

• 药材与资源 •

掌叶大黄转录组测序及蒽醌类合成关键酶基因差异表达分析

邹建珍¹, 胡晓晨², 李依民^{2*}, 郝紫彤¹, 张 岗^{1,2}, 刘蒙蒙^{1*}

1. 河北大学中医学院, 河北 保定 071002

2. 陕西中医药大学药学院/陕西省中医药管理局“秦药”研发重点实验室, 陕西 西安 712046

摘要: 目的 研究掌叶大黄 *Rheum palmatum* 蒽醌合成关键基因。方法 利用 RNA-seq 技术对掌叶大黄根、根茎、叶进行转录组测序和生物信息学分析。结果 经 Trinity 软件组装后获得 140 224 条 Unigenes, 全部 Unigenes 能被非冗余蛋白库、核酸序列数据库、京都基因组百科全书数据库、蛋白质序列数据库、蛋白家族数据库、基因本体联合数据库、同源蛋白簇数据库等公共数据库注释。差异基因分析结果显示, 掌叶大黄的根与叶相比, 差异表达基因总数为 4175 个, 根与根茎相比, 差异表达基因总数为 992 个, 叶与根茎相比, 差异表达基因总数为 4469 个。差异表达基因代谢途径分析发现, 苯丙烷生物合成通路在叶比根中被显著富集。蒽醌类化合物合成相关关键差异表达基因分析发现, 关键差异表达基因主要涉及甲羟戊酸 (mevalonic acid, MVA) 途径、甲基赤藓糖醇 (methylerythritol 4-phosphate, MEP)、莽草酸及聚酮途径的 13 个基因。结论 为进一步挖掘大黄有效成分生物合成过程中的关键基因, 为解析调控其有效成分生物合成途径提供研究基础。

关键词: 掌叶大黄; 蒽醌; 转录组; 基因表达; 甲羟戊酸途径; 甲基赤藓糖醇途径

中图分类号: R286.12 文献标志码: A 文章编号: 0253-2670(2023)07-2217-09

DOI: 10.7501/j.issn.0253-2670.2023.07.022

Transcriptome sequencing and expression analysis of anthraquinone synthesis key enzyme genes on *Rheum palmatum*ZOU Jian-zhen¹, HU Xiao-chen², LI Yi-min², HAO Zi-tong¹, ZHANG Gang^{1,2}, LIU Meng-meng¹

1. College of Traditional Chinese Medicine, Hebei University, Baoding 071002, China

2. Key Laboratory for Research and Development of “Qin Medicine” of Shaanxi Administration of Traditional Chinese Medicine, College of Pharmacy, Shaanxi University of Chinese Medicine, Xi'an 712046, China

Abstract: Objective To identify genes involved in anthraquinones biosynthesis in *Rheum palmatum*. **Methods** RNA-seq was used to investigate the gene expression profiles of leaves, rhizome, and root from *R. palmatum*. **Results** A total of 140 224 unigenes were obtained through the Trinity assembling. Functional annotation revealed that all unigenes were successfully annotated in the NR, NT, Swiss-port, PFAM, and KOG databases. Differentially expressed genes (DEGs) analysis suggested that 4 175 992 and 4469 genes were differentially expressed in root/leaf, root/rhizome and leaf/ rhizome, respectively. KEGG analysis showed that the Phen propane biosynthesis pathway was significantly enriched in leaf/root. The analysis of key differentially expressed genes related to anthraquinone synthesis revealed 13 genes mainly involved in the mevalonic acid (MVA) pathway, methylerythritol 4-phosphate (MEP), mangiferous acid and polyketide pathways. **Conclusions** The key genes are further explored in the biosynthesis of the active ingredients of rhubarb in order to provide a research basis for the analysis of the pathways regulating the biosynthesis of its active ingredients.

Key words: *Rheum palmatum* L.; anthraquinones; transcriptome; gene expression; mevalonic acid; methylerythritol 4-phosphate

大黄为蓼科植物掌叶大黄 *Rheum palmatum* L., 黄 *R. officinale* Baill. 的干燥根及根茎, 具有利湿退唐古特大黄 *R. tanguticum* Maxim. ex Balf. 或药用大

黄、凉血解毒、逐瘀痛经、清热泻火等功效。现代

收稿日期: 2022-09-09

基金项目: 国家自然科学基金资助项目 (82104334); 国家自然科学基金资助项目 (82204567); 国家自然科学基金资助项目 (8220142901); 陕西中医药大学“秦药”品质评价与资源开发创新团队项目 (2019-QN01); 咸阳市中青年科技创新领军人才 (L2022CXNLR009)

作者简介: 邹建珍, 女, 硕士研究生, 研究方向为药用植物生物技术与分子生物学。E-mail: 2648110750@qq.com

*通信作者: 刘蒙蒙, 讲师, 研究方向为中药资源学。E-mail: mmliu@hbu.edu.cn

李依民, 副教授, 研究方向为分子生药学。E-mail: 2051058@sntcm.edu.cn

研究表明, 大黄酸、大黄素、大黄酚、大黄素甲醚和芦荟大黄素 5 种蒽醌类物质是大黄主要的药效成分, 具有泻下、抗菌、抗炎、抗肿瘤、利尿、抗凝血、抗衰老、免疫调节等多种药理作用^[1-2]。由于蒽醌类物质在医药和生物领域的重要作用, 如何高效、快速地获得蒽醌类物质, 提高植物蒽醌类物质的合成效率, 已经成为中药资源学研究的热点方向。所以研究植物体内的蒽醌代谢路径及机制, 对蒽醌类物质的高效生产具有十分重要的意义。

植物体内蒽醌母核的生物合成主要涉及莽草酸(shikimic acid)和聚酮(polyketide)途径, 而 α -酮戊二酸(α -ketoglutarate, AKG)是莽草酸合成蒽醌的重要物质, 因此也涉及三羧酸循环(TCA cycle)^[3]。此外, 由于中间产物共用, 甲基赤藓糖醇(methylerythritol 4-phosphate, MEP)、甲羟戊酸(mevalonic acid, MVA)途径也参与其中^[4]。莽草酸途径中分支酸盐和 α -酮戊二酸经多步反应合成的邻琥珀酰苯甲酸酯(O-succinylbenzoate, OSB)构成蒽醌母核的 A、B 环, 而来自 MEP 或 MVA 的异戊烯焦磷酸(isopentenyl diphosphate, IPP)经异构反应构成 C 环^[5]。

在植物体的聚酮途径中, 目前已知乙酰辅酶 A 或丙二酰辅酶 A 经 III 型聚酮合成酶(polyketide synthase, PKS III)催化形成八肽聚酮长链, 再由聚酮环化酶(polyketide cyclase, PKC)催化得到蒽醌类, 经氧化最终得到蒽醌母核。大黄素型蒽醌母核主要来自聚酮途径, 而茜草型蒽醌母核主要来自莽草酸途径^[6], 这些途径之间又有中间产物的交叉。

药用植物蒽醌类成分生物合成途径及其调控研究刚起步, 在虎杖、何首乌、荞麦等中有关键酶基因相关报道。莽草酸途径中异分支酸合成酶(ICS)催化蒽醌合成关键分叉点上的代谢物异分支酸的合成, 聚酮途径 PKSIII 属于查耳酮合成酶家族(CHSs), 已发现有 CHS、二苯乙烯合酶、吡啶酮合酶、2-吡喃酮合酶和聚八酮合酶等^[7]。目前对掌叶大黄的转录组报道较少, 本研究利用 Illumina Hiseq 2000 高通量测序技术对掌叶大黄的根、根茎、叶分别进行转录组测序, 以期筛选出蒽醌类化合物生物合成途径相关的基因, 为初步阐明蒽醌类化合物生物合成途径奠定基础。

1 仪器与材料

1.1 材料

供试材料一年生掌叶大黄来源于甘肃省陇南市宕昌县, 经陕西中医药大学王继涛高级实验师鉴定为

掌叶大黄 *R. palmatum* L.。植物总 RNA 提取试剂盒(RN53-EASYspin Plus 多糖多酚/复杂植物 RNA 快速提取试剂盒)、反转录试剂盒(PC58-TRUEscript RT MasterMix)、荧光定量试剂盒(PC33-SYBR Green qPCR Mix)均购自北京艾德莱生物科技有限公司。

1.2 仪器与试剂

Waters 2695 型高效液相色谱仪, Empower 3 色谱工作站(Waters 公司, 美国), KQ-200KED 型超声波清洗机(江苏昆山公司), 实时荧光定量 PCR 仪(ABI StepOne7500, 美国)。

对照品没食子酸(批号 A0110)、儿茶素(批号 A0158)、番泻苷 B(批号 A0127)、大黄酚-8-O-葡萄糖苷(批号 A0762)、大黄素-8-O-葡萄糖苷(批号 A1502)、芦荟大黄素(批号 A0047)、大黄酸(批号 A0043)、大黄素(批号 A0044)、大黄酚(批号 A0046)、大黄素甲醚(批号 A0045)均购自成都曼斯特生物技术有限公司, 质量分数均大于 98%。

2 方法

2.1 掌叶大黄根、根茎、叶 HPLC 含量测定

取掌叶大黄植株若干, 随机选择 3 株混合, 3 次平行取样。分别取掌叶大黄根、根茎、叶, 用常规方法烘干。参考文献报道大黄 HPLC 含量测定方法^[8]进行实验。色精密称取没食子酸、儿茶素、番泻苷 B、大黄酚-8-O-葡萄糖苷、大黄素-8-O-葡萄糖苷、芦荟大黄素、大黄酸、大黄素、大黄酚、大黄素甲醚对照品适量, 分别置于 10 mL 量瓶, 用甲醇溶解并稀释至刻度, 摇匀, 得到质量浓度分别为 224.0、710.0、656.0、172.0、234.0、79.0、77.0、28.0、48.0、27.0 $\mu\text{g/mL}$ 的对照品储备液。分别精密量取各对照品储备液 1 mL, 甲醇稀释 10 倍, 得到相应质量浓度的混合对照品储备液。4 $^{\circ}\text{C}$ 保存备用。各成分的线性回归方程参考文献报道^[8]。

色谱柱为安捷伦 C₁₈ (250 mm $\times$ 4.6 mm, 5 μm) 色谱柱; 流动相由甲醇(A)和 0.2% 磷酸水溶液(B)组成, 梯度洗脱(0~5 min, 5%~15% A; 5~15 min, 15%~30% A; 15~25 min, 30%~35% A; 25~31 min, 35%~42% A; 31~46 min, 42%~53% A; 46~66 min, 53%~68% A; 66~75 min, 68%~100% A; 75~85 min, 100% A), 检测波长 260 nm, 柱温 30 $^{\circ}\text{C}$, 体积流量 1.0 mL/min。进样量为 10 μL 。在上述色谱条件下分析, 理论板数按各个成分计算均不低于 5000, 与相邻组分峰的分度均大于 1.5, 色谱峰对称因子均在 0.95~1.05。

2.2 总 RNA 提取

随机挑选 3 株掌叶大黄植株, 分别称 100 mg 左右新鲜的掌叶大黄的根、根茎、叶, 加液氮用研钵充分研磨, 采用植物总 RNA 提取试剂盒从新鲜的植物组织中提取总 RNA, 用 1.2% 琼脂糖凝胶电泳检测总 RNA 完整性, 用 ND2000 测定总 RNA 的 A_{260}/A_{280} 的值, 选择 A_{260}/A_{280} 为 1.8~2.0 的总 RNA 进行反转录。

2.3 cDNA 文库构建与序列组装

将掌叶大黄的根、根茎、叶的 RNA 由上海美吉生物医药科技有限公司构建测序文库, 使用 Illumina HiSeq 2000 测序平台进行转录组双末端测序。

2.4 Unigenes 的功能注释

去除原始测序序列中含有接头及低质量的序列后, 得到 clean reads。使用 Trinity^[9]对 clean reads 进行拼接, 然后使用序列聚类软件 Corset^[10]做进一步序列拼接和去冗余处理, 得到 Unigenes。将 Unigenes 序列与蛋白数据库 Nr、Nt、Swiss-Prot、Pfam、KEGG、COG 和 GO 做 BLASTX 比对 (e 值 $<0.000\ 01$), 得到 Unigenes 的蛋白功能注释信息。

2.5 基因表达差异分析

用 RSEM 软件^[11]对各样品进行基因表达水平分析。统计每个样品比对到参考序列对应基因的

reads 数量 (read count), 并将其转换成 FPKM (fragments per kilobase of exon model per million mapped reads)^[12]。随后用 TMM 对基因表达水平分析中得到的 read count 数据进行标准化处理, 再用 DEGseq^[13]筛选差异表达基因 (differentially expressed gene, DEG), 筛选标准为 $|\log_2(\text{Fold Change})| > 1$ 且 $q\text{-value} < 0.005$ 。

2.6 差异表达基因

通过基因本体联合数据库 (gene Ontology, GO) 数据库 (<http://geneontology.org>) 和京都基因和基因组百科全书 (Kyoto encyclopedia of genes and genomes, KEGG) 数据库 (<https://www.genome.jp/kegg/pathway.html>) 进行 GO 和 KEGG 分析, 利用 GOseq 和 KOBAS (2.0) 软件^[14-15]分别对 DEG 进行 GO 和 KEGG 富集分析。

2.7 qRT-PCR 分析

选取转录组差异基因进行 qRT-PCR 分析, 以验证转录组数据可靠性。qRT-PCR 引物利用 Primer Premier 6 软件设计, 引物列表详见表 1。将样品 RNA 用艾德莱反转录试剂盒反转录为 cDNA, 以掌叶大黄的 actin 基因为内参, 随后采用荧光定量 PCR 仪进行 qPCR 扩增。每个样品设置 3 个生物学重复, 采用 $2^{-\Delta\Delta C_t}$ 法计算基因相对表达量^[16]。

表 1 实时荧光定量 PCR 所用的引物

Table 1 Primers for qRT-PCR analysis

编号	基因	登录号	引物
1	<i>MenB2</i>	Cluster-7329.77811	5'-CGGTATTATCCCGCATCG-3' 5'-TTCACCCACAGCCTTCTCG-3'
2	<i>DXPS6</i>	Cluster-7329.42310	5'-GCCATAGTTCTAATAGCATCTCGG-3' 5'-GCATAGGTCTCCAATCGTCTT-3'
3	<i>MCT1</i>	Cluster-7329.52884	5'-CACCAGATAGAAGCACTGAAAGC-3' 5'-GACATCACTGTTGAGGACGAAAA-3'
4	<i>SMK1</i>	Cluster-7329.70484	5'-S-CCTCAAACGCCTCGCCTCACTTC-3' 5'-GCACCATTTCCAGCAATTCACA-3'
5	<i>SDH2</i>	Cluster-7329.78100	5'-CAAGCAATGCCCCGTATCCC-3' 5'-CCAACAGCCTGCCTAAAGAACA-3'
6	<i>BHLH4</i>	Cluster-7329.30018	5'-CCTCCTATCATACCAACCTCCTTT-3' 5'-GGCTTCGGCTATCTCCACTTT-3'
7	<i>ER4</i>	Cluster-7329.54336	5'-GGCGACGACGATGACGAA-3' 5'-TGCTCCGAATCAGCCAAAAG-3'
8	<i>NAC4</i>	Cluster-10565.0	5'-CAAACGACGGTCAACGACTACAT-3' 5'-TTCCACTTGGGCTCGCTCT-3'
9	<i>WRKY1</i>	Cluster-7329.88028	5'-GGCTTCAGAGGCGGTGG-3' 5'-AAGTTGCGGCTTTGGGTG-3'
10	<i>WRKY4</i>	Cluster-7329.53296	5'-CCACTCATCCTCAACAAACCCCT-3' 5'-CTTGCTGCGGCATACTACCC-3'
11	<i>MYB1</i>	Cluster-7329.58558	5'-CCTGGACATCTGAGGAGCATAG-3' 5'-CATTTGTTCTTCGGGCACTG-3'
12	<i>MYB12</i>	Cluster-7329.96022	5'-TTGAGGAGGATAAGGCATTTCG-3' 5'-CCCGAGAAGAAATAACCTGTGC-3'
13	<i>actin</i>	Previously identified	5'-AGGTCCAATGCTTTATC-3' 5'-CAGTCTTCTCCACCACAA-3'

3 结果与分析

3.1 掌叶大黄中有效成分的测定

掌叶大黄根、根茎、叶 HPLC 色谱分析结果如表 2 所示, 10 种对照品在掌叶大黄根、根茎、叶中均可检测到, 其中 5 种成分的含量在掌叶大黄的根、根茎、叶中差异不明显, 如儿茶素、大黄素-8-O-葡萄糖苷、芦荟大黄素、大黄酸、大黄素。大黄酚

的含量分析表明, 其在根、根茎、叶中的含量存在明显差异, 大黄酚-8-O-葡萄糖苷、没食子酸及番泻苷 B 在根及根茎的含量要显著大于其在叶中的含量, 大黄素甲醚在根茎与叶中的含量显著低于其在根中的含量。说明一年生的大黄次级代谢产物积累丰富, 其次级代谢活动活跃, 所采集材料可用于次级代谢途径关键基因分析。

表 2 掌叶大黄根、根茎、叶中 10 种指标成分分析

Table 2 Contents of 10 main chemical composition from leaves, rhizome, and roots of *R. palmatum*

部位	质量分数/%				
	没食子酸	儿茶素	番泻苷 B	大黄酚-8-O-葡萄糖苷	大黄素-8-O-葡萄糖苷
根	0.019 6±0.000 9a	1.066 9±0.183 0a	2.364 3±0.128 4a	0.582 4±0.183 9a	2.837 2±0.952 0a
根茎	0.012 2±0.002 2a	0.910 5±0.072 3a	2.263 1±0.057 6a	0.560 9±0.165 6a	2.893 2±0.216 2a
叶	0.009 2±0.000 8b	0.825 4±0.106 2a	1.085 3±0.021 8b	0.013 6±0.002 5b	1.592 2±0.827 0a

部位	质量分数/%				
	芦荟大黄素	大黄酸	大黄素	大黄酚	大黄素甲醚
根	0.005 2±0.000 7a	0.045 1±0.010 8a	0.046 5±0.001 8a	0.151 6±0.020 5a	0.436 6±0.017 7a
根茎	0.005 8±0.000 4a	0.039 5±0.027 1a	0.032 3±0.002 7a	0.087 3±0.027 2b	0.178 3±0.007 8b
叶	0.004 8±0.000 5a	0.017 8±0.008 5a	0.023 1±0.009 5a	0.006 8±0.001 2c	0.160 8±0.009 9b

不同字母表示差异性显著, * $P < 0.05$

different letters indicate significant differences, * $P < 0.05$

3.2 测序数据质控

利用 Illumina HiSeq™ 2000 测序平台, 共得到 28.04 Gb 的 clean read pairs 并且把所有数据都上传到 NCBI SRA 数据库, 得到登录号 SRP10855670。测序数据显示, 所有数据中的 GC 含量为 47.98%, Q_{20} 、 Q_{30} 值分别为 97.26%、92.59%。Trinity 无参组装获得 140 224 个 unigenes, 平均长度 1429 nt, 最长达到 15 389 nt, 最短序列为 201 nt, N_{50} 为 2045 nt。Unigenes 长度分布显示, 42 178 条 unigenes 长度超过 1000 nt, 33 679 条序列大于 2000 nt。根据统计结果可以发现测序结果真实可靠, 可以用于后续的分析。

3.3 Unigene 功能注释

将拼接得到的 unigene 与 7 大数据库进行比对, 注释结果如表 3 所示, 其中 Nr 数据库注释成功的 unigene 条数为 95 643 占总 unigene 数的 68.2%; Nt 数据库注释成功的 unigene 条数为 62 133 占总 unigene 数的 44.3%; KO 库注释成功的 unigene 条数为 39 641, 占总 unigene 数的 28.26%; SwissProt 库注释成功的 unigene 条数占 unigene 数的 52.98%; PFAM 库注释成功的 unigene 条数为 70 460, 占总 unigene 数的 50.24%; GO 库注释成功的 unigene 条数为 71 152, 占总 unigene 数的 50.74%, KOG 库注释成功的 unigene 条为 28 603 占总 unigene 数的 20.39%; 在 7 个数据库中至少有一个数据库注释成

表 3 各数据库 unigenes 数量注释信息统计

Table 3 BLAST analysis of unigene sequences in seven public databases

数据库	unigenes 数量	占比/%
非冗余蛋白库	95 643	68.20
核酸序列数据库	62 133	44.30
京都基因组百科全书数据库	39 641	28.26
蛋白质序列数据库	74 300	52.98
蛋白家族数据库	70 460	50.24
基因本体联合数据库	71 152	50.74
同源蛋白簇数据库	28 603	20.39
所有数据库中注释	15 683	11.18
至少在一个数据库中注释	102 626	73.18
unigenes 总量	140 224	100.00

功的 unigene 条数及占总 unigene 数的 73.18%。在所有以上 7 个数据库中同时注释成功的 unigene 条数占总 unigene 数的 11.18%。

3.4 基因表达水平分析

一般用 FPKM 值检测基因表达水平, FPKM 的优势在于把测序深度和基因长度对 reads 计数的影响都考虑进去。通过各样品基因的 FPKM 箱形图及密度分布图, 可以看出不同样品间基因总体表达量在分布度和离散度上表现出一定差异, 如图 1 所示, 说明掌叶大黄的根、根茎、叶的基因表达是有所差异的。

3.5 差异表达基因的筛选

根据 DEG 的筛选标准 $|\log_2(\text{Fold Change})| > 1$ 且

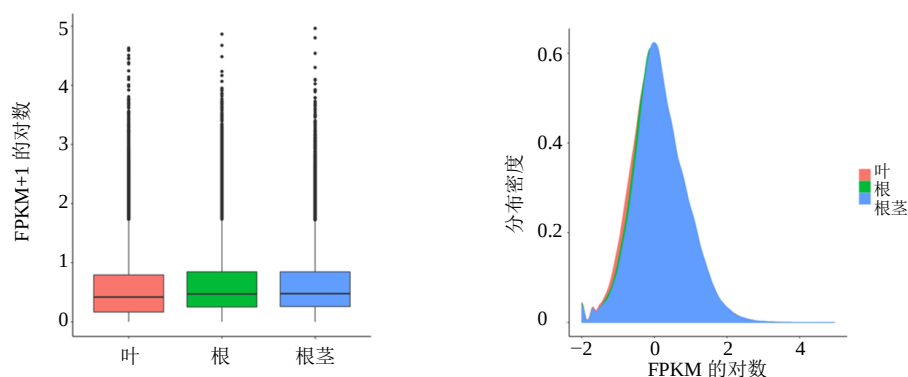


图 1 不同样品基因表达水平比对

Fig. 1 Comparison of genes expression of different samples of *R. palmatum*

q -value<0.005, 发现掌叶大黄的根与叶相比, 差异表达基因总数为 4175 个, 其中 2388 个基因表达上调, 1787 个基因表达下调; 根与根茎相比, 差异表达基因总数为 992 个, 其中 665 个基因表达上调, 327 个基因表达下调; 叶与根茎相比, 差异表达基因总数为 4469 个, 其中 2479 个基因表达上调, 1990 个基因表达下调 (图 2)。

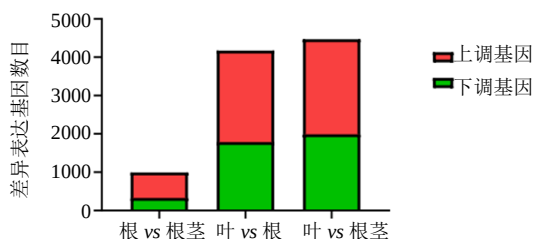


图 2 不同样品基因差异表达分析

Fig. 2 Gene differential expression analysis

3.6 差异基因 GO 功能富集分析

对掌叶大黄叶、根、根茎两两比对筛选出的 DEG 进行 GO 功能富集分析, 从生物过程 (biological process, BP) 和细胞组成 (cellular component, CC)、分子功能 (molecular function, MF) 3 个方面来描述差异基因的生物学功能。由于大黄的药用部位是根与根茎, 所以本实验只将叶与根及叶与根茎的对比。其中, 叶比根、叶比根茎中分别有 3142、3425 个差异基因被注释。在生物过程中, 差异基因都显著富集于代谢过程 (GO: 0008152) 和生物合成过程 (GO: 0009058), 叶和根间的差异基因还显著富集于单有机体代谢过程 (GO: 0044710); 差异基因在细胞组成中, 都显著富集于细胞 (GO: 0005623)、细胞组分 (GO: 0044464); 分子功能中, 差异基因主要都参与氧化还原酶活性 (GO: 0016491)、结构分子活性 (GO:

0005198)、核糖体的结构成分 (GO: 0003735)、四吡咯结合 (GO: 0046906)、血红素结合 (GO: 0020037) 等 (图 3)。

3.7 KEGG 富集分析差异表达基因

KEGG 富集结果显示, 在叶比根中, 差异基因被注释到 109 条通路中, 包括植物信号转导 (ko04075)、苯丙烷生物合成 (ko00940)、淀粉和蔗糖代谢 (ko00500)、苯丙氨酸代谢 (ko00360)、萜类骨架生物合成 (ko00900) 等; 叶比根茎中, 差异基因获得了 112 条通路的注释, 包括类胡萝卜素生物合成 (ko00906)、缬氨酸、亮氨酸和异亮氨酸生物合成 (ko00290)、磷酸肌醇代谢 (ko00562)、脂肪酸生物合成 (ko00061)、柠檬烯和蒎烯的降解 (ko00903) 等, 其中苯丙烷生物合成通路被显著富集, 33 个差异基因被注释到该通路中。差异基因中显著性富集的 Pathway 可能是蒽醌类化合物生物合成的主要生化代谢途径 (图 4)。

3.8 与蒽醌合成相关的关键差异表达基因筛选

参与大黄蒽醌类化合物骨架生物合成 MVA、MEP、莽草酸及聚酮途径的差异表达基因见表 4。MVA 途径中的关键酶有 AACT、HMGS、HMG-CoA HMGR 及 MPD。其中 AACT 和 MPD 在根中的表达量显著高于其在叶及根茎中的表达量, HMGS、HMGR 在叶中的表达量显著高于其在根及根茎中的表达量。

MEP 途径中, 关键酶的差异基因有 DXS、MCT/ispD、CMK/ispE 和 HDR/ispH。其中 DXS 和 MCT/ispD 在根中的表达量要显著高于其在叶及根茎中的表达量, MCT/ispD、HDR/ispH 在叶中的表达量要显著高于其在根及根茎中的表达量。

莽草酸途径中, 差异表达基因编码的酶有 DHQS、

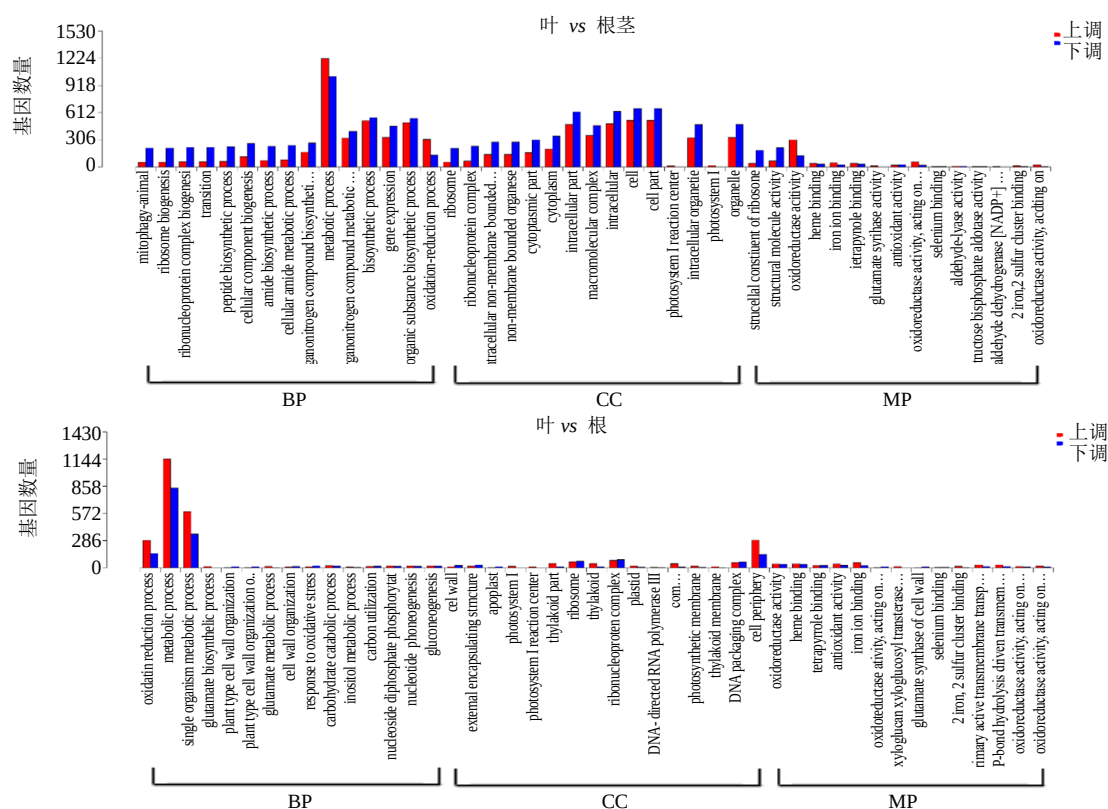


图 3 差异表达基因 GO 富集分析

Fig. 3 GO analysis of differentially expressed genes

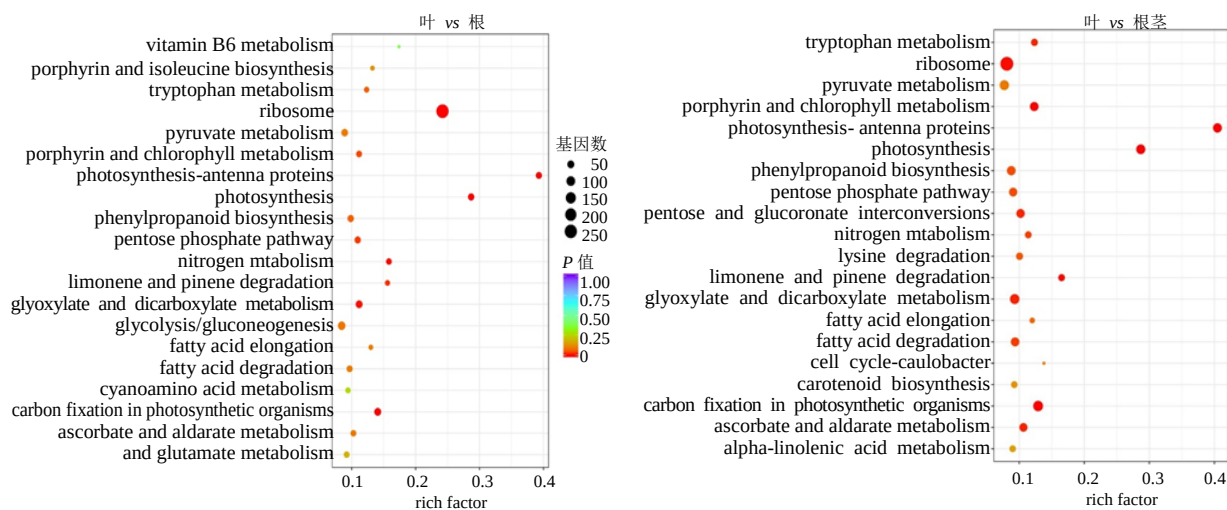


图 4 KEGG 通路富集散点图

Fig. 4 Scatter diagram of pathway enrichment analysis

SDH、SMK 及 MenB。这 4 个基因在叶中的表达量都显著高于其在根及根茎中表达量。聚酮途径中只检测到一个差异表达基因即 *PKSIII*，这个基因在叶中的表达量显著高于其在根及根茎中的表达量。

3.9 荧光定量检测分析

随机选取 12 个与蒽醌合成相关的基因作为检测对象，以掌叶大黄的 *actin* 基因作为内参基因，进行

qRT-PCR 检测。结果表明 11 个差异表达基因的 qRT-PCR 测定结果与转录组测序分析结果一致，*BHLH4* 的测定结果与转录组测序分析结果不一致(图 5)，可能是由于实验时使用了新提取的 RNA 样品，才导致使少数基因的表达趋势与差异表达分析中的 FPKM 值出现了差异。总体而言，本研究中的转录组分析数据得到的差异表达基因的信息是可信的。

表 4 关键差异基因的表达情况

Table 4 Expression of key differentially expressed genes

通路	基因名称	缩写	基因相对表达量		
			叶	根	根茎
MVA	乙酰-CoA 乙酰转移酶 (acetyl-CoA acetyltransferase)	AACT	0.41	7.50	4.47
	HMG-CoA 合酶 (HGM-CoA synthase)	HMGS	6.53	0	0
	HMG-CoA 还原酶 (HGM-CoA reductase)	HMGR	22.39	0.39	0.91
	MVPP 脱羧酶 (MVPP decarboxylase)	MPD	36.05	47.19	83.52
MEP	1-脱氧-D-葡萄糖-5-磷酸合成酶 (1-deoxy-D-xylulose-5-phosphate synthase)	DXS	10.66	33.84	28.35
	2-C-甲基-D-赤藓糖醇-4-磷酸胞苷酰基转移酶 (2-C-methyl-D-erythritol 4-phosphate cytidyltransferase)	MCT/ispD	12.19	56.43	33.05
	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/ispF	59.61	21.35	18.90
	4-羟基-3-甲基丁烯-2-烯基-1-二磷酸还原酶 (4-hydroxy-3-methylbut-2-enyl diphosphate reductase)	HDR/ispH	50.01	18.21	9.71
Shikimate	3-脱氧奎尼酸合成酶 (3-dehydroquinate synthase)	DHQS	22.60	0.14	0.13
	莽草酸脱氢酶 (shikimate dehydrogenase)	SDH	98.85	25.12	30.53
	莽草酸激酶 (shikimate kinase)	SMK	31.79	7.45	6.87
	萘酚酸合酶 (naphthoate synthase)	MenB	59.97	19.58	30.07
Polyketide	III 型聚酮合酶 (polyketide synthase III)	PKSIII	23.40	1.83	1.27

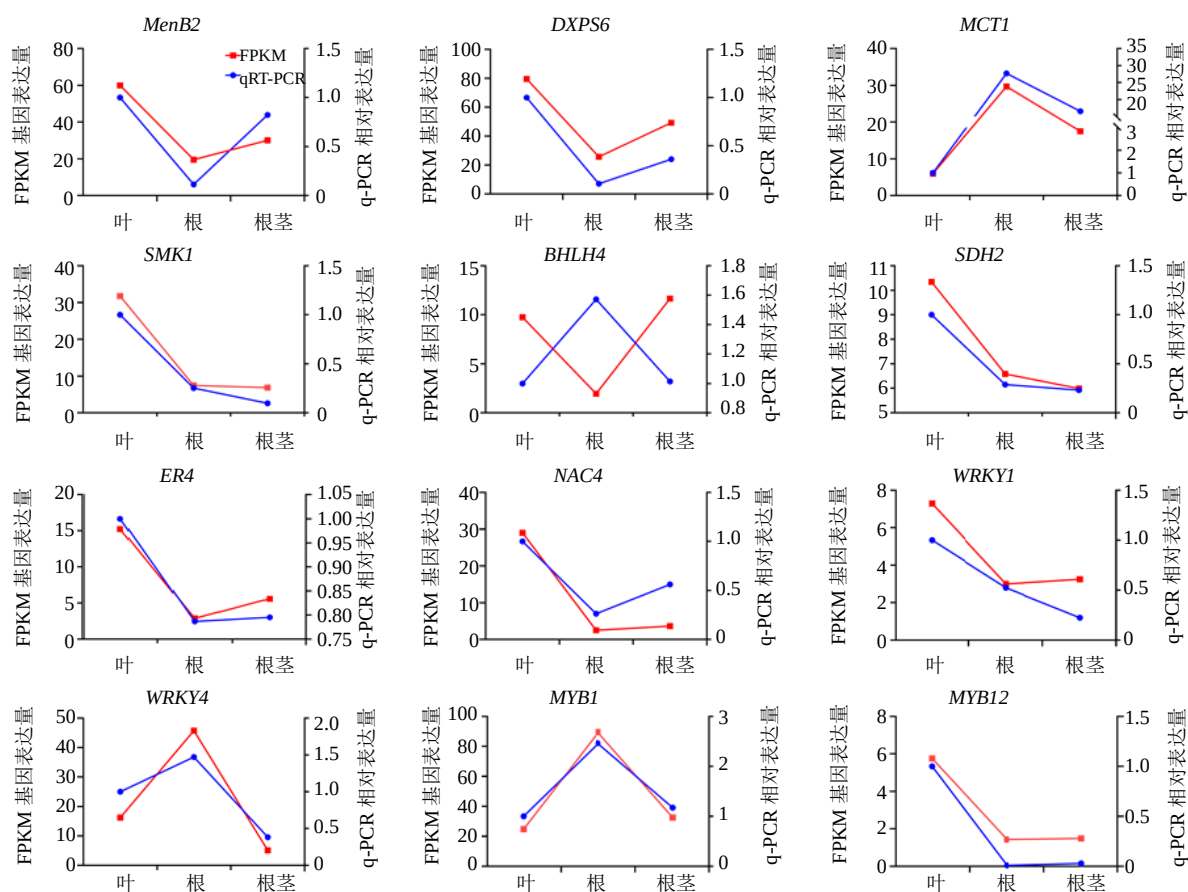


图 5 qRT-PCR 对转录组测序数据的验证

Fig. 5 Selective validation of transcriptome sequencing data by qRT-PCR

4 讨论

蒽醌类物质在生物体内的合成过程十分复杂, 整个过程涉及多条代谢途径产生的多种代谢产物。目前蒽醌类成分的生物合成途径尚未被完全阐明, 但普遍认为植物中蒽醌类物质生物合成主要源于 2 条代谢途径: 莽草酸/邻琥珀酰苯甲酸途径 (shikimate/osuccinylbenzoic acid route) 和聚酮途径 (polyketide pathway) [17-18]。本研究对一年生掌叶大黄的根、根茎、叶进行转录组测序, 筛选差异表达基因, 并结合 GO、KEGG 富集分析, 有助于在转录组水平上构建蒽醌类物质生物合成网络, 筛选参与蒽醌类物质生物合成关键基因。

莽草酸/邻琥珀酰苯甲酸途径合成蒽醌母核的前体来自于多种代谢途径, 该合成途径主要涉及莽草酸途径、TCA 循环、MVA 途径和 MEP 途径。研究发现 MVK、HMG-CoA 还原酶是 MVA 通路中的关键酶。有研究发现, 在植物体内过表达 MVK 可以提高异戊二烯类物质的含量 [19]。用抑制剂抑制 HMG-CoA 还原酶活性后, 番茄的生长速度会受到明显影响, 当加入 MVA 后, 番茄又恢复生长, 进一步说明 HMG-CoA 在 MVA 途径中起到重要作用 [20]。本研究中, 观察到了 HMG-CoA 还原酶编码基因的差异表达现象, 但并没有筛选到 MVK 基因的差异表达, 这可能与植物的生长状态有关。

有研究者认为, 可以通过调控 MEP 途径影响蒽醌的生物合成 [21]。在拟南芥中, 使 DXS 基因过表达, 发现植物中多种类异戊二烯终产物 (包括叶绿素、类胡萝卜素、生育酚和 ABA) 的含量均升高, 说明 DXS 是 MEP 途径中关键的限速酶 [22]。在 *Morinda citrifolia* 细胞系过表达 DXS 基因, 与对照细胞系相比, 过表达 DXS 基因的细胞系蒽醌产量提高了约 24%, 说明 DXS 对蒽醌的生物合成具有十分显著的影响 [23]。在掌叶大黄的转录组中, 我们发现了差异表达的 DXS 基因, 其在根中的表达量要显著高于在叶与根茎中的表达量, 这可能与蒽醌类化合物积累于掌叶大黄的根部有关系。

聚酮途径是由 PKS 参与反应。PKS 分为 3 大类: PKS I、PKS II 和 PKS III [24]。其中, PKS I、PKS II 多存在于真菌与细菌中, PKS III 不仅在细菌、真菌中广泛分布, 而且在植物体内也大量存在。PKS III 在生物体内主要参与具有生物活性的聚酮类次级代谢产物的合成, 例如黄酮类物质、二苯乙烯类物质的生物合成 [25-27]。在本研究中筛选得到差异表达的

PKS III 基因, 其在叶中的表达量要显著高于其在根与根茎中的表达量, 与测得的蒽醌类化合物的在根、茎与根茎中积累量没有显著相关性, 这可能与聚酮途径的复杂性有关, 说明 PKS III 单个基因的表达量并不能影响最终产物的累积量。

本研究利用 Illumina 二代高通量测序平台开展了掌叶大黄根、根茎及叶的转录组分析来研究掌叶大黄 3 个部位差异表达基因, 成功筛选出差异表达的编码 MVA、MEP、莽草酸及聚酮途径的多个关键酶差异表达基因, 这些差异基因与蒽醌的生物合成有重要关系, 为后蒽醌类物质生物合成过程中关键酶及其基因的分离、克隆和表达特征的研究及明确蒽醌类物质生物合成途径中调控因子奠定物质基础。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 李枫, 武红, 姚丽, 等. 大黄素下调膜联蛋白 A2 增强子宫内腺癌 HEC-1B 细胞顺铂敏感性 [J]. 中药材, 2022(08): 1949-195.
- [2] 李志清, 王善龙, 郭巍, 等. 异丹叶大黄素对前列腺癌细胞血管生成相关蛋白的影响 [J]. 辽宁中医杂志, 2021, 48(10): 155-158.
- [3] 王玉, 杨雪, 夏鹏飞, 等. 大黄化学成分、药理作用研究进展及质量标志物的预测分析 [J]. 中草药, 2019, 50(19): 4821-4837.
- [4] Shimizu Y, Ogata H, Goto S. Type III polyketide synthases: Functional classification and phylogenomics [J]. *Chembiochem*, 2017, 18(1): 50-65.
- [5] Tzin V, Galili G. New insights into the shikimate and aromatic amino acids biosynthesis pathways in plants [J]. *Mol Plant*, 2010, 3(6): 956-972.
- [6] Han Y S, van der Heijden R, Lefeber A W M, et al. Biosynthesis of anthraquinones in cell cultures of *Cinchona* 'Robusta' proceeds via the methylerythritol 4-phosphate pathway [J]. *Phytochemistry*, 2002, 59(1): 45-55.
- [7] 梁伟, 孙嘉辰, 郭凤霞, 等. 基于莽草酸/邻琥珀酰苯甲酸途径和聚酮途径合成蒽醌类化合物研究进展 [J]. 中草药, 2020, 51(7): 1939-1950.
- [8] 颜永刚, 尹立敏, 王红艳, 等. HPLC 法同时测定大黄叶中 9 种成分 [J]. 中草药, 2016, 47(13): 2360-2364.
- [9] Grabherr M G, Haas B J, Yassour M, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome [J]. *Nat Biotechnol*, 2011, 29(7): 644-652.

- [10] Davidson N M, Oshlack A. Corset: Enabling differential gene expression analysis for *de novo* assembled transcriptomes [J]. *Genome Biol*, 2014, 15: 7410.
- [11] Li B, Dewey C N. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome [J]. *BMC Bioinform*, 2011, 12(1): 323.
- [12] Mortazavi A, Williams B A, McCue K, *et al*. Mapping and quantifying mammalian transcriptomes by RNA-seq [J]. *Nat Methods*, 2008, 5(7): 621-628.
- [13] Anders S, Huber W. Differential expression analysis for sequence count data [J]. *Genome Biol*, 2010, 11(10): R106.
- [14] Young M D, Wakefield M J, Smyth G K, *et al*. Gene ontology analysis for RNA-seq: Accounting for selection bias [J]. *Genome Biol*, 2010, 11(2): R14.
- [15] Mao X Z, Cai T, Olyarchuk J G, *et al*. Automated genome annotation and pathway identification using the KEGG orthology (KO) as a controlled vocabulary [J]. *Bioinformatics*, 2005, 21(19): 3787-3793.
- [16] Livak K J, Schmittgen T D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method [J]. *Methods*, 2001, 25(4): 402-408.
- [17] Yamazaki M, Mochida K, Asano T, *et al*. Coupling deep transcriptome analysis with untargeted metabolic profiling in *Ophiorrhiza pumila* to further the understanding of the biosynthesis of the anti-cancer alkaloid camptothecin and anthraquinones [J]. *Plant Cell Physiol*, 2013, 54(5): 686-696.
- [18] Rama Reddy N R, Mehta R H, Soni P H, *et al*. Next generation sequencing and transcriptome analysis predicts biosynthetic pathway of sennosides from *Senna* (*Cassia angustifolia* Vahl.), a non-model plant with potent laxative properties [J]. *PLoS One*, 2015, 10(6): e0129422.
- [19] Anthony J R, Anthony L C, Nowroozi F, *et al*. Optimization of the mevalonate-based isoprenoid biosynthetic pathway in *Escherichia coli* for production of the anti-malarial drug precursor amorpha-4, 11-diene [J]. *Metab Eng*, 2009, 11(1): 13-19.
- [20] Narita J O, Grissem W. Tomato hydroxymethylglutaryl-CoA reductase is required early in fruit development but not during ripening [J]. *Plant Cell*, 1989, 1(2): 181-190.
- [21] 尹松松, 苏秀娟, 龚林涛, 等. 薰衣草 1-脱氧-D-木酮糖-5-磷酸还原异构酶基因 LaDXR 的克隆及表达分析 [J]. 分子植物育种, 2021, 19(7): 2193-2199.
- [22] 林谷音, 周佩娜, 尹梦娇, 等. 荆芥 1-脱氧-D-木酮糖-5-磷酸合成酶基因克隆及生物信息学分析 [J]. 中草药, 2021, 52(2): 527-537.
- [23] 魏麟, 伍贤进, 李胜华, 等. 鱼腥草 1-脱氧-D-木酮糖-5-磷酸合成酶 1 基因克隆与表达分析 [J]. 中草药, 2014, 45(11): 1607-1612.
- [24] Staunton J, Weissman K J. Polyketide biosynthesis: A millennium review [J]. *Nat Prod Rep*, 2001, 18(4): 380-416.
- [25] Austin M B, Noel J P. The chalcone synthase superfamily of type III polyketide synthases [J]. *Nat Prod Rep*, 2003, 20(1): 79-110.
- [26] Juvvadi P R, Seshime Y, Kitamoto K. Genomics reveals traces of fungal phenylpropanoid-flavonoid metabolic pathway in the filamentous fungus *Aspergillus oryzae* [J]. *J Microbiol*, 2005, 43(6): 475-486.
- [27] Li L, Aslam M, Rabbi F, *et al*. PpORS, an ancient type III polyketide synthase, is required for integrity of leaf cuticle and resistance to dehydration in the moss, *Physcomitrella patens* [J]. *Planta: An Internat J Plant Biology*, 2018, 25: 23.

[责任编辑 时圣明]