

质谱分子网络在天然产物结构研究中的应用

覃舒然, 刘海翠, 李大山*, 王文静*

云南中医药大学中药学院, 昆明 650500

摘要: 质谱分子网络 (molecular networking) 常用于新颖天然产物发现及鉴定、药物代谢及药物研发等研究。目前, 分子网络能够将具有相似二级质谱的化合物连接成簇, 在天然产物分析中得到了广泛的应用。本综述将简单介绍分子网络的基本概念和构建分子网络的步骤, 以及从化合物的鉴定及发现; 鉴定天然产物的代谢产物; 天然产物化学成分的定性及定量; 天然产物活性成分筛选、分离与制备; 天然产物的生物合成途径五个方面对分子网络的应用进行阐述。

关键词: 质谱分子网络; 天然产物; 代谢产物; 化学成分; 结构研究

中图分类号: O657.63 **文献标识码:** A

Application of mass spectrometry molecular networking in the study of natural product structure

QIN Shu-ran, LIU Hai-cui, LI Da-shan*, WANG Wen-jing*

College of Traditional Chinese Medicine, Yunnan University of Chinese Medicine, Kunming 650500, China

Abstract: Mass spectrometry molecular networking is often used in the discovery and identification of novel natural products, drug metabolism and drug research and development. At present, molecular networks can connect compounds with similar secondary mass spectra into clusters, which have been widely used in the analysis of natural products. This review will briefly introduce the basic concepts of molecular networks and the steps of constructing molecular networks, as well as expounds the application of molecular network from five aspects: the identification and discovery of compounds; identification of metabolites of natural products; qualitative and quantitative analysis of chemical composition of natural products; screening, separation and preparation of active components of natural products; the biosynthetic pathway of natural products.

Key words: mass spectrometry molecular network; natural products; metabolites; chemical composition; structural research

化合物的结构分析对于天然产物的发现至关重要, 特别是对于珍稀或名贵中药材的物质基础揭示。然而, 高度多样化的化合物结构解析需要耗费大量的时间和人力, 这是对天然产物进行表征的最大挑战之一, 因此将质谱分子网络技术应用于天然产物的结构研究中 [1]。

液相色谱串联质谱（LC-MS/MS）是代谢组学中最常用的分析方法之一，然而，解释这些复杂数据是天然产物研究的一大困难。近年来，分子网络等新型生物信息学方法的出现，为复杂基质中已知化合物的识别提供了新的思路和角度。其中全球天然产物分子网络集群数据库（Global Natural Products Social Molecular Networking, GNPS）是一个开放的串联质谱数据的数据库，是目前唯一能够实现分子网络的公共基础平台^[2]，能对 LC-MS/MS 生成的数据集进行分析^[3,4]。GNPS 中分子网络的可视化是将每个光谱表示为一个节点，而光谱到光谱的排列则表示节点之间的联系，可以于 GNPS 在线将相关分子图谱可视化为分子网络^[5]。GNPS 总共包括 235850 个谱图和 22644 个化合物。GNPS 不仅能够做到鉴定已知化合物、类似物及自动分析分子网络中化合物，同时还具备对多来源的二级质谱数据进行复制链接及存储管理的功能^[6]。

质谱分子网络（molecular networking, MN）是一个 MS/MS 数据组织和可视化平台。每个质谱被认为是一个向量，并使用余弦相似度与其他所有质谱进行比较，当两个质谱之间的相似度高于阈值时，它们就会在分子网络中连接在一起^[7,8]。2012 年 Pieter 教授首次提出分子网络技术^[9]，质谱分子网络是一种基于串联质谱分析（MS/MS）的方法，适用于各种天然产物的分析，其目的是快速识别已知化合物和确定各种各样未知的天然产物。简言之，该方法就是收集串联质谱数据，并基于片段相似性来构建网络^[10]。该技术已广泛应用于天然产物、代谢组学和药物发现等领域^[11,12]。

目前，Trivella 等^[13]和 Fox Ramos 等^[14]报道 MN 已被证明是一种非常有效的工具，可以快速识别复杂混合物中的天然产物并且有助于发现新的天然产物^[15]；而 Nothias-Esposito 等^[16]报道 MN 作为一种有效的工具用于以质谱为基础的代谢组学中来表征植物中特殊代谢物的生成；Kang 等^[17]运用 MN 能将次级代谢物分类为类似的簇，该技术在天然产物研究中被用于筛选分离目标；Lei 等^[18]采用 MN 观察到未加工附子及加工后附子的含量差异；除此之外，Tian 等^[6]报道了 MN 还可以清晰明了地观测到不同品种不同产地茶叶的化学成分的差别，从而对茶叶进行定性定量评价；通过分析碎片离子的类型和碎片离子的相对含量变化来分析天然产物分子结构的变化及其官能团的转化情况，来推断天然产物是否生成或转化。

1 质谱分子网络技术

MN 是基于串联质谱数据的组织和可视化，通过光谱相似图，揭示同源质谱片段的存在。由于结构类似的化合物具有相似的碎片光谱，它们的节点往往聚集在一起，形成类似

物簇^[19]。在相同的质谱分离条件下，若天然产物分子有类似的结构就会产生相似的质谱碎片离子，计算机可以根据系统的算法对进入高分辨质谱后形成的二级质谱碎片进行相似度计算，根据碎片相似度的高低，然后依照相似度将质谱碎片图整合为一种能够可视化的网络图谱。在 MN 中，密切相关的天然产物均具有相似的二级质谱片段，并根据相同的生物合成来源对这些天然产物进行分组，由此产生的簇允许对相同和类似分子进行可视化搜索，通过与天然产物数据库的比较，快速对已知代谢物进行检测，并基于系统碎片化对类似物进行识别^[3, 20]。MN 可以提供一种快速、灵敏的方法来同时处理分析某植物提取物的化学成分，有效节约人力时间和精力。

MN 在过去 5 年已经发展起来，成功地应用于天然产物化学结构研究，比如化合物鉴定、代谢物分离、生物研究、基因组测序和探索微生物菌株的优先排序。这种技术与 GNPS 结合，能开放地获取数据库，用于对 MS/MS 谱数据共享，处理，加强特定化学类别和化合物的识别，并有助于对样品进行进一步研究^[13]。

GNPS 平台利用串联质谱 (MS/MS) 数据对化合物进行注释，允许具有高度相似的质谱的节点组成簇，有助于使代谢组学中化学结构可视化^[21]。分子网络与特征检测方法相结合形成了基于特征的分子网络 (feature-based molecular networking, FBMN)。Phelan^[22]提供了使用 GNPS 和 FBMN 生成分子网络以注释微生物天然产物的方法。但 FBMN 对骨架差异显著的新化合物的有效鉴定仍然是困难的。He 等^[23]提出了一个基于生源砌块的分子网络 (blocks-based molecular network, BBMN) 策略，以进一步探索一叶萩型生物碱的未知化学空间，这一策略的应用能高效发现和鉴定 *Flueggea suffruticosa* 中的三种新型一叶萩型生物碱，即 suffranidines A~C。与经典 MN 和 FBMN 相比，BBMN 策略具有以下优点：1) 高选择性：选择性地识别有特定生物遗传特征片段的相关化合物；2) 更快速：将庞大的 MS² 数据优化为更简洁、更有针对性的数据，使后续的数据分析更加快速。这一策略不仅可以有效地发现植物中新的天然产物，而且可以广泛地应用于探索动物、微生物和海洋生物等。

2 建立质谱分子网络的步骤及其软件

建立 MN 主要有以下几个步骤：对样品进样分析，将进样后所得到的质谱数据用 Proteo Wizard 软件转化为 mzML、mzXML 和 mgf 等通用格式文件，但因处理后依然会有一些干扰离子影响分子网络的质量，所以需采用 Mzmine2 软件再次处理，上传至 GNPS 数据分析平台，通过计算机计算比较样品中所有化合物、标准化合物和 GNPS 收录的化合物

二级质谱图之间的相似度，通常以余弦值表示（相似度越大，则余弦值越大），然后根据相似度大小将所有分子离子数据整合成一张可视化的网络图谱（见图 1）。分子网络图中可直接看到样品中所有未知化合物与标准化合物之间的结构相似度，使用 Cytoscape 软件可视化分子网络，可以将这些网络具有的功能注释和数据库链接在一起研究，网络图中的节点代表分子，其中各节点之间由不同粗细的线段相连接，而这些线段代表这些分子之间相互作用的强弱^[24, 25]。



图 1 建立 MN 的步骤

Fig. 1 Steps of establishing MN

2.1 Proteo Wizard 软件简介

Proteo Wizard 是一款模块化用于质谱数据读取与格式转化的开放源代码软件，Proteo Wizard 可将质谱数据转换为 mzML、mzXML 和 mgf 等基于文本的通用格式，用于表示描述扫描数、前体和二级质谱离子和产生分子所需强度的质谱数据^[10]。

2.2 Mzmine2 软件简介

Mzmine2 是一个主要针对 LC-MS 数据，用于质谱数据处理的开源软件。Mzmine2 用于对批量或元 mzXML 文件来进行数据处理，把处理后的数据生成 csv 和 mgf 文件，并导出用于分子网络^[10]，且使用 Mzmine2 可以识别出同分异构体^[26]。Perruchona 等证明 Mzmine2 与分子网络的结合可以加速生物样品的去复制，进而关注潜在的未知化合物^[27]。

2.3 Cytoscape 软件简介

Cytoscape 是一个开放源代码软件平台，用于可视化分子相互作用网络和生物途径，并将这些网络与注释、基因表达谱和其他状态数据集成在一起，根据片段相似度对波谱进行聚集后，将数据导入到 Cytoscape 中，显示为节点和边缘组成的网络^[3]。Cytoscape 最初是为生物学研究设计的，现在已发展为一个复杂网络分析和可视化的通用平台。Cytoscape 核心分布为数据集成、分析和可视化提供了一组基本特征。

3 质谱分子网络技术的应用

3.1 鉴定复杂体系中天然产物的结构和发现新颖结构

由于生物信息学和分析化学的快速发展，特别是质谱的发展促进了天然产物的研究，串联质谱（MS/MS）和 MN 的结合已被证明是一种非常有效的工具，可以快速识别复杂混合物中的天然产物，并且有助于新颖天然产物的发现^[13, 14, 23]。

Yuan 等^[28]将 UPLC-LTQ-Orbitrap MS 与 MN 相结合对铁线莲属植物中三萜皂苷类成分进行鉴别，可快速、可视化地区分 6 种植物中三萜皂苷的差异，其中长花铁线莲、美花铁线莲、准噶尔铁线莲和甘川铁线莲这 4 种铁线莲相似度较高，可在民族用药中替代入药；银叶铁线莲三萜皂苷类成分较少，需谨慎替代入药；钝萼铁线莲中无三萜皂苷类成分，不可替代入药。Huang 等^[29]在 MN 中识别出 7 个生物碱类型的粒子簇，藉此在夏天无中共鉴定和推测出 52 个化合物，其中有 21 个潜在的新化合物，6 个化合物为首次在该植物中发现。Shen^[30]建立酸枣仁已知成分的本地数据库，对其成分进行分类总结，建立 MN，通过 UHPLC-Q-Orbitrap MS 分析技术，构建了酸枣仁中化学成分快速识别的分析策略，鉴定了其提取物中的 90 个化合物，分别为 64 个黄酮类成分、21 个生物碱类成分以及 5 个皂苷类成分。Nothias-Esposito 等^[16]采用 MN 比较了 *Euphorbia pithyusa* 和 *E. cupanii* 亲缘关系较近的两种大戟属植物中二萜类化合物分布情况，发现 10 个曼西醇型化合物，6 个 premyrsinol 衍生物和 4 个佛波醇酯。Pan 等^[31]以中药金银花为例，应用 MN 对其复杂多类型成分进行综合鉴定，从中共鉴定出 537 个化合物，其中包括类黄酮和咖啡酰奎宁酸的新型二聚体等潜在的新结构。Hu 等^[32]应用 MN 成功地鉴定了烘焙咖啡豆中 7 个 II 型二萜类化合物，并且发现了一系列新颖的含氮的对映贝壳杉烷型二萜。Lyu 等^[33]应用 MN，对荔枝果肉提取物中的 9 种原花青素和 11 种黄酮苷进行了定性分析。Areas 等^[34]利用 UHPLC-HRMS 和分子网络对 *Lessingianthus brevifolius* (Less.) H. Rob. (Asteraceae) 进行了去重复研究，鉴定出其中的 15 个已知化合物，包括两种倍半萜内酯和 13 种黄酮类化合物。Ruan 等^[35]根据 MS/MS 碎片模式的相似性创建 GNPS 分子网络，运用 Cytoscape 3.6.1 软件筛选结构相似的分子簇，在五味子中共鉴定出 31 个联苯环辛二烯型木脂素类化学成分，其中的 13 个为首次鉴定。Jiang 等^[36]基于 HPLC-Q-TOF-MS/MS 数据的分子网络分析来研究草乌中的二萜生物碱，九个标准品用于总结阳离子模式下的碎片特征，进一步用于确定两组未知化合物的结构。Brel 等^[37]应用 MN，允许通过串联质谱数据集组织把特定化学成分作为目标，对 57 株昆虫病原微生物进行了筛选，结果表明，粉棒束孢 BSNB-1250 乙酸乙酯 (EtOAc) 提取物表现出抗菌活性和潜在的新化学成分：paecilosetin 及 5 个新的类似物。McManus 等^[38]利用 MN 对 *Trichodesmium thiebautii* 提取物进行分析，发现其中充满了卤化化合物。随后通

过质谱分离得到了异毛霉素 C 和毛霉素 G~I。这些新的物质与之前描述的毛霉素具有结构的变化，通过比较研究能够发现与小鼠神经母细胞瘤细胞毒性相关的结构特征。Rivera-Mondragón 等^[4]采用基于 UPLC-MS/MS 的 MN 与常规分离和核磁共振相结合的非靶向代谢组学方法对 *Crescentia cujete* 果肉进行植物化学分析，分离并鉴定出 66 个化学成分。de Oliveira 等^[3]基于非靶向串联质谱的 MN 快速鉴定 *Adenocalymma imperatoris-maximilianii* 萃取物中 63 个已知的黄酮类化合物（含同分异构体），还利用 MN 首次从该植物中推导得到含异戊二烯基的黄酮衍生物。Reginaldo 等^[39]通过 HR-LC-ESI-MS/MS 分析脱水（旱季在其自然栖息地收获）和水合（植物通过水合作用后在实验室培育）的 *Selaginella* spp. 标本，首次运用 MN 对 *Selaginella* spp. 中 39 个化学成分进行结构推定，主要为卷柏素和黄酮类化合物。Purves 等^[40]利用基于 HR-MS/MS 的 MN 对发酵提取物进行比较，结果显示，40% 和 42% 的母离子分别是由南极和苏格兰的细菌分离物产生的，只有 8% 的物质在这些地点中是共有的，表明生物地理高度影响其化学成分；研究所得到的 MN 中包含超过 3500 个母离子，质荷比范围为 m/z 149~2 558，说明产生了丰富的次生代谢产物。Song 等^[41]采用 MN 对异喹啉类生物碱进行聚类，寻找可能的化学成分。Wang 等^[42]基于 MS/MS 的分子网络来研究 *Aspergillus caelatus* 的化学成分，从而发现了新的二酮哌嗪二聚体和曲霉素。Freire 等^[43]通过基于特征的分子网络、去复制和分离，对海绵 *Agelas dispar* 的甲醇部分进行研究，发现了新的溴吡咯衍生物。Zhao 等^[44]基于 UPLC-Q-TOF-MS、多元统计分析和 MN 的非目标代谢组学方法在 *Morchella* sp. 与对照组之间鉴定出的 50 个差异化学成分，此外，基于分子网络分析，鉴定了三种 desferri-ferricrocin 及其铁螯合物的同源化合物，包括 desferri-ferrichrome, desferri-ferrichrome C 和一个新注释的同系物。Caldas 等^[45]考虑到 *Gymnopilus imperialis* 可能含有大量化学成分，GNPS 具有寻找化学成分、节省时间以及对天然产物的分离和表征方面的功能，从 *G. imperialis* 的甲醇、二氯甲烷和乙酸乙酯提取物中注释了 24 个低聚异戊二烯，其中 4 个来自 GNPS 光谱库匹配，20 个来自基于分子网络的预测。Luo 等^[46]基于 MS/MS 的分子网络被用于开发 *Xylaria* sp. 的肽天然产物，从而发现了一种新的环七肽。Hu^[47]通过 UPLC-Q-TOF-MS 技术结合诊断离子与分子网络分析策略，从当归与酒当归中鉴定出 126 个化合物，包括 28 个苯酚类，32 个有机酸类，8 个香豆素类，8 个黄酮类，16 个氨基酸类，31 个其它化合物。

Wu 等^[48]采用基于 GNPS 特征网络的方法对黄金芽绿茶中 9 个黄酮苷、6 个儿茶素和 3 个黄酮醇进行了跟踪并利用 UPLC-MS² 数据和 GNPS 分子网络技术从黄金芽绿茶中发现具

有乙酰胆碱酯酶抑制活性的新木质素苷类化合物。Zhang 等^[49]采用基于 GNPS 的 MN 分析中根据碎片结构的相似性，将灯盏细辛注射液中的化合物分为 2 大类别，分别是咖啡酰酯衍生物和黄酮类。Xu^[50]利用基于 GNPS 平台 MN 分析 *Myrothecium roridum* IFB-E091 发酵产物组分的化学成分，以新环十肽 MR-1 和 MR-2 作为种子节点进行分子网络分析，结果发现，组分 Fr.5-3 中存在环十肽化合物，除已分离得到的 4 个新环十肽 MR-1~MR-4 外，还存在 3 个环十肽 MR-5~MR-7。

此外，MN 还可以与空间分辨特征、双电离模式 MS/MS 等相结合对天然产物进行结构分析，Kuo 等^[1]在空间分辨特征的指导下，将 MN 应用于沉香的结构分析中，对 3 大类的 2-(2-苯乙基)色酮及其类似物进行了表征。Li 等^[51]将基于特征的分子网络 (FBMN) 和双电离模式 MS/MS 相结合从银杏叶提取物和银杏叶中成功鉴定出 95 个化合物，其中包括黄酮苷、银杏内酯、木脂素苷、萜类和酚类化合物。

研究者们常采用 LC-MS 分析天然产物中的化学成分，采集各类样品中化合物的质谱数据，再通过 PCA、PLS-DA 等得到的载荷图和散点图等对样品进行分类总结，而采用了 MN 就能对其进行可视化处理，快速鉴定，得出结论。

3.2 鉴定天然产物在生物体内的代谢产物

质谱与代谢物的化学结构相联系，MN 能够根据生物体内的代谢物的结构相似性进行分组，通过检测代谢物之间相似性的关键信息，鉴定生物样品的代谢物^[16, 52]。

采用 MN 可以快速地对生物体内的代谢产物进行鉴定。Gao 等^[53]将串联质谱与 MN 结合，通过 GNPS 观察蛹虫草提取物处理的白血病细胞和对照组的代谢情况，对药用真菌提取物和癌细胞的代谢产物结构进行解析，并半定量地识别其变化。Said 等^[10]基于 LC-MS/MS 的 MN 方法鉴定人类尿液中的可可酚代谢物，目的是提出一种有生物活性的酚类代谢物鉴定的方法。Xiong 等^[54]建立了 UPLC-HRMS 结合质谱分子网络分析血浆中代谢物的方法，共定性检测到 187 个代谢物。Farg 等^[55]通过 GNPS 平台的分子网络，利用 LC 串联质谱技术，在体外培养人肠道微生物群后，共鉴定出生咖啡豆中的 36 个代谢产物，包括 4 个未报道的烷基肉桂酸酯和 6 个新的生物转化产物。

此外，MN 还能跟药代动力学相结合进行相关研究，Wang 等^[56]将质谱数据整合到分子网络中，以发现天然药物 *Ardisia japonica* (Thunb.) Blume 的活性成分，在大鼠血浆中共识别出 8 个吸收的原型成分。Seo 等^[57]基于 MS/MS 的分子网络系统研究了柘树叶的提取物中生物活性成分的药代动力学特征，使用 LC-QTOF/MS 获得血浆样品的产物离子质量数据，

并进行分子网络分析，由此产生的分子网络图表明，槲皮素和山奈酚的葡萄糖醛酸代谢物是主要的循环物种。

3.3 对天然产物化学成分的定量分析

基于 TOrTE (tandem-MS origin tracing engine) 的 MN 可以快速分析不同来源样品粗提取物中的目标化合物，并对其定量表征^[6]。先把每个节点的原始数据用 TOrTE 转化成为“cluster_info.txt”格式的文件，再将这些文件转化成“mzXML”格式的文件，而每个节点的色谱峰用 Proteo Wizard 软件进行提取，计算每个峰的峰面积，最后合并峰面积，并用 Cytoscape 软件可视化饼状图。每个节点代表一种化学成分，节点中不同颜色表示某一化学成分的不同来源，颜色在每个节点中占面积的大小表示某一化学成分在不同来源样品中量的多少，即面积越大，其含量越大；而面积越小，其含量越小^[24]（见图 2）。可有效避免使用大量标准品进行定量的限制，提升实验结果的准确度。

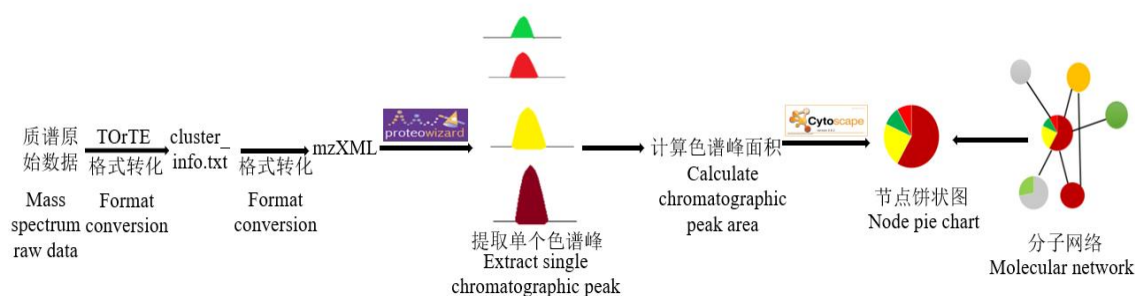


图 2 MN 的定量过程

Fig. 2 Quantitative process of MN

Lei 等^[18]以高通量、高灵敏度的 UPLC-Q-TOF/MS 为基础，采用 MN 对未加工附子及加工后附子的整体化学成分进行了综合检测和鉴定，鉴定出 145 个化合物，分析其含量差异。Winnikoff 等^[58]将培养的 20 种海洋蓝藻粗提物的 MS/MS 数据集导入 GNPS，生成分子网络，对几种具有生物活性和生物合成吸引力的脂肽 malyngamide C 进行定性和定量，其中产量最高的是巴拿马的 *Okeania hirsuta* (PAB10FEB10-01)，其含量达 2.4%。Calabrese 等^[59]基于特征的分子网络 (FBMN) 研究 *Zhanthoxylum heitzii* 中代谢物在植物器官中的分布，并发现潜在的新成分，总共检测到 143 个代谢物，分别为生物碱、木脂素、多酚、脂肪酸和氨基酸类的代谢物，并对不同器官进行了半定量分析。Ma 等^[60]运用液相色谱-串联质谱法以及基于质谱的分子网络进一步分析次生代谢产物的产量和多样性，在糙米培养基中分别加入金合欢醇 (1 $\mu\text{mol/L}$)、法尼焦磷酸铵盐 (5 $\mu\text{mol/L}$)、亚硫酸氢钠甲萘醌 (10 $\mu\text{mol/L}$)、3-氨基-L-酪氨酸 (1 mmol/L) 等 4 种诱导条件均可调控菌株 C23-3 次级代谢产

物发生较明显变化，其丁内酯、土震素和洛伐他汀类化合物的多样性和产量较为突出。Li 等^[61]基于分子网络，明显观察到炮制前后女贞子的化学成分在种类和含量上发生了很大变化。

3.4 天然产物中活性成分的筛选、分离与制备

MN 将不同的天然产物分类为类似的支架簇，该技术在天然产物研究中被用于筛选分离目标^[17]。

可将 MN 与生物活性结合对天然产物进行筛选分离：Nothias 等^[62]开发了以生物活性为基础的 MN，用于发现天然产物生物测定引导的分离中的药物先导物，提出了生物活性分子网络的概念。Olivon 等^[19]利用多信息分子网络分离具有生物活性的天然产物，使用大量的分子网络嵌入不同的信息层（生物活性和分类数据），以突出粗提物的化学多样性中潜在的生物活性支架，通过从两个植物来源（*Bocquillonia nervosa* 和 *Neoguillauminia cleopatra*）分离活性次生代谢产物瑞香烷型二萜原酸酯和 4 种 12-deoxyphorbols，它们抑制 Wnt 信号通路，并显示出对 CHIKV 病毒的抗病毒活性。Raheem 等^[63]采用代谢组学和生物活性指导的分离方法，并结合 MN，从英国风信子（*Hyacinthoides non-scripta*）中寻找具有抗锥虫体活性为 98.9%（20 $\mu\text{mol/L}$ ）的 norlanostane 型皂苷。Olivon 等^[64]通过多信息分子网络的方法，从 *Codiaeum peltatum* 中获得抗病毒化合物。Cabral 等^[65]基于 MN 靶向分离 *Conchocarpus fontanesianus* 中 6 个新的吡啶多喹啉类生物碱。Naman 等^[66]结合 MN 和生物活性测定指导下的分离，可以从 American Samoan 采集的 cf. *Symploca* sp. 样本中靶向分离出一种新的具有生物活性的环八肽——samoamide A，将生物活性筛选与 MN 的化学信息学策略相结合作为发现工具，加速了具有新的化学结构和预期生物活性的天然产物的靶向分离。Han 等^[67]基于 LC-MS/MS 分子网络的指导下，从 *Penicillium cellarum* 中分离出具有葡萄糖摄取刺激活性的七种新的被疣青霉素衍生物 penicicellarusins A~G，以及三种已知的类似物。

MN 与结构相结合对天然产物进行筛选分离：Wu 等^[68]提出了结构导向分离的概念，建立了一个抗病毒倍半萜分子网络，对 *Saussurea lappa* (Decne.) C.B Clarke 进行了研究，分离了七种新倍半萜，即 lappaterpenes A~G。Mi 等^[69]基于 MS/MS 质谱导向策略，将芫花花蕾的乙醇提取物通过硅胶粗分得到的馏分进行 LC-MS/MS 测试，将所获得的数据上传到 GNPS 质谱分析平台进行进一步分析，分子网络分析显示有明显与已知对照品化合物 yuanhuacine 节点连接的簇，这些节点则被视为瑞香烷型二萜类的衍生物，以此确定了富含

二萜类成分的馏分，最终分离得到了 4 个未被描述以及 8 个已知的瑞香烷型二萜。

MN 还能够与核磁共振波谱法（nuclear magnetic resonance spectroscopy, NMR）、层次聚类分析（hierarchical clustering analysis, HCA）和超临界流体色谱（supercritical fluid chromatography, SFC）等其他技术、方法相结合靶向分离目标化合物，尤其是与 NMR 结合：

表 1 MN 与 NMR、HCA 和 SFC 结合分离目标化合物

Table 1 MN was combined with NMR, HCA and SFC to separate the target compound

结合的技术	分离对象	分离的化合物
Combined technology	Separation object	Isolated compound
NMR	珊瑚源真菌产黄青霉	7 种新的 cyclohexadepsipeptide ^[70]
	藓类土壤源菌株可可链霉菌阿苏亚种 H2S5	ansamycin streacansamycin A~C ^[71]
	链霉菌 KY11784	skyllamycins A~C, 以及两个新的同属物 ^[72]
HCA	<i>Trachelospermum asiaticum</i>	木脂素 ^[73]
SFC	<i>Euphorbia semiperfoliata</i>	新的抗病毒二萜酯 ^[74]

采用 MN 与一些传统技术和方法相结合可以快速完成对复杂化学成分进行鉴定，对样品间的成分差异可视化，不仅能大大缩减了质谱、核磁等数据分析化合物结构的工作量，有效减少盲目性，而且能直观快速分析出化合物在不同样品中的分布。

3.5 天然产物的生物合成途径研究

Tian 等^[6]在 Cytoscape 软件上分析在绿茶和黑茶中茶氨酸和表没食子儿茶素没食子酸酯的相对含量变化，将非发酵茶的绿茶与发酵茶黑茶做对比发现，二者茶氨酸的相对含量均减小，而表没食子儿茶素没食子酸酯的相对含量却在增加。Kalinski 等^[75]基于 MN 分析，然后进行制备纯化，以绘制七个样本提取物中新的和已知的吡咯烷酮和相关化合物的多样性，分析了 *Tsitsikamma favus* 海绵的次级代谢产物库，其中两个样本表现出不同的形态和化学特征：不含 discorhabdins，只有微量的 tsitsikammamines，相反，有大量未分支和卤化的 makaluvamines，这项研究为推测 *latrunculid* 海绵中 pyrroloiminoquinone 的生物合成途径提供了新的线索。

上述研究通过运用 MN 分析不同基团的相对含量变化，推测各基团的取代和合成顺序，大大减少了生物合成途径的工作量，为研究天然产物的生物合成途径提供了参考。通过 MN 筛选出不同碎片离子，查找天然产物特征官能团的特征碎片离子，通过对比验证，建立相关性，并以此为方法通过分析碎片离子的种类变化和相对含量变化追踪天然产物分子结构变化及官能团的相互转化，从而推断天然产物的生成及转化。

4 结语

在过去的几年中，MN作为一种新兴工具^[43]被提出来用于简化天然产物的结构研究，已经成功应用于天然产物中化学成分及活性成分的发现、分离与制备、结构鉴定、定量分析。MN适用于对天然产物进行体内成分及其代谢产物的研究，可更好地阐释其药效物质基础和代谢过程^[24]。

运用MN比对特征碎片离子，分析其种类和相对含量变化，得知天然产物分子结构变化及官能团的相互转化，可用于推断天然产物的生源合成途径，但在文献调研中发现国内外相关报道尚不多见，这也许将会是MN用于天然产物研究的一个新方向。

综上所述，到目前为止，MN主要应用于天然产物的结构鉴定，其与其他技术相结合，是天然产物的结构研究特别是新颖结构和活性成分发现的一种高效快捷方法。体内代谢、定量分析以及生物合成途径方面的研究较少，是今后研究人员可以关注的方向。

参考文献

- 1 Kuo TH, et al. Mass spectrometry imaging guided molecular networking to expedite discovery and structural analysis of agarwood natural products[J]. *Anal Chim Acta*, 2019, 1080: 95-103.
- 2 Cheng TF, et al. Investigation of the chemical compounds in *Pheretima aspergillum* (E. Perrier) using a combination of mass spectral molecular networking and unsupervised substructure annotation topic modeling together with in silico fragmentation prediction[J]. *J Pharm Biomed Anal*, 2020, 184: 113197.
- 3 de Oliveira GG, et al. Dereplication of flavonoid glycoconjugates from *Adenocalymma imperatoris-maximilianii* by untargeted tandem mass spectrometry-based molecular networking[J]. *Planta Med*, 2017, 83: 636-646.
- 4 Rivera-Mondragón A, et al. UPLC-MS/MS-based molecular networking and NMR structural determination for the untargeted phytochemical characterization of the fruit of *Crescentia cujete* (Bignoniaceae)[J]. *Phytochemistry*, 2020, 177: 112438.
- 5 Wang MX, et al. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking[J]. *Nat Biotechnol*, 2016, 34: 828-837.
- 6 Tian MM. LC-MS/MS based molecular network analysis of the flavoalkaloids in six traditional tea categories[D]. Hefei: Anhui Agricultural University (安徽农业大学), 2018.
- 7 Chung HH, et al. Reaction tracking and high-throughput screening of active compounds in combinatorial chemistry by tandem mass spectrometry molecular networking[J]. *Anal Chem*, 2021, 93: 2456-2463.
- 8 Hollywood KA, et al. Metabolomics tools for the synthetic biology of natural products[J]. *Curr Opin Biotechnol*, 2018, 54: 114-120.
- 9 Watrous J, et al. Mass spectral molecular networking of living microbial colonies[J]. *Proc Natl Acad Sci*

- USA,2012,109:E1743-1752.
- 10 Said IH,et al.LC-MS/MS based molecular networking approach for the identification of cocoa phenolic metabolites in human urine[J].Food Res Int,2020,132:109119.
 - 11 Quinn RA,et al.Molecular networking as a drug discovery,drug metabolism,and precision medicine strategy[J].Trends Pharmacol Sci,2017,38:143-154.
 - 12 Zanzarin DM,et al.Metabolomics of soybean green stem and foliar retention(GSFR) disease using mass spectrometry and molecular networking[J].Rapid Commun Mass Spectrom,2020,34(Suppl 3):e8655.
 - 13 Trivella DBB,et al.The tripod for bacterial natural product discovery:genome mining,silent pathway induction,and mass spectrometry-based molecular networking[J].mSystems,2018,3:e00160-17.
 - 14 Fox Ramos AE,et al.Natural products targeting strategies involving molecular networking:different manners,one goal[J].Nat Prod Rep,2019,36:960-980.
 - 15 Duncan KR,et al.Molecular networking and pattern-based genome mining improves discovery of biosynthetic gene clusters and their products from *Salinispora* species[J].Chem Biol,2015,22:460-471.
 - 16 Nothias-Esposito M,et al.Investigation of premyrsinane and myrsinane esters in *Euphorbia cupanii* and *Euphorbia pithyusa* with MS2LDA and combinatorial molecular network annotation propagation[J].J Nat Prod,2019,82:1459-1470.
 - 17 Kang KB,et al.Targeted isolation of neuroprotective dicoumaroyl neolignans and lignans from *Sageretia theezans* using in silico molecular network annotation propagation-based dereplication[J].J Nat Prod,2018,81:1819-1828.
 - 18 Lei H,et al.A comprehensive quality evaluation of Fuzi and its processed product through integration of UPLC-QTOF/MS combined MS/MS-based mass spectral molecular networking with multivariate statistical analysis and HPLC-MS/MS[J].J Ethnopharmacol,2021,266:113455.
 - 19 Olivon F,et al.Bioactive natural products prioritization using massive multi-informational molecular networks[J].ACS Chem Biol,2017,12:2644-2651.
 - 20 Soares V,et al.Extending compound identification for molecular network using the LipidXplorer database independent method:a proof of concept using glycoalkaloids from *Solanum pseudoquina* A.St.-Hil[J].Phytochem Anal,2019,30:132-138.
 - 21 Carriot N,et al.Integration of LC/MS-based molecular networking and classical phytochemical approach allows in-depth annotation of the metabolome of non-model organisms - The case study of the brown seaweed *Taonia atomaria*[J].Talanta,2021,225:121925.

- 22 Phelan VV. Feature-based molecular networking for metabolite annotation[J]. *Methods Mol Biol*, 2020, 2104: 227-243.
- 23 He QF, et al. Discovery of neuritogenic securine alkaloids from *Flueggea suffruticosa* by a building blocks-based molecular network strategy[J]. *Angew Chem Int Ed Engl*, 2021, 60: 19609-19613.
- 24 Cheng TF, et al. LC-MS/MS-based molecular networking producing enlighten study of Chinese materia medica LC-MS/MS[J]. *Chin Tradit Herb Drugs(中草药)*, 2018, 49: 265-273.
- 25 Ozakin S, et al. Genome and metabolome mining of marine obligate *Salinispora* strains to discover new natural products[J]. *Turk J Biol*, 2019, 43: 28-36.
- 26 Hautbergue T, et al. Combination of isotope labeling and molecular networking of tandem mass spectrometry data to reveal 69 unknown metabolites produced by *Penicillium nordicum*[J]. *Anal Chem*, 2019, 91: 12191-12202.
- 27 Perruchon O, et al. Combination of UHPLC-MS/MS-molecular networking approach and FTICR-MS for the metabolic profiling of *Saccharomyces cerevisiae*[J]. *J Pharm Biomed Anal*, 2021, 195: 113857.
- 28 Yuan E, et al. Identification on triterpenoid saponins of *Clematis* species by UPLC-LTQ-Orbitrap MS combined with molecular network[J]. *Chin Tradit Herb Drugs(中草药)*, 2020, 51: 6157-6167.
- 29 Huang FF, et al. Rapid identification of alkaloids in the rhizomes of *Corydalis decumbens* by molecular networking base on HPLC-Q-TOF-MS/MS[J]. *J Chin Mass Spectr Soc(质谱学报)*, 2021, 42: 228-240.
- 30 Shen CX. Comprehensive chemical characterization and comparison of Ziziphi Spinosae Semen and Ziziphi Mauritiana Semen[D]. Taiyuan: Shanxi University(山西大学), 2020.
- 31 Pan HQ, et al. An integrated approach for global profiling of multi-type constituents: comprehensive chemical characterization of *Lonicerae Japonicae Flos* as a case study[J]. *J Chromatogr A*, 2020, 1613: 460674.
- 32 Hu GL, et al. Lactam ent-kaurane diterpene: a new class of diterpenoids present in roasted beans of *Coffea arabica*[J]. *J Agr Food Chem*, 2020, 68: 6112-6121.
- 33 Lyu Q, et al. Comprehensive structural characterization of phenolics in litchi pulp using tandem mass spectral molecular networking[J]. *Food Chem*, 2019, 282: 9-17.
- 34 Areas DRPC, et al. Chemical constituents and dereplication study of *Lessingianthus brevifolius*(Less.) H. Rob. (Asteraceae) by UHPLC-HRMS and molecular networking[J]. *Nat Prod Res*, 2022, 36: 1889-1892.
- 35 Ruan ZL, et al. Rapid identification of dibenzocyclooctadiene lignans in *Schisandra chinensis* with UHPLC Q-Exactive Orbitrap MS/MS and molecular network[J]. *Centr South Pharm(中南药学)*, 2022, 20: 499-506.
- 36 Jiang ZB, et al. Classification of diterpenoid alkaloids from *Aconitum kusnezoffii* Reichb. by liquid

- chromatography-tandem mass spectrometry-based on molecular networking[J].J Sep Sci,2022,45:739-751.
- 37 Brel O,et al.Paecilosetin derivatives as potent antimicrobial agents from *Isaria farinosa*[J].J Nat Prod,2020,83:2915-2922.
- 38 McManus KM,et al.Isolation of isotrichophycin C and trichophycins G-I from a collection of *Trichodesmium thiebautii*[J].J Nat Prod,2020,83:2664-2671.
- 39 Reginaldo FPS,et al.Molecular networking discloses the chemical diversity of flavonoids and selaginellins in *Selaginella convoluta*[J].Planta Med,2021,87:113-123.
- 40 Purves K,et al.Using molecular networking for microbial secondary metabolite bioprospecting[J].Metabolites,2016,6:2-19.
- 41 Song K,et al.Molecular network-guided alkaloid profiling of aerial parts of *Papaver nudicaule* L. using LC-HRMS[J].Molecules,2021,26:339-352.
- 42 Wang XH,et al.A molecular networking based discovery of diketopiperazine heterodimers and aspergillicins from *Aspergillus caelatus*[J].J Nat Prod,2022,85:25-33.
- 43 Freire VF,et al.Feature-based molecular networking discovery of bromopyrrole alkaloids from the marine sponge *Agelas dispar*[J].J Nat Prod,2022,85:1340-1350.
- 44 Zhao XY,et al.Combination of untargeted metabolomics approach and molecular networking analysis to identify unique natural components in wild *Morchella* sp.by UPLC-Q-TOF-MS[J].Food Chem,2022,366:130642.
- 45 Caldas LA,et al.Metabolomics of the wild mushroom *Gymnopilus imperialis*(Agaricomycetes,Basidiomycota) by UHPLC-HRMS/MS analysis and molecular network[J].Fungal Biol,2022,126:132-138.
- 46 Luo MN,et al.Molecular networking-based screening led to the discovery of a cyclic heptadepsipeptide from an endolichenic *Xylaria* sp.[J].J Nat Prod,2022,85:972-979.
- 47 Hu J.Study on chemical components of danggui buxue decoction based on UPLC-Q-TOF-MS combined with diagnostic ions and molecular network strategies[D].Taiyuan:Shanxi University(山西大学),2021.
- 48 Wu HY,et al.Discovery of neolignan glycosides with acetylcholinesterase inhibitory activity from Huangjinya green tea guided by ultra performance liquid chromatography-tandem mass spectrometry data and global natural product social molecular networking[J].J Agr Food Chem,2019,67:11986-11993.
- 49 Zhang YT,et al.Rapid identification of chemical constituents in erigeron breviscapus injection based on UHPLC-LTQ-Orbitrap MSⁿ and molecular network technology[J].J Chin Mass Spectr Soc(质谱学

- 报),2022,43:287-299.
- 50 Xu YP.Study on the secondary metabolites from *Myrothecium roridum* IFB-E091 of *Artemisia annua* based on molecular networking[D].Yangzhou:Yangzhou University(扬州大学),2021.
 - 51 Li YY,et al.Integrated molecular networking strategy enhance the accuracy and visualization of components identification:a case study of *Ginkgo biloba* leaf extract[J].J Pharm Biomed Anal,2022,209:114523.
 - 52 Allard PM,et al.Integration of molecular networking and in-silico MS/MS fragmentation for natural products dereplication[J].Anal Chem,2016,88:3317-3323.
 - 53 Gao YL,et al.Molecular networking as a dereplication strategy for monitoring metabolites of natural product treated cancer cells[J].Rapid Commun Mass Spectrom,2020,34(Suppl 1):e8549.
 - 54 Xiong Q,et al.High throughput analysis of metabolites in human plasma by UPLC-HRMS combined with mass spectrometry-based molecular networking[J].J Chin Mass Spectr Soc(质谱学报),2022,43:365-373.
 - 55 Farag MA,et al.Molecular networking based LC/MS reveals novel biotransformation products of green coffee by *ex vivo* cultures of the human gut microbiome[J].Metabolomics,2020,16:86-100.
 - 56 Wang HC,et al.Qualitative,quantitative,and pharmacokinetic study on the absorbed components of *Ardisia japonica*(Thunb.) Blume in rat plasma based on molecular networking combined with quadrupole time-of-flight LC/MS and triple quadrupole LC/MS[J].Biomed Chromatogr,2021,35:e5099.
 - 57 Seo JI,et al.Molecular networking-guided strategy for the pharmacokinetic study of herbal medicines:*Cudrania tricuspidata* leaf extracts[J].Biomed Pharmacother,2022,149:112895.
 - 58 Winnikoff JR,et al.Quantitative molecular networking to profile marine cyanobacterial metabolomes[J].J Antibiot(Tokyo),2014,67:105-112.
 - 59 Calabrese V,et al.Molecular networking and collision cross section prediction for structural isomer and unknown compound identification in plant metabolomics:a case study applied to *Zanthoxylum heitzii* extracts[J].Anal Bioanal Chem,2022,414:4103-4118.
 - 60 Ma XX,et al.LC-MS/MS based molecular network analysis of the effects of chemical regulation on the secondary metabolites and biological activities of a fungal strain *Aspergillus terreus* C23-3[J].Biotechnol Bull(生物技术通报),2021,37:95-110.
 - 61 Li HR,et al.Analysis of components of raw and wine-processed Ligustri Lucidi Fructus based on multivariate statistical analysis[J].Nat Prod Res Dev(天然产物研究与开发),2022,34:102-113.
 - 62 Nothias LF,et al.Bioactivity-based molecular networking for the discovery of drug leads in natural product bioassay-guided fractionation[J].J Nat Prod,2018,81:758-767.

- 63 Raheem DJ, et al. Application of metabolomics and molecular networking in investigating the chemical profile and antitrypanosomal activity of British bluebells (*Hyacinthoides non-scripta*) [J]. *Sci Rep*, 2019, 9: 2547-2560.
- 64 Olivon F, et al. Antiviral compounds from *Codiaeum peltatum* targeted by a multi-informative molecular networks approach [J]. *J Nat Prod*, 2019, 82: 330-340.
- 65 Cabral RS, et al. Targeted isolation of indolopyridoquinazoline alkaloids from *Conchocarpus fontanesianus* based on molecular networks [J]. *J Nat Prod*, 2016, 79: 2270-2278.
- 66 Naman CB, et al. Integrating molecular networking and biological assays to target the isolation of a cytotoxic cyclic octapeptide, samoamide A, from an American samoan marine cyanobacterium [J]. *J Nat Prod*, 2017, 80: 625-633.
- 67 Han JJ, et al. Exploring verrucosidin derivatives with glucose-uptake-stimulatory activity from *Penicillium cellarum* using MS/MS-based molecular networking [J]. *J Fungi (Basel)*, 2022, 8: 143-159.
- 68 Wu T, et al. Structure-based molecular networking for the discovery of anti-HBV compounds from *Saussurea lappa* (Decne.) C.B Clarke [J]. *Molecules*, 2022, 27: 2023-2042.
- 69 Mi SH, et al. Guided isolation of daphnane-type diterpenes from *Daphne genkwa* by molecular network strategies [J]. *Phytochemistry*, 2022, 198: 113144.
- 70 Hou XM, et al. Integrating molecular networking and (1)H NMR to target the isolation of chrysogeamides from a library of marine-derived *Penicillium* fungi [J]. *J Org Chem*, 2019, 84: 1228-1237.
- 71 Liu LL, et al. Molecular networking-based for the target discovery of potent antiproliferative polycyclic macrolactam ansamycins from *Streptomyces cacaoi* subsp. *asoensis* [J]. *Org Chem Front*, 2020, 7: 4008-4018.
- 72 Bracegirdle J, et al. Skyllamycins D and E, non-ribosomal cyclic depsipeptides from lichen-sourced *Streptomyces anulatus* [J]. *J Nat Prod*, 2021, 84: 2536-2543.
- 73 Lee J, et al. Targeted isolation of lignans from *Trachelospermum asiaticum* using molecular networking and hierarchical clustering analysis [J]. *Biomolecules*, 2020, 10: 378-387.
- 74 Nothias LF, et al. Environmentally friendly procedure based on supercritical fluid chromatography and tandem mass spectrometry molecular networking for the discovery of potent antiviral compounds from *Euphorbia semiperfoliata* [J]. *J Nat Prod*, 2017, 80: 2620-2629.
- 75 Kalinski JJ, et al. Molecular networking reveals two distinct chemotypes in pyrroloiminoquinone-producing *Tsitsikamma favus* sponges [J]. *Mar Drugs*, 2019, 17: 60-75.

收稿日期: 2022-04-14

接受日期: xxxx-xx-xx

基金项目：国家自然科学基金（82160744）

*通信作者 Tel: 15969493103, 13619610405; E-mail: lidashan@ynutcm.edu.cn, notoww@163.com