

基于转录组数据发掘银柴胡甾醇类成分生物合成的关键基因

李 军¹, 张 美², 周国玺³, 刘立轩¹, 徐 超¹, 郭鸿雁¹

1. 宁夏职业技术大学 宁夏中药材开发与利用工程技术研究中心, 宁夏 银川 750021

2. 宁夏回族自治区药品检验研究院, 宁夏 银川 750021

3. 甘肃陇神戎发药业股份有限公司, 甘肃 兰州 730102

摘 要: **目的** 研究银柴胡 *Stellaria dichotoma* 不同组织部位甾醇类成分含量及基因转录水平差异, 挖掘银柴胡甾醇类成分生物合成途径关键酶基因。 **方法** 采用高效液相色谱法测定银柴胡根、茎、叶和花中甾醇类成分的含量, 结合全长转录组测序与比较转录组学分析, 通过加权基因共表达网络分析 (weighted gene coexpression network analysis, WGCNA) 挖掘甾醇类成分生物合成的关键酶基因, 并利用 RT-qPCR 验证其表达水平。 **结果** 银柴胡中甾醇类成分在根、茎、叶和花中含量差异显著。转录组测序组装获得 372 483 条高质量全长转录本序列, 其中有 128 646 条全长转录本得到注释。差异基因分析结果显示, 发现根、茎、叶、花 4 个组织共有的基因有 43 354 条, 占基因总数的 50.02%。根与花相比, 差异表达基因总数为 9 647 条, 根和叶相比, 差异表达基因为 12 143 条, 叶与茎相比, 差异表达基因仅有 388 条。并从中鉴定出 119 个甾醇类成分合成的关键酶基因。进一步通过 WGCNA 鉴定到 43 个基因共表达网络模块, 在 MEblue 模块中鉴定了 *NP_NY_transcript_168676(FPPS)*、*NP_NY_transcript_335811(SQS)*、*NP_NY_transcript_44001(SQS)*、*NP_NY_transcript_246835(CAS)*、*NP_NY_transcript_328932(CAS)* 和 *NP_NY_transcript_183565(GPPS)* 6 个关键酶基因。对上述关键酶基因表达情况进行 RT-qPCR 验证, 结果与转录组数据表达趋势一致。 **结论** 揭示了银柴胡甾醇类成分合成的关键酶基因, 为解析调控其有效成分生物合成途径及分子机制提供研究基础。

关键词: 银柴胡; 转录组; 甾醇类成分; 生物合成; 加权基因共表达网络分析

中图分类号: R282.12 文献标志码: A 文章编号: 0253-2670(2026)16-6532-11

DOI: 10.7501/j.issn.0253-2670.2026.16.025

Discovery of key genes for biosynthesis of sterol biosynthesis of *Stellaria dichotoma* var. *lanceolata* based on transcriptome data

LI Jun¹, ZHANG Mei², ZHOU Guoxi³, LIU Lixuan¹, XU Chao¹, GUO Hongyan¹

1. Ningxia Traditional Chinese Medicine Development and Utilization Engineering Technology Research Center, Ningxia Vocational and Technical University, Yinchuan 750021, China

2. Ningxia Hui Autonomous Region Institute for Drug Control, Yinchuan 750021, China

3. Gansu Longshenrongfa Pharmaceutical Co., Ltd., Lanzhou 730102, China

Abstract: Objective To investigate the differences in sterol contents and gene transcription levels in different tissues of *Stellaria dichotoma* var. *lanceolata*, and explore key enzyme genes related to the biosynthetic pathway of sterols in this plant. **Methods** The contents of sterols in the roots, stems, leaves and flowers of *S. dichotoma* var. *lanceolata* was determined by high-performance liquid chromatography (HPLC). Combined with full-length transcriptome sequencing and comparative transcriptomic analysis, weighted gene coexpression network analysis (WGCNA) was used to mine key enzyme genes associated with sterol biosynthesis. The expression levels of these genes were further validated using RT-qPCR. **Results** The sterols contents in the roots, stems, leaves, and flowers of *S. dichotoma* var. *lanceolata* exhibited significant differences. A total of 372 483 high-quality full-length transcripts were obtained by transcriptome sequencing, of which 128 646 full-length transcripts were successfully annotated. Differential gene analysis revealed that 43 354 genes were expressed across the four tissues (roots, stems, leaves and flowers), accounting for 50.02% of all total genes. There were 9 647 differentially expressed genes (DEGs) between roots and flowers, 12 143 DEGs between roots and leaves, and only 388 DEGs between leaves and stems. A total of 119 key enzyme genes involved

收稿日期: 2026-02-03

基金项目: 宁夏回族自治区高等学校科学研究项目 (NYG2024299)

作者简介: 李 军 (1982—), 男, 山西大同人, 博士, 教授, 主要从事中药资源与开发利用研究。E-mail: lijunamy@126.com

in sterol biosynthesis were identified, and 43 gene co-expression modules were identified by WGCNA. In the MEblue module, six key enzyme genes involved in sterol biosynthesis were identified, including *NP_NY_transcript_168676 (FPPS)*, *NP_NY_transcript_335811(SQS)*, *NP_NY_transcript_44001(SQS)*, *NP_NY_transcript_246835 (CAS)*, *NP_NY_transcript_328932(CAS)* and *NP_NY_transcript_183565(GPPS)*. The expression levels of the above-mentioned key enzyme genes were validated by RT-qPCR, and the results showed expression trends consistent with those obtained from the transcriptome data. **Conclusion** This study revealed the key enzyme genes involved in sterol biosynthesis in *S. dichotoma* L. var. *lanceolata*, providing a research basis for analyzing the biosynthetic pathways and molecular mechanisms that regulate its active ingredients.

Key words: *Stellaria dichotoma* L. var. *lanceolata* Bunge; transcriptome; sterols; biosynthesis; WGCNA

银柴胡，别名银胡、山菜根、牛肚根、沙参儿、白根子等，为石竹科繁缕属植物银柴胡 *Stellaria dichotoma* L. var. *lanceolata* Bunge 的干燥根^[1]，主要分布在宁夏、内蒙古、陕西毗邻的干旱少雨沙生草原区，极耐干旱^[2]。银柴胡味甘性寒，具有清虚热、除疳热的功效，临床用于阴虚发热、骨蒸劳热、小儿疳热等症^[3]。其作为《中国药典》2025 年版收载的传统中药，在临床上应用比较广泛，尤其是作为主要成分的传统中成药乌鸡白凤丸，在临床上享有很好的声誉^[4]。银柴胡化学成分主要包括甾醇类、环肽类、生物碱类、挥发性物质等，其中甾醇类化合物为银柴胡主要活性成分，其次为生物碱类物质、环肽类化合物。而甾醇类成分中 α -菠甾醇和豆甾-7-烯醇含量较高，具有抗炎和解热作用^[5-6]。另外，有研究表明银柴胡还具有抗过敏、扩张血管、抗癌等药理作用^[7-8]。鉴于甾醇类物质在银柴胡及中医药中的重要作用，银柴胡中甾醇类物质的生物合成及调控研究，对甾醇类物质的高效生产具有非常重要的意义。

在不同的药用植物中，甾醇生物合成途径中的关键酶及调控机制都存在一定的差异，这些差异使得不同植物甾醇的组成和含量呈现出多样性，从而影响药用植物的生长和中药材的药效作用^[9]。但植物甾醇的生物合成过程均起始于三羧酸循环中的产物乙酰辅酶 A。该路径从乙酰辅酶 A 开始，经过经典的甲羟戊酸 (mevalonic acid, MVA) 通路，再合成异戊烯基焦磷酸 (isopentenyl pyrophosphate, IPP)，经异构化，缩合生成牻牛儿基焦磷酸 (geranyl pyrophosphate, GPP)。之后 IPP 与法尼基焦磷酸合成酶 (farnesyl pyrophosphate synthase, FPPS) 缩合，得到法尼基焦磷酸 (farnesyl pyrophosphate, FPP)。通过鲨烯合成酶 (squalene synthase, SQS) 缩合得到鲨烯，经鲨烯环氧酶 (squalene epoxidase, SQE) 和环阿屯醇合酶 (cycloartenol synthase, CAS) 催化生成植物甾醇生物合成的环阿屯醇。其中 3-羟基-3-甲基戊二酰辅酶 A 还原酶 (3-hydroxy-3-

methylglutaryl-CoA reductase, HMGR) 是 MVA 途径的限速酶，FPPS、SQS、CAS 等为环阿屯醇合成的关键酶^[10]。环阿屯醇作为植物甾醇生物合成的重要前体，经过一系列复杂的生化反应，包括甲基移除、双键还原和异构化等步骤，逐步转化为多种不同的植物甾醇，从而产生不同的作用。

关于药用植物甾醇类成分生物合成途径及其调控研究较少，在蕨麻、铁皮石斛、雷公藤等中有关键酶基因相关报道^[11]。何琴等^[12]利用转录组学和代谢组学分析结合 RT-qPCR 发现低温处理蕨麻后，根部的知母皂苷元含量上调，CAS 基因表达增强，推测 2,3-环氧角鲨烯流向 CAS 通路，提高蕨麻中甾醇含量。林江波等^[13]对铁皮石斛茎和叶进行转录组测序，比较其甾醇生物合成关键酶基因的表达，发现 CAS 在长期和成熟期的表达量差异不大，而成熟期茎、叶甾醇甲基转移酶 (sterol methyltransferase, SMT)，SMT1 的表达量比生长期高，SMT2 则生长期高于成熟期；同一时期，SMT1 和 SMT2 在叶的表达量都比茎高。国内外关于银柴胡的研究主要集中在其基原植物生物学特性、种质、人工种植、化学成分、药理作用及临床应用方面的研究，在分子生物学方面的研究较少，仅有叶绿体基因组、DNA 条形码方面的报道。

目前银柴胡基因组信息及全长转录组数据库未见报道，制约了其分子生药学相关研究的发展，且关于银柴胡甾醇类成分生物合成途径仍知之甚少。本研究基于 PacBio 三代测序平台构建了银柴胡高质量全长转录本数据库，通过对银柴胡根、茎、叶、花 4 个不同组织的比较转录组学分析，结合加权基因共表达网络 (weighted gene coexpression network analysis, WGCNA) 分析挖掘银柴胡中甾醇类成分生物合成途径关键酶基因，并通过 RT-qPCR 对其表达水平进行验证，为揭示银柴胡甾醇类成分生物合成相关基因及其分子机制奠定基础，对于银柴胡调控其有效成分生物合成途径、优良性状选育与抗逆相关的遗传改良具有重要意义。

1 材料与仪器

1.1 材料

试验材料二年生银柴胡采集于宁夏固原市隆德县(106.1244° N, 35.5048° E)。经宁夏职业技术大学郭鸿雁教授鉴定为植物银柴胡 *S. dichotoma* L. var. *lanceolata* Bunge。于 2025 年 6 月随机选取 3 株银柴胡, 设置 3 个生物学重复, 将银柴胡茎、嫩叶、老叶、花、根采摘后立即清洗、除杂, 一部分自然晾干、粉碎过筛测定含量, 另一部分置于冻存管中, 标记后液氮速冻, 运回实验室置-80 °C 超低温冰箱保存备用。

1.2 试剂

对照品 α -菠甾醇(批号 DSTDB013101, 成都德思特生物科技有限公司)、豆甾-7-烯醇(批号 202412257, 上海博尔森生物科技有限公司)质量分数均大于 98%。流动相所用甲醇为色谱纯, 冰醋酸为色谱纯, 水为娃哈哈纯净水; 提取溶剂为纯化水, 甲醇为分析纯。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)购自北京天根生物有限公司, HiScript Q RT SuperMix for qPCR (+gDNA wiper) 和 2×ChamQ SYBR COLOR qPCR Master Mix 试剂盒购自南京诺维赞生物科技有限公司。

1.3 仪器

Thermo scientific Vanquish-高效液相色谱仪配备 UV、CAD 检测器(赛默飞世尔有限公司); Mettler Toledo XPE26、XS-205 型电子分析天平(瑞士 Mettler 有限公司); Milli-Q 型超纯水机(默克密立博有限公司); T100 PCR 仪(美国伯乐有限公司); qTOWER3G 型实时荧光定量 PCR 仪(德国耶拿有限公司)。

2 方法

2.1 不同组织中甾醇类成分含量测定

2.1.1 对照品溶液的制备 精密称定 α -菠甾醇、豆甾-7-烯醇对照品适量, 分别置于 10 mL 量瓶, 甲醇溶解稀释至刻度, 摇匀, 得到质量浓度为 134.65 $\mu\text{g/mL}$ α -菠甾醇和 104.28 $\mu\text{g/mL}$ 豆甾-7-烯醇对照品溶液。

2.1.2 供试品溶液的制备 银柴胡根、茎、叶、花 4 种组织样品自然干燥后, 分别研磨过四号筛, 精密称取样品粉末约 5.0 g, 置于 100 mL 圆底烧瓶中, 精密加入甲醇 100 mL, 摇匀, 称定质量, 加热回流 1 h, 放冷, 再称定质量, 分别用上述相应溶剂补足缺失的质量摇匀, 滤过, 滤液蒸干, 残渣加水 30 mL

使溶解, 分别用乙醚提取 3 次, 分别为 30、30、20 mL; 合并提取液, 蒸干, 残渣加甲醇使溶解, 定容至 5 mL, 摇匀, 滤过, 滤液用 0.45 μm 滤膜滤过, 即得。

2.1.3 色谱条件 SHIMADZU Shimpack GIST C₁₈ 色谱柱(250 mm×4.6 mm, 5 μm); 流动相为甲醇-水(98:2); 检测波长 210 nm, 柱温 30 °C, 体积流量 1.0 mL/min, 进样量 20 μL 。

2.1.4 线性关系考察 按“2.1.1”项下方法制备得 α -菠甾醇和豆甾-7-烯醇混合对照品溶液, 按照“2.1.3”色谱条件分别进样 1、2、3、5、7、10 μL 测定峰面积, 以峰面积为纵坐标(Y), 质量浓度为横坐标(X)进行线性回归, 得 α -菠甾醇回归方程: $Y=1.487X+0.1808$, $R^2=0.9994$, 豆甾-7-烯醇回归方程: $Y=6.2957X+0.1607$, $R^2=0.9993$ 。结果表明, α -菠甾醇在 0.1795~1.7950 $\mu\text{g/mL}$, 豆甾-7-烯醇在 0.1390~1.3900 $\mu\text{g/mL}$, 线性关系良好。以信噪比(S/N)为 3 作为检出限, 以 S/N 为 10 作为定量限, α -菠甾醇的检出限为 0.06 $\mu\text{g/mL}$, 定量限为 0.18 $\mu\text{g/mL}$ 豆甾-7-烯醇的检出限为 0.04 $\mu\text{g/mL}$, 定量限为 0.14 $\mu\text{g/mL}$ 。

2.1.5 精密度试验 取对照品溶液, 连续进样 6 次, 每次进样 20 μL , 测得 α -菠甾醇和豆甾-7-烯醇的峰面积, 计算峰面积 RSD 分别为 1.22% 和 0.84%, 结果表明仪器精密度良好。

2.1.6 重复性试验 取银柴胡样品供试品 6 份, 按照“2.1.2”项下方法制备供试品溶液, 进样 20 μL 测定, 测定并计算 α -菠甾醇和豆甾-7-烯醇质量分数, 并计算 RSD 分别为 1.53% 和 1.14%。结果表明, 该方法重复性良好。

2.1.7 稳定性试验 取银柴胡样品, 按照“2.1.2”项下方法制备供试品溶液, 在 0、2、4、6、8、10、12、16、18、24 h 分别进样 20 μL 测定, 并测定 α -菠甾醇和豆甾-7-烯醇峰面积, 计算其 RSD 分别为 1.82% 和 2.47%。结果表明供试品溶液在 24 h 内稳定性良好。

2.1.8 加样回收试验 取已测定含量的银柴胡样品精密称定 2.5 g 粉末, 平行 3 组, 每组 3 份, 按 1:0.8、1:1、1:1.2 的比例分别加入 α -菠甾醇和豆甾-7-烯醇混合对照品, 按“2.1.2”项下方法制备银柴胡供试品溶液, 进样含量测定。 α -菠甾醇和豆甾-7-烯醇平均加样回收率分别为 100.17% 和 93.44%, RSD 分别为 1.76% 和 2.60%。

2.2 RNA 提取与测序

用磁珠法高效多糖多酚植物总 RNA 提取试剂盒按照说明书分别提取银柴胡根、茎、嫩叶、老叶、花的总 RNA，利用 Nanodrop2000 对所提 RNA 的浓度和纯度进行检测，琼脂糖凝胶电泳检测 RNA 完整性，Agilent2100 测定 RIN 值。将上述样品混合后基于 PacBio Sequel II 三代测序平台进行全长转录组测序，银柴胡根、茎、老叶、花分别于 Illumina HiSeq 2000 二代测序平台进行转录组测序，设置 3 个生物学重复。所有测序均由上海美吉生物医药科技有限公司完成。

2.3 测序数据质控与组装

去除测序所得原始读段中低质量、含接头和未知碱基 N 含量的 Reads，获得高质量的读段（clean reads）。使用 Trinity v2.15.1 对 clean reads 进行拼接，然后使用序列聚类软件 Corset v1.09 进一步序列拼接和去冗余处理，得到 Unigenes。

2.4 基因功能注释和分类

利用 BLAST 软件（<https://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST/>）将基因分别与非冗余蛋白质序列数据库（non-redundant protein sequence database, NR）、基因本体数据库（gene ontology, GO）、基因进化谱系：非监督直系同源群（evolutionary genealogy of genes: Non-supervised orthologous groups, eggNOG）、京都基因与基因组百科全书（Kyoto encyclopedia of genes and genomes, KEGG）、蛋白序列数据库（swiss protein knowledgebase, Swiss-Prot）、蛋白质家族数据库（protein families database, pfam）等进行比对注释^[14]，获得基因的注释信息。

2.5 基因表达差异分析

利用 RSEM v1.3.3 软件分析各组织样品的基因表达水平，首先统计比对至参考序列对应基因的 reads 数（read count），并计算其 TPM 值（Transcripts Per Million），基于 TPM 对 read count 数据进行标准化处理，并使用 R 语言 4.4.1 DESeq2 筛选差异表达基因，筛选阈值为 $|\log_2(\text{fold change})| > 1$ 且 q 值 < 0.005 。

2.6 加权基因共表达网络及 GO 和 KEGG 分析

利用 WGCNA v1.72-5，基于银柴胡根、茎、叶和花 4 个不同组织的转录组数据以及 2 种代谢物的含量信息，构建了加权基因共表达网络。具体参数如下：软阈值设置为 12，最小模块大小设置为 30，合并阈值为类似的模块设置为 0.35。根据动态聚类方法将具有相同表达模式的基因分类为一个模块。利用 clusterProfiler v4.10.0 对差异表达基因（differentially expressed genes, DEGs）分别进行 GO 功能注释与 KEGG 通路富集分析。

2.7 RT-qPCR 验证 DEGs

利用 Primer Premier 5.0 软件设计的基因特异性引物进行 RT-qPCR 验证，由生工生物工程有限公司合成（表 1）。利用 HiScript Q RT SuperMix for qPCR (+gDNA wiper) 试剂盒将从银柴胡根、茎、叶、花等不同组织样品中提取的 RNA 反转录成 cDNA。采用 2×ChamQ SYBR COLOR qPCR Master Mix 试剂盒，在荧光定量 PCR 仪上进行 RT-qPCR 检测基因表达，反应条件为两步法：95 ℃、3 min，1 个循环；95 ℃、30 s，56 ℃、30 s，72 ℃、40 s×35 个循环；采用 18 S RNA 作为内参基因，每个样品 3 个重复，采用 $2^{-\Delta\Delta C_t}$ 法计算 DEGs 的相对表达量。

表 1 RT-qPCR 引物序列
Table 1 Primers for RT-qPCR

基因	引物名称	引物序列 (5'-3')
<i>NP_NY_transcript_168676</i> (FPPS)	<i>NP_NY_transcript_168676-1F</i>	AAGGTTGGCATGATTGCTAC
	<i>NP_NY_transcript_168676-1R</i>	CAGTCTGAAATTCAACCTCG
<i>NP_NY_transcript_335811</i> (SQS)	<i>NP_NY_transcript_335811-2F</i>	GGTATTTTGCCTTCATCTG
	<i>NP_NY_transcript_335811-2R</i>	ACACAGCCACAGCAAAGAAG
<i>NP_NY_transcript_44001</i> (SQS)	<i>NP_NY_transcript_44001-3F</i>	TGAGCTTTTACAACCAGGAG
	<i>NP_NY_transcript_44001-3R</i>	TAGAGGATATGCAAGTTTGG
<i>NP_NY_transcript_246835</i> (CAS)	<i>NP_NY_transcript_246835-4F</i>	TCTTCTTTACAGAATCCCG
	<i>NP_NY_transcript_246835-4R</i>	GAAGCAGAGGTACACTCAAC
<i>NP_NY_transcript_328932</i> (CAS)	<i>NP_NY_transcript_328932-5F</i>	CATGTGAATTCCTGTTGTCG
	<i>NP_NY_transcript_328932-5R</i>	CTAAATGAGCTGTATTGCC
<i>NP_NY_transcript_183565</i> (GPPS)	<i>NP_NY_transcript_183565-6F</i>	CTGCTTGTGCTCTTGAAATG
	<i>NP_NY_transcript_183565-6R</i>	CCAAGAGGAAAAAGTGCATC
<i>18s rRNA</i>	<i>18s rRNA-F</i>	GGGCTCGAAGACGATCAGAT
	<i>18s rRNA-R</i>	GTTTCAGCCTTGCGACCATA

3 结果与分析

3.1 银柴胡中甾醇类成分的含量测定

银柴胡不同组织部位根、茎、叶、花 HPLC 色谱分析结果如表 2 所示, α -菠甾醇在花中的含量最高, 叶中的含量最低, 花和根中的含量差异极显著, 根与茎差异显著, 根和叶差异不显著。豆甾-7-烯醇在花中的含量最高, 与根中的含量差异极显著, 而根与茎及根与叶中的含量基本相同。

表 2 银柴胡不同组织部位甾醇类成分含量

Table 2 Sterol contents in different tissues of *S. dichotoma* var. *lanceolata*

组织部位	α -菠甾醇/%	豆甾-7-烯醇/%
根	0.003 4 $\pm$ 0.000 3 ^a	0.010 3 $\pm$ 0.000 7 ^a
茎	0.005 2 $\pm$ 0.000 2 ^b	0.012 2 $\pm$ 0.001 3 ^a
叶	0.002 4 $\pm$ 0.000 1 ^a	0.009 7 $\pm$ 0.000 7 ^a
花	0.019 1 $\pm$ 0.000 6 ^c	0.025 1 $\pm$ 0.001 1 ^c

不同字母表示差异显著, $P < 0.05$ 。

Different letters mean significant differences, $P < 0.05$.

3.2 银柴胡全长转录组组装与质量分析

通过 PacBio Sequel II 测序平台对银柴胡的根、茎、叶、花及嫩叶 5 个组织的混和样品进行测序, 共获得 5 514 628 条环形一致性序列, 平均长度为 1 444 bp。将全长非嵌合序列(full-length non-concatemer, FLNC) 转录组数据经过去冗余纠错后共获得 372 483 条高质量全长转录本序列, 平均长度 1 527 bp, N50 长度 1 753 bp。采用 BUSCO v5 软件对全长转录本序列进行质量完整性评估, 其完整性达到 96.4%。为了进一步评估银柴胡转录本的质量, 将得到的银柴胡全长转录本序列作为参考序列, 把银柴胡各组织样品的 clean reads 在参考序列上做 mapping 分析, 各组织样品的比对率都在

77.31%~89.93%。综上表明, 本研究获得的银柴胡全长转录本序列质量较高、数据可靠, 可以满足后续关键酶基因的分析要求。

3.3 银柴胡基因功能注释和功能分类

将银柴胡全长转录组与 6 大数据库比对, 获得注释信息。结果表明, 共有 128 546 条转录本得到注释, 占全部转录本的 91.74% (表 3)。其中在 eggNOG 和 NR 数据库中注释到的转录本数量最多, 分别为 128 471 和 117 969 条转录本, 占比为 91.62% 和 84.13%, 而在 KEGG 中注释到的转录本数量最少, 共注释到 67 241 条转录本序列, 占比为 47.95%。根据 Nr 数据库的比对结果, 统计比对率最高的前 10 个物种, 其中与银柴胡同源性序列比较高的物种有太阳花 *Heliosperma pusillum* (Waldst. & Kit.) Rchb (69 352, 53.98%)、甜菜 *Beta vulgaris* L. (13 867, 10.79%)、藜麦 *Chenopodium quinoa* Willd. (11 495, 8.95%)、菠菜 *Spinacia oleracea* L. (7 709, 6%)、苋 *Amaranthus tricolor* L. (5 483, 4.27%), 如图 1 所示。

表 3 银柴胡转录本的功能注释结果

Table 3 Functional annotation results of transcripts in *S. dichotoma* var. *lanceolata*

注释数据库	被注释的转录本数	占比/%
GO	107 150	76.41
KEGG	67 241	47.95
NR	128 471	91.62
Eggnog	117 969	84.13
Swiss-Prot	107 642	76.76
Pfam	104 214	74.32
总的被注释转录本数	128 546	91.74
总的转录本数	140 226	100.00

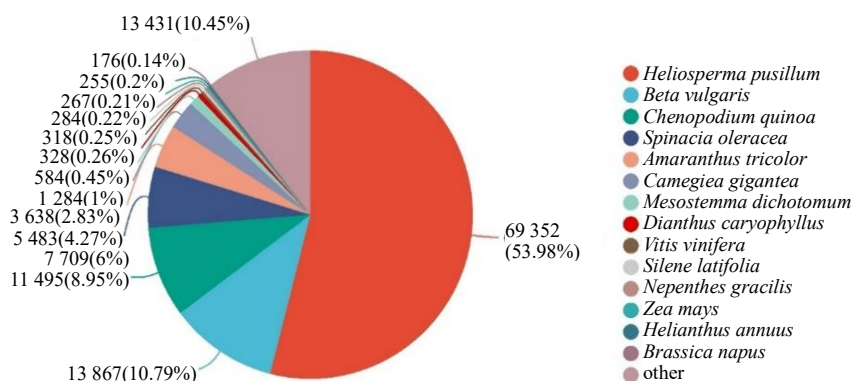


图 1 NR 注释统计物种分布图

Fig. 1 NR annotated statistical species distribution map

3.4 DEGs 分析

根据银柴胡不同组织部位的转录组数据, 绘制不同组织部位基因的韦恩图 (图 2-A), 发现根、茎、叶、花 4 个组织共有的基因有 43 354 条, 占基因总数的 50.02%。不同组织间的 DEGs 如图 2-B 所示, 叶 vs 茎存在 388 条 DEGs, 上调 174 条, 下调 214 条; 根 vs 茎存在 8 800 条 DEGs, 上调 2 840 条, 下调 5 960 条;

根 vs 花存在 9 647 条 DEGs, 上调 3 427 条, 下调 6 220 条; 叶 vs 花存在 3 221 条 DEGs, 上调 1 788 条, 下调 1 433 条; 茎 vs 花存在 1 838 条 DEGs, 上调 894 条, 下调 944 条; 叶 vs 根存在 12 143 条 DEGs, 上调 7 959 条, 下调 4 184 条。所有 DEGs 中叶和根的 DEGs 最多, 根与茎和根与花的次之, 叶和茎之间的 DEGs 最少。

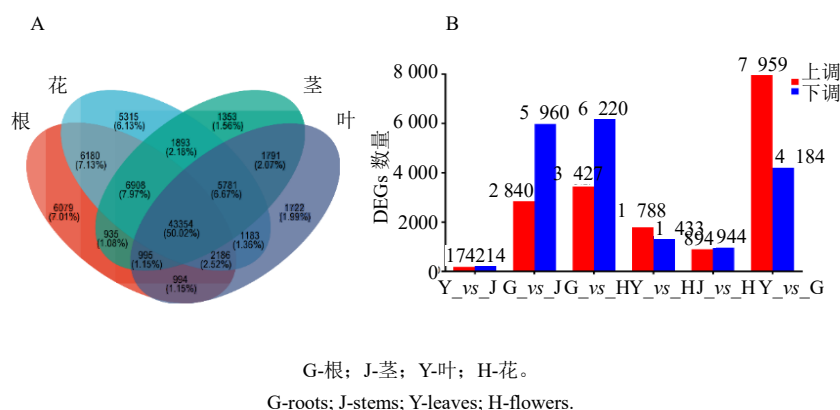


图 2 银柴胡不同组织部位中基因韦恩图 (A) 及差异表达分析 (B)

Fig. 2 Venn diagram (A) and differential expression analysis (B) of genes in different tissues of *S. dichotoma* var. *lanceolata*

3.5 银柴胡甾醇类生物合成相关酶基因表达分析

由于甾醇类化合物的合成起始于一条由乙酰辅酶 A 到环阿屯醇的代谢途径, 该途径共涉及 6 个关键酶, 分别是 HMGR、牻牛儿基焦磷酸合酶 (geranyl pyrophosphate synthase, GPPS)、FPPS、SQS、SQE 和 CAS。本研究基于银柴胡转录组数据库的基因注释结果, 从中鉴定出 119 个上述关键酶基因, 涉及 HMGR (30 个)、GPPS (14 个)、FPPS (14 个)、SQS (23 个)、SQE (10 个) 和 CAS (28 个) (图 3)。值得注意的是在这些关键酶基因中 HMGR 的转录本数量最多, 其作为 MVA 途径中的关键限速酶, 其活性不仅对植物甾醇的生物合成具有决定性作用, 还在植物正常生长发育进程中发挥着核心调控功能。

为了进一步分析上述关键酶基因在银柴胡不同组织表达水平, 对这些基因进行了组织特异性指数 (Tau) 计算。分析结果显示, 119 个关键酶基因中共有 19 个组织特异性基因 (Tau>0.8), 其中 9 个基因在花中高表达, 分别为 8 个 GPPS (NP_NY_transcript_132189、NP_NY_transcript_147927、NP_NY_transcript_195133、NP_NY_transcript_195975、NP_NY_transcript_241513、

NP_NY_transcript_311237、NP_NY_transcript_72354、NP_NY_transcript_125513)、1 个 SQS (NP_NY_transcript_220193); 5 个基因在根中高表达, 分别为 2 个 HMGR (NP_NY_transcript_31349、NP_NY_transcript_77176)、3 个 SQS (NP_NY_transcript_95705、NP_NY_transcript_135146 和 NP_NY_transcript_161510); 4 个基因在叶中高表达, 分别为 2 个 HMGR 基因 (NP_NY_transcript_160396、NP_NY_transcript_66779)、2 个 CAS 基因 (NP_NY_transcript_145417、NP_NY_transcript_311885); 1 个基因在茎中进行高表达, 1 个 GPPS 基因 (NP_NY_transcript_346261)。

3.6 WGCNA

为了深入挖掘在银柴胡甾醇类生物合成途径相关的关键酶基因, 通过 WGCNA 方法构建银柴胡 2 个代谢物即 α -菠甾醇和豆甾-7-烯醇的加权基因共表达网络, 按照动态剪切树法的标准化划分模块, 共得到 43 个基因共表达网络模块, 每个模块用不同的颜色代表 (图 4-A)。

根据加权基因共表达网络分析得出 MEblue 模块中的基因与 α -菠甾醇和豆甾-7-烯

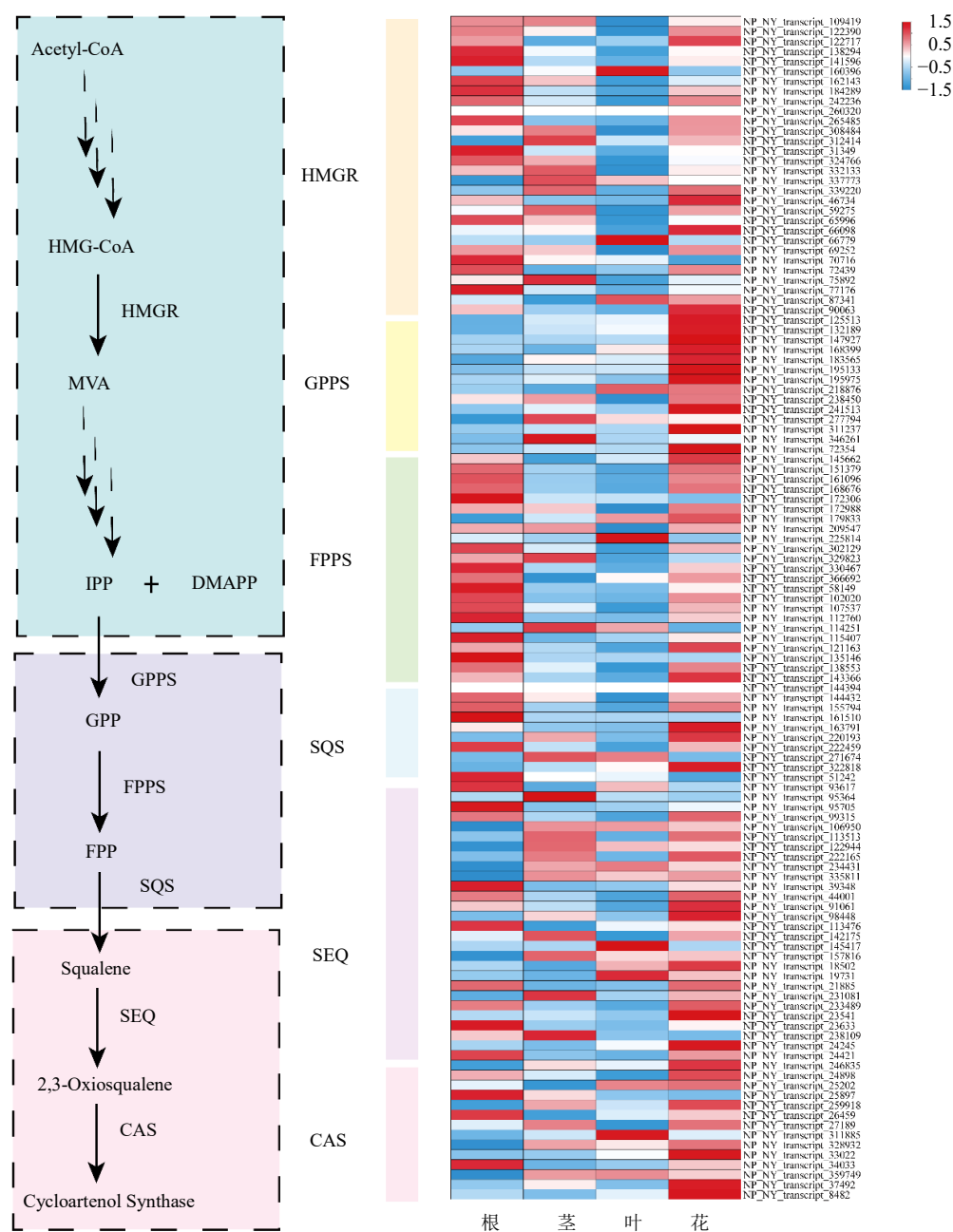


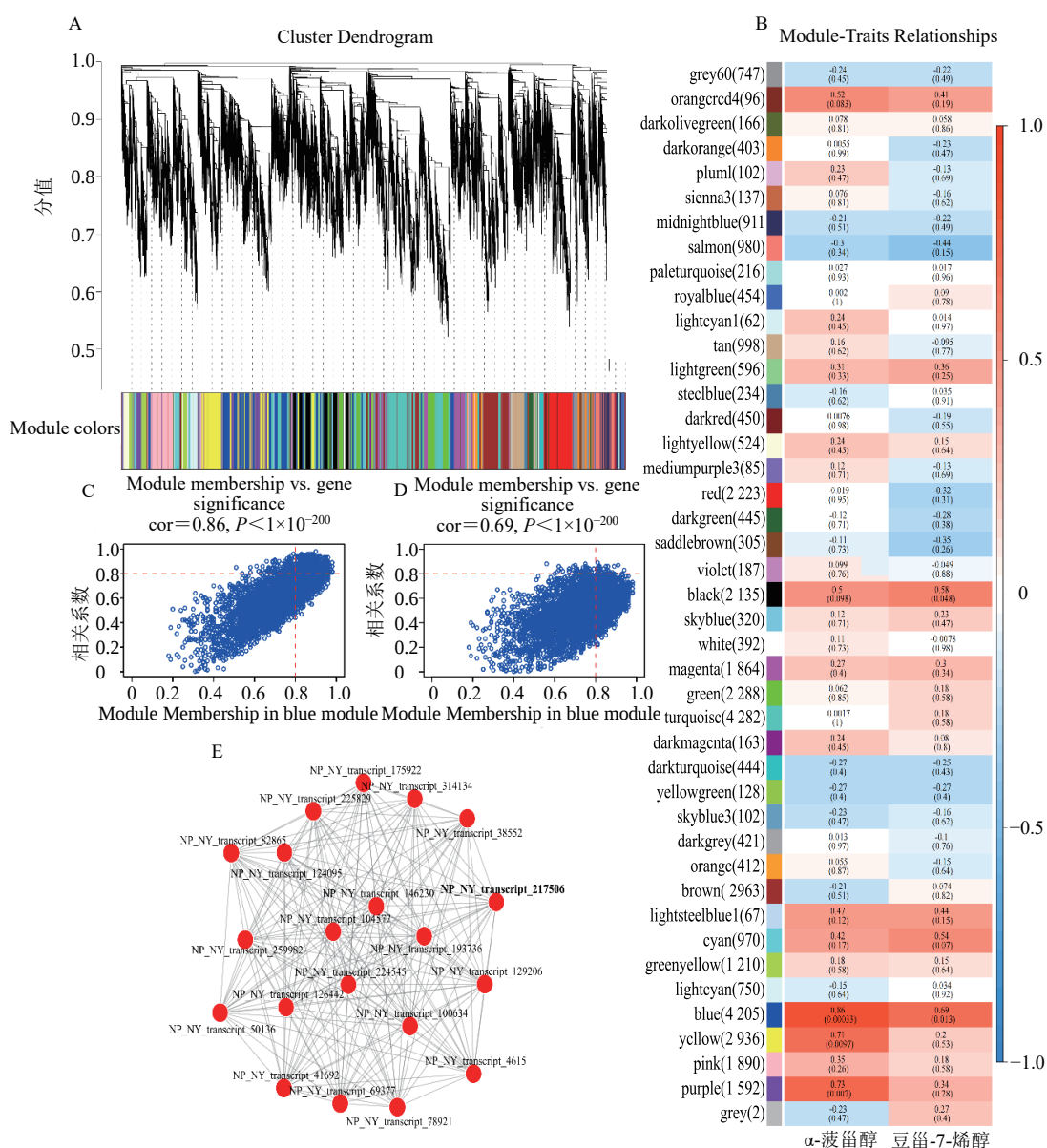
图 3 银柴胡甾醇类成分生物合成途径及其关键酶表达量热图

Fig. 3 Biosynthetic pathway of sterols and heatmap of key enzyme expression in *S. dichotoma* var. *lanceolata*

醇含量高度相关,相关性分别是 0.86 和 0.69(图 4-B~E)。对 MEblue 模块中的基因进行 GO 和 KEGG 富集分析。GO 富集结果显示,这些基因的生物过程主要富集在电子传递链和光系统 II 中的光和电子传递,值得注意的是有部分基因富集与丙酮酸来源的乙酰辅酶 A 生物合成途径、异戊二烯类生物合成途径以及麦角甾醇生物合成途径,这些都与甾醇类生物合成途径有

关(图 5-A)。KEGG 富集结果显示,有部分基因富集在萜类骨架生物合成途径上,这是合成所有萜类化合物(包括甾醇)的核心上游通路,可生成重要的前体分子,如 IPP 和二甲烯丙基焦磷酸(dimethylallyl diphosphate, DMAPP),这些是合成角鲨烯(甾醇的直接前体)的基石(图 5-B)。

MEblue 模块有 4 205 个基因,其中在甾醇类生



A-基因层次聚类树；B-基因模块与不同成分含量的关联热图；C-基因模块与 α -波甾醇含量的相关性；D-基因模块及和豆甾-7-烯醇的相关性；E-MEblue 模块的 Hub 基因共表达网络图。

A-gene hierarchical clustering tree; B-association heat map of gene module and contents of different components; C-correlation between gene modules and α -spinasterol content; D-correlation between gene modules and stigmasterol content; E-gene co-expression network diagram of Hub genes in MEblue module.

图 4 银柴胡不同组织部位 DEGs 的 WGCNA

Fig. 4 WGCNA of DEGs in different tissues of *S. dichotoma* var. *lanceolata*

物合成途径中鉴定到以下基因：NP_NY_transcript_168676(FPPS)、NP_NY_transcript_335811(SQS)、NP_NY_transcript_44001(SQS)、NP_NY_transcript_246835(CAS)、NP_NY_transcript_328932(CAS) 和 NP_NY_transcript_183565(GPPS)，这与转录组分析中获得的高表达甾醇类生物合成关键酶基因基本一致。表明

这 6 个基因可能为银柴胡甾醇类物质生物合成关键酶基因。

3.7 RT-qPCR 验证

为验证转录组数据的准确性，根据 WGCNA 分析结果，采用 RT-qPCR 对银柴胡甾醇类物质生物合成关键酶基因 NP_NY_transcript_168676 (FPPS)、

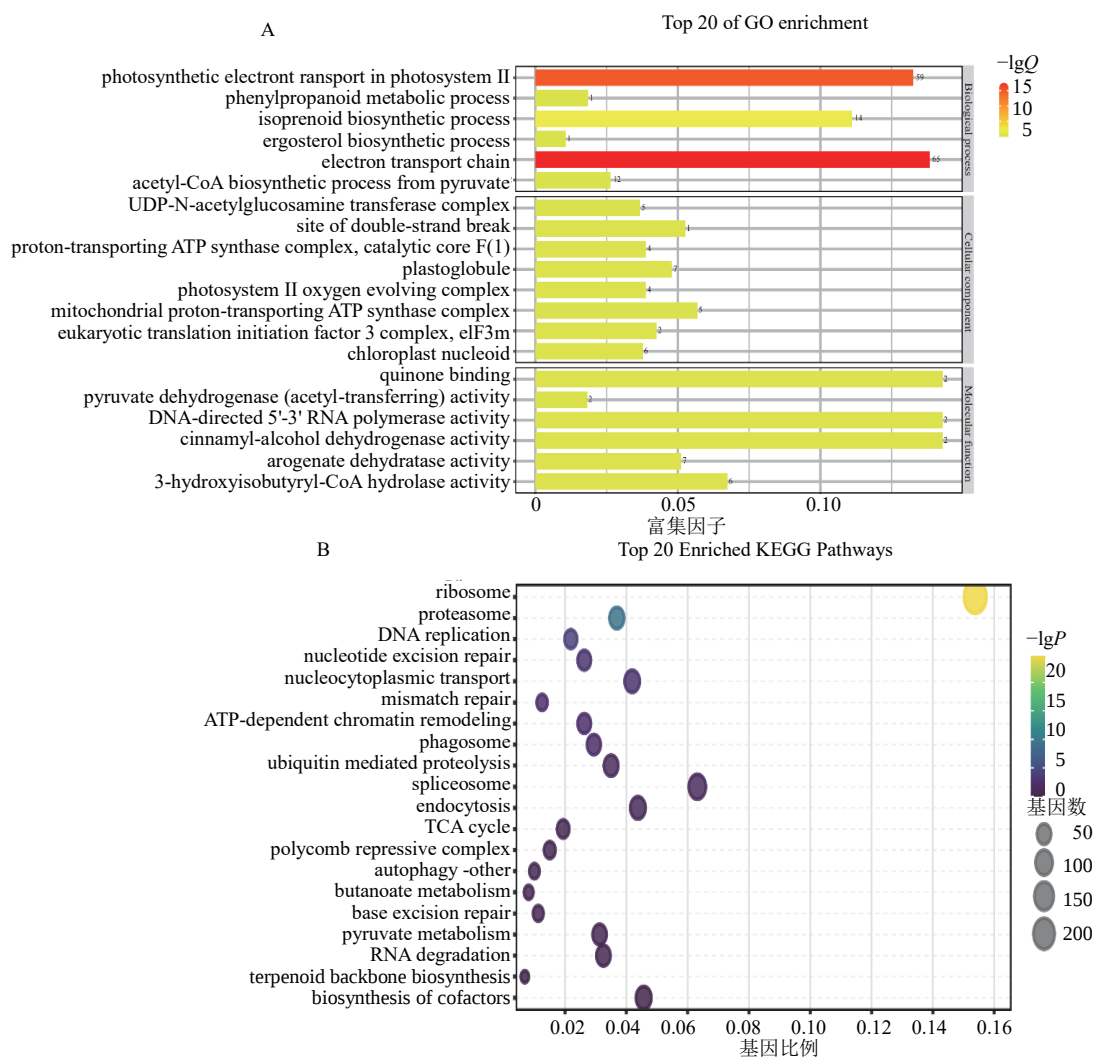


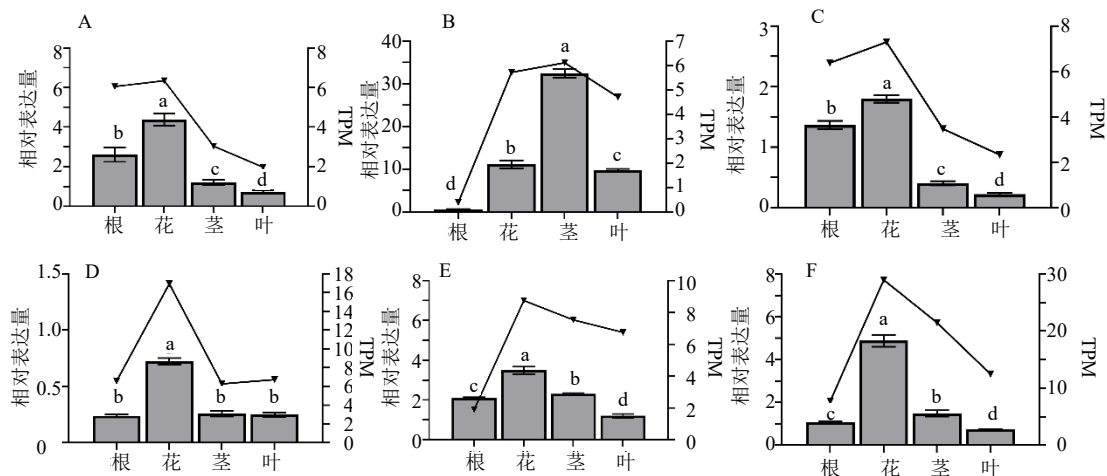
图5 银柴胡不同组织中差异表达基因 GO 富集分析 (A) 及 KEGG 通路富集散点图 (B)

Fig. 5 GO enrichment analysis (A) and KEGG pathway enrichment scatter plot (B) of differentially expressed genes in different tissues of *S. dichotoma* var. *lanceolata*

NP_NY_transcript_335811(SQS) 、 NP_NY_transcript_44001(SQS)、NP_NY_transcript_246835(CAS)、NP_NY_transcript_328932(CAS)和NP_NY_transcript_183565(GPPS) 的表达水平进行验证(图6)。结果显示,6个DEGs的RT-qPCR测定结果与转录组数据中TPM值的表达模式趋势保持一致,验证了转录组分析结果的可靠性。其中5个甾醇类合成关键酶基因 NP_NY_transcript_168676(FPPS)、NP_NY_transcript_44001(SQS)、NP_NY_transcript_246835(CAS)、NP_NY_transcript_328932(CAS)和NP_NY_transcript_183565(GPPS),均在花中表达量最高,而NP_NY_transcript_335811(SQS)在茎检测到较高的表达水平。

4 讨论

本研究首次利用高效液相色谱法测定了银柴胡不同组织部位(根、茎、叶和花)中 α -菠甾醇和豆甾-7-烯醇的含量,且根中 α -菠甾醇和豆甾-7-烯醇含量与先前研究报道的含量基本一致^[15-16]。另外,银柴胡中 α -菠甾醇和豆甾-7-烯醇在花中的含量高于其他部位,分别达到了0.0191%和0.0251%,根和花具有极显著差异,根与茎中的差异不明显。可能是由于银柴胡甾醇类成分的主要合成部位在花,后期累积部位在根部。有文献^[17]报道,次生代谢产物在植物不同组织器官中的积累往往存在器官异质性,这与植物的生长发育、环境适应及防御机制密切相关。本研究结果与叶芬等^[18]报道的一致,三七黄酮类化合物主要在根,其结果显示三七中黄酮



A-NP_NY_transcript_168676; B-NP_NY_transcript_335811; C-NP_NY_transcript_44001; D-NP_NY_transcript_246835; E-NP_NY_transcript_328932; F-NP_NY_transcript_183565; 不同小写字母表示组间差异显著 ($P < 0.05$)。

A-NP_NY_transcript_168676; B-NP_NY_transcript_335811; C-NP_NY_transcript_44001; D-NP_NY_transcript_246835; E-NP_NY_transcript_328932; F-NP_NY_transcript_183565; Different lowercase letters indicate significant difference between groups ($P < 0.05$).

图 6 银柴胡甾醇类成分生物合成途径关键酶基因 RT-qPCR 验证

Fig. 6 RT-qPCR verification of key enzyme genes in sterols synthesis pathway of *S. dichotoma* var. *lanceolata*

类成分在茎叶中含量较根高。

目前半夏、菝葜、甘草、决明子、小柴胡和淫羊藿等一大批我国传统药用植物都通过全长转录组测序对其有效成分合成关键基因和药用机制进行了分析与鉴定，部分已应用于实际^[19-23]。本研究获得了银柴胡的全长转录本，丰富了银柴胡的基因组数据。另外，不同组织的差异表达基因，发现叶和根的 DEGs 最多，根与花和根与茎的次之。基于银柴胡转录组数据库的基因注释结果，鉴定出 119 个上述关键酶基因，均为 MVA 途径中的关键限速酶，与之前报道一致^[24-25]。另外，基于上述关键酶基因的 Tau 指数筛选出 19 个组织特异性基因。由于已通过 RT-qPCR 对关键酶基因进行了验证，表明了转录组数据的可靠性。因此，基于 Tau 指数筛选出的组织特异性基因表达模式在后续银柴胡生物合成调控及分子机制中进一步验证。

WGCNA 技术可以更高效地整合大量的转录组数据，将全部的基因聚类为具有共表达特征的模块，继而深入挖掘这些模块与目的性状间的潜在关系^[26]。如张力岚等^[27]基于 WGCNA 筛选出 16 个茶树糖苷类香气前体含量性状相关的候选基因。Song 等^[28]运用 WGCNA 筛选出油茶树中参与开花时间调控的关键核心基因 *CoBMY7/8*、*CoTPP6/10* 和 *CoG8H7/11*。本研究利用银柴胡不同组织部位的转录组学数据和甾醇类成分含量通过 WGCNA 获得

NP_NY_transcript_168676(FPPS) 、 *NP_NY_transcript_335811*(SQS)等 6 个甾醇类成分生物合成途径关键酶基因，经 RT-qPCR 检测结果与转录组测序所得 TPM 趋势一致。这些关键酶基因与甾醇类成分的前体环阿屯醇合成密切相关^[29]。药用植物次生代谢合成途径关键催化酶基因鉴定和功能研究是解析药效成分生物合成机制的关键^[30]。曾梁秦^[31]利用转录组差异表达分析和 WGCNA 筛选出与大理天门冬甾体皂苷合成相关的候选基因，并通过分子对接模拟和功能验证阐明了下游关键酶的功能。谢彩侠等^[32]筛选出几个可能参与甾体皂苷合成的关键催化酶基因 *HMGR*、*GPPS*、*SQS* 等，为进一步解析甾体皂苷生物合成的分子机制奠定基础。另有报道从蕨麻全基因组鉴定得到 5 个蕨麻甾醇生物合成的 CAS 关键基因^[12]。而本研究的关键酶基因与上述报道中甾醇类成分合成关键酶基因类似，均有 CAS、GPPS 和 SQS 基因。表明这 6 个关键酶基因在银柴胡甾醇类化合物的生物合成中发挥重要作用，但各关键酶基因对银柴胡甾醇类成分的具体调控作用及机制尚需进一步深入研究。

综上，本研究首次利用 PacBio Sequel II 三代测序平台构建了银柴胡的三代全长转录组数据库，通过测定银柴胡根、茎、叶、花 4 个不同组织部位进行甾醇类成分的含量及比较转录组和 WGCNA 银柴胡甾醇类成分生物合成途径关键酶基因在不同

组织中的表达情况,挖掘了6个参与甾醇类成分生物合成通路的关键酶基因,为进一步研究银柴胡甾醇类成分生物合成的调控机制及后续银柴胡的分子育种和品种改良奠定了基础。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 李军, 张燃, 于淑萍, 等. 中药材银柴胡及其混伪品的DNA条形码鉴定研究 [J]. 药学报, 2019, 54(5): 937-943.
- [2] 刘王锁, 魏佳仪, 董文飞, 等. 宁夏野生银柴胡资源分布现状调查 [J]. 中药材, 2024, 47(9): 2186-2191.
- [3] 李振凯. 宁夏银柴胡道地性形成相关因子研究 [D]. 银川: 宁夏大学, 2023.
- [4] 中国药典 [S]. 一部. 2025: 293, 330.
- [5] 于凯强, 焦连魁, 任树勇, 等. 中药银柴胡的研究进展 [J]. 中国现代中药, 2015, 17(11): 1223-1229.
- [6] 宁娱, 李振凯, 王银洁, 等. 基于网络药理学和分子对接技术的银柴胡解热作用机制探讨 [J]. 中国现代中药, 2023, 25(5): 1056-1063.
- [7] Morita H, Iizuka T, Choo C Y, *et al.* Dichotomins J and K, vasodilator cyclic peptides from *Stellaria dichotoma* [J]. *J Nat Prod*, 2005, 68(11): 1686-1688.
- [8] 宋乐. 中药银柴胡超临界萃取工艺、萃取物鉴定及功效作用研究[D]. 银川: 宁夏大学, 2022.
- [9] 陈广辉, 李春妮, 寇博闻, 等. 甾体类产物合成酶的研究进展 [J]. 中国科学: 生命科学, 2025, 55(10): 2061-2079.
- [10] 赵曼, 孔珂昕, 薛露, 等. 植物甾醇及其衍生物的生物合成途径研究进展 [J]. 中草药, 2022, 53(18): 5884-5898.
- [11] 林江波, 王伟英, 邹晖, 等. 植物甾醇生物合成与关键酶基因研究进展 [J]. 农业科学, 2018, 10(3): 5-7.
- [12] 何琴, 陈秧, 田婷婷, 等. 低温和干旱对蕨麻甾醇生物合成及 CAS 基因表达模式的影响 [J]. 西北农业学报, 2025, 34(9): 1694-1703.
- [13] 林江波, 王伟英, 邹晖, 等. 基于转录组测序的铁皮石斛植物甾醇生物合成相关基因分析 [J]. 热带亚热带植物学报, 2019, 27(6): 693-701.
- [14] 王富华, 曾子参, 李端阳, 等. 基于转录组数据发掘秃叶黄檗异喹啉类生物碱合成的相关基因 [J]. 中草药, 2025, 56(9): 3263-3273.
- [15] 王秀芬, 李静, 方光明, 等. 宁夏地产银柴胡药材的生长年限与质量的相关性研究 [J]. 时珍国医国药, 2021, 32(8): 1992-1995.
- [16] 杨丽, 汪晶, 王晓波, 等. 多指标定量联合 PCA、OPLS-DA 及 CRITIC-TOPSIS 评价不同产地银柴胡质量差异 [J]. 中医药信息, 2025, 42(10): 28-35.
- [17] 代熙. 不同发育期金银花花器官各部位的化学成分分布规律研究 [D]. 北京: 北京中医药大学, 2022.
- [18] 叶芬, 孙森杰, 罗俊聪, 等. HPLC 法测定三七不同部位中 4 种黄酮类成分的含量 [J]. 山东化工, 2026, 55(3): 167-171.
- [19] 徐京亚, 郭海林, 姚祥, 等. 结缕草全长转录组测序及生物信息学分析 [J/OL]. 分子植物育种, (2025-02-21) [2025-12-24]. <https://link.cnki.net/urlid/46.1068.S.20230901.1133.016>.
- [20] 韦筱媚, 廖振锋, 胡营, 等. 岗梅全长转录组测序及其三萜皂苷生物合成相关基因的挖掘 [J]. 中国现代中药, 2023, 25(7): 1515-1528.
- [21] 张晗. 遮荫调控半夏生长的功能基因挖掘与研究 [D]. 淮北: 淮北师范大学, 2020.
- [22] Soorni A, Akrami A M, Abolghasemi R, *et al.* Transcriptome and phytochemical analyses provide insights into the organic sulfur pathway in *Allium hirtifolium* [J]. *Sci Rep*, 2021, 11: 768.
- [23] 秦双双, 周鹏程, 胡营, 等. 菝葜全长转录组测序及分析 [J]. 中国现代中药, 2023, 25(7): 1451-1462.
- [24] 江舟. 铁皮石斛植物甾醇生物合成相关基因研究进展 [J]. 现代园艺, 2025, 48(20): 35-37.
- [25] 杨叙, 张珂, 李雅静, 等. 独一味的全长转录组及环烯醚萜苷合成途径解析 [J]. 上海中医药大学学报, 2025, 39(4): 17-24.
- [26] 张娟娟, 刘雪梅, 李美欢, 等. 基于 WGCNA 挖掘鸡血藤儿茶素和表儿茶素合成途径关键基因 [J]. 中草药, 2025, 56(24): 9124-9134.
- [27] 张力岚, 杨军, 王让剑. 基于 WGCNA 发掘茶树糖苷类香气前体含量性状相关的候选基因 [J]. 作物学报, 2026, 52(2): 494-513.
- [28] Song H Y, Duan Z H, Huo H Q, *et al.* A global overview of transcriptome dynamics during the late stage of flower bud development in *Camellia oleifera* [J]. *BMC Plant Biol*, 2025, 25(1): 247.
- [29] 杨贵紫, 耿应敏, 周兴乾, 等. 基于全长转录组学研究昆明杯冠藤 C₂₁ 甾体类生物合成相关基因的分析 [J]. 时珍国医国药, 2025, 36(21): 4088-4094.
- [30] 李金梅, 覃艳红, 黄鼎, 等. 基于全长转录组挖掘八角果实中木脂素生物合成关键基因 [J]. 中草药, 2024, 55(23): 8153-8161.
- [31] 曾梁秦. 大理天门冬甾体皂苷合成途径关键酶基因的功能研究 [D]. 昆明: 云南农业大学, 2025.
- [32] 谢彩侠, 李伟峰, 龚海燕, 等. 低磷胁迫下盾叶薯蓣中甾体皂苷生物合成关键酶基因的筛选及鉴定 [J]. 中国中药杂志, 2022, 47(10): 2623-2633.

[责任编辑 时圣明]