浙 江 农 林 大 学 学 报, 2023, **40**(5): 970-981 Journal of Zhejiang A&F University doi: 10.11833/j.issn.2095-0756.20220737

## 基于代谢与转录水平的花榈木萜类合成候选基因分析

王佳琦,王 欣,邓小梅,吴蔼民

(华南农业大学林学与风景园林学院,广东广州 510642)

摘要:【目的】花榈木 Ormosia henryi 是重要的木本药用植物。研究花榈木萜类次生代谢物积累规律,分析萜类生物合成通路和相关的关键酶基因,对花榈木药用价值的开发具有重要意义。【方法】采用液相色谱-质谱联用技术与高通量转录组学测序,利用生物信息学分析方法从差异代谢产物和表达基因中寻找萜类生物合成的关联酶基因。【结果】代谢组数据中共检测到了 15 种萜类化合物,其中含有甘草次酸、齐墩果酸和芳樟醇等具有药用活性的物质,多数在叶片中相对含量高于其他组织部位。转录组数据筛选出 49 个与萜类合成有关的候选基因,分析后发现:甲羟戊酸 (MVA) 途径中的差异基因在幼叶中表达较高,甲基赤藓糖磷酸酯 (MEP) 途径的差异基因在老叶中表达较高。根据加权基因共表达网络 (WGCNA)分析,发现 MYB、WRKY、bHLH 和 HB-HD-ZIP 等转录因子家族在花榈木的萜类合成中起重要作用,并推测了 6 个可能与萜类物质生物合成有关的转录因子,即 HB-HD-ZIP (c64527.graph\_c1)、GRF (c76195.graph\_c0)、DBB (c66970.graph\_c2)、DBB (c75593.graph\_c0)、HB-HD-ZIP (c63393.graph\_c0)和 C3H (c70385.graph\_c1)。【结论】从花榈木代谢组和转录组数据库中初步获得了重要的萜类物质及可能参与萜类化合物生物合成途径的候选关键酶基因,为进一步阐明花榈木萜类化合物生物合成的分子机制奠定基础。图 8 表 1 参 36

中图分类号: S722 文献标志码: A 文章编号: 2095-0756(2023)05-0970-12

# Analysis of candidate genes for terpene synthesis in *Ormosia henryi* based on metabolome and transcriptome

WANG Jiaqi, WANG Xin, DENG Xiaomei, WU Aimin

(College of Forestry and Landscape Architecture, South China Agricultural University, Guangzhou 510642, Guangdong, China)

Abstract: [Objective] Ormosia henryi is an important woody medicinal plant. This study aims to explore the accumulation pattern of terpene secondary metabolites and analyze the terpene biosynthesis pathway and related key enzyme genes, which is of great significance to the development of medicinal value in O. henryi. [Method] Liquid chromatography-mass spectrometry and high-throughput transcriptomic sequencing were used to find the associated enzyme genes of terpene biosynthesis from differentially expressed genes based on bioinformatics analysis. [Result] A total of 15 terpene compounds were detected in the metabolomic data and they contained enoxolone, oleanolic acid, linalool and other substances with medicinal activity, most of which were relatively higher in leaves than in other tissue parts. Transcriptome data screened 49 candidate genes related to terpene synthesis, and after analysis it was found that differential genes in the MVA pathway were highly expressed in young leaves while differential genes in the MEP pathway were highly expressed in old leaves. Based on WGCNA analysis, it was found that transcription factor families such as MYB, WRKY, bHLH

收稿日期: 2022-11-28; 修回日期: 2023-03-09

基金项目:广东省林业科技创新项目(2017KJCX028)

作者简介: 王佳琦 (ORCID: 0000-0002-6444-7077),从事花榈木次生代谢研究。E-mail: wangjiaqi1990@163.com。通 信作者:吴蔼民 (ORCID: 0000-0001-8322-8833),教授,博士,从事林木生物学研究。E-mail: wuaimin@ scau.edu.cn

and HB-HD-ZIP played an important role in terpene synthesis of *O. henryi*, and 6 transcription factors that might be related to terpene biosynthesis content were predicted, namely HB-HD-ZIP (*c*64527.*graph\_c1*), GRF (*c*76195.*graph\_c0*), DBB (*c*66970.*graph\_c2*), DBB (*c*75593.*graph\_c0*), HB-HD-ZIP (*c*63393.*graph\_c0*) and C3H (*c*70385.*graph\_c1*). [Conclusion] Important terpenes and candidate key enzyme genes that may participate in the biosynthesis pathway of these terpenes have been preliminarily obtained from the database of metabolomics and transcriptome of *O. henryi*. [Ch, 8 fig. 1 tab. 36 ref.]

Key words: Ormosia henryi; metabolome; transcriptome; terpenoids; biosynthesis

花榈木 Ormosia henryi 为豆科 Leguminosae 红豆属 Ormosia 重要木本药用植物,主要产于东南亚和 南美洲,属于国家二级保护植物,在野外处于中度濒危状态<sup>[1]</sup>。花榈木常以根、根皮及茎、叶入药。《全 国中草药汇编》记述,其性味归经为辛、温,有毒,具有活血化瘀、祛风、消肿之功。研究发现:红豆 属植物含有生物碱、黄酮、苯丙素、萜类和挥发油等多种化学成分,其中萜类化合物是一类重要的活性 物质<sup>[2]</sup>。迄今为止,已从红豆属不同植物中发现大量的萜类化合物,如丁香酚、苯乙醇、榄香素和棕榈 酸等化合物,并发现多种物质具有消炎、抗氧化、抗菌的作用<sup>[3-4]</sup>。萜类物质是多种药用植物的主要活 性化合物,并且是天然化合物中规模最大种类最多的一类物质<sup>[5-6]</sup>,了解花榈木中萜类物质能够提高对 花榈木的药用认知。

植物萜类化合物主要由甲基赤藓糖磷酸酯 (MEP) 途径或甲羟戊酸 (MVA) 途径产生的二甲基烯丙基 二磷酸酯 (DMAPP) 和异戊烯基二磷酸酯 (IPP) 生成<sup>[7]</sup>。萜类化合物分子骨架是基于异戊二烯或其异构体 二甲基烯丙基焦磷酸 C5 单元构成的,根据骨架碳链长度的不同,主要分为半萜、单萜、二萜、三萜等 几种类型的萜类化合物<sup>[8]</sup>。MEP 途径仅发生在质体中,MVA 途径却分布在细胞质、内质网和过氧化物 酶体之间<sup>[9]</sup>。IPP 与 DMAPP 通过戊烯基转移酶的催化,缩合产生萜类化合物前体物质香叶酰二磷酸 (GPP)、法呢基二磷酸 (FPP) 和香叶基香叶基焦磷酸 (GGPP)。其中 GPP 是合成单萜类化合物前体物质, GGPP 是合成二萜前体物质,FPP 则是合成倍半萜和三萜化合物的前体物质<sup>[10-11]</sup>。随后通过萜类合成酶 (TPS) 和氧角鲨烯环化酶 (OSCs) 催化前体形成不同种萜类化合物,这 2 种酶的多样性使得萜类化合物的 种类有很多<sup>[12-13]</sup>。

TPS 基因家族主要分为 7 个亚家族,分别命名为 TPS-a、TPS-b、TPS-c、TPS-d、TPS-e/f、TPS-g 和 TPS-h。其中 TPS-a、TPS-b和 TPS-g 属于被子植物特有分支,这 3 个分支完全由专门的单萜、倍半萜或 二萜生物合成基因组成,主要在植物生态相互作用中起作用。TPS-d 是裸子植物特异性亚家族,主要包 含特异性代谢的裸子植物 TPSs<sup>[14]</sup>。TPS-c 和 TPS-e/f 是被子植物与裸子植物所共有的亚家族,TPS-c 分支 成员只包含"DXDD"序类,是单功能柯巴基焦磷酸合酶 (CPS)。与 TPS-c 类似,TPS-e/f 分支成员只包 含"DDXXD"序类,是单功能的贝壳杉烯合成酶 (KS)<sup>[15-16]</sup>。植物 OSC 家族是一个超基因家族,主要包 含达玛烯二醇合成酶 (DS)<sup>[17]</sup>、β-香树脂醇合成酶 (β-AS)<sup>[18]</sup>、α 香树脂醇合成酶 (α-AS)<sup>[19]</sup>和羽扇豆醇合 成酶 (LUS)<sup>[20]</sup>。近些年来对萜类的研究越来越广泛,但是豆科植物花榈木中的萜类化合物种类以及生物 合成情况却未见报道。

本研究利用代谢组学测序数据首次对花榈木中的萜类物质进行鉴定,并结合转录组测序结果,利用 生物信息学方法分析花榈木中萜类的主要合成情况,分析相关基因在花榈木不同组织部位中的表达水 平,为研究花榈木的药用活性物质以及花榈木中萜类的代谢调控途径奠定基础。

1 材料与方法

#### 1.1 材料

2021 年 4 月 23 日于华南农业大学 (中国广州) 苗圃中采集 3 年生花榈木新鲜的根、茎、树皮、老叶和幼叶 5 个组织部位样品,每样品重复 4 次,纯净水冲洗干净,立即在液氮中冷冻,并分成 2 个部分, 一部分立即在-80 ℃ 下保存用于总 RNA 提取,另一部分在真空下冷冻干燥用于代谢物提取。

1.2 代谢组测定

将采摘的根、茎、树皮、老叶和幼叶样品冻干磨粉后,精确称量各个样品粉末 0.5g,分别用 5

mL体积分数为 80%的 HPLC 级甲醇提取过夜,其中内标为 1 µmol·L<sup>-1</sup>的白杨素。4 ℃下 12 000 g 离心 30 min,取上清液装入样品瓶中,进行超高效液相色谱质谱 (UHPLC Q-TOF/MS)分析。为了分离花榈木 不同器官的代谢物,采用超高效液相色谱串联质谱法 (Q Exactive Plus)进行测定。为保证平行试验每个 样品平行 4 针。将 1 µL 样品进入 2.1 mm×100.0 mm、1.9 µm 粒径的 Hypersil GOLD 色谱柱,柱温 30 ℃,流速 0.3 mL·min<sup>-1</sup>。流动相 A 为体积分数 0.1%的 HPLC 级甲酸,流动相 B 是 HPLC 级乙腈。梯度 洗脱: 0~2.0 min, 0~10% 流动相 B; 2.0~10.0 min, 10%~55% 流动相 B; 10.0~10.1 min, 55%~80% 流动 相 B; 10.1~13.0 min, 80% 流动相 B; 13.0~14.0 min, 80%~95% 流动相 B; 14.0~18.0 min, 95%~10% 流 动相 B。全扫描质谱和数据关联扫描 (dd-MS2)的分辨率分别设置为 70 000 和 17 500,在正负离子模式 下均使用加热的 ESI 源。在正负模式下,喷雾电压被设定为 3.5 和 3.2 kV。毛细管温度被设定为 320 ℃,辅助气体加热器的温度被设定为 350 ℃。

使用 Compound Discoverer 3.2 对原始数据进行分析,然后根据综合分子量、质荷比 (*m/z*)、保留时间 (RT) 和二级光谱,与代谢物数据库 mzCloud、mzVault 和 Chemspider 进行对比。

#### 1.3 RNA 提取与文库构建

总 RNA 由 RNAprep Pure Assay 试剂盒提取, RNA 质量用 NanoPhotometer®分光光度计测定。为了 进行转录组测序,样品被送到百迈客生物技术公司。库的制备在 Illumina Hiseq 2000 平台上进行测序, 并产生成对的读长 (reads)。首先,使用内部的 perl 脚本处理 fastq 原始数据,以去除含适配器的读数、 含 ploy-N 的读数和低质量读数。经过 Trinity 软件组装获得一个非冗余的基因数据库 (universal gene, unigene),并进一步使用 HMMER软件与 Pfam 数据库比对,获得 unigene 的注释信息。

#### 1.4 差异基因和基因功能分析

以常用的基因表达水平估算方法中每千个碱基转录每百万映射读取的片断 (fragments per kilobase million, FPKM) 值进行表达量统计。使用错误发现率 (false discovery rate, FDR)  $\leq 0.01$  和  $|\log_2 C_F| \geq 2$  ( $C_F$  为差异倍数, fold change) 的阈值, 写入 R 中的 DESeq 2 包, 对花榈木不同部位样本进行差异基因表 达分析<sup>[21]</sup>。使用 TBtools 构建差异显著基因 (DEGs) 的热图, 京都基因和基因组百科全书 (KEGG) 对 DEGs 的富集分析是使用 R 中的 Cluster Profiler 包进行的, 加权基因共表达网络 (WGCNA) 是在 R 包 WGCNA 的帮助下进行的,并使用相关 P 值建立统计学意义。

### 2 结果与分析

#### 2.1 代谢组数据分析

根据总离子流图所得到的化合物精确相对分子质量、出峰时间、二级碎片信息,与mzCloud和mzVault数据库比对,共鉴定出15种萜类化合物(表1)。为了分析花榈木不同组织部位中萜类化合物的分布情况,对5个部位的萜类化合物的相对含量进行了热图分析(图1A),发现不同部位的萜类相对含量差别较大。能明显看出幼叶、老叶与根中萜类相对含量较高,与之相反的是茎与皮中相对含量比较少,并发现甘草次酸 (enoxolone)、齐墩果酸 (oleanolic acid)、芳樟醇 (linalool)和二氢丹参酮 I (dihydrotanshinone I)等具有抗氧化、抗炎免疫调节、抗心血管疾病和镇定作用的萜类化合物<sup>[22]</sup>,表明萜类化合物可能是花榈木中主要的药用活性物质。

#### 2.2 转录组数据分析

为了筛选花榈木中与萜类生物合成相关的潜在基因,利用 RNA-seq 对花榈木 5 个部位的所有转录组进行分析。使用 Illumina Hiseq 2000 平台进行测序,在去除低质量和短的读数后,总共获得了 1.9~2.4 M 的高质量待分析数据,共获得 96 302 个最长转录本,并与 NR、eggNOG、TrEMBL、Pfam、SwissProt、KEGG、COG、KOG及 GO 等 9 个数据库比对,确定编码序类有 47 809 条,并对其进行功能注释。计算出 FPKM 值,代表每个基因的表达水平。对花榈木 5 个部位的转录组进行差异基因分析,共得到 13 840 个差异显著基因 (DEGs)。在差异基因分析中发现,与其他 3 个部位样本相比,叶的特异差异基因 更多,茎中的特异差异基因最少 (图 2)。

#### 2.3 参与萜类生物合成相关基因表达分析

根据其他物种萜类生物合成信息,花榈木萜类物质合成途径大致可分为萜类前体物质的合成、萜骨

Table 1      Mass spectral information of terpenoids in O. henryi								
名称	分子式	质荷比	保留时间/min	峰面积				
				根	茎	老叶	皮	幼叶
齐墩果酸	C30H48O3	456.360 87	13.411	264 762 138	689 614	4 392 055	4 501 446	1 185 143
二氢丹参酮 I	$C_{18}H_{14}O_3$	278.093 90	7.125	293 469 237	564 453 110	77 252 839	415 872 759	45 119 942
Walleminone	$\mathrm{C_{15}H_{24}O_{3}}$	252.172 21	12.877	27 142 310	25 925 485	38 139 793	90 127 128	102 273 216
甘草次酸1	$C_{30}H_{46}O_4$	470.338 67	9.011	10 782 521	10 749 752	19 745 477	26 015 732	10 142 324
熊果酸	$C_{30}H_{48}O_3$	456.359 91	13.631	40 193 042	1 258 439	318 763 532	2 194 119	2 055 746
京尼平苷酸	$C_{16}H_{22}O_{10}$	374.121 60	1.932	16 460 987	112 195 727	49 129 697	97 661 105	52 007 928
甘草次酸2	$C_{30}H_{46} O_4$	470.338 61	13.353	6 685 919	1 285 086	34 589 990	2 917 822	2 213 863
劳丹醇酸	$C_{20}H_{36}O_{3}$	324.266 84	12.857	1 810 075	2 772 437	427 160	3 608 483	8 629 568
芳樟醇	C10H18O	154.135 49	12.167	2 423 647	2 992 083	1 182 137	4 399 836	13 100 226
脱落酸	$C_{15}H_{20}O_4$	264.136 42	8.304	1 324 567	8 782 319	6 127 101	13 051 696	24 808 273
积雪草苷	$C_{48}H_{78}O_{19}$	958.511 41	8.844	4 835 086	22 449 131	116 659 465	169 326 989	9 393 498
Lagochilin	$C_{20}H_{36}O_5$	356.255 49	12.385	1 032 268	511 236	837 844	660 680	27 515 252
木香烃内酯	$C_{15}H_{20}O_2$	232.146 14	9.727	600 726	697 407	6 046 546	4 143 536	4 973 618
鸡蛋花素	$C_{15}H_{14}O_6$	290.078 91	9.824	1 179 709	9 187 944	17 465 820	14 830 113	618 679
Prespatane	C <sub>15</sub> H <sub>24</sub>	204.187 43	12.689	15 917 262	1 479 579	2 907 710	4 359 663	8 310 627





A. 萜类化合物积累情况热图。B. 萜类化合物合成相关基因TPS、OSC的FPKM热图。使用"scale"函数将数据按行归一化。 用深红色和深蓝色表示代谢物的积累水平和基因的相对表达从高到低。

花榈木不同组织中萜类化合物积累和 TPS、OSC 基因表达量的热图

Figure 1 Heat map of terpenoid accumulation and TPS and OSC gene expression in different tissues of O. henryi

架的合成和后修饰 3 个阶段,其中萜类化合物的前体 GPP、FPP 和 GGPP 主要来自植物的 MVA 和 MEP 途径。MVA 通路共鉴定到 8 个基因,包括 2 个乙酰辅酶 A 乙酰转移酶 (AACT, c68943.graph c2, c73242.graph c0)、2个羟甲基戊二酰辅酶 A 合成酶 (HMGS, c57495.graph c0, c70166.graph c1)、1个羟 甲基戊二酰辅酶 A 还原酶 (HMGR, c68938.graph c1)、1个甲羟戊酸激酶 (MK, c70988.graph c0)、1个磷 酸甲羟戊酸激酶 (PMK, c71400.graph c0)、1个甲羟戊酸焦磷酸脱羧酶 (MVD, c76163.graph c2); MEP 通路共鉴定到 10 个基因,包括 4 个 1-脱氧-D-木酮糖-5-磷酸合酶 (DXS, c62607.graph c0, c70494.graph c2, c75298.graph c5, c75577.graph c0)、1个1-脱氧-D-木酮糖-5-磷酸还原酶 (DXR, c71705.graph c2), 1个4-焦磷酸胞苷-2-甲基-D-赤藓醇激酶 (CMK, c73279.graph c4)、1个2-甲基-D-赤藓醇-2,4-环焦磷酸合酶 (MDS, c62387.graph c0)、1个羟甲基丁烯基-4-焦磷酸合酶 (HDS, c56719.graph c0)、2个羟甲基丁烯基-4-磷酸还原酶 (HDR, c29394.graph c0, c70095.graph c0)。为了更直观地比较萜类生物合成途径主要基因 在不同组织中的表达情况,采用 FPKM 值对花榈木不同组织部位的合成途径基因表达量作图。发现参



不同颜色代表不同组织间的差异基因,重叠区域代表不同比较间具有共同调控模式的差异基因。

图 2 花榈木 5 个组织中的差异基因韦恩图 Figure 2 Venn diagram of differential genes in five organs of *O. henryi* 

与 MVA 途径的基因在幼叶中表达量普遍较高, MEP 途径相关基因在老叶中表达量较高, 与萜类相对含量比较对应,可能与叶中的萜类化合物的积累有关(图 3)。还检测出 2 个异戊烯基焦磷酸异构酶 (IPPI, c31804.graph\_c0, c89450.graph\_c0), 它们在叶中的表达量并不高,可能这些基因与萜类化合物相对含量并不密切相关。

MVA 和 MEP 这 2 个途径后生成萜类化合物的前体,还需要 3 个比较重要的酶参与:香叶基二磷酸合成酶 (GPPS)、法尼基二磷酸合成酶 (FPPS) 和香叶基香叶基焦磷酸合成酶 (GGPPS)。在花榈木的转录 组数据中共检测到 2 个 FPPS 基因 (c86328.graph\_c0,c75342.graph\_c0), 2 个 GPPS 基因 (c36059.graph\_c0, c76294.graph\_c0), 5 个 GGPPS 基因 (c55157.graph\_c0,c71336.graph\_c0,c65845.graph\_c2,c63610.graph\_c0,c69470.graph\_c0),表达情况如图 3 所示。2 个 GGPPS 基因 c55157.graph\_c0 和 c69470.graph\_c0 在幼叶中表达量比较高,多数二萜在幼叶中的积累量也是最多的 (图 1),可能这 2 个基因是花榈木二萜物质合成的关键酶基因。TPSs 是催化合成不同种单帖、二萜和倍半萜类化合物的酶,OSCs 可以催化合成三萜类化合物,在花榈木转录组中还鉴定出 6 个 TPSs 和 8 个 OSCs 基因,其表达谱如图 1B 所示。这些基因的表达在花榈木中出现了组织差异性,每个组织中都有其表达量高的基因,其中与 MVA 和 MEP 途径基因表达情况相似:在幼叶和老叶中表达量高的基因占比较大。这些 TPSs 和 OSCs 的组织差异表达可能是导致花榈木中萜类多样性和具有不同积累模式的主要原因。

为了进一步了解研究中 6 个 TPSs 的假设功能,根据其蛋白序列,将其与已经在拟南芥 Arabidopsis thaliana、葡萄 Vitis vinifera 和番茄 Lycopersicon esculentum 等物种中确定的 TPS 进行系统发育分析 (图 4)。 发现与之前报道相似,花榈木中的 TPSs 主要分布在 TPS-a、TPS-b、TPS-c、TPS-e 和 TPS-g 亚家族。 c70917.graph\_c0 和 c72335.graph\_c1 被分类到 TPS-b中,可能这 2 个 基因 在功能上比较类似; c64128.graph\_c1 被分到 TPS-e 亚家族,说明 c64128.graph\_c1 可能在花榈木中行使着 KS 酶功能; c73567.graph\_c1 则被分到 TPS-c 亚家族中,表明它可能有着 CPS 酶的功能。为了进一步推测花榈木 OSCs 的功能,对每个基因再次进行了美国国家生物技术信息中心数据库的比对注释,发现除 c71944.graph\_c0 基因是编码 LUC 的基因外,其他 7 个都是编码 β-AS 的基因。

#### 2.4 萜类生物合成相关基因共表达网络分析与鉴定

为进一步分析萜类化合物合成相关基因,利用 WGCNA 对确定的全部 DEGs 进行了共表达分析。这些 DEGs 被聚类为 10 个分支,每个不同颜色标记的模块代表了 1 个分支 (图 5)。模块是由具有相似表达模式的基因簇组成的,MEturquoise 模块包含的基因数量最多 (3 761 个基因),而 MEmagenta 模块的基因数量最少 (104 个基因)。其中有 7 个模块与一些特定的萜类化合物表现出明显的正相关 (相关系数> 0.80)(图 5)。除了 c63712.graph c0 基因外,共有 5 个 TPSs 和 8 个 OSCs 基因表达差异明显,推测这





使用 "scale" 函数将数据按行归一化。深红色和深蓝色表示基因的相对表达量从高到低。

图 3 花榈木萜类骨架生物合成途径

Figure 3 Terpene skeleton biosynthesis pathway of *O. henryi* 

13 个差异表达的 TPSs 和 OSCs 可能是导致花榈木 5 个组织部位萜类化合物组成差异的重要候选基因。 其中, MEturquoise 模块中共有 2 个 TPSs 和 4 个 OSCs 基因; MEblue 模块有 2 个 TPSs 和 3 个 OSCs 基因。MEturquoise 和 MEbrown 与三萜类化合物的相关性较强 (相关性系数>0.90), 而 MEblue、MEblack 和 MEpink 模块与倍半萜化合物有相关性,其中 MEblue 模块也与二萜和单萜具有相关性。二氢丹参酮 I 这个二萜化合物与 MEyellow 和 MEgrey 2 个模块都具有正相关性。这些正相关的模块表明:这些基因 在调节花榈木萜类化合物的生物合成中具有潜在的作用。

上述模块的分析结果表明: MEturquoise 模块富集了最多的 TPSs 和 OSCs,并且该模块与乌苏酸和 甘草次酸呈正相关,相关性分别为 0.95 和 0.93(图 5); MEblue 也富集到了较多的 TPSs 和 OSCs 基因, 并且与一些倍半萜和二萜成正相关,所以对这 2 个模块开展进一步的分析。在 MEturquoise 模块中发 现:基因主要富集在次生代谢和信号转导通路,共检测到 105 个转录因子可能参与 TPSs 的调控 (图 6 A 和图 6 B): bHLH、WRKY 和 MYB 家族数量最多,分别有 10、9 和 8 个,此外 C2H2、mTERF、 FAR1 和 bZIP 家族数量也达到了 6~7 个。通过对 MEblue 模块基因进行 KEGG 分析,发现基因主要富集 在 DNA 复制、同源重组和修复通路,并检测到 152 个转录因子,数量最多的是 bHLH 和 MYB 家族转 录因子 (图 7 A 和图 7 B)。为了进一步研究 MEturquoise 模块中转录因子与相应 TPSs 和 OSCs 基因的关



以鉴定的差异表达 TPS 基因的蛋白序列和已报道的 TPS 基因的蛋白序列构建系统发育树。TPS-a、TPS-b、TPS-c、TPS-g、TPS-e 和 TPS-f 家族被浅绿色、蓝色、红色、紫色、黄色和灰色覆盖。

图 4 花榈木 TPS 基因家族分析 Figure 4 Analysis of TPS gene family of O. henryi

系,构建了一个共表达的网络图,选择了与 MEturquoise 模块中 TPSs 和 OSCs 基因相关性较强的 34 个 转录因子 (边缘权重≥0.4),包括 MYB、WRKY、bHLH、C3H、DBB、HB-HD-ZIP 和一些其他家族的转录因子 (图 8)。最终利用 Cytoscape 插件 CytoHubba 的 Degree 算法分析识别到了 6 个转录因子,即 HB-HD-ZIP (c64527.graph\_c1)、GRF (c76195.graph\_c0)、DBB (c66970.graph\_c2)、DBB (c75593.graph\_c0)、HB-HD-ZIP (c63393.graph\_c0)和 C3H (c70385.graph\_c1)。这些转录因子在花榈木中可能与萜类化合物的合成基因有着密切的关系。

3 讨论

目前,天然植物次生代谢物已被广泛应用于抗癌药和治疗感染性疾病的药物<sup>[23]</sup>。根据合成起始分子 不同,植物次生代谢物可以分为生物碱、萜类、苯丙烷类三大类化合物<sup>[24]</sup>。萜类一直是天然产物中重要 的药用化合物,具有多种药理活性,如紫杉醇可以抗肿瘤,青蒿素属于抗疟疾特效药物,雷公藤内酯能 够抗炎等<sup>[25-26]</sup>,但药用木本植物花榈木中的萜类还没有过报道。本研究利用代谢组学技术分析了花榈木 中萜类物质在不同部位的积累情况,共检测出比较确定的15种萜类化合物,并发现了甘草次酸、齐墩 果、芳樟醇和二氢丹参酮 I 等具有抗氧化、抗炎免疫调节和镇定等药用活性的萜类化合物<sup>[27]</sup>。通过对花 榈木不同组织部位的萜类代谢物热图分析发现,萜类化合物积累具有明显组织特异性,主要积累在幼叶







Figure 5 Co-expression network analysis of genes and terpenoids in O. henryi

和老叶中,其他部位中积累较少,可能为了保护叶片免遭危害<sup>[28]</sup>。今后在提取和分析花榈木药用物质时 应该更多利用它的叶片。

近些年来,RNA-Seq高通量测序技术被越来越多地应用于药用植物基因信息解读、新基因发掘与基因功能研究中。人们已对药用植物连翘 Forsythia suspense、银杏 Ginkgo biloba、款冬 Tussilago farfara 等进行了转录组的研究,获得了大量有用的基因信息<sup>[29-30]</sup>。这使得阐明药用植物中活性物质的合成及积累规律成为可能,为增加次生代谢物积累、改善药用植物品质提供更多途径。本研究采用 RNA-Seq 对花榈木 5个不同组织部位进行无参转录组分析,共获得 96 302 个,在经过与数据库比对后共注释了47 809条转录本。利用 FPKM 值对基因表达量进行分析比较,共获得显著差异基因 13 840 个。韦恩图分析发现大量基因表达具有组织特异性,并通过注释信息在差异显著的基因中共鉴定出 49 个与萜类化合物生物合成相关的基因,包括 29 个萜类骨架生物合成途径的酶基因,6 个单帖、倍半萜和二萜生物合成酶基因,8 个三萜生物合成酶基因以及 6 个可能参与萜类生物合成调控的转录因子,对花榈木的萜类合成有了初步的认识,为后续研究提供了信息资源。

先前的研究表明: MVA 途径在许多植物的三萜生物合成中起主导作用,如人参 Panax ginseng、三七 P. notoginseng 和茶树 Camellia sinensis 等植物<sup>[31-32]</sup>, MEP 途径通常有助于单萜类化合物和二萜类化合物的生物合成<sup>[33]</sup>。但是本研究利用热图对萜类生物合成基因的表达分析发现: MVA 途径相关酶基因大



A. MEturquoise 模块基因的 KEGG 分析。纵坐标是富集通路的名称; 横坐标是富集因子数, 数字越大, 代表通路富集可能性越 大; q 值是通路富集分析校正后的 p 值, q 值越小通路富集可能性越高,所以图中越靠近右侧越红的通路,富集可能性越高。B. MEturquoise 模块富集到的转录因子。纵坐标是不同的转录因子家族,不同颜色的柱状图代表不同的转录因子家族; 横坐标是转 录因子的数量。



Figure 6 MEturquoise module gene analysis



A. MEblue 模块基因的 KEGG 分析。纵坐标是富集通路的名称;横坐标是富集因子数,数字越大,代表通路富集可能性越大; q 值是通路富集分析校正后的 p 值,q 值越小通路富集可能性越高;所以图中越靠近右侧越红的通路,富集可能性越高。 B. MEblue 模块富集到的转录因子。纵坐标是不同的转录因子家族,不同颜色的柱状图代表不同的转录因子家族;横坐标是转录因子的数量。



多数在幼叶中表达量较高,MEP途径多数相关基因则集中在老叶中表达。根据相应 FPPS、GPPS 和GGPPS 酶基因的表达情况,以及花榈木中三萜主要在老叶中积累,而二萜、倍半萜和单萜在幼叶中积 累较多的情况,推测在花榈木中三萜类化合物前体可能主要由 MEP 途径生成,而二萜、单帖和倍半萜 的前体物质则主要由 MVA 途径提供。由图 4 看出:花榈木的 TPS 基因在进化上相对于各种模式植物来 说是相对分离的,也进一步印证了其萜类的合成具有特殊性。也有可能由于转录组学的限制,部分基因



1个菱形节点代表1个酶基因,1个圆形节点代表1个转录因子。 图 8 萜类化合物合成相关酶基因与转录因子相关性网络图 Figure 8 Network diagram of terpenoid synthesis-related enzyme genes and transcription factor correlation

没有检测到,且研究中还没有检测到所有的萜类物质,从而导致判断出现误差,所以后续还需要大量的 试验来验证其功能,判断花榈木中萜类的具体合成情况。

基因的转录调控一直是植物代谢研究领域的热点。对不同植物的研究表明:参与萜类生物合成的转录因子主要分布在 bHLH、AP2/ERF、bZIP 和 WRKY 家族中,如在西洋参 *P. quinquefolius* 中转录因子 *PqWRKY*1 是三萜人参皂苷生物合成相关的正调节因子<sup>[34]</sup>,在艾叶 *Artemisia argyi* 中 *AarbHLHs* 的基因表达与 1,8-桉树烯或 β-石竹烯的含量变化呈显著相关<sup>[35]</sup>。本研究利用 WGCNA 对差异基因和萜类化合物进行了相关性分析,筛选出一些可能对萜类生物合成的关键酶基因表现重要调控作用的转录因子家族,如 MYB、WRKY、bHLH 和 HB-HD-ZIP 转录因子。有研究预测传统中药走马胎 *Ardisia kteniophylla* 中 AP2/ERF、MYB、WRKY 和 bHLH 转录因子可能调控萜类合成,预测赤桉 *Eucalyptus camaldulensis* 中 WRKY、MYB、NAC 和 bHLH 转录因子对萜类生物合成中的关键酶基因表现出重要的调控作用<sup>[36]</sup>。这 和本研究预测的结果基本相同。经过分析进一步筛选出 6 个转录因子处于共表达网络的中心位置,推测 这些候选转录因子可能调控了萜类化合物的生物合成,后续还需要进一步的试验证明。

4 结论

本研究通过代谢组与转录组学的分析,发现花榈木叶中萜类相对含量最高,并鉴定出 49 个与萜类 化合物生物合成相关的基因。预测了可能调控萜类化合物生物合成的上游转录因子。本研究为花榈木资 源活性成分萜类化合物的积累状况、生物合成及调控提供了大量的信息,弥补了花榈木萜类合成研究中 的空白,为进一步开展花榈木的主要药用活性物质研究提供基础。

#### 5 参考文献

[1] 桂平,龙鹏.珍稀树种花榈木研究进展[J].贵州农业科学, 2021, 49(7): 98-106.

979

GUI Ping, LONG Peng. Research progress on rare tree species of *Ormosia henryi* [J]. *Guizhou Agricultural Science*, 2021, **49**(7): 98 – 106.

- [2] 张琳婧, 周文娟, 倪林, 等. 红豆属植物化学成分及其药理活性研究进展[J]. 中草药, 2021, 52(14): 4433 4442. ZHANG Linjing, ZHOU Wenjuan, NI Lin, *et al.* A review on chemical constituents and pharmacological activities of *Ormosia* [J]. *Chinese Traditional and Herbal Drugs*, 2021, 52(14): 4433 - 4442.
- [3] 翟大才,姚建林,王文娟,等. 红豆树叶挥发油化学成分及其抗氧化和抑菌活性研究[J]. 天然产物研究与开发, 2019, 31(5): 814-819.
  ZHAI Dacai, YAO Jianlin, WANG Wenjuan, *et al.* Chemical constituents of the volatile oil from *Ormosia hosiei* leaves and

its antioxidant and antimicrobial activity [J]. *Natural Product Research and Development*, 2019, **31**(5): 814 – 819.

- [4] 倪斌,张伟,符杰雄,等. 花梨木叶挥发油化学成分的GC-MS分析[J]. 广东林业科技, 2012, 28(2): 59-62.
  NI Bin, ZHANG Wei, FU Jiexiong, et al. GC-MS analysis of chemical constituents of the volatile oil from leaves of Ormosia henryi Prain [J]. Guangdong Forestry Science and Technology, 2012, 28(2): 59-62.
- [5] ZHOU Fei, PICHERSKY Eran. More is better: the diversity of terpene metabolism in plants [J]. Current Opinion in Plant Biology, 2020, 55: 1 – 10.
- [6] HOLOPAINEN J K, GERSHENZON J. Multiple stress factors and the emission of plant VOCs [J]. Trends in Plant Science, 2010, 15(3): 176 – 184.
- [7] ABBAS F, KE Yanguo, YU Rangcai, *et al.* Volatile terpenoids: multiple functions, biosynthesis, modulation and manipulation by genetic engineering [J]. *Planta*, 2017, 246(5): 803 – 816.
- [8] 陈妤, 朱沛煌, 李荣, 等. 植物异戊烯基转移酶研究进展[J]. 生物技术通报, 2021, 37(2): 149-161.
  CHEN Yu, ZHU Peihuang, LI Rong, *et al.* Research progress of plant prenyltransferases [J]. *Biotechnology Bulletin*, 2021, 37(2): 149-161.
- [9] VRANOVA E, COMAN D, GRUISSEM W. Network analysis of the MVA and MEP pathways for isoprenoid synthesis [J]. *Annual Review of Plant Biology*, 2013, **64**: 665 700.
- [10] WEI Guo, TIAN Peng, ZHANG Fengxia, *et al.* Integrative analyses of nontargeted volatile profiling and transcriptome data provide molecular insight into VOC diversity in cucumber plants (*Cucumis sativus*) [J]. *Plant Physiology*, 2016, **172**(1): 603-618.
- [11] LIU Songyu, SHAN Bingqi, ZHOU Xiaomiao, et al. Transcriptome and metabolomics integrated analysis reveals terpene synthesis genes controlling linalool synthesis in grape berries [J]. Journal of Agricultural and Food Chemistry, 2022, 70(29): 9084 – 9094.
- [12] DEGENHARDT J, KOLLNER T G, GERSHENZON J. Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants [J]. *Phytochemistry*, 2009, 70(15/16): 1621 – 1637.
- [13] CHEN Feng, THOLL D, BOHLMANN J, et al. The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom [J]. Plant Journal, 2011, 66(1): 212 – 229.
- [14] 朱沛煌, 陈妤, 季孔庶. 松科植物萜类合成酶及其基因家族研究进展[J]. 南京林业大学学报(自然科学版), 2021, 45(3): 233-244.

ZHU Peihuang, CHEN Yu, JI Kongshu. A review of terpene synthases and genes in Pinaceae [J]. *Journal of Nanjing Forestry University (Natural Sciences Edition)*, 2021, **45**(3): 233 – 244.

- [15] PARVEEN I, WANG Mei, ZHAO Jianping, *et al.* Investigating sesquiterpene biosynthesis in *Ginkgo biloba*: molecular cloning and functional characterization of (E, E)-farnesol and α-bisabolene synthases [J]. *Plant Molecular Biology*, 2015, 89(4/5): 451 462.
- [16] MARTIN D M, AUBOURG S, SCHOUWEY M B, et al. Functional annotation, genome organization and phylogeny of the grapevine (*Vitis vinifera*) terpene synthase gene family based on genome assembly, FLcDNA cloning, and enzyme assays [J/OL]. *BMC Plant Biology*, 2010, 10: 226[2022-10-28]. doi: 10.1186/1471-2229-10-226.
- [17] HU Zhongfeng, GU Andi, LIANG Lanli, et al. Construction and optimization of microbial cell factories for sustainable production of bioactive dammarenediol- II glucosides [J]. Green Chemistry, 2019, 21(12): 3286 – 3299.
- [18] YIN Yanchao, ZHANG Xiaodong, GAO Zhiqiang, *et al.* Over-expressing root-specific β-amyrin synthase gene increases glycyrrhizic acid content in hairy roots of *Glycyrrhiza uralensis* [J]. *Chinese Herbal Medicines*, 2019, **11**(2): 192 – 199.
- [19] LUCHNIKOVA N A, GRISHKO V V, IVSHINA I B. Biotransformation of oleanane and ursane triterpenic acids [J/OL]. *Molecules*, 2020, 25(23): 5526[2022-10-28]. doi: 10.3390/molecules25235526.
- [20] BACHORIK J, URBAN M. Biocatalysis in the chemistry of lupane triterpenoids [J/OL]. Molecules, 2021, 26(8):

2271 [2022-10-28]. doi: 10.3390/molecules26082271.

- [21] LIU Yi, YIN Qi, DAI Baojia, *et al.* The key physiology and molecular responses to potassium deficiency in *Neolamarckia cadamba* [J/OL]. *Industrial Crops and Products*, 2021, **162**: 113260[2022-10-28]. doi. 10.1016/j. inderop. 2021.113260.
- [22] DARSHANI P, SARMA S S, SRIVASTAVA A K, et al. Anti-viral triterpenes: a review [J]. Phytochemistry Reviews, 2022, 21(6): 1761 – 1842.
- [23] CRAGG G M, NEWMAN D J. Nature: a vital source of leads for anticancer drug development [J]. *Phytochemistry Reviews*, 2009, 8(2): 313 331.
- [24] REJEB I B, PASTOR V, MAUCH-MANI B. Plant responses to simultaneous biotic and abiotic stress: molecular mechanisms [J]. *Plants*, 2014, 3(4): 458 – 475.
- [25] WAN Yan, LIU Dong, XIA Jia, *et al.* Ginsenoside CK, rather than Rb1, possesses potential chemopreventive activities in human gastric cancer via regulating PI3K/AKT/NF-κB signal pathway [J/OL]. *Frontiers in Pharmacology*, 2022, 13: 977539[2022-10-28]. doi. 10.3389/fphar. 2022.977539.
- [26] STASHENKO E, GONZALE A F, MARTINEZ J R, et al. Hallazgo de diclofenaco en un producto fitoterapéutico a base de caléndula comercializado en Colombia [J]. Salud UIS, 2020, 52(3): 261 – 284.
- [27] DIOMEDE L, BEEG M, GAMBA A, *et al.* Can antiviral activity of licorice help fight COVID-19 infection? [J/OL]. *Biomolecules*, 2021, **11**(6): 855[2022-10-18]. doi:10.3390/biom11060855.
- [28] HUANG Ancheng, OSBOURN A. Plant terpenes that mediate below-ground interactions: prospects for bioengineering terpenoids for plant protection [J]. Pest Management Science, 2019, 75(3): 2368 – 2377.
- [29] SUN Luchao, RAI A, RAI M, et al. Comparative transcriptome analyses of three medicinal Forsythia species and prediction of candidate genes involved in secondary metabolisms [J]. Journal of Natural Medicines, 2018, 72(4): 867 – 881.
- [30] GUO Ying, WANG Tongli, FU Fang, *et al.* Temporospatial flavonoids metabolism variation in *Ginkgo biloba* leaves [J/OL]. *Frontiers in Genetics*, 2020, **11**: 589326[2022-10-28]. doi: 10.3389/fgene.2020.589326.
- [31] THIMMAPPA R, GEISLER K, LOUVEAU T, *et al.* Triterpene biosynthesis in plants [J]. *Annual Review of Plant Biology*, 2014, **65**: 225 257.
- [32] CHEN Cong, ZHU Huanqing, KANG Jiaxin, *et al.* Comparative transcriptome and phytochemical analysis provides insight into triterpene saponin biosynthesis in seeds and flowers of the tea plant (*Camellia sinensis*) [J/OL]. *Metabolites*, 2022, 12(3): 204[2022-10-28]. doi. 10.3390/metabo12030204.
- [33] ZHOU Hanchen, SHAMALA L F, YI Xingkai, et al. Analysis of terpene synthase family genes in *Camellia sinensis* with an emphasis on abiotic stress conditions [J/OL]. *Scientific Reports*, 2020, 10(1): 933[2022-10-28]. doi: 10.1038/s41598-020-57805-1.
- [34] SUN Yongzhen, NIU Yunyun, XU Jiang, et al. Discovery of WRKY transcription factors through transcriptome analysis and characterization of a novel methyl jasmonate-inducible PqWRKY1 gene from Panax quinquefolius [J]. Plant Cell, Tissue and Organ Culture, 2013, 114(2): 269 – 277.
- [35] YI Xiaozhe, WANG Xingwen, WU Lan, et al. Integrated analysis of basic helix loop helix transcription factor family and targeted terpenoids reveals candidate AarbHLH genes involved in terpenoid biosynthesis in Artemisia argyi [J/OL]. Frontiers in Plant Science, 2022, 12: 811166[2022-10-28]. doi: 10.3389/fpls.2021.811166.
- [36] ZHAN Ni, HUANG Lanhong, WANG Zhen, et al. Expression of genes encoding terpenoid biosynthesis enzymes during leaf development of Eucalyptus camaldulensis [J]. Biologia Plantarum, 2022, 66: 146 – 154.