

## • 药材与资源 •

## 基于转录组测序揭示适度干旱胁迫对丹参基因表达的调控

李晓艳, 周敬雯, 严铸云, 陈 新\*

成都中医药大学药学院, 四川 成都 611137

**摘要:**目的 对丹参进行转录组测序, 分析丹参不同组织响应适度干旱胁迫的分子机制。方法 以栽培 4 个月的丹参为材料, 设置适度干旱胁迫组(土壤相对含水量 55%~60%)和对照组(土壤相对含水量 75%~80%), 应用 Illumina HiSeq 2000 分别对根和叶进行转录组测序, 对测序后的结果进行基因功能注释、差异表达基因(differentially expressed genes, DEGs)筛选和共表达网络分析等。结果 转录组测序共获得 58 085 条 Unigene。其中 28 846 条被注释, 根和叶中的 DEGs 分别有 1 853、1 457 个。GO 富集结果表明, 丹参根和叶中的 DEGs 在 GO 功能部位中的分布基本一致, 均在代谢过程、刺激响应、细胞结构、催化活性等功能中得到显著富集。KEGG 途径富集分析表明, 根中 DEGs 显著富集在 DNA 复制、植物激素信号转导、植物-病原菌相互作用、胡萝卜素合成等途径, 叶中 DEGs 则主要富集在氨基酸、生物碱、苯丙烷类生物合成等途径。适度干旱胁迫能促进丹参叶中苯丙烷类、根中萜类化合物生物合成途径关键酶基因的表达, 是促进丹参有效成分累积的基础。AP2/ERF、bHLH、bZIP、WRKY、MYB 类转录因子在根和叶中差异表达均较显著, 基因共表达网络分析预测了在适度干旱胁迫下可能参与调控萜类基因表达的转录因子。结论 通过高通量转录组测序, 揭示了适度干旱胁迫对丹参不同组织基因表达的调控特征, 可为深入研究丹参药效成分的生物合成机制和在栽培中的合理灌溉提供科学依据。

**关键词:** 丹参; 干旱胁迫; 转录组测序; 差异表达基因; 转录因子

中图分类号: R286.12 文献标志码: A 文章编号: 0253-2670(2020)06-1600-09

DOI: 10.7501/j.issn.0253-2670.2020.06.029

## Sequencing and analysis of transcriptome to reveal regulation of gene expression in *Salvia miltiorrhiza* under moderate drought stress

LI Xiao-yan, ZHOU Jing-wen, YAN Zhu-yun, CHEN Xin

School of Pharmacy, Chengdu University of Traditional Chinese Medicine, Chengdu 611137, China

**Abstract: Objective** To analyze the molecular mechanism of different tissues of *Salvia miltiorrhiza* in response to moderate drought stress by transcriptome sequencing. **Methods** The roots and leaves of 4-month-old *S. miltiorrhiza* from moderate drought stressed group and control group (the relative water content in soil was 55%—60% and 75%—80%, respectively) were used as the test material. And the transcriptome sequencing analysis was carried out by using Illumina HiSeq 2000. After obtaining transcriptome data, gene function annotation, differentially expressed genes (DEGs) screening and and co-expression network analysis were performed. **Results** A total of 58 085 Unigenes were obtained by transcriptome sequencing. Among them, 28 846 Unigenes were annotated, and there were 1 853 and 1 457 DEGs in roots and leaves, respectively. The GO enrichment results showed that the DEGs of the roots and leaves were both significantly enriched in metabolic process, stimulus response, cell structure, and catalytic activity, etc. The KEGG pathway analysis showed that DEGs in roots were significantly enriched in DNA replication, plant hormone signal transduction, plant-pathogen interaction, and carotenoid biosynthesis, etc. And the DEGs in leaves were mainly concentrated in amino acids, alkaloids, and phenylpropanoid biosynthesis. The genes of key enzymes involved in phenylpropanoid and terpenoid biosynthesis were up-regulated by moderate drought stress, which might be the basis for accelerating the accumulation of active ingredients in leaves and roots of *S. miltiorrhiza*. AP2/ERF, bHLH, bZIP, WRKY, and MYB transcription factors were significantly differentially expressed in roots and leaves. Gene co-expression network analysis predicted transcription factors that may be involved in regulating the expression of terpenoid genes under moderate drought stress. **Conclusion** The high-throughput transcriptome sequencing revealed the regulatory

收稿日期: 2019-10-06

基金项目: 国家自然科学基金资助项目(81573537)

作者简介: 李晓艳(1996—), 女, 硕士研究生, 研究方向为药用植物胁迫生理生态学。E-mail: 1143563094@qq.com

\*通信作者 陈 新, 男, 博士, 教授, 主要从事药用植物分子生态学研究。E-mail: chenxin@cdutcm.edu.cn

characteristics of moderate drought stress on gene expression in different tissues of *S. miltiorrhiza*, which could provide scientific basis for further research on the biosynthesis mechanism of medicinal components of *S. miltiorrhiza* and reasonable irrigation in cultivation.

**Key words:** *Salvia miltiorrhiza* Bge.; drought stress; transcriptome sequencing; differentially expressed genes; transcription factors

丹参 *Salvia miltiorrhiza* Bunge 为唇形科多年生草本植物, 以干燥根及根茎入药, 具有祛瘀止痛、养心除烦、活血通经等传统功效, 现多用于治疗冠心病、心绞痛、心悸、胸闷等症<sup>[1]</sup>。丹参是我国传统大宗药材, 药用历史悠久且市场需求量大, 目前以人工栽培为主。其主要活性成分为脂溶性的丹参酮类化合物和水溶性的丹酚酸类化合物, 分别由萜类代谢途径和苯丙烷类代谢途径合成<sup>[2]</sup>。丹参中萜类生物合成途径主要包括由 2-C-甲基-D-赤藓糖醇-4-磷酸 (2-C-methyl-D-erythritol 4-phosphate, MEP)、甲羟戊酸 (mevalonate, MVA)、异戊烯基二磷酸酯 (isopentenyl diphosphate, IPP)、牻牛儿基牻牛儿基焦磷酸 (geranylgeranyl diphosphate, GGPP) 组成的萜类骨架途径, 以及下游的单萜、二萜、三萜、倍半萜、胡萝卜素、玉米素、类固醇、N-聚糖、泛醌等萜类物质生成途径<sup>[3]</sup>。苯丙烷类生物合成途径则主要包括 2 条支路途径: 苯丙氨酸合成途径和酪氨酸衍生途径<sup>[4]</sup>。

植物的次生代谢产物具有重要的生态学意义, 其产生和积累与环境变化密不可分<sup>[5]</sup>, 当植物遭受干旱、高温、盐或紫外辐射等非生物胁迫时, 体内部分代谢通路可能被激活, 从而影响代谢产物的累积<sup>[6]</sup>。研究发现, 适度的干旱胁迫 (土壤相对含水量为 55%~60%) 不仅能促进丹参根部丹参酮成分的合成, 还有利于丹参叶片中总酚酸含量的累积, 严重干旱和水分过多均不利于丹参的生长<sup>[7-8]</sup>。因此在丹参的栽培过程中, 通过合理灌溉来控制药材的产量和质量显得尤为重要。

近年来随着高通量转录组测序技术的快速发展, 通过比较 2 个或多个样本间的差异表达基因, 来探讨植物响应干旱等环境胁迫的分子机制已成为目前研究的热点<sup>[9]</sup>。干旱胁迫对丹参形态结构、生理生化反应、次生代谢产物含量积累的影响已有深入研究, 但对次生代谢调控机制的研究还较少<sup>[10-12]</sup>。本研究以丹参为材料, 分别设置对照组和适度干旱胁迫组, 利用转录组测序和差异表达基因 (differentially expressed genes, DEGs) 分析等方法, 从转录组水平上揭示适度干旱胁迫对丹参不同组织基因表达的调控特征, 为实际生产中丹参的合理灌溉以及分子育种提供理论依据。

## 1 材料与仪器

### 1.1 材料

以根段繁殖方式培育丹参幼苗, 所用丹参根段采自四川省中江县, 经成都中医药大学陈新教授鉴定为丹参 *Salvia miltiorrhiza* Bunge。

### 1.2 仪器

Plant Total RNA Isolation Kit Plus 试剂盒, RE-05021, 四川成都福际生物技术有限公司; NovoScript® 1st Strand cDNA Synthesis Kit 试剂盒, E041-01A, 上海近岸蛋白质科技有限公司; NovoScript® SYBR One-Step qRT-PCR Kit 试剂盒, E092-01A, 上海近岸蛋白质科技有限公司。NanoPhotometer™ P-Class 微量分光光度计, P330, 德国 Implen 公司; CFX96 实时荧光定量 PCR (qRT-PCR) 仪, Bio-Rad 公司。

## 2 方法

### 2.1 样品处理

盆栽土为营养土和珍珠岩的混合土, 体积比 2:1, 按常规栽培方法插入根段后将其置于温室大棚内。待植株生长 4 个月后, 选取生长状况一致的随机分为 2 组。对照组的土壤相对含水量控制为 75%~80%, 于每日 18:00 称质量补水, 适度干旱胁迫组停止浇水使其自然干旱, 并每日采用称重法测定土壤相对含水量, 处理 7 d 后达到 55%~60%, 分别采收对照组和适度干旱胁迫组的丹参根和叶, 洗净后用液氮速冻, 置-80 °C 保存备用。

### 2.2 总 RNA 的提取与转录组测序

分别取丹参根和叶在液氮中研成粉末, 用 Plant Total RNA Isolation Kit Plus 试剂盒提取总 RNA, 并用 NanoPhotometer™ P-Class 微量分光光度计检测总 RNA 的纯度与浓度。每处理 3 份生物学重复。将每处理 3 份 RNA 样品等量混合, 委托北京百迈客生物科技有限公司构建测序文库, 并使用 Illumina HiSeq 2000 测序平台进行转录组测序。

### 2.3 测序数据组装及 Unigene 功能注释

对原始测序数据进行过滤, 去除测序接头、引物以及低质量的序列后, 获得高质量的 clean reads。使用 Trinity 软件将 clean reads 组装成 contig, 利用 De Bruijn 进一步简化后获得 Unigene。然后使用 BLAST 软件将 Unigene 序列与

NR、Swiss-Prot、GO、COG、KOG、eggNOG4.5、KEGG 数据库比对，并在预测完 Unigene 的氨基酸序列之后使用 HMMER 软件与 Pfam 数据库比对，获得 Unigene 的注释信息。

#### 2.4 差异表达基因分析

使用 FPKM (fragments per kilobase of transcript per million mapped reads) 计算 Unigene 的表达量，利用 EBseq 平台分别筛选丹参根和叶中的差异表达基因，筛选条件为错误发现率 (false discovery rate, FDR)  $\leq 0.05$  且差异倍数 (fold change, FC)  $\geq 1.5$ 。并对筛选出的 DEGs 进行

GO、KEGG 富集分析。

#### 2.5 qRT-PCR 验证

分别将不同处理下的丹参根和叶的总 RNA 用 NovoScript® 1st Strand cDNA Synthesis Kit 试剂盒逆转录成 cDNA，将模板稀释 20 倍用于后续 qRT-PCR 反应。选取 KEGG 途径中与次生代谢产物生物合成相关的 5 个 Unigene 和内参 Actin 进行 qRT-PCR 验证，利用预测的 CDS 序列和 Primer6 软件设计特异性引物 (表 1)，反应程序参照 NovoScript® SYBR One-Step qRT-PCR Kit 试剂盒。基因相对表达量用  $2^{-\Delta\Delta Ct}$  方法计算。

表 1 qRT-PCR 扩增特异引物

Table 1 Specific primers of qRT-PCR amplification

| 基因名称 (基因编号)            | 正向引物 (5'→3')              | 反向引物 (5'→3')              |
|------------------------|---------------------------|---------------------------|
| Actin                  | AGGAACCACCGATCCAGACA      | GGTGCCCTGAGGTCCTGTT       |
| HDR2 (c33737.graph_c0) | ATCGTATTGGGTTGACAGCGAGAAA | TATGCCGTTTGCAATACTTCATAGC |
| DXS4 (c29762.graph_c0) | CGTAGACAGGCAGAAGATT       | AACTGGTCGGTCCCTCAAT       |
| HMGR (c35323.graph_c3) | TTCTTGACACCAACCAGTT       | CAGTGACAATGCCATCCA        |
| 4CL9 (c20556.graph_c0) | GGAGCACTGTTGTCATAT        | GGTATTGTAGCCACGAGTT       |
| TAT (c28345.graph_c0)  | GGCAGACCACAATACAGT        | TAGACCTCATCGGCTATCA       |

#### 2.6 转录因子分析

分别从丹参根和叶的 DEGs 中筛选出萜类基因和转录因子，通过在 R 软件中运行 t.GCN.r<sup>[13]</sup>脚本构建萜类基因和转录因子间的共表达网络。具体步骤：首先，将萜类基因 FPKM 值作为指导基因表达谱文件，转录因子 FPKM 值作为查询基因表达谱文件，运用 cor 函数计算指导基因和查询基因间的 Pearson 相关系数，得到基因表达谱之间的相似性矩阵。然后对其进行显著性检验，选择  $P < 0.05$  的基因对作为显著共表达的基因，并根据基因间相似性系数构建基因共表达网络。最后，在 Cytoscape 软件中进行基因共表达网络的可视化。

### 3 结果与分析

#### 3.1 测序数据与基因注释

过滤原始数据后，各样品的 clean data 均达到 5.96 Gb，Q30 碱基百分比在 95%以上，且 GC 含量均在 50%左右，见表 2。将 clean reads 进行组装共得到 58 085 条 Unigene，Unigene 的 N<sub>50</sub> 为 1 778，其中长度在 1 kb 以上的 Unigene 有 14 367 条，占 Unigene 总数的 67.08%，表明组装的完整性较高。通过选择 BLAST 参数  $E$  值  $\leq 1 \times 10^{-5}$  和 HMMER 参数  $E$  值  $\leq 1 \times 10^{-10}$ ，最终获得

28 846 个有注释信息的 Unigene，约占 Unigene 总数的 49.66%。Unigene 与 Nr 数据库的 BLAST 比对结果见图 1，其中丹参转录组测序得到的 Unigene 与芝麻 *Sesamum indicum* L. 基因的同源性最高，达到 30.91%，而与丹参直接比对上的 Unigene 只有 757 条，占总数的 2.73%，表明目前能获得的丹参基因的注释信息还较少。

#### 3.2 差异表达基因的 GO 富集分析

与对照组相比，适度干旱胁迫下丹参根中的 DEGs 共有 1 853 个，其中上调和下调表达的基因分别为 1 222、631 个。丹参叶中的 DEGs 共有 1 457 个，其中上调和下调表达的基因分别为 811、664 个。图 2-A、B 分别显示了适度干旱胁迫下丹参根、叶中 DEGs 在 GO 二级节点的注释结果，从中可以看出，丹参根、叶中的 DEGs 在生物过程、细胞组分、分子功能的分布基本一致，都主要被注释到代谢过程、细胞过程、刺激响应、细胞部位、细胞膜、催化活性、结合等功能中。但丹参叶中有 Unigene 在细胞核和细胞外部发生了差异表达，其中注释到细胞核中的 DEGs 是类果糖激酶 1 基因 (fructokinase-like 1, FRK1)、未鉴定的蛋白

表 2 各样品的 clean data 统计

Table 2 Clean data statistics of each sample

| 组别     | 组织 | clean reads 数 | clean bases 数 | GC/%  | ≥Q30/% |
|--------|----|---------------|---------------|-------|--------|
| 对照     | 叶  | 20 065 215    | 5 964 689 082 | 48.44 | 95.33  |
|        | 根  | 21 218 836    | 6 317 748 212 | 48.84 | 95.48  |
| 适度干旱胁迫 | 叶  | 26 181 629    | 7 805 247 040 | 48.40 | 95.26  |
|        | 根  | 28 354 775    | 8 444 066 754 | 49.71 | 95.23  |

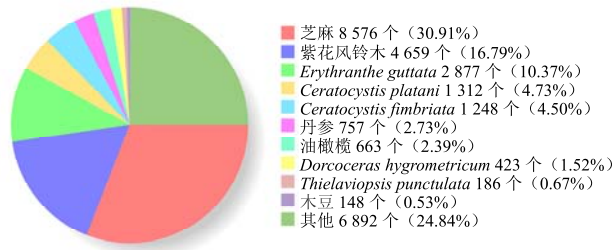


图 1 丹参 Unigene 与 Nr 数据库 BLAST 比对的物种分布

Fig. 1 BLAST comparison between *S. miltiorrhiza* Unigene and species in Nr database

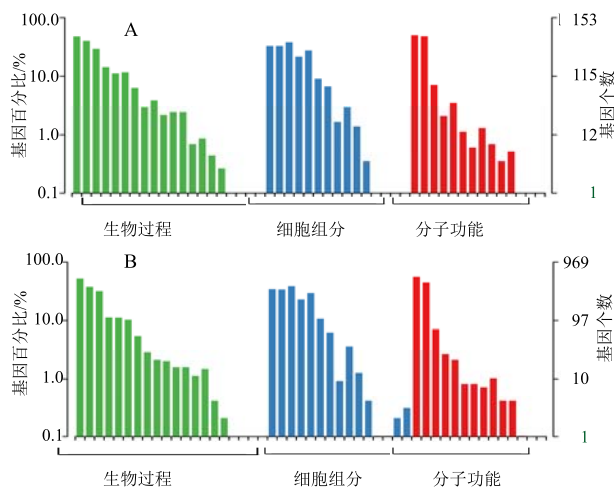


图 2 适度干旱胁迫下丹参根 (A) 和叶 (B) 中 DEGs 的 GO 功能分类

Fig. 2 GO function classification of DEGs in roots (A) and leaves (B) of *S. miltiorrhiza* under moderate drought stress

LOC105176065 亚型 X1 基因, 注释到细胞外部的则是丝氨酸羧肽酶基因 (serine carboxypeptidases, SCP)、金属蛋白酶基因 (PREDICTED: metalloproteinase 1-like)、组织蛋白酶 L 基因 (cysteine proteinase cathepsin L, CTSL)、蛋白质转录活性 10 基因。

### 3.3 差异表达基因的 KEGG 富集分析

分别将丹参根和叶中的 DEGs 进行 KEGG 途径富集分析,  $P \leq 0.05$  的通路被定义为 DEGs 显著富集的通路, 见表 3。适度干旱胁迫下, 丹参根中显著富集到的 KEGG 途径分为 3 类, 第 1

表 3 差异表达基因显著富集的 KEGG 途径

Table 3 Significantly enriched KEGG pathways of DEGs

| 组织    | KEGG 途径         | P 值                   |                       |
|-------|-----------------|-----------------------|-----------------------|
| 根     | DNA 复制          | $2.42 \times 10^{-7}$ |                       |
|       | 植物激素信号转导        | $6.84 \times 10^{-6}$ |                       |
|       | 植物-病原互作         | 0.000 209             |                       |
|       | 胡萝卜素合成          | 0.000 330             |                       |
|       | $\alpha$ -亚麻酸代谢 | 0.002 420             |                       |
|       | 淀粉和蔗糖代谢         | 0.006 623             |                       |
|       | 叶               | 苯丙氨酸、酪氨酸和色氨酸的生物合成     | $4.25 \times 10^{-5}$ |
|       |                 | 异喹啉生物碱合成              | $5.31 \times 10^{-5}$ |
|       |                 | 三聚氰胺、嘧啶和吡啶生物碱的生物合成    | 0.000 580             |
|       |                 | 酪氨酸代谢                 | 0.000 734             |
| 甘油酯代谢 |                 | 0.003 302             |                       |
|       | 精氨酸与脯氨酸代谢       | 0.005 009             |                       |
|       | 苯丙烷类生物合成        | 0.006 025             |                       |
|       | 淀粉和蔗糖代谢         | 0.006 975             |                       |

类为 DNA 复制途径, 第 2 类为信号转导途径, 包括植物激素信号转导、植物与病原菌的相互作用, 第 3 类则是胡萝卜素合成、不饱和脂肪酸代谢以及淀粉和糖代谢途径。丹参叶中显著富集到的途径主要涉及到一些代谢过程, 包括氨基酸、生物碱、甘油酯类、淀粉和糖代谢以及苯丙烷类生物合成。并且由 KEGG 注释结果可知, 适度干旱胁迫可促进丹参叶中苯丙烷类生物合成途径一些关键酶基因的表达, 如苯丙酸支路途径的 4-香豆素 CoA 连接酶基因 (4-coumarate : coenzyme A ligase 9, 4CL9), 酪氨酸支路的酪氨酸氨基转移酶基因 (tyrosine aminotransferase, TAT) 等, 这些基因在根中的表达则未受到干旱胁迫的影响。

### 3.4 适度干旱胁迫对丹参萜类代谢途径基因表达的影响

丹参转录组数据中注释到萜类途径的 Unigene 共有 328 条, 适度干旱胁迫下, 丹参根中上调和下调表达的萜类途径基因分别有 17、7 个 (表 4), 叶中上调和下调的萜类基因分别有 14、8 个 (表 5), 表明适度干旱胁迫主要能促进丹参萜类代谢途径基因的表达。其中胡萝卜素途



表 4 丹参根中受适度干旱胁迫调控的萜类代谢途径基因

Table 4 Genes of terpenoid metabolic pathway regulated by moderate drought stress in roots of *S. miltiorrhiza*

| 基因编号            | 表达变化 | 基因注释   | KEGG 途径      |
|-----------------|------|--|--------------|
| c29089.graph_c0 | 下调   | 甲羟戊酸激酶 (mevalonate kinase, MK)   | MVA          |
| c35323.graph_c3 | 上调   | 3-羟基-3-甲基戊二酰辅酶 A 还原酶 (3-hydroxy-3-methylglutaryl-coenzyme A reductase, HMGR)   | MVA          |
| c20558.graph_c1 | 下调   | 3-羟基-3-甲基戊二酰辅酶 A 还原酶 1 (3-hydroxy-3-methylglutaryl-coenzyme A reductase 1, HMGR1)  | MVA          |
| c20558.graph_c0 | 下调   | HMGR1, partial   | MVA          |
| c29762.graph_c0 | 上调   | 1-脱氧-D-木糖-5-磷酸合成酶 4 (putative 1-deoxy-D-xylulose 5-phosphate synthase precursor, DXS4)   | MEP          |
| c34914.graph_c0 | 上调   | 铁/抗坏血酸氧化还原酶 (Iron/ascorbate family oxidoreductase)   | 二萜           |
| c31423.graph_c0 | 下调   | 赤霉素生物合成相关蛋白 KAO2   | 二萜           |
| c17845.graph_c0 | 上调   | 赤霉素生物合成相关蛋白 GA3ox2   | 二萜           |
| c29734.graph_c0 | 下调   | 类胡萝卜素裂解双加氧酶 7 (carotenoid cleavage dioxygenase 7)  | 类胡萝卜素        |
| c31738.graph_c0 | 上调   | 预测: 类胡萝卜素裂解双加氧酶 8 同系物 (carotenoid cleavage dioxygenase 8 homolog B)  | 类胡萝卜素        |
| c23592.graph_c0 | 上调   | 预测: $\beta$ -胡萝卜素羟化酶 2 (PREDICTED: beta-carotene hydroxylase 2)  | 类胡萝卜素        |
| c26541.graph_c0 | 上调   | 细胞色素 P450 CYP707A99  | 类胡萝卜素        |
| c21507.graph_c0 | 上调   | $\beta$ -胡萝卜素羟化酶 (beta-carotene hydroxylase)   | 类胡萝卜素        |
| c35680.graph_c0 | 上调   | 八氢番茄红素去饱和酶 (Phytoene desaturase)   | 类胡萝卜素        |
| c9406.graph_c0  | 上调   | 9-顺式环氧双加氧酶 (9-cis epoxy-carotenoid dioxygenase)  | 类胡萝卜素        |
| c19023.graph_c0 | 上调   | 预测: 玉米黄质环氧化酶 (PREDICTED: zeaxanthin epoxidase)   | 类胡萝卜素        |
| c31278.graph_c0 | 上调   | 细胞色素 P450 CYP707A102   | 类胡萝卜素        |
| c29344.graph_c0 | 上调   | 预测: 细胞分裂素脱氢酶 9 (PREDICTED: cytokinin dehydrogenase 9-like)   | 玉米素          |
| c26420.graph_c0 | 下调   | 玉米素 O-葡萄糖基转移酶 (zeatin O-glucosyltransferase)   | 玉米素          |
| c33390.graph_c1 | 上调   | 鲨烯单加氧酶 (squalene monooxygenase, SQLE)  | 类固醇, 三萜, 倍半萜 |
| c31295.graph_c0 | 下调   | 异戊基半胱氨酸 $\alpha$ -羧基甲基酯酶 (isoprenylcysteine alpha-carbonyl methyl esterase)  | 法呢醛          |
| c19045.graph_c0 | 上调   | $\beta$ -1,4-甘露糖基糖蛋白 4 $\beta$ -N-乙酰氨基葡萄糖基转移酶 (beta-1,4-mannosyl-glycoprotein 4beta-N-N-聚糖 acetylglucosaminyltransferase-like) |              |
| c27665.graph_c0 | 上调   | 未鉴定的蛋白 LOC105168318  | N-聚糖         |
| c24793.graph_c0 | 上调   | $\beta$ -1,4-甘露糖基糖蛋白 4 $\beta$ -N-乙酰氨基葡萄糖基转移酶 (beta-1,4-mannosyl-glycoprotein 4-beta-N-acetylglucosaminyltransferase)          | N-聚糖         |

径的 DEGs 较多, 并且大部分基因的表达呈上调趋势, 表明适度干旱胁迫更有利于丹参中胡萝卜素的合成。根中有 3 条差异表达的 Unigene 均被注释为丹参 MVA 途径的关键酶基因 3-羟基-3-甲基戊二酰辅酶 A 还原酶 (3-hydroxy-3-methylglutaryl-coenzyme A reductase, HMGR), 其中 c35323.graph\_c3 (HMGR) 的表达上调, c20558.graph\_c0、c20558.graph\_c1 (HMGR1) 的表达下调, 与叶中 HMGR1 的表达正好相反。MEP 途径在根中受干旱胁迫正向调控的是 1-脱氧-D-木糖-5-磷酸合成酶 4 (putative 1-deoxy-D-xylulose 5-phosphate synthase precursor, DXS4) 基因, 而叶中则是 4-羟基-3-甲基-2-烯基二磷酸还原酶 2 (4-hydroxy-3-methylbut-2-enyl diphosphate reductase 2, HDR2)。适度干旱胁迫抑制丹参叶中二萜途径的关键酶基因二磷酸铜酰合酶

(copalyl diphosphate synthase, CPS)、贝壳杉烯合酶 (kaurene synthase, KS) 等的表达, 但能诱导根中其他二萜途径基因的表达。类固醇、三萜、倍半萜途径的关键酶基因鲨烯单加氧酶 (squalene monooxygenase, SQLE) 在根和叶中均受适度干旱胁迫的正向调控。IPP、GGPP、泛醌途径基因只在叶中发生差异表达, 单萜途径基因在根和叶中均未受干旱胁迫影响。

### 3.5 qRT-PCR 验证基因表达

为验证转录组数据中差异表达基因的可靠性, 选取了丹参萜类生物合成相关基因 HDR2、DXS4、HMGR 以及苯丙烷类生物合成关键酶基因 4CL9、TAT 进行 qRT-PCR 分析。结果表明 (图 3), 适度干旱胁迫下的 DXS4、HMGR 基因在根中的表达上调, 4CL9、TAT 基因在叶中的表达上调, 与转录组测序结果中基

表 5 丹参叶中受适度干旱胁迫调控的萜类代谢途径基因

Table 5 Genes of terpenoid metabolic pathway regulated by moderate drought stress in leaves of *S. miltiorrhiza*

| 基因编号            | 表达变化 | 基因注释   | KEGG 途径    |
|-----------------|------|--|------------|
| c29089.graph_c0 | 下调   | 甲羟戊酸激酶 (mevalonate kinase, MK)   | MVA        |
| c20558.graph_c1 | 上调   | 3-羟基-3-甲基戊二酸-辅酶 A 还原酶 1 (3-hydroxy-3-methylglutaryl-coenzyme A reductase 1, HMGR1)                                       | MVA        |
| c20558.graph_c0 | 上调   | HMGR1, partial   | MVA        |
| c33737.graph_c0 | 上调   | 4-羟基-3-甲基-2-烯基二磷酸还原酶 2 (4-hydroxy-3-methylbut-2-enyl diphosphate reductase 2, HDR2)                                      | MEP        |
| c26661.graph_c0 | 下调   | 异戊烯基焦磷酸异构酶 (isopentenyl pyrophosphate isomerase, IPP)  | IPP        |
| c18427.graph_c0 | 上调   | 牻牛儿基牻牛儿基焦磷酸还原酶 (geranylgeranyl diphosphate reductase)  | GGPP       |
| c29141.graph_c0 | 下调   | 二磷酸铜酰合酶 (copalyl diphosphate synthase, CPS)  | 二萜         |
| c31122.graph_c0 | 下调   | 贝壳杉烯合酶 (kaurene synthase, KS)  | 二萜         |
| c22036.graph_c0 | 下调   | 赤霉素生物合成相关蛋白 (gibberellin biosynthesis-related protein GA20ox2)   | 二萜         |
| c19023.graph_c0 | 上调   | 预测: 玉米黄质环氧化酶 (PREDICTED: zeaxanthin epoxidase)   | 类胡萝卜素      |
| c9406.graph_c0  | 上调   | 9-顺式环氧双加氧酶 (9-cis epoxycarotenoid dioxygenase)   | 类胡萝卜素      |
| c23592.graph_c0 | 上调   | 预测: $\beta$ -胡萝卜素羟化酶 2 (PREDICTED: beta-carotene hydroxylase 2)  | 类胡萝卜素      |
| c26541.graph_c0 | 上调   | 细胞色素 P450 CYP707A99  | 类胡萝卜素      |
| c26420.graph_c0 | 下调   | 玉米素 O-葡萄糖基酰基转移酶 (zeatin O-glucosyltransferase)   | 玉米素        |
| c36084.graph_c1 | 上调   | 鲨烯单加氧酶 (squalene monooxygenase, SQLE)  | 类固醇、三萜、倍半萜 |
| c21867.graph_c0 | 下调   | 细胞色素 P450 710A11   | 类固醇        |
| c33158.graph_c0 | 上调   | 羽扇豆醇合酶 (lupeol synthase)   | 类固醇        |
| c22442.graph_c0 | 上调   | $\beta$ -1,4-甘露糖糖蛋白 4 $\beta$ -N-乙酰氨基葡萄糖基转移酶 (beta-1,4-mannosyl-glycoprotein 4beta-N-N-聚糖 acetylglucosaminyltransferase) |            |
| c33421.graph_c0 | 上调   | 细胞色素 P450 CYP73A120  | 泛醌         |
| c30422.graph_c0 | 下调   | 预测: probable NAD (P) H dehydrogenase (quinone) FQR1-like 3 isoform X1  | 泛醌         |
| c20556.graph_c0 | 上调   | 4-香豆酸: 辅酶 A 连接酶 9 (4-coumarate:coenzyme A ligase 9, 4CL9)  | 泛醌         |
| c28345.graph_c0 | 上调   | 酪氨酸转氨酶 (tyrosine aminotransferase, TAT)  | 泛醌         |

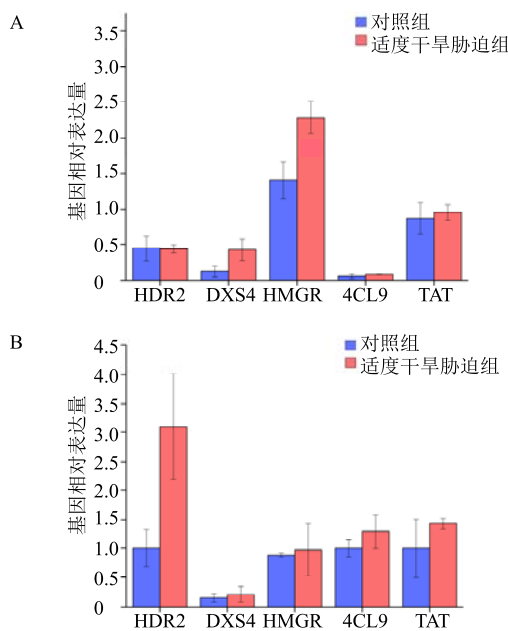


图 3 适度干旱胁迫下 5 个 Unigene 在丹参根 (A) 和叶 (B) 中的相对表达量

Fig. 3 Relative expression of five terpenoid genes in roots (A) and leaves (B) of *S. miltiorrhiza* under moderate drought stress

因的表达变化趋势基本一致。

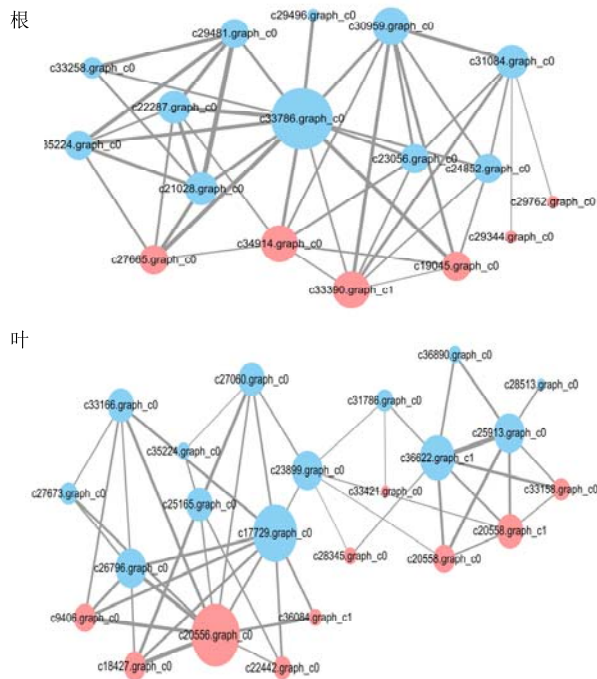
### 3.6 适度干旱胁迫对转录因子的影响

丹参转录组数据中共有 1 128 条 Unigene 被注释为转录因子。适度干旱胁迫下, 丹参根中有来自 38 个家族的 126 个转录因子基因的表达发生变化, 叶中有来自 28 个家族的 73 个转录因子基因的表达发生变化, 其中上调的转录因子分别占总数的 84%和 65%。从表 6 可知, 根和叶中差异表达较多的转录因子类型主要是 AP2/ERF、bHLH、bZIP、WRKY、MYB 等, 其中 AP2/ERF 的数目最多, 在根中的表达全部上调。HSF 转录因子在根和叶中均表现为上调, 而 C2H2、GRAS 等转录因子对适度干旱胁迫的响应具有组织特异性。

从与已鉴定的途径中关键酶基因共表达的基因中筛选未知关键酶或转录因子, 是目前分离候选基因的有效手段。通过构建差异表达的萜类基因与转录因子间的共表达网络发现, 适度干旱胁迫下, 主要是表达上调的萜类基因与转录因子共表达, 而表达下调的萜类基因与差异表达的转录因子间无显著的相关性。根和叶中有不同转录因子家族基因与上调萜类基因间形成共表达网络 (图 4), 根中包括

表 6 适度干旱胁迫下主要差异表达的转录因子  
Table 6 Major differentially expressed transcription factors under moderate drought stress

| TF 家族       | 根        |          | 叶        |          |
|-------------|----------|----------|----------|----------|
|             | 上调 TF 数目 | 下调 TF 数目 | 上调 TF 数目 | 下调 TF 数目 |
| AP2/ERF     | 13       | 0        | 7        | 3        |
| HSF         | 12       | 0        | 3        | 0        |
| C2H2        | 8        | 2        | 0        | 0        |
| bHLH        | 6        | 3        | 4        | 2        |
| bZIP        | 6        | 3        | 4        | 1        |
| WRKY        | 8        | 1        | 3        | 2        |
| GRAS        | 6        | 0        | 0        | 0        |
| HB-HD-ZIP   | 5        | 1        | 3        | 1        |
| NAC         | 5        | 1        | 1        | 0        |
| MYB         | 3        | 1        | 2        | 3        |
| MYB-related | 5        | 0        | 3        | 2        |
| LOB         | 2        | 2        | 2        | 0        |



每个节点代表 1 个 Unigene, 蓝色节点为转录因子, 红色节点为萜类基因; 线的粗细程度代表基因间相关性, 线越粗, 相关性越高

Each node represents a Unigene. The blue nodes represent TF genes, and the red nodes represent terpenoid biosynthetic genes; The thickness of the lines represents the correlation between genes. The thicker a line is, the higher the correlation is

图 4 萜类基因与转录因子间的共表达网络

Fig. 4 Co-expression networks between terpenoid biosynthetic genes and transcription factors

WRKY ( c30959.graph\_c0 、 c29496.graph\_c0 、 c22287.graph\_c0 、 c23056.graph\_c0 )、 AP2/ERF ( c33786.graph\_c0 、 c24852.graph\_c0 )、 GRAS ( c31084.graph\_c0 、 c21028.graph\_c0 )、 NAC ( c33258.graph\_c0 ) 等转录因子, 叶中则是 WRKY ( c31786.graph\_c0 )、 Dof ( c23899.graph\_c0 、 c36622.graph\_c1 )、 bHLH ( c27673.graph\_c0 )、 bZIP ( c33166.graph\_c0 )、 YABBY ( c25913.graph\_c0 )、 C3H ( c26796.graph\_c0 )、 LOB ( c17729.graph\_c0 ) 等转录因子。

#### 4 讨论

干旱是影响植物生长发育最重要的环境因素之一, 植物在适应干旱胁迫的过程中形成了复杂的基因表达调控网络<sup>[14]</sup>。利用高通量测序技术能够为植物基因表达的全面分析提供合理的数据资源, 对揭示植物耐旱机制、发掘耐旱基因、培育抗性品种具有重要意义<sup>[15]</sup>。

本研究对适度干旱胁迫下的丹参不同组织进行了转录组测序和差异表达基因分析。GO 富集结果表明, 适度干旱胁迫下, 丹参根和叶中 DEGs 在 GO 功能部位中的分布基本一致, 均富集到代谢过程、细胞结构、刺激响应、催化活性等功能中, 和甘草根在适度干旱胁迫下的转录组 GO 分析结果类似<sup>[16]</sup>, 表明丹参对于干旱胁迫初期的响应可能是涉及整个植株的。但叶细胞核中的 FRK1 基因的表达下调, 胞外区部位的 SCP 和 CTSL 基因的表达上调。FRK 作为己糖激酶家族的一员, 具有双重功能, 它既能催化果糖的磷酸化又能调节糖代谢和激素之间的关系<sup>[17]</sup>。SCP 主要涉及到植物次级代谢产物中的酰基转移、油菜素内酯信号转导途径、伤害应答、细胞程序性死亡等<sup>[18]</sup>。CTSL 是溶酶体半胱氨酸蛋白酶家族的主要成员, 能参与蛋白质水解、细胞凋亡等多种生理活动<sup>[19]</sup>。适度干旱胁迫对这 3 种酶的不同调控, 表明干旱胁迫初期可通过抑制丹参叶中的糖代谢, 诱导初生壁细胞的程序性死亡等方式, 进而抑制丹参地上部分的生长。

KEGG 途径富集分析表明, 丹参根和叶中受适度干旱胁迫调控的途径存在显著差异。根中 DNA 复制、信号途径、初生和次生代谢途径基因均受到适度干旱胁迫的不同影响, 而叶中只有多种代谢途径基因受干旱胁迫的调控, 表明根对于干旱胁迫的响应更为复杂。土壤适度干旱可提高丹参叶片 TAT 酶活性<sup>[20]</sup>, 本研究中苯丙烷类生物合成途径

的关键酶基因 4CL9、TAT 等在丹参叶中的表达上调,可能是适度干旱胁迫促进丹参叶中丹酚酸类成分积累的基础。丹参酮是一类二萜醌类化合物,主要在丹参根中合成<sup>[21]</sup>。赵志刚等<sup>[22]</sup>的研究表明,丹参根中有效成分在丹参生长前期不断积累,达到最高值后会持续降低到枯黄期,在第 2 年随着新芽萌动才又逐渐升高。本实验室先前研究也表明,丹参根中活性成分开始积累后大概在 2 个月内会保持增长的趋势,达到积累高峰后开始下降<sup>[23]</sup>。本研究开始胁迫处理时丹参根已有活性成分积累(丹参根变红),属于丹参酮积累的第一个时期。研究发现,在丹参毛状根中过表达 SmHMGR<sup>[24]</sup>、SmDXS<sup>[25]</sup>等基因均可以显著提高丹参酮类成分含量。所以适度干旱胁迫下,丹参幼苗中丹参酮类成分的积累增加可能与根中萜类生物合成关键酶基因 HMGR、DXS4 的表达上调有关。

当遭受干旱等非生物胁迫时,植物可通过转录因子间形成的级联调控网络开启抗逆基因的表达以应对外界的不利环境<sup>[26]</sup>。本研究表明适度干旱胁迫可调控丹参中多种转录因子的表达,且以上调方式为主。其中 AP2/ERF、bHLH、bZIP、WRKY、MYB 等是在根和叶中表达量变化均较显著的几类转录因子,这些转录因子在干旱胁迫应答调控网络中发挥重要作用,已在拟南芥<sup>[27]</sup>、水稻<sup>[28]</sup>、大豆<sup>[29]</sup>等多种植物中得到证实。适度干旱胁迫能诱导丹参中热激转录因子 HSF 的表达,与干旱胁迫下油菜中 HSF 的表达一致<sup>[30]</sup>。WRKY、AP2/ERF、bHLH、bZIP 等转录因子还广泛参与植物萜类代谢途径的调控<sup>[31]</sup>,许多学者通过分析转录因子、关键酶基因、次生代谢产物积累变化之间的相关性,成功从丹参基因组中筛选出了多个可能参与丹参酮生物合成调控的 WRKY<sup>[32]</sup>、AP2/ERF<sup>[33]</sup>、bHLH<sup>[34]</sup>、GRAS<sup>[35]</sup>类转录因子。C2H2、C3H、C2C2 (Dof、CO-LIKE、YABBY、GATA、LSD) 是植物中的锌指蛋白类转录因子,能够调控耐逆相关基因的表达,在植物发育和抗逆反应中具有重要意义<sup>[36]</sup>。适度干旱胁迫下,WRKY、AP2/ERF、GRAS、bHLH、bZIP、Dof 等转录因子分别与根和叶中表达上调的萜类基因形成基因共表达网络,表明这些转录因子可能是适度干旱胁迫促进丹参萜类生物合成研究重要的分子线索。

本研究通过高通量转录组测序,初步揭示了适

度干旱胁迫对丹参不同组织基因表达的调控特征,对其中一些关键基因和转录因子进行深入研究,可为丹参药效成分的生物合成机制和在栽培中的合理灌溉提供科学依据。

#### 参考文献

- [1] 董帅,王辉,谢治深.丹参功用本草考证及现代药理认识[J].辽宁中医药大学学报,2019,21(11):152-155.
- [2] 姜雪,史磊.丹参活性成分及药理作用研究进展[J].药学研究,2017,36(3):166-169.
- [3] 时敏,王瑶,周伟,等.药用植物萜类化合物的生物合成与代谢调控研究进展[J].中国科学:生命科学,2018,48(4):8-20.
- [4] 韩立敏.丹参酚酸类成分生物合成途径研究进展[J].安徽农业科学,2014(20):6562-6564.
- [5] 黄璐琦,郭兰萍.环境胁迫下次生代谢产物的积累及道地药材的形成[J].中国中药杂志,2007,32(4):277-280.
- [6] 李光跃,罗晓雅,孙窗舒,等.干旱胁迫对黄芪植株生长中黄酮类成分积累的影响[J].西北植物学报,2017,37(1):138-143.
- [7] 刘大会,郭兰萍,黄璐琦,等.土壤水分含量对丹参幼苗生长及有效成分的影响[J].中国中药杂志,2011,36(3):321-325.
- [8] 梁宗锁,杨东风,杨宗岐,等.不同土壤水分对丹参叶片总酚酸类成分积累及相关酶活性的影响[J].浙江理工大学学报,2013,30(4):573-578.
- [9] Zheng J, Fu J, Gou M, et al. Genome-wide transcriptome analysis of two maize inbred lines under drought stress[J]. *Plant Mol Biol*, 2009, 72(4/5): 407-421.
- [10] 高扬,梁宗锁,刘燕.不同土壤水分条件下丹参耗水特征与水分利用率的研究[J].西北植物学报,2004,24(13):2221-2227.
- [11] 王微.土壤水分条件对丹参生长和化学成分含量的影响、机制及其应用基础[D].上海:复旦大学,2014.
- [12] 左小容.干旱对丹参的生长及有效成分积累的影响[D].杨凌:西北农林科技大学,2011.
- [13] Coman D, Rütimann P, Gruissem W. A flexible protocol for targeted gene co-expression network analysis[J]. *Methods Mol Biol*, 2014, 1153: 285-299.
- [14] 胡守景.植物应答干旱胁迫的基因表达调控[J].中国农业科技导报,2008,10(4):1-6.
- [15] 王丹,龚荣高.干旱胁迫下枇杷叶片的转录组分析[J].华北农学报,2017,32(1):60-67.
- [16] 张春荣,桑雪雨,渠萌,等.基于转录组测序揭示适度干旱胁迫对甘草根基因表达的调控[J].中国中药杂志,2015,40(24):4817-4823.

- [17] 陈毅鸿, 张积森. 植物果糖激酶基因家族演化及其功能 [J]. 分子植物育种, 2016, 14(2): 359-369.
- [18] 叶玲飞, 罗光宇, 向建华, 等. 丝氨酸羧肽酶及其类蛋白的研究进展 [J]. 植物生理学报, 2013, 49(2): 122-126.
- [19] 杨东辉, 刘宇, 肖蓉, 等. 组织蛋白酶 L 的结构与功能 [J]. 中国生物化学与分子生物学报, 2012, 28(12): 1093-1099.
- [20] 刘晓蕾. 干旱下丹参酚酸类化合物积累机制及其抗旱性研究 [D]. 杨凌: 西北农林科技大学, 2011.
- [21] 李坤, 林爱真, 杨媛, 等. 丹参酮生物合成分子调控机制的研究进展 [J]. 浙江农林大学学报, 2018, 35(2): 367-375.
- [22] 赵志刚, 郜舒蕊, 谢景, 等. 不同生长时期丹参生物量积累及活性成分的变化规律 [J]. 中国现代中药, 2015, 17(11): 1171-1176.
- [23] 穆帝秀. 不同磷水平下 AM 真菌对丹参生长发育的影响 [D]. 成都: 成都中医药大学, 2018.
- [24] Zhubo D, Cui G, Zhou S, *et al.* Cloning and characterization of a novel 3-hydroxy-3-methylglutaryl coenzyme A reductase gene from *Salvia miltiorrhiza* involved in diterpenoid tanshinone accumulation [J]. *J Plant Physiol*, 2011, 168(2): 148-157.
- [25] Shi M, Luo X, Ju G, *et al.* Increased accumulation of the cardio-cerebrovascular disease treatment drug tanshinone in *Salvia miltiorrhiza* hairy roots by the enzymes 3-hydroxy-3-methylglutaryl CoA reductase and 1-deoxy-d-xylulose 5-phosphate reductoisomerase [J]. *Funct Integrat Genom*, 2014, 14(3): 603-615.
- [26] Zhang J Z, Zhu R A C A. From laboratory to field. using information from *Arabidopsis* to engineer salt, cold, and drought tolerance in crops [J]. *Plant Physiology*, 2004, 135(2): 615-621.
- [27] 王凯悦, 陈芳泉, 黄五星. 植物干旱胁迫响应机制研究进展 [J]. 中国农业科技导报, 2019, 21(2): 19-25.
- [28] 张磊, 许泉. 水稻干旱胁迫相关转录因子研究进展 [J]. 江西农业学报, 2014, 26(10): 5-11.
- [29] 李雁杰, 朱丹华, 董德坤. 大豆抗旱相关转录因子研究进展 [J]. 长江大学学报: 自科版, 2014, 11(5): 50-56.
- [30] Zhu X, Huang C, Zhang L, *et al.* Systematic analysis of Hsf family genes in the *Brassica napus* genome reveals novel responses to heat, drought and high CO<sub>2</sub> stresses [J]. *Front Plant Sci*, 2017, 8: 1174-1180.
- [31] 赵恒伟, 葛锋, 孙颖, 等. 植物萜类物质生物合成的相关转录因子及其应用前景 [J]. 中草药, 2012, 43(12): 2512-2519.
- [32] Hai Z Y, Guo W, Yang D, *et al.* Transcriptional profiles of SmWRKY family genes and their putative roles in the biosynthesis of tanshinone and phenolic acids in *Salvia miltiorrhiza* [J]. *Int J Mol Sci*, 2018, 19(6): 1593-1600.
- [33] 吕海洲. AP 2% 2FERF 转录因子调控丹参活性成分生物合成的功能研究 [D]. 北京: 北京协和医学院, 2017.
- [34] Zhang X, Luo H, Xu Z, *et al.* Genome-wide characterisation and analysis of bHLH transcription factors related to tanshinone biosynthesis in *Salvia miltiorrhiza* [J]. *Sci Rep*, 2015, 5: 11244-11250.
- [35] Bai Z, Xia P, Wang R, *et al.* Molecular cloning and characterization of five SmGRAS genes associated with tanshinone biosynthesis in *Salvia miltiorrhiza* hairy roots [J]. *PLoS One*, 2017, 12(9): e185322.
- [36] 向建华, 李灵之, 陈信波. 植物非生物逆境相关锌指蛋白基因的研究进展 [J]. 核农学报, 2012, 26(4): 666-672.