

天然生物活性分子高效发现的新策略和方法研究进展

吴振龙^{1,2}, 王英^{1,2}, 叶文才^{1,2*}

(1. 暨南大学天然生物活性分子与创新药物研究中心, 广东 广州 510632; 2. 广东省中药药效物质基础与创新药物研究重点实验室, 广东 广州 510632)

[摘要] 天然生物活性分子是创新药物的重要源泉。然而, 长期以来天然生物活性分子发现研究的基本理论和方法并没有根本性的提升, 仍存在发现周期较长、人工消耗大、发现过程带有盲目性等问题, 严重制约了源于天然生物活性分子的创新药物发现研究。近年来, 随着现代分析技术、代谢组学、生物信息学、人工智能等的快速发展, 国际上出现了基于质谱技术、核磁共振技术等多学科交叉的天然生物活性分子发现的去重策略和方法, 如分子网络技术、小分子精准识别技术等。这些新策略、新方法的应用极大地提高了新颖结构天然生物活性分子的发现效率, 也逐渐推动传统的研究策略和方法发生变革。对于基于质谱技术、核磁共振技术的天然生物活性分子高效发现新策略和新方法进行总结, 为上述技术的进一步开发和利用提供参考。

[关键词] 天然生物活性分子; 质谱; 核磁共振; 去重

[中图分类号] R93

[文献标志码] A

[文章编号] 1001-5094 (2022) 03-0163-10

Advances in Research on Novel Strategies and Approaches for the Efficient Discovery of Bioactive Natural Molecules

WU Zhenlong^{1,2}, WANG Ying^{1,2}, YE Wencai^{1,2}

(1. Center for Bioactive Natural Molecules and Innovative Drugs Research, Jinan University, Guangzhou 510632, China; 2. Guangdong Province Key Laboratory of Pharmacodynamic Constituents of TCM & New Drugs Research, Guangzhou 510632, China)

[Abstract] Bioactive natural molecules are important resources for the drug development. However, for a long time, no fundamental breakthroughs in the basic theories and investigation methods for research on natural products were achieved. The discovery processes for bioactive natural molecules are still time-consuming, labor-intensive, and unpredictable, which seriously restricted the research efficiency of bioactive natural molecules for innovative drugs development. In recent years, with the rapid development of modern analytical technology, metabonomics, bioinformatics, and artificial intelligence, the research approaches for bioactive natural molecules discovery are revolutionizing, and a number of mass spectrometry (MS)- and nuclear magnetic resonance (NMR)-based dereplication strategies, such as molecular networking (MN) and small molecule accurate recognition technology (SMART), are emerging and applying. This review summarized the MS- and NMR-based novel strategies and methods for efficient discovery of bioactive natural products, aiming to provide some reference for the further development and utilization of these technologies.

[Key words] bioactive natural molecule; mass spectrometry; nuclear magnetic resonance; dereplication

天然生物活性分子因其多样的化学结构和广泛的药理活性, 一直是药物先导物的重要源泉。据统计, 从1981年1月至2019年9月, FDA共批准了1881个新药, 其中约有一半直接或间接来源于天然化合物^[1]。然而, 近年来, 大型制药公司和研发机构对天然生物活性分子的研究兴趣逐渐减少, 相关的研发投入也呈逐年下降趋势^[2-4]。究其原因, 一方

面是由于传统的天然生物活性化合物发现过程漫长、费时耗力, 且定向性不强, 分离过程的盲目性导致大量的已知成分被重复地发现; 另一方面, 与发现新天然化合物的数量逐年增加相反, 近年来所发现化合物的结构新颖性却在逐渐降低, 适用于首创新药 (first-in-class) 研发的新颖结构小分子化合物仍然匮乏^[5-6]。针对上述不足和挑战, 如何实现已知化合物的去重 (dereplication), 并快速、定向地发现结构新颖的天然生物活性分子成为了天然药物化学学科关注的热点和难点问题。

近年来, 随着现代分析技术、代谢组学、生

接受日期: 2021-12-20

* 通信作者: 叶文才, 教授;

研究方向: 中药及天然药物活性成分;

Tel: 020-85220004; E-mail: chywc@aliyun.com

物信息学、人工智能技术等快速发展以及各学科之间的交叉融合, 天然生物活性分子的发现研究也涌现出了一批新的策略和方法^[7-10], 如基于液相色谱串联二级质谱(LC-MS/MS)的分子网络技术(molecular networking, MN)和基于核磁共振(NMR)的小分子精确识别技术(small molecule accurate recognition technology, SMART)等。这些新策略和新方法的出现为靶向挖掘结构新颖的天然生物活性分子带来了新的研究思路, 目前已广泛应用于植物、微生物、海洋天然产物的发现研究中, 掀起了天然产物研究的新浪潮。本文将对部分具代表性的新技术和方法进行总结, 以期为进一步开发和利用天然生物活性分子的高效发现策略和方法提供参考。

1 基于 LC-MS/MS 的天然生物活性分子高效发现策略

液相色谱串联质谱(LC-MS)因具有较高的灵敏度、分辨率以及较快的分析速度等优点, 可满足对天然产物提取物进行准确的定性和定量分析的要求。前期, 已有学者借助化合物在 LC-MS 中的特征紫外吸收峰和相对分子质量, 实现了对天然产物提取物或馏分中目标化合物的靶向分离。但由于特征紫外吸收通常仅适用于分离度较好、样品组成较为简单的组分, 而对成分复杂的粗提取物则效果不理想。此外, 对于一些结构新颖的化合物, 其相对分子质量可能与已知成分相同, 故在 LC-MS 中也难以区分。因此, 单独使用 LC-MS 技术无法对复杂天然提取物中的未知成分进行高效识别。

2012年, 美国加利福尼亚大学圣地亚哥分校的 Pieter Dorrestein 教授团队首次提出了基于二级质谱(MS/MS)的分子网络技术, 并将其运用于天然生物活性分子的发现研究中^[11]。分子网络技术是综合运用现代质谱、生物信息学、计算机等技术所开发的一种二级质谱数据可视化策略。其原理是: 结构相同或相近的化合物在同一条件下可产生相同或相似的二级质谱碎片, 反之亦然。按照一定的算法将 LC-MS/MS 图谱中的每张二级质谱图进行比对, 并计算各图谱的相似性, 以余弦值(0~1)表示,

相似性越高, 则余弦值越大。根据计算所得余弦值的大小, 可将相似性较高的图谱分别聚集成簇, 而相似性较低的图谱则单独成簇, 最终可将所有的二级质谱图整合成为一张可视化的分子网络图(见图1)。在该网络图中, 每一个节点代表一个化合物, 其所包含的二级质谱信息可用节点的名称、颜色、大小、形状等来表示。节点与节点之间的连线则表示两个化合物结构的相关性, 其大小可用连线的粗细来表示。通过构建分子网络图, 可将复杂的二级质谱信息十分直观地呈现出来, 这些信息包括了测试样品中所含有的全部化合物, 以及化合物与化合物之间化学结构的关联性。近期, Nothias 等^[12]在运用 MS/MS 数据构建经典分子网络的基础上, 首次引入了液相色谱的保留时间以及一级质谱数据的信息, 开发了基于特征的分子网络技术(feature-based molecular networking, FBMN)。该技术一方面解决了传统分子网络技术在鉴定同分异构体方面的局限性, 可区分同分异构体; 另一方面, 还可通过整合一级质谱数据的离子丰度等信息, 对目标化合物进行相对定量分析。

在传统分子网络技术中, 化合物的指认通常是通过比对待分析物与已知化合物的二级质谱数据完成, 利用该方法所发现的新化合物多为已知化合物的同系物或类似物, 其结构的新颖性往往不强。因此, 如何实现靶向识别具有新颖骨架的化合物依然面临挑战。天然化合物骨架结构复杂、类型多样, 但它们通常来自于简单的生源前体, 经转化形成特定的生源砌块, 再经进一步的生物合成途径而形成。而天然化合物的生源砌块所对应的结构片段往往会在化合物的二级质谱图中显示出特征性的子离子碎片峰或(和)中性丢失碎片峰, 可用于生源砌块的识别。为克服传统分子网络技术在鉴定具有新颖骨架结构化合物方面的不足, 笔者所在课题组最近通过集成生源砌块识别技术和分子网络技术, 首次提出了基于生源砌块的分子网络策略(building blocks-based molecular network, BBMN)^[13]。相比于传统的分子网络技术, BBMN 策略在发现新颖骨架结构化合物方面具有明显的优势: 一方面, BBMN 策略可根据目标化合物的结构特点对复杂提取物中所包

含的生源砌块进行快速识别, 因而对分析的化合物具有强选择性; 另一方面, 针对二级质谱数据量庞大、冗杂的特点, BBMN 策略在选择性过滤的基础上可简化待分析物的数据集, 并通过分子网络技术对数据进行可视化分析, 方便研究人员快速锁定目标化

合物 (见图 2)。天然化合物大多由特定的生源砌块组合而成, 因此, BBMN 策略可广泛应用于其他结构类型天然化合物的发现研究中, 有望为传统天然活性化合物的发现研究带来新变革。

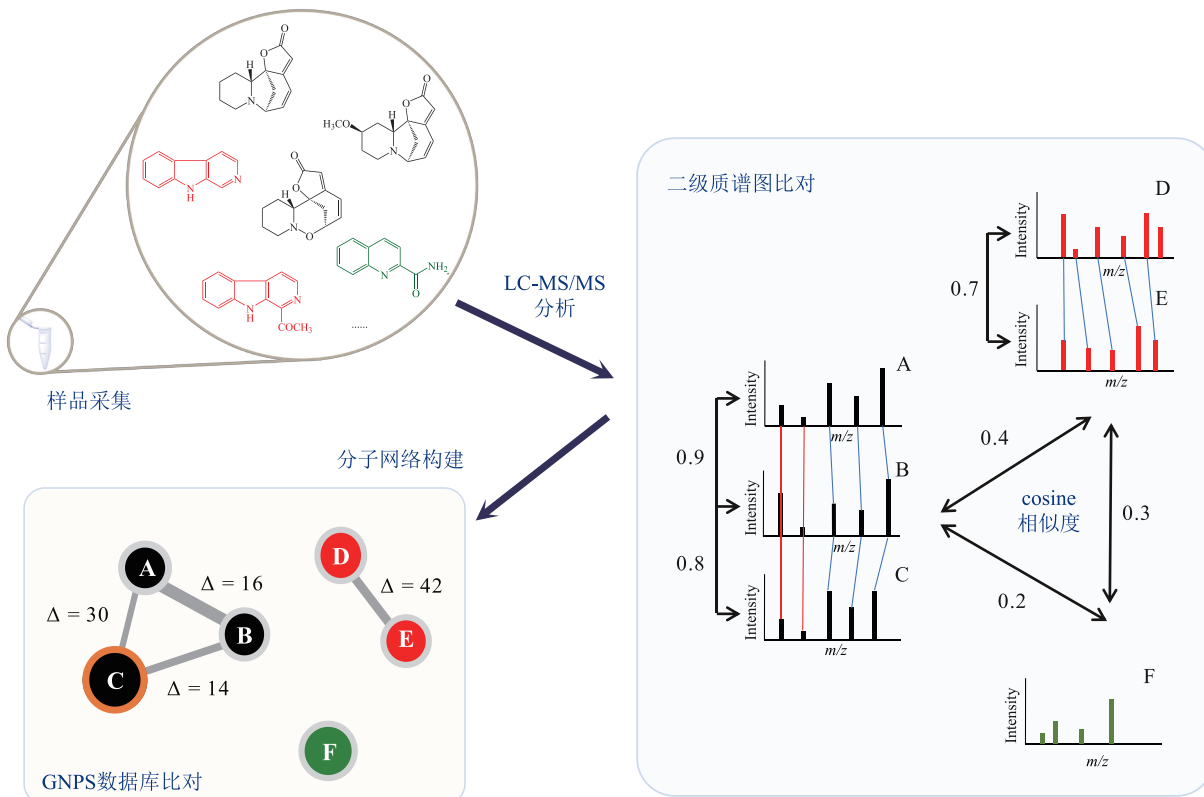


图 1 分子网络原理示意图

Figure 1 A schematic diagram of molecular networking

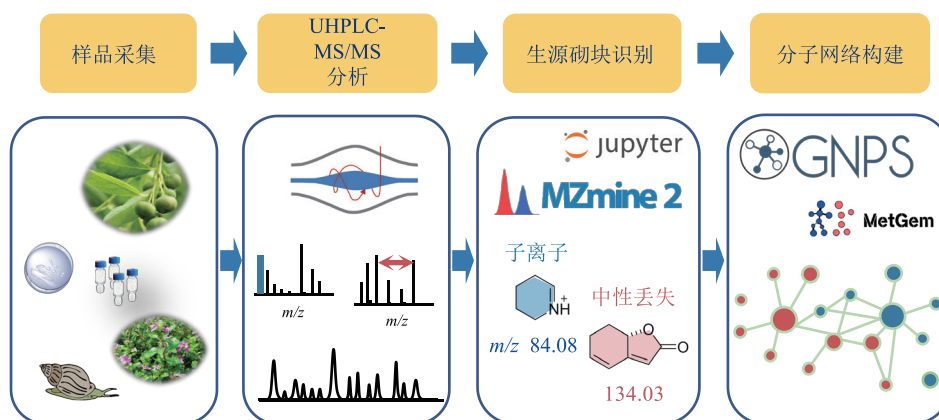


图 2 基于生源砌块的分子网络构建流程图

Figure 2 A diagram for the building blocks-based molecular network construction

1.1 运用分子网络技术可实现快速去重及发现结构新颖化合物

为避免重复分离得到已知化合物, 在对复杂的天然提取物进行分离纯化前, 快速指认出提取物中所包含的已知化合物尤为重要, 该过程通常被称作去重^[8,14-15]。2016年, Dorrestein教授和Bandeira教授共同建立了基于LC-MS/MS数据的全球天然产物交互分子网络平台(Global Natural Products Society molecular networking, GNPS)(<http://gnps.ucsd.edu>)^[16], 旨在为全世界天然产物化学家提供一个可共享使用的二级质谱数据存储及分析平台。在数据存储方面, GNPS平台提供了开放性的免费数据库Mass Spectrometry Interactive Virtual Environment(MassIVE), 可通过该数据库查找化合物的原始MS/MS质谱数据, 并进行数据共享。该数据库不仅囊括了FDA Library、PhytoChemical Library、NIH Clinical Collection、MassBank Spectral Library、HMDB Spectral Library、CASMI Spectral Library等第三方数据库的化合物二级质谱数据信息, 还收录了Dorrestein教授实验室及其合作实验室所采集的大量化合物二级质谱数据。目前, MassIVE数据库共收录了大约2万余个天然化合物的共计约23万余张二级质谱图。除存储收录二级质谱数据外, GNPS平台还利用加州大学圣地亚哥分校的计算质谱中心(Center for Computational Mass Spectrometry, CCMS), 开发了一系列可处理二级质谱数据的计算机算法和软件, 如Dereplicator+、VarQuest等^[17-18]。利用这些算法和软件, 可将待分析物的二级质谱数据与数据库中的化合物质谱数据进行比对, 快速指认出待分析物中所包含的已知化合物。

借助GNPS平台, 国内外研究人员已运用分子网络技术对微生物、海洋以及植物来源的天然产物成功开展了去重研究, 在快速识别已知化合物的基础上, 发现了一系列结构新颖的天然化合物。Yang等^[19]运用分子网络技术对一系列海洋和陆地来源的微生物样本进行了研究, 通过分别采集各样品的LC-MS/MS图谱, 并运用一些已知结构的单体化合物作为“种子”, 构建了这些样本的分子网络图。借助于“种子”化合物在网络图中的指示作用, 作

者快速地识别了12个已知化合物和46个结构类似物, 并最终实现了对上述样本中化合物的针对性分离。Moore课题组^[20]运用分子网络技术对146种盐孢菌属和链霉菌属菌株在不同培养条件下的次生代谢产物进行了快速检测, 从603个样品中获得了近180万张二级质谱的谱图。通过分析比对以上次生代谢产物的分子网络图谱, 发现同一菌株在不同培养环境下的生物合成途径较为多样, 其次生代谢产物有较大差异。基于以上发现, 高效、快速地确定了部分菌株的最优培养条件, 并从中获得了一些结构新颖的天然化合物。Fox Ramos等^[21]在对夹竹桃科植物*Geissopermum laeve*的吲哚型生物碱类成分的研究过程中, 运用分子网络技术对该植物的茎皮部位开展了进一步的化学成分挖掘。通过一个自建的单萜吲哚生物碱二级质谱数据库(monoterpene indole alkaloid database, MIADB), 作者对分子网络图中已知化合物的节点进行了指认, 并指导分离了3个新的单萜吲哚生物碱类化合物。此外, Zhu等^[22]运用分子网络技术对植物致病真菌*Epicoccum nigrum* 09116进行了化学成分探索, 成功分离获得了8个新的二苯并螺缩酮类化合物。有趣的是, 该研究除利用分子网络技术进行去重研究外, 还利用此技术成功鉴别出了上述新化合物的生物合成中间体。

运用基于生源砌块的分子网络策略, 笔者课题组对大戟科白饭树属药用植物一叶萩(*Flueggea suffurtiocsia*)的总生物碱部位进行了深入挖掘^[13]。首先, 通过搜索总生物碱部位中所包含的特征生源砌块的离子碎片, 即质荷比为84.08的子离子碎片和(或)质荷比为134.06的中性丢失碎片, 二者分别对应为一叶萩型生物碱骨架中的哌啶环和苯并咪喃酮结构单元, 可从一叶萩总生物碱部位的LC-MS/MS谱图中快速识别含有一叶萩型生物碱生物合成砌块的离子峰。随后, 结合化学信息学手段和分子网络技术, 成功构建了一叶萩总生物碱部位的基于生源砌块的分子网络图。借助于该网络图, 从一叶萩总生物碱部位中快速识别并分离鉴定了3个目标化合物。其中, 化合物suffranidine A具有一个十分罕见的8/5/6/5/6/6/6/6八环骨架, 并包含一个独特的

笼状氮杂双环 [6.4.0.0^{3,11}] 十二烷结构单元。化合物 suffranidines B 和 C 是 2 个高度官能团化的一叶萩型生物碱二聚体, 二者分别在一叶萩碱基本母核结构基础上插入了一个额外的 C6 单元, 构建成了 2 类全新骨架的一叶萩型生物碱。

1.2 运用分子网络技术发现天然生物活性化合物

在天然产物研究中, 除发现结构新颖的化合物之外, 如何快速寻找具有生物活性的化合物是天然产物化学家所关注的另外一个十分重要的问题^[23-28]。从前期的研究经验来看, 获得天然生物活性化合物的途径通常有以下 2 条。1) 以结构为导向的分离策略。通过对复杂提取物进行系统的化学成分研究, 从中分离获得单体化合物, 再对所获得的化合物进行系统的生物活性评价, 从而发现具有生物活性的化合物。2) 以活性为导向的追踪分离策略。首先对粗提物进行活性评价, 再针对活性部位或馏分开展进一步分离, 然后评价各次级馏分的生物活性, 依次反复, 直到获得活性化合物。虽然以上方法均有成功获得药物分子的案例, 如紫杉醇、长春碱、青蒿素的发现, 但是由于该过程需要耗费巨大的时间及人力、物力成本, 无法与现代高通量筛选技术相兼容, 已逐渐被制药企业和研究机构所弃用。另外, 在近年来的研究中, 通过活性追踪方法所分离获得的天然产物大多为已知化合物^[29], 并经常发生活性“消失”的情况, 造成研究往往以失败告终。因此, 运用传统的方法高效获取结构新颖的活性天然产物依然面临巨大的挑战。

随着现代质谱技术及生物信息学的快速发展, 运用基于 LC-MS/MS 的分子网络技术指导发现活性天然产物已成为可能^[30]。例如, Naman 等^[31]运用分子网络技术对蓝藻细菌 *Symploca* sp. 的细胞毒活性成分进行了研究。在此之前, 研究人员已从该细菌中分离鉴定了多个具有显著体外细胞毒活性的化合物。为了从该细菌中获得其他结构新颖的活性化合物, 作者采集了不同地域的蓝藻细菌样品并培养富集其次级代谢产物, 经初步分离, 分别测试各馏分的二级质谱数据及体外细胞毒活性, 最终将生物活性结果整合到分子网络图中。在这个带有生物活性标签的分子网络图指导下, 作者首先从众多样

本中识别细胞毒活性较强的馏分, 随后从中排除已报道具有生物活性的已知化合物, 最终成功获得了一个新的环状八肽化合物。体外细胞毒活性测试结果显示, 该化合物具有较强的抗大细胞肺癌细胞株 H460 活性, 其 IC₅₀ 值为 1.1 μmol·L⁻¹。该方法虽然可有效避免重复分离已知的活性化合物, 但仍然需要对各次馏分进行反复的活性测试, 故研究周期相对较长。

2018 年, Nothias 等^[32]首次提出了基于生物活性的分子网络 (bioactivity-based molecular network) 策略, 并运用该策略从大戟科植物 *Euphorbia dendroides* 的提取物中靶向获得了 2 个新的具显著抗基孔肯雅病毒 (CHIKV) 活性的二萜类化合物。基于生物活性的分子网络策略主要包括以下 3 个步骤: 1) 采集待分析物各馏分的 LC-MS/MS 数据, 运用 MZmine、OpenMS 等质谱数据处理软件提取谱图中的母离子峰及碎片峰信号, 并对各母离子峰进行相对定量分析; 2) 测试各馏分的生物活性, 根据样品质谱数据中母离子峰强度和所测得的样品活性数据结果, 运用 Pearson 相关性分析计算各母离子峰的生物活性分值; 3) 将活性预测分值整合到质谱数据中, 并运用 GNPS 平台生成基于生物活性的分子网络图。与 Naman 等^[31]的方法相比, 基于生物活性的分子网络策略不需要通过反复的分离和活性测试, 便可从复杂的天然产物提取物中快速识别出具有生物活性的化合物, 显示出了较强的靶向性。

2 基于核磁共振技术的天然化合物高效发现策略

目前, NMR 技术是天然产物结构表征最常用、最可靠的方法之一, 它提供了测试分子在原子层面的信息, 在未知化合物结构鉴定方面具有十分重要的地位^[33]。与 LC-MS/MS 相比较, NMR 技术的灵敏度低、测试时间长, 但同时 NMR 也具有 LC-MS/MS 技术所无法比拟的优点, 如测试时不破坏样品、数据重现性高、可对难离子化的样品进行测试、可区分同分异构体等。除此之外, 通过 NMR 技术所获取的样品结构信息较 MS 更为丰富, 可对复杂样品中所包含化合物的结构进行准确鉴定^[34]。

近年来, 随着商业化 NMR 仪器磁场强度的逐渐提高, 以及超低温探头的普遍使用, NMR 仪器的分辨率和灵敏度均得到了极大提升, 可对微克级别的样品进行测试, 并同时大大缩短测试时间^[35]。另外, 二维 NMR 技术 (2D NMR) 克服了一维 NMR (1D NMR) 谱图信号重叠严重的问题, 更适用于复杂天然产物的结构分析^[36]。最近, 研究人员开发了一系列功能强大的 2D NMR 脉冲序列 (如 diffusion-ordered spectroscopy, DOSY)、非均一采样技术 (nonuniform sampling, NUS)、协方差核磁共振技术 (covariance NMR) 等^[37-42], 进一步提高了仪器的分辨率, 并缩短了测试时间, 使 NMR 技术可更加广泛地应用于复杂和微量成分的研究中。

2.1 核磁共振技术指导发现结构新颖的天然化合物

在天然产物研究领域, 研究人员开发了多种谱图比对算法, 利用 NMR 谱图 (尤其是 2D NMR 谱图) 进行去重研究^[43-44]。然而, 由于样品浓度、杂质峰信号、溶剂效应、官能团之间的相互作用等对化学位移值的影响, 上述算法对复杂的 NMR 谱图信号的识别效率和准确性较差, 限制了它们在天然产物去重研究中的推广和应用。为克服上述不足, 深度学习 (deep learning) 等人工智能技术 (artificial intelligence, AI) 逐渐被应用到复杂 2D NMR 图谱的精确识别中。相比于传统的机器学习方法, 深度学习不需要研究人员的任何设计和参与, 在训练过程中即可创建最合适的特征集, 尤其适合庞大且未知的数据集^[45-46]。

2017 年, Zhang 等^[47] 运用非均一采样技术 (non-uniform sampling, NUS) 和卷积神经网络技术 (convolutional neural network, CNN), 开发了基于异核单量子相关谱 (heteronuclear single quantum coherence, HSQC) 的天然产物高效发现新策略, 并将其命名为小分子精确识别技术 (small molecule accurate recognition technology, SMART)。该技术利用 siamese 神经网络架构对一个包含了 2 054 张天然产物 HSQC 谱图的数据集进行深度卷积神经网络训练^[48], 构建了一个可区分谱图相似度的节点空间。在这个节点空间中, 结构相似的化合物在空间上相

近, 而结构相差较大的化合物在空间上则相离较远。随后, 通过一个非均一采样序列, 快速采集待分析物的 HSQC 图谱, 并提交至上述训练后的深度卷积神经网络。在接下来的数据分析中, 可根据待分析物的 HSQC 谱图信号在节点空间中的分布来实现对复杂提取物中所含有的已知化合物和新化合物的快速区分 (见图 3)。

为进一步提高 SMART 技术识别天然产物 HSQC 谱图的准确性, Reher 等^[49] 运用 JEOL 数据库 (<https://www.j-resonance.com/en/nmrdb>) 中 25 434 张天然产物的 HSQC 谱图以及利用 ACD/Labs 软件所预测的 27 642 张天然产物的 HSQC 谱图进行了卷积神经网络学习, 将 SMART 技术升级到 2.0 版本。相比于之前的版本, SMART 2.0 使用了更多的天然化合物的 HSQC 谱图作为训练集, 这些化合物大约占目前已知天然产物的 15%。因此, 通过增加训练集的数目极大地提升了 SMART 技术在识别不同结构类型天然产物 HSQC 谱图方面的能力。为验证该技术的适用性, 作者随后运用该工具对蓝藻细菌 *Symploca* sp. 的提取物进行了分析。首先, 作者使用 1.7 mm TCI 微量核磁管溶解了 1 mg 的馏分样品, 运用 NUS-ASAP-HSQC 序列快速测试了该样品的 HSQC 谱图 (600 MHz, 13 min)。随后, 运用 SMART 2.0 对该谱图进行分析, 并结合质谱靶向分离技术, 从中快速识别并获得了一个结构新颖的大环内酯类化合物 *symplocolide A*。

除 SMART 技术外, 近期还有学者基于 HSQC 谱开发了 metabolomics and dereplication by two-dimensional experiments (MADByTE) 和 atomic novelty scoring 技术^[50-51]。前者可通过 HSQC 及 TOCSY 谱图信息获得复杂混合物的自旋耦合体系特征, 并通过构建相关的网络图实现对混合物中的共有质子自旋耦合体系的识别, 从而实现去重。后者则根据 human metabolome database (HMDB) 和 BioMagResBank 库中的 HSQC 数据列出 10 308 个信号峰, 随后通过计算待分析物中各核磁信号峰与库中相邻最近的信号峰的距离, 并进行打分, 便可从复杂混合物体系中快速识别与库中化合物结构差异较大的新颖结构化合物。

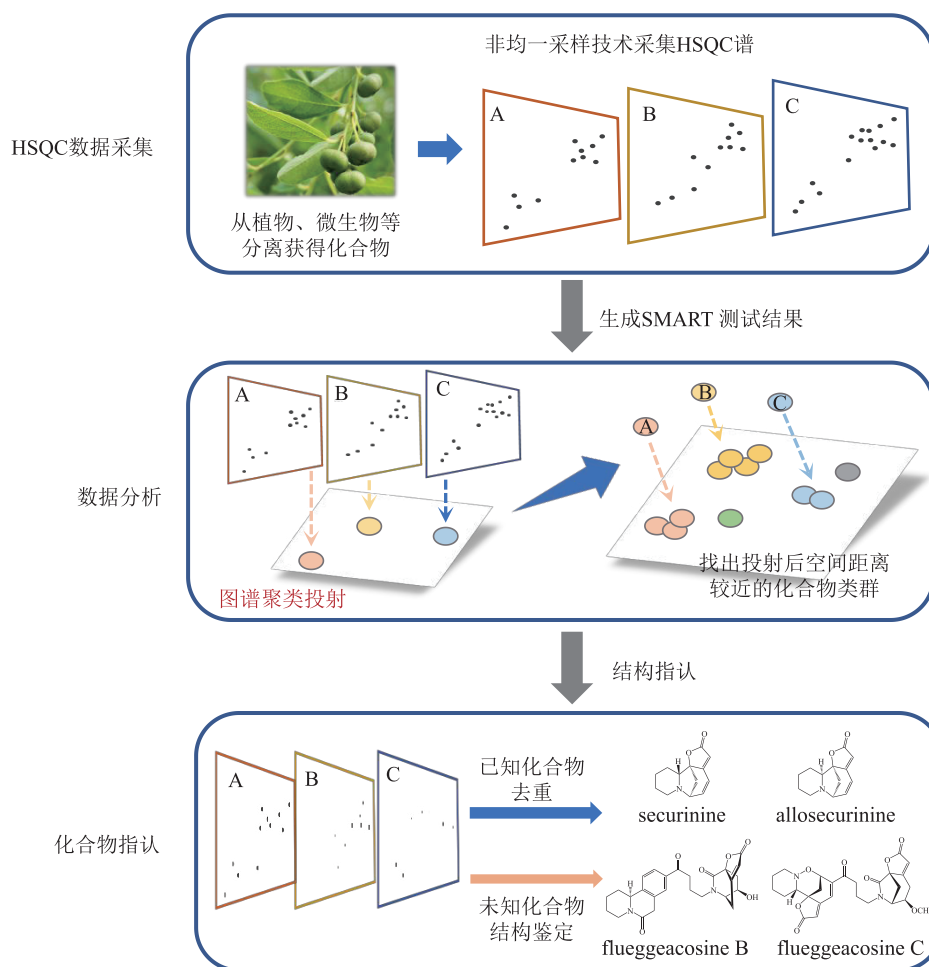


图3 小分子精确识别技术工作流程示意图

Figure 3 Workflow for the small molecule accurate recognition technology

2.2 核磁共振技术指导天然生物活性化合物的分离

与基于MS技术的活性成分靶向分离策略类似,利用NMR技术定向地寻找活性成分也是天然生物活性化合物发现研究的重要方向之一,但目前在该领域的研究报道相对较少。迄今,仅有学者基于统计泛多样性分析(statistical heterocovariance analysis, HetCA)方法,将小分子化合物的NMR数据与生物学功能相关联。其中,ELINA是Eliciting Nature's Activities的简称,是由Grienke等^[52]基于¹H NMR谱的HetCA统计分析所开发的一种活性化合物追踪分离策略。在该策略中,根据提取物或馏分活性测试的结果,对其¹H NMR谱信号中的特征峰进行正面(hot)或负面(cold)评分,从而可在未分离前判断活性化合物可能含有的特征信号峰。运

用该策略,作者从多孔真菌*Fomitopsis pinicola*的复杂提取物中发现了具有甾醇硫脂酶抑制活性的羊毛甾烷三萜类化合物。此外,Delsuc课题组^[53]开发了一种名为Plasmodesma的计算机程序(<https://plasmodesma.igbmc.science>),可将复杂天然产物提取物的1D和2D NMR谱图数据进行自动化处理,进而可从中提取出活性相关成分的NMR指纹图谱,从而实现药效团结构进行快速排查。

3 结语与展望

近年来,受益于各种现代技术的迅猛发展,天然生物活性分子的发现研究涌现出了一大批基于LC-MS/MS和NMR技术,并集成生物信息学、代谢组学、计算机科学等多学科技术手段的新策略和

新方法。通过综合运用这些新策略和新方法, 天然药物化学家已开展了诸多创新性的研究工作, 并取得了丰硕的研究成果。相比于传统的提取-分离-纯化-活性测试手段, 这些新策略和新方法具有更强的靶向性, 因而研究效率更高, 更加符合现代药物研发对快速筛选获得先导化合物的需求。迄今为止, 天然生物活性分子高效发现的新策略和新方法研究仍

处于高速发展阶段。随着质谱、NMR 等分析技术的分辨率和灵敏度的进一步提升, 天然化合物质谱及 NMR 谱图数据库的逐步扩充、完善并实现资源共享, 以及更加精准的计算机算法和更加“聪明”的人工智能技术的不断推出, 这些新技术和新方法在天然产物研究中的应用将会更加广泛, 并进一步助力天然生物活性分子的高效发现。

[参考文献]

- [1] Newman D J, Cragg G M. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019[J]. *J Nat Prod*, 2020, 83(3): 770–803.
- [2] Rishton G M. Natural products as a robust source of new drugs and drug leads: past successes and present day issues[J]. *Am J Cardiol*, 2008, 101(10): S43–S49.
- [3] Drewry D H, Macarron R. Enhancements of screening collections to address areas of unmet medical need: an industry perspective[J]. *Curr Opin Chem Biol*, 2010, 14(3): 289–298.
- [4] Lachance H, Wetzel S, Kumar K, et al. Charting, navigating, and populating natural product chemical space for drug discovery[J]. *J Med Chem*, 2012, 55(13): 5989–6001.
- [5] Pye G R, Bertin M J, Lokey R S, et al. Retrospective analysis of natural products provides insights for future discovery trends[J]. *Proc Natl Acad Sci USA*, 2017, 114(22): 5601–5606.
- [6] Li J W, Vederas J C. Drug discovery and natural products: end of an era or an endless frontier[J]. *Science*, 2009, 325: 161–165.
- [7] Wolfender J L, Nuzillard J M, van der Hoof J J, et al. Accelerating metabolite identification in natural product research: toward an ideal combination of liquid chromatography–high-resolution tandem mass spectrometry and NMR profiling, in silico databases, and chemometrics[J]. *Anal Chem*, 2019, 91(1): 704–742.
- [8] Gaudencio S P, Pereira F. Dereplication: racing to speed up the natural products discovery process[J]. *Nat Prod Rep*, 2015, 32(6): 779–810.
- [9] Beniddir M A, Kang K B, Genta-Jouve G, et al. Advances in decomposing complex metabolite mixtures using substructure-and network-based computational metabolomics approaches[J/OL]. *Nat Prod Rep*, 2021[2021-10-01]. <http://pubs.rsc.org/en/content/articlelanding/2021/np/d1np00023c>. DOI: 10.1039/D1NP00023C.
- [10] Atanasov A G, Zotchev S B, Dirsch V M, et al. Natural products in drug discovery: advances and opportunities[J]. *Nat Rev Drug Discov*, 2021, 20: 200–216.
- [11] Watrous J, Roach P, Alexandrov T, et al. Mass spectral molecular networking of living microbial colonies[J]. *Proc Natl Acad Sci USA*, 2012, 109(26): E1743–E1752.
- [12] Nothias L F, Petras D, Schmid R, et al. Feature-based molecular networking in the GNPS analysis environment[J]. *Nat Methods*, 2020, 17(9): 905–908.
- [13] He Q F, Wu Z L, Li L, et al. Discovery of neuritogenic *Securinega* alkaloids from *Flueggea suffruticosa* by a building blocks-based molecular network strategy[J]. *Angew Chem Int Ed*, 2021, 60(36): 19609–19613.
- [14] Corley D G, Durley R C. Strategies for database dereplication of natural products[J]. *J Nat Prod*, 1994, 57(11): 1484–1490.
- [15] Hubert J, Nuzillard J M, Renault J H. Dereplication strategies in natural product research: How many tools and methodologies behind the same concept[J]. *Phytochem Rev*, 2017, 16(1): 55–95.
- [16] Wang M, Carver J J, Phelan V V, et al. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking[J]. *Nat Biotechnol*, 2016, 34(8): 828–837.
- [17] Mohimani H, Gurevich A, Mikheenko A, et al. Dereplication of peptidic natural products through database search of mass spectra[J]. *Nat Chem Biol*, 2017, 13: 30–37.
- [18] Gurevich A, Mikheenko A, Shlemov A, et al. Increased diversity of peptidic natural products revealed by modification-tolerant database search of mass spectra[J]. *Nat Microbiol*, 2018, 3: 319–327.
- [19] Yang J Y, Sanchez L M, Rath C M, et al. Molecular networking as a dereplication strategy[J]. *J Nat Prod*, 2013, 76(9): 1686–1699.
- [20] Crusemann M, O'Neill E C, Larson C B, et al. Prioritizing natural

- product diversity in a collection of 146 bacterial strains based on growth and extraction protocols[J]. *J Nat Prod*, 2017, 80(3): 588–597.
- [21] Fox Ramos A E, Alcover C, Evanno L, *et al.* Revisiting previously investigated plants: a molecular networking-based study of *Geissospermum leave*[J]. *J Nat Prod*, 2017, 80(4): 1007–1014.
- [22] Zhu G, Hou C, Yu W, *et al.* Molecular networking assisted discovery and biosynthesis elucidation of the antimicrobial spiroketals epicospirocins[J]. *Chem Comm*, 2020, 56: 10171–10174.
- [23] Chow S, Liver S, Nelson A. Streamlining bioactive molecular discovery through integration and automation[J]. *Nat Rev Chem*, 2018, 2: 174–183.
- [24] Bernardini S, Tiezzi A, Masci V L, *et al.* Natural products for human health: an historical overview of the drug discovery approaches[J]. *Nat Prod Res*, 2018, 32(16): 1928–1950.
- [25] David B, Wolfender J L, Dias D A. The pharmaceutical industry and natural products: historical status and new trends[J]. *Phytochem Rev*, 2015, 14: 299–315.
- [26] Thomford N E, Sentehebane D A, Rowe A, *et al.* Natural products for drug discovery in the 21st century: innovations for novel drug discovery[J]. *Int J Mol Sci*, 2018, 19(6): 1578. DOI:10.3390/ijms19061578.
- [27] Li G, Lou H X. Strategies to diversify natural products for drug discovery[J]. *Med Res Rev*, 2018, 38(4): 1255–1294.
- [28] Kingston D G. Modern natural products drug discovery and its relevance to biodiversity conservation[J]. *J Nat Prod*, 2011, 74(3): 496–511.
- [29] Agarwal G, Carcache P J B, Addo E M, *et al.* Current status and contemporary approaches to the discovery of antitumor agents from higher plants[J]. *Biotech Adv*, 2020, 38: 107337. DOI:10.1016/j.biotechadv.2019.01.004.
- [30] Fox Ramos A E, Evanno L, Poupon E, *et al.* Natural products targeting strategies involving molecular networking: different manners, one goal[J]. *Nat Prod Rep*, 2019, 36: 960–980.
- [31] Naman C B, Rattan R, Nikoulina S E, *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(3): 625–633.
- [32] Nothias L F, Nothias-Esposito M, da Silva R, *et al.* Bioactivity-based molecular networking for the discovery of drug leads in natural product bioassay-guided fractionation[J]. *J Nat Prod*, 2018, 81(4): 758–767.
- [33] Edison A S, Colonna M, Gouveia G J, *et al.* NMR: Unique strengths that enhance modern metabolomics research[J]. *Anal Chem*, 2021, 93(1): 478–499.
- [34] Markley J L, Bruschweiler R, Edison A S, *et al.* The future of NMR-based metabolomics[J]. *Curr Opin Biotechnol*, 2017, 43: 34–40.
- [35] Giraudeau P. NMR-based metabolomics and fluxomics: developments and future prospects[J]. *Analyst*, 2020, 145: 2457–2472.
- [36] Ernst R R, Bodenhausen G, Wokaun A. Principles of nuclear magnetic resonance in one and two dimensions[M]. *Oxford: Oxford Science Publications*, 1987: 358–485.
- [37] Morris K F, Johnson C S. Diffusion-ordered two-dimensional nuclear magnetic resonance spectroscopy[J]. *J Am Chem Soc*, 1992, 114(8): 3139–3141.
- [38] Farjon J. How to face the low intrinsic sensitivity of 2D heteronuclear NMR with fast repetition techniques: go faster to go higher[J]. *Magn Reson Chem*, 2017, 55: 883–892.
- [39] Vitorge B, Bieri S, Humam M, *et al.* High-precision heteronuclear 2D NMR experiments using 10-ppm spectral window to resolve carbon overlap[J]. *Chem Commun*, 2009, 8: 950–952.
- [40] Mobli M, Hoch J C. Nonuniform sampling and non-Fourier signal processing methods in multidimensional NMR[J]. *Prog Nucl Magn Reson Spectrosc*, 2014, 83: 21–41.
- [41] Snyder D A, Bruschweiler R. Generalized indirect covariance NMR formalism for establishment of multidimensional spin correlations[J]. *J Phys Chem A*, 2009, 113(46): 12898–12903.
- [42] Snyder D A. Covariance NMR: Theoretical concerns, practical considerations, contemporary applications and related techniques[J]. *Prog Nucl Magn Reson Spectrosc*, 2021, 122: 1–10.
- [43] Robinette S L, Ajredini R, Rasheed H, *et al.* Hierarchical alignment and full resolution pattern recognition of 2D NMR spectra: application to nematode chemical ecology[J]. *Anal Chem*, 2011, 83(5): 1649–1657.
- [44] Smurnyy Y D, Blinov K A, Churanova T S, *et al.* Toward more reliable C-13 and H-1 chemical shift prediction: a systematic comparison of neural-network and least-squares regression based approaches[J]. *J Chem Inf Model*, 2008, 48: 128–134.
- [45] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. *Nature*, 2015, 521:

- 436-444.
- [46] Schmidhuber J. Deep learning in neural networks: an overview[J]. *Neural Networks*, 2015, 61: 85-117.
- [47] Zhang C, Idelbayev Y, Roberts R, *et al.* Small molecule accurate recognition technology (SMART) to enhance natural products research[J]. *Sci Rep*, 2017, 7(1): 14243. DOI: 10.1038/s41598-017-13923-x.
- [48] Chopra S, Hadsell R, LeCun Y. Learning a similarity metric discriminatively, with application to face verification[J]. *Proc CVPR IEEE*, 2005, 1: 539-546.
- [49] Reher R, Kim H W, Zhang C, *et al.* A convolutional neural network-based approach for the rapid annotation of molecularly diverse natural products[J]. *J Am Chem Soc*, 2020, 142(9): 4114-4120.
- [50] Egan J M, van Santen J A, Liu D Y, *et al.* Development of an NMR-based platform for the direct structural annotation of complex natural products mixtures[J]. *J Nat Prod*, 2021, 84(4): 1044-1055.
- [51] Duggan B M, Cullum R, Fenical W, *et al.* Searching for small molecules with an atomic sort[J]. *Angew Chem Int Ed*, 2020, 59(3): 1144-1148.
- [52] Grienke U, Foster P A, Zwirchmayr J, *et al.* ¹H NMR-MS-based heterocovariance as a drug discovery tool for fishing bioactive compounds out of a complex mixture of structural analogues[J]. *Sci Rep*, 2019, 9: 11113. DOI:10.1038/s41598-019-47434-8.
- [53] Marguerite L, Markov P, Chiron L, *et al.* Automatic differential analysis of NMR experiments in complex samples[J]. *Magn Reson Chem*, 2018, 56(6): 469-479.



【专家介绍】叶文才：教授、博士生导师，现任暨南大学副校长。曾获长江学者特聘教授、国家杰出青年科学基金。一直从事中药和天然药物活性成分及创新药物等研究工作。先后主持了国家重点研发计划等科研项目 40 余项。已在 *JACS*、*J Clin Invest*、*Angew Chem Int Ed*、*Hepatology*、*J Extracell Vesicles* 等国际知名杂志发表 SCI 收录论文 510 余篇；获授权国内外发明专利 60 余项，研发的 2 个候选新药已完成 III 期临床试验；以第一完成人获国家科技进步奖二等奖 1 项、广东省科学技术奖一等奖 3 项，并获中国专利优秀奖 1 项、中国产学研合作创新成果一等奖 2 项，主要成果入选 2017 中国十大医学进展。获第二届全国创新争先奖、中国药发展奖、广东省丁颖科技奖、国务院政府特殊津贴，并获全国模范教师、全国优秀科技工作者称号。

《药学进展》杂志 2022 年征订启事

《药学进展》杂志是国家教育部主管、中国药科大学和中国药学会共同主办的药学类科技期刊，1959 年创刊，2014 年全新改版，国内外公开发行人，是专注于医药科技前沿与产业动态的专业信息媒体。本刊以科学前沿与国家战略需求为宗旨，以综述、评述、行业发展报告为特色，以药理学学科进展、技术进展、新药研发各环节技术信息为重点，主要报道药理学学科链、研发技术链、医药产业链的国内外研究前沿与进展；围绕新药研发产业链，聚焦药理学学科进展、全球研发前沿、科研思路方法、靶点机制探讨、新药研发报告、临床用药分析、技术政策动态；希冀以期刊与论坛为平台，整合行业资源，发挥协同作用，推动新药研发与产业发展。

《药学进展》编委会由国家重大专项化学药总师陈凯先院士担任主编，编委由新药研发技术链政府监管部门、高校科研院所、制药企业、临床医院、CRO、金融资本及知识产权相关机构 200 余位极具影响力的专家组成。

《药学进展》设有“药咖论坛”“前沿与进展”“医药行业报告”“热点透视”“知识产权”“业界关注”“文献计量”“临床药学”等栏目。改版至今，组稿策划了“肿瘤药理学研究进展”“聚焦心脑血管疾病药物”“糖尿病药物研发策略”“靶向纳米递药系统的创新药物制剂设计”“化学探针在药学领域中的应用”“聚焦抗体药物研发”“新型麻醉药和麻醉相关药的研发”“聚焦肿瘤心脏病学”“神经药理学研究”“新药研发的新进展”等 60 余个专题，刊载了数十篇报道行业领域进展且极具学术价值的综述类文章，多位院士评述，充分发挥了《药学进展》作为专业媒体引领学术发展、服务科技的作用；与科睿唯安（原汤森路透）信息机构合作，独家分享全球新药研发报告，因内容全面、资料权威、视角新颖、观点独到、数据翔实、时效性强广受好评。目前刊物已在药理学学科进展、科研思路方法、靶点机制探讨、新药研发报告、临床用药分析、国际医药前沿等方面形成特色。

《药学进展》杂志为月刊，每期 80 页，铜版纸全彩印刷，国内外公开发行人，每期定价 40 元，全年定价 480 元。CN 32-1109/R，ISSN 1001-5094，国内邮发代号：28-112，欢迎广大读者向本刊编辑部或当地邮局订阅。

编辑部地址：南京市童家巷 24 号中国药科大学《药学进展》编辑部； 邮编：210009
电话 / 传真：025-83271227； E-mail: yxjz@163.com