基于转录组挖掘重楼皂苷合成关键基因

田密密[1#], 徐 萍[1,2#], 章晓叶[1], 周君哲[1], 崔茂应[3*], 郑国伟[1*]

(1. 云南中医药大学中药学院, 云南省南药可持续利用研究重点实验室, 云南 昆明 650500; 2. 西藏大学生态环境学院, 西藏自治区 拉萨 850000; 3. 云南中医药大学信息学院, 云南 昆明 650500)

摘要: 为挖掘重楼皂苷合成下游通路中的关键基因, 本研究采用HPLC测定了2组滇重楼根茎中8种重楼皂苷含量, 进一步进行转录组测序分析。结果显示, 2组滇重楼根茎中8种重楼皂苷的含量存在显著差异, 其中高皂苷含量组 (HS) 与低皂苷含量组 (LS) 的总皂苷含量分别为24.81和5.99 mg·g^{-1}。此外, LS组中未检测到重楼皂苷II、III和V。结合皂苷含量分析, 2组滇重楼总皂苷含量差异显著的主要原因是薯蓣皂苷类皂苷在HS组中大量积累。转录组测序共获得了11 399个差异基因, HS组与LS组相比, 有5 708个基因表达上调, 5 691个为基因表达下调。研究成功筛选出7个参与重楼皂苷合成的候选基因, 并对其进行了蛋白结构预测。通过互作网络分析, 筛选出80个与候选基因的相关系数在0.95或以上的转录因子, 其中与2个或2个以上的候选基因有相关性的转录因子共有8个。本研究为预测重楼皂苷合成的相关基因、探索重楼皂苷生物合成途径及其积累规律提供了理论依据。

关键词: 重楼皂苷; 滇重楼; 转录组; 生物合成途径; 基因

中图分类号: R931 **文献标识码**: A **文章编号**: 0513-4870(2026)05-1852-09

Mining the transcriptome for key genes related to polyphyllin biosynthesis

TIAN Mi-mi[1#], XU Ping[1,2#], ZHANG Xiao-ye[1], ZHOU Jun-zhe[1], CUI Mao-ying[3*], ZHENG Guo-wei[1*]

(*1. Yunnan Key Laboratory of Southern Medicinal Utilization, College of Traditional Chinese Medicine, Yunnan University of Chinese Medicine, Kunming 650500, China; 2. School of Ecology and Environment, Xizang University, Lhasa 850000, China; 3. Institute of Information of Yunnan University of Chinese Medicine, Kunming 650500, China*)

Abstract: To identify the key genes involved in the synthesis of polyphyllin within downstream pathways, this study employed high-performance liquid chromatography (HPLC) to quantify the content of eight types of polyphyllins in two groups of *Paris polyphylla* var. *yunnanensis* rhizomes, followed by transcriptome sequencing analysis. The results revealed significant differences in the content of the eight types of polyphyllins between the two groups. Specifically, the total saponin content in the high saponin content group (HS) was 24.81 mg·g^{-1}, whereas that in the low saponin content group (LS) was 5.99 mg·g^{-1}. Notably, polyphyllin II, III, and V were undetectable in the LS group. The substantial difference in total saponin content between the two groups can be primarily attributed to the considerable accumulation of diosgenin-type saponins in the HS group. A total of 11 399 differentially expressed genes were identified through transcriptome sequencing, the comparison between HS and LS, with 5 708 genes exhibiting upregulation and 5 691 genes exhibiting downregulation. This study successfully screened seven candidate genes involved in the synthesis of polyphyllin and conducted protein structure prediction for these genes. Through

收稿日期: 2025-11-12; 修回日期: 2026-01-30.

基金项目: 国家自然科学基金资助项目 (82160718); 中医联合专项重点项目 (202101AZ070001-163); 云南省高层次人才项目 (YNWR QNBJ-2020-257, XDYC-QNRC-2022-0277); 云南省科技人才和平台计划 (202105AG070012); 云南道地药材繁育与利用导师团队项目.

#共同第一作者.

*通讯作者 E-mail: cmy@ynucm.edu.cn; gwzhengkm@163.com

DOI: 10.16438/j.0513-4870.2025-1342

interaction network analysis, 80 transcription factors with a correlation coefficient of 0.95 or higher with the candidate genes were identified, among which a total of 8 transcription factors were correlated with two or more candidate genes. The findings provide a theoretical foundation for predicting genes related to polyphyllin synthesis, exploring the biosynthetic pathway of polyphyllin, and understanding their accumulation patterns.

Key words: polyphyllin; *Paris polyphylla* var. *yunnanensis*; transcriptome; biosynthetic pathway; gene

滇重楼 *Paris polyphylla* Smith var. *yunnanensis* (Franch.) Hand.-Mazz. 为藜芦科 (Melanthiaceae) 重楼属 (*Paris* L.) 多年生草本植物, 具有清热解毒、消肿止痛、凉肝定惊的功效, 其主要成分为重楼皂苷。现代研究表明, 重楼皂苷I对高转移人乳头状浆液性卵巢癌细胞具有明显的体外抑制生长作用[1], 重楼皂苷II在抑制肿瘤血管方面有良好的活性[2], 重楼皂苷III可抑制结肠癌细胞HCT116的体外增殖并促进HCT116细胞凋亡[3]; 重楼皂苷I～III、V～VII对红色毛癣菌、絮状表皮癣菌、石膏小孢子菌以及耐氟康唑白色念珠菌均表现出显著抑制活性[4]; 重楼皂苷VI可以通过调节炎症和氧化应激反应改善纤维化[5], 重楼皂苷VII、重楼皂苷H以及纤细薯蓣皂苷对乳腺癌细胞MCF-7和MDA MB-231均有一定的细胞毒活性[6]。据统计, 国内滇重楼的使用量每年超过1 000 t, 约有80%的野生滇重楼资源已被利用[7], 导致野生滇重楼已濒临灭绝。重楼根茎现已成为80种专利药物的主要成分[8]。由于野生资源减少及栽培周期较长, 滇重楼在研究和药用领域的使用受到限制。因此, 通过合成生物学与酶催化技术获取重楼皂苷已成为当前研究热点。

重楼皂苷的上游通路已得到充分研究[9,10], 目前关于重楼皂苷生物合成的研究聚焦于参与糖基化过程的尿苷二磷酸葡萄糖醛酸转移酶 (UGT) 与细胞色素P450酶 (CYP)。PpCYP90G4与PpCYP94D108、PpCYP94D109或PpCYP72A616等酶形成的协同催化体系, 是胆固醇合成重楼皂苷前体薯蓣皂苷元的关键[11,12]。

在PpUGT73C1[13]或PpUGT91BP1[14]的继续催化下, 薯蓣皂苷元与偏诺皂苷元的C3-OH糖基化, 分别转化为延龄草苷与偏诺皂苷元-3-*O*-葡萄糖苷, 同时已有研究报道PpUGT80A40[4]、PpUGT80A49[15]也参与薯蓣皂苷元向延龄草苷的转化, PpUGT80A41[4]与PpUGT6[16]则可促进偏诺皂苷元合成偏诺皂苷元-3-*O*-葡萄糖苷。当前重楼皂苷II、III、V～VII的生物合成途径已明确, 然而薯蓣皂苷元的C-17如何发生*α*-羟基化形成偏诺皂苷元, 以及其他重楼皂苷的下游合成机制尚不清楚, 2′-*O*-鼠李糖基转移酶PpUGT73CE1[4]分别催化延龄草苷与偏诺皂苷元-3-*O*-葡萄糖苷转化为重楼皂苷V与重楼皂苷VI, 在PpUGT73YD1与PpUGT738A2[17]的调控下, 重楼皂苷VI的连续两次糖基化形成重楼皂苷VII, 在相同酶的催化下, 重楼皂苷V首先形成重楼皂苷III, 继而转化为重楼皂苷II, 另外, 促进重楼皂苷V合成重楼皂苷III的鼠李糖基转移酶有PpUGT73DY2[18]、PpUGT91T1和PpUGT91T2[19]。

合成重楼皂苷的关键基因与重楼皂苷含量的联系密切相关。转录组反映了基因的表达情况, 是连接基因表达与表型的重要纽带。同种植物不同个体细胞中的转录本存在差异, 这为植物的变异与进化提供了重要依据。本研究对采自云南省保山市与昭通市两地的滇重楼样本, 进行了8种重楼皂苷的含量测定与转录组分析。通过整合皂苷含量与转录组数据, 旨在筛选与重楼皂苷生物合成相关的候选基因。该研究对于揭示滇重楼甾体皂苷的生物合成与积累规律具有重要参考价值。

材料与方法

实验材料 重楼样品分别采自保山市腾冲县北海乡 (北纬25.29°, 东经98.48°, 海拔1 765.2 m) 和昭通市镇雄县木卓镇 (北纬27.68°, 东经104.83°, 海拔1 505.1 m), 每组3个生物学重复, 所采集的材料为种子萌发生长7年的滇重楼 (图1)。新鲜滇重楼根茎样品经过洗净、除去须根、切片后, 一部分样品放入液氮中迅速冷冻存入−80 ℃超低温冰箱中, 用于转录组测序; 剩余样品置于45 ℃烘箱烘干至恒重, 粉碎后过三号筛 (50目筛), 密封置于阴凉干燥处保存, 用于重楼皂苷含量检测。

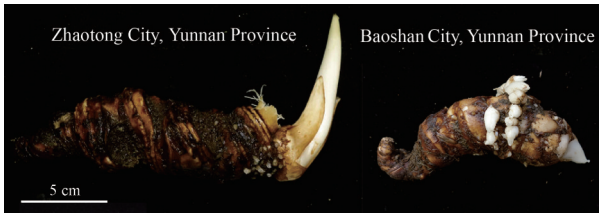


Figure 1 Phenotypes of 7-year-old *Paris polyphylla* var. *yunnanensis* in two groups

试剂与仪器 重楼皂苷I (AFCC1451, 纯度为98%)、重楼皂苷II (AFBG1350, 纯度为98%)、重楼皂苷III (AFCA0455, 纯度为98%)、重楼皂苷V (AF20050105, 纯度为98%)、重楼皂苷VI (AFBH2201, 纯度为98%)、

重楼皂苷VII (AF21121305, 纯度为98%)、重楼皂苷H (AFBJ0816, 纯度为98%)、纤细薯蓣皂苷 (AFCC0255, 纯度为98%) 均购于成都埃法生物科技有限公司; 乙腈 (色谱纯, 纯度 ≥ 99.7%) 购于云南罗德宝生物科技有限公司; 纯净水购于昆明娃哈哈启力饮料有限公司。

高效液相色谱仪 (Agilent 1200, 美国安捷伦公司)、电子天平 (PTY-224/323, 华志电子科技有限公司)、电热恒温水浴锅 (HWS-26, 上海一恒科学仪器有限公司)、超声波清洗仪 (CQ-400A, 上海跃进医用光学器械厂)。

对照品溶液的制备 分别称取重楼皂苷I～III、V～VII、H以及纤细薯蓣皂苷对照品适量, 色谱纯甲醇溶解并定容至2.0 mL, 依次制成各标准品溶液 (浓度分别为2.0、2.0、3.0、2.1、2.1、1.6、1.0、2.0 mg·mL^{-1})。

供试品溶液的制备 精密称取各重楼粉末样品0.5 g, 分别加入5.0 mL甲醇, 并记录各样品初始质量, 55 ℃条件下超声提取1 h后, 对各样品进行补重。5 000 r·min^{-1}的条件下离心2 min后, 经0.22 μm微孔滤膜过滤, 即得。

液相色谱条件 参考文献[20]方法测定, 色谱柱为Agilent SB-C18色谱柱 (250 mm × 4.6 mm, 5 μm), HPLC的分析系统为Agilent 1200系统, 流动相为水−乙腈 (60∶40), 流速为0.7 mL·min^{-1}, 进样量为10 μL, 柱温25 ℃。

转录组文库构建 按照试剂盒说明书步骤分别提取各样品总RNA, 并对提取的RNA进行完整性检测、RNA浓度测定。确保符合后要求后, 将Oligo(dT) 磁珠富集含有PloyA的mRNA反转录为cDNA, 对cDNA进行PCR扩增, 以获得足够的cDNA总量。修复cDNA连接SMRT哑铃头建转录组文库, 经核酸外切酶消化得完整文库。

数据处理及分析 使用CD-HIT对转录组数据进行去冗余与数据校正, 利用GO、KOG、与KEGG平台对数据进行注释与差异分析, 并使用iTAK结合hmmscan鉴定转录因子。为探究候选基因与各重楼皂苷的合成以及与转录因子间的关系, 利用迈维云平台 (https://cloud.metware.cn) 分析各重楼皂苷与对应候选基因的相关性, 使用 SWISS-MODEL (https://swissmodel. expasy. org/) 对候选基因进行蛋白同源建模, 再将筛选的候选基因与其相关性阈值达到0.95或以上且$P < 0.05$的转录因子进行交互网络分析。

结果

1 两组重楼样品中8种重楼皂苷的测定

HPLC检测结果显示, 采自保山市的滇重楼样品中, 8种皂苷的平均总含量为24.81 mg·g^{-1}, 是昭通样品 (5.99 mg·g^{-1}) 的4.1倍 (图2)。据此显著差异, 本研究将保山市的样品定义为高皂苷含量组 (high saponin, HS), 而将昭通市的样品定义为低皂苷含量组 (low saponin, LS)。其中重楼皂苷I、II、VII是中国药典规定的重楼质控成分[21], HS与LS中重楼皂苷I、II和VII的总含量分别为20.19、0.62 mg·g^{-1}。HS的重楼皂苷I、II和VII总含量比LS高19.57 mg·g^{-1}。HS中的重楼皂苷I～III和VII比LS分别高8.05、9.18、0.93、2.34 mg·g^{-1}, 重楼皂苷V、VI、H与纤细薯蓣皂苷在两组样品中的含量无显著差异。另外, LS中不含薯蓣皂苷类重楼皂苷II、III和V。结果表明, 薯蓣皂苷类重楼皂苷在HS中大量累积导致总皂苷含量高于LS。

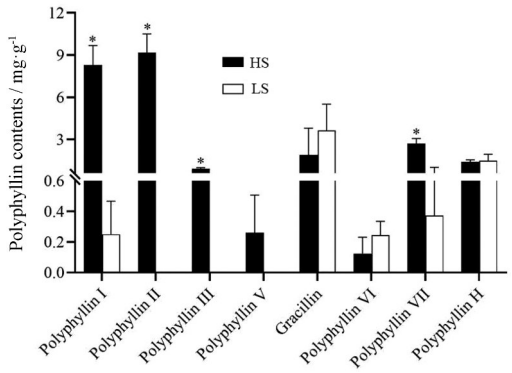


Figure 2 Comparison of the contents of 8 polyphyllins in the two groups of *Paris polyphylla* var. *yunnanensis*. $n = 6$, $\bar{x} \pm s$. $^*P < 0.05$ *vs* LS. HS: High saponin; LS: Low saponin

2 转录组测序质量评估

为进一步分析导致重楼根茎中重楼皂苷差异累积的分子机制以及筛选关键候选基因, 对两组样品进行转录组测序, 每组3个生物学重复, 共6个转录组文库。具体信息如表1所示, 6个重楼样品的原始测序读段 (raw reads) 数量分布在89 204 038～98 512 422条之间, 高质量测序读段 (clean reads) 数量分布在81 879 402～94 710 206条之间, 高质量碱基 (clean base) 在12.28～14.21 Gb之间, 且所有样品的Q30均在93.09%以上, 错误率均不超过0.03%, 说明所有样品得到了高质量的原始测序读段, 数据可靠, 满足分析要求。

3 主成分分析及高层次聚类分析

PCA分析结果显示, PC1和PC2分别解释了43.11%与16.86%的样品变异, 合计解释度为59.97% (图3A)。HS与LS在PC1与PC2方向上呈明显分离趋势, 表明两组样品存在显著差异。对两个组别的样品进行高层聚类分析 (图3B), 发现各基因在两组样品的表达差异明显, 表明两组样品可能存在明显差异。

Table 1 Sequencing data statistics of two groups of *Paris polyphylla* Smith var. *yunnanensis*

Name	Raw reads	Clean reads	Clean base/Gb	Error rate/%	Q20/%	Q30/%	GC content/%
HS-1	97 163 032	94 710 206	14.21	0.03	97.67	93.32	51.12
HS-2	96 736 860	94 180 970	14.13	0.03	97.57	93.09	50.70
HS-3	96 282 134	94 001 216	14.10	0.03	97.76	93.49	51.42
LS-1	98 512 422	89 978 722	13.50	0.03	97.75	93.61	50.58
LS-2	89 204 038	81 879 402	12.28	0.03	97.77	93.52	49.88
LS-3	90 356 546	84 722 642	12.71	0.03	97.80	93.66	49.79

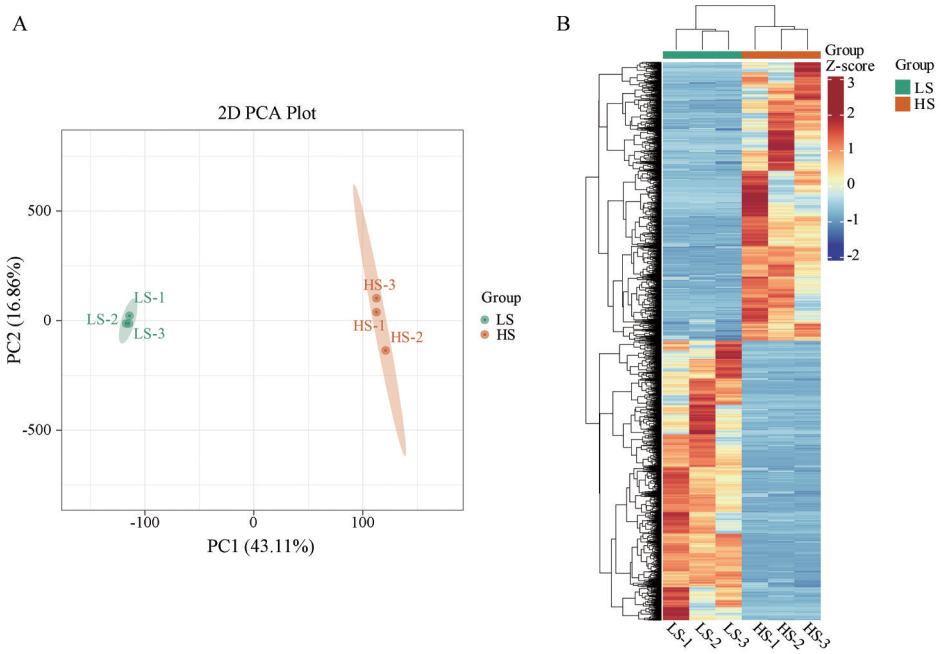


Figure 3 The PCA plot (A) and high-throughput heatmap (B) of HS *vs* LS

4 转录本注释和功能归类

通过差异基因分析, HS与LS比较, 共有11 399个差异基因 (DEGs), 其中5 708个上调, 5 691个下调 (图4A)。为了验证这些差异基因的功能, 分别对这些基因进行了GO、KOG、KEGG富集分析。GO富集分析结果显示 (图4B), 3 963个单基因在GO数据库得到了注释, 一共涉及43项, 图中列出32项, 生物学过程方面, DEGs主要聚集于“cellular process”, 共7 539个单基因; 细胞组分方面, “cellular anatomical entity”富集的DEGs最多; 分子功能方面, 富集最多基因的项为“binding”, 共6 821个单基因, 0.035%的单基因富集在“biological process involved in intraspecies interaction between organisms”, 仅4个转录本。

KOG富集分析结果显示 (图4C), 共有12 007条unigenes得到注释, “仅一般功能预测 (R)”项富集了1 723个滇重楼转录本, 数量最多, 其次为“信号转导机制 (T)”, 转录本数量为902个, 最少为“细胞运动 (N)”, 仅8个转录本。

经KEGG富集分析, 14 921个被注释的高质量转录本被归类为139条通路中 (图4D), 涉及代谢、遗传信息处理、有机系统、环境信息处理和细胞过程等5个领域, 其中富集明显的通路有“次生代谢物的生物合成”和“淀粉和蔗糖代谢”。KEGG富集结果显示, 与重楼皂苷生物合成密切的通路有萜类骨架生物合成途径 (ko00900) 和甾体生物合成途径 (ko00100), 进一步分析, 发现63条unigenes参与萜类骨架生物合成途径, 55条unigenes参与甾体生物合成途径。

5 重楼皂苷下游合成途径中相关基因的挖掘

基于重楼皂苷下游合成通路中相关酶的文献[4,11-13]报道, 对比本地转录组数据库, 本研究共鉴定出26个CYP基因和20个UGT基因, 包括9条*PpCYP90G4*、2条*PpCYP94D108*、7条*PpCYP94D109*、8条*PpCYP72A616*、15条*PpUGT73C1*和5条*PpUGT80A40*。另外, 两组样品的表达量水平结果显示*PacBio_transcript_47741*、*PacBio_transcript_47840*、*PacBio_transcript_47945*、*PacBio_transcript_47736*、*PacBio_transcript_48072*、*PacBio_transcript_48081*、*PacBio_transcript_48462*、*PacBio_transcript_48495*、*PacBio_transcript_45627* 均

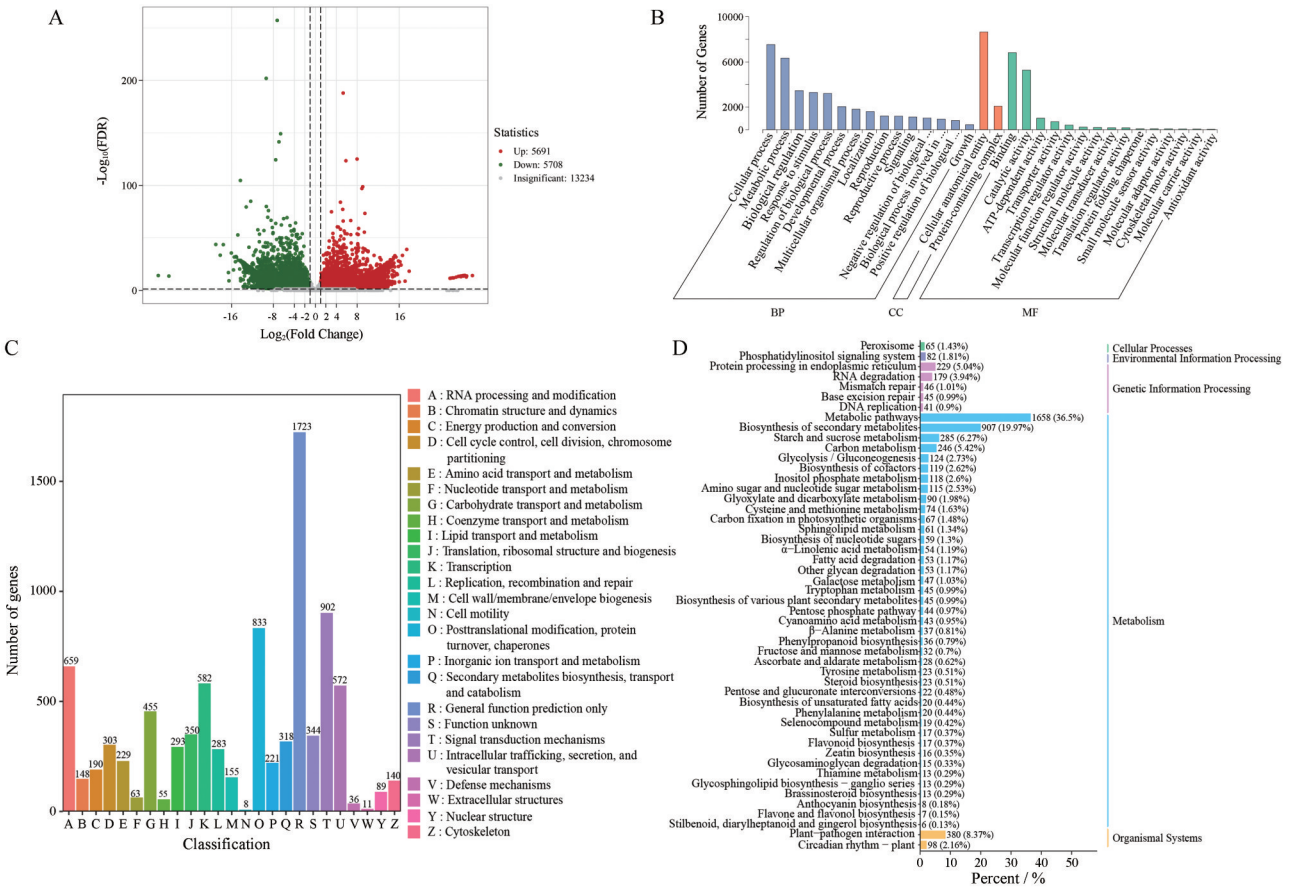


Figure 4 The volcano plot (A), GO annotation plot (B), KOG annotation plot (C) and KEGG annotation plot (D) of HS *vs* LS

在HS中上调, 在LS中下调 (图5A)。

通过对8种重楼皂苷含量与候选基因进行相关性分析, 保留相关系数绝对值大于0.8的结果, 共得到27个候选基因, 与重楼皂苷呈现显著正相关的基因共有14个 (图5B)。多数候选基因与重楼皂苷I、II、III和VII呈现显著相关性, 均与之呈现显著正相关的基因共有7个。与重楼皂苷I～III、V～VII和纤细薯蓣皂苷分别呈现显著正相关的基因依次为8、8、9、3、2、9和1个。

PpCYP90G4是调控胆固醇向16,22-羟基胆固醇转化的重要酶, PpCYP94D108、PpCYP94D109和PpCYP72A616参与催化16,22-羟基胆固醇合成薯蓣皂苷元 (图5C), 本研究中1条 *PpCYP90G4* (*PacBio_transcript_47247*) 与重楼皂苷VI呈现显著正相关, 但 *PacBio_transcript_47247* 和其余4条 *PpCYP90G4* (*PacBio_transcript_46319*、*PacBio_transcript_46952*、*PacBio_transcript_47018*、*PacBio_transcript_47605*) 均与重楼皂苷I、II、III、VII呈现负相关。1条 *PpCYP94D108* (*PacBio_transcript_38108*) 与纤细薯蓣皂苷呈现显著正相关; 1条 *PpCYP94D109* (*PacBio_transcript_47741*) 与重楼皂苷I、II、III、V均呈现显著正相关, 2条 *PpCYP94D109* (*PacBio_transcript_47840*、*PacBio_transcript_47945*) 与重楼皂苷I、II、III、VII均呈现显著正相关。1条 *PpCYP72A616* (*PacBio_transcript_19726*) 与重楼皂苷V的含量呈现显著正相关; 1条 *PpCYP72A616* (*PacBio_transcript_47674*) 与重楼皂苷VI呈现显著正相关。PpUGT73C1参与合成偏诺皂苷元-3-*O*-葡萄糖苷与延龄草苷 (图5C), 本研究共鉴定了9条与重楼皂苷有显著相关性的 *PpUGT73C1*, 其中5条 (*PacBio_transcript_47736*、*PacBio_transcript_48072*、*PacBio_transcript_48081*、*PacBio_transcript_48462*、*PacBio_transcript_48495*) 与重楼皂苷I、II、III、VII呈现显著正相关; 2条 *PpUGT80A40* 与重楼皂苷合成具有相关性的基因中, 1条 (*PacBio_transcript_45627*) 与重楼皂苷VII呈现显著正相关, 另1条 (*PacBio_transcript_41795*) 与重楼皂苷III、VII均呈现显著正相关 (图5B)。

6 蛋白结构预测

为了提高候选基因的功能可信度与筛选效率, 减少非功能基因或低可信度序列的干扰, 本研究排除上述与重楼皂苷含量呈显著正相关但开放阅读框 (ORF) 长度不足1 000 bp的基因后, 最终获得7个基因的蛋白质结构预测图 (图6)。这些基因被确定为重楼皂苷合成的关键候选基因。

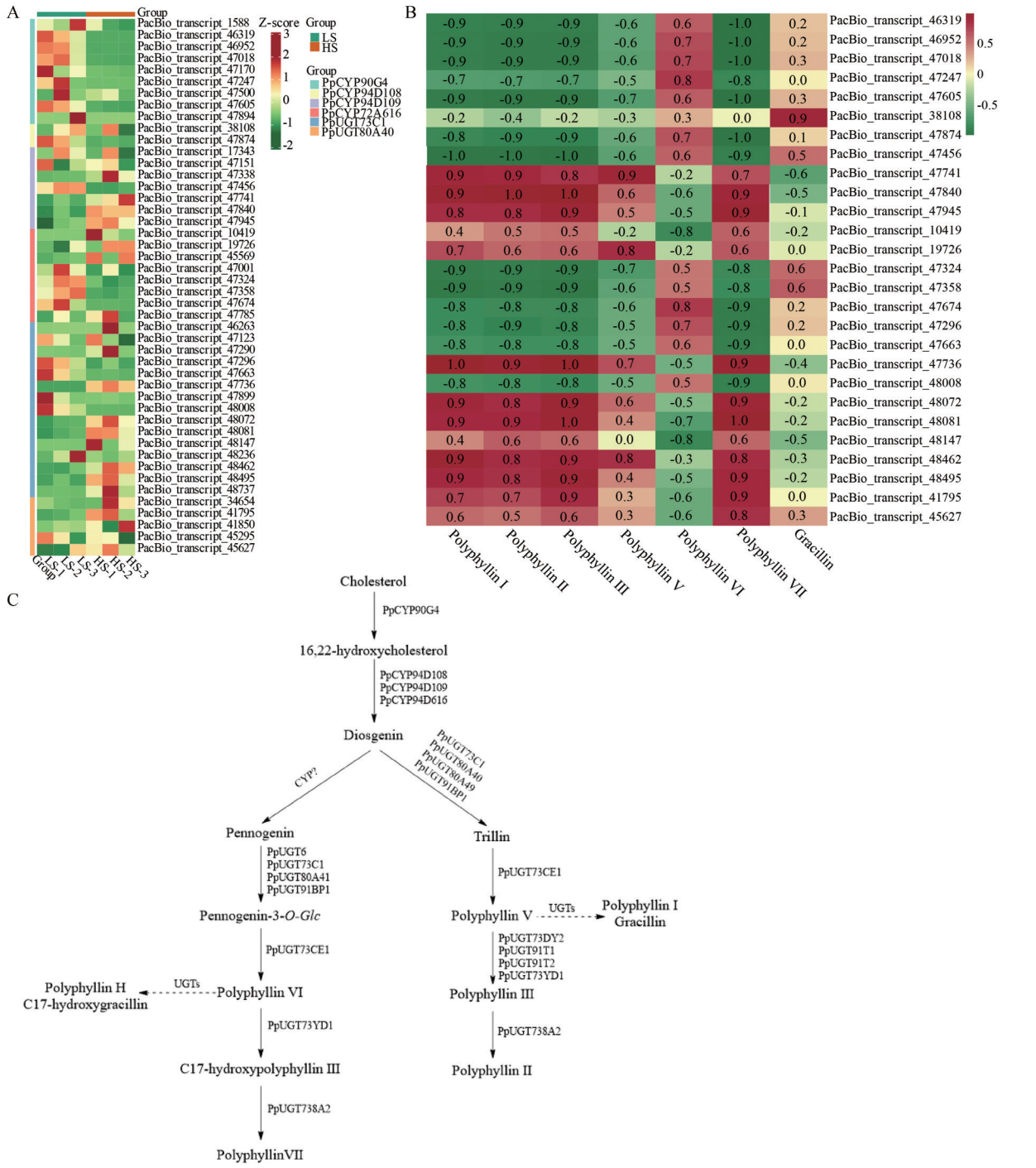


Figure 5 Expression pattern of genes related to polyphyllin biosynthesis in two groups of Paris (A), correlation analysis of the contents of polyphyllins and the expression of related genes in the downstream synthesis pathway (B) and the downstream biosynthetic pathway of polyphyllins (C)

7 转录因子的鉴定及与候选基因的互作网络分析

为了挖掘滇重楼中可能参与甾体皂苷合成调控的转录因子, 使用了转录因子预测生物信息学方法。HS与LS比较, 共鉴定了737个转录因子 (图7A), 其中SNF2占比较大, 一共85个, 其次是FAR1 (51个)、C3H (43个)、PHD (42个)、SET (38个)、B3-ARF (32个)、GRAS (28个)、MYB-related (26个) 和Jumonji (24个)。另外, 共鉴定到14个MYB、10个AP2/ERF、10个bHLH、6个WRKY和9个bZIP。与上述7个基因呈正相关且相关阈值在0.95或以上, $P < 0.05$的转录因子共有80个, 分别来自40个家族, 其中2个转录因子未被归属。与2个及以上的候选基因呈正相关且相关阈值在0.95或

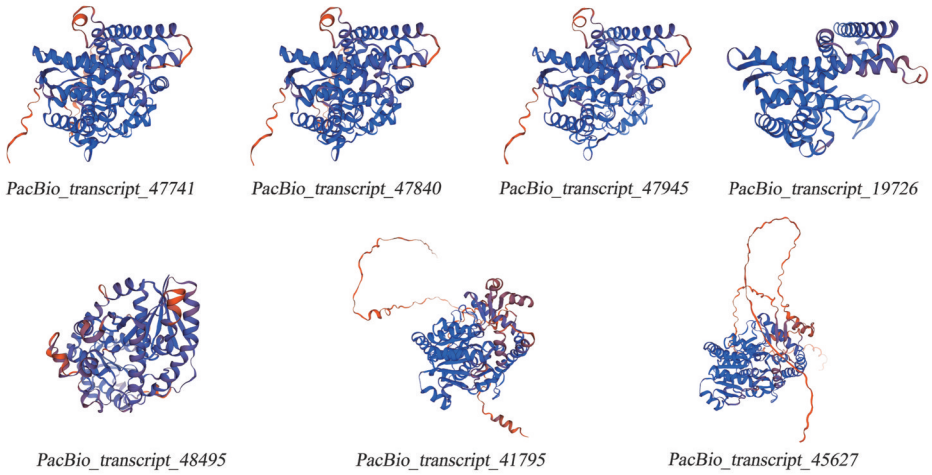


Figure 6 Protein structure predictions of 7 genes

以上的转录因子共有 8 个 (图 7B), 分别来自 GRAS、FAR1、SNF2、bZIP、AP2/ERF-ERF、TUB 和 Pseudo ARR-B 家族, 其中 1 个转录因子未被归属。此外, 与候选基因 *PacBio_transcript_45627* 呈正相关且相关阈值在 0.95 或以上, $P < 0.05$ 的转录因子仅有 1 个 (*PacBio_transcript_34220*), 来自 LUG 家族。

讨论

根据苷元 C-17 位是否含有氧取代基, 重楼皂苷可分为薯蓣皂苷类和偏诺皂苷类。其中, 重楼皂苷VI、VII、H 属于偏诺皂苷类, 而重楼皂苷I～III、V和纤细薯蓣皂苷属于薯蓣皂苷类[22]。通过对 HS 与 LS 中 8 种重楼皂苷含量的测定, 结果显示两组样本中偏诺皂苷类皂苷含量分别为 4.23 和 2.03 mg·g^{-1}, 差异不显著, 而 HS 与 LS 中薯蓣皂苷类皂苷含量分别为 20.58 和 3.90 mg·g^{-1}, 分别占总皂苷含量的 83% 和 65%。这表明两组总皂苷含量显著差异的主要源于薯蓣皂苷类皂苷在 HS 中大量累积, 其中重楼皂苷I～III的含量差异尤为突出。

由于滇重楼基因组的复杂性与资源匮乏[23], 探索重楼皂苷的生物合成途径面临诸多困难, 转录组测序成为研究重楼皂苷生物合成途径的重要方法。挖掘重楼皂苷下游合成通路中负责前体皂苷元糖基化的基因, 仍是该研究的重点。近年来重楼基因挖掘的报道中, Wang 等[24]筛选出一个参与重楼皂苷生物合成的基因 *PpUGT2*, Zhang 等[25]发现 2 种共生菌能显著促进重楼皂苷的合成, 通过将内生菌接种到重楼根部, 后期运用转录组分析筛选出 17 个促进重楼皂苷合成的 *CYPs* 和 2 个 *UGTs*, 曲别阿香等[26]筛选到 17 条 unigenes 参与薯蓣皂苷元的合成, 其中 2 条 *CYP90G4*、1 条 *CYP94D108*、1 条 *CYP94D109* 和 1 条 *CYP72A616*; Xu 等[27]研究中筛选到 4 条 *PpCYP90G4*、1 条 *PpCYP94D108*、6 条 *PpCYP94D109* 和 10 条 *PpUGT73C1*。本研究筛选出 3 条 *PpCYP94D109* (*PacBio_transcript_47741*、*PacBio_transcript_47840*、*PacBio_transcript_47945*) 和 5 条 *PpUGT73C1* (*PacBio_transcript_47736*、*PacBio_transcript_48072*、*PacBio_transcript_47081*、*PacBio_transcript_48462*、*PacBio_transcript_48495*), 这些基因在 HS 中均呈现上调趋势, 且与重楼皂苷I～III、V、VII呈正相关, 表明这 8 条转录本可能参与重楼皂苷I～III、V、VII的生物合成, 另外, 1 条 *PpCYP94D108* (*PacBio_transcript_38108*) 可能参与了纤细薯蓣皂苷的合成; *PacBio_transcript_47247* (*PpCYP90G4*)、*PacBio_transcript_47674* (*PpCYP72A616*) 可能参与重楼皂苷VI的合成; *PacBio_transcript_19726* (*PpCYP72A616*) 可能参与重楼皂苷V的合成; 1 条 *PpUGT80A40* (*PacBio_transcript_45627*) 可能参与重楼皂苷VII的合成, 另 1 条 *PpUGT80A40* (*PacBio_transcript_41795*) 可能参与重楼皂苷III、VII的合成。通过表达量分析和相关性分析, 并排除 ORF 较短的基因后, 对候选基因进行了蛋白结构预测, 最终筛选出 7 条关键候选基因。结果表明, 这 7 条候选基因与重楼皂苷的生物合成高度相关。

转录因子在调节植物生理生化过程中发挥着重要作用, 本研究鉴定的转录因子大多数来自 SNF2、FAR1、C3H、PHD、B3-ARF、GRAS 与 MYB-related 等家族。已有研究表明 AP2/ERF、bHLH、MYB、WRKY 和 bZIP 等转录因子可能参与甾体皂苷的合成[28-30], HD-ZIP I亚家族的转录因子 DZHDZ32 可能是调控薯蓣皂苷元生物合成的重要潜在因子[31], bZIP 家族转录因子 PpTGA3 在重楼中表达后, 重楼皂苷II和重楼皂苷VI的含量显著增加[32]。本研究共鉴定到了 14 个 MYB、10 个 AP2/ERF、10 个 bHLH、6 个 WRKY 和 9 个 bZIP, 7 个候选基因

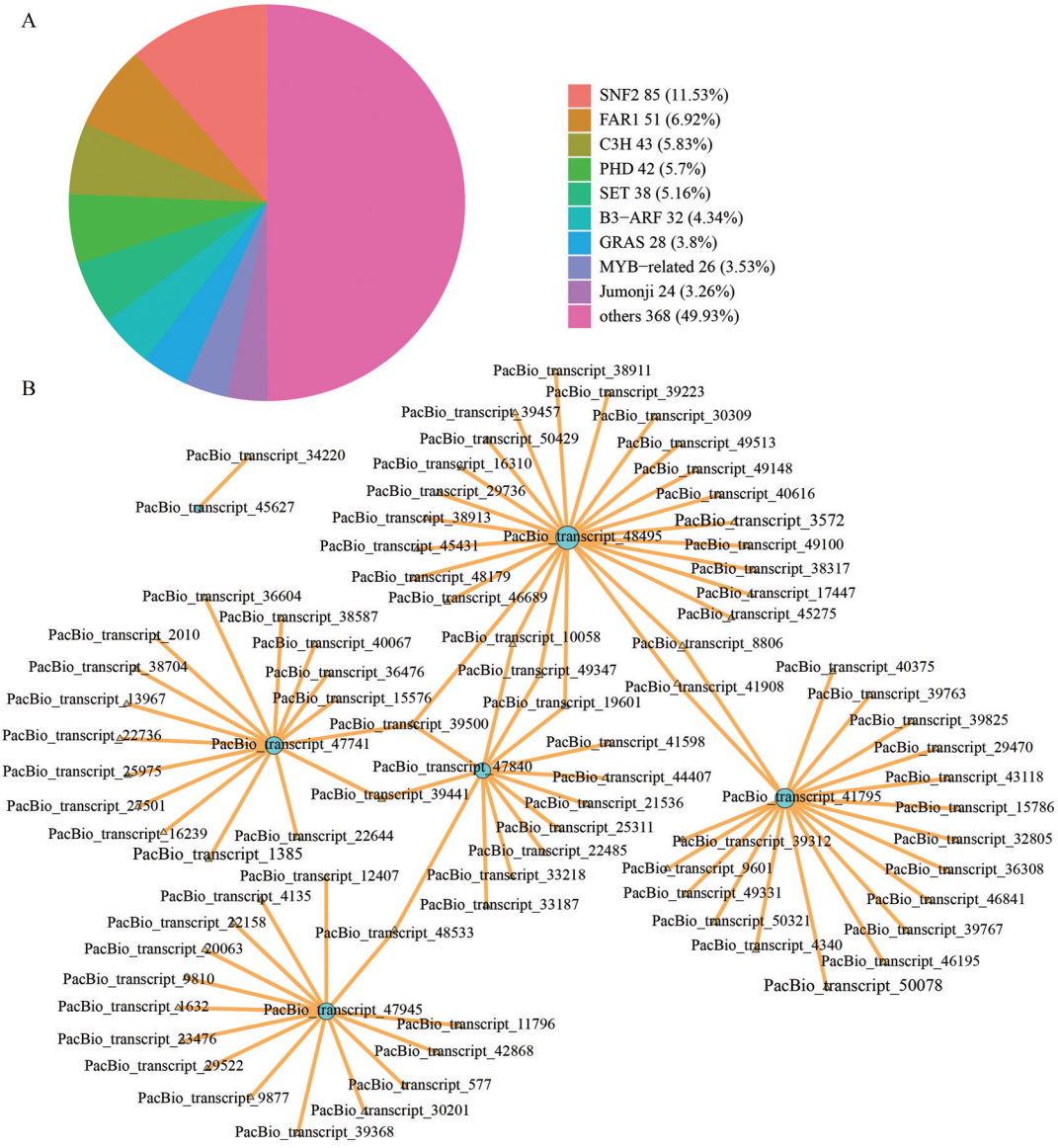


Figure 7 Distribution of transcription factor types (A) and correlation analysis of the key genes of downstream synthesis of polyphyllins with transcription factors (B)

与转录因子的互作网络分析结果显示, 80个转录因子与候选基因呈现正相关关系且相关系数在0.95或以上 ($P < 0.05$), 8个转录因子与2个及以上的候选基因都呈现显著正相关 (相关阈值 ≥ 0.95, $P < 0.05$), 这8个转录因子极可能参与了重楼皂苷的合成, 排除未被归类的转录因子, 分别来自7个家族, 其中1个属于bZIP家族、1个属于AP2/ERF-ERF。

作者贡献: 田密密与徐萍为共同第一作者, 负责文献调研、数据整理及初稿撰写; 章晓叶与周君哲参与部分数据处理与图表制作; 郑国伟与崔茂应为通讯作者, 负责论文框架设计、修改及最终定稿。所有作者均对论文内容进行了审阅并同意最终版本。

利益冲突: 所有作者均声明不存在利益冲突。

References

[1] Gu LH, Feng JG, Qian LJ, et al. Research on proliferation inhibitory effect of *Paris* saponin I on high metastatic human ovarian cancer cell line HO-8910PM *in vitro* [J]. Chin Arch Tradit Chin Med (中华中医药学刊), 2012, 30: 2212-2215.

[2] Cheng H, Long J, Su J, et al. Mechanism of *Paris polyphylla* saponin II inducing autophagic to inhibit angiogenesis of cervical cancer [J]. Naunyn Schmiedebergs Arch Pharmacol, 2024, 397: 3179-3194.

[3] Luo Y, Jiang YL, Zeng Q, et al. Effects and mechanism of polyphyllin III on proliferation, cycle and apoptosis of colon cancer cell HCT116 [J]. Chin J Inf Tradit Chin Med (中国中医药信息杂志), 2025, 32: 80-85.

[4] Chen Y, Yan Q, Ji Y, et al. Unraveling the serial glycosylation in

the biosynthesis of steroidal saponins in the medicinal plant *Paris polyphylla* and their antifungal action [J]. Acta Pharm Sin B, 2023, 13: 4638-4654.

[5] Xie Y, Gan C, Liu H, et al. Polyphyllin VI ameliorates pulmonary fibrosis by suppressing the MAPK/ERK and PI3K/AKT signaling pathways *via* upregulating DUSP6 [J]. Phytother Res, 2024, 38: 5930-5948.

[6] Miu JX. Inhibitory Effects and Mechanisms of 30 Saponins on the Proliferation of ER*α* -Positive and Triple-Negative Human Breast Cancer Cells (30种皂苷类化合物对ER*α*阳性及三阴性人乳腺癌细胞的增殖抑制作用及机制研究) [D]. Kunming: Kunming Medical University, 2023.

[7] Zhou N, Xu L, Park SM, et al. Genetic diversity, chemical components, and property of biomass *Paris polyphylla* var. *yunnanensis* [J]. Front Bioeng Biotechnol, 2021, 9: 713860.

[8] Gao X, Su Q, Li J, et al. RNA-Seq analysis reveals the important co-expressed genes associated with polyphyllin biosynthesis during the developmental stages of *Paris polyphylla* [J]. BMC Genomics, 2022, 23: 559.

[9] Yin X, Liu J, Kou C, et al. Deciphering the network of cholesterol biosynthesis in *Paris polyphylla* laid a base for efficient diosgenin production in plant chassis [J]. Metab Eng, 2023, 76: 232-246.

[10] Guo SY, Yin Y, Lei T, et al. A cycloartenol synthase from the steroidal saponin biosynthesis pathway of *Paris polyphylla* [J]. J Asian Nat Prod Res, 2021, 23: 353-362.

[11] Onoda K, Kato M, Tsunematsu Y, et al. Biosynthetic gene expression and tissue distribution of diosgenin in *Dioscorea japonica* [J]. J Agric Food Chem, 2023, 71: 4292-4297.

[12] Christ B, Xu C, Xu M, et al. Repeated evolution of cytochrome P450-mediated spiroketal steroid biosynthesis in plants [J]. Nat Commun, 2019, 10: 3206.

[13] Hua X, Song W, Wang K, et al. Effective prediction of biosynthetic pathway genes involved in bioactive polyphyllin in *Paris polyphylla* [J]. Commun Biol, 2022, 5: 50.

[14] Li S, Chen S, Fu C, et al. Identification and characterization of PpUGT91BP1 as a trillin synthase from *Paris polyphylla* [J]. Plant Cell Physiol, 2025, 66: 249-259.

[15] He M, Guo SY, Yin Y, et al. Gene clone and functional identification of sterol glycosyltransferases from *Paris polyphylla* var. *yunnanensis* [J]. China J Chin Mater Med (中国中药杂志), 2023, 48: 3774-3785.

[16] He M, Guo S, Yin Y, et al. A novel sterol glycosyltransferase catalyses steroidal sapogenin 3-*O* glucosylation from *Paris polyphylla* var. *yunnanensis* [J]. Mol Biol Rep, 2023, 50: 2137-2146.

[17] Song W, Li T, Yan S, et al. Two rhamnosyl transferases for *de novo* biosynthesis of polyphyllin in *Nicotiana benthamiana* [J]. Plant J, 2025, 122: e70237.

[18] Zhao CX, Yan YF, Zhao LX, et al. Characterization of a 4'-*O*-rhamnosyltransferase and *de novo* biosynthesis of bioactive steroidal triglycosides from *Paris polyphylla* [J]. Plant Commun, 2025, 6: 101257.

[19] Xie Z, Yu H, Peng S, et al. Comparative transcriptome analysis reveals key genes responsible for the differences in polyphyllin composition in two *Paris polyphylla* species [J]. Gene, 2025, 946: 149325.

[20] Xu P, Chen J, Mi Q, et al. Metabolomics and transcriptomics profiling of three *Paris* species with varied polyphyllin compositions [J]. Sci Data, 2024, 11: 1293.

[21] Chinese Pharmacopoeia Commission. Pharmacopoeia of the People's Republic of China (中华人民共和国药典) [S]. Beijing: China Medical Science Press, 2020: 271.

[22] Wang XX, Jia TY, Yang CJ, et al. Content comparison of nine saponins in different parts of two origin plants of Paridis Rhizoma [J]. Mod Chin Med (中国现代中药), 2022, 24: 2132-2140.

[23] Li J, Lv MQ, Du L, et al. An enormous *Paris polyphylla* genome sheds light on genome size evolution and polyphyllin biogenesis [J]. bioRxiv, 2020. DOI: https://doi. org/10.1101/2020.06.01.126920.

[24] Wang LY, Wang JH, Zhou HM, et al. Cloning and prokaryotic expression of the UDP-glycosyltransferase gene *PpUGT2* from *Paris polyphylla* var. *chinensis* [J]. J Agric Biotechnol (农业生物技术学报), 2025, 33: 1246-1258.

[25] Zhang Q, Chang S, Yang Y, et al. Endophyte-inoculated rhizomes of *Paris polyphylla* improve polyphyllin biosynthesis and yield: A transcriptomic analysis of the underlying mechanism [J]. Front Microbiol, 2023, 14: 1261140.

[26] Qubie AX, Qubie JZ, Feng MK, et al. Identification of genes involved in biosynthesis of spirostane-type polyphyllin in *Paris fargesii* based on transcriptome analysis [J]. Chin Tradit Herb Drugs (中草药), 2023, 54: 6798-6812.

[27] Xu P, Mi Q, Zhang X, et al. Dissection of transcriptome and metabolome insights into the polyphyllin biosynthesis in *Paris* [J]. BMC Plant Biol, 2025, 25: 206.

[28]]Li Y, Yang H, Li Z, et al. Advances in the biosynthesis and molecular evolution of steroidal saponins in plants [J]. Int J Mol Sci, 2023, 24: 2620.

[29] Xu P, Zhang X, Yu M, et al. Full-length transcriptome sequencing and bioinformatics analysis of *Paris* [J] Jiangsu J Agric Sci (江苏农业学报), 2025, 41: 1412-1421.

[30] Lin Y, Hu Q, Ye Q, et al. Diosgenin biosynthesis pathway and its regulation in *Dioscorea cirrhosa* L [J]. PeerJ, 2024, 12: e16702.

[31] Yang H, Li Y, Hu Z, et al. An HD-ZIP I transcription factor DZHDZ32 upregulates diosgenin biosynthesis in *Dioscorea zingiberensis* [J]. Int J Mol Sci, 2025, 26: 4185.

[32] Pu XL, Zhang LY, Zhang JY, et al. Identification of a bZIP transcription factor PpTGA3 regulating polyphyllin biosynthesis in *Paris polyphylla* [J]. Ind Crops Prod, 2025, 225: 120492.