

# 基于生物合成途径结合 LC-MS 技术快速发现近缘稀少植物星毛冠盖藤中化学成分

陆国寿<sup>1,2</sup>, 胡筱希<sup>1\*</sup>, 黄周锋<sup>1</sup>, 张赟赟<sup>1</sup>, 黄建猷<sup>1</sup>

1. 广西壮族自治区中医药研究院, 广西 南宁 530022

2. 广西中药质量标准研究重点实验室, 广西 南宁 530022

**摘要:** **目的** 探讨以生物合成途径结合高分辨液质联用技术, 为近亲缘稀少植物星毛冠盖藤自建成分数据库, 实现其化学成分的快速发现。 **方法** 以星毛冠盖藤为研究对象, 根据京都基因与基因组百科全书 (Kyoto encyclopedia of genes and genomes, KEGG) 生物合成基因挖掘、BioNavi-NP 生物逆合成预测及同属植物化学成分研究等手段获得星毛冠盖藤中潜在的化学成分, 并自建成分数据库。将自建成分数据库与高分辨质谱数据进行比对, 统计整理满足质量误差小于  $6 \times 10^{-6}$  的化合物, 对其进行裂解规律分析及二级碎片离子比对。对鉴定得到的未知化合物采用 reaxys 网站进行检索。 **结果** 星毛冠盖藤转录组 Unigenes 次生代谢 KEGG 通路主要包括苯丙素生物合成、氨酰-tRNA 生物合成、N-聚糖生物合成及萜类骨架生物合成等通路; 从瑶药星毛冠盖藤中快速鉴定出 76 个成分, 其中包括 1 个香豆素类新化合物。 **结论** 以生物合成关联性结合 LC-MS 技术, 能快速高效、低成本的对瑶药星毛冠盖藤中潜在的化学成分进行分析鉴定, 一定程度上解决近亲缘稀少植物药的化学成分库建库难的问题, 为植物药的化学成分快速发现提供一种新思路及方法。

**关键词:** 生物合成途径; 生物逆合成; 高分辨液质联用; 星毛冠盖藤; 近缘稀少植物; 香豆素; 伞形花内酯; 阿魏酸; 柳杉酚  
**中图分类号:** R284.1 **文献标志码:** A **文章编号:** 0253-2670(2024)20-6862-12

**DOI:** 10.7501/j.issn.0253-2670.2024.20.004

## Rapid identification of chemical constituents in close relative lacking plants *Pileostegia tomentella* based on biosynthetic pathway combined with LC-MS technology

LU Guoshou<sup>1,2</sup>, HU Xiaoxi<sup>1</sup>, HUANG Zhoufeng<sup>1</sup>, ZHANG Yunyun<sup>1</sup>, HUANG Jianyou<sup>1</sup>

1. Guangxi Institute of Traditional Medical and Pharmaceutical Sciences, Nanning 530022, China

2. Guangxi Key Laboratory of Traditional Chinese Medicine Quality Standards, Nanning 530022, China

**Abstract: Objective** To establish a component database for close relative lacking plants *Pileostegia tomentella*, and achieve rapid discovery of its chemical components by biosynthetic association and LC-MS. **Methods** Taking *P. tomentella* as the research object, the potential chemical components of *P. tomentella* were obtained by means of KEGG biosynthesis gene mining, BioNavi NP reverse biosynthesis prediction, and Phytochemistry component research of the same genus, and the component database was built. Compare the self built composition database with high-resolution mass spectrometry data, and statistically organize compounds with quality errors less than  $6 \times 10^{-6}$ , their fragmentation patterns were analyzed and secondary fragments ions were compared. The unknown compounds were identified using the reaxys website. **Results** The KEGG pathway of secondary metabolism of unigenes in the transcriptome of *P. tomentella* mainly includes phenylpropanoids biosynthesis, aminoacyl tRNA biosynthesis, N-glycan biosynthesis and terpene skeleton biosynthesis. 76 components were identified from the *P. tomentella*, including a new coumarin compound. **Conclusion** The potential chemical components in *P. tomentella* can be quickly, efficiently, and low-cost analyzed and identified by combining biosynthetic correlation with LC-MS technology, which to some extent solves the problem of difficulty in building a

收稿日期: 2024-03-18

基金项目: 广西自然科学基金面上项目 (2024GXNSFAA010279); 广西科技计划项目 (桂科 AD17195002); 广西自然科学基金面上项目 (2020GXNSFAA259067); 广西中医药适宜技术开发与推广项目 (GZSY20-04)

作者简介: 陆国寿, 主任药师, 从事中药化学研究。Tel: (0771)5868986 E-mail: luguoshou@foxmail.com

\*通信作者: 胡筱希, 助理研究员, 从事中药化学研究。Tel: (0771)5868986 E-mail: huxiaoxi0124@126.com

chemical component library of close relative lacking plant drugs, and provides a new approach and method for rapid discovery of chemical components in plant drugs.

**Key words:** biosynthetic pathway; biological reverse synthesis; high resolution liquid chromatography-mass spectrometry; *Pileostegia tomentella* Hand. -Mazz.; close relative lacking plants; coumarin; umbelliferone; ferulic acid; sugiol

瑶药消瘤藤为虎耳草科植物星毛冠盖藤 *Pileostegia tomentella* Hand. -Mazz. 的干燥根和茎。瑶医认为其性味微甜，微寒。属风打相兼药。具有化痰止咳、清热消肿、散结、活血、抗癌的功效。用于阿毒（癌症）、呐哈（咳嗽）<sup>[1-2]</sup>。星毛冠盖藤为近年在广西瑶族地区被发掘用于治疗肿瘤的瑶族药，被收载入广西壮族自治区瑶药材质量标准（第二卷）。本研究在利用液质联用技术（LC-MS）对星毛冠盖藤进行化学成分分析时发现，若采用查阅近亲缘植物的研究文献的自建化学成分数据库，难以建立起内容丰富且关联性强的成分数据库，究其原因是因为星毛冠盖藤的近亲缘稀少。据《中国植物志》记载<sup>[3]</sup>，星毛冠盖藤原植物的同属植物仅有 2 种，其属于近亲缘稀少的植物，类似于星毛冠盖藤亲缘特点的植物也不是个例，如瑶药田皂角（豆科植物合萌 *Aeschynomene indica* Linn.）、黑草等也存在近亲缘稀少的问题。对于此类植物药，如何自建内容丰富且关联性强的成分数据库来辅助高分辨液质联用技术检测分析，是目前尚存在的一个尚未解决的问题。

转录组学与代谢组学联合或其他多组合联合运用的方式已经在生物合成途径分析、植物代谢途径调控等领域取得了不菲的成果，许多林作物、农作物、中草药的生物合成途径已经得到了诠释<sup>[4]</sup>。植物代谢产物的分类、代谢途径、代谢机制等方面的研究正在逐步完善并建立了多个相关的生物代谢反应数据库，涉及到的天然产物数量以万个为单位。目前文献报道<sup>[5-6]</sup>的相关研究思路，基本限于转录组学与代谢组学或其他组学联合的方式来展开生物合成途径分析、植物代谢途径调控方面的研究。本研究在参考学习已有研究思路方法的基础上，采取逆向思维的方法，以转录组学、生物合成途径分析、代谢成分获取、液质联用测定分析的思路来实现中草药中化学成分的快速发现。通过中药原植物的转录组学研究来获得其主要的生物合成途径，再以其生物合成途径所涉及的代谢产物为切入点建立化学成分库，再结合高分辨液质联用技术对近亲缘稀少的中药进行化学成分研究。研究步骤如下：（1）通过生物合成途径基因挖掘技术，获取中药原植物生物合成代谢途径所关

联的化学成分数据；（2）以中药原植物中已知化学成分为底物，采用生物逆合成预测工具 BioNavi-NP 进行逆向预测，获取中草药原植物生物合成代谢途径所关联的化学成分数据；（3）综合多种手段获取的化学成分数据自建化学成分数据库，结合 UPLC-LTQ-Orbitrap-MS 检测数据进行中药化学成分分析。

本研究基于上述思路，以瑶药星毛冠盖藤为实验对象，通过生物合成途径基因及生物逆合成预测的手段建立起星毛冠盖藤化学成分数据库，最后结合 UPLC-LTQ-Orbitrap-MS 技术从星毛冠盖藤中快速分析鉴定出了 76 个化合物中，采用文献查询建库部分匹配鉴定获得 5 个化合物，采用生物合成途径关联预测、生物逆合成预测等手段建库部分匹配鉴定获得 71 个化合物。瑶药星毛冠盖藤的研究结果让本研究提出的研究思路得到了试验验证。在一定程度上为近亲缘稀少植物在 LC-MS 检测分析中化学成分建库建库难的问题提供了一种解决方法。

## 1 仪器与材料

### 1.1 仪器

Q-Exactive 型液相色谱-线性轨道离子阱组合质谱系统（美国热电公司）；SHZ-DIII 型真空泵（巩义予华仪器有限公司）；FA1004 型电子天平（上海精密科学仪器有限公司）。

### 1.2 材料

消瘤藤药材经广西壮族自治区中医药研究院黄云峰主任药师鉴定为虎耳草科植物星毛冠盖藤 *P. tomentella* Hand. -Mazz.。对照品当药苷（批号 B21644，上海源叶生物科技有限公司）、茵芋苷（批号 B21573，上海源叶生物科技有限公司），质量分数均 >98%。乙腈、甲醇、甲酸、乙酸铵为色谱纯（德国 Merck 公司）；水为超纯水。

## 2 方法与结果

### 2.1 星毛冠盖藤转录组测序与生物合成相关基因的挖掘<sup>[7-9]</sup>

**2.1.1 RNA 的提取及测序** 将样品在冰上融化后粉碎，充分混匀后离心，取适量上清液，用 Trizol 试剂分离提取总 RNA，用 Oligotex mRNA 试剂盒对 RNA 进行纯化。利用 Agilent 2100 检测 RNA 样品

的浓度和完整性, 3 个星毛冠盖藤 RNA 样品均满足转录组测序的建库要求。满足建库要求的 RNA 样品送至北京百迈客生物科技有限公司, 使用 Illumina NovaSeq 平台进行转录组测序。

**2.1.2 数据过滤及组装** 对转录组测序得到原始测序数据 (raw reads) 后进行数据过滤, 去除掉低质量、包含接头和未知碱基 N 含量过高的 reads, 得到高质量测序数据 (clean reads)。利用 Trinity 软件对 clean reads 进行重头组装, 用 cd-hit 软件去除完全一样的序列, 然后使用 tgiel 进行聚类, 合并相似度大于 90%, overlap 长度大于 35 的序列, 最后得到 Unigenes。使用 Illumina NovaSeq 高通量测序平台对 3 个星毛冠盖藤样品进行转录组测序, 共获得 19.52 Gb Clean Data, 各样品 Clean Data 均达到 6.12 Gb, Q30 碱基百分比在 93.66% 及以上, 平均每个样品的 GC 含量占总碱基数的 45.47%。组装后共获得 57 149 条 Unigene。其中长度在 1 kb 以上的 Unigene 有 15 660 条。

**2.1.3 Unigenes 的注释和分类** 使用 DIAMOND 软件将 Unigenes 序列与非冗余蛋白序列 (NCBI non-redundant protein sequences, NR)、注释和修订蛋白序列数据库 (a manually annotated and reviewed protein sequence database, Swiss-Prot)、基因本体论 (gene ontology, GO)、直系同源蛋白簇 (clusters of orthologous groups, COG)、真核同源群簇 (clusters of eukaryotic orthologous groups, KOG)、京都基因与基因组百科全书 (Kyoto encyclopedia of genes and genomes, KEGG) 等数据库进行比对; 使用 KOBAS 得到 Unigenes 在 KEGG 中的 KEGG Orthology 结果; 预测完 Unigenes 的氨基酸序列之后使用 HMMER 软件与 Pfam (protein family) 数据库比对, 获得 Unigenes 的注释信息。共获得 24 195 条 Unigene 的注释结果, 占全部 Unigene 的 42.33%。其中, NR 数据库中注释到的 Unigene 最多, 有 23 838 条, 占总 Unigene 的 41.71%。Unigenes 注释同源基因的物种在相似序列匹配度较高的物种中, 蓝果树 *Nyssa sinensis* Oliv. 所占比例最高, 8 961 条, 其次为茶树 *Camellia sinensis* (L.) O. Ktze. var. *sinensis*, 3 145 条, 中华猕猴桃 *Actinidia chinensis* Planch., 1 401 条, 山梨猕猴桃 *Actinidia rufa* (Sieb. & Zucc.) Planch. ex Miq., 819 条, 葡萄 *Vitis vinifera* L., 734 条等。见图 1-A。

**2.1.4 Unigenes 的 KOG 分类结果** 对星毛冠盖藤转录组数据进行 KOG 数据库功能注释和分类, 共有 12 889 条 Unigenes 被注释到 25 种 KOG 分类中,

共有 3 440 条 Unigenes 被注释到 “一般功能预测 (general function prediction only)”, 其次是 “翻译后修饰, 蛋白质周转, 伴侣 (posttranslational modification, protein turnover, chaperones)” 1 342 条, 还有 “信号转导机制 (signal transduction mechanisms)” 1 328 条, 此外还有 “碳水化合物的运输和新陈代谢 (carbohydrate transport and metabolism)” 761 条等被注释到 KOG 分类中。见图 1-B。

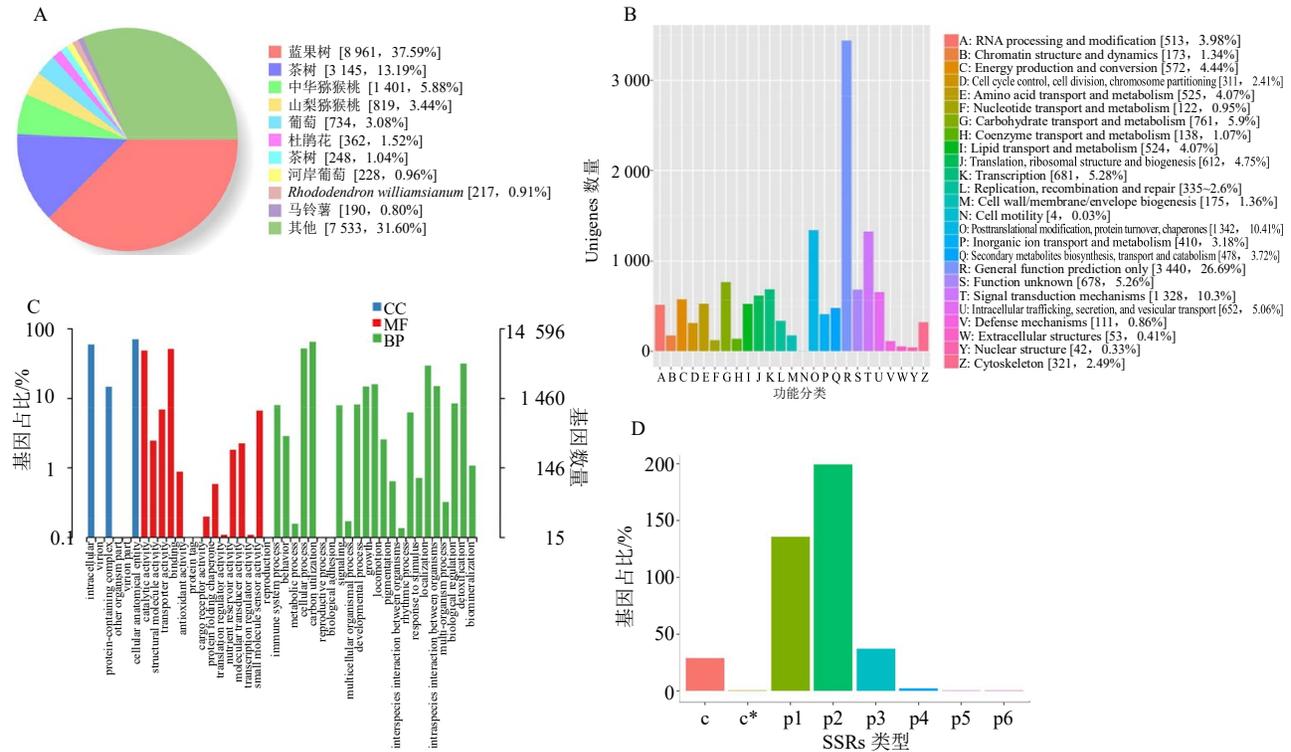
**2.1.5 Unigenes 的 GO 分析结果** 利用 GO 数据库对 Unigenes 进行功能注释, 14 596 条 Unigenes 被注释到细胞组成 (cellular component), 分子功能 (molecular function) 和生物过程 (biological processes) 3 大类中。主要的细胞组成为细胞内 (intracellular) 和蛋白质复合物 (protein-containing complex), 主要的分子功能为催化活性 (catalytic activity) 和结合蛋白 (binding), 主要的生物过程为代谢过程 (metabolic process) 和细胞过程 (cellular process)。见图 1-C。

**2.1.6 SSR 位点分析** 利用鉴定单重复序列的软件 MISA (microsatellite identification tool) 对转录组数据中的 Unigenes 进行 SSR 位点的检测。按照单碱基、双碱基、三碱基重复分别至少 12、6、5 次, 四碱基、五碱基、六碱基都不少于 4 次的标准, 进行检索分析。最后对获得的 SSR 数据进行分类统计分析。利用 MISA (MicroSatellite identification tool) 软件鉴定简单重复序列 (simple sequence repeat, SSR)。对星毛冠盖藤转录组的 Unigenes 进行 SSRs 分析, 共发现 13 040 个 SSRs 位点, 其中 p1 (完美单碱基重复) 4 374 个, p2 (完美双碱基重复) 6 420 个, p3 (完美三碱基重复) 1 195 个, p4 (完美四碱基重复) 74 个, p5 (完美五碱基重复) 11 个, p6 (完美六碱基重复) 17 个, c (混合 SSR) 931 个。见图 1-D。

**2.1.7 Unigenes 次生代谢 KEGG 通路注释及关联化合物整理** 根据转录组测序与次生代谢产物合成相关基因挖掘结果, 通过 KEGG 数据库的通路分析发现其 Unigenes 主要参与的生物合成相关次生代谢通路, 收集整理通路 ID 所涉及的化学成分。KEGG 通路分析发现 Unigenes 参与苯丙素、氨酰 t-RNA、N-聚糖、各类型的 N-聚糖、萜类等生物合成相关的 53 个次生代谢通路。苯丙素的生物合成代谢通路 (ko00940) Unigenes 数量最多, 为 217 条; 氨酰-tRNA 生物合成 (ko00970) Unigenes 数量次之, 为 108 条; 与各类型的 N-聚糖生物合成 (ko00510) 有关的

Unigenes 有 92 条；各类型的 *N*-聚糖生物合成 (ko00513) 的 Unigenes 数量有 84 条。选择表 1 前

13 个相关通路 ID，将它们代谢途径所涉及的化学成分整理到星毛冠盖藤自建化学成分库中。



A-Unigenes 与 NR 数据库匹配物种分布；B-KOG 功能分类；C-GO 功能分类；D-SSR 不同重复类型和重复次数的数量分布。

A-species distribution of *Pileostegia tomentella* Unigenes against NR database; B-KOG functional classification; C-GO functional classification; D-quantitative distribution of different motif lengths and repeats of SSR.

图 1 星毛冠盖藤转录组测序与生物合成相关基因的挖掘结果

Fig. 1 Results of genes related to biosynthesis and sequencing of transcriptome of *P. tomentella*

表 1 星毛冠盖藤转录组 Unigenes 次生代谢 KEGG 通路注释

Table 1 Biosynthetic pathway of secondary metabolites involved in major active substances of *P. tomentella*

编号	通路	通路 ID	Unigenes 数量
1	苯丙素生物合成 (phenylpropanoid biosynthesis)	ko00940	217
2	氨酰-tRNA 生物合成 (aminoacyl-tRNA biosynthesis)	ko00970	108
3	<i>N</i> -聚糖生物合成 ( <i>N</i> -glycan biosynthesis)	ko00510	92
4	各类型的 <i>N</i> -聚糖生物合成 (various types of <i>N</i> -glycan biosynthesis)	ko00513	84
5	萜类骨架生物合成 (terpenoid backbone biosynthesis)	ko00900	70
6	泛醌和其他萜醌生物合成 (ubiquinone and other terpenoid-quinone biosynthesis)	ko00130	67
7	苯丙氨酸、酪氨酸和色氨酸生物合成 (phenylalanine, tyrosine and tryptophan biosynthesis)	ko00400	65
8	类黄酮生物合成 (flavonoid biosynthesis)	ko00941	61
9	脂肪酸生物合成 (fatty acid biosynthesis)	ko00061	59
10	类胡萝卜素生物合成 (carotenoid biosynthesis)	ko00906	54
11	二萜生物合成 (diterpenoid biosynthesis)	ko00904	50
12	倍半萜和三萜生物合成 (sesquiterpenoid and triterpenoid biosynthesis)	ko00909	45
13	鞘糖脂生物合成-神经节系列 (glycosphingolipid biosynthesis - ganglio series)	ko00604	44
14	泛酸和 CoA 生物合成 (pantothenate and CoA biosynthesis)	ko00770	44

表 1 (续)

编号	通路	通路 ID	Unigenes 数量
15	角质、软脂和蜡生物合成 (cutin, suberine and wax biosynthesis)	ko00073	41
16	精氨酸生物合成 (arginine biosynthesis)	ko00220	40
17	玉米素生物合成 (zeatin biosynthesis)	ko00908	39
18	类固醇生物合成 (steroid biosynthesis)	ko00100	38
19	异喹啉生物碱生物合成 (isoquinoline alkaloid biosynthesis)	ko00950	38
20	叶酸合成 (folate biosynthesis)	ko00790	37
21	托烷、哌啶和吡啶生物碱生物合成 (tropane, piperidine and pyridine alkaloid biosynthesis)	ko00960	34
22	二苯醚类、二芳庚烷类和姜辣素生物合成 (stilbenoid, diarylheptanoid and gingerol biosynthesis)	ko00945	33
23	其他类型的 O-聚糖生物合成 (other types of O-glycan biosynthesis)	ko00514	32
24	糖基磷脂酰肌醇生物合成 [glycosylphosphatidylinositol (GPI)-anchor biosynthesis]	ko00563	31
25	不饱和脂肪酸的生物合成 (biosynthesis of unsaturated fatty acids)	ko01040	30
26	异黄酮生物合成 (isoflavonoid biosynthesis)	ko00943	21

## 2.2 BioNavi-NP 生物逆合成预测

通过 <http://biopathnavi.qmclab.com/> 网站利用在线采用生物逆合成预测工具 BioNavi-NP 进行逆向预测: 以星毛冠盖藤中已知的化合物为目标化合物(底物), 在“Target compound”框输入化合物结构或在线绘制结构, “Building blocks” “Quick settings”等参数选默认值, 提交 (Submit), 待网站自动分析预测完成后, 下载分析结果, 获得该底物生物合成途径可能涉及的化合物的文件 (pathway.txt) 及逆合成示意图。收集整理 BioNavi-NP 预测得到的生物合成途径中各化合物的化学结构图、化学式、相对分子质量等。以文献报道的星毛冠盖藤及其同属植物中的化学成分为底物进行 BioNavi-NP 生物逆合成预测, 获得了 14 种底物的生物逆合成预测结果。以化合物断马钱子苷半缩醛内酯 (vogeloside) 为底物进行预测, 包含 5 个生物合成步骤, 涉及 6 个化合物, 其中所涉及的化合物 2、4 在 Reaxys 网站未检索到, 故推测为潜在的新化合物。根据系统命名法, 将化合物 2 命名为 4-((2-acetylbut-3-enoyl)oxy)-2-hydroxy-butanoic acid, 化合物 4 命名为 3,6-dihydroxy-5-vinyl-4,4a,5,6-tetrahydro-

1*H*,3*H*-pyrano[3,4-*c*]pyran-1-one。Vogeloside 生物合成途径逆向预测图见图 2。

## 2.3 星毛冠盖藤的 UPLC-LTQ-Orbitrap-MS 法检测

**2.3.1** 对照品溶液及供试品溶液的制备 取当药苷、茵芋苷对照品, 分别精密称定, 置 25 mL 量瓶中, 以 95%乙醇为溶剂配制成质量浓度为 100、104  $\mu\text{g/mL}$  的对照品溶液。取星毛冠盖藤药材粉末 10 g, 加 100 mL 95%乙醇回流提取 2 次, 每次 30 min, 滤过, 合并滤液, 回收溶剂, 滤液蒸干后用无水乙醇定容于 100 mL 量瓶中, 得供试液 (以生药计质量浓度为 100 mg/mL)。

**2.3.2** 液相色谱条件 XTerra C<sub>18</sub> 色谱柱 (100 mm  $\times$  2.1 mm, 3.5  $\mu\text{m}$ ); 流动相 A 为 0.1%甲酸 + 2.5 mmol/L 乙酸铵水, 流动相 B 为乙腈; 梯度洗脱: 0~15 min, 90%~70% A; 15~25 min, 70%~30% A; 25~30 min, 30%~20% A; 30~32 min, 20%~10% A; 32~37 min, 10%~10% A; 37~37.10 min, 10%~90% A; 37.10~40 min, 90% A; 体积流量 0.2 mL/min; 柱温 25  $^{\circ}\text{C}$ ; 进样量 2  $\mu\text{L}$ 。

**2.3.3** 质谱条件 ESI 离子源; 正离子检测; 毛细管温度 320  $^{\circ}\text{C}$ ; 脱气溶剂温度 350  $^{\circ}\text{C}$ ; 脱气溶剂气流

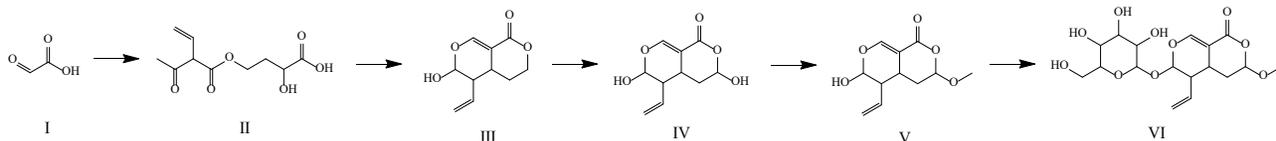


图 2 化合物 vogeloside 生物合成途径逆向预测示意图

Fig. 2 Schematic diagram of reverse prediction of vogeloside biosynthesis pathway of compounds

量 10 L/min; 扫描范围  $m/z$  150~2 000; 雾化器压力 35 000 kPa; 毛细管电压 4 kV; 锥孔电压 100 V。

### 2.4 化学成分数据库建立及数据分析处理

综合整理通过文献查询<sup>[10-15]</sup>、生物合成途径关联预测、生物逆合成预测等手段获得的关联化合物,以化学结构图、化学式、相对分子质量等内容自建星毛冠盖藤化学成分库。利用 Compound Discoverer 3.3、Thermo Xcalibur 2.0、proteoWizard 等软件进行质谱数据处理,得到一级质谱精确相对分子质量,结合自建的化学成分数据库进行成分比对,再通过推测化合物可能的裂解规律与二级质谱中

的主要碎片离子匹配,鉴定出星毛冠盖藤中潜在的化学成分。

综合整理通过文献查询、生物合成途径关联预测、生物逆合成预测等手段获得的关联化合物自建星毛冠盖藤化学成分库与 UPLC-LTQ-Orbitrap-MS 法测定质谱数据进行了对比,共匹配鉴定出 76 个化合物,其中包括 1 个潜在的新化合物,命名为化合物 2 (3,6-dihydroxy-5-vinyl-4,4a,5,6-tetrahydro-1*H*,3*H*-pyrano[3,4-*c*]pyran-1-one); 鉴定出的化合物中主要包括有机酸类、香豆素类、生物碱类、氨基酸类等类化合物。结果见图 3 和表 2。

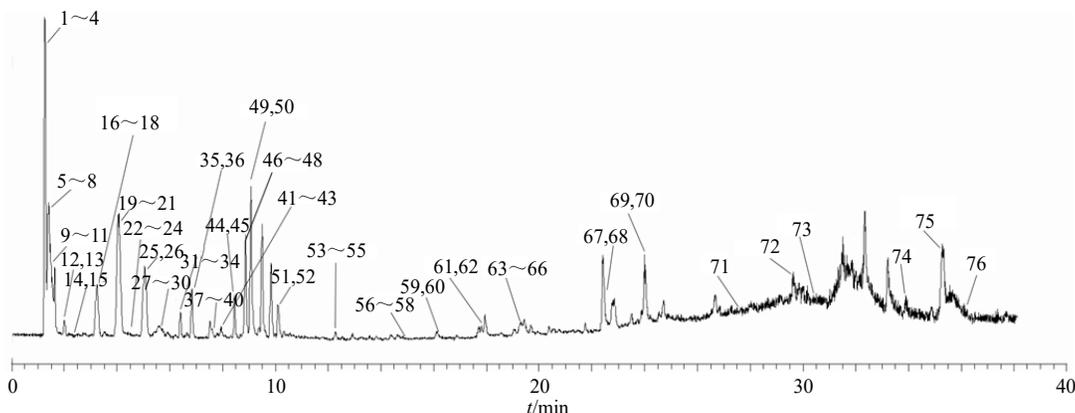


图 3 星毛冠盖藤 95%乙醇提取物的总离子流图 (正离子模式)

Fig. 3 Total ion flow diagram of 95% ethanol extract of *P. tomentella* (positive ion mode)

表 2 星毛冠盖藤样品化学数据库比对分析结果

Table 2 Secondary MS data of *P. tomentella*

序号	$t_R$ /min	电离模式	理论值 ( $m/z$ )	测得值 ( $m/z$ )	特征离子	分子式	化合物
1	1.29	[M]	192.063 3	192.062 0	146.029 5, 128.019 7	$C_7H_{12}O_6$	奎宁酸
2	1.30	[M+H] <sup>+</sup>	175.119 4	175.119 2	158.095 9, 128.922 1, 110.911 4	$C_6H_{14}N_4O_2$	精氨酸
3	1.36	[M+H] <sup>+</sup>	123.044 5	123.040 7	107.060 8, 89.035 2	$C_7H_6O_2$	4-羟基苯甲醛
4	1.38	[M]	351.069 5	351.069 0	303.092 7, 258.076 5	$C_{17}H_{18}ClNO_3S$	<i>N</i> -(对甲苯磺酰基)- <i>L</i> -苯丙氨酸酰胺 甲基氯酮
5	1.40	[M]	204.089 8	204.086 9	159.065 4	$C_{11}H_{12}N_2O_2$	色氨酸/同分异构体
6	1.41	[M]	360.157 2	360.150 3	330.139 6, 298.073 3, 282.588 7, 268.102 9	$C_{20}H_{24}O_6$	(-)-落叶松醇
7	1.44	[M+H] <sup>+</sup>	156.102 4	156.102 2	125.386 6, 108.958 7	$C_8H_{13}NO_2$	倒千里光裂碱
8	1.48	[M+H] <sup>+</sup>	147.092 1	147.095 9	118.086 7	$C_9H_{10}N_2$	(1 <i>H</i> -indol-3-yl)methanamine
9	1.53	[M]	188.094 9	188.092 0	158.069 7, 144.080 7	$C_{11}H_{12}N_2O$	<i>N</i> -methyl-1-(5-phenyl-1,2-oxazol-3-yl)methanamine
10	1.55	[M+H] <sup>+</sup>	182.081 6	182.081 5	152.061 9, 136.953 2	$C_9H_{11}NO_3$	酪氨酸
11	1.56	[M+H] <sup>+</sup>	222.112 9	222.112 7	190.986 6, 180.065 8	$C_{12}H_{15}NO_3$	老头掌碱
12	2.17	[M+H] <sup>+</sup>	167.070 7	167.070 5	149.023 4, 120.024 0	$C_9H_{10}O_3$	托品酸
13	2.26	[M]	392.199 9	392.191 9	372.066 7, 343.081 8	$C_{22}H_{29}FO_5$	secondary alcohol/同分异构体
14	2.40	[M+H] <sup>+</sup>	220.118 4	220.118 2	203.142 9, 172.952 6	$C_9H_{17}NO_5$	泛酸

表2 (续)

序号	$t_R$ /min	电离模式	理论值 ( $m/z$ )	测得值 ( $m/z$ )	特征离子	分子式	化合物
15	2.57	[M+H] <sup>+</sup>	137.060 2	137.060 0	121.065 4, 107.048 9, 91.054 7	C <sub>8</sub> H <sub>8</sub> O <sub>2</sub>	2-(4-hydroxyphenyl) acetaldehyde
16	3.10	[M+H] <sup>+</sup>	179.070 7	179.070 5	146.035 7, 130.982 0, 103.055 0	C <sub>10</sub> H <sub>10</sub> O <sub>3</sub>	松柏醛/同分异构体
17	3.18	[M+H] <sup>+</sup>	151.075 8	151.075 5	133.974 2, 118.943 2, 103.047 6	C <sub>9</sub> H <sub>10</sub> O <sub>2</sub>	对香豆醇/同分异构体
18*	3.21	[M+H] <sup>+</sup>	359.134 1	359.133 8	197.081 6, 179.070 6, 151.075 4	C <sub>16</sub> H <sub>22</sub> O <sub>9</sub>	獐牙菜苷
19	3.51	[M]	217.106 2	217.107 3	194.943 9	C <sub>8</sub> H <sub>15</sub> N <sub>3</sub> O <sub>4</sub>	<i>N</i> -acetyl-l-citrulline
20*	3.83	[M+H] <sup>+</sup>	163.039 4	163.039 2	134.024 7, 51.047 5	C <sub>9</sub> H <sub>6</sub> O <sub>3</sub>	6-羟基香豆素
21*	4.04	[M+H] <sup>+</sup>	325.092 3	325.091 9	163.038 5, 119.049 3	C <sub>15</sub> H <sub>16</sub> O <sub>8</sub>	茵芋苷
22	4.07	[M]	370.141 6	370.149 6	252.372 6, 152.725 3	C <sub>21</sub> H <sub>22</sub> O <sub>6</sub>	裂榄脂素
23	4.09	[M]	353.073 6	353.070 5	257.049 5, 143.085 0, 119.085 9	C <sub>9</sub> H <sub>16</sub> N <sub>5</sub> O <sub>8</sub> P	2,5-diamino-6-hydroxy-4-(5-phosphoribosylamino) pyrimidine
24	4.21	[M+H] <sup>+</sup>	179.034 4	179.034 2	163.033 6, 146.035 7	C <sub>9</sub> H <sub>6</sub> O <sub>4</sub>	秦皮乙素
25	4.45	[M+H] <sup>+</sup>	193.086 4	193.086 2	151.035 2, 122.036 6	C <sub>11</sub> H <sub>12</sub> O <sub>3</sub>	醋酸香豆素酯
26	4.69	[M+H] <sup>+</sup>	197.081 4	197.081 1	167.076 2, 141.054 8	C <sub>10</sub> H <sub>12</sub> O <sub>4</sub>	5-羟基茴香醇
27	4.69	[M+H] <sup>+</sup>	225.196 6	225.196 4	142.981 2	C <sub>13</sub> H <sub>24</sub> N <sub>2</sub> O	aniferine
28	5.05	[M+H] <sup>+</sup>	355.102 8	355.102 6	193.049 8, 163.039 5	C <sub>16</sub> H <sub>18</sub> O <sub>9</sub>	东莨菪苷
29	5.37	[M+H] <sup>+</sup>	211.060 6	211.060 4	181.064 7, 165.070 5, 136.904 2	C <sub>10</sub> H <sub>10</sub> O <sub>5</sub>	5-羟基阿魏酸
30	5.52	[M+H] <sup>+</sup>	165.091 5	165.091 3	148.087 7, 133.941 7	C <sub>10</sub> H <sub>12</sub> O <sub>2</sub>	丁香酚
31	5.64	[M+H] <sup>+</sup>	213.076 2	213.076 0	212.051 9, 170.097 2, 151.038 7, 126.055 1	C <sub>10</sub> H <sub>12</sub> O <sub>5</sub>	化合物II
32	5.75	[M+H] <sup>+</sup>	195.065 7	195.065 5	149.059 9, 121.028 9	C <sub>10</sub> H <sub>10</sub> O <sub>4</sub>	阿魏酸
33	5.84	[M+H] <sup>+</sup>	223.060 6	223.060 3	207.028 1, 162.031 1	C <sub>11</sub> H <sub>10</sub> O <sub>5</sub>	2-琥珀酰苯甲酸酯
34	6.13	[M+H] <sup>+</sup>	179.070 7	179.070 5	164.039 5, 130.982 0	C <sub>10</sub> H <sub>10</sub> O <sub>3</sub>	松柏醛/同分异构体
35	6.39	[M+H] <sup>+</sup>	211.097 0	211.096 7	194.071 7, 179.070 8, 147.945 9, 131.048 9	C <sub>11</sub> H <sub>14</sub> O <sub>4</sub>	芥子醇
36	6.40	[M]	413.147 4	413.142 1	349.111 9, 222.270 9	C <sub>22</sub> H <sub>23</sub> NO <sub>7</sub>	芥子碱
37	6.58	[M+H] <sup>+</sup>	331.154 5	331.154 1	218.866 3, 116.995 1	C <sub>19</sub> H <sub>22</sub> O <sub>5</sub>	5-脱氧独脚金醇/同分异构体
38	6.89	[M+H] <sup>+</sup>	147.044 5	147.044 2	132.080 9, 123.964 4, 86.060 7	C <sub>9</sub> H <sub>6</sub> O <sub>2</sub>	香豆素
39	6.91	[M+H] <sup>+</sup>	339.107 9	339.107 7	165.054 4, 147.041 1, 119.049 6	C <sub>16</sub> H <sub>18</sub> O <sub>8</sub>	对香豆素奎尼酸
40	7.07	[M+H] <sup>+</sup>	177.055 1	177.034 2	134.108 2, 108.958 9, 95.049 3	C <sub>10</sub> H <sub>8</sub> O <sub>3</sub>	7-甲氧基香豆素
41	7.45	[M+H] <sup>+</sup>	347.149 4	347.149 4	249.810 2, 114.091 4	C <sub>19</sub> H <sub>22</sub> O <sub>6</sub>	独脚金醇
42	7.58	[M]	450.193 6	450.197 5	417.025 8, 274.495 9	C <sub>20</sub> H <sub>36</sub> O <sub>7</sub> P <sub>2</sub>	geranylgeranyl pyrophosphate
43	7.58	[M+H] <sup>+</sup>	225.076 2	225.075 9	155.981 5, 144.996 8	C <sub>11</sub> H <sub>12</sub> O <sub>5</sub>	芥子酸
44	7.77	[M+H] <sup>+</sup>	165.055 1	165.054 9	119.951 2, 94.042 2	C <sub>9</sub> H <sub>8</sub> O <sub>3</sub>	香豆素酸
45*	7.84	[M+H] <sup>+</sup>	163.039 4	163.039 2	134.977 1, 116.966 4, 107.049 0	C <sub>9</sub> H <sub>6</sub> O <sub>3</sub>	伞形花内酯
46	7.92	[M+H] <sup>+</sup>	181.086 4	181.086 2	166.985 8, 152.065 1	C <sub>10</sub> H <sub>12</sub> O <sub>3</sub>	松柏醇
47	8.09	[M+H] <sup>+</sup>	164.047 3	164.042 5	136.974 3, 123.079 1, 108.052 8	C <sub>9</sub> H <sub>8</sub> O <sub>3</sub>	咖啡因醛
48	8.19	[M+H] <sup>+</sup>	153.055 1	153.054 9	119.957 4, 110.035 9	C <sub>8</sub> H <sub>8</sub> O <sub>3</sub>	3,4-dihydroxyphenyl acetaldehyde
49	9.04	[M]	194.057 9	194.053 0	165.070 5, 138.066 5	C <sub>10</sub> H <sub>10</sub> O <sub>4</sub>	5-羟基茴香醛
50	9.09	[M]	178.026 6	178.026 3	136.021 6, 108.959 0	C <sub>9</sub> H <sub>6</sub> O <sub>4</sub>	祖师麻甲素
51	10.31	[M+H] <sup>+</sup>	179.070 7	179.070 5	136.021 5, 123.043 9	C <sub>10</sub> H <sub>10</sub> O <sub>3</sub>	松柏醛/同分异构体
52	10.42	[M+H] <sup>+</sup>	219.102 0	219.101 7	190.990 2, 134.999 3	C <sub>13</sub> H <sub>14</sub> O <sub>3</sub>	benzyl 2-(3-oxocyclobutyl) acetate

表 2 (续)

序号	$t_R$ /min	电离模式	理论值 ( $m/z$ )	测得值 ( $m/z$ )	特征离子	分子式	化合物
53	12.91	[M+H] <sup>+</sup>	437.341 9	437.341 9	391.337 5, 159.117 0	C <sub>30</sub> H <sub>44</sub> O <sub>2</sub>	2-phytyl-1,4-naphthoquinone
54	13.18	[M+H] <sup>+</sup>	402.155 2	402.155 0	307.060 1, 162.060 7	C <sub>21</sub> H <sub>23</sub> NO <sub>7</sub>	4'-O-desmethylpapaveroxine
55	13.40	[M]	450.349 7	450.343 0	257.014 6, 224.987 1, 184.991 6	C <sub>31</sub> H <sub>46</sub> O <sub>2</sub>	维生素 K1
56	14.81	[M]	202.215 7	202.216 8	185.041 7, 170.069 6	C <sub>10</sub> H <sub>26</sub> N <sub>4</sub>	精胺
57	14.91	[M]	290.260 9	290.269 2	219.016 8, 206.944 1, 192.076 4	C <sub>20</sub> H <sub>34</sub> O	香叶基香叶醇
58	15.22	[M+H] <sup>+</sup>	247.097 0	247.093 6	186.935 1, 145.101 3	C <sub>14</sub> H <sub>14</sub> O <sub>4</sub>	二氢欧山芹醇
59	16.39	[M+H] <sup>+</sup>	291.195 9	291.195 6	255.673 4	C <sub>18</sub> H <sub>26</sub> O <sub>3</sub>	3-[(2E)-3,7-dimethylocta-2,6-dienyl]-5-methoxy-2-methylbenzene-1,4-diol
60	16.83	[M]	255.158 3	255.159 2	131.085 9, 111.081 2	C <sub>12</sub> H <sub>21</sub> N <sub>3</sub> O <sub>3</sub>	吡咯赖氨酸
61	17.94	[M+H] <sup>+</sup>	318.206 8	318.206 5	256.261 8	C <sub>19</sub> H <sub>27</sub> NO <sub>3</sub>	原吐根碱/同分异构体
62	18.35	[M]	272.250 4	272.258 6	192.910 0, 84.960 4	C <sub>20</sub> H <sub>32</sub>	蓖麻烯
63	19.28	[M]	460.261 3	460.269 5	391.144 8, 147.064 8	C <sub>30</sub> H <sub>36</sub> O <sub>4</sub>	4,4'-diapolycopenedioic Acid
64	19.66	[M+H] <sup>+</sup>	295.190 9	295.190 5	250.050 0, 137.096 0	C <sub>17</sub> H <sub>26</sub> O <sub>4</sub>	姜辣素
65	19.85	[M+H] <sup>+</sup>	297.242 9	297.242 6	267.000 5, 194.943 8	C <sub>18</sub> H <sub>32</sub> O <sub>3</sub>	18-oxooleic acid/同分异构体
66	19.90	[M+H] <sup>+</sup>	167.070 7	167.070 5	121.994 3, 95.049 3	C <sub>9</sub> H <sub>10</sub> O <sub>3</sub>	3-(2-羟基苯基)丙酸
67	22.36	[M]	211.070 5	211.075 6	194.071 7, 165.070 5	C <sub>7</sub> H <sub>9</sub> N <sub>5</sub> O <sub>3</sub>	5,6,7,8-tetrahydropterin-6-carboxylic acid
68	22.75	[M+H] <sup>+</sup>	301.216 7	301.216 2	283.281 4, 259.169 6, 163.075 1	C <sub>20</sub> H <sub>28</sub> O <sub>2</sub>	柳杉酚
69	23.98	[M+H] <sup>+</sup>	235.169 7	235.169 3	188.924 2	C <sub>15</sub> H <sub>22</sub> O <sub>2</sub>	germacrene acid
70	24.37	[M+H] <sup>+</sup>	277.180 3	277.180 1	235.169 4, 222.067 4	C <sub>17</sub> H <sub>24</sub> O <sub>3</sub>	2-[(2E)-3,7-dimethylocta-2,6-dienyl]-6-methoxy-benzene-1,4-diol
71	27.93	[M+H] <sup>+</sup>	439.357 5	439.357 6	393.349 6, 264.963 5	C <sub>30</sub> H <sub>46</sub> O <sub>2</sub>	去甲基叶啉醇 1
72	29.87	[M+H] <sup>+</sup>	361.237 8	361.234 4	288.985 4, 172.115 6	C <sub>22</sub> H <sub>32</sub> O <sub>4</sub>	大麻萜酚酸
73	30.45	[M]	204.089 8	204.087 7	158.080 8, 131.974 4	C <sub>11</sub> H <sub>12</sub> N <sub>2</sub> O <sub>2</sub>	色氨酸/同分异构体
74	33.82	[M+H] <sup>+</sup>	425.378 3	425.378 3	380.185 7, 300.000 9	C <sub>30</sub> H <sub>48</sub> O	环己烯酮
75	35.66	[M]	610.189 7	610.184 9	489.061 7, 285.007 8	C <sub>28</sub> H <sub>34</sub> O <sub>15</sub>	新橙皮苷
76*	36.03	[M+H] <sup>+</sup>	429.373 2	429.373 6	385.091 0, 343.079 4	C <sub>29</sub> H <sub>48</sub> O <sub>2</sub>	5- $\alpha$ -stigmastane-3,6-dione

\*为采用文献查询建库部分匹配鉴定获得的 5 个化合物, 其余为通过生物合成途径并结合 LC-MS 鉴定的化合物。

\*The five compounds obtained by partial match identification using literature query library building, and the rest were identified by biosynthetic route and combined with LC-MS.

(1) 化合物 2 的鉴定: ESI<sup>+</sup>模式下给出的准分子离子峰为  $m/z$  212.051 9 [M+H]<sup>+</sup>, 碎片离子  $m/z$  170.097 2 是准分子离子发生断裂, 脱去 1 分子 C<sub>2</sub>H<sub>5</sub>OH 而产生, 碎片离子  $m/z$  126.055 1 和  $m/z$  151.038 7 是继续分别脱去 1 分子 COOH 和 1 分子 H<sub>2</sub>O 产生的, 质谱元素组成计算得出分子式为 C<sub>10</sub>H<sub>12</sub>O<sub>5</sub>, 根据该化合物的碎片离子及其裂解规律分析, 鉴定其为化合物 II(3,6-dihydroxy-5-vinyl-4,4a,5,6-tetrahydro-1H,3H-pyrano[3,4-c]pyran-1-one), 见图 4。

(2) 香豆素类成分(伞形花内酯)的鉴定: ESI<sup>+</sup>模式下给出的准分子离子峰为  $m/z$  163.039 2 [M+H]<sup>+</sup>, 二级质谱图中, 准分子离子连续脱去 2 分子 CO 形成碎片离子  $m/z$  134.977 1 [M+H-CO]<sup>+</sup>和  $m/z$  107.049 0 [M+H-2CO]<sup>+</sup>, 另一个碎片离子  $m/z$  116.966 4 是由准分子离子脱去 1 分子 CO 和 1 分子 H<sub>2</sub>O 而产生, 质谱元素组成计算得出分子式为 C<sub>9</sub>H<sub>6</sub>O<sub>3</sub>, 根据该化合物的碎片离子, 并结合文献报道<sup>[16]</sup>, 鉴定其为伞形花内酯。

(3) 有机酸成分(阿魏酸)的鉴定: ESI<sup>+</sup>模式

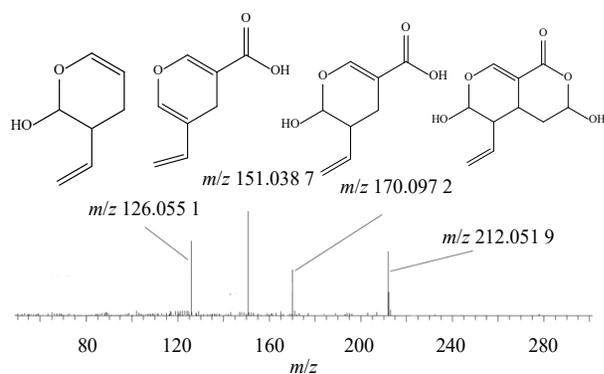


图 4 化合物 2 的二级碎片裂解规律图

Fig. 4 Pyrolysis pattern of secondary fragments of compound 2

下给出的准分子离子峰为  $m/z$  195.065 5  $[M+H]^+$ ，二级质谱图中，准分子离子连续脱去 1 分子  $-CH_3$  形成碎片离子  $m/z$  179.029 9  $[M+H-CH_3]^+$ ，另一个碎片离子  $m/z$  149.059 9 是由准分子离子脱去 1 分子  $-COOH$  而产生，在脱去 1 分子的  $-CH_3$  得到碎片离子  $m/z$  138.065 5，质谱元素组成计算得出分子式为  $C_{10}H_{10}O_4$ ，根据该化合物的碎片离子，并结合文献报道<sup>[17]</sup>，鉴定其为阿魏酸。

(4) 有机酚类成分（柳杉酚）的鉴定：ESI<sup>+</sup>模式下给出的准分子离子峰为  $m/z$  301.216 2  $[M+H]^+$ ，二级质谱图中，准分子离子连续脱去 1 分子  $H_2O$  形成碎片离子  $m/z$  283.281 4  $[M+H-H_2O]^+$ ，继而发生断裂，形成  $m/z$  259.169 6 的碎片离子，同时，准分子离子发生重排，得到  $m/z$  240.233 7 的碎片离子， $m/z$  163.075 1 是由于准分子离子直接产生断裂生成的碎片离子，质谱元素组成计算得出分子式为  $C_{20}H_{28}O_2$ ，根据该化合物的碎片离子，鉴定其为柳杉酚。

(5) 生物碱类成分（原吐根碱）的鉴定：ESI<sup>+</sup>模式下给出的准分子离子峰为  $m/z$  308.206 5  $[M+H]^+$ ，二级质谱图中，准分子离子连续脱去 2 分子  $CO$  形成碎片离子  $m/z$  262.143 3  $[M+H-H_2O]^+$ ，再脱去 1 分子羰基  $O$ ，形成  $m/z$  247.120 4 的碎片离子。质谱元素组成计算得出分子式为  $C_{19}H_{27}NO_3$ ，根据该化合物的碎片离子，并结合文献报道<sup>[18]</sup>，鉴定其为原土根碱。

(6) 星毛冠盖藤中苯丙素类成分的部分生物合成途径：根据 KEGG 途径分析结果，并结合星毛冠盖藤化学成分分析结果，对星毛冠盖藤中苯丙素途径的代谢产物成分的生物合成途径进行分析。5-Hydroxyferulic acid 是一种羟基肉桂酸，是苯丙烷

途径的代谢产物。其中涉及多种酶的参与，如苯丙氨酸解氨酶（phenylalanine ammonia-lyase, PTAL）、香豆酸-3-羟基化酶（*p*-coumarate 3-hydroxylase）咖啡酰辅酶 A-O-甲基转移酶（caffeic acid 3-O-methyltransferase, COMT）、阿魏酸-5-羟基化酶（ferulate-5-hydroxylase, F5H）。对香豆酸（*p*-coumaric acid）在 4-香豆酸：辅酶 A 连接酶（4-coumarate:CoA ligase, 4CL）的催化下，形成对香豆酰辅酶 A（*p*-coumaroyl-CoA），在 2,4-二羟基肉桂酰辅酶 A 还原酶（2,4-dihydroxy-cinnamoyl-CoA）的催化下形成伞形花内酯（umbelliferone）。芥子醇（sinapyl alcohol）可由 4-羟基肉桂醛（*p*-coumaraldehyde）经过香豆素-3-羟基化酶（*p*-coumarate 3-hydroxylase）、COMT、F5H 及肉桂醇脱氢酶（cinnamyl-alcohol dehydrogenase, CAD）一系列催化下得到，见图 5。

(7) 星毛冠盖藤中生物碱类成分的部分生物合成途径：根据 KEGG 途径分析结果，并结合星毛冠盖藤化学成分相关报道，对星毛冠盖藤中的生物碱成分原吐根碱（protoemetine）的生物合成途径进行预测。以 *L*-络氨酸（*L*-tyrosine）为底物，在多种酶的催化下，如络氨酸脱羧酶（tyrosine decarboxylase, TYDC）、单酚氧化酶（monophenol）、脱乙酰基异吲哚啉合酶（deacetylisoipecoside synthase）、*O*-甲基转移酶（*O*-methyltransferase, OMT）、 $\beta$ -葡萄糖苷酶（ipecoside  $\beta$ -*D*-glucosidase, GLU1），经过一系列的反应得到，见图 6。

### 3 讨论

通过查询同属植物的化学成分研究文献，借鉴同属植物化学成分研究结果来自建化学成分库，是高分辨液质联用技术对植物药化学成分研究常采用的思路之一，但这一思路对同属近亲缘稀少的植物药来说并不适用。随着转录组学及代谢组学的不断发展，以转录组学入手，对药材原植物进行转录组学生物合成途径挖掘获得与该药材原植物相关的代谢途径，并通过生物合成代谢数据库查询来获得相关的化合物进行自建化学成分库。基于该思路自建的化学成分库与相应药材具有较好的关联性，同时这一建库思路更适合近亲缘稀少的植物药。

本研究主要采用这一建库思路，结合高分辨液质联用技术从瑶药星毛冠盖藤中共鉴定出的 76 种化合物当中，仅有 5 种化合物是采用文献查询建库部分匹配鉴定获得，其余都是通过生物合成途径关联预测、生物逆合成预测等手段建库部分匹配鉴定

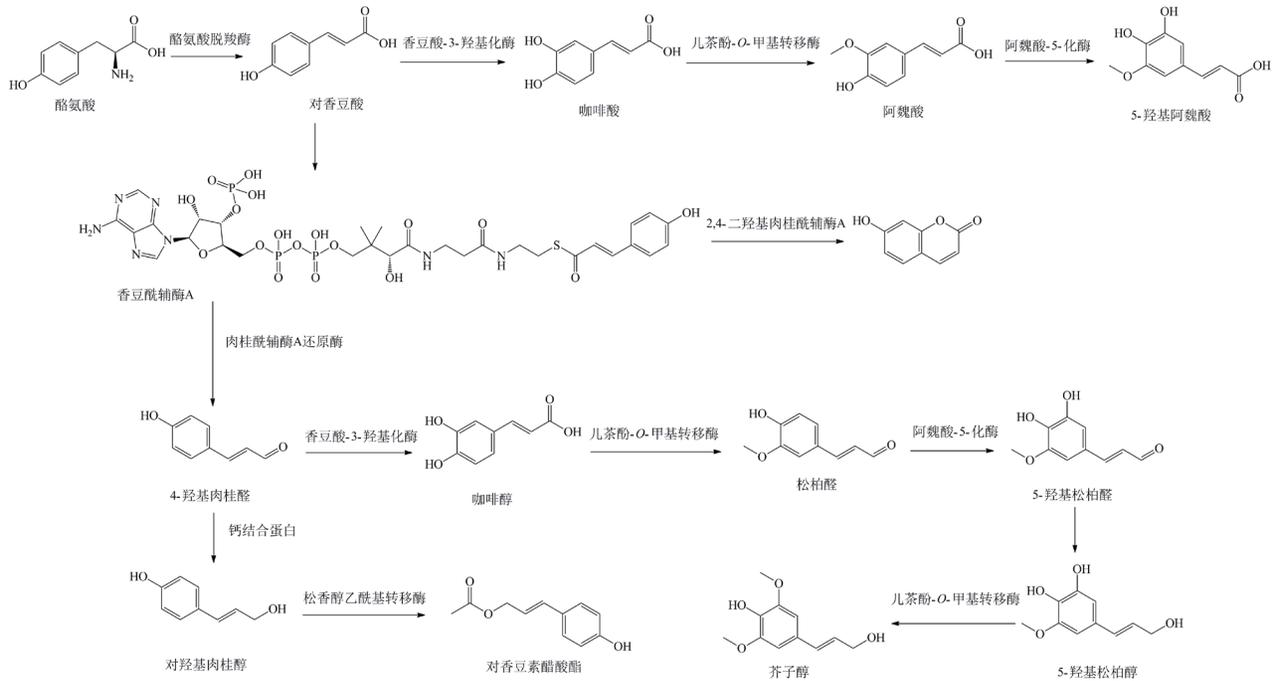


图5 星毛冠盖藤中苯丙素类成分部分合成途径

Fig. 5 Partial phenylpropanoid biosynthesis from *P. tomentella*

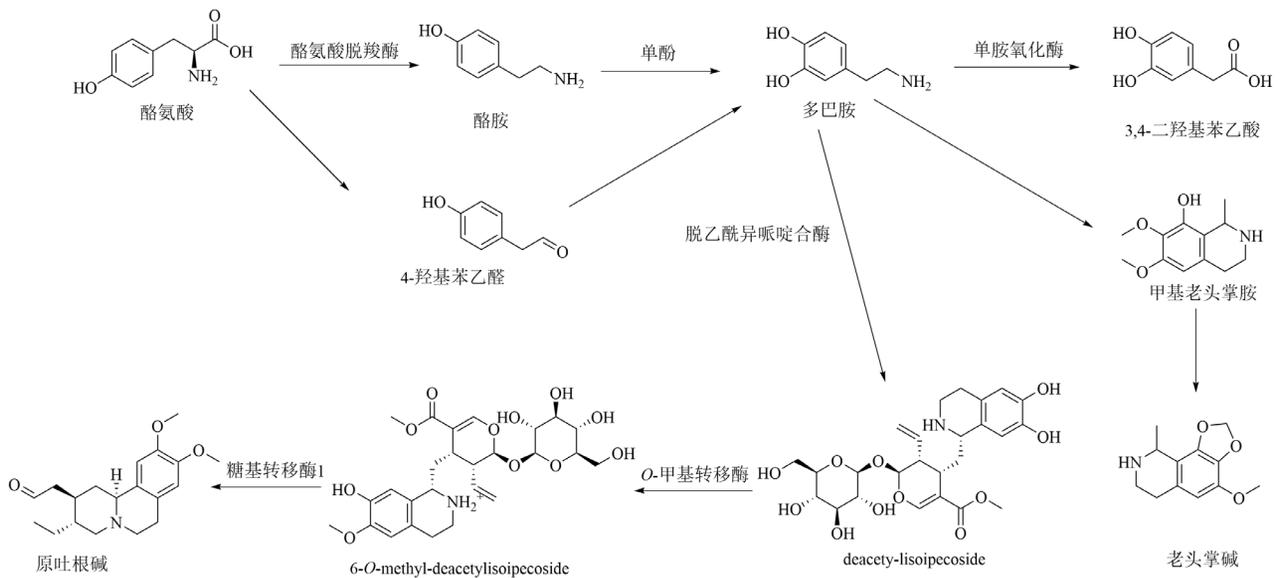


图6 星毛冠盖藤中生物碱类成分部分合成途径

Fig. 6 Partial alkaloid biosynthesis from *P. tomentella*

获得。这一鉴定结果说明采用常规的文献查询建库对近亲缘稀少植物有较大的局限性，而通过生物合成途径关联预测、生物逆合成预测等手段建库可以较好地丰富建库数据。同时，以生物合成途径为思路进行建库，相应的化学成分数据与相应植物有较好的关联性。本研究团队在对星毛冠盖藤开展研究工作时，已发现星毛冠盖藤中富含香豆素类成分<sup>[10,19-20]</sup>，

本研究从瑶药星毛冠盖藤中鉴定出的化合物中，亦匹配鉴定出了9个香豆素类或香豆素生物合成的前体化合物，其中包括茵芋苷、6-羟基香豆素、对香豆醇、秦皮乙素、对香豆素奎尼酸、7-甲氧基香豆素、香豆素酸、伞形花内酯、祖师麻甲素等。这一结果与前期研究结果相吻合，说明本研究方法具有较好的准确度及可行性。

在研究过程中发现,若仅通过转录组学的生物合成途径基因挖掘来获得关联成分也存在一定的局限性,若药材主要代谢通路与 KEGG 等数据库收录的代谢通路相匹配的通路较少,则获得的关联化学成分也会偏少,难建立起丰富的关联性高的化学成分数据库。为了弥补这一方面的不足,本研究又兼顾采用了生物逆合成预测的手段来获得更多的关联成分用于自建化学成分库。生物逆合成预测是以拟进行生物合成为目标物(底物),利用基于 AI 超算技术深度学习驱动的生物逆合成路径导航工具逆向预测该目标物(底物)潜在的生物合成途径及相应的关联化合物,可作为生物合成路线设计研工作的辅助。Zheng 等<sup>[21]</sup>建立起的生物逆合成路径导航工具 BioNavi-NP。BioNavi-NP 可以识别 368 种测试化合物中 90.2%的生物合成途径,并以 72.8%的测试集恢复报告的构件,比现有的常规基于规则的方法精确 1.7 倍。Delépine 等<sup>[22]</sup>构建了一个基于广义反应规则的自动开源逆向合成工作流程工具 RetroPath2.0。该规则通过高效且控制良好的协议从底物到目标执行逆向合成搜索,其易用性和用途的多样性使其可作为生物工程师工作的宝贵补充。本研究利用相应的生物逆合成预测工具,来分析预测瑶药星毛冠盖藤生物合成中有可能存在的化合物,除了可以弥补单通过转录组学生物合成途径基因挖掘获得关联化合物的不足之外,也在一定程度上使生物逆合成路径导航工具的应用领域得到了拓展。此外,利用过转录组学生物合成途径基因挖掘获得的关联化合物均为已知化合物。而生物逆合成路径导航工具逆向预测则有可能预测到未知化合物。本研究以星毛冠盖藤中已知化合物 vogeloside 为目标物(底物),通过利用 BioNavi-NP 生物逆合成工具进行预测,预测到 vogeloside 的生物合成过程中涉及到的 2 个化合物应为新化合物,并结合高分辨液质联用数据分析,得到了 1 个与预测结果相匹配的化合物。

本研究相应工作的开展及完成,可以在一定程度上对瑶药星毛冠盖藤实现了快速、高效、准确的化学成分发现,也为近亲缘稀少植物的化学成分研究工作提供了一种思路。

**利益冲突** 所有作者均声明不存在利益冲突

#### 参考文献

[1] 国家中医药管理局《中华本草》编委会. 中华本草-9 [M]. 上海: 上海科学技术出版社, 1999: 40-41.

- [2] 广西壮族自治区药品监督管理局. 广西壮族自治区药品监督管理局关于广西壮族自治区瑶药材质量标准(第二卷)的公示 [EB/OL]. <http://yj.j.gxzf.gov.cn/yp/ypgzwj/t10910167.shtml>, 2021-12-06/2022-12-14.
- [3] 中国科学院中国植物志编辑委员会. 中国植物志(第 35 卷) [M]. 北京: 科学出版社, 1995: 173-175.
- [4] 王莉, 张艳霞, 史玲玲, 等. 功能基因组学和代谢组学技术在植物次生代谢物合成及调控研究中的应用 [J]. 北京林业大学学报, 2007, 29(5): 153-159.
- [5] Zhang R Z, Zhao J T, Wang W Q, et al. Metabolomics-based comparative analysis of the effects of host and environment on *Viscum coloratum* metabolites and antioxidative activities [J]. *J Pharm Anal*, 2022, 12(2): 243-252.
- [6] 李琦, 高晓悦, 陈万生, 等. 黄花蒿中青蒿素生物合成相关转录因子研究进展 [J]. 中草药, 2021, 52(6): 1827-1833.
- [7] 刘畅, 俸婷婷, 刘雄伟, 等. 苗药八爪金龙转录组测序与次生代谢产物合成相关基因的挖掘 [J]. 中草药, 2021, 52(5): 1434-1447.
- [8] 朱昀昊, 张梦佳, 李璐, 等. 夏枯草的转录组测序与次生代谢产物生物合成相关基因的挖掘 [J]. 中草药, 2019, 50(5): 1220-1226.
- [9] 朱昀昊, 董诚明, 郑晓珂, 等. 基于转录组测序的山茱萸次生代谢生物合成相关基因的挖掘 [J]. 中国中药杂志, 2017, 42(2): 213-219.
- [10] 陆国寿, 卢文杰, 黄周锋, 等. 瑶族药消瘤藤的化学成分分析 [J]. 中国实验方剂学杂志, 2017, 23(14): 78-81.
- [11] 王云卿, 马国需, 赵丹, 等. 瑶药星毛冠盖藤正丁醇部位化学成分研究 [J]. 中国药学杂志, 2017, 52(18): 1580-1584.
- [12] 王云卿, 马国需, 梁琼平, 等. 瑶药“肿瘤藤”(星毛冠盖藤)的化学成分研究 [J]. 中草药, 2016, 47(24): 4326-4330.
- [13] 李琴雯. 柔毛冠盖藤根茎正丁醇部分化学成分及其活性研究 [D]. 长沙: 湖南中医药大学, 2013.
- [14] 邹菊英, 陈胜璜, 李琴雯, 等. 柔毛冠盖藤化学成分研究 [J]. 中药材, 2012, 35(5): 724-725.
- [15] 李小军, 袁燕, 李芝, 等. 苗药冠盖藤的化学成分研究 [J]. 中草药, 2014, 45(8): 1052-1055.
- [16] 顿珠次仁, 邓敏芝, 朱根华, 等. 基于 UPLC-Q-TOF-MS 技术快速分析裂叶独活中香豆素类成分 [J]. 中药药理与临床药理, 2022, 33(1): 105-113.
- [17] 吴吉祥, 高方圆, 叶晓岚, 等. 创新药物川阿格雷及其拼合分子阿魏酸和川穹嗪的质谱裂解规律 [J]. 第二军医大学学报, 2012, 33(7): 755-758.
- [18] 孙磊, 胡晓茹, 刘丽娜, 等. 实时直接分析-串联质谱法

- (DART-MS/MS) 快速检测吐根中生物碱 [J]. 中国中药杂志, 2012, 37(10): 1426-1430.
- [19] 黄建猷, 胡筱希, 麦琬婷, 等. 指纹图谱及多成分定量结合化学模式识别法评价不同产地消瘤藤质量 [J]. 中草药, 2021, 52(14): 4334-4340.
- [20] 刘瑛, 陆国寿, 胡筱希, 等. 基于网络药理学的消瘤藤总香豆素抗肿瘤潜在机制探讨 [J]. 中药药理与临床, 2019, 35(6): 57-63.
- [21] Zheng S J, Zeng T, Li C T, *et al.* Deep learning driven biosynthetic pathways navigation for natural products with BioNavi-NP [J]. *Nat Commun*, 2022, 13(1): 3342.
- [22] Delépine B, Duigou T, Carbonell P, *et al.* RetroPath2.0: A retrosynthesis workflow for metabolic engineers [J]. *Metabolic Engineering*, 2018, 45: 158-170.

[责任编辑 王文倩]