

DOI: 10.11931/guihaia.gxzw202109057

陈博雯, 李开祥, 曾祥艳, 等, 2023. 基于转录组的八角挥发油合成相关基因挖掘与分析 [J]. 广西植物, 43(2): 303–314.
CHEN BW, LI KX, ZENG XY, et al., 2023. Discovery and analysis of volatile oil synthesis related genes in *Illicium verum* based on transcriptome sequencing [J]. *Guihaia*, 43(2): 303–314.



基于转录组的八角挥发油合成相关基因挖掘与分析

陈博雯, 李开祥*, 曾祥艳, 黄开顺, 梁文汇, 陈迎迎, 杨卓颖

(广西壮族自治区林业科学研究院, 广西特色经济林培育与利用重点实验室, 国家林业和草原局八角肉桂工程技术研究中心, 广西木本香料工程技术研究中心, 南宁 530002)

摘要: 为更好地挖掘八角 (*Illicium verum*) 挥发油合成相关基因, 该文对挥发油性状差异显著的优良无性系桂角 69 号及普通品种砧 01 号叶片进行了转录组测序及组装注释, 并对差异表达基因进行了 GO 分类和 KEGG 通路分析。结果表明: (1) 转录本经组装后获得 84 182 条序列, 使用 NR、NT、Swiss-Prot、KEGG、KOG、GO 和 Pfam 数据库进行序列比对, 共注释了 59 161 条序列, 筛选出 30 572 个差异表达基因。与砧 01 号相比, 桂角 69 号叶片中上调基因有 15 025 个, 下调基因有 15 547 个。(2) GO 分类结果显示共有 20 287 个差异基因被注释。KEGG 分析结果表明, 有 21 600 个差异基因被注释到 133 条 KEGG 通路上, 其中挥发油合成相关的单萜生物合成通路、萜类骨架生物合成通路、苯丙素合成通路中的芳樟醇合酶、月桂烯合酶、香叶基香叶基焦磷酸合酶、肉桂酰辅酶 A 还原酶、咖啡酸 3-O-甲基转移酶、肉桂醇脱氢酶等关键酶基因呈差异表达。(3) 转录因子分析发现差异表达基因分布于 31 个转录因子家族, 其中 MYB 家族序列数量最多。该文利用转录组测序技术分析八角优良无性系与普通品种叶片的差异基因及其相关功能和代谢通路, 获得的候选基因为深入探究挥发油特征组分的合成机制以及八角的分子育种提供了参考。

关键词: 八角, 叶片, 挥发油合成, 草蒿脑, 转录组分析

中图分类号: Q943 文献标识码: A 文章编号: 1000-3142(2023)02-0303-12

Discovery and analysis of volatile oil synthesis related genes in *Illicium verum* based on transcriptome sequencing

CHEN Bowen, LI Kaixiang*, ZENG Xiangyan, HUANG Kaishun,
LIANG Wenhui, CHEN Yingying, YANG Zhuoying

(Guangxi Forestry Research Institute, Guangxi Key Laboratory of Characteristic Non-wood Forest Cultivation & Utilization, National Forestry and Grassland Administration Engineering Technology Research Center of Anise & Cinnamon, Guangxi Engineering Technology Research Center of Woody Spices, Nanning 530002, China)

Abstract: In order to identify volatile oil synthesis related genes in *Illicium verum*, transcriptome sequencing and assembly annotation were carried out on the leaves of the superior clone Guijiao 69 and common variety Zhen 01, two

收稿日期: 2022-02-28

基金项目: 广西重点研发计划项目(桂科 AB18221040); 广西博士后创新支持计划项目(桂科财字[2020]21号); 广西特色经济林培育与利用重点实验室课题(19-A-01-04)。

第一作者: 陈博雯(1983-), 博士, 高级工程师, 主要从事林木生物技术研究, (E-mail) gfri_bwchen@163.com。

*通信作者: 李开祥, 博士, 教授级高级工程师, 主要从事经济林培育研究, (E-mail) lkx202@126.com。

varieties with significant difference in volatile oil content. Then GO and KEGG pathways analyses of differentially expressed genes (DEGs) were performed. The results were as follows: (1) A total of 84 182 Unigenes were obtained after transcripts assembly, and 59 161 Unigenes were annotated in NR, NT, Swissprot, KEGG, KOG, GO, and Pfam databases. A total of 30 572 DEGs were obtained after filtering the low abundance genes, and 15 025 up-regulated and 15 547 down-regulated genes were identified in Guijiao 69 compared to the common variety Zhen 01. (2) GO classification results showed that 20 287 DEGs were annotated. In addition, 21 600 DEGs were involved in 133 KEGG pathways. Among them, genes encoding the key enzymes related to volatile oil synthesis such as linalool synthase, myrcene synthase, geranyl geranyl pyrophosphate synthase, cinnamoyl CoA reductase, caffeic acid 3-O-methyltransferase and cinnamyl alcohol dehydrogenase were differentially expressed and enriched in monoterpene biosynthesis pathway, terpene skeleton biosynthesis pathway and phenylpropanoid biosynthesis pathway. (3) Transcription factor analysis showed that the DEGs were distributed in 31 transcription factor families, of which MYB family had the most Unigenes. In this study, transcriptome sequencing was performed to analyze the DEGs superior clones and common varieties of *I. verum* as well as their related functions and pathways. The candidate genes obtained provide the references for further exploring the synthesis mechanism of characteristic components of volatile oil and molecular breeding of *I. verum*.

Key words: *Illicium verum*, leaf, volatile oil synthesis, estragole, transcriptome analysis

八角 (*Illicium verum*) 又称大茴香、八月珠, 是中国传统香料和药用植物, 属于八角科 (Illiciaceae) 八角属 (*Illicium*)。它原野生于中国南亚热带地区, 主要分布于中国云南、广西、广东等省区以及越南 (马锦林等, 2006; Wang et al., 2011; Sun et al., 2019; 黄开顺等, 2020a)。

八角的主要利用部位是果实和枝叶, 可干制后整颗使用或制成五香粉, 是优质的烹饪调味香料。其干燥果实也被广泛认为是医药原料, 可作为食疗中药材, 具有镇痛、止咳、健胃、祛寒湿之效, 并且还被认为可用于治疗神经衰弱及消化不良等症 (马锦林等, 2006; 何冬梅等, 2009)。以八角叶片和果实为原料提取的挥发油又称茴香油, 是食品工业中的重要添加香料, 亦是香皂、香水、牙膏及化妆品等日化用品中不可缺少的常用加香剂 (王琴等, 2005; Chouksey et al., 2010; 叶志敏, 2011; 李萍等, 2016)。此外, 在工业上茴香油也被用作涂料填充剂及无氰电镀添加剂等 (马锦林等, 2006)。八角中还含有多种医用有效成分, 如八角中富含的莽草酸是合成禽流感防治药物“达菲” (磷酸奥司他韦) 的重要原料 (王祖磊和朱祥瑞, 2010), 草蒿脑则可用于合成新一代抗抑郁药“文拉法新” (梁忠云和王国聪, 2010)。

广西和云南是最早开展八角育种工作的主要省区。20 世纪 80 年代, 广西林业科学院就已经开始了八角资源调查、品种分类以及优株筛选等工作。随着一些八角优良单株被初步筛选培育出来以及嫁接技术的进步, 八角无性系造林试验在广西主要的八角产区陆续启动。

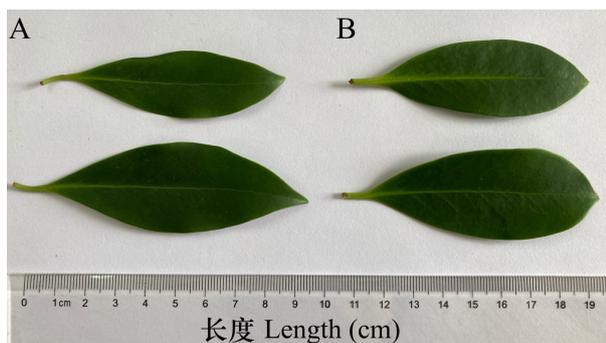
八角是“药食两用”的经济树种, 其挥发油含量及组分是重要的经济性状之一, 将其作为指标进行优株选育, 进一步筛选所需化学型的八角品种是较为传统的做法 (叶志敏, 2011)。现代育种技术借助于分子生物学手段实施分子设计育种, 虽然具有指向性强、育种速度快等优势, 但需要有一定的研究基础方能开展。目前, 八角分子生物学层面研究报道极少, 挥发油合成代谢途径及其调控机制不明, 极大地制约了八角分子育种的发展。转录组测序技术因其能够从特定基因组的整体层次上研究基因结构及基因表达量, 能揭示特定生物学过程中的分子机理, 所以在植物分子育种等领域已被广泛应用。耿秀文等 (2021) 利用转录组数据对东紫苏 (*Elsholtzia bodinieri*) 单萜合成相关基因进行研究, 系统地分析了单萜合成代谢通路并成功验证了 6 个单萜合成相关基因, 为东紫苏挥发油单萜合酶的克隆提供了基因资源。时小东等 (2018) 采用转录组测序方法对厚朴 (*Houpoa officinalis*) 次级代谢途径进行研究, 总结出厚朴苯丙素和萜类基因调控网络, 明确了关键酶的基因信息, 为在分子水平上改造厚朴代谢途径和目标产物的定向合成提供基础。杨国等 (2019) 借助高通量测序技术对白术 (*Atractylodes macrocephala*) 倍半萜合成途径展开分析, 确定了其中的关键酶基因, 极大地推动了白术分子育种的进展。这种基于转录组测序技术的合成代谢途径分析也在白芷 (*Angelica dahurica*) (吴萍等, 2020)、五味子 (*Schisandra Chinensis*) (陈春宇, 2020) 等中草药植物中取得了较好的研究结果, 证明了这种方法的有效性和可靠性。

本研究以挥发油含量与组分差异显著的优良无性系桂角 69 号与普通品种砧 01 号为研究对象, 采用高通量测序技术对两个品种叶片进行转录组测序, 通过转录组组装、注释、GO 分类和 KEGG 通路分析等方法对叶片中差异表达基因进行挖掘, 拟探讨以下问题: (1) 挥发油表型差异显著的两品种在转录水平上有哪些差异; (2) 这些差异表达的基因参与哪些生物学过程、代谢通路, 其中有哪些与挥发油合成相关。本研究有助于深入发掘八角挥发油合成相关的关键基因, 为下一步八角分子设计育种提供候选基因。

1 材料与方 法

1.1 材料

供试材料为八角优良无性系桂角 69 号和普通品种砧 01 号的 15 年生苗木, 采自广西林科院八角种质资源收集圃。砧 01 号叶片椭圆形, 叶尖渐尖; 桂角 69 号叶片椭圆形, 叶尖凸尖(图 1)。选择 10 月中旬晴朗天气采样, 在树冠外围东、西、南、北 4 个方向各采集成熟、完整、无病虫害叶片 15 片, 混合后置于液氮中速冻带回实验室保存, 用于挥发油检测及转录组测序。



A. 砧 01 号; B. 桂角 69 号。下同。
A. Zhen 01; B. Guijiao 69. The same below.

图 1 砧 01 号与桂角 69 号叶片形态

Fig. 1 Leaf morphology of Zhen 01 and Guijiao 69

1.2 方法

1.2.1 挥发油提取及测定 称取鲜叶装入圆底烧瓶中, 连接挥发油提取器和冷凝器, 蒸馏速率控制在 $2\sim 3\text{ mL}\cdot 100\text{ min}^{-1}$, 蒸馏至挥发油不再增加后停止加热, 冷却 10 min, 读取刻度管中挥发油毫升数, 计算挥发油含量。

取挥发油按照 GB/T 15068—2008《八角茴香(精)油》所述方法使用 Agilent6890N 气相色谱仪

进行分析, 具体分析条件为: HP-FFAP 柱 ($30\text{ m}\times 320\text{ }\mu\text{m}\times 0.5\text{ }\mu\text{m}$), 柱温 $60\text{ }^{\circ}\text{C}$, 汽化室 $250\text{ }^{\circ}\text{C}$, 检测室 $300\text{ }^{\circ}\text{C}$, $2\text{ }^{\circ}\text{C}\cdot\text{min}^{-1}$ 程序升温 $60\text{ }^{\circ}\text{C}$ (1 min) 至 $220\text{ }^{\circ}\text{C}$ (20 min), 进样量 $0.2\text{ }\mu\text{L}$, 分流比 100:1, 载气流速 $1.0\text{ mL}\cdot\text{min}^{-1}$ 。采用面积归一化法计算挥发油组分含量。试验均重复 3 次, 使用 SPSS 19.0 软件进行分析。

1.2.2 RNA 提取及转录组测序 取叶片使用 TRzol Universal 总 RNA 提取试剂盒(天根)按说明书操作提取总 RNA。RNA 的完整性通过凝胶电泳的方法检测, RNA 的浓度及纯度通过使用 NanoDrop ND-2000 进行检测。检测合格后的 RNA 样品送至华大基因公司进行 cDNA 文库构建并使用 BGISEQ-500 平台测序。

1.2.3 转录本组装及注释 将测序得到的原始数据使用 Trimmomatic 进行过滤得到 Clean Reads (Bolger et al., 2014), 使用 Trinity 软件进行组装 (Kim et al., 2015)。使用 Tgicl 软件将组装好的转录本进行聚类除去冗余以提高 Unigene 质量 (Mehra et al., 2020)。对供试材料的 6 个样品转录本进行分析, 桂角 69 号的 3 个生物学重复分别为标记为 Guijiao 69_1、Guijiao 69_2 和 Guijiao 69_3, 砧 01 号的 3 个生物学重复标记为 Zhen 01_1、Zhen 01_2 和 Zhen 01_3。

将得到的 Unigene 序列使用 BLAST、Blast2GO、HMMSCAN 软件 (Altschul et al., 1997; Buchfink et al., 2015) 与 EuKaryotic Orthologous Groups (KOG)、Swiss-Prot、NR、NT、Kyoto Encyclopedia of Genes and Genomes (KEGG)、Pfam、Gene Ontology (GO) 7 个数据库进行对比, 分析获得基因的注释信息。

1.2.4 差异表达基因分析 使用 RSEM 对各个样品的基因表达水平进行比对计算 (Li & Dewey, 2011), 依据基因表达水平进行样本相关性及 PCA 分析。采用 DESeq (Love et al., 2014) 对差异表达基因进行分析, 过滤掉平均表达量 <0.5 的基因后, 按照 Q 值 <0.05 和 $\log_2|\text{Fold change}| \geq 1$ 的标准筛选差异基因, 并对差异基因分别进行 GO 富集分析和 KEGG 通路富集分析。

2 结果与分析

2.1 叶片挥发油含量及组分分析

两品种叶片的挥发油气相色谱检测结果显示其主成分差异较大, 桂角 69 号主成分为草蒿脑, 砧 01 号则为反式茴香脑(图 2)。桂角 69 号中挥发油的总量和草蒿脑组分的含量都显著高于砧 01

号,分别为后者的 1.8 倍和 68 倍;而其顺式茴香脑和反式茴香脑的含量显著低于砧 01 号,分别是前者的 1/4 和 1/3;两个品种中的水分含量差别不大,均为 70%左右(表 1)。

2.2 转录组测序及组装注释

参试样本共获得 38.18 Gb 数据,各样品过滤后 Clean Reads 数均高于 42 M,碱基数均在 6.3 Gb 以上,各样品的 GC 含量稳定在 45%~46% 范围内,并且 Q30 质量值均在 90% 以上(表 2)。参试 6 个样本

Clean Reads 数据 SRA 平台 (<http://trace.ncbi.nlm.nih.gov/Traces/sra/>) 登录号为 PRJNA788720。

组装后共获得 84 182 条转录本序列,平均长 1 212 bp, N50 长 1 987 bp,平均 GC 含量为 42.85%。编码基因数量为 47 416,平均长 1 218 bp, N50 长 1 581 bp,平均 GC 含量为 45.13%。将转录本序列与 NR、NT、Swiss-Prot、KEGG、KOG、GO 和 Pfam 数据库的比对,共获得 59 161 个基因序列的注释结果,注释基因比例为 70.28%。分别有 55 153 (65.52%)、

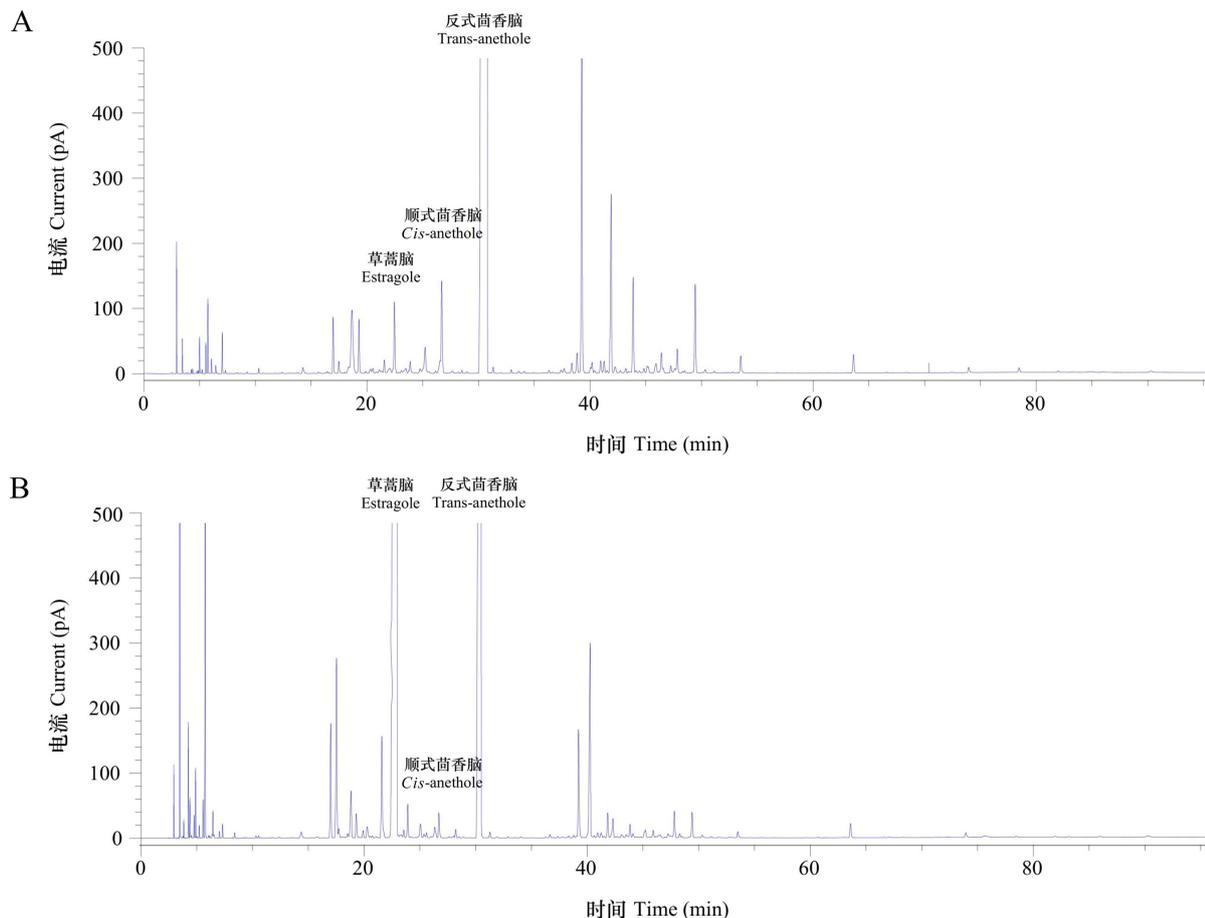


图 2 挥发油成分气相色谱图

Fig. 2 GC chromatography of volatile oil components

表 1 挥发油含量测定结果

Table 1 Test result of volatile oil contents

样品编号 Sample number	挥发油含量 Volatile oil content mL · 100 g ⁻¹	叶片水分含量 Leaf water content (%)	草蒿脑 Estragole (%)	顺式茴香脑 Cis-anethole (%)	反式茴香脑 Trans-anethole (%)
砧 01 号 Zhen 01	0.81±0.12 a	69.12±0.74 a	0.75±0.30 a	1.09±0.16 a	81.41±1.12 a
桂角 69 号 Guijiao 69	1.45±0.03 b	72.52±1.27 b	51.34±1.29 b	0.25±0.33 b	26.78±0.81 b

注: 同列不同字母表示 0.05 水平差异显著。

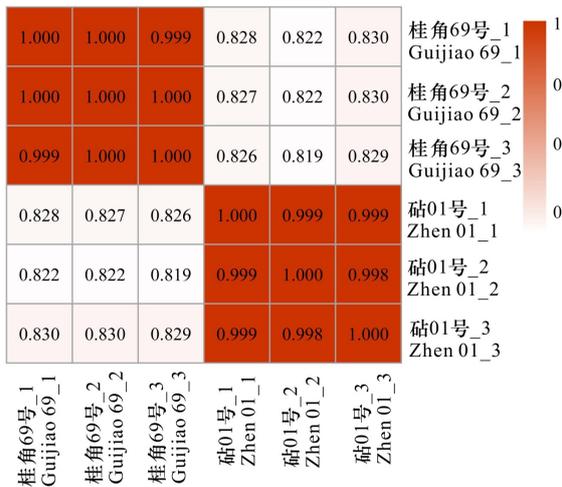
Note: Different letters in the same column represent significant differences at 0.05 level.

表 2 测序数据质量评估

Table 2 Quality statistics of sequencing data

样本编号 Sample number	数据量 Reads number (M)	碱基数 Base number (Gb)	GC 含量 GC content (%)	Q30 质量值 Q30 value (%)
桂角 69 号_1 Guijiao 69_1	42.31	6.35	45.71	90.12
桂角 69 号_2 Guijiao 69_2	42.21	6.33	45.73	90.38
桂角 69 号_3 Guijiao 69_3	42.44	6.37	45.78	90.05
砧 01 号_1 Zhen 01_1	42.10	6.32	45.89	90.33
砧 01 号_2 Zhen 01_2	43.28	6.49	45.98	90.62
砧 01 号_3 Zhen 01_3	42.14	6.32	45.97	90.79

A



38 241 (45.43%)、40 388 (47.98%)、43 136 (51.24%)、43 176 (51.29%)、43 097 (51.20%)、43 030 (51.12%) 个基因在 NR、NT、Swiss-Prot、KEGG、KOG、Pfam、GO 数据库中获得了解释。

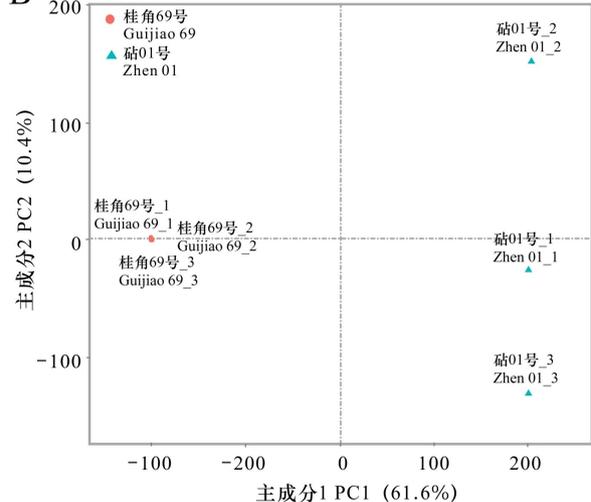
2.3 差异表达基因的筛选及分析

样本相关分析结果显示,同组样本基因表达模式趋于一致且相关系数均在 0.99 以上,两个品种之间的相关系数约为 0.80(图 3:A),主成分分析结果表明两组样品之间相互独立(图 3:B)。

2.4 差异基因 GO 功能注释及富集分析

两个品种样本中共筛选获得 30 572 个差异基因(DEGs),其中上调 DEGs 15 025 个,下调 DEGs 15 547 个。对 DEGs 进行 GO 功能注释和分析,确定其重要的生物学功能。本研究结果表明,共有 20 287 个 DEGs 获得注释,并划分为 3 大类、54 小

B



A. 样品相关性分析; B. 样品主成分分析。

A. Sample correlation analysis; B. Principal component analysis of samples.

图 3 样品相关性和主成分分析

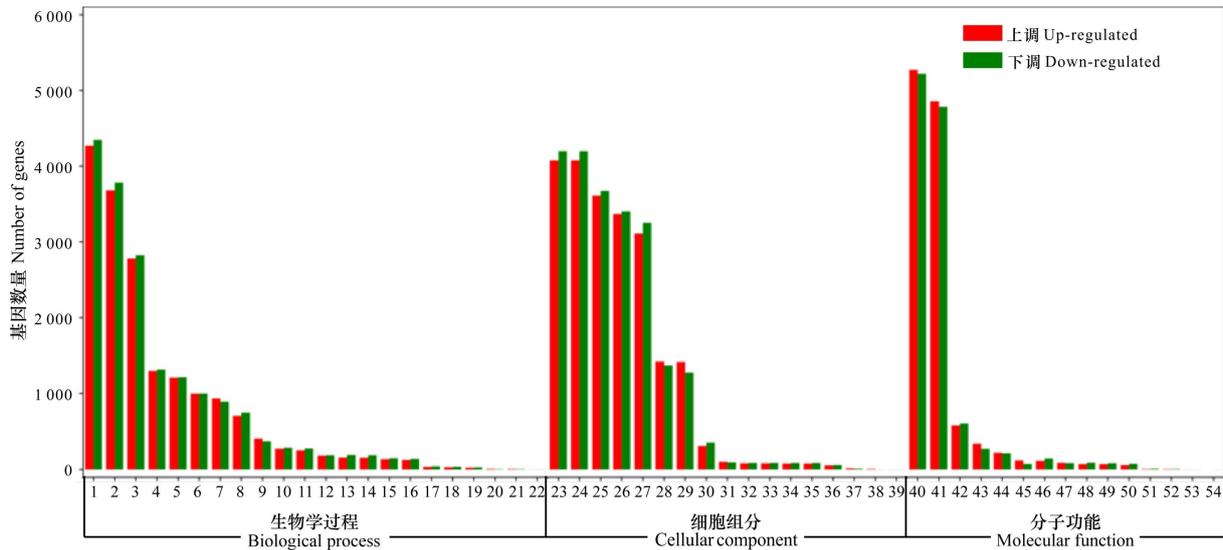
Fig. 3 Sample correlation and principal component analysis

类(图 4)。生物学过程类别主要涉及细胞过程、代谢过程、单一机体过程;细胞组分类别主要涉及细胞、细胞部件、细胞膜、细胞膜部件、细胞器 5 种细胞组分;分子功能类别主要涉及结合和催化活性 2 种分子功能。DEGs 富集分析结果显示,差异显著的只有生物学过程和分子功能(图 5)。生物学过程中基因表达(GO:0010467, gene expression)富集基因较多,与砧 01 号相比,桂角 69 号中有 1 327 个基因上调,1 292 个基因下调。差异最显著的 2 个 GO 条目分别与饥饿和营养相关(GO:0042594, response

to starvation; GO: 0031667, response to nutrient levels)。在分子功能方面,翻译起始因子活性(GO: 0003743, translation initiation factor activity)和 RNA 结合翻译因子活性(GO:0008135, translation factor activity, RNA binding)为最显著富集 GO 条目。

2.5 差异基因 KEGG 功能注释及富集分析

将 30 572 个 DEGs 与 KEGG 数据库进行比对,结果共有 21 600 个 DEGs 注释在 133 条通路上(图 6:A)。KEGG 生化代谢通路共分成 5 个生物过程,包括代谢、遗传信息过程、环境信息过程、



1. 细胞过程; 2. 代谢过程; 3. 单一机体过程; 4. 生物调节; 5. 生物过程的调节; 6. 刺激响应; 7. 细胞成分组织或生物发生; 8. 定位; 9. 信号; 10. 发展过程; 11. 多细胞生物过程; 12. 积极调节生物过程; 13. 复制; 14. 生殖过程; 15. 生物过程的负调控; 16. 多生物过程; 17. 生长; 18. 免疫系统过程; 19. 节律过程; 20. 解毒; 21. 运动; 22. 生物黏附; 23. 细胞; 24. 细胞部分; 25. 细胞膜; 26. 细胞膜部件; 27. 细胞器; 28. 细胞器部分; 29. 大分子复合物; 30. 膜封闭腔; 31. 细胞外区; 32. 细胞连接; 33. 共质体; 34. 其他生物体; 35. 其他有机体部分; 36. 超分子纤维; 37. 胞外区部分; 38. 类核; 39. 细胞外基质; 40. 结合; 41. 催化活性; 42. 转运体活性; 43. 结构分子活性; 44. 核酸结合转录因子活性; 45. 抗氧化活性; 46. 分子功能调节; 47. 电子载体活性; 48. 信号传感器活性; 49. 分子传感器活性; 50. 转录因子活性, 蛋白质结合; 51. 营养库活动; 52. 蛋白质标签; 53. 金属伴侣活性; 54. 翻译调节活性。

1. Cellular process; 2. Metabolic process; 3. Single-organism process; 4. Biological regulation; 5. Regulation of biological process; 6. Response to stimulus; 7. Cellular component organization or biogenesis; 8. Localization; 9. Signaling; 10. Developmental process; 11. Multicellular organismal process; 12. Positive regulation of biological process; 13. Reproduction; 14. Reproductive process; 15. Negative regulation of biological process; 16. Multi-organism process; 17. Growth; 18. Immune system process; 19. Rhythmic process; 20. Detoxification; 21. Locomotion; 22. Biological adhesion; 23. Cell; 24. Cell part; 25. Membrane; 26. Membrane part; 27. Organelle; 28. Organelle part; 29. Macromolecular complex; 30. Membrane-enclosed lumen; 31. Extracellular region; 32. Cell junction; 33. Symplast; 34. Other organisms; 35. Other organism parts; 36. Supramolecular fiber; 37. Extracellular region part; 38. Nucleoid; 39. Extracellular matrix; 40. Binding; 41. Catalytic activity; 42. Transporter activity; 43. Structural molecule activity; 44. Nucleic acid binding transcription factor activity; 45. Antioxidant activity; 46. Molecular function regulator; 47. Electron carrier activity; 48. Signal transducer activity; 49. Molecular transducer activity; 50. Transcription factor activity, protein binding; 51. Nutrient reservoir activity; 52. Protein tag; 53. Metallochaperone activity; 54. Translation regulator activity.

图 4 差异基因的 GO 功能分类

Fig. 4 GO function classification of DEGs

细胞学过程和有机体系统。代谢中除去总的代谢目录,以糖代谢通路注释的基因最多(1 518),氨基酸代谢通路和脂质代谢通路次之,注释基因数分别为 817、728。而在遗传信息过程中,翻译通路所含基因最多(1 513),折叠、分类和降解通路中含有 1 048 个 DEGs。虽然另外 3 个生物过程所注释的通路较其他分类少,但是一些通路中含有的 DEGs 较多,如有 991 个 DEGs 位于信号转导通路,887 个 DEGs 位于环境适应通路。值得注意的是,有 63 个 DEGs 位于不饱和脂肪酸合成相关通路,另有 256 个 DEGs 位于苯丙素合成通路。

通过 KEGG 富集分析,确定 DEGs 富集的最主要通路(图 6:B)。结果显示,脂肪酸生物合成和

代谢通路都被富集。一些次生物质合成通路也被富集,如类固醇生物合成、类胡萝卜素合成、单萜生物合成、苯丙素生物合成等。部分次生代谢途径也有富集,如甘油磷脂代谢、生物素代谢、硫辛酸代谢、嘌呤代谢等。同时,与防御相关的激素通路如植物-病原菌相互作用、植物激素信号转导也被富集。

2.6 挥发油物质合成途径分析

八角叶片挥发油主要由单萜、倍半萜化合物以及苯丙素类芳香族化合物组成,故对与之合成相关的单萜生物合成、萜类骨架生物合成、苯丙素合成通路中的 DEGs 进行分析。

单萜生物合成途径中发现芳樟醇合酶(*LS*)、

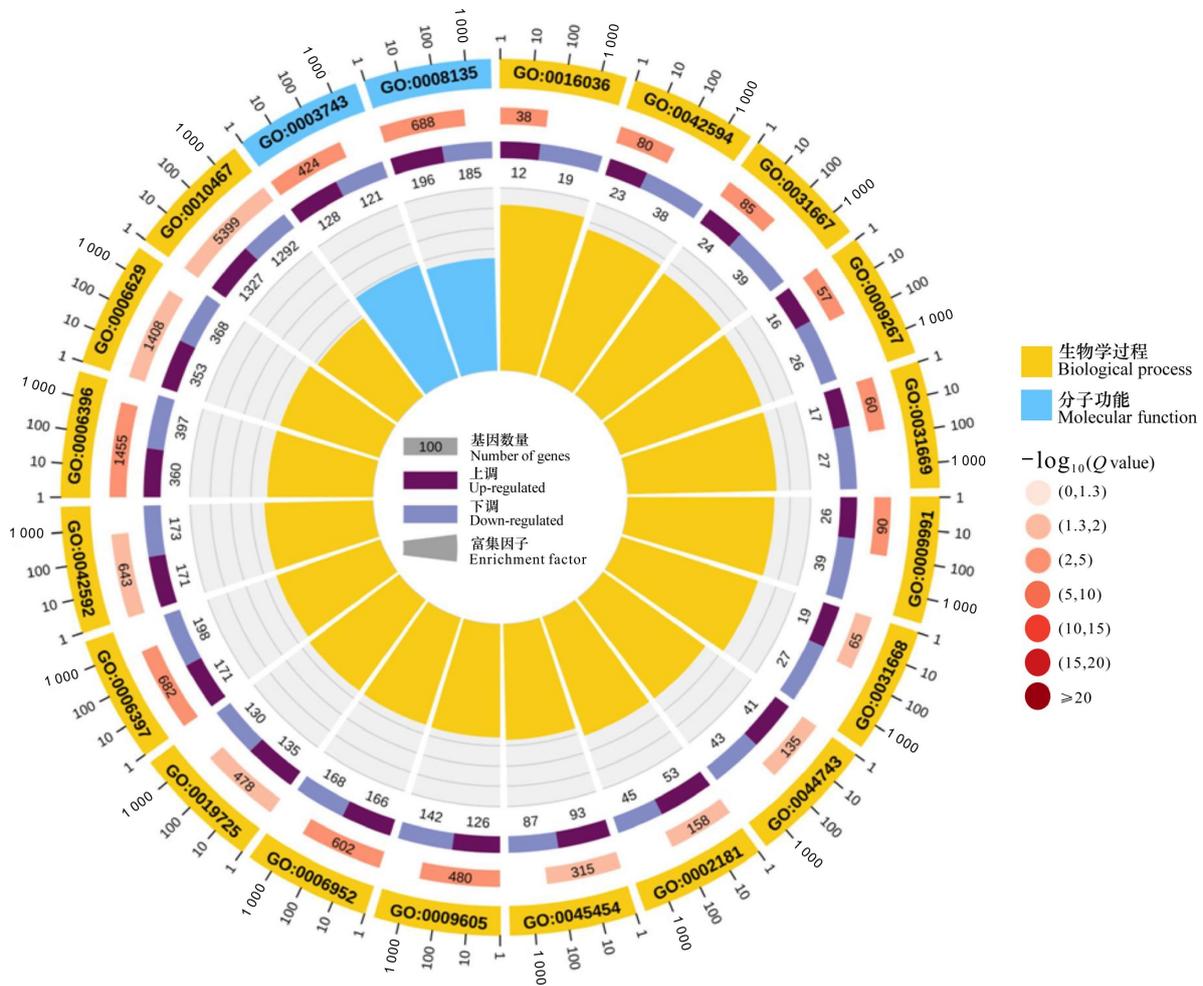


图 5 差异基因的 GO 功能富集

Fig. 5 GO function enrichment of DEGs

月桂烯合酶(*MYS*)、 β -萜澄茄烯合酶(*CBS*)基因呈现差异表达,其中*LS*在桂角 69 号中表达量远远高于砧 01 号(表 3)。萜类骨架生物合成途径中 4 个关键酶基因呈现差异表达,分别参与 MEP、MVA 等合成途径,相关 Unigene 共 10 个。其中,与砧 01 号相比,香叶基香叶基焦磷酸合酶(*GGPPS*)基因在桂角 69 号中高表达。苯丙素合成通路中的肉桂酰辅酶 A 还原酶(*CCR*)、咖啡酸 3-*O*-甲基转移酶(*COMT*)、肉桂醇脱氢酶(*CAD*)等 7 个关键酶基因呈差异表达,其中咖啡酰辅酶 A *O*-甲基转移酶(*CCoAOMT*)基因在桂角 69 号中表达显著上调,而 *CYP73A* 基因则在砧 01 号中上调最为显著。

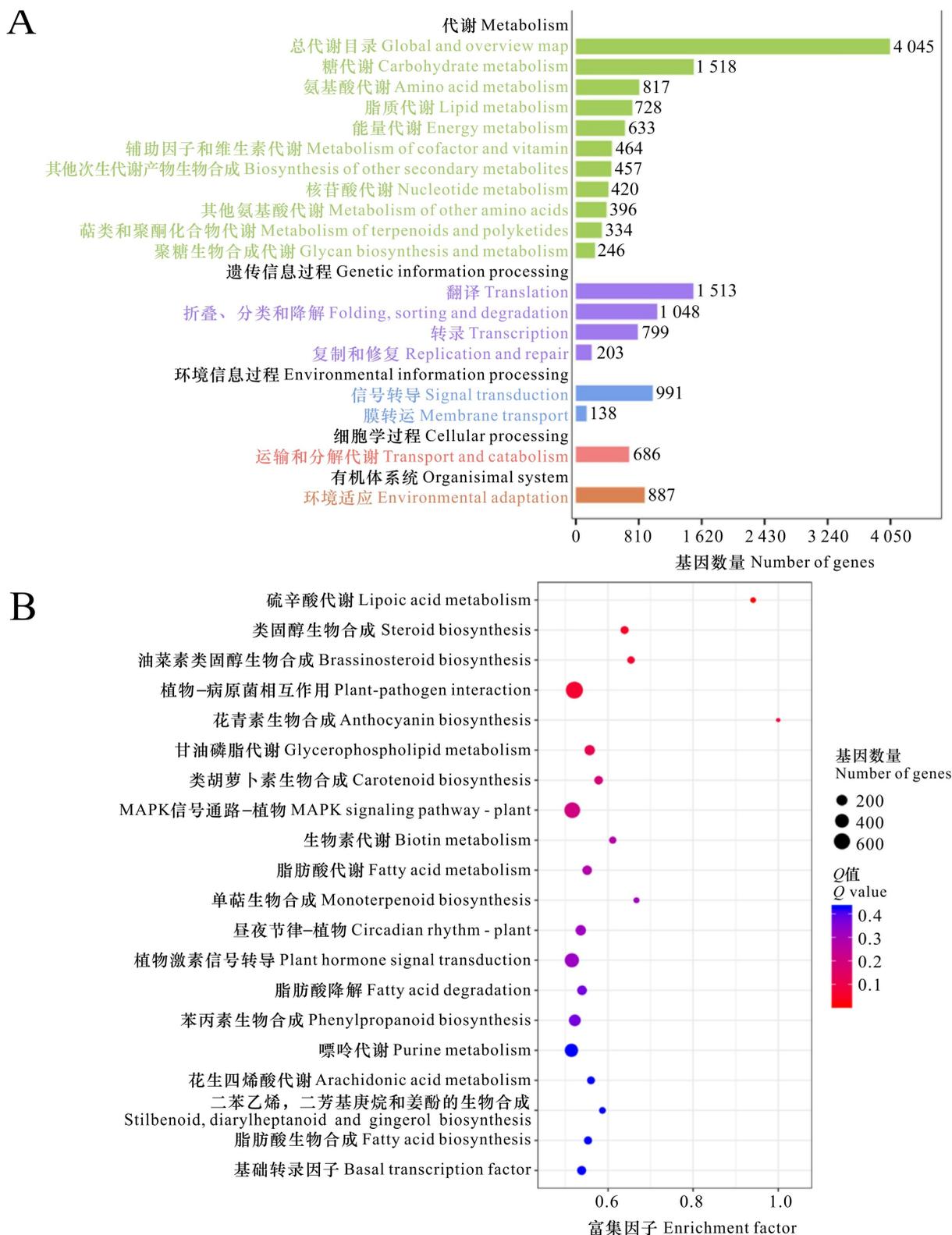
2.7 转录因子差异表达基因分析

根据基因注释结果,对其中差异表达的转录因子基因进行分析发现,DEGs 分布于 31 个转录因子家族,其中 MYB 家族 Unigene 数量最多,共 13

个(表 4)。其他与挥发油合成相关的转录因子家族如 WRKY、bHLH、AP2-EREBP、bZIP、C2H2、C3H 中分别具有 9、6、4、3、10、9 个 Unigenes。值得注意的是,注释为 bHLH 转录因子 MYC2 的 Unigene9952 以及注释为 MYB 转录因子 MYB111 的 Unigene15870 在桂角 69 号中上调表达。

3 讨论与结论

八角茴香油在香精、食品工业以及医药中应用广泛,普遍认为反式茴香脑是茴香油的主要成分。茴香油中的草蒿脑,又名对烯丙基茴香醚,含量仅次于反式茴香脑,同为苯丙素类挥发性成分,具有浓郁的茴香香气和甜味,是构成茴香油特定香味的主要化合物之一,因此也常被用于生产食用香料(缪剑华等,2008;梁忠云和王国聪,2010;



A. KEGG 通路注释; B. KEGG 富集前 20 的通路。

A. KEGG pathway annotation; B. Top 20 pathways of KEGG enrichment.

图 6 差异基因的 KEGG 通路注释和富集

Fig. 6 KEGG pathway annotation and enrichment of DEGs

表 3 挥发油合成代谢通路及相关基因分析
Table 3 Analysis of pathway and related genes of volatile oil synthesis

代谢通路 Metabolic pathway	相关基因 Related gene	Unigene 数量 Number of Unigenes	基因表达量 (FPKM) Gene expression	
			桂角 69 号 Guijiao 69	砧 01 号 Zhen 01
单萜生物合成 Monoterpenoid biosynthesis	芳樟醇合酶 <i>Linalool synthase (LS)</i>	5	1.76	2.56
	月桂烯合酶 <i>Myrcene synthase (MYS)</i>	1	27.30	0.68
	β -葑澄茄烯合酶 <i>Beta-cubebene synthase (CBS)</i>	1	0.90	3.97
萜类骨架生物合成 Terpenoid backbone biosynthesis	1-脱氧-D-木酮糖-5-磷酸合成酶 <i>1-Deoxy-D-xylulose-5-phosphate synthase (DXS)</i>	5	14.07	3.08
	香叶基香叶基焦磷酸合酶 <i>Geranyl geranyl pyrophosphate synthase (GGPPS)</i>	2	124.72	9.31
	法尼基焦磷酸合酶 <i>Farnesyl pyrophosphate synthase (FPS)</i>	2	2.46	6.04
	3-羟基 3-甲基戊二酰辅酶 A 还原酶 <i>3-Hydroxy-3-methylglutaryl-coenzyme A reductase (HMGR)</i>	1	20.30	99.79
	肉桂酰辅酶 A 还原酶 <i>Cinnamoyl-CoA reductase (CCR)</i>	14	31.26	35.59
苯丙素合成 Phenylpropanoids biosynthesis	咖啡酸 3-O-甲基转移酶 <i>Caffeic acid 3-O-methyltransferase (COMT)</i>	3	16.57	46.18
	肉桂醇脱氢酶 <i>Cinnamyl alcohol dehydrogenase (CAD)</i>	2	1.62	4.09
	4-香豆酸-CoA 连接酶 <i>4-Coumarate-CoA ligase (4CL)</i>	2	7.31	8.31
	苯丙氨酸解氨酶 <i>Phenylalanine ammonia-lyase (PAL)</i>	1	6.76	19.50
	咖啡酰辅酶 A O-甲基转移酶 <i>Caffeoyl-CoA O-methyltransferase (CCoAOMT)</i>	1	110.92	79.18
	CYP73A	1	1.52	27.34

表 4 转录因子分析
Table 4 Analysis of transcription factor

转录因子 Transcription factor	Unigene 数量 Number of Unigenes	上调 Up- regulated	下调 Down- regulated	转录因子 Transcription factor	Unigene 数量 Number of Unigenes	上调 Up- regulated	下调 Down- regulated
MYB	13	5	8	GeBP	2	1	1
C2H2	10	7	3	GRAS	2	1	1
C3H	9	3	6	mTERF	2	1	1
FAR1	9	3	6	PBF-2-like	2	1	1
WRKY	9	4	5	Trihelix	2	2	0
C2C2-Dof	7	6	1	Alfin-like	1	0	1
ARF	6	3	3	ARR-B	1	0	1
bHLH	6	4	2	BES1	1	1	0
G2-like	6	4	2	C2C2-CO-like	1	1	0
NAC	6	4	2	C2C2-YABBY	1	1	0
AP2-EREBP	4	2	2	LIM	1	0	1
ABI3VP1	3	1	2	MADS	1	0	1
bZIP	3	0	3	SBP	1	0	1
C2C2-GATA	3	1	2	TAZ	1	1	0
EIL	3	2	1	VOZ	1	0	1
HSF	3	2	1				

阳小勇和唐荣平,2018)。此外,由于草蒿脑具有较强的杀菌、退热、止咳、健胃、醒脑等生物活性而被广泛应用于农药、医药等领域(梁忠云和王国聪,2010)。近年来,草蒿脑多次被报道用于合成医药中间体茴香醛和对甲氧基苯乙酸,后者是合成血管扩张剂盐酸硫氮卓酮和新一代抗抑郁药“文拉法新”的重要原料,具有较高的应用前景(梁忠云和王国聪,2010)。因此,在丰富的八角种质资源中选育具有特色挥发油组分的八角品种,有助于八角茴香油的多元化开发利用,对于八角产业发展具有重要意义。

对于八角果实,无论是挥发油总量还是反式茴香脑等重要组分含量,不同采收期均存在一定的差异,且果实发育过程中,其生理过程也通常发生一系列变化,而叶片的挥发油物质含量相对较为稳定,特别是在秋季(王琴等,2011;黄开顺等,2020b),因此本研究选择在10月份采集叶片进行挥发油分析其研究结果最具代表性。本研究发现,桂角69号除了在果产量、果角完好率等生产性状上优于普通品种砧01号以外,其叶片中总挥发油含量也显著高于后者,特别是其挥发油中主成分草蒿脑含量显著高于对照品种,具有非常鲜明的化学型特征。

植物挥发油主要由萜类、苯丙素类等次级代谢产物组成。萜类生物合成途径已较为清晰,MVA途径和MEP途径提供的IPP经萜类合酶分别合成单萜、倍半萜、二萜等复杂多样的萜类化合物。*HMGR*和*DXS*分别是MVA途径和MEP途径中的关键酶基因,两者在桂角69号和砧01号中表达模式相反,提示砧01号中可能以MVA途径为主,而桂角69号则更倾向于通过MEP途径合成IPP。单萜生物合成途径中分析发现月桂烯合酶基因在桂角69号中大幅上调表达,有报道提出月桂烯合酶的催化产物有月桂烯、桉烯、沉香醇、柠檬烯等多种形式(徐应文等,2009),因此月桂烯合酶的高表达意味着桂角69号挥发油组分中可能具有更多、更丰富的单萜组分。另外,为二萜合成提供前体的*GGPPS*基因在桂角69号中高表达,则表明其可能含有更高的二萜组分,在后续研究中可进行针对性检测。在对两种叶片进行转录组分析中发现差异基因中有256个富集于苯丙素合成通路基因,其中包括*CCR*、*COMT*、*CAD*、*4CL*、*PAL*、*CCoAOMT*、*CYP73A*等关键酶基因。这些基因中大部分都在桂角69号中呈下调表达,仅有*CCoAOMT*在桂角69号中上调表达,且为大幅上调表达,可能与其高草蒿脑合成相关联。这里值得注意的

是,*O*-甲基转移酶是苯丙素合成过程的重要的酶,可利用*S*-腺苷甲硫氨酸作为甲基供体合成苯丙素类挥发性成分草蒿脑及茴香脑。一些茴香属植物的*O*-甲基转移酶的功能已被鉴定,例如在茴香(*Foeniculum vulgare*)中,对丙烯基苯酚*O*-甲基转移酶将反式-对丙烯基苯酚甲基化为反式茴香脑,而对烯丙基苯酚被烯丙基苯酚*O*-甲基转移酶甲基化后形成草蒿脑(Gross et al., 2002)。目前,对草蒿脑及茴香脑生物合成的机制仍然知之甚少,*CCoAOMT*不失为一个候选的研究切入点。转录因子调控是植物调节次生代谢的重要手段,bHLH、WRKY、MYB等转录因子广泛参与到萜类合成途径当中。本研究发现,注释为转录因子MYC2的Unigene9952以及注释为转录因子MYB111的Unigene15870在桂角69号中上调表达。MYC2通过茉莉酸信号转导在倍半萜化合物的生物合成中发挥作用(蒋玲,2017),MYB111则被报道正向调节黄酮醇生物合成(朱昀昊等,2019)。两者潜在的调控功能及其在桂角69号中的特异上调或许与桂角69号的高挥发油含量相关联,可作为候选基因深入研究。

此外,叶片对八角植株的生长发育至关重要,不仅是八角合成和贮存养分的主要器官,还可以通过向花、果等器官输送养分平衡各器官间养分含量(黄开顺等,2012)。糖类物质代谢、氨基酸代谢及脂质代谢均是植物体内重要的新陈代谢过程,对植物的生长至关重要。在一定的程度上,糖代谢能反映植株的生长状态(吴晓俊等,2000;胡珍兰等,2014),并且叶片糖代谢的变化能够影响果实的生长发育,例如,葡萄叶片糖含量及其相关代谢酶活性会对酿酒葡萄的品质和产量有显著影响(刘丽媛,2016;胡紫璟,2016)。氨基酸主要被用于蛋白质合成,也可以作为次生代谢产物生物合成的前体和能量来源。它在植物的中枢代谢中具有许多重要作用,并且还在某些生理代谢途径中被认为是最终代谢物的中间产物,并参与调控多种代谢过程和其他各种生理生化途径,从而影响植物生长发育(Amir et al., 2018; Yang et al., 2018; Yang et al., 2020)。植物脂类包括广泛的化合物,如脂肪酸、甘油磷脂、半乳脂、鞘脂、甾醇和中性脂(例如,三酰基甘油、甾醇酯、角质、蜡和软木脂)(Ohlrogge & Browse, 1995; Ischebeck, 2016)。脂类作为主要的生物分子,在大多数生物中发挥着结构成分、可持续的碳及能量存储、活跃的信号传感器和表面覆盖物等重要作用(Li-Beisson et al., 2016)。植物叶片的角质层蜡也通

常被定义为表面覆盖的脂类(Kunst & Samuels, 2009)。桂角 69 号的五年平均产量、单果平均鲜重及果角完好率均远高于砧 01 号,且两者的转录组差异表达基因中包含大量糖代谢、氨基酸代谢及脂质代谢的相关基因,可从中筛选候选基因,鉴定其与八角植株生长间的关系,为分子辅助选育八角良种提供参考。

本研究采用转录组测序技术对八角叶片基因表达特征进行分析,挖掘其中挥发油合成代谢通路及相关基因,获得的候选基因为下一步探究草蒿脑、茴香脑等挥发油特征组分的合成机制提供了参考,并为开展分子设计育种培育特定化学型的八角品种提供了研究基础。

参考文献:

- ALTSCHUL SF, MADDEN TL, SCHÄFFER AA, et al., 1997. Gapped BLAST and PSI BLAST: a new generation of protein database search programs [J]. Nucl Acid Res, 25(17): 3389-3402.
- AMIR R, GALILI G, COHEN H, 2018. The metabolic roles of free amino acids during seed development [J]. Plant Sci, 275: 11-18.
- BOLGER AM, LOHSE M, USADEL B, 2014. Trimmomatic: a flexible trimmer for Illumina sequence data [J]. Bioinformatics, 30(15): 2114-2120.
- BUCHFINK B, XIE C, HUSON DH, 2015. Fast and sensitive protein alignment using DIAMOND [J]. Nat Methods, 12(1): 59-60.
- CHEN CY, 2020. Study on the lignan and flavonoid biosynthetic pathway in *Schisandra chinensis* by transcriptomics and metabolomics [D]. Changchun: Jilin Agricultural University: 128-129. [陈春宇, 2020. 基于转录组和代谢组解析五味子木脂素和黄酮生物合成途径 [D]. 长春: 吉林农业大学: 128-129.]
- CHOUKSEY D, SHARMA P, PAWAR RS, 2010. Biological activities and chemical constituents of *Illicium verum* Hook fruits (Chinese star anise) [J]. Der Pharm Sin, 1(3): 1-10.
- GENG XW, ZHANG AL, TANG RH, et al., 2021. High-throughput transcriptome sequencing of *Elsholtzia bodinieri* and excavation of genes related to monoterpene biosynthesis [J]. Chin Trad Herb Drugs, 52(11): 3373-3382. [耿秀文, 张爱丽, 唐仁华, 等, 2021. 利用转录组数据挖掘东紫苏单萜生物合成相关基因 [J]. 中草药, 52(11): 3373-3382.]
- GROSS M, FRIEDMAN J, DUDAI N, et al., 2002. Biosynthesis of estragole and t-anethole in bitter fennel (*Foeniculum vulgare* Mill. var. *vulgare*) chemotypes. Changes in SAM: phenylpropene O-methyltransferase activities during development [J]. Plant Sci, 163(5): 1047-1053.
- HE DM, LIU HX, HUANG CS, et al., 2009. Studies on extracting and component of star anise oil from *Illicium verum* Hook. f in Guangxi [J]. Food Res Develop, 30(6): 3-6. [何冬梅, 刘红星, 黄初升, 等, 2009. 广西八角茴油的提取及其成分研究 [J]. 食品研究与开发, 30(6): 3-6.]
- HU ZJ, 2016. The study on the influence of different nitrogen levels on metabolism of carbon and nitrogen in wine grape 'cabernet gemischet' tree [D]. Lanzhou: Gansu Agricultural University: 61-62. [胡紫璟, 2016. 不同氮素水平对酿酒葡萄'蛇龙珠'植株碳氮代谢影响的研究 [D]. 兰州: 甘肃农业大学: 61-62.]
- HU ZL, ZHANG HW, HU CX, et al., 2014. Effects of molybdenum and nitrogen combined application on the net photosynthesis and soluble sugar in flue-cured tobacco [J]. Acta Agric Boreal-Sin, 29(3): 199-203. [胡珍兰, 张海伟, 胡承孝, 等, 2014. 钼氮配施对烤烟光合作用及可溶性糖含量的影响 [J]. 华北农学报, 29(3): 199-203.]
- HUANG KS, LI GQ, AN JC, et al., 2020a. Developing situation of processing and utilization industry of characteristic resources of star anise (*Illicium verum*) [J]. Biomass Chem Eng, 54(6): 6-12. [黄开顺, 黎贵卿, 安家成, 等, 2020a. 八角特色资源加工利用产业发展现状 [J]. 生物质化学工程, 54(6): 6-12.]
- HUANG KS, LIAO JM, QIN Y, et al., 2012. Characteristics of flowering fruiting and nutrient requirement of *Illicium verum* [J]. J SW For Coll, 32(3): 97-99. [黄开顺, 廖健明, 覃毓, 等, 2012. 八角开花结果与养分需求特性 [J]. 西南林业大学学报, 32(3): 97-99.]
- HUANG KS, ZENG XY, LI BC, et al., 2020b. Seasonal variation of contents of essential oil and chemical composition in *Illicium verum* leaves under different fertilizer treatments [J]. Nanfang For Sci, 48(4): 4-10. [黄开顺, 曾祥艳, 李宝财, 等, 2020b. 不同肥料处理下八角叶挥发油和化学成分含量的季节变化 [J]. 南方林业科学, 48(4): 4-10.]
- ISCHEBECK T, 2016. Lipids in pollen-they are different [J]. BBA-Mol Cell Biol Lipids, 1861(9 Pt. B): 1315-1328.
- JIANG L, 2017. Transcriptomic analysis of the endangered medicinal plants *Atractylodes lancea* (Thunb.) DC [D]. Nanjing: Nanjing University of traditional Chinese Medicine: 37-39. [蒋玲, 2017. 濒危药用植物茅苍术转录组学研究 [D]. 南京: 南京中医药大学: 37-39.]
- KIM HS, LEE BY, WON EJ, et al., 2015. Identification of xenobiotic biodegradation and metabolism-related genes in the copepod *Tigriopus japonicus* whole transcriptome analysis [J]. Mar Genom, 24 Pt.3: 207-208.
- KUNST L, SAMUELS L, 2009. Plant cuticles shine: advances in wax biosynthesis and export [J]. Curr Opin Plant Biol, 12(6): 721-727.
- LI B, DEWEY CN, 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome [J]. BMC Bioinformatics, 12: 323.
- LI P, SHU Z, SHEN XX, et al., 2016. Comparison of star anise oil extracted by three methods [J]. Food Sci Technol, 41(12): 213-219. [李萍, 舒展, 申晓霞, 等, 2016. 3 种

- 方法提取的八角茴香油的比较研究 [J]. 食品科技, 41(12): 213-219.]
- LIANG ZY, WANG GC, 2010. Research progress of estragole resources and bioactivity application [J]. Guangxi For Sci, 39(1): 49-51. [梁忠云, 王国聪, 2010. 草蒿脑资源及生物活性应用研究进展 [J]. 广西林业科学, 39(1): 49-51.]
- LI-BEISSON Y, NAKAMURA Y, HARWOOD J, 2016. Lipids: from chemical structures, biosynthesis, and analyses to industrial applications [J]. Subcell Biochem, 86: 1-18.
- LIU LY, 2016. The physiological study on the sugar and acid accumulations and metabolic regulation mechanisms of *Vitis amurensis* rupestris. Grape [D]. Yangling: Northwest University of Agriculture and Forestry Science and Technology: 92-93. [刘丽媛, 2016. 山葡萄糖酸积累规律及代谢调控机理研究 [D]. 杨凌: 西北农林科技大学: 92-93.]
- LOVE MI, HUBER W, ANDERS S, 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2 [J]. Genome Biol, 15(12): 550.
- MA JL, ZHANG RQ, LI KX, 2006. Literature review of research on star anise [J]. Non-Wood For Res, (3): 59-61. [马锦林, 张日清, 李开祥, 2006. 广西八角良种研究综述 [J]. 经济林研究, (3): 59-61.]
- MEHRA R, JASROTIA RS, MAHAJAN A, et al., 2020. Transcriptome analysis of snow mountain garlic for unraveling the organosulfur metabolic pathway [J]. Genomics, 112(1): 99-107.
- MIU JH, GUO Y, SONG YJ, et al., 2008. GC-MS analysis of chemical components of volatile oil extracted from star anise fruit, branches and leaves by subcritical CO₂ [J]. Chin J Chin Mat Med, 33(9): 1090-1092. [缪剑华, 郭勇, 宋芸娟, 等, 2008. 八角果实及枝叶亚临界 CO₂ 萃取挥发油化学成分的 GC-MS 分析 [J]. 中国中药杂志, 33(9): 1090-1092.]
- OHLROGGE J, BROWSE J, 1995. Lipid biosynthesis [J]. Plant Cell, 7(7): 957-970.
- SHI XD, GU YX, DAI J, et al., 2018. Analysis of secondary metabolism pathway genes based on *Magnolia officinalis* transcriptome [J]. Lishizhen Med Mater Med Res, 29(1): 247-250. [时小东, 顾雨熹, 代娇, 等, 2018. 基于转录组的厚朴次级代谢产物途径基因挖掘及分析 [J]. 时珍国医国药, 29(1): 247-250.]
- SUN WL, SHAHRAJABIAN MH, CHENG Q, 2019. Anise (*Pimpinella anisum* L.), a dominant spice and traditional medicinal herb for both food and medicinal purposes [J]. Cogent Biol, 5(1673688): 1-25.
- WANG GW, HU WT, HUANG BK, et al., 2011. *Illicium verum*: A review on its botany, traditional use, chemistry and pharmacology [J]. J Ethnopharmacol, 136(1): 10-20.
- WANG Q, JIANG L, WEN QB, 2005. Advances in studies on *Illicium verum* [J]. Chin Cond, (5): 18-22. [王琴, 蒋林, 温其标, 2005. 八角茴香的研究进展 [J]. 中国调味品, (5): 18-22.]
- WANG Q, OU ZB, JIANG L, et al., 2011. Study on optimum harvest time of *Illicium verum* [J]. Sci Technol Food Ind, 32(4): 85-86. [王琴, 区子弁, 蒋林, 等, 2011. 八角果实最佳采收期的研究 [J]. 食品工业科技, 32(4): 85-86.]
- WANG ZL, ZHU XR, 2010. The progress in research on the use of *Illicium verum* Hook.f. and its extract shikimic acid [J]. Bull Sci Technol, 26(4): 531-535. [王祖磊, 朱祥瑞, 2010. 八角茴香及其提取物莽草酸的应用进展 [J]. 科技通报, 26(4): 531-535.]
- WU XJ, LIU D, HU ZB, 2000. UDP-glucose pyrophosphorylase [J]. Plant Physiol Comm, 36(3): 193-200. [吴晓俊, 刘涤, 胡之璧, 2000. 尿苷二磷酸葡萄糖焦磷酸化酶 [J]. 植物生理学通讯, 36(3): 193-200.]
- WU P, WANG XY, LI QM, et al., 2020. Analysis of differentially expressed genes of roots of *Angelica dahurica* with high-throughput RNA-seq technology [J]. SW Chin J Agric Sci, 33(2): 233-240. [吴萍, 王晓宇, 李青苗, 等, 2020. 基于高通量转录组测序的白芷差异表达基因分析 [J]. 西南农业学报, 33(2): 233-240.]
- XU YW, LÜ JJ, WU W, et al., 2009. The progress of the research on plant monoterpene synthases [J]. Acta Ecol Sin, 29(6): 3188-3197. [徐应文, 吕季娟, 吴卫, 等, 2009. 植物单萜合酶研究进展 [J]. 生态学报, 29(6): 3188-3197.]
- YANG QQ, ZHAO DS, LIU QQ, 2020. Connections between amino acid metabolisms in plants: lysine as an example [J]. Front Plant Sci, 11: 928.
- YANG QQ, ZHAO DS, ZHANG CQ, et al., 2018. A connection between lysine and serotonin metabolism in rice endosperm [J]. Plant Physiol, 176(3): 1965-1980.
- YANG XY, TANG RP, 2018. Study on the chemical components and application of *Illicium verum* [J]. Chin Cond, 43(8): 194-195. [阳小勇, 唐荣平, 2018. 八角茴香的化学成分及应用研究 [J]. 中国调味品, 43(8): 194-195.]
- YANG G, LI HH, JIN YF, et al., 2019. Analysis of genes related to biosynthesis of sesquiterpene in *Atractylodes macrocephala* by transcriptome [J]. Plant Physiol J, 55(12): 1827-1838. [杨国, 李鸿慧, 金叶飞, 等, 2019. 基于转录组分析白术倍半萜生物合成的相关基因 [J]. 植物生理学报, 55(12): 1827-1838.]
- YE ZM, 2011. Present situation and development countermeasures of star anise breeding in China [J]. J Yangtze Univ (Nat Sci Ed), 8(10): 229-232. [叶志敏, 2011. 我国八角育种现状及发展对策 [J]. 长江大学学报 (自然科学版), 8(10): 229-232.]
- ZHU YH, ZHANG MJ, LI L, et al., 2019. Analysis of MYB transcription factors related to triterpenoids and phenolic acids biosynthesis of *Prunella vulgaris* [J]. Chin Trad Herb Drugs, 50(9): 2165-2171. [朱昞昊, 张梦佳, 李璐, 等, 2019. 夏枯草三萜和酚酸类合成相关的 MYB 转录因子的挖掘及分析 [J]. 中草药, 50(9): 2165-2171.]

(责任编辑 李 莉 王登惠)