

# 基于UPLC-Q-TOF-MS/MS与Python算法的伤科接骨片成分自动解析

张新佳<sup>1</sup>, 王戟森<sup>2</sup>, 梅宇翔<sup>2</sup>, 周志刚<sup>2</sup>, 肖雪<sup>3,4\*</sup>, 章弘扬<sup>1\*</sup>

(1. 华东理工大学 化学与分子工程学院, 上海 200237; 2. 九江市第一人民医院 骨科, 江西 九江 332000; 3. 广东药科大学 中医药研究所(广东省代谢病中西医结合研究中心), 广东 广州 510006; 4. 国家药品监督管理局药品快速检验技术重点实验室(广东省药品检验所) 广东 广州 510663)

**摘要:** 该研究基于超高效液相色谱-四极杆-飞行时间串联质谱(UPLC-Q-TOF-MS/MS), 结合自建化学成分分子式库和特征碎片离子库, 采用Python中的NumPy矩阵运算和Matplotlib绘图技术, 开发了一套特征碎片离子  $m/z$  理论值自动比对算法, 实现了对伤科接骨片中多成分的快速筛查与自动解析。结合标样比对与文献分析, 深入探讨了各类成分的MS/MS裂解规律。利用该方法, 从伤科接骨片中鉴定出77个化合物, 包括28个三萜类、16个皂苷类、9个氨基酸类、9个黄酮苷类、6个生物碱类、6个脂肪酸类、2个环烯醚萜苷类和1个色素类成分。性能验证显示, 该Python算法的检出率为100%, 相对标准偏差(RSD)小于3.0%、误匹配率为2.5%。此研究可为该中药复方的药效物质基础和质量评价研究提供科学数据与技术支撑。

**关键词:** 伤科接骨片; 超高效液相色谱-四极杆-飞行时间串联质谱; Python算法; 成分自动解析; 质谱裂解机理

**中图分类号:** O657.7; R284.1 **文献标识码:** A **文章编号:** 1004-4957(2026)06-0001-12

## Automated Identification of Chemical Constituents in Shangke Jiegu Tablet Based on UPLC-Q-TOF-MS/MS and Python-based Algorithms

ZHANG Xin-jia<sup>1</sup>, WANG Ji-sen<sup>2</sup>, MEI Yu-xiang<sup>2</sup>, ZHOU Zhi-gang<sup>2</sup>, XIAO Xue<sup>3,4\*</sup>, ZHANG Hong-yang<sup>1\*</sup>

(1. School of Chemistry and Molecular Engineering, East China University of Science and Technology, Shanghai 200237, China; 2. Department of Orthopaedics, The First People's Hospital of Jiujiang, Jiujiang 332000, China; 3. Institute of Traditional Chinese Medicine(Guangdong Metabolic Disease Research Center of Integrated Chinese and Western Medicine), Guangdong Pharmaceutical University, Guangzhou 510006, China; 4. NMPA Key Laboratory for Rapid Testing Technology of Drugs(Guangdong Institute for Drug Control), Guangzhou 510663, China)

**Abstract:** Based on ultra-high performance liquid chromatography-quadrupole time-of-flight tandem mass spectrometry(UPLC-Q-TOF-MS/MS), this study developed an automated matching algorithm for the theoretical  $m/z$  values of characteristic fragment ions, utilizing NumPy matrix operations and Matplotlib visualization in Python. By integrating self-built databases of molecular formulas and characteristic fragment ions, the platform achieved rapid screening and automated identification of multi-component in Shangke Jiegu Tablet. The MS/MS fragmentation patterns of various compounds were systematically investigated and validated through comparison with the reference standards and literatures. A total of 77 compounds were identified, including 28 triterpenoids, 16 saponins, 9 amino acids, 9 flavonoid glycosides, 6 alkaloids, 6 fatty acids, 2 iridoid glycosides, and 1 pigment. Performance validation demonstrates that this Python algorithm achieves the detection rate of 100%, relative standard deviations(RSDs) less than 3.0%, and false matching rate of 2.5%. This

收稿日期: 2026-02-05; 修回日期: 2026-03-04

基金项目: 中国仪器仪表学会科学仪器托举计划项目(CISTJ2024); 国家药品监督管理局药品快速检验技术重点实验室开放课题(KF2022006)

\* 通讯作者: 肖雪, 博士, 副研究员, 研究方向: 中药质量研究、药品快检与过程分析, E-mail: erxiaohappy@163.com; 章弘扬, 博士, 副教授, 研究方向: LC-MS复杂体系分析新方法, E-mail: hongyang\_zhang@ecust.edu.cn

research provides scientific data and technical support for the study of the pharmacodynamic material basis and quality evaluation of this traditional Chinese medicine prescription.

**Key words:** Shangke Jiegu Tablet; UPLC-Q-TOF-MS/MS; Python-based algorithm; automated identification of constituents; mass spectrometry fragmentation mechanism

伤科接骨片是一种主治骨科损伤类疾病的中药复方制剂,由三七、红花、甜瓜子、马钱子(粉)、土鳖虫、乳香(炙)、没药(炙)、海星(炙)、鸡骨(炙)、冰片、朱砂、自然铜(煅)12味中药材组成,具有活血化瘀、消肿止痛、舒筋壮骨等功效<sup>[1-2]</sup>。其化学组成十分复杂,包含皂苷、黄酮、生物碱、萜类及脂肪酸等多种活性成分<sup>[3-5]</sup>,各成分极性跨度极大,且存在不少同分异构体。这种“多成分、宽极性”的物质基础特征,在赋予其协同药效优势的同时,也为该复方快速、全面的化学表征与质量控制带来了挑战。

目前,伤科接骨片的质量评价主要依赖高效液相色谱(HPLC)指纹图谱技术<sup>[6-10]</sup>。但受限于常规液相色谱峰容量和检测器灵敏度,普遍存在分析周期长、分离度不足、低丰度成分难检出等问题。此外,单一色谱方法仅能提供保留时间和紫外吸收信息,缺乏能揭示化合物结构的质谱信息,导致各峰物质缺乏明确的化学归属,无法满足复方整体表征的需求。近年来,超高效液相色谱-四极杆-飞行时间串联质谱(UPLC-Q-TOF-MS/MS)技术凭借其高分离度、高分辨率和高灵敏度的特点,成为中药复方多成分鉴定强有力的工具<sup>[11-13]</sup>。然而,该技术的引入也带来了一定挑战:单次分析即可产生包含大量化合物保留时间、*m/z*值和峰强度的多维数据<sup>[14]</sup>。面对如此庞杂的信息,传统人工解谱不仅工作量大、效率低,还易造成错检、漏检。随着计算机和人工智能技术的飞速发展,基于算法、软件的自动解析策略逐渐成为中药复杂成分定性的关键手段<sup>[15-17]</sup>。

为此,本研究在自建化学成分质谱库的基础上,建立了一种UPLC-Q-TOF-MS/MS分析结合Python编程的新方法,实现伤科接骨片中多成分的快速筛查与自动鉴定。

## 1 实验部分

### 1.1 仪器与试剂

Agilent 1290 UPLC-6530 Q-TOF 液质联用仪(美国 Agilent 公司); AL 104 分析天平(瑞士梅特勒-托利多仪器有限公司); 玛瑙研钵(上海新诺仪器集团有限公司); EPDE-E1-T 型超纯水机(南京易普易达科技发展有限公司); 40 HZ 300 W 数控功率可调加热超声波清洗机(上海声彦超声波仪器有限公司); 200、1000  $\mu$ L 移液枪(德国 Eppendorf 公司)。

*L*-色氨酸(99%, 上海泰坦科技股份有限公司); 马钱苷酸(98%, 成都植标化纯生物技术有限公司); 土的宁、马钱子碱(均为98%, 北京奇松生物科技有限公司); 三七皂苷 R1、人参皂苷 Rg1、人参皂苷 Rb1(均为98%, 南通飞宇生物科技有限公司)、亚油酸(98%, 上海阿拉丁生化科技股份有限公司)、11-羰基- $\beta$ -乙酰乳香酸(98%, 上海毕得医药科技股份有限公司); 伤科接骨片(大连美罗中药厂有限公司, 批号: 20190919); HPLC 级乙腈、甲醇(德国 Merck 公司); 甲酸(HPLC 级, 上海阿拉丁生化科技股份有限公司)。

### 1.2 标准品及样品溶液制备

分别精密称取9种标准品,加入甲醇溶解,配制成混合标准溶液,质量浓度依次为:*L*-色氨酸, 100  $\mu$ g/mL; 马钱苷酸, 50  $\mu$ g/mL; 土的宁、马钱子碱、11-羰基- $\beta$ -乙酰乳香酸, 均为 35.7  $\mu$ g/mL; 三七皂苷 R1、人参皂苷 Rg1、亚油酸, 均为 357  $\mu$ g/mL; 人参皂苷 Rb1, 892  $\mu$ g/mL。该混标溶液置于-20  $^{\circ}$ C 冻存。

取伤科接骨片4片,刮去包衣后,放入玛瑙研钵研细。精密称取粉末约1.0 g,置于10 mL离心管中,加入5 mL甲醇,超声30 min,静置冷却后取上层清液,过0.22  $\mu$ m 滤膜(PTFE, Pall Corporation),即得。

### 1.3 UPLC-Q-TOF-MS/MS 条件

采用 Agilent Eclipse Plus C<sub>18</sub> 色谱柱(2.1 mm $\times$ 100 mm, 1.8  $\mu$ m), 流动相为0.02%甲酸水溶液(A)-乙腈(B), 梯度洗脱程序: 0~7 min, 5%~13% B; 7~8 min, 13%~15% B; 8~13 min, 15%~27% B; 13~14

min, 27% B; 14~20 min, 27%~35% B; 20~26 min, 35%~60% B; 26~28 min, 60% B; 28~30 min, 60%~74% B; 30~36 min, 74%~75% B; 36~37 min, 75%~80% B; 37~41 min, 80%~84% B; 41~43 min, 84%~100% B。柱温40 °C, 流速0.3 mL/min, 进样量2  $\mu$ L。

质谱检测分别在正、负模式下进行。干燥气温度350 °C; 干燥气流速10 L/min, 雾化气压力207 kPa, 毛细管电压3 500 V(正模式)、3 000 V(负模式), 毛细管出口电压125 V, 截取锥电压65 V, 八级杆射频电压750 V, 质量扫描范围 $m/z$  50~1 500。内标校正离子 $m/z$  121.050 9、922.009 8(正模式)和 $m/z$  112.985 6、1 033.988 1(负模式)。二级质谱采用Auto-MS/MS模式, 碰撞电压设为20 V。

#### 1.4 伤科接骨片成分分子式库建立

为尽可能覆盖所有可能的化学成分, 通过检索TCMSP(<https://old.tcmsp-e.com/tcmsp.php>)、ChemSpider(<http://www.chemspider.com>)、PubChem(<https://pubchem.ncbi.nlm.nih.gov/>)等开源数据库, 并结合大量文献调研, 建立了伤科接骨片化学成分理论库。库中包含氨基酸类、色素类、环烯醚萜苷类、黄酮苷类、生物碱类、皂苷类、脂肪酸类、三萜类等923个化合物, 将其名称、分子式、理论分子量等信息整理成csv格式文件, 用于一级质谱分子式自动匹配。

#### 1.5 数据处理

使用Agilent MassHunter 10.0软件分析采集到的样品总离子流图(TIC), 以自建的伤科接骨片成分分子式库进行初步筛查。设定匹配参数: 得分 $\geq$ 80、质量数匹配容差 $\pm$ 10 ppm。将匹配结果导出为包含名称、分子式、精确准分子离子 $m/z$ 值和保留时间等信息的化合物列表。在此基础上, 采用Auto-MS/MS模式得到伤科接骨片各成分的碎片离子信息。通过研究其裂解规律, 构建相应的特征碎片离子理论库, 结合Python算法辅助, 实现伤科接骨片成分的自动鉴定。

## 2 结果与讨论

### 2.1 伤科接骨片一级质谱分析

伤科接骨片提取液在正、负模式下的TIC结果见图1。经分子式匹配后发现, 绝大多数化合物在正、负模式下分别以 $[M+H]^+$ 和 $[M+H]^-$ 形式存在, 少数化合物产生 $[M+Na]^+$ 和 $[M+HCOO]^-$ 加合离子。黄酮苷、脂肪酸、部分皂苷和少数氨基酸类在负模式下信号更高, 其余化合物则在正模式下响应更强。

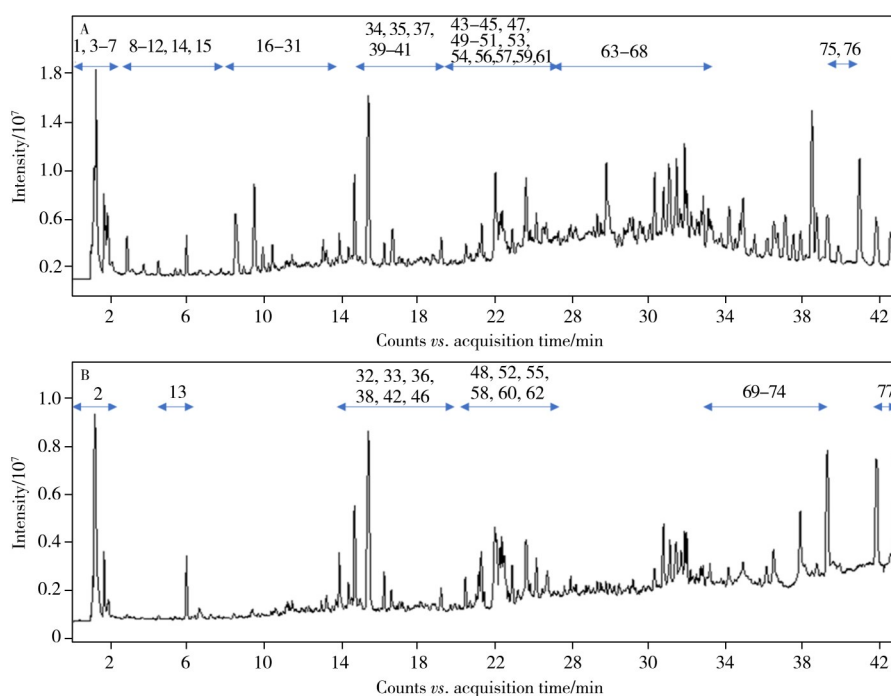


图1 伤科接骨片样品的总离子流图

Fig. 1 Total ion chromatograms of Shangke Jiegu Tablet sample

A: positive mode, B: negative mode; peak numbers in this figure are consistent with those in Table 1

## 2.2 Python 算法辅助伤科接骨片成分鉴定

利用自建特征碎片离子库结合 Python 编程, 建立了质谱数据自动解析流程(见图 2)。首先, 以 MS Convert 软件将原始数据转为 .mzML 格式, 并导入 Jupyter Notebook 运行环境中。Python 算法采用模块化设计, 自动对数据进行去噪预处理, 保留相对丰度大于 1% 的离子。随后, 基于精确准分子离子峰  $m/z$  在理论库中进行候选化合物锚定。在此基础上, 核心模块利用 Python 的 NumPy 矩阵运算功能, 对候选物的实测碎片与理论碎片进行高通量比对。算法内置了加权评分模型, 综合考量了特征离子的匹配数目、质量偏差及碎片强度, 尽可能排除假阳性干扰。计算如下:

$$\text{Score} = \left( \frac{N_{\text{match}}}{N_{\text{theo}}} \right) \times 40 + \left( 1 - \frac{\overline{\Delta\text{ppm}}}{20} \right) \times 30 + \left( \frac{\sum I_{\text{match}}}{I_{\text{base}}} \right) \times 30 \quad (1)$$

式中,  $(N_{\text{match}}/N_{\text{theo}})$  表示实际碎片匹配率(权重为 40%),  $N_{\text{match}}$  和  $N_{\text{theo}}$  分别为实际和理论的碎片数目;  $(1 - \overline{\Delta\text{ppm}}/20)$  用于评估碎片平均质量偏差(权重为 30%), 20 ppm 为设定的碎片质量容差上限;  $(\sum I_{\text{match}})/I_{\text{base}}$  表示实际匹配碎片在谱图中的丰度贡献(权重为 30%),  $\sum I_{\text{match}}$  和  $I_{\text{base}}$  分别为实际碎片强度总和与基峰强度。

为排除假阳性, 本算法要求候选化合物须同时满足以下条件: (1) 准分子离子质量偏差  $\leq 10$  ppm; (2) 匹配综合评分  $S \geq 70$ 。低于此阈值的候选物被判为低置信度, 并自动剔除。采用 Matplotlib 库对结果进行可视化, 以实测/理论谱图对比模式呈现特征碎片离子  $m/z$ 、相对强度和注释名称。上述算法的部分代码如图 3 所示。最终从伤科接骨片中鉴定了 77 个化合物, 包括 9 个氨基酸类、1 个色素类、2 个环烯醚萜苷类、9 个黄酮苷类、6 个生物碱类、16 个皂苷类、6 个脂肪酸类和 28 个三萜类。各成分的保留时间、分子式、准分子离子  $m/z$  值、特征碎片离子归属及匹配得分见表 1。

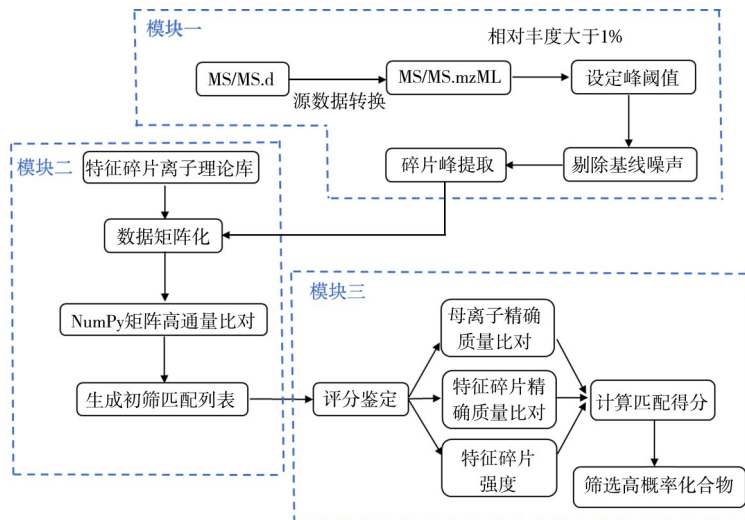


图 2 本研究质谱数据自动解析算法流程

Fig. 2 Workflow of the automatic interpretation algorithm for the MS data in this study

表 1 算法辅助的伤科接骨片多成分质谱鉴定结果

Table 1 Algorithm-assisted MS identification results of the multi-component in Shangke-Jiegu Tablet

No.	Compound	$t_r/\text{min}$	Formula	$m/z$	MS/MS fragment	Score
1	精氨酸	0.78	$\text{C}_6\text{H}_{14}\text{N}_4\text{O}_2$	175.1187 [M+H] <sup>+</sup>	60.0560 [CH <sub>6</sub> N <sub>3</sub> ] <sup>+</sup> , 70.0654 [C <sub>4</sub> H <sub>8</sub> N] <sup>+</sup> , 116.0701 [M+H-CH <sub>3</sub> N <sub>3</sub> ] <sup>+</sup> , 158.0920 [M+H-NH <sub>3</sub> ] <sup>+</sup>	89.38
2	牛磺酸	0.89	$\text{C}_2\text{H}_7\text{NO}_3\text{S}$	124.0083 [M-H] <sup>-</sup>	79.9563 [M-H-C <sub>2</sub> H <sub>6</sub> N] <sup>-</sup> , 80.9690 [M-H-C <sub>2</sub> H <sub>6</sub> N] <sup>-</sup>	94.74
3	缬氨酸	0.95	$\text{C}_5\text{H}_{11}\text{NO}_2$	118.0876 [M+H] <sup>+</sup>	55.0550 [M+H-NH <sub>3</sub> ] <sup>+</sup> , 72.0814 [M+H-HCOOH] <sup>+</sup>	96.05
4	亮氨酸	1.10	$\text{C}_6\text{H}_{13}\text{NO}_2$	132.1014 [M+H] <sup>+</sup>	86.0952 [M+H-HCOOH] <sup>+</sup> , 69.0690 [M+H-HCOOH-NH <sub>3</sub> ] <sup>+</sup>	92.64

(续表1)

No.	Compound	$t_R$ /min	Formula	$m/z$	MS/MS fragment	Score
5	酪氨酸	1.19	$C_9H_{11}NO_3$	182.0819 [M+H] <sup>+</sup>	91.0549 [C <sub>7</sub> H <sub>7</sub> ] <sup>+</sup> , 107.0498 [M+H-C <sub>2</sub> H <sub>5</sub> NO <sub>2</sub> ] <sup>+</sup> , 123.0442 [M+H-NH <sub>3</sub> -C <sub>2</sub> H <sub>2</sub> O] <sup>+</sup> , 136.0759 [M+H-HCOOH] <sup>+</sup> , 165.0541 [M+H-NH <sub>3</sub> ] <sup>+</sup>	97.65
6	异亮氨酸	1.64	$C_6H_{13}NO_2$	132.1014 [M+H] <sup>+</sup>	86.0962 [M+H-HCOOH] <sup>+</sup> , 69.0710 [M+H-HCOOH-NH <sub>3</sub> ] <sup>+</sup>	81.07
7	脯氨酸	1.78	$C_5H_9NO_2$	116.0709 [M+H] <sup>+</sup>	70.0658 [M+H-HCOOH] <sup>+</sup> , 72.0792 [M+H-CO <sub>2</sub> ] <sup>+</sup>	72.85
8	苯丙氨酸	2.12	$C_9H_{11}NO_2$	166.0860 [M+H] <sup>+</sup>	120.0804 [M+H-HCOOH] <sup>+</sup> , 103.0541 [M+H-NH <sub>3</sub> -HCOOH] <sup>+</sup> , 91.0536 [C <sub>7</sub> H <sub>7</sub> ] <sup>+</sup> , 77.0389 [C <sub>6</sub> H <sub>5</sub> ] <sup>+</sup>	98.73
9 <sup>*</sup>	色氨酸	3.30	$C_{11}H_{12}N_2O_2$	205.0974 [M+H] <sup>+</sup>	188.0703 [M+H-NH <sub>3</sub> ] <sup>+</sup> , 170.0594 [M+H-NH <sub>3</sub> -H <sub>2</sub> O] <sup>+</sup> , 159.0911 [M+H-HCOOH] <sup>+</sup> , 118.0651 [C <sub>8</sub> H <sub>7</sub> N+H] <sup>+</sup>	97.85
10	4-羟基土的宁	4.56	$C_{21}H_{22}N_2O_3$	351.1707 [M+H] <sup>+</sup>	323.1390 [M+H-C <sub>2</sub> H <sub>4</sub> ] <sup>+</sup> , 280.1317 [M+H-C <sub>3</sub> H <sub>5</sub> NO] <sup>+</sup> , 238.1182 [M+H-C <sub>3</sub> H <sub>5</sub> NO-C <sub>2</sub> H <sub>2</sub> O] <sup>+</sup>	74.92
11	羟基红花黄色素 A	4.98	$C_{27}H_{32}O_{16}$	613.1741 [M+H] <sup>+</sup>	595.1613 [M+H-H <sub>2</sub> O] <sup>+</sup> , 577.1589 [M+H-2H <sub>2</sub> O] <sup>+</sup>	81.94
12	槲皮素-3, 7, 4'-三-O-葡萄糖苷	5.42	$C_{33}H_{40}O_{22}$	789.2108 [M+H] <sup>+</sup>	627.1552 [M+H-Glu] <sup>+</sup> , 465.1022 [M+H-2Glu] <sup>+</sup> , 303.0484 [M+H-3Glu] <sup>+</sup>	95.95
13 <sup>*</sup>	马钱苷酸	5.97	$C_{16}H_{24}O_{10}$	375.1319 [M-H] <sup>-</sup>	331.1373 [M-H-CO <sub>2</sub> ] <sup>-</sup> , 195.0649 [M-H-Glu-H <sub>2</sub> O] <sup>-</sup> , 161.0450 [Glu-H] <sup>-</sup>	86.94
14	马钱苷	6.10	$C_{11}H_{16}O_5$	229.1070 [M+H] <sup>+</sup>	211.0952 [M+H-H <sub>2</sub> O] <sup>+</sup> , 179.0686 [M+H-H <sub>2</sub> O-CH <sub>2</sub> OH] <sup>+</sup> , 151.0753 [M+H-H <sub>2</sub> O-CH <sub>2</sub> OH-CO] <sup>+</sup>	79.33
15	19-N-甲基土的宁	6.33	$C_{22}H_{24}N_2O_2$	349.1907 [M+H] <sup>+</sup>	321.1615 [M+H-C <sub>2</sub> H <sub>4</sub> ] <sup>+</sup> , 248.1069 [M+H-C <sub>3</sub> H <sub>11</sub> NO] <sup>+</sup>	73.14
16 <sup>*</sup>	土的宁	8.31	$C_{21}H_{22}N_2O_2$	335.1803 [M+H] <sup>+</sup>	317.1624 [M+H-H <sub>2</sub> O] <sup>+</sup> , 198.0899 [M+H-C <sub>8</sub> H <sub>11</sub> NO] <sup>+</sup>	78.62
17	马钱子碱氮氧化物	8.53	$C_{23}H_{26}N_2O_5$	411.1908 [M+H] <sup>+</sup>	393.1804 [M+H-H <sub>2</sub> O] <sup>+</sup> , 383.1582 [M+H-C <sub>2</sub> H <sub>4</sub> ] <sup>+</sup>	76.18
18	19-N-甲基羟基土的宁	9.03	$C_{22}H_{24}N_2O_3$	365.1861 [M+H] <sup>+</sup>	350.1568 [M+H-CH <sub>3</sub> ] <sup>+</sup> , 332.1524 [M+H-CH <sub>3</sub> -H <sub>2</sub> O] <sup>+</sup> , 250.0905 [M+H-C <sub>6</sub> H <sub>13</sub> NO] <sup>+</sup>	76.88
19	6-羟基山奈酚-3-芸香苷-6-葡萄糖苷	9.12	$C_{33}H_{40}O_{21}$	773.2124 [M+H] <sup>+</sup>	627.1564 [M+H-Rha] <sup>+</sup> , 611.1584 [M+H-Glu] <sup>+</sup> , 465.1034 [M+H-Glu-Rha] <sup>+</sup>	95.92
20	6-羟基山奈酚-3, 6-二糖苷	9.20	$C_{27}H_{30}O_{17}$	627.1563 [M+H] <sup>+</sup>	465.1013 [M+H-Glu] <sup>+</sup> , 303.0490 [M+H-2Glu] <sup>+</sup>	91.96
21 <sup>*</sup>	马钱子碱	9.55	$C_{23}H_{26}N_2O_4$	395.1963 [M+H] <sup>+</sup>	367.1653 [M+H-C <sub>2</sub> H <sub>4</sub> ] <sup>+</sup> , 244.0986 [M+H-C <sub>9</sub> H <sub>13</sub> NO] <sup>+</sup>	80.66
22	槲皮素	9.79	$C_{15}H_{10}O_7$	303.0494 [M+H] <sup>+</sup>	285.0501 [M+H-H <sub>2</sub> O] <sup>+</sup> , 257.0468 [M+H-H <sub>2</sub> O-CO] <sup>+</sup> , 229.0516 [M+H-H <sub>2</sub> O-CO-CO] <sup>+</sup>	77.48
23	6-羟基山奈酚-3-O-葡萄糖苷	10.23	$C_{21}H_{20}O_{12}$	465.1036 [M+H] <sup>+</sup>	303.0497 [M+H-Glu] <sup>+</sup> , 121.0307 [C <sub>7</sub> H <sub>5</sub> O <sub>2</sub> ] <sup>+</sup>	79.99
24	人参皂苷 Rh1	10.57	$C_{36}H_{62}O_9$	639.4463 [M+H] <sup>+</sup>	621.4336 [M+H-H <sub>2</sub> O] <sup>+</sup> , 441.3708 [M+H-2H <sub>2</sub> O-Glu] <sup>+</sup>	99.91
25	芦丁	10.71	$C_{27}H_{30}O_{16}$	611.1615 [M+H] <sup>+</sup>	465.1027 [M+H-Rha] <sup>+</sup> , 303.0486 [M+H-Glu-Rha] <sup>+</sup>	89.84
26	山奈酚-3, 7-二-O-葡萄糖苷	10.72	$C_{27}H_{30}O_{16}$	611.1615 [M+H] <sup>+</sup>	449.0986 [M+H-Glu] <sup>+</sup> , 287.0559 [M+H-2Glu] <sup>+</sup>	71.61

(续表 1)

No.	Compound	$t_R$ /min	Formula	$m/z$	MS/MS fragment	Score
27	芹菜素-2-O-葡萄糖苷	11.73	$C_{27}H_{30}O_{15}$	595.1656 [M+H] <sup>+</sup>	433.1154 [M+H-Glu] <sup>+</sup> , 121.0279 [C <sub>7</sub> H <sub>5</sub> O <sub>2</sub> ] <sup>+</sup>	73.69
28	异鼠李素-3-O-(2-O-鼠李糖基)- 葡萄糖苷	11.98	$C_{28}H_{32}O_{16}$	625.1765 [M+H] <sup>+</sup>	479.1150 [M+H-Rha] <sup>+</sup> , 147.0639 [Rha+H] <sup>+</sup>	71.43
29	$\alpha$ -香树脂酮醇/异构体	13.50	$C_{30}H_{48}O_2$	441.3726 [M+H] <sup>+</sup>	423.3610 [M+H-H <sub>2</sub> O] <sup>+</sup> , 235.2021 [C <sub>16</sub> H <sub>26</sub> O+H] <sup>+</sup> , 191.1791 [C <sub>14</sub> H <sub>22</sub> +H] <sup>+</sup>	82.91
30	没药萜 N/异构体	13.86	$C_{15}H_{18}O_3$	247.1327 [M+H] <sup>+</sup>	229.1206 [M+H-H <sub>2</sub> O] <sup>+</sup> 91.0552 [C <sub>7</sub> H <sub>7</sub> ] <sup>+</sup>	97.48
31	$\alpha$ -香树脂酮醇/异构体	13.92	$C_{30}H_{48}O_2$	441.3719 [M+H] <sup>+</sup>	423.3633 [M+H-H <sub>2</sub> O] <sup>+</sup> , 235.2073 [C <sub>16</sub> H <sub>26</sub> O+H] <sup>+</sup>	81.84
32	人参皂苷 Rd/异构体	14.02	$C_{48}H_{82}O_{18}$	991.5488 [M+HCOO] <sup>-</sup>	783.4926 [M-H-Glu] <sup>-</sup>	73.14
33	人参皂苷 Rf/异构体	14.09	$C_{42}H_{72}O_{14}$	845.4909 [M+HCOO] <sup>-</sup>	781.4764 [M-H-H <sub>2</sub> O] <sup>-</sup> , 637.4341 [M-H-Glu] <sup>-</sup> , 475.3816 [M-H-2Glu] <sup>-</sup> , 391.2836 [C <sub>24</sub> H <sub>40</sub> O <sub>4</sub> -H] <sup>-</sup>	79.99
34 <sup>*</sup>	三七皂苷 R1	15.28	$C_{47}H_{80}O_{18}$	955.5256 [M+Na] <sup>+</sup>	801.5018 [M+H-Rib] <sup>+</sup> , 783.4910 [M+H-Rib-H <sub>2</sub> O] <sup>+</sup> , 753.4789 [M+H-Glu-H <sub>2</sub> O] <sup>+</sup> , 621.4368 [M+H-Glu-H <sub>2</sub> O-Rib] <sup>+</sup> , 459.3843 [M+H-2Glu-H <sub>2</sub> O-Rib] <sup>+</sup>	86.94
35 <sup>*</sup>	人参皂苷 Rg1	15.86	$C_{42}H_{72}O_{14}$	801.5029 [M+H] <sup>+</sup>	783.4903 [M+H-H <sub>2</sub> O] <sup>+</sup> , 765.4784 [M+H-2H <sub>2</sub> O] <sup>+</sup> , 639.4471 [M+H-Glu] <sup>+</sup> , 621.4371 [M+H-Glu-H <sub>2</sub> O] <sup>+</sup> , 603.4278 [M+H-Glu-2H <sub>2</sub> O] <sup>+</sup> , 459.3834 [M+H-2Glu-H <sub>2</sub> O] <sup>+</sup> , 441.3744 [M+H-2Glu-2H <sub>2</sub> O] <sup>+</sup>	83.28
36	人参皂苷 F3/异构体	15.90	$C_{41}H_{70}O_{13}$	815.4787 [M+HCOO] <sup>-</sup>	637.4397 [M-H-Rib] <sup>-</sup> , 619.4268 [M-H-Rib-H <sub>2</sub> O] <sup>-</sup> , 475.3705 [M-H-Rib-Glu] <sup>-</sup>	74.08
37	$\alpha$ -香树脂酮醇/异构体	16.22	$C_{30}H_{48}O_2$	441.3683 [M+H] <sup>+</sup>	423.3636 [M+H-H <sub>2</sub> O] <sup>+</sup> , 191.1786 [C <sub>14</sub> H <sub>22</sub> +H] <sup>+</sup>	76.65
38	人参皂苷 Rf/异构体	18.64	$C_{42}H_{72}O_{14}$	845.4936 [M+HCOO] <sup>-</sup>	637.4366 [M-H-Glu] <sup>-</sup> , 475.3860 [M-H-2Glu] <sup>-</sup>	74.10
39	没药萜 N/异构体	18.76	$C_{15}H_{18}O_3$	247.1340 [M+H] <sup>+</sup>	229.1226 [M+H-H <sub>2</sub> O] <sup>+</sup> , 109.0659 [C <sub>7</sub> H <sub>9</sub> O] <sup>+</sup> , 91.0541 [C <sub>7</sub> H <sub>7</sub> ] <sup>+</sup>	89.61
40	松香酸/异构体	18.81	$C_{20}H_{30}O_2$	303.2306 [M+H] <sup>+</sup>	257.2276 [M+H-HCOOH] <sup>+</sup> , 187.1456 [C <sub>14</sub> H <sub>18</sub> +H] <sup>+</sup>	76.32
41	$\alpha$ -香树脂酮醇/异构体	19.08	$C_{30}H_{48}O_2$	441.3724 [M+H] <sup>+</sup>	423.3615 [M+H-H <sub>2</sub> O] <sup>+</sup> , 191.1781 [C <sub>14</sub> H <sub>22</sub> +H] <sup>+</sup>	72.33
42	人参皂苷 F3/异构体	19.08	$C_{41}H_{70}O_{13}$	815.4811 [M+HCOO] <sup>-</sup>	637.4340 [M-H-Rib] <sup>-</sup> , 619.4253 [M-H-Rib-H <sub>2</sub> O] <sup>-</sup> , 475.3816 [M-H-Rib-Glu] <sup>-</sup> , 457.3758 [M-H-Rib-Glu-H <sub>2</sub> O] <sup>-</sup>	72.39
43	没药萜 N/异构体	19.24	$C_{15}H_{18}O_3$	247.1318 [M+H] <sup>+</sup>	229.1215 [M+H-H <sub>2</sub> O] <sup>+</sup> , 91.0547 [C <sub>7</sub> H <sub>7</sub> ] <sup>+</sup>	90.61
44	松香酸/异构体	19.69	$C_{20}H_{30}O_2$	303.2365 [M+H] <sup>+</sup>	257.2271 [M+H-HCOOH] <sup>+</sup> , 187.1469 [C <sub>14</sub> H <sub>18</sub> +H] <sup>+</sup>	74.71
45	脱氢松香酸	19.76	$C_{20}H_{28}O_2$	301.2160 [M+H] <sup>+</sup>	255.2115 [M+H-HCOOH] <sup>+</sup> , 69.0696 [C <sub>5</sub> H <sub>8</sub> +H] <sup>+</sup>	74.86
46	人参皂苷 F2/异构体	19.93	$C_{42}H_{72}O_{13}$	829.4979 [M+HCOO] <sup>-</sup>	621.4374 [M-H-Glu] <sup>-</sup>	75.97
47	羽扇豆酮/异构体	20.08	$C_{30}H_{48}O$	425.3771 [M+H] <sup>+</sup>	407.3658 [M+H-H <sub>2</sub> O] <sup>+</sup> , 369.3130 [M+H-C <sub>4</sub> H <sub>8</sub> ] <sup>+</sup>	78.62
48	人参皂苷 F1/异构体	20.16	$C_{36}H_{62}O_9$	683.4411 [M+HCOO] <sup>-</sup>	475.3832 [M-H-Glu] <sup>-</sup> , 391.2835 [C <sub>24</sub> H <sub>40</sub> O <sub>4</sub> -H] <sup>-</sup>	72.23
49	$\alpha$ -香树脂酮醇/异构体	21.04	$C_{30}H_{48}O_2$	441.3729 [M+H] <sup>+</sup>	423.3604 [M+H-H <sub>2</sub> O] <sup>+</sup> , 191.1785 [C <sub>14</sub> H <sub>22</sub> +H] <sup>+</sup>	82.46

(续表1)

No.	Compound	$t_R$ /min	Formula	$m/z$	MS/MS fragment	Score
50	没药萜 N/异构体	21.09	$C_{15}H_{18}O_3$	247.1323 [M+H] <sup>+</sup>	229.1225 [M+H-H <sub>2</sub> O] <sup>+</sup> , 109.0641 [C <sub>7</sub> H <sub>9</sub> O] <sup>+</sup>	87.12
51 <sup>*</sup>	人参皂苷 Rb1	21.93	$C_{54}H_{92}O_{23}$	1131.5633 [M+Na] <sup>+</sup>	1091.6028 [M+H-H <sub>2</sub> O] <sup>+</sup> , 929.5469 [M+H-H <sub>2</sub> O-Glu] <sup>+</sup> , 767.4938 [M+H-H <sub>2</sub> O-2Glu] <sup>+</sup> , 443.3883 [M+H-H <sub>2</sub> O-4Glu] <sup>+</sup>	88.20
52	人参皂苷 F1/异构体	22.02	$C_{36}H_{62}O_9$	683.4402 [M+HCOO] <sup>-</sup>	475.3818 [M-H-Glu] <sup>-</sup>	74.66
53	(24S)-20, 24-环氧-25-羟基达玛树脂-3-酮	22.05	$C_{30}H_{50}O_3$	459.3823 [M+H] <sup>+</sup>	441.3743 [M+H-H <sub>2</sub> O] <sup>+</sup> , 423.3594 [M+H-2H <sub>2</sub> O] <sup>+</sup>	75.43
54	羽扇豆酮/异构体	22.36	$C_{30}H_{48}O$	425.3767 [M+H] <sup>+</sup>	407.3659 [M+H-H <sub>2</sub> O] <sup>+</sup> , 369.3134 [M+H-C <sub>4</sub> H <sub>8</sub> ] <sup>+</sup>	80.91
55	人参皂苷 Rd/异构体	22.39	$C_{48}H_{82}O_{18}$	991.5478 [M+HCOO] <sup>-</sup>	927.5232 [M-H-H <sub>2</sub> O] <sup>-</sup> , 783.4890 [M-H-Glu] <sup>-</sup> , 621.4362 [M-H-2Glu] <sup>-</sup>	73.69
56	没药萜 M	22.44	$C_{16}H_{22}O_4$	279.1615 [M+H] <sup>+</sup>	261.1390 [M+H-H <sub>2</sub> O] <sup>+</sup> , 247.1360 [M+H-CH <sub>3</sub> OH] <sup>+</sup>	71.84
57	羽扇豆酮/异构体	22.95	$C_{30}H_{48}O$	425.3766 [M+H] <sup>+</sup>	407.3669 [M+H-H <sub>2</sub> O] <sup>+</sup> , 369.3182 [M+H-C <sub>4</sub> H <sub>8</sub> ] <sup>+</sup>	83.52
58	人参皂苷 Rd/异构体	22.96	$C_{48}H_{82}O_{18}$	991.5514 [M+HCOO] <sup>-</sup>	621.4378 [M-H-2Glu] <sup>-</sup>	74.15
59	9-甲氧基没药萜/异构体	23.37	$C_{16}H_{18}O_3$	259.1326 [M+H] <sup>+</sup>	227.1057 [M+H-CH <sub>3</sub> OH] <sup>+</sup> , 109.0640 [C <sub>7</sub> H <sub>9</sub> O] <sup>+</sup>	77.67
60	人参皂苷 F2/异构体	24.63	$C_{42}H_{72}O_{13}$	829.4961 [M+HCOO] <sup>-</sup>	621.4437 [M-H-Glu] <sup>-</sup> , 459.3802 [M-H-2Glu] <sup>-</sup>	74.31
61	松香酸/异构体	24.72	$C_{20}H_{30}O_2$	303.2319 [M+H] <sup>+</sup>	187.1464 [C <sub>14</sub> H <sub>18</sub> +H] <sup>+</sup>	72.45
62	人参皂苷 F2/异构体	25.49	$C_{42}H_{72}O_{13}$	829.4934 [M+HCOO] <sup>-</sup>	621.4332 [M-H-Glu] <sup>-</sup>	75.63
63	鞣酮	25.60	$C_{19}H_{28}O_2$	289.2170 [M+H] <sup>+</sup>	121.0644 [C <sub>8</sub> H <sub>8</sub> O+H] <sup>+</sup>	73.92
64	9-甲氧基没药萜/异构体	25.81	$C_{16}H_{18}O_3$	259.1316 [M+H] <sup>+</sup>	244.1099 [M+H-CH <sub>3</sub> ] <sup>+</sup> , 227.1057 [M+H-CH <sub>3</sub> OH] <sup>+</sup>	72.57
65	松香酸/异构体	28.26	$C_{20}H_{30}O_2$	303.2301 [M+H] <sup>+</sup>	257.2263 [M+H-HCOOH] <sup>+</sup> , 187.1485 [C <sub>14</sub> H <sub>18</sub> +H] <sup>+</sup>	82.99
66	油酸/异构体	31.29	$C_{18}H_{34}O_2$	283.2635 [M+H] <sup>+</sup>	265.2544 [M+H-H <sub>2</sub> O] <sup>+</sup>	74.87
67	11-羰基-β-乳香酸	32.31	$C_{30}H_{46}O_4$	471.3470 [M+H] <sup>+</sup>	453.3363 [M+H-H <sub>2</sub> O] <sup>+</sup> , 435.3273 [M+H-2H <sub>2</sub> O] <sup>+</sup> , 425.3427 [M+H-HCOOH] <sup>+</sup> , 407.3310 [M+H-H <sub>2</sub> O-HCOOH] <sup>+</sup>	88.59
68	松香酸/异构体	33.23	$C_{20}H_{30}O_2$	303.2303 [M+H] <sup>+</sup>	257.2255 [M+H-HCOOH] <sup>+</sup> , 187.1489 [C <sub>14</sub> H <sub>18</sub> +H] <sup>+</sup>	79.06
69	亚麻酸	33.44	$C_{18}H_{30}O_2$	277.2183 [M-H] <sup>-</sup>	59.0145 [C <sub>2</sub> H <sub>4</sub> O <sub>2</sub> -H] <sup>-</sup>	73.21
70	α-乳香酸/异构体	33.52	$C_{30}H_{48}O_3$	455.3559 [M-H] <sup>-</sup>	437.3332 [M-H-H <sub>2</sub> O] <sup>-</sup> , 409.3527 [M-H-HCOOH] <sup>-</sup>	83.56
71	十六碳烯酸	34.52	$C_{16}H_{30}O_2$	253.2185 [M-H] <sup>-</sup>	235.2062 [M-H-H <sub>2</sub> O] <sup>-</sup>	73.24
72 <sup>*</sup>	亚油酸	37.89	$C_{18}H_{32}O_2$	279.2346 [M-H] <sup>-</sup>	59.0139 [C <sub>2</sub> H <sub>4</sub> O <sub>2</sub> -H] <sup>-</sup>	73.3
73	棕榈酸	38.55	$C_{16}H_{32}O_2$	255.2343 [M-H] <sup>-</sup>	237.2226 [M-H-H <sub>2</sub> O] <sup>-</sup>	76.88
74	α-乳香酸/异构体	39.24	$C_{30}H_{48}O_3$	455.3564 [M-H] <sup>-</sup>	409.3449 [M-H-HCOOH] <sup>-</sup>	81.23
75	油酸/异构体	39.39	$C_{18}H_{34}O_2$	283.2635 [M+H] <sup>+</sup>	265.2515 [M+H-H <sub>2</sub> O] <sup>+</sup> , 247.2431 [M+H-2H <sub>2</sub> O] <sup>+</sup>	72.41
76 <sup>*</sup>	11-羰基-β-乙酰乳香酸	39.50	$C_{32}H_{48}O_5$	513.3587 [M+H] <sup>+</sup>	495.3567 [M+H-H <sub>2</sub> O] <sup>+</sup> , 453.3368 [M+H-CH <sub>3</sub> COOH] <sup>+</sup> , 407.3311 [M+H-CH <sub>3</sub> COOH-HCOOH] <sup>+</sup>	98.15
77	α-乳香酸/异构体	42.09	$C_{30}H_{48}O_3$	455.3563 [M-H] <sup>-</sup>	437.3484 [M-H-H <sub>2</sub> O] <sup>-</sup> , 409.3441 [M-H-HCOOH] <sup>-</sup>	84.48

\*: confirmed by standard samples

### 2.3 各类化合物质谱裂解规律

为阐明表1中各类化合物的二级质谱裂解规律,分别选取色氨酸(9号)、羟基红花黄色素A(11号)、马钱苷酸(13号)、6-羟基山奈酚-3-芸香苷-6-葡萄糖苷(19号)、马钱子碱(21号)、人参皂苷Rg1(35号)、亚油酸(72号)和11-羰基-β-乙酰乳香酸(76号)为代表性成分进行讨论。

```

A
def find_candidates(self, exp_precursor_mz):
    """一级锚定"""
    if self.library is None: return pd.DataFrame()
    lib_mzs = self.library['Precursor_MZ'].values
    ppm_errors = np.abs(lib_mzs - exp_precursor_mz) / lib_mzs * 1e6
    mask = ppm_errors <= self.ms1_tol

    candidates = self.library[mask].copy()
    candidates['ppm_error'] = ppm_errors[mask]
    return candidates

def match_fragments_matrix(self, exp_mz, exp_int, theo_mz):
    """二级匹配(ppm)"""
    if len(exp_mz) == 0 or len(theo_mz) == 0: return []

    abs_diff_matrix = np.abs(exp_mz[:, np.newaxis] - theo_mz)
    ppm_matrix = (abs_diff_matrix / theo_mz) * 1e6
    match_mask = ppm_matrix <= self.ms2_tol

    exp_idx, theo_idx = np.where(match_mask)
    return [
        {
            'exp_mz': exp_mz[e],
            'exp_int': exp_int[e],
            'theo_mz': theo_mz[t],
            'mass_error': ppm_matrix[e, t]
        } for e, t in zip(exp_idx, theo_idx)]

B
def visualize_custom_mirror(self, matches, theo_mz_list, title_info, save_path, precursor_data=None):
    """生成镜像对比图"""
    frag_mz = [m['exp_mz'] for m in matches]
    frag_int = [m['exp_int'] for m in matches]
    base_peak_int = np.max(frag_int) if len(frag_int) > 0 else 1.0

    plot_mz, plot_int_norm, is_precursor_list = [], [], []

    # 处理碎片数据
    for mz, intensity in zip(frag_mz, frag_int):
        plot_mz.append(mz)
        plot_int_norm.append((intensity / base_peak_int) * 100)
        is_precursor_list.append(False)

    # 处理母离子数据
    if precursor_data:
        pre_mz, pre_real_int = precursor_data
        if not any(abs(m - pre_mz) < 0.01 for m in plot_mz):
            plot_mz.append(pre_mz)
            plot_int_norm.append((pre_real_int / base_peak_int) * 100)
            is_precursor_list.append(True)

    fig, ax = plt.subplots(figsize=(15, 9))
    bar_width = 0.4

    # 绘制实验谱图 (上方)
    for mz, intensity, is_pre in zip(plot_mz, plot_int_norm, is_precursor_list):
        display_height = 40 if (is_pre and intensity > 70) else intensity
        ax.bar(mz, display_height, width=bar_width, color='black')

        label_text = f"{mz:.4f}" if is_pre else f"{mz:.4f}\n(intensity:.1f)%"
        if display_height > 2:
            ax.text(mz, display_height + 2, label_text, ha='center', va='bottom', fontsize=12, color='black')

    # 绘制理论谱图 (下方)
    theo_y = np.full(len(theo_mz_list), -10)
    ax.bar(theo_mz_list, theo_y, width=bar_width, color='red')
    for mz in theo_mz_list:
        ax.text(mz, -12, f"{mz:.4f}", ha='center', va='top', fontsize=12, rotation=90, color='red')

    # 设置坐标轴与图例
    ax.axhline(0, color='black', linewidth=1)
    all_display_mzs = np.concatenate([plot_mz, theo_mz_list])
    if len(all_display_mzs) > 0:
        margin = max((np.max(all_display_mzs) - np.min(all_display_mzs)) * 0.1, 10)
        ax.set_xlim(np.min(all_display_mzs) - margin, np.max(all_display_mzs) + margin)

    ax.set_ylim(-40, 115)
    ax.set_yticks([0, 20, 40, 60, 80, 100])
    ax.set_ylabel("Relative Intensity (%)", fontsize=12)
    ax.set_xlabel("m/z", fontsize=12)
    ax.set_title(f"Match: {title_info['name']} | Score: {title_info['score']}", fontsize=16, fontweight='bold')

    ax.legend(handles=[
        Line2D([0], [0], color='black', lw=2, label='Experimental'),
        Line2D([0], [0], color='red', lw=2, label='Theoretical')
    ], loc='upper right', fontsize=11)

    plt.tight_layout()
    plt.savefig(save_path, dpi=300)
    plt.close()

```

图 3 本研究部分算法代码截图

Fig. 3 Screenshots of the partial algorithm codes in this study

A. primary anchoring and secondary matching; B. Matplotlib visualization (A. 一级锚定和二级匹配; B. Matplotlib 可视化)

2.3.1 氨基酸类 氨基酸类在正模式下碎裂常伴有 HCOOH 中性丢失, 还会丢失 NH<sub>3</sub>、H<sub>2</sub>O 等小分子, 产生的特征碎片可作为该类物质的定性依据。以色氨酸为例, 其准分子离子 *m/z* 205.097 4 脱去 C<sub>2</sub> 位氮

基产生碎片离子  $m/z$  188.0703, 随后失水产生碎片离子  $m/z$  170.0594。碎片离子  $m/z$  159.0911 是母离子丢失  $\text{HCOOH}$  (46.0055 Da) 所得。此外, 特征碎片  $m/z$  118.0651 表明该化合物中含典型的吡啶(苯并吡咯)结构<sup>[18]</sup>。经标准品比对, 确定其为色氨酸。

**2.3.2 色素类** 伤科接骨片中的色素类以红花黄色素类为主, 属查耳酮类化合物, 其结构中含多种官能团, 易与糖连接形成糖苷。以羟基红花黄色素 A 为例, 正模式下其准分子离子  $m/z$  613.1741 依次失水产生  $m/z$  595.1613 和 577.1589 的碎片离子。采用算法将这些特征碎片  $m/z$  的实测值与数据库中的理论值自动匹配, 结合相关文献鉴定该物质为羟基红花黄色素 A。

**2.3.3 环烯醚萜苷类** 环烯醚萜苷类常发生糖基丢失及其他碎裂, 可作为该类化合物的定性依据。如马钱苷酸在负模式下断裂糖苷键脱去外侧的葡萄糖基, 继而失水产生  $m/z$  195.0649 的特征碎片。此外, 碎片离子  $m/z$  331.1373 是其母离子脱羧丢失  $\text{CO}_2$  所得,  $m/z$  161.0450 为负模式下葡萄糖基的特征离子。经标准品比对, 确定该物质为马钱苷酸, 其实测/理论二级质谱图见图4。

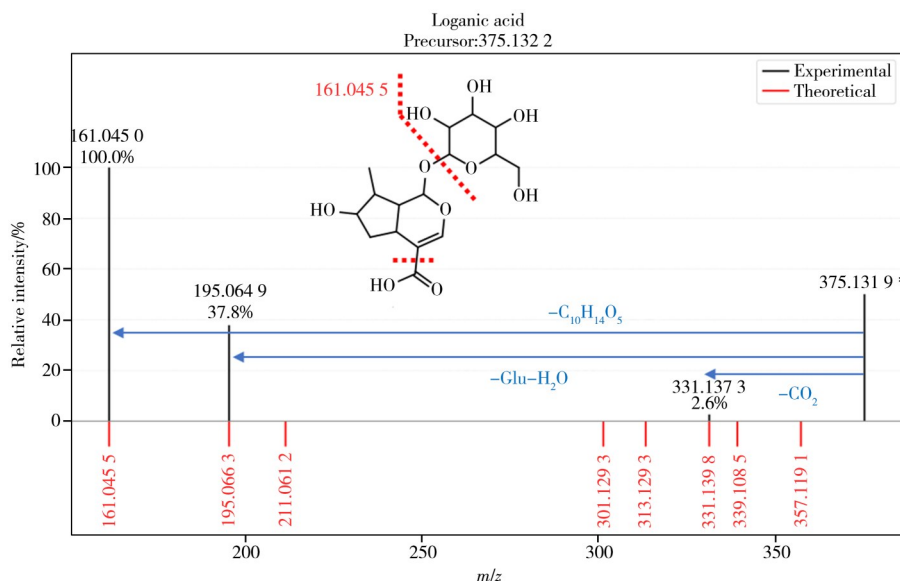


图4 马钱苷酸的实测(上)与理论(下)二级质谱图

Fig. 4 Experimental(up) and theoretical(down) MS/MS spectra of loganic acid

**2.3.4 黄酮苷类** 黄酮苷类是由2个苯环通过3个碳原子相互连接衍生出的系列化合物, 以  $\text{C}_6-\text{C}_3-\text{C}_6$  为基本骨架, 在伤科接骨片中主要以糖苷形式存在。以6-羟基山柰酚-3-芸香苷-6-葡萄糖苷为例, 正模式下其准分子离子  $m/z$  773.2124 分别丢失葡萄糖基和鼠李糖基, 产生  $m/z$  611.1584 和 627.1564 的碎片离子。而碎片离子  $m/z$  465.1034 由母离子连续丢失葡萄糖基和鼠李糖基后产生, 证实其结构中含葡萄糖基和鼠李糖基。依据上述讨论和相关文献<sup>[19]</sup>, 鉴定该物质为6-羟基山柰酚-3-芸香苷-6-葡萄糖苷。

**2.3.5 生物碱类** 伤科接骨片中的生物碱类以吡啶型为主, 在正模式下通常发生  $\text{C}_2\text{H}_4$ 、 $\text{C}_9\text{H}_{13}\text{NO}$  等中性丢失<sup>[20]</sup>。如马钱子碱的  $[\text{M}+\text{H}]^+$  母离子  $m/z$  395.1954 分别丢失  $\text{C}_2\text{H}_4$  和  $\text{C}_9\text{H}_{13}\text{NO}$ , 形成  $m/z$  367.1666 和  $m/z$  244.0974 的特征碎片, 表明其具有吡啶型生物碱结构。经标准品比对, 确定该化合物为马钱子碱, 其实测/理论二级质谱图见图5。

**2.3.6 皂苷类** 达玛烷型四环三萜类皂苷是伤科接骨片中一类重要活性成分, 主要包括原人参二醇及三醇型等, 具有相似的骨架结构及质谱裂解规律<sup>[21]</sup>。如正模式下人参皂苷 Rg1 的准分子离子  $m/z$  801.5029 依次失水, 分别得到碎片离子  $m/z$  783.4903 和 765.4784。特征碎片  $m/z$  639.4471 由母离子丢失1个葡萄糖基所得, 随后依次失水, 分别产生碎片离子  $m/z$  621.4371 和 603.4278。此外, 碎片离子  $m/z$  459.3834 和 441.3744 由母离子丢失2个葡萄糖基后, 再依次失水产生。依据上述讨论结合标准品比对, 确定该物质为人参皂苷 Rg1, 其实测/理论二级质谱图见图6。

**2.3.7 脂肪酸类** 脂肪酸类以长链烷烃和羧基为结构特征, 在负模式下响应更好。伤科接骨片中脂肪酸类结构的区别主要为链长和不饱和度的差异。因此, 通常将中性丢失长链烷基所产生的特征碎片作为该类物质的定性依据。如亚油酸的准分子离子  $m/z$  279.2346 经长碳链断裂, 产生高丰度的碎片离

子  $m/z$  59.0139 ( $C_2H_3O_2^-$ )。经标准品比对, 确定该化合物为亚油酸。

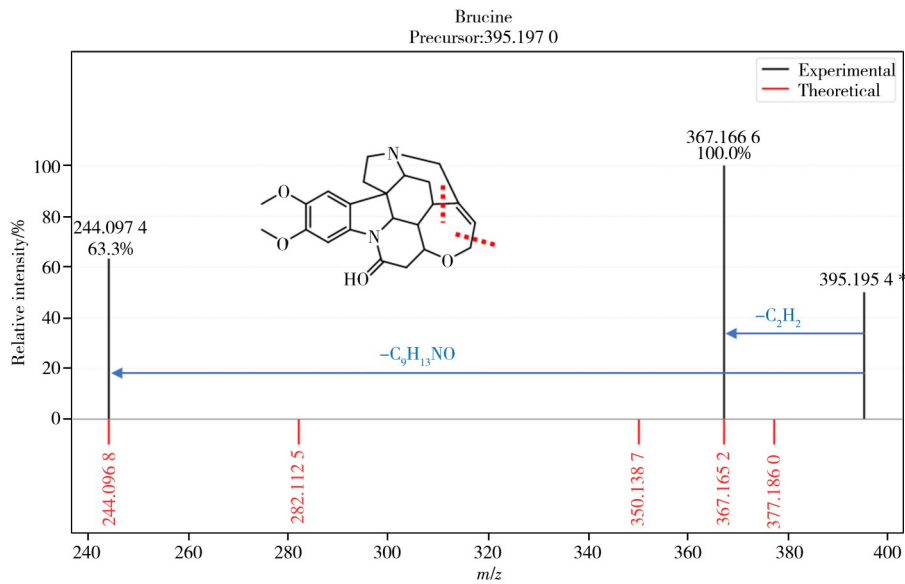


图 5 马钱子碱的实测(上)与理论(下)二级质谱图

Fig. 5 Experimental(up) and theoretical(down) MS/MS spectra of brucine

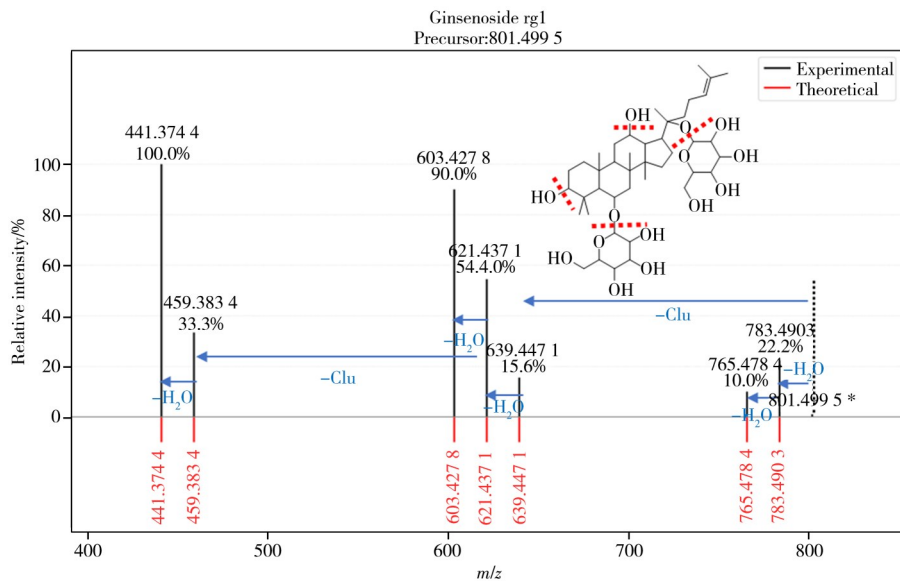


图 6 人参皂苷Rg1的实测(上)与理论(下)二级质谱图

Fig. 6 Experimental(up) and theoretical(down) MS/MS spectra of ginsenoside Rg1

**2.3.8 三萜类** 伤科接骨片中的三萜类主要源于乳香药材(代表性活性成分为乳香酸类), 其主要母核结构包括齐墩果烷型、乌苏烷型和甘遂烷型等。参考前期文献<sup>[22]</sup>, 乳香酸类在正模式下常会产生  $[M+H-CH_3COOH]^+$  特征离子, 可作为该类物质的定性依据。以 11-羰基- $\beta$ -乙酰乳香酸为例, 其准分子离子  $m/z$  513.3565 丢失乙酰基得到碎片离子  $m/z$  453.3341, 进而丢失 HCOOH 产生碎片离子  $m/z$  407.3311。此外, 母离子还可失水得到碎片离子  $m/z$  495.3567。最终经标准品比对, 鉴定该化合物为 11-羰基- $\beta$ -乙酰乳香酸, 其实测/理论二级质谱图见图 7。

## 2.4 Python 解析算法性能验证

**2.4.1 检出率** 为评估算法鉴定的准确性, 实验采集了 9 个混合标准溶液的 UPLC-Q-TOF-MS/MS 数据, 并输入 Python 算法中运行。结果显示, 算法以 100% 检出率成功匹配这 9 个物质, 且综合评分  $S$  均大于 85, 未出现漏检。

**2.4.2 重复性** 为评估算法的重复性, 实验连续采集了 5 次伤科接骨片样品数据, 分别输入 Python 算法中运行。结果显示, 在 5 个独立数据集中, 算法均输出了相同的 79 个化合物(表 1 中及 2 个假阳性),

其综合评分 $S$ 的相对标准偏差(RSD)均小于3.0%。

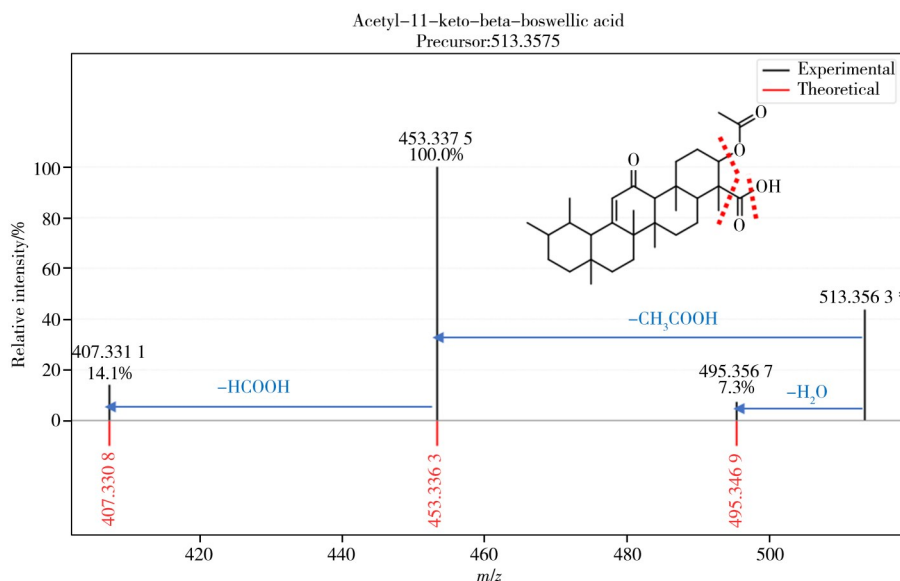


图7 11-羧基- $\beta$ -乙酰乳香酸的实测(上)与理论(下)二级质谱图

Fig. 7 Experimental(up) and theoretical(down) MS/MS spectra of acetyl-11-keto-beta-boswellic acid

**2.4.3 误匹配率** 对算法输出结果进行了人工复核,发现2个违背色谱流出规律的假阳性匹配。如算法在8.53 min处同时匹配到马钱子碱及其氮氧化物,而在反相色谱中,马钱子碱氮氧化物应早于其对应的原生物碱出峰。实验发现,马钱子碱氮氧化物极易发生ESI源内裂解脱去氧原子,产生与马钱子碱完全一致的准分子离子和二级碎片。据此计算,本算法解析的误匹配率仅为2.5%。

## 2.5 本方法与现有方法对比

文献调研显示,目前仅采用高效液相色谱-紫外检测器联用(HPLC-UV)定量了伤科接骨片中3~4个指标成分<sup>[8, 10]</sup>,尚无该复方全成分的液相色谱-质谱联用技术(LC-MS)鉴定报道。此外,还将本方法与其他中药复方UPLC-Q-TOF-MS/MS分析方法进行了比较,结果见表2。

表2 伤科接骨片现有分析方法与本方法的对比

Table 2 Comparison between existing analytical methods of Shangke Jiegu tablet and our method

Evaluation item	Existing analytical method	This method
分析对象	瘀血痹胶囊 <sup>[23]</sup> 、四逆散 <sup>[24]</sup>	伤科接骨片
数据解析模式	人工经验解谱及查阅文献	高通量自动匹配及结果可视化
分析成分数目	< 60个	8大类、77个
数据分析周期	数天以上/单个样品	< 1 min/单个样品
解析准确率	错检/漏检率高	基于量化加权评分,准确率高

由表可见,传统LC-MS数据人工解析主要依赖文献和经验,存在低效耗时、错检/漏检率高等缺陷。本方法通过引入Python自动解析算法,大幅提升了数据解析效率和鉴定数量。利用量化加权评分模型,有效排除了假阳性干扰,保证了鉴定准确率,为伤科接骨片及其他中药复方的成分解析提供了新范式。

## 3 结论

本研究针对中药复方成分复杂、解析困难的现状,建立了一种基于UPLC-Q-TOF-MS/MS技术结合Python算法的伤科接骨片多成分快速筛查与自动鉴定的新策略。该方法摒弃了传统低效的人工解谱模式,通过自建化学成分分子式库和特征碎片离子库,采用自编Python脚本实现了目标化合物的快捷精准匹配。从伤科接骨片中鉴定出77个化合物,涵盖了氨基酸类(9个)、色素类(1个)、环烯醚萜苷类(2个)、黄酮苷类(9个)、生物碱类(6个)、皂苷类(16个)、脂肪酸类(6个)和三萜类(28个)8大类,并深入探讨了其质谱裂解规律,验证了该自动解析策略在中药复杂体系分析中的适用性。

尽管本研究在中药复杂成分解析方面展现出显著优势,但仍存在一定局限性:首先,本研究侧重

于成分的定性表征,未对复方中重要成分进行定量分析;其次,受限于质谱数据分析的固有缺陷,本算法对部分同分异构体的区分能力仍有待提升。未来研究将在此基础上进一步深化:一方面,将色谱保留时间参数引入机器学习算法中,突破异构体化合物的鉴定瓶颈;另一方面,本研究构建的“液质联用+算法解析”智能化框架具有高度可移植性,可拓展至其他中药复方成分解析与复杂生物样本的代谢组学分析中。

#### 参考文献:

- [1] National Pharmacopoeia Commissio. *Pharmacopoeia of the People's Republic of China*. Vol I. Beijing: China Medical Science Press(国家药典委员会. 中华人民共和国药典. 一部. 北京: 中国医药科技出版社), **2025**.
- [2] Zhang C, Zhang J Y, Chen L, Liu Y D, Li X J, Xu M Z, Wu Z Y, Xi H Y, Zhang N, Zhou B Q, Lin N, Zhang Y Q. *China J. Chin. Mater. Med.*(张楚, 张景媛, 陈凛, 刘毓东, 李晓洁, 胥明珠, 吴卓耘, 奚灏灏, 张宁, 周宝强, 林娜, 张彦琼. 中国中药杂志), **2025**, 50(17): 4888-4896.
- [3] Fan L, Zhao H Y, Xu M, Zhou L, Guo H, Han J, Wang B R, Guo D A. *J. Chromatogr. A*, **2009**, 1216(11): 2063-2070.
- [4] Li S J, Zhang M Y, Hou P Y, Zhang R W, Hou C Z, Bi K S, Chen X H. *RSC Adv.*, **2015**, 5(95): 77689-77698.
- [5] Xu J, Guo Y Q, Zhao P, Guo P, Ma Y G, Xie C F, Jin D Q, Gui L P. *Fitoterapia*, **2012**, 83(4): 801-805.
- [6] Cheng L, Yuan Z M, Lü J, Zhao L, Li R H, Li C Y. *Chin. Tradit. Pat. Med.*(程岚, 袁子民, 吕佳, 赵琳, 李瑞海, 李春艳. 中成药), **2005**, 27(11): 1275-1277.
- [7] Qiao L, Wang J, Yuan Z M. *China Med. Her.*(乔丽, 王静, 袁子民. 中国医药导报), **2009**, 6(31): 38-40.
- [8] Huang D Z. *Strait Pharm. J.*(黄德忠. 海峡药理学), **2010**, 22(12): 84-85.
- [9] Hu Y S, Zhang H W. *Prog. Mod. Biomed.*(胡毓诗, 张挥武. 现代生物医学进展), **2010**, 10(15): 2958-2960.
- [10] Xue H Y, Qin Z, Zou L. *World Latest Med. Inf.*(薛恒跃, 秦峥, 邹亮. 世界最新医学信息文摘), **2019**, 19(63): 326-329.
- [11] Zhu T Y, Li N, Dong Y X, Li Z, Zhao L Q Q, Chen J H, Zhao C Y, Xiong H. *J. Liq. Chromatogr. Relat. Technol.*, **2025**, 48(16/20): 623-636.
- [12] Jiang J J, Liu X C, Zhang Y Y, Zhang M M, Liu J. *J. Instrum. Anal.*(姜娟娟, 刘晓闯, 张艳艳, 张明媚, 刘健. 分析测试学报), **2024**, 43(6): 814-825.
- [13] Xiao J, Song N N, Lu T, Pan Y N, Song J Y, Chen G, Sun L, Li N. *J. Pharm. Biomed. Anal.*, **2018**, 156: 284-296.
- [14] Want E. *Bioanalysis*, **2009**, 1(4): 805-819.
- [15] Huang F Q, Dong X S, Yin X J, Fan Y, Fan Y M, Mao C C, Zhou W. *J. Chromatogr. A*, **2020**, 1625: 461296.
- [16] Li T, Zhou Z Z, Zhang K, Ma W, Chen W, Tu P F, Li J, Song Q Q, Song Y L. *J. Pharm. Biomed. Anal.*, **2021**, 204: 114281.
- [17] Chen Z, Yan G L, Yang Y, Sun H, Zhang A H, Han Y, Wang S G, Wang X J. *J. Sep. Sci.*, **2022**, 45(18): 3431-3442.
- [18] Chen Y B, Rao B, Qin L. *Amino Acids Biotic Resour.*(陈永波, 饶斌, 覃兰. 氨基酸和生物资源), **2000**, (1): 55-58.
- [19] Jin Y, Xiao Y S, Zhang F F, Xue X Y, Xu Q, Liang X M. *J. Pharm. Biomed. Anal.*, **2008**, 46(3): 418-430.
- [20] Gu L Q, Hou P Y, Zhang R W, Liu Z Y, Bi K S, Chen X H. *J. Chromatogr. B*, **2016**, 1033: 157-165.
- [21] Xing R, Zhou L J, Xie L, Hao K, Rao T, Wang Q, Ye W, Fu H X, Wang X W, Wang G J, Liang Y. *Anal. Chim. Acta*, **2015**, 867: 56-66.
- [22] Wang Z J, Wang L. *Phys. Test. Chem. Anal.: Chem. Anal.*(王志嘉, 王璐. 理化检验-化学分册), **2017**, 53(4): 425-428.
- [23] Zhang X Y, Dong X Y, Zhang R H, Zhou S F, Wang W, Yang Y, Wang Y F, Yu H J, Ma J, Chai X. *Heliyon*, **2024**, 10(7): e28736.
- [24] Tian T, Xu X, Li X, Zhang W H, Lu H T. *J. Pharm. Biomed. Anal.*, **2021**, 194: 11381.

(责任编辑: 丁 岩)