

不同生长年限巴戟天蒽醌类含量差异比较及转录组学挖掘蒽醌类生物合成相关基因

梁 雪, 黄 鼎, 黄雪慧, 黄 勇, 张超裕, 王素娥, 田 慧*

广西中医药大学药学院, 广西 南宁 530000

摘 要: 目的 研究不同生长年限巴戟天 *Morinda officinalis* 根部蒽醌类化合物的含量差异及生物合成相关基因。方法 采用 HPLC 测定 1 年生和 6 年生巴戟天根中蒽醌类化合物成分, 并进行比较转录组分析, 采用 qRT-PCR 验证基因表达结果。结果 HPLC 检测结果显示在 6 年生根中蒽醌类物质显著增加, 通过比较转录组分析, 从 1 年生和 6 年生巴戟天的根中筛选到 8 619 条差异表达基因, KEGG 分析表明差异基因主要富集在各种次生代谢物的生物合成、苯丙素生物合成、泛醌和其他萜类醌生物合成等通路中。此外, 共鉴定到 13 条蒽醌生物合成途径相关的差异表达 Unigenes, 其中 *ICS*、*SK*、*CS*、*DHNA-CoA synthase* 均在 6 年生根中上调表达, 而 *DXS*、*DXR*、*DAHPS*、*IDH* 均下调表达, *DHQP/SDH* 既有上调表达的又有下调表达的 Unigenes。进一步选择 6 个候选基因进行 qRT-PCR 验证, 验证结果与转录组数据一致。结论 为鉴别参与巴戟天蒽醌生物合成的相关基因提供参考, 为后期进一步深入研究蒽醌类生物合成的累积规律分子机制提供基础数据。

关键词: 巴戟天; 比较转录组; 蒽醌类生物合成; 差异表达基因; 生长年限

中图分类号: R286.2 文献标志码: A 文章编号: 0253-2670(2024)11-3845-10

DOI: 10.7501/j.issn.0253-2670.2024.11.026

Comparison of anthraquinone content in *Morinda officinalis* at different growth years and transcriptomics mining of anthraquinone biosynthesis related genes

LIANG Xue, HUANG Ding, HUANG Xue-hui, HUANG Yong, ZHANG Chaoyu, WANG Sue, TIAN Hui

College of Pharmacy, Guangxi University of traditional Chinese medicine, Nanning 530000, China

Abstract: Objective To explore differences of anthraquinones content, the genes related to biosynthesis of anthraquinones in the roots of Bajitian (*Morinda officinalis* How) in different growth years. **Methods** The components of anthraquinones by HPLC in the roots of 1-year-old and 6-year-old *M. officinalis* roots were determined, the transcriptomes were sequenced and compared, genes expression were verified by RT-qPCR. **Results** The results of HPLC showed that anthraquinones increased significantly after six years of rooting. By comparative transcriptome analysis, 8 619 differentially expressed genes were screened from 1-year-old and 6-year-old *M. officinalis*, KEGG analysis showed that the differential genes were mainly concentrated in biosynthesis of various secondary metabolites-part 2, phenylpropanoid biosynthesis, ubiquinone and other terpenoid-quinone biosynthesis and etc. A total of 13 differentially expressed Unigenes were identified from the anthraquinone biosynthesis pathway, the expression of *ICS*, *SK*, *CS*, *DHNA-CoA synthase* were significantly up-regulated in 6-year-old *M. officinalis*, *DXS*, *DXR*, *DAHPS*, *IDH* were significantly down-regulated. *DHQP/SDH* were up-regulated and down-regulated Unigenes. RT-qPCR was performed on six candidate genes, and the gene expression results were consistent with transcriptome sequencing. **Conclusions** This experiment provided reference for identifying the relevant genes involved in anthraquinone biosynthesis, and laid the foundation for further in-depth research on the molecular mechanism of accumulation of anthraquinone biosynthesis in *M. officinalis*.

Key words: *Morinda officinalis* How; comparative transcriptomics; anthraquinone biosynthesis; differential expression gene; growth years

收稿日期: 2023-12-05

基金项目: 国家自然科学基金资助项目 (82160715); 中央本级重大增减支项目 (2060302); 2022 年国家级广西大学生创新创业训练计划项目 (202210600004); 广西研究生教育创新计划项目 (YCSW2023393)

作者简介: 梁 雪, 硕士研究生, 研究方向为中药资源学。E-mail: 2739990547@qq.com

*通信作者: 田 慧, 教授, 研究方向为中药质量评价与分子生物学。E-mail: 377244732@qq.com

巴戟天来源于茜草科巴戟天 *Morinda officinalis* How 的干燥根^[1], 又名紫巴戟、鸡肠风等^[2], 性味甘、辛、微温, 归肾、肝经, 具有补肾阳, 强筋骨, 祛风湿的功效, 为我国“四大南药”之一。适用于治疗肾虚腰膝疼痛、筋骨痿软无力、风湿痹证等。研究表明, 寡糖、蒽醌、环烯醚萜苷类是巴戟天中主要的化学成分, 现已从巴戟天中鉴定分离出 40 多种蒽醌^[3], 药理研究表明巴戟天蒽醌具有治疗类风湿性关节炎、调节骨代谢和抗骨质疏松的作用^[4-5], 是巴戟天传统功效“祛风湿、强筋骨”药效物质基础之一; 还能抗癌、抗氧化和降血糖等^[6-7]; 其中的蒽醌类甲基异茜草素-1-甲醚具有抗衰老、降血压作用^[8-9]。可见, 蒽醌类是巴戟天的重要的活性成分类别之一, 有必要对其合成途径基因的挖掘, 这也是深入理解并调控巴戟天有效成分的生物合成过程的先决条件。

在植物体内蒽醌类成分生物合成途径较为复杂, 目前普遍认为其主要来源于 2 条代谢途径, 包括莽草酸/邻琥珀酰苯甲酸途径 (shikimate/o-succinylbenzoic acid route) 和聚酮途径 (polyketide pathway)。茜草科植物体内的蒽醌生物合成来源于莽草酸/邻琥珀酰苯甲酸途径, 该途径涉及莽草酸途径、三羧酸 (tricarboxylic acid cycle, TCA) 循环、甲羟戊酸途径 (mevalonate pathway, MVA) 途径和甲基赤藓醇磷酸 (methyl erythritol phosphate, MEP) 途径分别提供异分支酸 (isochorismic acid, IC)、 α -酮戊二酸 (α -ketoglutarate, AKG)、二甲基烯丙基焦磷酸 (3,3-dimethylallyl diphosphate, DMAPP) 为物质基础, 经过该途径产生的蒽醌也称之为茜草型蒽醌。蒽醌生物合成途径包括几十步的连续酶促反应, 关键酶有异分支酸合成酶 (isochorismate synthase, ICS)、莽草酸激酶 (shikimate kinase, SK)、脱氧木酮糖磷酸合成酶 (1-deoxy-D-xylulose-5-phosphate synthase, DXS)、脱氧木酮糖磷酸还原异构酶 (1-deoxy-D-xylulose-5-phosphate reductoisomerase, DXR) 等^[10-11]。已有报道对巴戟天的 *MoDXR*^[12]、*MoDXS*^[13]、*MoTPS*^[14]、*PAL*、*C4H*^[15]、*ICS*^[16]等基因分别单独进行克隆, 研究其结构和功能。此外, 生长年限对巴戟天蒽醌的累积有明显影响, 有学者研究发现 1~6 年生巴戟天总蒽醌含量随着生长年限的增加呈现出较大幅度的增加^[17-18], 但巴戟天蒽醌类生物合成途径酶基因表达在不同生长年限的表达分析还未见报道。

本研究对同一产地不同生长年限的巴戟天样品

的 4 种蒽醌类物质含量进行测定, 并通过比较转录组分析, 筛选不同年份之间的差异表达基因, 并从中挖掘到参与巴戟天蒽醌生物合成相关的差异表达基因, 最后通过实时荧光定量 PCR 法 (qRT-PCR) 验证了这些差异表达基因的表达规律, 为巴戟天蒽醌类化合物的生物合成调控和累积规律的分子机制提供基础数据。

1 材料与试剂

1.1 材料

试验用材料于 2022 年 3 月采自广东省肇庆市德庆县 (111°79'15"E, 23°15'13"N, 海拔 140 m), 1 年生 (编号 YN) 及 6 年生 (编号 LN) 的巴戟天根各 3 个重复 (每个重复 5 株), 分组为 YN1、YN2、YN3 和 LN1、LN2、LN3, 液氮速冻, 经广西中医药大学梁子宁教授鉴定为茜草科植物巴戟天 *M. officinalis* How, 用于蒽醌含量检测及总 RNA 提取。

1.2 仪器

Waters e2695 高效液相色谱仪 (美国 Waters 公司); CPA 225D 型十万分之一电子分析天平 (德国赛多利斯科学仪器有限公司); KQ5200DE 型数控超声清洗器 (昆山市超声仪器有限公司); DHG-9240A 电热鼓风干燥箱 (上海齐欣科学仪器有限公司)。RNA prep pure 植物总 RNA 提取试剂盒 (TIANGEN 公司); QuantstudioTM5 荧光定量 PCR 仪 (美国 ABI 公司); NANO DROP 8000 型紫外分光光度计 (美国 Thermo 公司); DYY-6C 型凝胶电泳仪 (北京六一生物科技有限公司); Hiscript QRT super Mix for qPCR (+gDNA wiper) 逆转录试剂盒 (南京诺唯赞生物科技股份有限公司); SYBR® Select Master Mix (2X) 定量 PCR 试剂盒 (美国 ABI 公司)。

1.3 试剂

对照品 2-羟基-1-甲氧基蒽醌 (批号 MUST-21120122)、甲基异茜草素-1-甲醚 (批号 MUST-21083010)、甲基异茜草素 (批号 MUST-21052511) 均购自成都曼思特生物科技有限公司; 蒽醌-2-羧酸 (批号 A77951) 购自上海吉至生化科技有限公司, 质量分数均 $\geq 98\%$; 甲醇、磷酸均为色谱纯; 水为超纯水 (Milli-Q 超纯水系统制备); 其他试剂均为分析纯。

2 方法

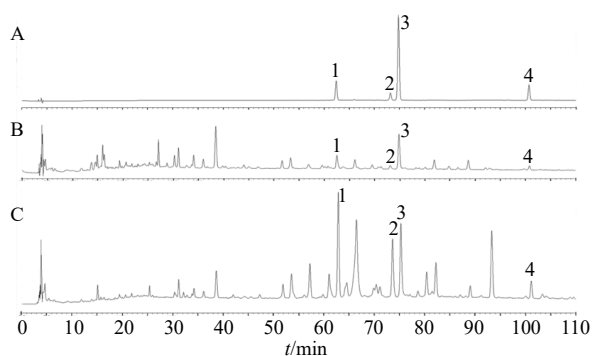
2.1 蒽醌类成分定量测定

2.1.1 供试品溶液的制备 称取每组 1 年生及 6 年生巴戟天新鲜根 10 g, 置于烘箱 50 °C 去芯烘干,

研磨粉碎混匀,精密称取巴戟天粉末 0.5 g(过 50 目筛),加 25 mL 甲醇回流提取 2 h,滤过,蒸干,残渣加甲醇溶解,转移至 5 mL 量瓶中,加甲醇至刻度,摇匀,过 0.45 μm 微孔滤膜,取续滤液作为供试品溶液。

2.1.2 对照品溶液的制备 精密称取 2-羟基-1-甲氧基蒽醌、蒽醌-2-羧酸、甲基异茜草素-1-甲醚、甲基异茜草素对照品适量,用甲醇溶解制成质量浓度分别为 56.000、50.000、15.600、3.800 $\mu\text{g/mL}$ 的混合对照品溶液。

2.1.3 色谱条件 Waters XSelect HSS T3 色谱柱(250 mm $\times$ 4.6 mm, 5 μm), 甲醇(A)-0.2%磷酸水(C)为流动相,梯度洗脱: 0~5 min, 20%A; 5~20 min, 20%~42%A; 20~42 min, 42%~50%A; 42~65 min, 50%~60%A; 65~110 min, 60%~80%A, 检测波长 275 nm, 体积流量 0.7 min/mL, 柱温 35 $^{\circ}\text{C}$, 进样量 10 μL 。色谱图见图 1。



1-2-羟基-1-甲氧基蒽醌; 2-蒽醌-2-羧酸; 3-甲基异茜草素-1-甲醚; 4-甲基异茜草素。
1-2-hydroxy-1-methoxyanthraquinone; 2-anthraquinone-2-carboxylic acid; 3-rubiadin-1-methyl ether; 4-rubiadin.

图 1 混合对照品(A)、1年生巴戟天(B)和6年生巴戟天(C)的HPLC图

Fig. 1 HPLC of mixed reference solution (A), YN sample (B) and LN sample (C)

2.1.4 线性方程的绘制 分别精密吸取“2.1.2”项下的混合对照品溶液 7.00、5.00、2.00、1.00、0.50、0.25、0.05 mL 置于 10 mL 量瓶中,用甲醇稀释到刻度,得到系列质量浓度的溶液。按“2.1.3”项下色谱条件进样测定峰面积,以混合对照品的系列质量浓度为横坐标,峰面积为纵坐标,绘制标准曲线,结果显示 4 个成分线性关系良好。见表 1。

2.1.5 精密度试验 取“2.1.2”项下的混合对照品溶液,按“2.1.3”项下色谱条件连续进样 6 次。2-

表 1 线性方程

Table 1 Linear equations

成分	回归方程	<i>r</i>
2-羟基-1-甲氧基蒽醌	$Y=64\ 509\ 538 X-15\ 594$	0.999 2
蒽醌-2-羧酸	$Y=28\ 022\ 667 X-24\ 823$	0.999 3
甲基异茜草素-1-甲醚	$Y=103\ 045\ 938 X-42\ 230$	0.999 3
甲基异茜草素	$Y=77\ 131\ 412 X-25\ 998$	0.999 3

羟基-1-甲氧基蒽醌、蒽醌-2-羧酸、甲基异茜草素-1-甲醚、甲基异茜草素质量分数的 RSD 分别为 0.80%、0.58%、0.58%、1.10%,表明方法精密度良好。

2.1.6 重复性试验 按“2.1.1”项下方法平行制备巴戟天(LN)供试品溶液 6 份,进样测定。2-羟基-1-甲氧基蒽醌、蒽醌-2-羧酸、甲基异茜草素-1-甲醚、甲基异茜草素质量分数的 RSD 分别为 2.77%、2.26%、2.99%、2.28%,表明方法重复性良好。

2.1.7 稳定性试验 取巴戟天(LN)供试品溶液分别在制备后 0、3、6、12、16、24 h 进样测定。2-羟基-1-甲氧基蒽醌、蒽醌-2-羧酸、甲基异茜草素-1-甲醚、甲基异茜草素峰面积的 RSD 分别为 2.14%、2.09%、1.99%、1.66%,表明供试品溶液在 24 h 内稳定。

2.1.8 加样回收率试验 精密称取巴戟天(LN)粉末 0.25 g,平行 6 份,分别精密加入与样品中待测成分含量相当的混合对照品溶液,按“2.1.1”项下方法制备巴戟天供试品溶液,进样测定。2-羟基-1-甲氧基蒽醌、蒽醌-2-羧酸、甲基异茜草素-1-甲醚、甲基异茜草素的平均加样回收率分别为 100.96%、97.53%、98.75%、100.10%,RSD 分别为 2.62%、1.77%、1.52%、3.99%。

2.1.9 含量测定 取“2.1.1”项下制得的各组巴戟天供试品溶液,按“2.1.3”项下色谱条件测定。

2.2 转录组测序分析

2.2.1 RNA 提取与文库构建 取各组新鲜巴戟天根部冻干打粉,用 RNA prep pure 植物总 RNA 提取试剂盒(TIANGEN 公司)制备总 RNA, Nanophotometer 分光光度计、Qubit 2.0 荧光计和 Agilent 2100 分别检测 RNA 纯度、浓度和完整性;通过 Oligo(dT)磁珠富集带有 polyA 尾的 mRNA,随后加入分裂缓冲液(fragmentation buffer)将 RNA 打断成短片段,以短片段 RNA 为模板,用六碱基随机引物(random hexamers)合成第一链 cDNA,然后加入缓冲液、dNTPs 和 DNA polymerase I 合成二链 cDNA,随后利用 AMPure XP beads 纯化双链

cDNA。纯化的双链 cDNA 再进行末端修复、加 A 尾并连接测序接头，然后用 AMPure XP beads 进行片段大小选择，最后进行 PCR 富集得到最终的 cDNA 文库，使用 Qubit2.0 和 qRT-PCR 方法对文库进行定量检测（文库有效浓度 >2 nmol/L）。

2.2.2 转录组测序及功能注释 库检合格后，不同文库按照目标下机数据量进行 pooling，用 illumina novaseq 6000 平台进行测序。高通量测序仪得到的图像数据经 CASAVA 碱基识别转化为大量高质量的 Data，称为原始数据（raw data）。使用 fastp 软件对数据进行过滤，测序错误率检查、GC 含量分布检查，得到高质量的 reads（clean reads）。采用 Trinity 对 clean reads 进行拼接，然后使用 corset 去冗余得到 Unigene 序列。使用 DIAMOND BLASTX 软件将 Unigene 序列与 KEGG、NR、Swiss-Prot、GO、KOG、Trembl 数据库比对，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。

2.2.3 基因序列比对、差异表达基因挖掘 采用了 RSEM 软件，将 Trinity 组装并去冗余之后的 Unigene 作为参考序列，将每个样品的 clean reads 往参考序列上进行比对，采用 FPKM（fragments Pper kilobase of transcript per million fragments mapped）表达量作为衡量转录本或基因表达水平的指标。使用 DESeq 进行样

品组间的差异表达分析，用 Benjamini-Hochberg 方法对假设检验概率（P value）进行多重假设检验校正，得到错误发现率（false discovery rate, FDR），并计算差异倍数（fold change, FC），差异基因的筛选条件为 $|\log_2 FC| \geq 1$ ，且 $FDR < 0.05$ ，得到的差异基因进行注释及 KEGG 富集分析。

2.3 蒽醌生物合成基因筛选与 qRT-PCR 验证

依据文献提出的高等植物蒽醌类生物合成途径^[10,19]，检索比对差异表达基因的 KEGG、NR、Swissport、TrEMBL、GO、KOG、Pfam 数据库注释信息，并比较其在 1 年生和 6 年生巴戟天的 FPKM 表达量，确定巴戟天蒽醌类生物合成的相关基因。

取“2.2.1”项下提取所得的各组巴戟天总 RNA 采用 Hiscript QRT super Mix for qPCR（+gDNA wiper, Vazyme）试剂盒说明书进行逆转录反应，合成 cDNA。用得到的 cDNA 模板和引物，按照 SYBR® Select Master Mix（2X, ABI）定量 PCR 试剂盒方法进行内参基因筛选预实验。而后对筛选到的 1 个 Actin 内参基因继续按照定量 PCR 试剂盒方法，利用 Quantstudio™ 5 荧光定量 PCR 仪进行实时荧光定量 PCR 实验，每个样品进行 3 次技术重复。通过 $2^{-\Delta\Delta Ct}$ 法对 ICS、SK、CS、DHNA-CoA synthase、DXS、DXR 基因进行相对表达量分析。相关引物全部引物由武汉天一辉远生物科技有限公司合成。见表 2。

表 2 实时荧光定量 PCR 引物信息

Table 2 Primers for real-time fluorescence quantitative PCR

基因名称	引物名称	引物序列（5'-3'）
ICS1	Cluster-24221.5-F	AAGGATGAGACAAAACGGAGA
	Cluster-24221.5-R	TGGAGAGGAAAGAACCGC
ICS2	Cluster-14601.0-F	TCATCATTCGCTCTTTTCATTTG
	Cluster-14601.0-R	GGTGCTGAGCGTGAGTATTTTC
SK1	Cluster-26492.5-F	CGAAATCCCTCCTCATCCT
	Cluster-26492.5-R	AACAGTTCCATCAAATCCAATC
SK2	Cluster-18646.2-F	ATAGTTGAGCCTTGAACAGACC
	Cluster-18646.2-R	AAGCCAAGTTAGAACGAGAGG
CS	Cluster-17119.10-F	GAAGTGGTCCCTCCAGTCC
	Cluster-17119.10-R	GAACAAATGCCAATAGCCC
DHNA-CoA synthase	Cluster-22418.29-F	AACGGTCAATAAAATCACGC
	Cluster-22418.29-R	CAGATGGGTTCTCCCTCAGT
DXS	Cluster-12750.0-F	TCCAGTTGAGTTTGCCACC
	Cluster-12750.0-R	GGATCCGCAGACTAGAGCC
DXR	Cluster-11858.5-F	TAGAATGTATGATGGATTGGGG
	Cluster-11858.5-R	GGGGGTGCTTTTAGGGAT
Actin	Actin-F	CATTGTGCTCAGTGGTGGC
	Actin-R	AATCTTCATGCTACTCGGGG

3 结果与分析

3.1 1 年生和 6 年生巴戟天蒽醌类 HPLC 含量测定

1 年生和 6 年生巴戟天的 HPLC 图谱分析结果见表 3，2 种年限的巴戟天中蒽醌类含量存在明显

差异，与 1 年生相比，6 年生巴戟天蒽醌含量显著升高（ $P < 0.01$ ）。6 年生中的 2-羟基-1-甲氧基蒽醌、蒽醌-2-羧酸、甲基异茜草素-1-甲醚、甲基异茜草素的含量分别是 1 年生的 14、29、2、2 倍。

表 3 巴戟天 4 个蒽醌类成分含量测定结果 ($\bar{x} \pm s, n = 3$)

Table 3 Determination results of four anthraquinones in *M. officinalis* ($\bar{x} \pm s, n = 3$)

成分	质量分数/($\mu\text{g} \cdot \text{g}^{-1}$)	
	1 年生	6 年生
2-羟基-1-甲氧基蒽醌	13.848 $\pm$ 0.095	203.467 $\pm$ 1.122**
蒽醌-2-羧酸	10.008 $\pm$ 0.011	293.840 $\pm$ 8.281**
甲基异茜草素-1-甲醚	57.819 $\pm$ 1.680	100.518 $\pm$ 2.320**
甲基异茜草素	14.858 $\pm$ 0.275	28.872 $\pm$ 0.858**

与一年生比较: ** $P < 0.01$ 。

** $P < 0.01$ indicate significant differences.

3.2 RNA 提取与质量控制

对 1 年生和 6 年生巴戟天进行总 RNA 提取, 检测得到各样本 RNA 质量浓度均大于 41.00 ng/ μL , $A_{260}/A_{280} > 1.86$, $A_{260}/A_{230} > 1.81$, 总量大于 1.43 ng/ μL , RNA 完整数量 (RIN) 大于 8.40, 说明 RNA 质量和完整性较高, 可以进行下一步高通量测序。

3.3 测序与组装结果

高通量测序共获得 43.90G clean Base, 各样品 clean Base 均达到 6 G, Q20 碱基比都在 97% 以上, Q30 都在 92% 以上, GC 含量为 43.44%~43.74%, 说明测序质控良好, clean reads 质量合格, 见表 4。Trinity 对 clean reads 进行拼接得到 107 732 条 Transcript, 进一步拼接得到 65 939 条 Unigene, Unigene 序列碱基总数为 137 901 438, 平均长度为 2 091 bp, N50、N90 分别为 2 982、1 095 bp, $\geq 2 000$ bp 的 Transcript 和 Unigene 最多, 组装长度最长的序列占比最大, 可见组装完整性较高, 能够满足后续生物信息学研究。

将测序数据 clean reads 与组装结果 Unigene 进行序列比对, 以 FPKM 值对基因表达定量, 得到映射读数 (mapped reads) 所占百分比均不小于 87.48%, 说明转录组有效数据充分。见表 5。

表 4 巴戟天测序数据统计

Table 4 *M. officinalis* sequencing data statistics

样品	原始数据/条	纯净数据/条	纯净数据总量/G	Q20/%	Q30/%	GC/%
YN1	52 762 268	51 126 204	7.67	97.11	92.23	43.49
YN2	53 095 274	51 264 190	7.69	97.40	92.81	43.54
YN3	44 859 726	43 887 330	6.58	97.40	92.52	43.52
LN1	50 832 626	47 159 504	7.07	97.44	92.93	43.74
LN2	51 760 686	50 449 146	7.57	97.50	93.05	43.57
LN3	50 542 310	48 784 940	7.32	97.51	93.08	43.44

表 5 基因比对率统计

Table 5 Gene comparison rate statistics

样品	有效读数	有效匹配数 (匹配率/%)	单一匹配数 (匹配率/%)	多匹配数 (匹配率/%)
YN1	25 563 102	22 464 696 (87.88)	5 877 365 (22.99)	16 587 331 (64.89)
YN2	25 632 095	22 424 171 (87.48)	5 877 365 (22.99)	16 750 077 (65.35)
YN3	21 943 665	19 597 131 (89.31)	4 804 454 (21.89)	14 792 677 (67.41)
LN1	23 579 752	20 841 173 (88.39)	4 994 127 (21.18)	15 847 046 (67.21)
LN2	25 224 573	22 165 198 (87.87)	5 341 190 (21.17)	16 824 008 (66.70)
LN3	24 392 470	21 481 765 (88.07)	5 084 672 (20.85)	16 397 093 (67.22)

3.4 巴戟天的 Unigene 功能注释

将巴戟天的 Unigene 与 TrEMBL、NR、基因本体论 (gene ontology, GO)、SwissProt、Pfam、KOG、KEGG 等公共数据库进行比对, 结果见表 6, Unigene 注释到这些数据库的占比由大到小分别为 TrEMBL、NR、GO、SwissProt、Pfam、KOG、KEGG, 在 TrEMBL 数据库找到的相似序列最多, 为 53 399 条 (占 80.98%), 其次为 NR 数据库, 为 53 229 条 (占 80.72%), 在 KEGG 找到的相似序列最少, 为 33 225 条 (占 50.39%), 在最少一个数据库被注释的 Unigene

表 6 注释统计

Table 6 Annotation statistics

数据库	Unigene/条	占比/%
TrEMBL	53 399	80.98
NR	53 229	80.72
GO	47 297	71.73
SwissProt	40 194	60.96
Pfam	39 950	60.59
KOG	33 884	51.39
KEGG	33 225	50.39
annotated in at least one database	54 014	81.92
total unigenes	65 939	100.00

为 54 014 条, 占 81.92%。

Unigene 序列与 NR 数据库中的序列作比对, 得到近缘物种注释信息, 巴戟天与中粒咖啡 *Coffea canephora* Pierre ex A. Froehner 相似序列匹配度最高, 为 36 977 条 (69.47%), 其后依次是油橄榄 *Olea europaea* var. *sylvestris* L. (2.63%), 芝麻 *Sesamum indicum* L. (2.08%), 栓皮栎 *Quercus suber* Blume (1.36%), 紫花风铃木 *Handroanthus impetiginosus* (Mart. ex DC.) Mattos (1.31%), 葡萄 *Vitis vinifera* L. (1.27%) 等。

3.5 差异表达基因分析

以 $|\log_2FC| \geq 1$, 且 $FDR < 0.05$ 作为筛选标准, 共获得 8 619 个差异表达基因, 对 1 年生和 6 年生巴戟天进行转录本的上调和下调表达分析, 发现上调表达基因有 3 597 个, 下调表达基因有 5 022 个。

GO 可以按照生物学过程 (biological process, BP)、细胞组分 (cellular component, CC)、分子功能 (molecular function, MF) 这 3 大类别从宏观上对巴戟天的差异表达基因进行功能限定和描述, 结果表明, 共有 7 039 条基因序列得到了注释, 在生物学过程分类中, 主要聚集于代谢过程和细胞过程; 在细胞组分中, 主要聚集于含蛋白质复合物和细胞解剖实体; 在 MF 中, 主要聚焦于催化活性和结合。见图 2。

此外, 根据 P 值 < 0.05 选取差异表达基因富集最显著的 KEGG 通路共 29 条, 发现差异表达基因主要富集在各种次生代谢物的生物合成-第 2 部分, 苯丙素生物合成, 淀粉和蔗糖代谢, 戊糖磷

酸途径, 泛醌和其他萜类醌生物合成, 糖酵解/糖异生等。见图 3。

3.6 蒽醌类生物合成相关基因的鉴定

从差异表达基因中筛选与蒽醌类生物合成途径相关的基因, 共筛选到 10 种 (13 条) 差异表达的相关基因, 然后将蒽醌含量测定结果和基因表达水平都整合到通路图中, 见图 4。相关基因包括 ICS (2 条 Unigene)、DXS (1 条 Unigene)、DXR (1 条 Unigene)、DAHPS (1 条 Unigene)、DHQD/SDH (双功能基因, 3 条 Unigene)、SK (2 条 Unigene)、CS (1 条 Unigene)、IDH (1 条 Unigene)、DHNA-CoA synthase (1 条 Unigene)。其中, 6 年生中有 4 种 (6 条) 蒽醌类生物合成基因的表达式远高于 1 年生, 包括 *ICS1* (Unigene 编号: Cluster-24221.5)、*ICS2* (Unigene 编号: Cluster-14601.0)、*SK1* (Unigene 编号: Cluster-26492.5)、*SK2* (Unigene 编号: Cluster-18646.2)、*DHNA-CoA synthase* (Unigene 编号: Cluster-22418.29)、*CS* (Unigene 编号: Cluster-17119.10)。此外, 有 4 种 (4 条) 相关合成基因均为下调表达, 包括 *DXS* (Unigene 编号: Cluster-12750.0)、*DXR* (Unigene 编号: Cluster-11858.5)、*DAHPS* (Unigene 编号: Cluster-11170.0)、*IDH* (Unigene 编号: Cluster-24058.7)。DHQD/SDH (Unigene 编号: Cluster-23411.2, Cluster-15733.0, Cluster-24763.0) 既有上调表达的 Unigene, 又有下调表达的 Unigene。

3.7 候选基因的 qRT-PCR 验证

进一步选择与蒽醌合成相关的 6 条上调表达基

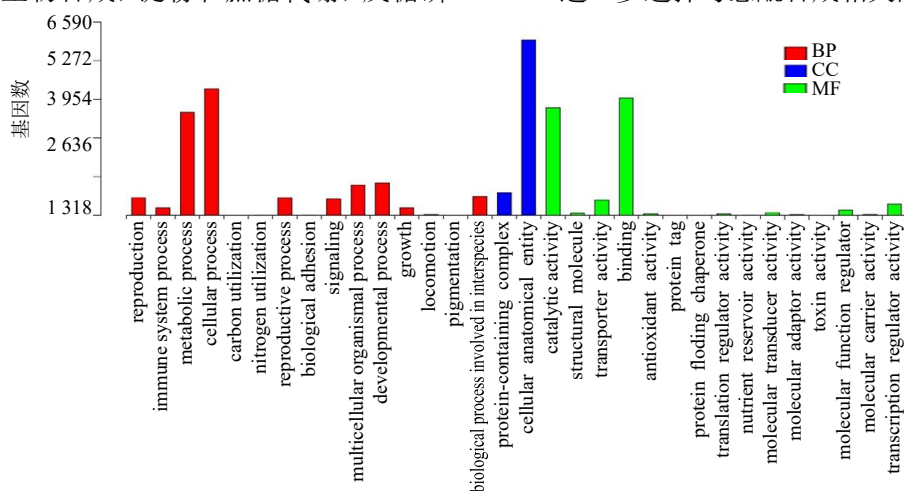


图 2 差异表达基因 GO 分类结果

Fig. 2 GO classification results of DEGs

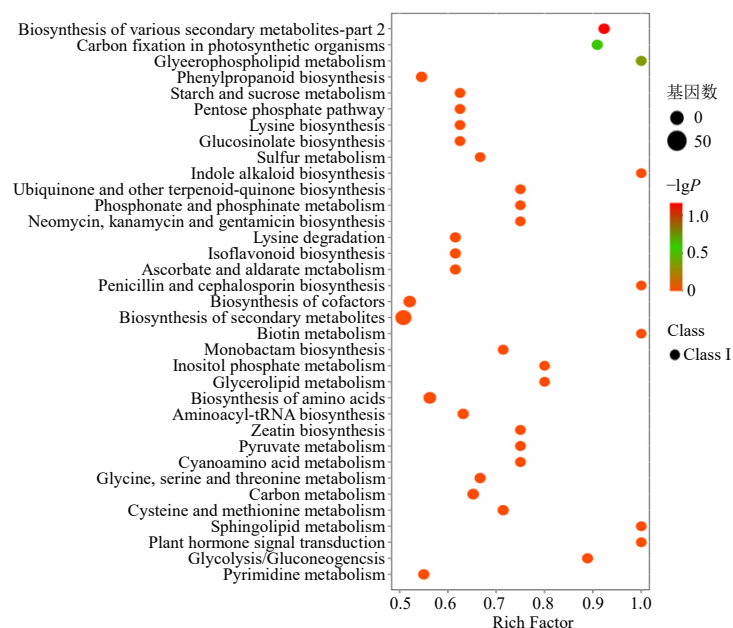
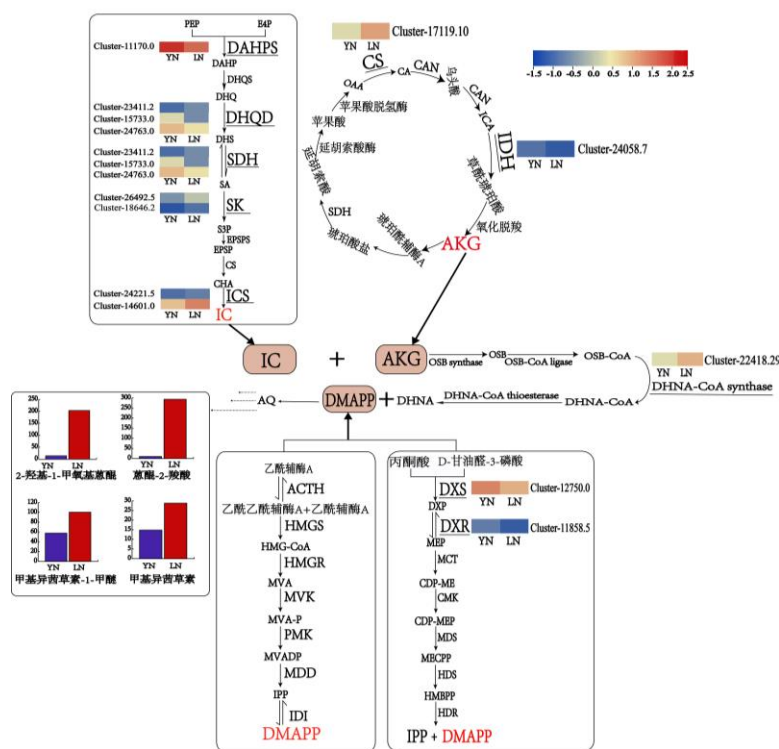


图 3 差异表达基因 KEGG 富集图

Fig. 3 KEGG pathways enriched for DEGs



AQ-蒽醌母核；DAHPS-3-脱氢-D-阿拉伯糖庚酮糖-7-磷酸合成酶；DHQD-3-脱氢喹啉酸脱水酶；SDH-莽草酸脱氢酶；SK-莽草酸激酶；ICS-异分支酸合成酶；CS-柠檬酸合成酶；IDH-异柠檬酸脱氢酶；DHNA-1,4-二羟基-2-苯甲酸-CoA 合成酶；DXS-脱氧木酮糖磷酸合成酶；DXR-脱氧木酮糖磷酸还原异构酶。

AQ-Anthraquinone mother nucleus; DAHPS-3-deoxy-D-arabino-hetulosonate-7-phosphate synthase; DHQD-3-dehydroquininate dehydratase; SDH-shikimate dehydrogenase; SK-shikimate kinase; ICS-isochorismate synthase; CS-citrate synthase; IDH-isocitrate dehydrogenase; DHNA-CoA synthase-1,4-dihydroxy-2-naphthoic acid coenzyme A synthase; DXS-1-deoxy-D-xylulose-5-phosphate synthase; DXR-1-deoxy-D-xylulose-5-phosphate reductoisomerase.

图 4 巴戟天蒽醌类生物合成途径的基因表达和含量测定分析

Fig. 4 Genes expression of anthraquinone biosynthesis pathway and content determination analysis

因和2条下调表达基因进行qRT-PCR,分别为*ICS1*、*ICS2*、*SK1*、*SK2*、*CS*、*DHNA-CoA synthase*、*DXS*、*DXR*,以验证转录组学数据的准确性。结果表明,除*DXS*外,样本候选基因的相对表达量和FPKM值结果一致,表明此次实验测定的转录组数据是可靠性较高,见图5。

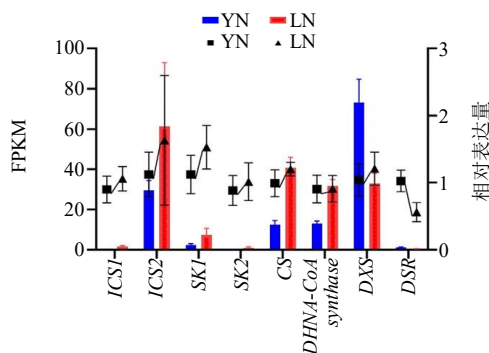


图5 候选酶基因的表达情况 ($\bar{x} \pm s$, $n=3$)

Fig. 5 Expression of differently genes of key enzymes ($\bar{x} \pm s$, $n=3$)

4 讨论

本研究采用 illumina novaseq 6000 平台结合生物信息分析技术,对巴戟天根进行转录组测序和功能注释分析。测序数据经 Trinity 软件组装,获得 65 939 条 Unigene 序列,平均长度 2 091 bp, N50 为 2 982 bp,转录组测序数据质控水平较高,组装效果好,所得的 Unigene 序列长度较长,利用生物信息学对 Unigene 与 NR、GO、KOG 等 7 个数据库进行比对与注释,结果显示注释率为 82.91%,较优于同含有蒽醌的药用植物转录组,如茜草 (80.44%)^[20]、虎杖 (61.04%)^[21]、大黄^[22] (78.27%) 等。以 NR 数据库注释信息为基础研究巴戟天的物种序列相似性,发现中粒咖啡 (69.47%) 的相似序列匹配度最高。上述基因序列的发现极大地丰富了巴戟天的转录组信息,扩充了巴戟天的生物信息,可为后续的基因功能研究提供参考。本实验进一步研究 1 年生和 6 年生巴戟天基因表达量差异,筛选得到 8 619 个差异表达基因,差异基因富集程度较高的 KEGG 代谢通路主要包括各种次生代谢物的生物合成-第二部分,苯丙素生物合成,淀粉和蔗糖代谢,戊糖磷酸途径,泛醌和其他萜类醌生物合成,糖酵解/糖异生等。表明 1 年生和 6 年生巴戟天基因差异表达较大,这暗示了两者在次生代谢物的组成和含量方面可能存在较大的差异。

转录组测序已应用于西洋参^[23]、红豆杉^[24]、红

花^[25]等中药的功能基因研究上,基因的表达调控对次级代谢产物的形成具有重要的调节作用,如在小茜草中,参与蒽醌生物合成的相关基因 *OSB-CoA ligase*、*ICS*、*DXS* 等在根中表达量均高于茎、叶,其根中蒽醌类成分含量更高因而选择根部作为药用部位^[26]。有学者发现抑制 *HMGR* 的表达能 90% 以上抑制蒽醌类紫草素的积累^[27]。已有研究表明,通过转录组学进一步研究植物发育过程的代谢通路中基因表达,有利于了解植物生长发育规律和次生代谢产物合成和累积的调控。周豪楠等^[28]对 6 个生长时期的粉葛葛根素含量与葛根素生物合成基因表达量进行关联分析,发现 *I2'H*、*IF7GT*、*IFS1* 均与葛根素的含量呈现正向相关,即均正向调控葛根素的生物合成;二、三年生蒙古黄芪中差异表达皂苷合成基因和 *CYP*、*UGT* 基因共同参与黄芪皂苷的合成、加工与修饰,这些差异基因大多表达出既可以上调又可以下调,但上调表达数目明显大于下调表达数,三年生三萜皂苷含量明显上升,可能与这些差异表达基因调控有关^[29]。

目前,对巴戟天中蒽醌类成分的研究主要在含量测定、药材质量、化学成分、药理作用方面^[3],而巴戟天蒽醌物质的生物合成途径却少有研究,本研究整合了巴戟天蒽醌生物合成的可能途径,分析了该途径中相关基因在巴戟天根部不同年份的表达情况,其中 *ICS*、*SK*、*CS*、*DHNA-CoA synthase* 在 6 年生的表达量均明显增加,结合对巴戟天蒽醌类含量进行测定,6 年生蒽醌含量明显上升,基因表达量与含量趋势较为一致,表明这些基因在蒽醌的合成和累积中可能发挥重要的调控作用。从整体看,从 1~6 年生,巴戟天不但蒽醌含量发生了累积,蒽醌组成和合成速率可能也在变化。在巴戟天根发育的前、中期,维管形成层的活动旺盛,向外分裂出大量含有代谢产物的次生韧皮部细胞;而在根发育的后期,由于根逐渐老化,形成层活动以产生死细胞的木质部导管和纤维为主,此时,相较于 1 年生巴戟天, *DXS*、*DXR*、*DAHPS*、*IDH* 表达有所下调,推测某些种类蒽醌在 6 年生巴戟天中合成速率较慢。前人研究了其中一些差异表达基因发现, *ICS* 是莽草酸途径中的限速酶,在蒽醌合成调控中发挥重要作用,文献报道在海巴戟细胞培养液中蒽醌的产生总是伴随着 *ICS* 活性的增加而增加,即使在分支酸浓度较低,不足以维持正常生长的极端条件下,蒽醌积累速率与 *ICS* 活性之间仍然存在相关性^[16]。 *SK* 是莽

草酸途径的关键酶, 位于蒽醌合成途径的早期阶段。决明中的 SK 催化蒽醌合成反应主要发生在根与子叶等组织, 形成的中间产物运输到种子等组织中, 完成生物合成的后期阶段并积累蒽醌终产物^[30]。此外, 有学者指出, MEP 途径中的起始步骤由 DXS 催化丙酮酸和甘油醛发生缩合反应生成 DXP, DXS 是 MEP 途径中关键的限速酶, Quevedo 等^[31]通过培育过表达 DXS 基因的海巴戟细胞系, 发现与对照细胞系相比, 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。樊星等^[32]分别在茜草细胞悬浮培养体系中添加 MVA 途径抑制剂洛伐他汀和 MEP 途径抑制剂氯胺, 结果洛伐他汀组羟基茜草素和大叶茜草素产量显著增加, 氯胺组羟基茜草素和大叶茜草素量下降, 暗示茜草中蒽醌类成分可能是通过 MEP 途径合成的。但巴戟天中蒽醌类成分合成途径中的大多数基因如何影响蒽醌累积尚不清楚。

目前, 利用高通量测序技术对巴戟天转录组的研究还处于起步阶段。本实验利用比较转录组分析了不同生长年限巴戟天的基因数据间差异性, 并从差异表达基因中挖掘到了参与巴戟天有效成分蒽醌类生物合成途径的相关基因, 这些基因可能在巴戟天蒽醌合成中发挥重要的作用, 也为研究巴戟天根部中蒽醌类成分的合成调控及其遗传改良提供重要的理论基础。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 中国药典 [S]. 一部. 2020: 83.
- [2] 程超寰. 本草释名考订 [M]. 北京: 中国中医药出版社, 2013: 96.
- [3] 黄清霞, 覃川娴, 何泽源, 等. 巴戟天化学成分、药理作用及质量标志物预测分析 [J]. 中华中医药学刊, 2022, 40(7): 251-258.
- [4] 廖慧君, 张叶金, 王钰, 等. 基于网络药理学探讨巴戟天治疗类风湿性关节炎的物质基础及作用机制 [J]. 按摩与康复医学, 2021, 12(15): 47-52.
- [5] 何玉琼. 巴戟天环烯醚萜苷及蒽醌类成分调节骨代谢的作用及机制研究 [D]. 上海: 中国人民解放军海军军医大学, 2019.
- [6] 杨振民, 欧亚非, 伊勇涛, 等. 巴戟天中蒽醌类化合物及生物活性研究 [J]. 天然产物研究与开发, 2019, 31(1): 87-92.
- [7] 沈巍, 孙艺琦, 张奇, 等. 不同产地和生长年限对巴戟天中环烯醚萜苷类成分形成和积累的影响 [J]. 中草药, 2022, 53(17): 5484-5490.
- [8] 王阳. 甲基异茜草素-1-甲醚的降血压作用及机制研究 [D]. 新乡: 新乡医学院, 2020.
- [9] 王羚邨. 巴戟天抗衰老活性成分及作用机制研究 [D]. 广州: 广州中医药大学, 2011.
- [10] 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.
- [11] Han Y S, Van der Heijden R, Verpoorte R. Biosynthesis of anthraquinones in cell cultures of the Rubiaceae [J]. *Plant Cell Tissue Organ Cult*, 2001, 67(3): 201-220.
- [12] 谢德金, 叶友杰, 杨德明, 等. 巴戟天 MoDXR 基因及其启动子的克隆与分析 [J]. 药科学报, 2020, 55(2): 335-344.
- [13] 谢德金, 周成城, 杨柯, 等. 巴戟天 MoDXS 基因及其启动子的克隆与分析 [J]. 园艺学报, 2021, 48(3): 577-589.
- [14] 周成城, 谢德金, 叶友杰, 等. 巴戟天 MoTPS 基因及其启动子克隆与分析 [J]. 分子植物育种, 2021, 19(9): 2889-2898.
- [15] 周成城, 谢德金, 万志庭, 等. 巴戟天苯丙烷类代谢途径相关基因的克隆与表达 [J]. 应用与环境生物学报, 2021, 27(1): 158-166.
- [16] Stalman M, Koskamp A M, Luderer R, et al. Regulation of anthraquinone biosynthesis in cell cultures of *Morinda citrifolia* [J]. *J Plant Physiol*, 2003, 160(6): 607-614.
- [17] 姚辉, 吴鸿, 冯承浩, 等. 巴戟天根的结构及其与蒽醌类化合物积累的关系 [J]. 实验生物学报, 2004, 37(2): 96-102.
- [18] 林美珍, 陈伟绩. 巴戟天的生物学特性和繁殖方式对其品质的影响 [J]. 亚热带农业研究, 2009, 5(1): 22-25.
- [19] 廖海, 周嘉裕. 高等植物蒽醌生物合成途径相关酶及其基因研究进展 [J]. 安徽农业科学, 2009, 37(24): 11388-11391.
- [20] 彭亮, 张岗, 颜永刚, 等. 茜草转录组分析及其蒽醌类化合物合成相关基因的挖掘 [J]. 中草药, 2018, 49(8): 1890-1896.
- [21] 郝大程, 马培, 穆军, 等. 中药植物虎杖根的高通量转录组测序及转录组特性分析 [J]. 中国科学: 生命科学, 2012, 42(5): 398-412.
- [22] 黑小斌, 李欢, 李依民, 等. 药用大黄幼苗转录组高通量测序及蒽醌类生物合成基因筛选 [J]. 中国药学杂志, 2019, 54(7): 526-535.
- [23] Wu D, Austin R S, Zhou S J, et al. The root transcriptome for North American ginseng assembled and profiled across seasonal development [J]. *BMC Genomics*, 2013, 14: 564.
- [24] Hao D C, Ge G B, Xiao P G, et al. The first insight into the

- tissue specific taxus transcriptome via Illumina second generation sequencing [J]. *PLoS One*, 2011, 6(6): e21220.
- [25] Wang R, Ren C X, Dong S, *et al.* Integrated metabolomics and transcriptome analysis of flavonoid biosynthesis in safflower (*Carthamus tinctorius* L.) with different colors [J]. *Front Plant Sci*, 2021, 12: 712038.
- [26] Yi S Y, Lin Q W, Zhang X J, *et al.* Selection and validation of appropriate reference genes for quantitative RT-PCR analysis in *Rubia yunnanensis* Diels based on transcriptome data [J]. *Biomed Res Int*, 2020, 2020: 5824841.
- [27] Singh R S, Gara R K, Bhardwaj P K, *et al.* Expression of 3-hydroxy-3-methylglutaryl-CoA reductase, p-hydroxybenzoate-m-geranyltransferase and genes of phenylpropanoid pathway exhibits positive correlation with shikonins content in *Arnebia Arnebia euchroma* (Royle) Johnston [J]. *BMC Mol Biol*, 2010, 11(1): 88.
- [28] 周豪楠, 马喜玲, 涂志伟, 等. 基于转录组学的粉葛葛根素生物合成相关基因挖掘 [J]. *江苏农业科学*, 2023, 51(5): 30-37.
- [29] 王祥, 王育朋, 安佳, 等. 不同生长年限蒙古黄芪转录组学分析及三萜皂苷合成关键基因挖掘 [J]. *中草药*, 2023, 54(3): 915-925.
- [30] 李洋洋, 李卓彬, 俞继华, 等. 决明不同组织莽草酸激酶基因的表达模式与蒽醌含量分析 [J]. *生物学杂志*, 2016, 33(3): 30-33.
- [31] Quevedo C, Perassolo M, Alechine E, *et al.* Increasing anthraquinone production by overexpression of 1-deoxy-d-xylulose-5-phosphate synthase in transgenic cell suspension cultures of *Morinda citrifolia* [J]. *Biotechnol Lett*, 2010, 32(7): 997-1003.
- [32] 樊星, 胡高升, 李娜, 等. 洛伐他丁、氯胺和茉莉酸甲酯处理对茜草悬浮培养细胞中紫色素和大叶茜草素积累的作用 [J]. *中国天然药物*, 2013, 11(4): 396-400.

[责任编辑 时圣明]