

两品种当归转录组分析及抗旱关键基因挖掘

康舒淇¹, 朱田田^{1,2,3,4*}, 晋玲^{1,2,3,4}, 刘天乐¹, 张菁¹, 张明惠¹, 徐丽¹, 张帅¹ (1. 甘肃中医药大学药学院, 兰州 730000; 2. 西北中藏药省部共建协同创新中心, 兰州 730000; 3. 甘肃省珍稀中药资源评价与保护利用工程研究中心, 兰州 730000; 4. 陇药产业创新研究院, 兰州 730000)

摘要:目的 对当归 [*Angelica sinensis* (Oliv.) Diels] 主栽品种“岷归 1 号”和“岷归 2 号”进行转录组测序差异比较分析, 挖掘当归响应干旱胁迫关键基因。方法 以两品种当归的新鲜叶片和根组织为研究材料构建 cDNA 文库, 利用二代高通量测序平台 Illumina HiSeq™ 4000 进行测序分析, 从差异表达基因中筛选响应干旱胁迫的关键基因。结果 转录组测序共获得 584 423 236 条高质量序列, Q20 (碱基量 ≥ 20%) 与 Q30 (碱基量 ≥ 30%) 占比分别在 97.47%、92.64% 以上, 鸟嘌呤和胞嘧啶所占比例 (GC) 含量是 42.78% ~ 43.15%。两品种当归叶片和根中共筛选得到 1 894 个差异表达基因 (DEGs), 数目分别为 674 和 1 220 个, 共有差异表达基因为 338 个。基因本体 (GO) 富集结果表明, 两品种当归在同一组织部位中的 DEGs 注释分类主要包括细胞过程、代谢过程和催化活性等功能。京都数据库与基因组百科全书 (KEGG) 分析发现差异表达基因均在植物-病原互作、植物-MAPK 信号通路、苯丙烷生物合成和植物激素信号转导等通路中显著富集, 精细分类注释结果与 GO 及 KEGG 分析结果趋势相符合。基于功能注释结果, 挖掘到抗旱相关基因 60 个, 选取 *HVA22C*、*KRPI*、*PUB23*、*DREB1B*、*Bp10* 通过实时荧光定量聚合酶链反应 (qRT-PCR) 验证其表达量, 结果表明, 其基因表达水平与转录组测序基因表达趋势一致。结论 两品种当归在脱落酸调节、渗透调节、活性氧清除和其他功能蛋白质调节等抗旱途径方面存在一定差异性, 筛选出的抗旱基因可为进一步研究当归响应干旱胁迫的分子机制提供数据参考。

关键词: 当归; 品种; 差异表达基因分析; 抗旱基因; 实时荧光定量聚合酶链反应

doi:10.11669/cpj.2024.12.005 中图分类号: R282 文献标志码: A 文章编号: 1001-2494(2024)12-1100-10

Transcriptome Analysis of Two Cultivars of *Angelica sinensis* and Excavation of Key Gene for Drought Resistance

KANG Shuqi¹, ZHU Tiantian^{1,2,3,4*}, JIN Ling^{1,2,3,4}, LIU Tianle¹, ZHANG Jing¹, ZHANG Minghui¹, XU Li¹, ZHANG Shuai¹ (1. College of Pharmacy, Gansu University of Traditional Chinese Medicine, Lanzhou 730000, China; 2. Northwest Collaborative Innovation Center for Traditional Chinese Medicine Coconstructed by Gansu Province and MOE of PRC, Lanzhou 730000, China; 3. Engineering Research Center for Evaluation, Protection, and Utilization of Rare Traditional Chinese Medicine Resources, Lanzhou 730000, China; 4. Gansu Pharmaceutical Industry Innovation Research Institute, Lanzhou 730000, China)

ABSTRACT: OBJECTIVE To excavate the key genes of drought stress response of *A. sinensis*, and carry out the comparative analysis with the transcriptome data of the main cultivars ‘Mingui 1’ and ‘Mingui 2’ of *Angelica sinensis* (Oliv.) Diels. **METHODS** With the fresh leaf and root tissues of two cultivars of *A. sinensis* as materials, a cDNA library was constructed. The Illumina HiSeq™ 4000, a second-generation high-throughput sequencing platform, was used for sequencing analysis, and key enzyme genes in response to drought stress were screened from differentially expressed genes (DEGs). **RESULTS** A total of 584 423 236 clean reads were obtained from transcriptome sequencing, in which the percentage of Q20 (base amount ≥ 20%) and Q30 (base amount ≥ 30%) were above 97.47% and 92.64%, and the GC content ranged from 42.78% to 43.15%. A total of 1 894 DEGs were screened from the leaves and roots of two cultivars of *A. sinensis*, the numbers of which were 674 and 1 220, respectively, and they had 338 shared DEGs. The GO enrichment results showed that the annotation classification of DEGs of two cultivars of *A. sinensis* in the same tissue part mainly included cellular process, metabolic process and catalytic activity. KEGG analysis found that the DEGs were significantly enriched in plant-pathogen interaction, MAPK signaling pathway-plant, phenylpropanoid biosynthesis and plant hormone signal transduction pathways. The detailed classification annotation results were consistent with the trends of GO and KEGG analysis. Based on the

基金项目: 甘肃省高校青年博士基金项目资助 (2023QB-094); 甘肃省科技重大专项资助 (23ZDFA013-1); 甘肃省教育厅“双一流”科研重点项目资助 (GSSYLXM-05); 甘肃省科技计划项目资助 (20JR5RA182); 甘肃中医药大学科学研究与创新基金项目资助 (2021KCZD-4); 西北中藏药省部共建协同创新中心开放基金项目资助 (Xbzzy202207)

作者简介: 康舒淇, 女, 硕士研究生 研究方向: 中药资源保护、评价与可持续利用研究 * **通讯作者:** 朱田田, 女, 副教授, 硕士生导师 研究方向: 中药资源评价与分子生药学研究 Tel: (0931) 5161169

functional annotation results, 60 drought resistance genes were excavated. *HVA22C*, *KRPI*, *PUB23*, *DREB1B* and *Bp10* were selected to verify their expression levels by qRT-PCR. and the results showed their genes expression level were consistent with the transcriptome sequencing gene expression trends. **CONCLUSION** The two cultivars of *A. sinensis* have some differences in drought resistance pathways such as abscisic acid regulation, osmoregulation, scavenging of reactive oxygen species and regulation of other functional proteins, and the screened drought resistance genes can provide data references for further research on the molecular mechanisms of *A. sinensis* in response to drought stress.

KEY WORDS: *A. sinensis*; cultivar; DEGs analysis; drought resistance gene; qRT-PCR

当归为伞形科植物当归 [*Angelica sinensis* (Oliv.) Diels] 的干燥根, 是我国传统大宗中药材之一, 具有补血活血、调经止痛、润燥通肠的功效^[1]。当归目前以栽培品种为主, 产区主要分布在甘肃、云南、四川等省^[2], 其中甘肃岷县种植历史悠久、产量大, 药材质量优, 为著名的道地药材主产区^[3]。“岷归 1 号”和“岷归 2 号”是通过系统育种法选育出的优良品种, “岷归 1 号”茎呈紫色, 为目前种植区域最大的品种, “岷归 2 号”茎呈绿色, 具高产稳产、抗逆能力强等特点, 是近年来被逐步推广种植的品种^[4-5]。研究发现两品种当归茎叶色差异与类黄酮代谢物有关, 而此类化合物在植物发育和防御响应方面也存在重要作用^[6]。在抗性适应能力方面, “岷归 1 号”对辐照胁迫敏感性较强, 但“岷归 2 号”对其耐受性更强^[7]。

外界环境给予的各种生物和非生物胁迫都会对药用植物造成一定负面影响, 其中干旱作为极易发生的胁迫之一, 可能会严重制约植物生长发育^[8]。在当归种植过程中, 除了抽薹开花及根际土壤病原微生物导致连作障碍等问题会对其产量和品质造成影响^[9-10], 近年来全球气候变暖、极端气候等问题的出现均导致干旱风险增加, 且栽培环境属西北干旱和半干旱气候区, 故干旱也是限制当归生长发育及产量的胁迫因素之一^[11], 会使得植物叶片萎蔫、相对含水量降低及活性氧代谢失常等, 进而导致氧化损伤至死亡; 同时干旱胁迫也可能改变植物内源性激素合成代谢、信号转导通路及各种生物代谢调控等, 进而影响药材质量^[12]。当前对当归的研究主要集中于种质资源评价、药效成分和药理作用研究等方面^[13], 针对当归的干旱胁迫还未见系统性研究。

课题组前期对干旱胁迫后的两品种当归植株形态做差异对比, 发现在同一程度干旱胁迫下, “岷归 1 号”叶片卷曲程度及植株下垂状况较“岷归 2 号”更为明显, 推测“岷归 2 号”抗旱适应能力可能略强于“岷归 1 号”(图 1 A)。本研究通过对两品种当归“岷归 1 号”和“岷归 2 号”进行转录组测序比较分

析, 从中挖掘当归抗旱关键基因, 可为研究当归响应干旱胁迫的分子机制提供候选基因, 旨在为当归药材生态种植评价、品质提升及耐旱性品种培育提供研究基础。

1 材料

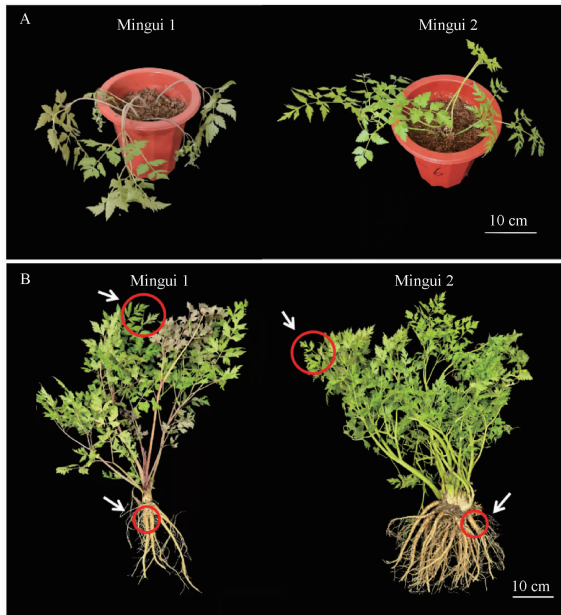
1.1 植物材料

当归全长转录组参考序列数据已上传于国家生物技术信息中心 (NCBI) (Access: PRJNA782300)^[6]; 选取同一时间种植且生境条件相同的成药生长期“岷归 1 号”和“岷归 2 号”新鲜叶片、根组织为研究材料, 取样部位为顶端叶片及根部膨大处 (图 1B), 样品均于 2022 年 9 月 18 日采自甘肃省定西市临洮县 (海拔 2 708 m; 35°37'52.22''N, 104°7'10.13''E)。经甘肃中医药大学中药资源教研室晋玲教授鉴定为伞形科植物当归 [*A. sinensis* (Oliv.) Diels]。样本组织采集后用水清洗干净, 投入液氮中带回实验室置于 -80 °C 超低温冰箱保存, 每个组织样品设置 3 个生物学重复, 用于转录组分析及差异表达基因验证, 样品编号信息见表 1。

1.2 仪器与试剂

1.2.1 仪器 FTC-3000P 荧光定量 PCR 仪 (加拿大 FunglynBiotech 公司); DYY-7C 微型转移电泳仪 (北京六一仪器厂); c300 多功能分子成像系统 (美国 Azure Biosystems 公司); MYSPI6 型台式高速冷冻离心机 (美国 Thermo Fisher Scientific 公司); Tone 96g 梯度聚合酶链式反应 (PCR) 扩增仪 (德国 analytik-jena 公司); EXF40086V 超低温冰箱 (北京五洲东方科技发展有限公司)。

1.2.2 试剂 Kit R6827 植物 RNA 提取试剂盒 (美国 Omega 公司); FastKing RT Kit-KR116 试剂盒 (北京天根生化科技有限公司); 合成引物 (上海生物工程股份有限公司); Q341 型逆转录聚合酶链式反应 (RT-PCR) 试剂盒 (南京诺唯赞生物科技股份有限公司)。



A - 植株形态对比; B - 取样部位, 箭头所指为取样部位。
A - plant morphological comparison; B - sampling parts, the arrow points to the sampling parts.

图1 “岷归1号”和“岷归2号”在同一程度干旱胁迫下的植株形态对比及取样部位

Fig. 1 Plant morphological comparison of ‘Mingui 1’ and ‘Mingui 2’ under the same degree of drought stress and sampling parts

表1 当归样品信息表

Tab. 1 Information sheet for *A. sinensis*

Cultivars name	Tissue site	Sample number (n = 3)
Mingui 1 (M1)	Leaf	M1L-1, M1L-2, M1L-3
	Root	M1R-1, M1R-2, M1R-3
Mingui 2 (M2)	Leaf	M2L-1, M2L-2, M2L-3
	Root	M2R-1, M2R-2, M2R-3

2 方法

2.1 总 RNA 提取、cDNA 文库构建及 RNA-seq 测序

采用 Trizol 法提取两品种当归叶片、根组织总 RNA, Agilent 2100 Bioanalyzer 检测总 RNA 样品的质量, 并使用琼脂糖凝胶电泳检验是否存在污染, 检验合格的 RNA 样品用于 cDNA 文库的构建。将 mRNA 打断为片状并以之为模板反转录为双链 cDNA, QiaQuick-PCR 试剂盒纯化片段, 经修复末端, 加 A 尾连接测序接头, 筛选片段进行 PCR 扩增以完成当归叶片及根组织样品文库构建。最后, 利用 Gene Denovo 生物技术公司的 Illumina HiSeq™ 4000 进行测序分析。

2.2 测序数据质量过滤、组装及基因表达分析

利用 fastp^[14] (Version 0.18.0) 将原始序列数据 (raw reads) 过滤得到质量值 Q20 (碱基量 ≥ 20%) 与 Q30 (碱基量 ≥ 30%) 合格、鸟嘌呤和胞嘧啶所占比例 (GC) 含量及分布符合预期的高质量序列 (clean reads); clean reads 使用 Trinity 软件^[15] 进行拼接组装获得 unigenes; 以课题组前期实验得到的全长转录本为参考, 运用 RSEM (Version 1.2.19) 软件^[16] 对各样品进行基因表达定量, 相对表达水平以 RPKM 值 (reads per kilobase per million mapped reads) 表示, 计算出的基因表达量可直接比较两品种当归间基因表达差异, 用于后续分析。

2.3 差异表达基因分析

使用 DESeq2 (version 1.20.0)^[17] 软件对两品种当归同一组织部位进行转录水平差异比较, 本研究设置 2 个差异组别, 叶片组织为 M1L vs M2L, 根组织为 M1R vs M2R, 以 $|\log_2(\text{Fold-Change})| > 1$ 和错误发现率 (false discovery rate, FDR) < 0.05 为差异表达基因 (differentially expressed genes, DEGs) 筛选条件。基于筛选结果, 通过基因本体 (Gene Ontology, GO) 数据库与京都基因和基因组百科全书 (Kyoto Encyclopedia of Genes and Genomes, KEGG) 数据库对 DEGs 进行 GO 功能分类统计和 KEGG 通路富集分析^[18-19]。为了进一步挖掘注释 DEGs 种类功能, 将 DEGs 在 Swiss-Prot (Swissprot Protein Sequence Database) 数据库中比对, 利用 Uniprot (Universal Protein) 数据库对已鉴定和去重后 DGEs 的生物学功能和分子功能进行精细分类统计。

2.4 实时荧光定量聚合酶链式反应 (qRT-PCR) 表达量验证

为明确转录组数据的可靠性, 基于差异分析结果, 选取 5 个关键 DGEs (*HVA22C*、*KRP1*、*PUB23*、*DREB1B*、*Bp10*) 进行 qRT-PCR 验证。运用 NCBI Primer-BLAST 网站设计引物, 委托生工生物工程 (上海) 股份有限公司合成, 引物详情见表 2。研究选用 *ACT* (*actin*) 作为内参基因^[20], 根据诺唯赞 RT-PCR 试剂盒 (Q341) 说明书配制反应体系及设置反应程序以完成 qRT-PCR 检测, 体系含 2 × ChamQ SYBR qPCR Master Mix 10 μL, 正、反引物各 0.4 μL, 双蒸 H₂O 5.2 μL, cDNA 模板 4 μL, 总体积为 20 μL; 反应程序为 95 °C 90 s, 95 °C 5 s, 60 °C 15 s, 72 °C 20 s, 40 个循环。采用 2^{-ΔΔCt} 法^[21] 计算基因相对表达量。

表 2 实时荧光定量聚合酶链式反应 (qRT-PCR) 扩增特异引物

Tab. 2 Specific primers of qRT-PCR amplification

Gene name	Primer sequence (5-3)	Amplification length/bp
<i>HVA22C</i>	F:TTCAACGGAGCTAGTATTGCCTACG R:TCTCCAAAGCCTCAGTTCATTCTC	175
<i>KRP1</i>	F:GCTGATGTGAATGGAGATGGAGTTC R:CTTCTTCAAGCCAAACCTCAGATTC	106
<i>PUB23</i>	F:ACGGCATCCAAAGAATCCCAACTC R:TAACCCGCTCTGCTGCTTCC	175
<i>DREB1B</i>	F:GTGTTGCCAGCTTCCAGTAATCC R:TCTACGCCTCACTCCTCTGTA AAC	93
<i>Bp10</i>	F:TCCAAGACCCAACCCACAAGG R:AGAGGTATCCGATGCTATGAAGG	147
<i>ACT</i>	F:TGGTATTCTGCTGGATTCTGCT R:TGAGATCACCAACCAAGG	109

表 3 当归转录组测序数据质量分析

Tab. 3 Quality analysis of transcriptome sequencing data of *A. sinensis*

No.	Raw reads	Clean reads/%	Clean reads/bp	Q20/%	Q30/%	GC/%	Mapped reads	Mapped ratio/%
M1L-1	45 403 718	45 123 184 (99.38)	6 673 603 863	97.97	93.75	42.83	40 424 439	89.59
M1L-2	56 479 970	56 136 434 (99.39)	8 310 128 766	97.79	93.39	42.95	50 542 876	90.04
M1L-3	51 450 794	51 103 374 (99.32)	7 610 806 624	97.64	93.00	42.91	45 316 441	88.68
M2L-1	49 972 658	49 665 860 (99.39)	7 374 089 668	97.87	93.53	43.14	44 735 242	90.07
M2L-2	47 873 956	47 583 270 (99.39)	7 024 011 837	97.91	93.66	43.10	41 434 239	87.08
M2L-3	49 948 026	49 632 084 (99.37)	7 375 539 235	97.47	92.64	42.93	44 482 152	89.62
M1R-1	46 193 770	45 851 916 (99.26)	6 790 909 357	97.60	92.96	43.15	37 046 874	80.80
M1R-2	48 730 572	48 436 458 (99.40)	7 068 058 117	97.95	93.82	43.04	38 902 366	80.32
M1R-3	46 580 816	46 266 062 (99.32)	6 862 726 173	97.69	93.13	42.91	37 217 870	80.44
M2R-1	43 229 068	42 923 476 (99.29)	6 341 975 820	97.58	92.88	43.01	34 686 284	80.81
M2R-2	55 754 104	55 400 420 (99.37)	8 232 259 847	97.58	92.90	42.78	42 144 671	76.07
M2R-3	46 570 526	46 300 698 (99.42%)	6 884 153 864	97.69	93.10	42.95	36 909 107	79.72

3.2 差异表达基因分析

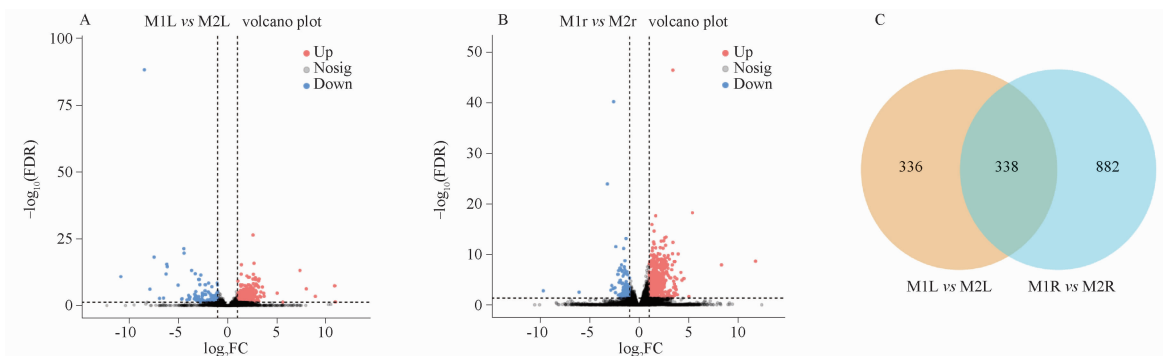
火山图中, $\log_2 FC > 1$ 为表达上调, $\log_2 FC < -1$ 则为表达下调。由图 2A 和 B 可见, M1L vs M2L 中上调表达基因 (up-regulated, UR) 553 个, 下调表达基因 (down-regulated, DR) 121 个; M1R vs M2R 中 UR 1 069

3 结果与分析

3.1 测序数据质量分析

对两品种当归叶片、根组织进行测序分析, 经测序共获得 584 423 236 条 clean reads, 均占 raw reads 的 99.26% 以上。所有测序数据 Q20 与 Q30 占比分别在 97.47%、92.64% 以上; GC 含量范围是 42.78% ~ 43.15%; 将各样品 clean reads 进行序列比对, 两品种当归叶片和根中分别有 136 283 756、130 651 633、113 167 110 和 113 740 062 条 mapped reads 可定位到基因上, mapped ratio 均在 88.68%、87.08%、80.32% 及 76.07% 以上, 说明过滤后碱基组成及分布较平衡, 数据质控良好, 可用于后续准确分析 (表 3)。

个, DR 151 个, 共有 DEGs 中 UR 324 个, DR 14 个, 说明相同生态环境下的 M1 和 M2 差异表达较大, 同一组织部位中 UR 的数量略多于 DR。基于筛选条件, 共检测获得 1 894 个 DEGs, 分布于叶片中有 674 个 DEGs, 根中有 1 220 个, 其中 338 个为共有 DEGs (图 2C)。



A - 叶片组织差异组别; B - 根组织差异组别; C - 韦恩图; 红色、蓝色和灰色分别表示显著上调、显著下调及非显著 DEGs。

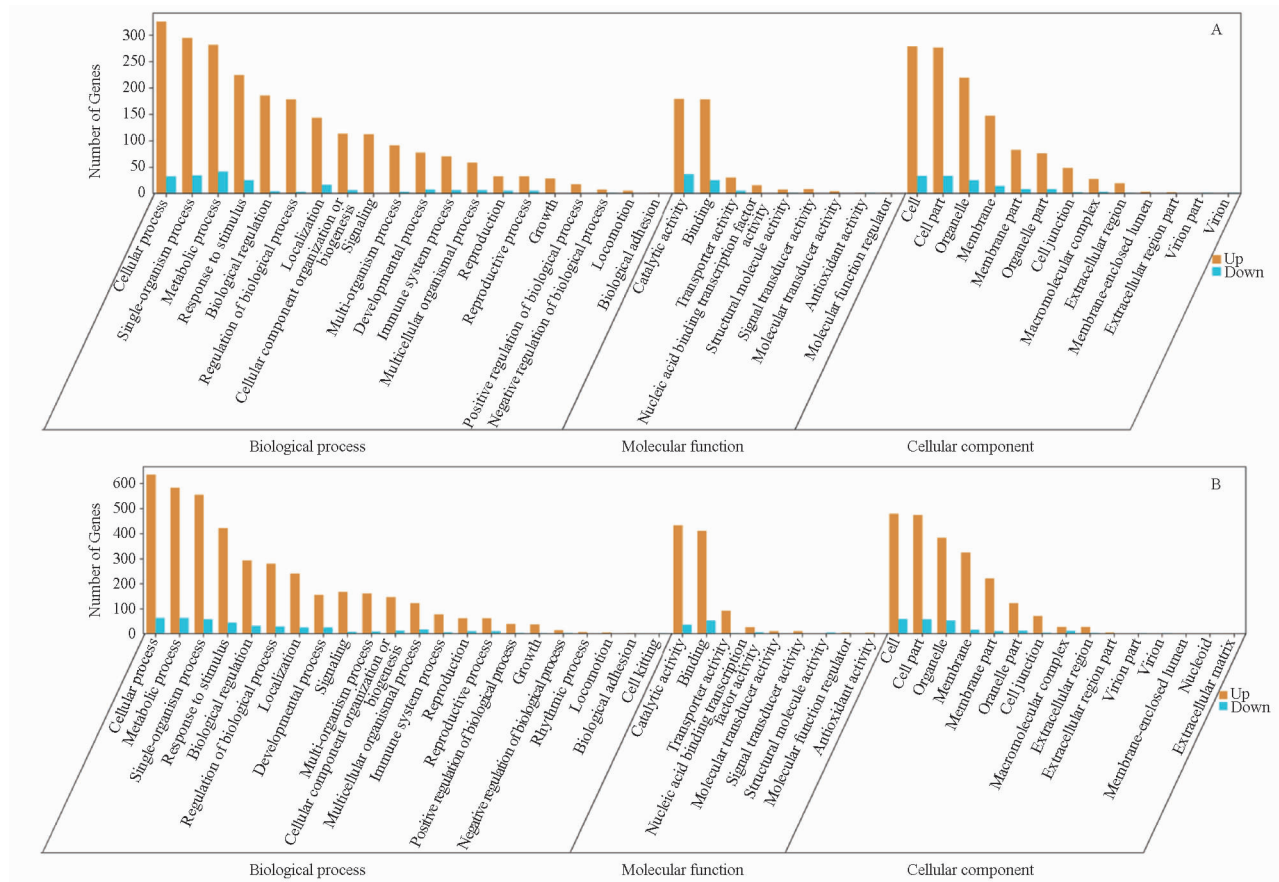
A - Leaf tissue difference group; B - Root tissue differential group; C - Venn diagram; Red, blue and grey indicate significantly up-regulated, significantly down-regulated and non-significant DEGs, respectively.

图 2 “岷归 1 号”和“岷归 2 号”差异表达基因 (DEGs) 火山图和韦恩图

Fig. 2 Volcano plot and Venn diagram of DEGs in ‘Mingui 1’ and ‘Mingui 2’

3.2.1 GO 富集分析 基于 GO 数据库注释基因功能,比较不同差异组别间的差异倍数,将差异基因集(M1L vs M2L、M1R vs M2R)比对注释于参与生物过程(biological process)、细胞组分(cellular component)、分子功能(molecular function)3个大类别。通过GO功能富集分析,可进一步明确两品种当归DEGs主要集中哪些功能类别。GO富集分析结果显示(图3),M1L vs M2L与M1R vs M2R全部DEGs分为42和46种功能亚类别,其中参与生物过程在两差异组别注释数目为20和22种,主要分布于细胞过程(cellular process)、代谢过程(metabolic process)和单一有机体过程(single-organism process),M1L vs M2L中UR与DR数目分别为325、

281、294及32、41、34个,M1R vs M2R中UR与DR数目分别为634、582、554及63、63、58个。两差异组别注释于细胞组分数目为13和15种,排名前三的为细胞(cell)、细胞部分(cell part)及细胞器(organelle),M1L vs M2L中UR与DR数目分别为278、276、219及33、33、25个,M1R vs M2R中UR与DR数目分别为478、473、382及59、58、52个。两差异组别比对到分子功能的富集数目均为9种,主要集中在催化活性(catalytic activity)、结合蛋白(binding)及转运活性(transporter activity)中,M1L vs M2L中UR与DR数目分别为179、178、30及36、25、5个,M1R vs M2R中UR与DR数目分别为432、410、92及36、52、2个。



A - 叶片组织差异组别; B - 根组织差异组别; 横坐标为 GO 功能类别, 纵坐标为该功能基因数量(黄色表示上调, 蓝色表示下调)。

A - leaf tissue difference group; B - root tissue differential group; the horizontal coordinate is the GO functional category, and the vertical coordinate is the number of the functional genes (yellow means up-regulated, blue means down-regulated).

图3 “岷归1号”和“岷归2号”差异表达基因GO功能分类

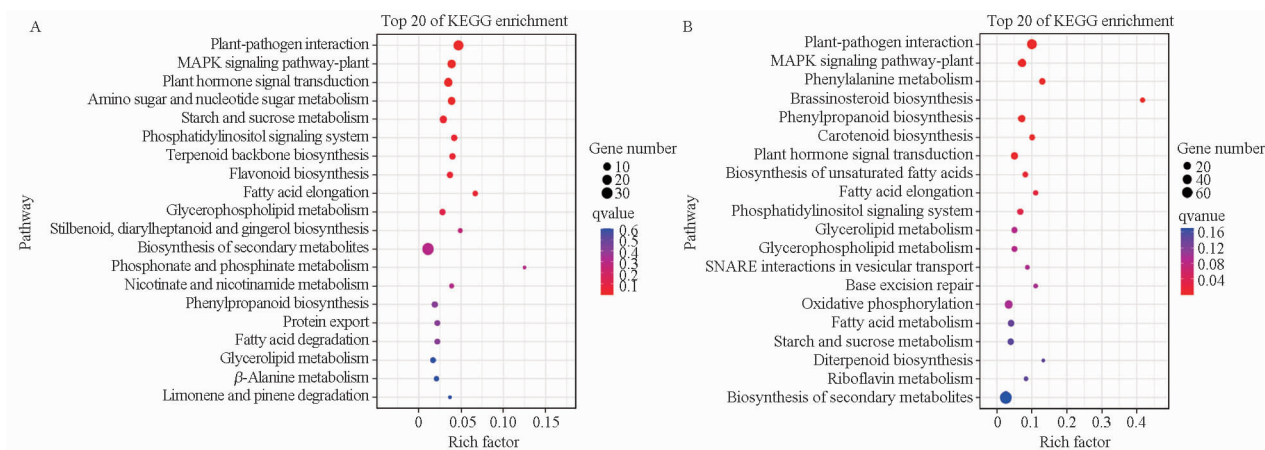
Fig. 3 GO functional classification of DEGs of ‘Mingui 1’ and ‘Mingui 2’

3.2.2 KEGG 代谢通路富集分析 根据 KEGG 分类与注释结果(图4),M1L vs M2L与M1R vs M2R富集通路共分为5类,涉及通路数量由多到少分别是代谢(metabolism)、遗传信息处理(genetic

information processing)、环境信息处理(environmental information processing)、细胞过程(cellular processes)和有机系统(organismal systems)。M1L vs M2L与M1R vs M2R参与代谢通路(pathway)分别为56和

63 条,根据 Q -value(FDR 校正后 P 值)筛选出前 20 个通路,其中主要包括植物-病原互作(plant-pathogen interaction)、植物-丝裂原活化蛋白激酶(MAPK)信号通路(MAPK signaling pathway-plant)、苯丙烷生物合成(phenylpropanoid biosynthesis)、植

物激素信号转导(plant hormone signal transduction)、氧化磷酸化(oxidative phosphorylation)、淀粉与蔗糖代谢(starch and sucrose metabolism)以及次生代谢物的生物合成(biosynthesis of secondary metabolites)等,其余通路中 DEGs 参与较少。



A - 叶片组织差异组别;B - 根组织差异组别;横坐标为富集因子(该 pathway 中 DEGs 数目除以该 pathway 中总数目,纵坐标为注释 pathway)。

A - leaf tissue difference group; B - root tissue differential group; Rich factor (number of DEGs in this pathway divided by total number of DEGs in this pathway, ordinate is annotation pathway).

图4 “岷归1号”和“岷归2号”差异表达基因 KEGG 富集通路

Fig. 4 KEGG enrichment pathway of DEGs of ‘Mingui 1’ and ‘Mingui 2’

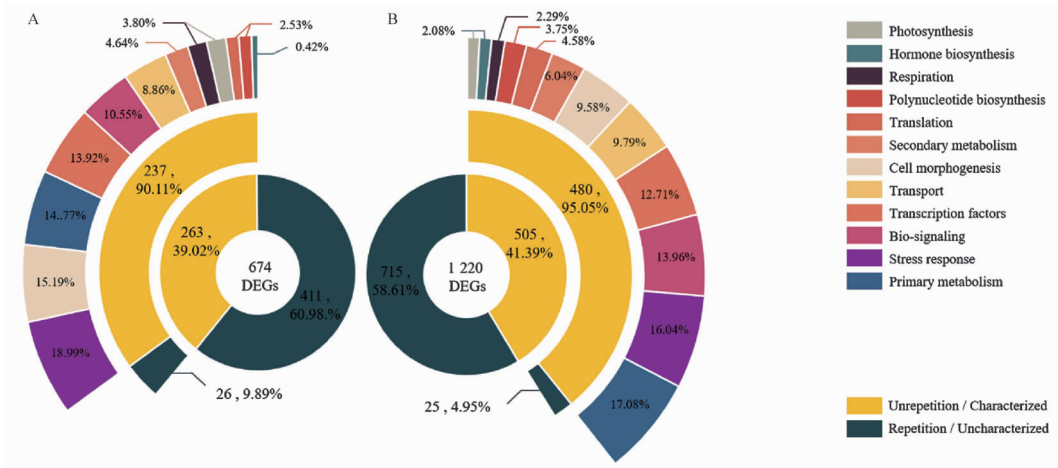
3.2.3 差异表达基因精细分类功能注释 基于 Swiss-Prot 数据库比对结果,通过 Uniprot 蛋白质数据库对筛选获得的 M1L vs M2L 与 M1R vs M2R 差异基因转录数据进行精细分类功能注释。两差异组别各包含 411 和 715 个重复基因,剩余已鉴定基因数目分别为 263 和 505 个。在已鉴定 DEGs 中, M1L vs M2L 中有 26 个注释无生物学特性, M1R vs M2R 则为 25 个,根据生物学功能和分子功能可将 237 和 480 个 DEGs 注释为 12 类:光合作用(photosynthesis, 9, 10)、呼吸作用(respiration, 9, 10)、初级代谢(primary metabolism, 35, 82)、次生代谢(secondary metabolism, 11, 29)、激素生物合成(hormone biosynthesis, 1, 10)、生物信号(bio-signaling, 25, 67)、多核苷酸生物合成(polynucleotide biosynthesis, 6, 18)、细胞形态发生(cell morphogenesis, 36, 46)、转录因子(transcription factors, 33, 61)、翻译(translation, 6, 22)功能、转运(transport, 21, 47)功能以及逆境响应(stress response, 45, 77),见图 5。M1L vs M2L 精细功能分类占比居前五位的是代谢、逆境响应、细胞形态发生、转录因子和生物信号,比例依次为 19.41%、18.99%、15.19%、13.92% 和 10.55%。M1R vs M2R 主要

分布于代谢、逆境响应、生物信号、转录因子和转运功能,比例依次为 23.12%、16.94%、13.96%、12.71% 和 9.79%,注释结果与 GO 及 KEGG 功能富集结果趋势相符合。

3.2.4 抗旱相关差异表达基因筛选 根据功能注释及富集代谢途径结果对两差异组别中相同 DEGs 进一步精细筛选,筛选得到有关脱落酸(ABA)调节、渗透调节、活性氧清除和其他功能蛋白质调节的抗旱相关基因共 60 个。其中属于 DREB 转录因子亚家族的关键基因主要包括 *DREB1B*、*DREB1D*,在当归叶片和根中均呈上调表达模式,表明该类基因在 M2 中上调,且 *DREB1B* 在根中表达量略高于叶片,*DREB1D* 则在叶片中高表达。脱水诱导早期反应 *ERD* 基因(*ERD7*、*ERD15*)参与胁迫相关激素诱导,在叶片中的表达量高于根中。涉及 ABA 信号转导途径的 *HVA22* 基因(*HVA22A*、*HVA22C*)中,*HVA22A* 在叶片和根中表达量相近, \log_2FC 值分别为 1.82 和 1.84,*HVA22C* 则在根中显著高表达, \log_2FC 值分别为 1.94 和 2.87。植物-MAPK 信号通路中的丝裂原活化蛋白激酶 3(*MPK3*)基因具有调节气孔细胞呼吸的功能,在叶片中显著高表达。U-box 基因家族 E3 泛素蛋白连接酶(*PUB9*、*PUB23*)可协同调

控干旱信号通路, *PUB9* 在叶片中表达量高于根中, *PUB23* 则在根中表达量高于叶中。参与渗透调节的关键基因主要包括阿拉伯半乳蛋白 (*AtAGP41*)、天冬氨酸蛋白酶 (*APF2*、*ASPI*、*At3g56050*、*At4g16563*)、钙信号蛋白 (*ERD4*、*RXW8*、*KRP1*)、脯氨酸受体样蛋白激酶 (*PERK1*)、海藻糖-6-磷酸磷酸酶 (*TPPJ*) 及海藻糖-6-磷酸合成酶 (*TPS7*), 其基因表达量在叶片和根中均呈不同程度上调。氧化还原关键酶包括 2 个抗

坏血酸氧化酶 (*AAO*、*Bp10*)、2 个脱氢酶还原酶 (*At3g08610*、*SDR2a*)、1 个氧化还原酶 (*CYP76A2*)、1 个天冬氨酸氧化酶 (*AO*), 除 *At3g08610* 基因表达下调之外, 其余基因均表达上调。从差异表达量角度分析, 筛选出的抗旱相关基因在叶片和根组织差异组中均呈现 M2 大于 M1 的情况, 且上调 DEGs 多于下调 DEGs, 只有少数筛选基因下调, 结果表明, M2 的抗旱性可能略强于 M1 (表 4)。



A - 叶片组织差异组别; B - 根组织差异组别。

A - leaf tissue difference group; B - root tissue differential group.

图 5 “岷归 1 号”和“岷归 2 号”差异表达基因精细分类图

Fig. 5 Detailed classification of DEGs of ‘Mingui 1’ and ‘Mingui 2’

表 4 当归抗旱相关 DEGs 分析

Tab. 4 Analysis of DEGs related to drought resistance of *A. sinensis*

Sort	Subclassification	Gene abbreviation	Comment name	log ₂ FC		Expression pattern
				Leaf	Root	
Abscisic acid regulation		<i>DREB1B</i>	Dehydration-responsive element-binding protein 1B	2.96	3.03	UR
		<i>DREB1D</i>	Dehydration-responsive element-binding protein 1D	2.82	2.30	UR
		<i>ERD7</i>	Protein early-responsive to dehydration 7	1.58	1.18	UR
		<i>ERD15</i>	Protein early-responsive to dehydration 15	1.54	1.41	UR
		<i>HVA22A</i>	HVA22-like protein a	1.82	1.84	UR
		<i>HVA22C</i>	HVA22-like protein c	1.94	2.87	UR
		<i>MPK3</i>	Mitogen-activated protein kinase 3	2.16	2.02	UR
		<i>FER</i>	Receptor-like protein kinase FERONIA	1.83	1.11	UR
		<i>PUB23</i>	E3 ubiquitin-protein ligase PUB23	1.99	2.49	UR
		<i>PUB9</i>	U-box domain-containing protein 9	1.85	1.20	UR
Osmotic regulation		<i>AGP41</i>	Arabinogalactan protein 41	1.46	1.48	UR
		<i>APF2</i>	Aspartyl protease family protein 2	1.10	2.05	UR
		<i>ASPI</i>	Aspartic proteinase Asp1	1.82	1.65	UR
		<i>At3g56050</i>	Probable inactive receptor-like protein kinase At3g56050	1.28	1.29	UR
		<i>At4g16563</i>	Probable aspartyl protease At4g16563	1.63	1.54	UR
		<i>ERD4</i>	CSC1-like protein ERD4	1.31	1.35	UR
		<i>NTL9</i>	Protein NTM1-like 9 (Calmodulin-binding NAC protein)	1.15	1.37	UR
		<i>PERK1</i>	Proline-rich receptor-like protein kinase PERK1	1.70	2.64	UR
		<i>KRP1</i>	Calcium-binding protein KRP1	3.68	1.78	UR
		<i>RALFL33</i>	Protein RALF-like 33	1.56	1.17	UR
	<i>RXW8</i>	CSC1-like protein RXW8	1.43	2.03	UR	

续表 4 (continued)

Sort	Subclassification	Gene abbreviation	Comment name	log ₂ FC		Expression pattern	
				Leaf	Root		
Reactive oxygen scavenging		<i>TPPJ</i>	Probable trehalose-phosphate phosphatase J	2.59	1.16	UR	
		<i>TPS7</i>	Probable a-trehalose-phosphate synthase	1.16	1.23	UR	
		<i>AAO</i>	L-ascorbate oxidase (ASO)	1.62	2.33	UR	
		<i>AO</i>	L-aspartate oxidase	1.04	1.21	UR	
		<i>At3g08610</i>	NADH dehydrogenase [ubiquinone]	-1.98	-2.36	DR	
		<i>Bp10</i>	L-ascorbate oxidase homolog	3.55	4.50	UR	
		<i>CYP76A2</i>	Cytochrome P450 76A2	1.29	2.51	UR	
		<i>SDR2a</i>	Short-chain dehydrogenase reductase 2a	2.15	2.08	UR	
	Regulation of other functional proteins	Aquaporin	<i>SYPI21</i>	Protein Syntaxin of plants 121	2.70	1.89	UR
		Response protein	<i>At2g40140</i>	Zinc finger CCCH domain-containing protein 29	1.78	1.01	UR
		<i>At4g27290</i>	G-type lectin S-receptor-like serine/threonine-protein kinase At4g27290	2.32	1.25	UR	
		<i>ATL6</i>	E3 ubiquitin-protein ligase ATL6	1.18	1.46	UR	
		<i>BAP2</i>	BON1-associated protein 2	2.30	1.54	UR	
		<i>DGK1</i>	Diacylglycerol kinase 1	1.01	1.49	UR	
		<i>DGK2</i>	Diacylglycerol kinase 2	2.64	2.19	UR	
		<i>FLS2</i>	LRR receptor-like serine/threonine-protein kinase FLS2	1.41	2.93	UR	
		<i>GATA8</i>	GATA transcription factor 8	2.80	1.17	UR	
		<i>NCL</i>	Sodium/calcium exchanger NCL	2.59	2.69	UR	
		<i>NHL10</i>	NDR1/HIN1-like protein 10	1.55	1.36	UR	
		<i>NHL13</i>	NDR1/HIN1-like protein 13	2.32	1.02	UR	
		<i>PBL3</i>	Probable serine/threonine-protein kinase PBL3	1.56	1.51	UR	
		<i>PCR2</i>	Protein plant cadmium resistance 2	1.56	1.58	UR	
		<i>PCRK1</i>	Serine/threonine-protein kinase PCRK1	1.24	1.01	UR	
		<i>RPV1</i>	Disease resistance protein RPV1	2.22	2.17	UR	
		<i>WRKY40</i>	Probable WRKY transcription factor 40	1.58	2.49	UR	
		<i>WRKY53</i>	Probable WRKY transcription factor 53	1.13	1.53	UR	
	Cell wall construction	<i>CDC20-1</i>	Cell division cycle 20.1, cofactor of APC complex	1.92	1.23	UR	
		<i>CESA2</i>	Cellulose synthase A catalytic subunit 2	3.21	2.21	UR	
		<i>CHIT3</i>	Hevamine-A	1.48	1.13	UR	
		<i>CSLC12</i>	Probable xyloglucan glycosyltransferase 12	1.14	1.27	UR	
		<i>DCR</i>	BAHD acyltransferase DCR	1.65	1.82	UR	
		<i>MUC121</i>	Xylan glycosyltransferase MUC121	1.66	1.51	UR	
		<i>TBR</i>	Protein trichome birefringence	1.39	1.07	UR	
		<i>THE1</i>	Receptor-like protein kinase THESEUS 1	1.67	1.08	UR	
		<i>XTH15</i>	Xyloglucan endotransglucosylase/hydrolase 15	2.59	1.73	UR	
		<i>XTH22</i>	Xyloglucan endotransglucosylase/hydrolase protein 22	2.85	1.27	UR	
		<i>XTH23</i>	Probable xyloglucan endotransglucosylase/hydrolase protein 23	3.07	2.35	UR	
		<i>XXT3</i>	Probable xyloglucan 6-xylosyltransferase 3	1.60	1.66	UR	
		<i>XXT5</i>	Probable xyloglucan 6-xylosyltransferase 5	1.63	1.26	UR	

注: UR - 上调; DR - 下调。

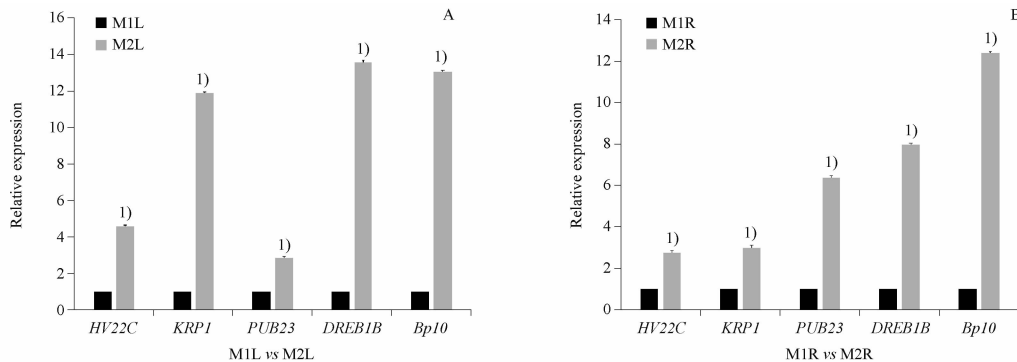
Note: UR - up-regulated; DR - means down-regulated.

3.3 差异表达基因的 qRT-PCR 验证

根据当归转录组测序结果对部分关键基因进行定量分析, 随机选取抗旱相关基因 *HVA22C*、*KRP1*、*PUB23*、*DREB1B*、*Bp10* 进行 qRT-PCR 验证 ($n=3$)。结果见图 6, 5 个关键基因在两品种当归叶片和根中表达变化情况与转录组测序基因表达趋势一致, 说明本研究测序数据可靠性较高, 参考性较强。

4 讨论

干旱是影响植物生长发育的主要限制因素之一, 植物响应干旱胁迫是一个复杂信号调控过程, 其转录网络受多个基因调节, 涉及多种代谢产物及合成途径。利用转录组测序技术可使基因表达量化, 挖掘关键功能基因, 是培育品种优良性状和研究基因功能的重要手段^[22]。本研究通过对两品种当归进行转录组测序, 结合 GO、KEGG 和精细功能注释



A - 叶片组织差异组别;与 M1L 组比, ¹⁾ $P < 0.001$; B - 根组织差异组别;与 M1R 组比, ¹⁾ $P < 0.001$ 。

A - leaf tissue difference group, ¹⁾ $P < 0.001$, vs M1L; B - root tissue differential group; ¹⁾ $P < 0.001$, vs M1R.

图 6 当归中关键基因的 qRT-PCR 表达量验证。 $n = 3, \bar{x} \pm s$

Fig. 6 qRT-PCR expression verification of key genes in *A. sinensis*. $n = 3, \bar{x} \pm s$

对筛选获得的 DEGs 分类,发现差异功能主要集中在生长代谢、抗逆响应、苯丙烷类、黄酮类等化合物生物合成及各类生物反应等过程,推测两品种当归可能在生长发育、抗胁迫能力和药效成分合成等方面均存在一定差异性。

当前研究干旱胁迫的调节途径主要包括 ABA 调节、渗透调节、活性氧清除和其他功能蛋白质调节等,涉及脱落酸、过氧化氢酶等多种物质的合成、代谢及信号转导通路^[23]。部分转录因子(transcription factor, TF)可通过结合特异元件来调节功能基因表达以介导非生物胁迫信号转导。DREB 转录因子属于 AP2/ERF 转录因子家族的一个亚家族,该亚家族基因可增强抗逆能力^[24]。ERD 基因最早发现于拟南芥脱水诱导,也受其他低温、高盐、ABA 等胁迫影响,ERD15 是 ABA 反应和依赖水杨酸(SA)防御途径的重要调节器^[25]。HVA22 家族基因在 ABA 和应激诱导下,编码一类具有保守 TB2/DP1/HVA22 结构域的应激反应蛋白,可结合干旱响应(MYB)防御和胁迫响应元件(TC-rich repeats)以及激素响应元件(ABRE、ERE、SARE 等)^[26]。PUB9、PUB23 属于植物 U-box E3 泛素蛋白连接酶,对干旱胁迫表现出超敏性,泛素化可能会改善植株抗旱能力^[27]。在当归转录数据中,以上关键基因在叶片和根中均为上调表达模式,表明其可能正向调控当归抗旱过程,响应 ABA 调节引起叶片气孔关闭,直接性降低蒸腾作用导致的水分流失,有效影响当归抵御干旱并提高其抗旱性。

渗透调节也是药用植物抵抗干旱的重要生理机制之一,本研究涉及的渗透调节物质包括脯氨酸(Pro)、可溶性碳水化合物、钙信号蛋白等。耐旱植物可通过积累蔗糖、海藻糖等来提高抗旱性,

海藻糖的主合成途径为海藻糖-6-磷酸合成酶-海藻糖-6-磷酸磷酸酶(trehalose-6-phosphate synthase-trehalose-6-phosphate phosphatase, TPS-TPP)途径,TPS 与 TPP 受环境诱导表达,促使代谢物生成,目前已在多个物种间被鉴定^[28]。ERD4、RXW8、KRPI 等钙信号蛋白基因在当归叶片和根中呈现上调表达模式,表明其跨膜转导作用可调节相对含水量以维持细胞渗透压,增加渗透势,以此反映抗旱性^[29]。药用植物在干旱胁迫下可产生有害活性氧,抗氧化系统可发挥抵御或保护水分能力,而本研究中抗氧化酶基因 AAO、Bp10、At3g08610、SDR2a、CYP76A2、AO 的上调表达也在一定程度说明其具有清除活性氧功能,以此在干旱胁迫中达到动态平衡抵御干旱的作用^[30]。

此外,还有其他功能蛋白质也可能参与干旱胁迫,在当归共有 DEGs 中,涉及分类主要包括水通道蛋白、响应蛋白及细胞壁建构等,基因数目分别为 1、17 和 13 个。水通道蛋白可通过增加细胞膜通透性来进行高效物质转运,从而提高对各种胁迫的耐受性,SYPI21 在当归叶片中表达量显著高于根中,说明该基因可能在叶片中响应胁迫程度大于根中;有学者对植株干旱胁迫发现,其叶片抗病基因与抗旱基因表达模式一致^[31],本研究中各类响应蛋白,如发育蛋白基因(At4g27290、WRKY53)、抗病蛋白基因(ATL6、BAP2、FLS2、NHL10、NHL13、PBL3、PCRK1、RPV1、WRKY40)、盐胁迫蛋白基因(At2g40140、NCL)、冷胁迫蛋白基因(DGK1、DGK2)及解毒蛋白基因(PCR2)等均与直接参与抗旱的基因表达一致,间接反映以上基因可能与抗旱性存在交叉协同调控作用;植物细胞壁建构合成与木质素、角质、蜡质合成相关,CESA2、CHIT3、CSLC12、DCR

等基因通过调控变化促使植物表层形成屏障,与耐旱性形成有一定联系,且在药用植物生长发育和抗性方面发挥显著作用。

本研究对相同生境条件下的两品种当归进行转录比较,经 qRT-PCR 初步筛选验证抗旱关键基因 *HVA22C*、*KRPI*、*PUB23*、*DREB1B*、*Bp10*,为研究当归响应干旱胁迫的分子机制提供候选基因。同时,此结果也为后续进一步探究当归基因功能验证提供基因资源。

REFERENCES

- [1] Ch. P (2020) Vol I(中国药典 2020 年版.一部)[S]. 2020:139.
- [2] YAN H, ZHANG X B, ZHU S D, *et al.* Production regionalization study of Chinese angelica based on MaxEnt model. n [J]. *China J Chin Mater Med* (中国中药杂志), 2016, 41 (17): 3139-3147.
- [3] LIU F Z, LI Y B, WANG J, *et al.* Systematic evaluation and analysis on Dao-di Herbs *Angelica sinensis* [J]. *Mod Tradit Chin Med Mater Med-World Sci Technol*(世界科学技术-中医药现代化), 2018, 20(9):1531-1539.
- [4] JIA J N. Breeding report of new strain 90-01 of *Angelica sinensis* [J]. *J Chin Med Mater* (中药材), 2005, 28(8):649-650.
- [5] XIE H M, LIU X R, JIN Y M, *et al.* Breeding of Danggui 90-02 in Gansu province [J]. *Mod Chin Med* (中国现代中药), 2007, 9(3):37-38.
- [6] ZHU T T, ZHANG M H, SU H Y, *et al.* Integrated metabolomic and transcriptomic analysis reveals differential mechanism of flavonoid biosynthesis in two cultivars of *Angelica sinensis* [J]. *Molecules*, 2022, 27(1):306-323.
- [7] PENG T, WANG Y, YANG T, *et al.* Physiological and biochemical responses and comparative transcriptome profiling of two *Angelica sinensis* cultivars under enhanced ultraviolet-B radiation [J]. *Front Plant Sci*, 2021, 12:805407. DOI:10.3389/fpls.2021.805407.
- [8] ANDEREGG W R L, ANDEREGG L D L. Hydraulic and carbohydrate changes in experimental drought-induced mortality of saplings in two conifer species [J]. *Tree Physiol*, 2013, 33(3):252-260.
- [9] LI J, LI M, ZHU T T, *et al.* Integrated transcriptomics and metabolites at different growth stages reveals the regulation mechanism of bolting and flowering of *Angelica sinensis* [J]. *Plant Biol*, 2021, 23(4):574-582.
- [10] YANG T, YAO Y Y, ZHANG Y T, *et al.* Effects of different plant hormones and microorganisms on yield and quality of *Angelica sinensis* in repeated cropping fields [J]. *J Chin Med Mater* (中药材), 2021, 44(5):1065-1070.
- [11] ADITI G, ANDRÉS R, I A C. The physiology of plant responses to drought [J]. *Science*, 2020, 368(6488):266-269.
- [12] LI Y X, CHEN X X, DAI X M, *et al.* Study on programmed cell death of plants in environmental stress [J]. *Biotechnol Bull*(生物技术通报), 2009, (4):7-11.
- [13] LIU M L, ZHANG A, FENG W, *et al.* Comparison and analysis of three kinds of components in *Angelicae sinensis* radix pieces and Danggui Jianzhong decoction by UHPLC-Q-TOF-MS [J]. *Chin Pharm J*(中国药学杂志), 2022, 57(8):613-622
- [14] CHEN S, ZHOU Y, CHEN Y, *et al.* Fastp: an ultra-fast all-in-one FASTQ preprocessor [J]. *Bioinformatics*, 2018, 34(17):i884-i890.
- [15] GRABHEEE M G, HASS 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.
- [16] 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):323-338.
- [17] LOVE M I, HUBER W, ANDERS S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2 [J]. *Gen Biol*, 2014, 15(12):550-570.
- [18] KANEHISA M, ARAKI M, GOTO S, *et al.* KEGG for linking genomes to life and the environment [J]. *Nucleic Acids Res*, 2008, 36(D1):D480-D484.
- [19] GENE ONTOLOGY CONSORTIUM. The gene ontology resource: enriching a Gold mine [J]. *Nucleic Acids Res*, 2021, 49(D1):D325-D334.
- [20] XU R, XU J, LI Y C, *et al.* Integrated chemical and transcriptomic analyses unveils synthetic characteristics of different medicinal root parts of *Angelica sinensis* [J]. *Chin Herb Med*, 2020, 12(1):19-28.
- [21] WILLEMS E, LEYNS L, VANDESOMPELE J. Standardization of real-time PCR gene expression data from independent biological replicates [J]. *Anal Biochem*, 2008, 379(1):127-129.
- [22] J R, LIU C C, SUN C H, Plant stress RNA-seq Nexus: a stress-specific transcriptome database in plant cells [J]. *BMC Genom*, 2018, 19(1):966-982.
- [23] ULLAH A, MANGHWAR H, SHABAN M, *et al.* Phytohormones enhanced drought tolerance in plants;a coping strategy [J]. *Environ Sci Pollut Res Int*, 2018, 25(33):33103-33118.
- [24] ZHANG L F, HOU Q D, CAI X W, *et al.* Cloning of *DREB1D* gene from pitaya and its functional analysis in transgenic *Arabidopsis thaliana* [J]. *J Fruit Sci*(果树学报), 2023, 40(7):1330-1341.
- [25] CAO S S, WU N, WANG L P, *et al.* Cloning, bioinformatics analysis and function identification of two soybean *ERD15* genes [J]. *Chin J Oil Crop Sci*(中国油料作物学报), 2022, 44(4):790-797.
- [26] ZHANG H J, YUAN Y C, XING H X, *et al.* Genome-wide identification and expression analysis of the *HVA22* gene family in cotton and functional analysis of *GhHVA22E1D* in drought and salt tolerance [J]. *Front Plant Sci*, 2023, 14:1139526. DOI:10.1389/fpls.2023.1139526.
- [27] ZHAO J, ZHAO L, ZHANG M, *et al.* Arabidopsis E3 ubiquitin ligases *PUB22* and *PUB23* negatively regulate drought tolerance by targeting ABA receptor *PYL9* for degradation [J]. *Int J Mol Sci*, 2017, 18(9):1841-1858.
- [28] DU J L, LI X L, MA Y W, *et al.* Research progress in plant trehalose-6-phosphate synthase genes [J]. *Plant Sci J*(植物科学学报), 2023, 41(3):411-420.
- [29] KATHRIN T. Calcium-nutrient and messenger. [J]. *Front Plant Sci*, 2019, 10:440.
- [30] HE A, DEAN J, LODHI I J, *et al.* Peroxisomes as cellular adaptors to metabolic and environmental stress. [J]. *Trends Cell Biol*, 2021, 31(8):656-670.
- [31] ZHANG H, HU H Y, LI H X, *et al.* Physiological response and transcriptome analysis of the desert steppedominant plant *Lespedeza potaninii* to drought stress [J]. *Acta Pratacult Sin*(草业学报), 2023, 32(7):188-205.

(收稿日期:2023-10-09)