# 上海有机所发展出多层代谢网络技术

作者：writer 来源：中国科学院

本文原地址：https://www.iikx.com/news/progress/20775.html

*本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！*

**上海有机所发展出多层代谢网络技术**

。中国科学院上海有机化学研究所生物与化学交叉研究中心研究员朱正江课题组在《自然-通讯》(Nature Communications)上，在线发表了题为Metabolite Annotation from Knowns to Unknowns through Knowledge-guided Multi-layer Metabolic Networking的研究论文。该工作发展出知识驱动的多层代谢网络技术(Knowledge-guided multi-layer networking，KGMN)，在复杂生物样本中实现了未知代谢物的大规模鉴定。

生命体代谢组包含内源性产生的已知代谢物，以及微生物菌群、植物、食物和其他来源的未知代谢物。这些未知代谢物对生命活动具有重要的调控作用。然而，如何大规模鉴定未知代谢物是当前代谢组学研究的热点和难点。前期研究中，朱正江课题组发展了基于代谢反应网络的代谢组学技术MetDNA(Nature Communications,2019, 10: 1516)。该技术利用代谢反应网络中产物和底物存在结构相似性与二级质谱图相似性的基本原理，设计了代谢物二级质谱图“ 谱图借用 ”、“ 多次迭代 ”算法，利用标准谱图库鉴定出的代谢物作为种子，依靠代谢反应网络进行代谢物注释的迭代和传递，突破了标准二级质谱图库的覆盖度限制，实现大规模的已知代谢物鉴定。然而，该技术仍存在一些局限：代谢物鉴定传递的过程仅限于已知代谢反应网络，因而无法用于发现新的未知代谢物;复杂质谱数据中的大量冗余信号(如同位素峰、加合物峰、中性丢失和源内裂解等)会对代谢物的鉴定造成假阳性。

针对以上问题，在MetDNA2中，科研人员进一步发展了知识驱动的多层代谢网络技术(KGMN)，实现了从已知代谢物鉴定未知代谢物的能力，并显著提升了代谢物鉴定的准确度。该技术首次整合了3层代谢网络——知识驱动的代谢反应网络、知识引导的二级质谱图相似性网络、全局代谢峰相关性网络。研究利用理论代谢反应对已知的代谢反应网络进行扩展，从而构建了包含已知和未知代谢物的扩展代谢反应网络(KMRN，网络1)。MetDNA2从标准谱图库鉴定出的种子代谢物出发，基于扩展代谢反应网络和“ 谱图借用 ”策略，构建二级质谱图相似性网络(网络2)。该策略可通过多次迭代和循环扩增的算法，将代谢组学质谱数据中所有已知和未知代谢物连接，直至没有新的注释代谢物。在网络2中，代谢物节点之间的连接有四个限制条件——MS1 m/z、保留时间、MS/MS谱图相似性和代谢反应转化(metabolic biotransformation)。对于注释的每一个代谢物，MetDNA2会进一步通过靶向检索其相关的冗余质谱特征峰(如同位素峰、加合物峰、中性丢失和源内裂解等)，并构建全局代谢峰相关性网络(网络3)。利用全局代谢峰相关性网络，MetDNA2对注释的代谢物结果进行全局优化，提升代谢物鉴定的准确度，去除假阳性注释结果返回最终鉴定结果。整个数据处理流程全程自动化，无需人工干预，提升了数据分析的效率。

利用上述技术，MetDNA2对已知代谢物的鉴定准确性从~70%提升至>95%。同时，在不同的生物样本中，MetDNA2还可鉴定~100-300个未知代谢物。MetDNA2对单个生物样本鉴定的代谢物数

目在2000-5000个左右。在MetDNA2中，每一个鉴定结果均根据国际代谢组学协会标准指定特定的可信度。此外，MetDNA2还包含一系列重要的更新和升级，如全面升级的标准代谢物谱图数据库(>2000个代谢物);两种不同色谱体系的保留时间数据库(HILIC和C18体系);适配所有厂商的高分辨二级质谱数据等。

为了相关研究便捷的应用这一工具，课题组提供了用户友好型的界面和网站MetDNA2(http://metdna.zhulab.cn/)，学术用户可免费注册使用。该工作开发的KGMN技术已申请国家发明专利和国家软件著作权。相关技术和软件的商业用途需要联系朱正江进行授权使用。研究工作得到国家自然科学基金、科技部、中科院、上海市科学技术委员会等的支持。

论文链接

研究团队单位：上海有机化学研究所

更多 科学进展 请访问 https://www.iikx.com/news/progress/

本文版权归原作者所有，请勿用于商业用途，爱科学iikx.com转发