个,其中表达量上调的基因为 2 360 个,表达量下调的基因为 2 491 个。将 DEGs 单基因序列注释到 NR、Swiss-Prot、Pfam、GO、COG、KOG、KEGG 数据库,共 4 059 条单基因数据得到注释,被 GO 数据库注释到的为 2 291 条,COG 数据库注释到的为 1 826 条,KEGG 数据库注释到的为 1 302 条。

**2.3.2 差异表达基因的 GO 功能注释** 滇山茶花叶及绿叶样品间差异表达基因的 GO 分类统计显示,差异表达基因主要被分为 31 个功能组(表 4),在

表 4 差异表达基因 GO 功能注释

Table 4 DEGs GO annotation

| 项目 Item                      |                                                               | 单基因序列数量<br>Number of unigene |
|------------------------------|---------------------------------------------------------------|------------------------------|
| 细胞组分<br>Cellular component   | 细胞成分 Cell part                                                | 332                          |
|                              | 细胞 Cell                                                       | 329                          |
|                              | 细胞器 Organelle                                                 | 252                          |
|                              | 膜 Membrane                                                    | 143                          |
|                              | 大分子复合物 Macromolecular complex                                 | 140                          |
|                              | 细胞器成分 Organelle part                                          | 112                          |
|                              | 膜成分 Membrane part                                             | 75                           |
|                              | 细胞连接 Cell junction                                            | 18                           |
|                              | 胞外区域 Extracellular region                                     | 15                           |
|                              | 膜封闭腔 Membrane-enclosed lumen                                  | 9                            |
| 分子功能水平<br>Molecular function | 催化活性 Catalytic activity                                       | 283                          |
|                              | 整合 Binding                                                    | 252                          |
|                              | 结构分子活性 Structural molecule activity                           | 69                           |
|                              | 转运活性 Transporter activity                                     | 44                           |
|                              | 核酸结合转录因子活性 Nucleic acid binding transcription factor activity | 13                           |
|                              | 电子载体活性 Electron carrier activity                              | 13                           |
|                              | 酶调节活性 Enzyme regulator activity                               | 7                            |
| 生物学过程<br>Biological process  | 代谢过程 Metabolic process                                        | 409                          |
|                              | 细胞过程 Cellular process                                         | 331                          |
|                              | 单一生物过程 Single-organism process                                | 281                          |
|                              | 刺激应答 Response to stimulus                                     | 97                           |
|                              | 定位 Localization                                               | 87                           |
|                              | 生物调控 Biological regulation                                    | 82                           |
|                              | 细胞成分或生物合成 Cellular component organization or biogenesis       | 57                           |
|                              | 发育过程 Developmental process                                    | 52                           |
|                              | 多细胞生物过程 Multicellular organismal process                      | 25                           |
|                              | 信号传导 Signaling                                                | 20                           |
|                              | 繁殖过程 Reproductive process                                     | 17                           |
|                              | 多生物过程 Multi-organism process                                  | 15                           |
|                              | 生长 Growth                                                     | 7                            |
|                              | 繁殖 Reproduction                                               | 6                            |

表 5 差异表达基因 COG 功能注释  
Table 5 DEGs COG annotation

| COG | 功能分类 Function classification                                                 | 单基因序列数量<br>Number of unigene |
|-----|------------------------------------------------------------------------------|------------------------------|
| A   | RNA 的加工和修饰 RNA processing and modification                                   | 5                            |
| B   | 染色体结构和动态变化 Chromatin structure and dynamics                                  | 4                            |
| C   | 能源的产生和转化 Energy production and conversion                                    | 205                          |
| D   | 细胞周期调控,细胞分裂,染色体分离 Cell cycle control, cell division, chromosome partitioning | 21                           |
| E   | 氨基酸转运与代谢 Amino acid transport and metabolism                                 | 158                          |
| F   | 核苷酸转运和代谢 Nucleotide transport and metabolism                                 | 20                           |
| G   | 碳水化合物转运和代谢 Carbohydrate transport and metabolism                             | 205                          |
| H   | 辅酶转运与代谢 Coenzyme transport and metabolism                                    | 40                           |
| I   | 脂质转运与代谢 Lipid transport and metabolism                                       | 111                          |
| J   | 翻译、核糖体结构和生物合成 Translation, ribosomal structure and biogenesis                | 315                          |
| K   | 转录 Transcription                                                             | 98                           |
| L   | 复制,重组和修复 Replication, recombination and repair                               | 92                           |
| M   | 细胞壁/细胞膜的生物发生 Cell wall/membrane/envelope biogenesis                          | 44                           |
| N   | 细胞运动 Cell motility                                                           | 2                            |
| O   | 翻译后修饰,蛋白质翻转,伴侣 Posttranslational modification, protein turnover, chaperones  | 178                          |
| P   | 无机离子转运与代谢 Inorganic ion transport and metabolism                             | 147                          |
| Q   | 次生代谢产物的合成、转运和代谢 Secondary metabolites biosynthesis, transport and catabolism | 139                          |
| R   | 一般功能预测 General function prediction only                                      | 485                          |
| S   | 未知功能 Function unknown                                                        | 103                          |
| T   | 信号转导机制 Signal transduction mechanisms                                        | 84                           |
| U   | 胞内运输、分泌和囊泡运输 Intracellular trafficking, secretion, and vesicular transport   | 17                           |
| V   | 防御机制 Defense mechanisms                                                      | 21                           |
| Z   | 细胞骨架 Cytoskeleton                                                            | 19                           |

细胞组分、分子功能水平和生物学过程三大类中分别主要包含了 10、7 和 14 个功能小类。在细胞组分这一类中,差异表达基因在细胞成分、细胞及细胞器中所占比例最高;在分子功能水平这一类中,差异表达基因在催化活性、整合中所占比例最高;在生物学过程这一类中,差异表达基因在代谢过程、细胞过程及单一有机体过程中所占比例最高。

**2.3.3 差异表达基因的 COG 功能注释** 将注释到 COG 数据库的 1 826 条差异表达基因进行直系同源分类,获得 23 个功能类别(表 5),分别与代谢、翻译、转录等相关。其中表达数目最多的基因类别是 R(一般功能预测),其次分别是 J(翻译、核糖体结构和生物合成)、C(能源产生与转换)、G(碳水化合物转运和代谢)和 O(翻译后修饰、蛋白翻转、伴侣基因)。

**2.3.4 差异表达基因的 KEGG 功能注释** 将差异表达基因通过与 KEGG 数据库比对,有 1 363 条基

因得到注释,富集到 131 条代谢通路,包括合成代谢通路、花青素代谢途径、植物病原互作通路等,差异表达基因注释序列最多的 15 条代谢通路为:核糖体(Ribosome)、丙酮酸代谢(Pyruvate metabolism)、RNA 转运(RNA transport)和光合作用(Photosynthesis)等(表 6)。

从以上结果可以看出:在 GO 注释中生物学过程这一类别内,代谢过程所占比例较高;COG 注释中差异表达基因大量富集在 R(一般功能预测)和 J(翻译、核糖体结构和生物合成)功能类。通过转录组的代谢途径分析,本研究选择与叶色相关的代谢途径且差异表达基因注释序列较多的途径作为切入点进行分析。结合 KEGG 通路分析结果,最终确定从卟啉和叶绿素代谢途径、类黄酮生物合成途径两条和叶色直接相关的途径来寻找滇山茶花叶中发生差异表达的关键基因。

表 6 差异表达基因 KEGG 功能注释

Table 6 Differential expression genes KEGG function annotation

| 编号      | KEGG 功能注释 KEGG function annotation                     | 单基因序列数量<br>Number of unigene | 比例<br>Percentage/% |
|---------|--------------------------------------------------------|------------------------------|--------------------|
| ko03010 | 核糖体 Ribosome                                           | 125                          | 9.17               |
| ko00010 | 糖酵解途径 Glycolysis / Gluconeogenesis                     | 73                           | 5.36               |
| ko00190 | 氧化磷酸化 Oxidative phosphorylation                        | 51                           | 3.74               |
| ko04141 | 内质网中的蛋白质加工 Protein processing in endoplasmic reticulum | 41                           | 3.01               |
| ko00620 | 丙酮酸代谢 Pyruvate metabolism                              | 41                           | 3.01               |
| ko03013 | RNA 转运 RNA transport                                   | 40                           | 2.93               |
| ko04145 | 吞噬体 Phagosome                                          | 33                           | 2.42               |
| ko03040 | 剪接体 Spliceosome                                        | 30                           | 2.20               |
| ko00680 | 甲烷代谢 Methane metabolism                                | 30                           | 2.20               |
| ko04146 | 过氧化物酶体 Peroxisome                                      | 25                           | 1.83               |
| ko00020 | 柠檬酸循环 Citrate cycle                                    | 24                           | 1.76               |
| ko00500 | 淀粉与蔗糖代谢 Starch and sucrose metabolism                  | 24                           | 1.76               |
| ko00630 | 乙醛酸盐代谢 Glyoxylate and dicarboxylate metabolism         | 24                           | 1.76               |
| ko00071 | 脂肪酸代谢 Fatty acid metabolism                            | 23                           | 1.69               |
| ko00195 | 光合作用 Photosynthesis                                    | 22                           | 1.61               |

表 7 卟琳和叶绿素代谢途径的差异表达基因信息

Table 7 Information of differential expression gene of porphyrin and chlorophyll metabolism

| 酶 Enzyme | 基因名称 Gene name                                                 | 基因表达<br>Gene expression | 基因 ID<br>Gene ID                    |
|----------|----------------------------------------------------------------|-------------------------|-------------------------------------|
| 4.4.1.17 | 细胞色素 c 血红素裂和合酶 Cytochrome c heme-lyase (HCCS)                  | 下调 Down                 | c101920.graph_c0                    |
| 4.2.1.24 | δ 氨基乙酰丙酸脱水酶 Delta-aminolevulinic acid dehydratase (δ-ALAD)     | 下调 Down                 | c102122.graph_c0                    |
| 6.1.1.17 | 谷氨酰 tRNA 合成酶 Glutamyl-tRNA synthetase (GluRS)                  | 下调 Down                 | c108648.graph_c0                    |
| 1.3.7.4  | 铁氧还蛋白氧化还原酶 Ferredoxin oxidoreductase (FOR)                     | 上调 Up                   | c113343.graph_c0                    |
| 1.3.3.3  | UDP 糖基转移酶 UDP-glycosyltransferase (UDP-Gly)                    | 上调 Up                   | c126922.graph_c0                    |
| 1.2.1.70 | 谷氨酰-tRNA 还原酶家族蛋白 Glutamyl-tRNA reductase family protein (HemA) | 上调 Up                   | c126978.graph_c0                    |
| 2.3.1.37 | 5-氨基酮戊酸合成酶 5-aminolevulinate synthase, mitochondrial (ALAS)    | 下调 Down                 | c36002.graph_c0                     |
| 4.99.1.1 | 亚铁螯合酶 Ferrochelatase (FECH)                                    | 下调 Down                 | c94567.graph_c0<br>c106937.graph_c0 |

2.4 叶色相关途径及关键基因分析

**2.4.1 卟琳和叶绿素代谢途径关键基因分析** 滇山茶卟琳和叶绿素代谢途径中,谷氨酸-tRNA 合成酶(GluRS)、δ 氨基乙酰丙酸脱水酶(δ-ALAD)的基因发生下调,导致 5-氨基乙酰丙酸(ALA)合成量减少,而 ALA 的合成与转换是叶绿素合成的主要控制点<sup>[10]</sup>;5-氨基酮戊酸合成酶(ALAS)的基因发生下调,导致胆色素原及原卟琳 IX 合成量减少,以上是叶绿素合成的第一大步。在第二大步中,原卟琳 IX 合成量在第一大步合成中大量减少,从而使得镁原卟琳 IX 的合成量减少,抑制了叶绿素 a 的合成;亚铁螯合酶(FECH)基因发生下调,导致原卟琳 IX

向正铁血红素的转化途径受阻,使得血红素和光敏色素的合成减少,进而抑制了叶绿素的合成(表 7)。

**2.4.2 类黄酮生物合成途径关键基因分析** 滇山茶类黄酮生物合成途径中,查尔酮合成酶(CHS)的基因表达量上调,使得柚皮素、黄酮醇生物合成量增多,但花青素合成酶(ANS)的基因表达量发生下调,使得花青素含量减少,无色花青素含量增加;且无色花青素还原酶(LAR)、黄酮醇合成酶(FLS)的基因表达量发生上升,从而使得儿茶素及表儿茶素合成量增多,儿茶素及表儿茶素结合生成无色的原花青素,因此细胞内原花青素的含量增多,使叶色产生变化(表 8)。

表 8 类黄酮生物合成途径的差异表达基因信息  
Table 8 Information of differential expression gene of flavonoid biosynthesis

| 酶 Enzyme      | 基因名称 Gene name                              | 基因表达 Gene expression | 基因 ID Gene ID                         |
|---------------|---------------------------------------------|----------------------|---------------------------------------|
| 2. 3. 1. 74   | 查尔酮合成酶 Chalcone synthase (CHS)              | 上调 Up                | c128727. graph_c0                     |
| 1. 14. 11. 23 | 黄酮醇合成酶 Flavonol synthase (FLS)              | 上调 Up                | c102601. graph_c0                     |
| 1. 17. 1. 3   | 无色花青素还原酶 Leucoanthocyanidin reductase (LAR) | 上调 Up                | c123030. graph_c0                     |
| 1. 14. 11. 19 | 花青素合成酶 Anthocyanidin synthase (ANS)         | 下调 Down              | c114104. graph_c0<br>c80746. graph_c0 |

3 讨 论

3.1 滇山茶的花叶形成原因探讨

植物叶片细胞中的色素包括叶绿素、类胡萝卜素和花青素,这几种色素在细胞中的含量及分布决定了植物叶片的颜色<sup>[11]</sup>。一般来讲,花叶、金叶的产生主要是由于叶绿素含量降低导致<sup>[12-14]</sup>。叶绿素的合成从谷氨酸到叶绿素的合成与循环,经过 19 步、18 个酶和对应的 29 个基因参与<sup>[15]</sup>。其中任意一个酶基因表达量的上调或下调都可能使相关酶的活力发生改变,影响蛋白质的合成,从而导致叶绿素合成过程受阻,使叶色表现出不同程度的变化。本研究中,滇山茶叶绿素代谢途径中多个酶基因表达量发生变化,导致叶绿素合成受阻,使得滇山茶叶色发生变化。

除此之外,叶绿体发育异常会导致叶绿素降解<sup>[16]</sup>,而叶绿体的发育是一个非常复杂的过程,受多种因素影响,叶绿素合成途径受阻是影响叶绿体发育的一个重要因素,它将导致光合器官中积累的叶绿素含量降低,不利于叶绿体的发育<sup>[17]</sup>。光敏色素对叶绿素合成中基因表达调控起着重要的作用,研究表明,在拟南芥中,光敏色素调控了叶绿素合成途径中一些关键基因的表达,还影响了叶绿体的发育<sup>[18]</sup>。滇山茶叶绿素代谢途径中,FECH 基因发生下调,使得血红素和光敏色素的合成减少。叶绿体代谢状况以代谢物为信号调控核基因的表达,镁原卟啉 IX 是影响叶绿体核基因表达的主要卟啉类物质,镁原卟啉 IX 的积累为叶绿体调控核基因的表达所需<sup>[19]</sup>。滇山茶叶绿素代谢途径中,ALAS 基因表达量发生下调,阻碍了原卟啉 IX 的合成,导致镁原卟啉 IX 的合成减少,从而影响了叶绿体核基因的表达,抑制了叶绿体的发育。且通过对花叶滇山茶金叶区和绿叶区的色素含量测定发现,绿叶区中叶绿素的含量为 0.475 7 mg/g,金叶区中叶绿素的含量仅为 0.01 mg/g,下降了 47.57 倍。

植物花叶发生的主要方式有自然突变、T-DNA

插入突变、甲基磺酸乙酯(EMS)诱变和病毒诱导等。其中,病毒诱导的花叶是病毒入侵和寄主防御形成的一种平衡关系。一些观赏价值较高的花叶植物正是由于病毒诱导所致<sup>[20]</sup>,如‘洒金’桃叶珊瑚(*Aucuba japonica*)、‘花叶’扶桑(*Hibiscus rosa-sinensis*)等<sup>[21-22]</sup>。本研究中,在疑似病毒的 276 个 contigs 中,有 49 个 contigs 与西瓜银斑病毒(WSMV)序列相似度极高,有 16 个 contigs 与番茄褪绿病毒(ToCV)序列相似,有 15 个 contigs 与南方番茄病毒(STV)序列相似度较高,这就说明滇山茶的花叶形状有可能是病毒入侵引起的。

病毒入侵寄主后,寄主的内在反应为:病毒诱导的基因沉默(virus induced gene silencing, VIGS),具有同源性依赖性<sup>[23]</sup>,包括转录基因沉默(TGS)和转录后基因沉默(PTGS),这是病毒诱导的花叶中细胞形态和生理变化的直接原因。所以可以推测,滇山茶花叶的形成是由于病毒入侵叶片后,使控制叶色色素合成的相关同源性基因沉默,使得叶绿素无法正常合成,色素含量发生变化,从而形成花叶。

综上,滇山茶的花叶形成是由于与叶色相关代谢通路中大量的基因表达量发生变化引起的,并由单基因突变所致,病毒诱导的多个代谢途径基因沉默可能是滇山茶花叶呈色的主要原因,但病毒的种类还有待进一步的分子鉴定。

3.2 滇山茶的转录组测序丰富了滇山茶的基因序列信息

滇山茶转录组测序共获得 228 862 条单基因序列信息,与公共数据库进行对比,获得注释信息的单基因序列共有 137 455 条,占 60.06%,在一定程度上丰富了滇山茶的序列资源,也为后续研究提供了基础数据。例如,通过基于转录组的 SSR 分析,发现包含 SSR 的序列共 14 230 条,发生频率为 10.35%,共筛选到 7 187 个滇山茶 SSR 位点。由于 SSR 标记具有高度重复性、共显性、操作技术简单性等优点,已被广泛用于遗传多样性分析和分子标记辅助育种等方面<sup>[10]</sup>,因此,下一步若需要利用

SSR 标记对滇山茶展开深入研究,可利用基于转录组的 SSR 位点信息,为 SSR 引物设计提供参考。

此外,仍有 40% 的单基因序列没有得到注释,造成这种现象的原因是滇山茶生物学信息研究的空白以及数据库信息的不完整,有待进一步丰富数据

参考文献:

[1] 洪 丽,王金刚,龚束芳. 彩叶植物叶色变化及其相关影响因素研究进展[J]. 东北农业大学学报,2010,**41**(6):152-156.  
HONG L, WANG J G, GONG S F. Progress in color changes and its influencing factors of color leaf plants[J]. *Journal of Northeast Agriculture University*, 2010,**41**(6): 152-156.

[2] 薛英利,赵庆师,黄由安,等. 滇山茶的花色类型与细胞内环境关系初探[J]. 云南农业大学学报,2015,**30**(3):455-463.  
XUE Y L, ZHAO Q S, HUANG Y A, *et al.* Research on relationship between floral colors and intracellular environment of *Camellia reticulata* Lindl[J]. *Journal of Yunnan Agricultural University*, 2015,**30**(3): 455-463.

[3] 杨桂英,何 瀚,曹子林,等. 三种不同倍性滇山茶的耐寒性研究[J]. 云南大学学报(自然科学版),2016,**38**(2):335-343.  
YANG G Y, HE H, CAO Z L, *et al.* A study on cold resistance in diploid, tetraploid and hexaploid of *Camellia reticulata* (Theaceae)[J]. *Journal of Yunnan University (Natural Sciences Edition)*, 2016,**38**(2): 335-343.

[4] 昆明市科协. 滇之奇葩[M]. 昆明:云南科学技术出版社,2010.

[5] 常 青. 滇山茶的叶片特性研究[M]. 昆明:昆明理工大学,2014:35-40.

[6] GAILHOFER M, THALER I, MILICIC D. Occurrence of *Camellia leaf yellow mottle virus* (CLYMV) on east adriatic coast[J]. *Acta Horticulturae*, 1988, 234: 385-392.

[7] 王兴春,杨致荣,王 敏,等. 高通量测序技术及其应用[J]. 中国生物工程杂志,2012,**32**(1):109-114.  
WANG X C, YANG Z R, WANG M, *et al.* High-throughput sequencing technology and its application[J]. *China Biotechnology*, 2012,**32**(1): 109-114.

[8] LANGMEAD B, TRAPNELL C, *et al.* Ultrafast and memory-efficient alignment of short DNA sequences to the human genome[J]. *Genome Biology*, 2009,**10**(3): 1-10.

[9] GROTEWOLD E. The genetics and biochemistry of floral pigments[J]. *Annual Review of Plant Biology*, 2006,**57**(1): 761-780.

[10] 马春雷,姚明哲,王新超,等. 茶树叶绿素合成相关基因克隆及在白叶 1 号不同白化阶段的表达分析[J]. 作物学报,2015,**41**(2):240-250.  
MA C L, YAO M Z, WANG X C, *et al.* Cloning and expression of three genes involved in chlorophyll biosynthesis at different albescent stages of tea plant variety “Baiye 1”[J]. *Acta Agronomica Sinica*, 2015,**41**(2): 240-250.

[11] BALL P. Nature’s palette: the science of plant color[J]. *Nature*, 2007, 449: 982-983.

库信息。

本研究为滇山茶花叶呈色调控基因的挖掘、克隆及功能验证等后续研究积累了资料,为最终实现滇山茶彩叶品种的培育提供了基础研究数据。

[12] LI S J, DENG X M, MAO H Z, *et al.* Enhanced anthocyanin synthesis in foliage plant *Caladium bicolor*[J]. *Plant Cell Reports*, 2005,**23**(10): 716-720.

[13] GUO D P, GUO Y P, ZHAO J P, *et al.* Photosynthetic rate and chlorophyll fluorescence in leaves of stem mustard (*Brassica juncea* var. *tsatsai*) after turnip mosaic virus infection[J]. *Plant Science*, 2005, **168**(1): 57-63.

[14] 胡海姿,张 睿,尚爱芹,等. 金叶植物色素含量对光强的响应[J]. 园艺学报,2007,**34**(3):717-722.  
HU H Z, ZHANG R, SHANG A Q, *et al.* Response of pigment content of golden leaf plants of light intensity[J]. *Acta Horticulturae Sinica*, 2007,**34**(3): 717-722.

[15] WINKEL SB, SHIRLEY B W, *et al.* Flavonoid biosynthesis: A colorful model for genetics, biochemistry, cell biology, and biotechnolology[J]. *Plant Physiology*, 2001,**126** (2): 485-493.

[16] ZHOU S, SAWICKI A, WILLOWS R D, *et al.* C-terminal residues of *Oryza sativa* GUN4 are required for the activation of the *ChlH* subunit of magnesium chelatase in chlorophyll synthesis[J]. *Febs Letters*, 2012,**586**(3): 205-210.

[17] 邹国兴. 三个水稻叶色突变体的基因克隆与功能分析[D]. 北京:中国农业科学院,2010:36-39.

[18] STEPHENSON P G, CHRISTIAN F, MATTHEW J T. PIF3 is repressor of chloroplast development[J]. *Proc Natl Acad Sci USA*, 2009, 105: 7 654-7 659.

[19] 李保珠,赵孝亮,彭 雷. 植物叶绿体发育及调控研究进展[J]. 植物学报,2014,**49**(3):337-345.  
LI B Z, ZHAO X L, PENG L. Research advances in the development and regulation of plant chloroplasts[J]. *Bulletin of Botany*, 2014,**49**(3):337-345.

[20] HULL R. Matthews’ Plant Virology[M]. Beijing: Science Press, 2007.

[21] 倪德祥,蔡同润,张丕方,等. 光质对花叶芋组织培养形态发生的影响[J]. 园艺学报,1987,**14**(4):271-273.  
NI D X, CAI T R, ZHANG P F, *at el.* The effect of light quality on morphogenesis of *caladium bicolor in vitro* [J]. *Acta Horticulturae Sinica*, 1987,**14**(4): 271-273.

[22] 侯元凯. 彩叶植物研究进展[J]. 世界林业研究,2010,**6**(23): 24-28.  
HOU Y K. Study advances in color-leaf plants[J]. *World Forestry Research*, 2010,**6**(23): 24-28.

[23] BAULCOMBE D. RNA silencing in plants[J]. *Nature*, 2004, 431: 356-363.