

利用高通量测序挖掘柠檬醛型野草香中的香叶醇脱氢酶基因

徐玉琪, 梁春茂, 成佳蕙, 李园园, 普春霞*

(云南中医药大学中药学院, 云南昆明 650500)

摘要: 以药食两用植物柠檬醛型野草香为原料进行精油提取及 GC-MS 成分测定, 并利用 Illumina NovaSeq 6000 高通量测序获得其转录组数据信息, 挖掘其精油合成关键酶基因。结果表明, 该野草香种质柠檬醛含量在 80% 以上, 高通量测序后共获得 22.08 GB 数据, 147 186 998 条高质量序列, 利用 Trinity 组装获得 33 720 条 unigenes, 平均长度 1 799 bp, 均在 7 大公共数据库中成功获得功能注释。通过 KEGG 代谢通路分析显示有 2 条单萜代谢相关途径, 为萜类骨架生物合成 (编号为 ko00900) 和单萜类生物合成 (编号为 ko00902), 相关 unigenes 分别有 109 条和 32 条; 扩增获得 1 143 bp 的香叶醇脱氢酶基因 (GeDH), 编码 380 个氨基酸, 具有保守的 Zn²⁺、NADP 结合位点和类似于紫苏属植物的 Zn1 结构基团, 为进一步研究香蒿属植物特色精油中单萜生物合成途径解析寻找新的切入点。GeDH 基因作为调控柠檬醛精油合成的关键酶, 其挖掘为天然柠檬醛精油的开发利用提供了一定的帮助, 为拓宽食用调味品、天然食品抗菌剂的种类奠定了基础。

关键词: 高通量测序; 野草香; 柠檬醛; 精油; 香叶醇脱氢酶

文章编号: 1673-9078(2024)12-57-66

DOI: 10.13982/j.mfst.1673-9078.2024.12.1390

Mining the Geraniol Dehydrogenase Gene in a Citral Chemotype of *Elsholtzia cypriani* via High-throughput Transcriptome Sequencing

XU Yuqi, LIANG Chunmao, CHENG Jiahui, LI Yuanyuan, PU Chunxia*

(College of Traditional Chinese Medicine, Yunnan University of Chinese Medicine, Kunming 650500, China)

Abstract: The essential oil of a medicinal and edible citral chemotype of *Elsholtzia cypriani* was extracted and its components were subsequently identified using GC-MS. Transcriptome data were obtained using Illumina high-throughput sequencing (NovaSeq 6000), and key enzyme-coding genes involved in essential oil synthesis were identified. The results showed that the citral content of the *E. cypriani* germplasm exceeded 80%. After high-throughput sequencing, a total of 22.08 GB of data and 147 186 998 clean reads were obtained. In total, 33 720 unigenes with an average length of 1799 bp were generated using the Trinity assembly, and all were successfully functionally annotated in seven major open databases. KEGG metabolic pathway analysis revealed that the unigenes were enriched in two monoterpene pathways, with 109 unigenes

引文格式:

徐玉琪, 梁春茂, 成佳蕙, 等. 利用高通量测序挖掘柠檬醛型野草香中的香叶醇脱氢酶基因[J]. 现代食品科技, 2024, 40(12):57-66.

XU Yuqi, LIANG Chunmao, CHENG Jiahui, et al. Mining the geraniol dehydrogenase gene in a citral chemotype of *Elsholtzia cypriani* via high-throughput transcriptome sequencing [J]. Modern Food Science and Technology, 2024, 40(12): 57-66.

收稿日期: 2023-11-21

基金项目: 国家自然科学基金项目 (32260102); 云南省院士专家工作站项目 (202305AF150030)

作者简介: 徐玉琪 (1999-), 女, 硕士研究生, 研究方向: 中药资源开发与利用, E-mail: 1403985236@qq.com

通讯作者: 普春霞 (1977-), 女, 博士, 副教授, 研究方向: 中药资源开发与利用, E-mail: puchunxia@ynucm.edu.cn

related to terpenoid backbone biosynthesis (No. ko00900) and 32 unigenes associated with monoterpenoid biosynthesis (No. ko00902). The 1143 bp geraniol dehydrogenase (*GeDH*) gene encoding 380 amino acids was amplified. The *GeDH* protein contained conserved Zn²⁺- and NADP-binding sites and a Zn1 structural group similar to that in *Perilla*. Our findings provide a new breakthrough point for further research on the biosynthesis pathway of monoterpenoids in the characteristic essential oils of *Elsholtzia*. The mining of *GeDH*, which encodes a key enzyme that regulates the synthesis of citral essential oils, will prove invaluable in the development and utilization of natural citral essential oils and lays the foundation for increasing the variety of edible flavored oils and natural antibacterial agents for use in food products.

Key words: high-throughput sequencing; *Elsholtzia cypriani* (Pavolini) S. Chow ex P. S. Hsu; citral; essential oils; geraniol dehydrogenase

野草香 (*Elsholtzia cypriani* (Pavolini) S. Chow ex P. S. Hsu) 为唇形科香薷属植物, 又名木姜花、野香薷、鱼香菜等, 主产于云南、贵州、湖南、广西等地^[1]。全草或叶入药, 有特殊清香, 味辛, 性凉, 具有清热发表, 解毒截疟的功效, 花穗可止血或作调味使用^[1,2]。在云南民间, 野草香可作种子油^[3]和蔬菜^[4]食用。香薷属植物大部分可药食两用^[5], 是秋冬季重要的蜜源植物, 其精油主要为单萜和倍半萜化合物, 具有作为食品天然抗菌剂和食品防腐剂的潜能^[6]。野草香精油特色成分与同属植物相同, 一般为去氢香薷酮^[3,7], 但分布在贵州地区的野草香被称为“木姜花”, 其精油主要成分为柠檬醛^[8], 课题组在接近贵州的云南昭通地区也发现了此类型的野草香, 在当地其被作为蔬菜进行栽培食用。

柠檬醛由顺式柠檬醛 (橙花醛) 和反式柠檬醛 (香叶醛) 2 种异构体组成, 因其具有浓郁的柠檬香味, 被广泛运用于食品、香精、香料的制造和美容行业中。柠檬醛具有抑菌^[9]、抗炎^[10]、抗氧化^[11]、抗肿瘤^[12]等生物活性, 还有平喘抗过敏^[13]、治疗心血管疾病^[14]等作用。虽然柠檬醛已有化学合成的方法, 但鉴于天然香精在安全性和价格上与合成香精存在差异, 以柠檬醛为主成分的野草香精油品种具有开发前景。同时, 与同样富含柠檬醛的山鸡椒^[15] (多年生落叶灌木或小乔木植物, 主产于广东、广西、贵州等地) 和香茅草^[16] (多年生草本植物, 广东、海南、台湾等热带地区栽培) 相比, 柠檬醛型野草香作为一年生草本植物, 主要分布在云南、贵州等地, 更适宜在滇中地区进行大规模栽培。

在课题组之前的研究中推测出一条尚未被证实的香薷属特色单萜成分代谢通路^[17], 该通路以香叶醇合酶 (Geraniol Synthases, GES) 催化香叶基二磷酸 (Geranyl Diphosphate, GPP) 形成的线形单萜-香叶醇为起点, 香叶醇由香叶醇脱氢酶 (Geraniol Dehydrogenase, *GeDH*) 氧化为柠檬醛 (Citral)^[18],

柠檬醛经未被证实的细胞色素 P450 (Cytochrome P450, CYP) 基因转化为玫瑰呋喃、去氢香薷酮, 以香薷酮为最终产物。在香薷属特色单萜成分代谢通路研究中, 作为起始作用酶的香叶醇合酶已被成功克隆^[17], 而香叶醇脱氢酶尚未被发现。

柠檬醛型野草香作为药食两用植物的研究大多集中在精油上, 而基因资源研究鲜有报道。本研究通过对柠檬醛型野草香进行转录组测序及数据分析, 初步挖掘了柠檬醛型野草香精油合成关键酶, 为香薷属植物单萜类生物合成途径解析以及功能验证的研究提供了基因资源和理论基础, 同时为天然柠檬醛精油的开发利用提供基础资源。

1 材料与方法

1.1 试验材料

1.1.1 原料

柠檬醛型野草香种子采自云南省昭通鲁甸 (纬度: 27°11'59"; 经度: 103°32'26"; 海拔: 1 923 m); 原植物由云南中医药大学普春霞副教授鉴定为野草香 [*E. cypriani* (Pavolini) S. Chow ex P. S. Hsu]。2022 年 5 月于云南中医药大学呈贡校区 (纬度: 24°50'26"; 经度: 102°49'11"; 海拔: 1 916 m) 进行播种栽培, 2022 年 8 月采集新鲜嫩叶于液氮中进行速冻后, 转移至 -80 °C 冰箱保存备用。

1.1.2 试剂与仪器

柠檬醛标准品 GC≥95%, 上海源叶生物科技有限公司; 环己烷、无水硫酸钠 (色谱纯), 天津市致远化学试剂有限公司; 蒸馏水; MiniBEST Plant RNA Extraction Kit 试剂盒、PrimeScriptTMII 1st Strand cDNA Synthesis Kit 试剂盒、DL2000 DNA Maker, 日本 TaKaRa 公司; 2X FidCycle Fast 高保真 PCR Mix (含蓝染料), 生工生物工程 (上海) 有限公司;

CP214 电子天平, 上海奥豪斯仪器有限公司;

7890A-5975C 气相色谱-质谱联用仪, 美国安捷伦科技有限公司; Illumina NovaSeq 6000 测序平台, 武汉迈特维尔生物科技有限公司; AB Veriti PCR 扩增仪、基础电泳电流仪, 美国伯乐公司 (Bio-Rad Laboratories)。

1.2 试验方法

1.2.1 野草香精油提取

参照 2020 版中国药典 (一部) 附录 XD 采取水蒸气蒸馏法提取柠檬醛型野草香精油: 取新鲜叶片 10 g (无需破碎), 提取温度 80 °C, 提取时间 6 h, 设置三个生物学平行重复。收集精油, 用无水硫酸钠进行过滤干燥, 并用环己烷萃取, 得到具有浓郁香味的精油, 置于棕色瓶中 4 °C 保存^[7]。柠檬醛野草香叶片的精油得率约为 0.78%~0.87%。

1.2.2 野草香精油的气相色谱-质谱 (GC-MS) 测定

1.2.2.1 GC条件

色谱柱: Agilent 19091J-115, HP-5 (5% Phenyl Methyl Siloxan, 50 m×0.32 mm×0.52 μm); 进样口温度 250 °C, 检测器 (氢火焰离子检测器, FID) 温度 250 °C; 柱箱升温程序: 50 °C 升到 250 °C, 每分钟升 10 °C; 进样量 1 μL; 分流比 20:1; 进样口压力 102.28 kPa; 体积流量 2.1 mL/min; 载气: 高纯度 He^[7,19]。

1.2.2.2 MS条件

质谱条件: 电子轰击离子源 (EI), 电子能量 70 eV, 扫描范围 50~550 *m/z*; 离子源温度 230 °C; 四极杆温度 150 °C。溶剂延迟时间设置为 6 min^[7,19]。定量方法: 手动积分, 相对峰面积归一化法; 谱库: NIST14。

1.2.3 转录组测序及文库构建

将柠檬醛野草香新鲜嫩叶从 -80 °C 超低温冰箱取出, 委托武汉迈特维尔生物科技有限公司利用 Illumina NovaSeq 6000 高通量测序平台进行转录组测序, 并构建文库。设置三个生物学平行重复。

1.2.4 转录组数据组装及功能注释

为保证数据分析的质量及可靠性, 利用 fastp 软件对所得原始测序数据进行过滤, 去除带接头 (adapter)、含 N (N 表示无法确定碱基信息) 的以及低质量 (Qphred ≤ 20 的碱基数占整个 read 长度的 50% 以上的 reads) 的 reads, 得到 clean reads。由于野草香基因组为无参考序列, 所以采用 Trinity 对

clean reads 进行拼接组装, 最后得到 unigenes。将组装得到的 unigenes 序列比对到 NCBI 官方的蛋白序列数据库 (RefSeq Non-redundant Proteins, NR)、京都基因与基因组百科全书 (Kyoto Encyclopedia of Genes and Genomes, KEGG)、基因本体论数据库 (Gene Ontology, GO)、真核生物蛋白质同源簇数据库 (Clusters of Orthologous Groups for Eukaryotic Complete Genomes, KOG) 等数据库进行基因功能注释。

通过分析 KEGG 代谢通路, 从中筛选出野草香精油单萜合成相关候选基因, 获得基因序列和开放阅读框 (Open Reading Frame, ORF), 利用 MEGA 11.0 对编码关键酶的基因序列进行多重比对, 以此来对基因进行归类。

1.2.5 野草香叶醇脱氢酶 (GeDH) 基因克隆

筛选出 GeDH 基因, 获得基因序列和 ORF, 设计扩增引物。引物由生工生物工程 (上海) 股份有限公司合成 (表 1)。

利用 TaKaRa 公司的 MiniBEST Plant RNA Extraction Kit 试剂盒提取柠檬醛型野草香鲜叶的总 RNA, 经 1.5% 凝胶电泳检测以及浓度测定合格后, 取 4 μL 总 RNA, 按照 TaKaRa 公司的 PrimeScriptTMII 1st Strand cDNA Synthesis Kit 试剂盒说明书进行反转录反应, 合成 cDNA。以 cDNA 为模板, 反应体系 (50 μL): 模板 2.5 μL, 正反向引物各 2 μL, 2X FidCycle Fast 高保真 PCR Mix (含蓝染料) 25 μL, 无菌水 18.5 μL。PCR 反应程序为 98 °C、30 s; 98 °C、10 s, 55 °C、5 s, 72 °C、5 s, 35 个循环; 72 °C、2 min, 12 °C 保温。PCR 产物经电泳分析后, 送生工生物工程 (上海) 股份有限公司测序。

表 1 GeDH基因与引物序列

Table 1 GeDH gene and primer sequences

基因 ID	引物序列 (5'-3')
Cluster-22310.0	F: ATGGCTCAACTTTTCATTTC AAC R: TTACTCAGCTTTCAAAGATTCCCCAC

1.3 数据分析

将 GC-MS 检测得到的数据利用 NIST14 谱库检索、质谱分析等对成分进行检索定性, 并采用手动积分, 相对峰面积归一化法计算各组分的相对百分含量。利用 ORF Finder (<https://www.ncbi.nlm.nih.gov/orffinder/>) 获得基因的 ORF, 并对扩增序列进行翻译获得氨基酸序列。使用 DNAMAN 软件对氨基酸序列进行多重比对。使用 MEGA11.0 构建系统发育树。

2 结果与讨论

2.1 精油GC-MS成分测定

表 2 柠檬醛型野草香精油分析结果

Table 2 Analysis results of essential oil from citral chemotype of *E. cyprianii*

编号	保留时间/min	化合物	分子式	相对百分含量/%		
				EC-1	EC-2	EC-3
1	11.40	桉稀	C ₁₀ H ₁₆	0.27	0.30	0.31
2	11.63	甲基庚烯酮	C ₈ H ₁₄ O	0.47	0.51	0.52
3	11.78	月桂烯	C ₁₀ H ₁₆	0.16	0.17	0.18
4	14.28	苯己酮	C ₈ H ₈ O	0.19	0.19	0.19
5	15.14	芳樟醇	C ₁₀ H ₁₈ O	0.97	0.88	0.89
6	15.52	马鞭草烯醇	C ₁₀ H ₁₆ O	0.14	0.23	0.12
7	16.84	2,2-二甲基-3,4-辛二醛	C ₁₀ H ₁₆ O	0.83	1.26	0.72
8	17.11	异橙花醛	C ₁₀ H ₁₆ O	0.96	1.00	0.99
9	17.64	异香叶醛	C ₁₀ H ₁₆ O	1.36	1.44	1.39
10	18.11	α-松油醇	C ₁₀ H ₁₈ O	0.13	0.21	0.14
11	18.97	香茅醇	C ₁₀ H ₂₀ O	0.27	0.32	0.27
12	19.05	橙花醇	C ₁₀ H ₁₈ O	2.17	2.60	2.17
13	19.18	异香叶醇	C ₁₀ H ₁₈ O	0.17	0.21	0.19
14	19.52	顺式-柠檬醛	C ₁₀ H ₁₆ O	38.26	38.50	38.40
15	19.75	香叶醇	C ₁₀ H ₁₈ O	0.69	0.59	0.61
16	20.02	胡椒酮	C ₁₀ H ₁₆ O	0.11	0.12	0.13
17	20.35	反式-柠檬醛	C ₁₀ H ₁₆ O	44.81	43.89	45.15
18	22.20	未知		0.18	0.23	0.25
19	22.44	反式香叶酸	C ₁₀ H ₁₆ O ₂	0.33	0.40	0.38
20	22.81	丁香酚	C ₁₀ H ₁₂ O ₂	0.15	0.15	0.13
21	23.17	未知		0.20	0.21	0.23
22	24.77	石竹烯	C ₁₅ H ₂₄	3.56	3.02	3.07
23	25.66	葎草烯	C ₁₅ H ₂₄	1.22	1.06	1.08
24	26.34	大牛儿烯	C ₁₅ H ₂₄	0.43	0.29	0.36
25	26.69	2,4-二叔丁基	C ₁₄ H ₂₂ O	0.18	0.15	0.16
26	28.96	石竹素	C ₁₅ H ₂₄ O	0.74	0.96	0.80
27	29.58	环氧化蛇麻烯	C ₁₅ H ₂₄ O	0.19	0.24	0.20
		合计		99.12	99.12	99.03

注: EC表示野草香 *E. cyprianii*, 设置三个生物学平行重复, 表3同。

按设定的 GC-MS 条件分析柠檬醛型野草香精油的化学成分, 共检测出柠檬醛型野草香精油 35 个

成分, 剔除数据值小于 0.1% 且出现频率低于 50% 的变量, 余下 27 个成分 (表 2), 其中相对百分含量最高的成分为反式-柠檬醛 (44.81%、43.89%、45.15%), 其次是顺式-柠檬醛 (38.26%、38.50%、38.40%)。柠檬醛型野草香精油中的反式-柠檬醛和顺式-柠檬醛已用柠檬醛标准品进行比对 (图 1)。

朱金莲等^[7]对云南省文山州和宣威市的野草香 (*E. cyprianii*) 精油进行 GC-MS 测定分析, 二者的精油主要成分均为去氢香薷酮, 与周维书等^[3]的研究结果一致, 结合本研究推测云南地区的野草香精油可能主要分为去氢香薷酮型和柠檬醛型两种化学型。本研究中柠檬醛型野草香叶片精油中柠檬醛含量为 82.39%~83.55%, 较贵州省地区的木姜花^[8](*E. cyprianii*) 叶片中柠檬醛的含量 (68%) 高。

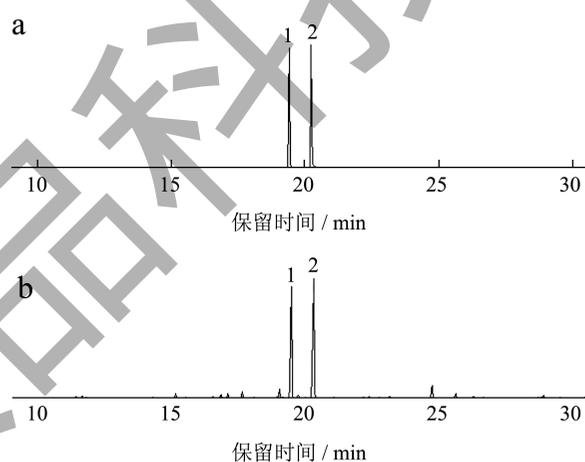


图 1 柠檬醛标准品 (a) 和柠檬醛型野草香精油 (b) GC-MS

Fig.1 GC-MS chart of citral standard (a) and essential oil from citral chemotype of *E. cyprianii* (b)

注: 1 为反式-柠檬醛, 2 为顺式-柠檬醛。

2.2 转录组数据组装与质量分析

利用 Illumina NovaSeq 6000 高通量测序平台对柠檬醛型野草香嫩叶进行转录组测序, 总过滤去除冗余片段后获得 147 186 998 条 clean reads, 共计 22.08 GB 的有效数据, Q30 (碱基量 ≥ 30%) 分别为 93.72%、94.28%、93.74%, GC 含量分别为 49.71%、49.56%、50.28%, 表明测序质控良好, clean reads 质量合格 (表 3)。运用 Trinity 软件将质控后得到的高质量序列进行组装, 得到 60 151 个 transcripts, 获得 33 720 个 unigenes, 其中 unigenes 的总长度为 60 678 337 bp, 平均长度为 1 799 bp, N50 为 2 306 bp。图 2 显示, 11 549 条 unigenes 长度超过 1 000 bp, 12 117 条 unigenes 长度超过 2 000 bp,

表明组装具有较高的完整性，能够进行后续的注释分析。转录组数据已上传至 NCBI 公共数据库，登录号为 SRR25346811~SRR25346813。

表 3 野草香转录组数据产出质量

Table 3 Quality of output of *E. cyprianii* transcriptome data

样本	Raw Reads/条	Clean Reads/条	Clean Base/GB	Q30/%	GC 含量/%
EC-1	47 333 170	45 005 970	6.75	93.72	49.71
EC-2	46 639 188	45 106 154	6.77	94.28	49.56
EC-3	58 827 156	57 074 874	8.56	93.74	50.28

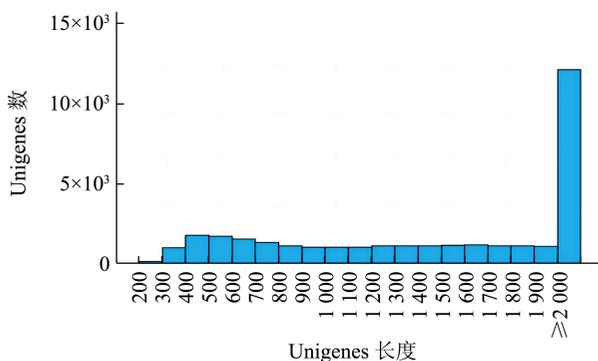


图 2 野草香转录组 unigenes 长度分布图

Fig.2 Length distribution of unigenes in transcriptome of *E. cyprianii*

2.3 转录组 unigenes 的功能注释

使用 Blast 将拼装所得的 33 720 条 unigenes 分别与 KEGG、NR、GO、KOG 等数据库进行比对分析，对各数据库注释的 unigenes 数目进行统计，进而获得野草香转录组 unigenes 的功能注释信息。结果表明，在 NR 数据库得到成功注释的 unigenes 最多，为 29 301 条 (86.9%)，在 KEGG、Swiss-port、TrEMBL、KOG、GO、PFAM 数据库获得注释的 unigenes 数目依次为 17 827 条 (52.87%)、22 667 条 (67.22%)、29 084 条 (86.25%)、17 233 条 (51.11%)、25 660 条 (76.1%)、23 740 条 (70.4%)。至少在 1 个数据库注释成功的 unigenes 有 29 406 条 (87.21%)。

将 unigenes 在 NR 数据库中进行比对注释，可以看出与其他物种序列的相似性，在相似序列匹配度较高的物种中，与粗毛紫苏 *Perilla frutescens* var. *hirtella* 的匹配度最高，为 40.83%，其次是紫苏 *Perilla frutescens* var. *frutescens* 为 37.56%，一串红 *Salvia splendens* 为 5.37%、白花泡桐 *Paulownia fortunei* 为 5.26%、西班牙鼠尾草 *Salvia hispanica* 为 3.05%、芝麻 *Sesamum indicum* 为 1.02%、紫花风铃木 *Handroanthus impetiginosus* 为 0.61%，其他物种为 6.29%。

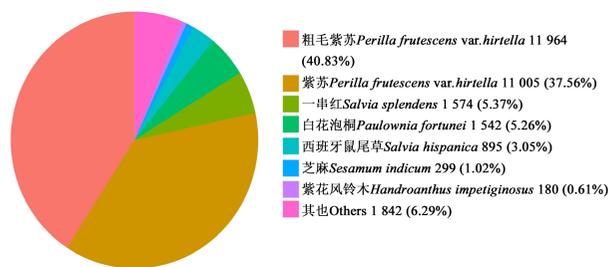


图 3 野草香转录组 unigenes 的物种分布图

Fig.3 Species classification of unigenes in transcriptome of *E. cyprianii*

在柠檬醛型野草香转录组中，共有 25 660 条 unigenes 根据 GO 功能分别被注释到生物过程 (Biological Process, BP)、细胞组分 (Cellular Component, CC) 和分子功能 (Molecular Function, MF) 3 大类中 (图 4)。在生物过程中基因主要聚集在细胞过程 (Cellular Process)、代谢过程 (Metabolic Process)，分别有 17 500、14 558 条。在细胞组分中细胞解剖实体 (Cellular Anatomical Entity) 相关基因数量最多，有 22 016 条，其次是蛋白复合物 (Protein-containing Complex)，有 3 629 条。分子功能中具有结合功能 (Binding) 和催化活性 (Catalytic Activity) 的基因数量较多，分别为 14 751 和 12 007 条。

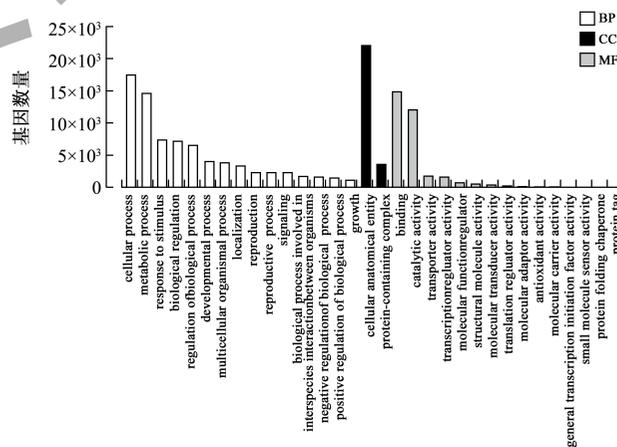


图 4 野草香转录组 unigenes 的 GO 功能分类注释

Fig.4 GO functional classification annotation of unigenes of *E. cyprianii* transcriptome

KOG 功能分类分析显示，得到 25 个不同的 KOG 功能类群 (图 5)。其中一般功能预测 (General Function Prediction Only) 有 3 489 条，数量最多；其次是信息转导机制 (Signal Transduction Mechanisms) 有 2 144 条；细胞运动性 (Cell Motility) 仅有 5 条，数量最少。

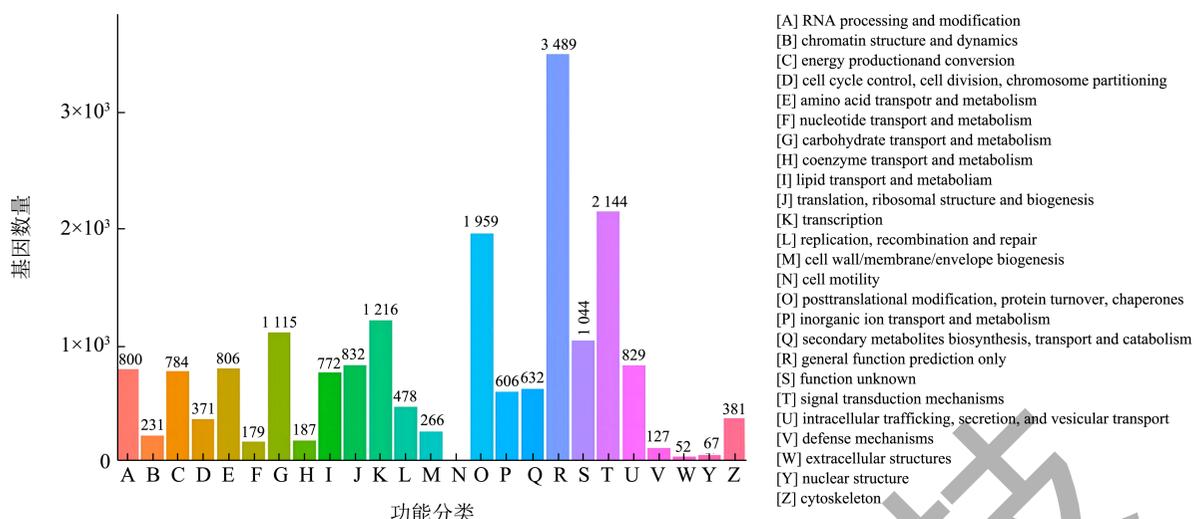


图 5 野草香转录组 unigenes 的 KOG 功能分类注释

Fig.5 KOG functional classification annotation of unigenes of *E. cyprianii* transcriptome

有 17 827 条 unigenes 参与 KEGG 代谢通路，分布于 129 条代谢通路中。KEGG 代谢通路分为 5 大分支：细胞过程 (Cellular Processes, A) 578 条、环境信息处理 (Environmental Information Processing, B) 1 610 条、遗传信息处理 (Genetic Information Processing, C) 3 072 条、代谢 (Metabolism, D) 15 012 条和有机系统 (Organismal Systems, E) 1 184 条 (图 6)。其中注释数量最多的是代谢途径 (Metabolic Pathways)，为 3 981 条，其次是次级代谢产物的生物合成 (Biosynthesis of Secondary Metabolites) 为 2 259 条，数量最少的是咖啡因代谢 (Caffeine Metabolism)，为 2 条。

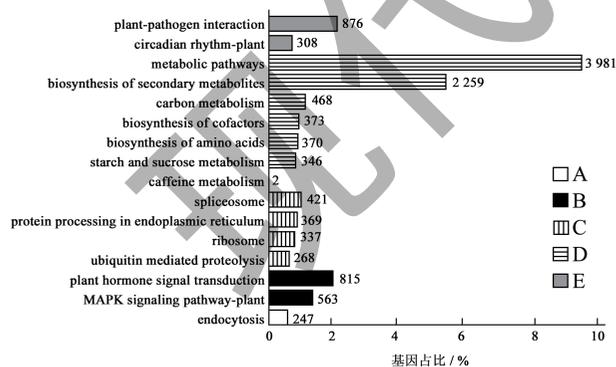


图 6 野草香转录组 unigenes 的 KEGG 分类注释

Fig.6 KEGG classification annotation of unigenes of *E. cyprianii* transcriptome

2.4 柠檬醛型野草香单萜代谢途径与相关基因

单萜类化合物是柠檬醛型野草香精油中的主要组成。在 KEGG 代谢通路分析中，与野草香精油萜类生物合成相关的代谢通路有 3 条，分别是：

萜类骨架生物合成 (编号为 ko00900)，相关基因有 109 条；单萜类生物合成 (编号为 ko00902)，相关基因有 32 条；倍半萜和三萜生物合成，(编号为 ko00909)，相关基因有 35 条。其中与柠檬醛型野草香单萜合成相关的代谢通路是萜类骨架生物合成和单萜类生物合成。

萜类化合物生物合成的前体来源于胞浆甲羟戊酸 (Mevalonic Acid, MVA) 和质体甲基赤藓糖磷酸 (Methylerythritol Phosphate, MEP) 途径，通过分析注释数据，获得 MVA 途径与萜类合成相关的关键酶 6 个，MEP 途径与萜类合成相关的关键酶 7 个，萜类骨架生物合成途径关键酶 4 个，与单萜合成相关的关键酶 4 个。野草香精油单萜代谢途径的酶与相关基因具体情况见表 4。

在进行目的基因筛选时不仅需要根据其代谢相关途径进行查找，同时也需要利用参照植物在不同的数据库中进行广泛查找，以此来提高目的基因的获取概率。在本研究中，从 KEGG 数据库中没有筛选到香叶醇合酶 (*GES*) 基因和香叶醇脱氢酶 (*GeDH*) 基因，但在 KEGG 数据库中有 1 条注释为 TPS-cin 的序列 (Cluster-7263.9) 在 NR 等数据库被注释为 *GES* 基因，其平均 FPKM 值为 98.62 (表 5)，利用 NCBI Blast 序列比对，发现该条序列与课题组前期克隆成功并公布的 3 条香蒿属 (*Elsholtzia* Willd.) *GES* 合酶基因序列^[17] 相似度达到 88% 以上，与 Ito 等^[20] 和 Masumoto 等^[21] 公布的紫苏属 (*Perilla*) *GES* 基因序列相似度达到 84% 以上 (图 7)。在 KEGG 数据库中有 1 条来自于苯丙素合成途径 (编号为 ko00940) 注释为肉桂醇脱氢

酶 (Cinnamyl-alcohol dehydrogenase, CAD) 的序列 (Cluster-22310.0) 在 NR 等数据库被注释为 *GeDH* 基因, 其平均 FPKM 值为 117.19 (表 5), 利用 NCBI Blast 序列比对, 发现该条序列与 Sato-Masumoto 等^[22]

公布的紫苏属 (*Perilla*) *GeDH* 基因序列相似度达到 92% 以上 (图 8), 具有较高同源性。该 *GeDH* 基因可能与之前报道的罗勒^[18]、紫苏^[22]、生姜^[23]的 *GeDH*s 一样, 属于肉桂醇脱氢酶类群。

表 4 野草香精油单萜代谢途径的酶与相关基因

Table 4 Enzymes and related genes of terpenoid metabolic pathways in essential oil from *E. cyprianii*

代谢通路	酶名称	KO 编号	EC 编号	基因数量/条
胞浆甲羟戊酸 (Mevalonic acid, MVA)	乙酰乙酰-CoA 硫解酶 (Acetyl-CoA C-acetyltransferase, AACT)	K00626	2.3.1.9	9
	羟甲基戊二酰-CoA 合成酶 (Hydroxymethylglutaryl-CoA synthase, HMGS)	K01641	2.3.3.10	1
	3-羟基-3-甲基戊二酰-CoA 还原酶 (Hydroxymethylglutaryl-CoA reductase, HMGR)	K00021	1.1.1.34	5
	甲羟戊酸激酶 (Mevalonate kinase, MVK)	K00869	2.7.1.36	1
	磷酸甲羟戊酸激酶 (Phosphomevalonate kinase, PMK)	K00938	2.7.4.2	6
	二磷酸甲羟戊酸脱羧酶 (Diphosphomevalonate decarboxylase, MVD)	K01597	4.1.1.33	1
	质体甲基赤藓糖磷酸 (Methylerythritol Phosphate, MEP)	1-脱氧-D-木酮糖合成酶 (1-Deoxy-D-xylulose-5-phosphate synthase, DXS)	K01662	2.2.1.7
1-脱氧-D-木酮糖-5-磷酸还原酶 (1-Deoxy-D-xylulose-5-phosphate reductoisomerase, DXR)		K00099	1.1.1.267	1
2-甲基-D-赤藓糖醇-4-磷酸胞苷基转移酶 (2-C-methyl-D-erythritol 4-phosphate cytidyltransferase, MCT)		K00991	2.7.7.60	1
4-二磷酸胞苷-2-甲基-D-赤藓糖醇激酶 (4-Diphosphocytidyl-2-C-methyl-D-erythritol kinase, CMK)		K00919	2.7.1.148	1
2-C-甲基-D-赤藓糖醇 2,4-环二磷酸合酶 (2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase, MDS)		K01770	4.6.1.12	1
(E)-4-羟基-3-甲基丁-2-烯基二磷酸合成酶 (E)-4-Hydroxy-3-methylbut-2-enyl-diphosphate synthase, HDS)		K03526	1.17.7.1/3	2
4-羟基-3-甲基丁-2-烯基二磷酸还原酶 (4-Hydroxy-3-methylbut-2-enyl diphosphate reductase, HDR)		K03527	1.17.7.4	10
其他萜类骨 架生物合成 Biosynthesis of other terpenoid backbone	异戊烯基焦磷酸异构酶 (Isopentenyl-diphosphate delta-isomerase, IDI)	K01823	5.3.3.2	2
	法呢基焦磷酸合酶 (Farnesyl diphosphate synthase, FPPS)	K00787	2.5.1.1/10	3
	香叶基二磷酸合成酶 (Geranyl diphosphate synthase, GPS)	K14066	2.5.1.1	9
单萜生物合成 Monoterpenoid biosynthesis	香叶基香叶基二磷酸合酶 (Geranylgeranyl diphosphate synthase, type II, GGPS)	K13789	2.5.1.1/10/29	7
	1,8-桉叶素合酶 (1,8-Cineole synthase, TPS-cin)	K07385	4.2.3.108	12
	月桂烯/奥辛烯合酶 (Myrcene/ocimene synthase, MYS)	K12467	4.2.3.15	6
	(3S)-芳樟醇合成酶 (3S)-Linalool synthase, LIS)	K15086	4.2.3.25	8
	(+)-新薄荷醇脱氢酶 (+)-Neomenthol dehydrogenase, NDS)	K15095	1.1.1.208	6

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> Eisholtzia winitiana voucher BXR0001 geraniol synthase (GES) mRNA, complete cds	Eisholtzia winit...	2 787	2 787	98%	0.0	94.86%	1 782	MZ221236.1
<input checked="" type="checkbox"/> Eisholtzia heterophylla voucher YYXR0001 geraniol synthase (GES) mRNA, complete cds	Eisholtzia heter...	2 154	2 154	98%	0.0	88.45%	1 791	MZ221235.1
<input checked="" type="checkbox"/> Eisholtzia ciliata voucher XR0001 geraniol synthase (GES) mRNA, complete cds	Eisholtzia ciliata	2 121	2 121	98%	0.0	88.15%	1 785	MZ221234.1
<input checked="" type="checkbox"/> Perilla frutescens var. crispa strain 79 geraniol synthase mRNA, complete cds	Perilla frutesce...	1 790	1 790	100%	0.0	84.64%	1 812	DQ897973.1
<input checked="" type="checkbox"/> Perilla frutescens strain 1864 geraniol synthase mRNA, complete cds	Perilla frutescens	1 790	1 790	100%	0.0	84.64%	2 119	DQ234300.1
<input checked="" type="checkbox"/> Perilla setoyensis geraniol synthase (Tps-5031G8) mRNA, complete cds	Perilla frutesce...	1 784	1 784	100%	0.0	84.57%	2 071	FJ644545.1
<input checked="" type="checkbox"/> Perilla citriodora strain 1861 geraniol synthase mRNA, complete cds	Perilla frutesce...	1 784	1 784	100%	0.0	84.60%	2 019	DQ234299.1
<input checked="" type="checkbox"/> Perilla citriodora strain 4935 geraniol synthase mRNA, complete cds	Perilla frutesce...	1 784	1 784	100%	0.0	84.60%	2 109	DQ234298.1
<input checked="" type="checkbox"/> Perilla citriodora geraniol synthase mRNA, complete cds	Perilla frutesce...	1 784	1 784	100%	0.0	84.60%	2 057	DQ088667.1
<input checked="" type="checkbox"/> Perilla frutescens var. hirtella geraniol synthase (Tps-5073G30) mRNA, complete cds	Perilla frutesce...	1 760	1 760	99%	0.0	84.39%	1 998	FJ644547.1

图 7 GES 基因序列比对

Fig.7 Alignment of the predicted sequences of GES gene

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> Perilla setoyensis strain 5031 geraniol dehydrogenase mRNA, complete cds	Perilla frutesce...	926	926	99%	0.0	92.20%	1 083	JX855838.1
<input checked="" type="checkbox"/> Perilla citriodora strain 87 geraniol dehydrogenase mRNA, complete cds	Perilla frutesce...	920	920	99%	0.0	92.05%	1 083	JX855837.1
<input checked="" type="checkbox"/> Perilla frutescens strain 32 geraniol dehydrogenase mRNA, complete cds	Perilla frutescens	920	920	99%	0.0	92.05%	1 083	JX855836.1

图 8 GeDH 基因序列比对

Fig.8 Alignment of the predicted sequences of GeDH gene

表 5 野草香精油合成关键基因表达量

Table 5 Expression of key genes involved in the synthesis of characteristic essential oil from *E. cyprianii*

基因 ID	KO 注释	NR 注释	平均 FPKM 值
Cluster-7263.9	TPS-cin	GES	98.62
Cluster-22310.0	CAD	GeDH	117.19

2.5 香叶醇脱氢酶 (GeDH) 基因克隆

本研究利用柠檬醛型野草香 cDNA, 结合表 1 中的 *GeDH* 基因的扩增引物进行 PCR 扩增, 经电泳分析, 产生明亮的目标条带 (图 9), 经测序分析, 拼接后的核苷酸序列长度为 1 143 bp, 编码 380 个氨基酸。利用 DNAMAN 软件对柠檬醛型野草香 *GeDH* 基因和 3 条紫苏属 *GeDH* 基因^[21] 的氨基酸序列进行比对 (图 10), 柠檬醛型野草香 *GeDH* 显示出保守的 Zn2 (GD(X)_{9,10}C(X)₂C(X)₂C(X)₇C)、NADP (GXG(X)₂G) 结合位点和类似于紫苏属植物的 Zn1 (GHE(X)₂G(X)₅G(X)₂V) 结合位点。该 *GeDH* 基因的核苷酸序列及推测的氨基酸序列已提交给 GenBank, 登录号为: OR529418。

香薷属植物单萜生物合成途径中 *GES* 基因被筛选出并已克隆成功^[17], 在之前香薷属精油研究中^[17,24], 柠檬醛的含量非常低, 仅在个别样品中发现, 因此并未发现 *GeDH* 基因, 而在唇形科植物罗

勒^[18]、紫苏^[22]等植物中, *GeDH* 基因被成功克隆, 作为香薷属特色精油合成通路上促进中间产物柠檬醛产生的关键作用酶, *GeDH* 基因对香薷属单萜生物合成途径的解析有着重要意义。已有研究表明与植物精油中高含量成分对应的单萜合酶基因其克隆成功率更高, 主要原因可能在于基因表达量与其产物的积累量是一致的^[25]。柠檬醛型野草香中 *GeDH* 基因的成功克隆, 意味着研究香薷属特色精油合成通路上的后续未知基因时选择合适的精油化学型材料, 对香薷属单萜生物合成途径的解析可能是更高效的。

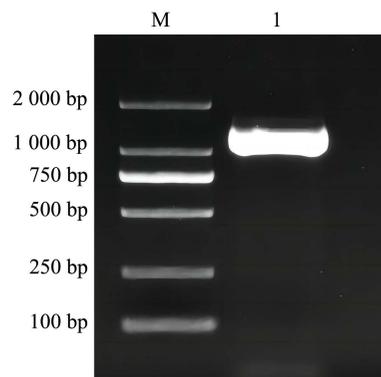


图 9 野草香 GeDH 基因克隆凝胶电泳图

Fig.9 Gel electrophoresis of GeDH gene from *E. cyprianii* cloning

注: M 为 DL2000 DNA Maker, 1 为 *GeDH* 基因。

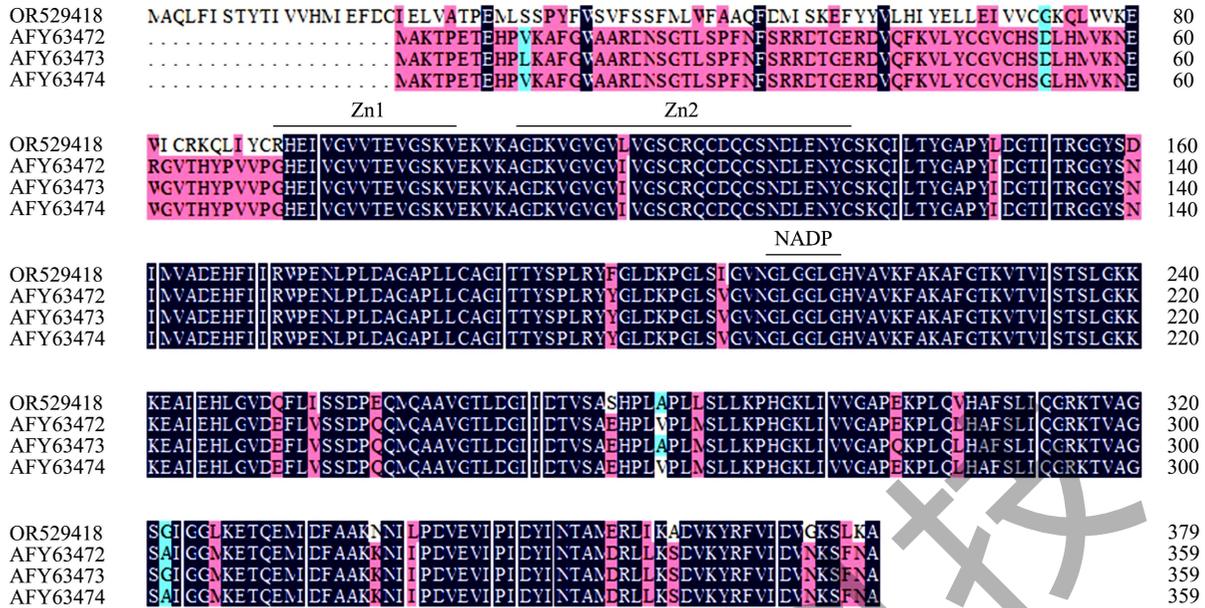


图 10 *GeDH* 基因氨基酸多序列比对

Fig.10 Amino acid sequence alignment of *GeDH*

在 Iijima 等^[22]的研究中, 生姜的 *GeDH* 基因与多种植物的 CADs 具有高度的同源性, 本研究从 NCBI 下载其他植物 *GeDH* 基因与 *CAD* 基因的氨基酸序列进行聚类分析 (图 11), 可以看到野草香与紫苏属植物亲缘关系最近, 表明单萜类化合物生物合成途径中的相关基因片段对揭示唇形科属间系统演化关系是有利的。

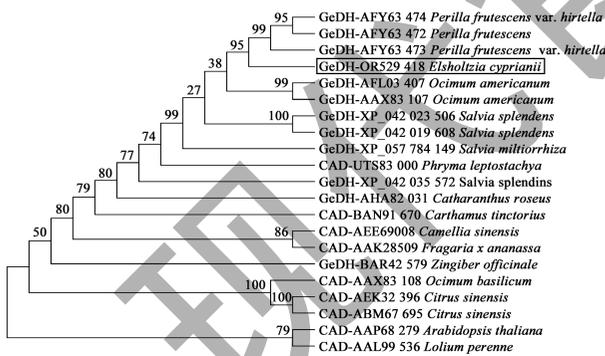


图 11 *GeDH* 基因和 *CAD* 基因的氨基酸序列进化树分析

Fig.11 Nucleotide sequence phylogenetic tree of *GeDH* and *CAD*

CAD 通过加氢反应将肉桂醛变成肉桂醇^[26], 而 *GeDH* 通过脱氢反应将香叶醇变成香叶醛 (即反式-柠檬醛), 两种酶的作用相反。在罗勒^[27]香叶醇的氧化研究中, *CAD1* 编码的酶可逆地氧化香叶醇以产生香叶醛, *CAD1* 在 $NADP^+$ 用作氧化方向的辅因子时, 氧化肉桂醛和香叶醇, 但在 NAD^+ 用作辅因子时仅氧化肉桂醛而不氧化香叶醇。所有真正的

CAD 蛋白都含有一些共同的保守结构域, 即 Zn1 结合标签、Zn2 结合位点和富含甘氨酸的 NADP 结合位点^[26], *GeDH* 基因也同样含有这些保守结构域, 由此可推测 *GeDH* 基因与 *CAD* 基因可能存在着趋同进化的关系^[28]。

3 结论

本研究以柠檬醛型野草香新鲜叶片为原料, 对其精油进行提取及成分测定分析, 并通过 Illumina NovaSeq 6000 高通量测序获得柠檬醛型野草香转录组数据, 挖掘 *GeDH* 基因。柠檬醛型野草香精油中含柠檬醛 82.39%~83.55%, 是目前野草香精油成分研究中柠檬醛含量最高的。转录组测序质控良好, 拼接质量高, 将拼装所得的 unigenes 进行功能注释及功能分类, 与野草香精油单萜类生物合成相关的代谢通路有 2 条, 分别是: 萜类骨架生物合成 (编号为 ko00900) 和单萜类生物合成 (编号为 ko00902)。利用粗毛紫苏在 NR 等其他数据库中查找, 分别找到 *GES* 基因和 *GeDH* 基因各 1 条, 与 NCBI 公布的基因序列进行比对具有较高的同源性。从柠檬醛型野草香转录组中扩增获得 1 143 bp 的 *GeDH* 基因序列, 编码 380 个氨基酸, 显示出保守的 Zn2、NADP 结合位点和类似于紫苏属植物的 Zn1 结构基团。同源比对和聚类分析发现, 野草香与紫苏属植物亲缘关系最近, 该 *GeDH* 与 *CAD* 基因在保守结构域上具有相似性, 推测两者可能存在

着趋同进化的关系。本研究将有助于通过分子手段调控精油中柠檬醛的产量,为天然柠檬醛精油的开发利用提供一定的帮助,为拓宽食用调味油、天然食品抗菌剂和防腐剂的种类奠定基础。

参考文献

- [1] 国家中医药管理局,《中华本草》编委会.中华本草[M].上海:上海科学技术出版社,1999.
- [2] 《贵州植物志》编委会.贵州植物志(第八卷)[M].成都:四川民族出版社,1988.
- [3] 周维书,朱奕丽,周震.野草香精油成分的研究[J].中国药理学杂志,1992,2:79-80.
- [4] 许本汉,德宏傣蔬——野草香[J].植物杂志,1995,2:23.
- [5] 王艺舟,谷荣辉,刘博.滇西北藏区野生香料植物资源调查[J].北方园艺,2017(16):112-120.
- [6] 普春霞,徐玉琪,朱金莲,等.云南7种药食香薷属植物基于精油化学型的抗菌活性研究[J].食品科技,2022,47(9):183-189.
- [7] 朱金莲,杨礼攀,唐仁华,等.6种香薷属植物精油化学型分析及其系统学意义探究[J].中药材,2022,45(1):108-114.
- [8] 侯颖辉,李德文,于二汝,等.木姜花和木姜子精油成分比较[J].中国调味品,2017,42(7):139-142.
- [9] 张晶晶,彭锐,杜春贵,等.柠檬醛抑制真菌作用的研究进展及其在食品防霉中的应用前景[J].中国调味品,2020,45(7):186-190,200.
- [10] CAMPOS C A, LIMA B S, TRINDADE G G G, et al. Anti-hyperalgesic and anti-inflammatory effects of citral with β -cyclodextrin and hydroxypropyl- β -cyclodextrin inclusion complexes in animal models [J]. Life Sciences, 2019, 229: 139-148.
- [11] LING Q, ZHANG B, WANG Y, et al. Chemical composition and antioxidant activity of the essential oils of citral-rich chemotype *Cinnamomum camphora* and *Cinnamomum bodinieri* [J]. Molecules, 2022, 27(21): 7356.
- [12] CHAOUKI W, LEGER D Y, LIAGRE B, et al. Citral inhibits cell proliferation and induces apoptosis and cell cycle arrest in MCF-7 cells [J]. Fundamental & Clinical Pharmacology, 2010, 23(5): 549-556.
- [13] 殷志勇,王秋娟,贾莹.山苍子水提物柠檬醛抗哮喘作用的实验研究[J].中国临床药理学与治疗学,2006,2:197-201.
- [14] 罗晶,田鑫,刘波,等.芳香中药精油成分在心血管疾病中的应用[J].中草药,2020,51(1):245-255.
- [15] THIELMANN J, MURANYI P. Review on the chemical composition of *Litsea cubeba* essential oils and the bioactivity of its major constituents citral and limonene [J]. Journal of Essential Oil Research, 2019, 31(5): 361-378.
- [16] VALKOVÁ V, ĎURANOVÁ H, GALOVIČOVÁ L, et al. *Cymbopogon citratus* essential oil: Its application as an antimicrobial agent in food preservation [J]. Agronomy, 2022, 12(1): 155.
- [17] GENG X W, TANG R H, ZHANG A L, et al. Mining, expression, and phylogenetic analysis of volatile terpenoid biosynthesis-related genes in different tissues of ten *Elsholtzia* species based on transcriptomic analysis [J]. Phytochemistry, 2022, 203: 113419.
- [18] IJIM Y, GANG D R, FRIDMAN E, et al. Characterization of geraniol synthase from the peltate glands of sweet basil [J]. Plant Physiology, 2004, 134(1): 370-379.
- [19] 耿秀文,张爱丽,唐仁华,等.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,2021,52(11):3373-3382.
- [20] ITO M, HONDA G. Geraniol synthases from *Perilla* and their taxonomical significance [J]. Phytochemistry, 2007, 68(4): 446-453.
- [21] MASUMOTO N, KORIN M, ITO M. Geraniol and linalool synthases from wild species of *Perilla* [J]. Phytochemistry, 2010, 71(10): 1068-1075.
- [22] SATO-MASUMOTO N, ITO M. Isolation and characterization of isopiperitenol dehydrogenase from piperitenone-type *Perilla* [J]. Biological and Pharmaceutical Bulletin, 2014, 37(5): 847-852.
- [23] IJIMA Y, KOEDUKA T, SUZUKI H, et al. Biosynthesis of geraniol, a potent aroma compound in ginger rhizome (*Zingiber officinale*): Molecular cloning and characterization of geraniol dehydrogenase [J]. Plant Biotechnology, 2014, 31(5): 525-534.
- [24] CHEN S Q, CHEN J X, XU Y F, et al. *Elsholtzia*: A genus with antibacterial, antiviral, and anti-inflammatory advantages [J]. Journal of Ethnopharmacology, 2022, 297(28): 115549.
- [25] BENABDELKADER T, GUITTON Y, PASQUIER B, et al. Functional characterization of terpene synthases and chemotypic variation in three lavender species of section *Stoechas* [J]. Physiologia Plantarum, 2015, 153(1): 43-57.
- [26] 张鲁斌,谷会,弓德强,等.植物肉桂醇脱氢酶及其基因研究进展[J].西北植物学报,2011,31(1):8.
- [27] IJIMA Y, WANG G, FRIDMAN E, et al. Analysis of the enzymatic formation of citral in the glands of sweet basil [J]. Archives of Biochemistry and Biophysics, 2006, 448(1-2): 141-149.
- [28] LIU C C, SMIT S J, DANG J J, et al. A chromosome-level genome assembly reveals that a bipartite gene cluster formed via an inverted duplication controls monoterpenoid biosynthesis in *Schizonepeta tenuifolia* [J]. Molecular Plant, 2023, 16(3): 533-548.