上海有机所开发出四维代谢组学精准分析技术Met4DX

作者：writer 来源：中国科学院

本文原地址：https://www.iikx.com/news/progress/22817.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！

上海有机所开发出四维代谢组学精准分析技术Met4DX

。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*A mass spectrum-oriented computational method for ion mobility-resolved untargeted metabolomics*的研究论文。该工作针对基于离子淌度质谱的四维代谢组学技术开发了一种端到端的精准数据分析技术Met4DX，用于四维代谢组学数据的四维峰检测、峰对齐和峰定量，并结合四维数据库实现代谢物的四维精准匹配和鉴定（图1）。

离子淌度质谱相较于传统质谱增加了根据离子尺寸、形状以及电荷分离的离子淌度，有效提升了质谱的分离能力，特别是代谢物同分异构体的分辨能力，其跟液相色谱联用形成多维分离分析技术，可进一步提高复杂生物样本分析的分离度和峰容量。一次四维代谢组学分析能够同时对代谢物离子进行四个维度的表征，包括精确质量数（MS1）、二级质谱图(MS/MS)、色谱保留时间（RT）和离子淌度碰撞截面积（CCS），能有效提升对复杂生物样品中代谢物定性和定量分析的覆盖度和准确度。然而，四维代谢组数据的高度复杂性对数据的高效精准分析提出了挑战，尤其是四维质谱峰的检测仍是难点。因此，四维代谢组数据分析技术和工具相对有限。目前，少量工具如MS-DIAL与MZmine等均采用自上而下压缩数据的降维策略进行峰检测。该降维策略可以降低数据的维数和复杂性，但降维过程也不可避免低引入了信号掩蔽及干扰，显著降低了四维峰检测的灵敏度。

液相色谱-离子淌度-质谱依次从液相、离子淌度以及质谱维度对代谢物实现多维分离，所需要的时间也逐级减小。受到该分离方式的启发，本研究开发了从一张质谱图出发的自下而上峰组装算法（mass spectrum oriented bottom-up assembly algorithm）用于四维代谢组学数据中四维峰的检测（图2）。该技术的特点是将每一张质谱图作为四维数据中的最小数据单元，采用逆向工程的策略依次构建其在离子淌度和液相色谱上的峰形。具体地，对于每一张质谱图，该算法会依次进行其前体离子的搜索、离子淌度流出峰的组装与检测、离子淌度流出峰的拓展、色谱流出峰的组装与检测和四维峰积分等五个步骤实现四维代谢峰检测与峰定量。本工作开发的自下而上的峰组装算法避免了数据压缩与降维，有效地提高了四维峰检测的覆盖度与灵敏度。以上述算法为核心，科研人员进一步开发了适用于四维代谢组学的端到端的精准数据分析技术Met4DX，通过二级谱图去冗余模块、自下而上的峰组装模块、四维峰对齐以及分组模块、代谢物的多维匹配与鉴定模块等实现了的四维复杂代谢组的精准定性和精确定量分析。

Met4DX技术能够实现高覆盖的四维质谱峰检测，定量精密度高。与同类技术相比（如MS-DIAL和MetaboScape），Met4DX能够提升四维峰检测的覆盖度2-3倍，提升准确定量代谢物的数目2-5倍。Met4DX在代谢物同分异构体识别上具有优异的性能。以在小鼠肝脏代谢组为例，Met4DX精准识别代谢物同分异构体数目高达3033对，比同类技术显著提升3.6倍，且可准确识别出CCS差异为1%的共流出同分异构体。同时，该研究还收集了HMDB和KEGG中的超过13万个代谢物，建立了目前最全面的四维代谢物数据库用于代谢物的多维匹配与鉴定。

目前，Met4DX支持包括布鲁克timsTOF和安捷伦DTIM-MS等仪器平台采集的四维代谢组学数据。为了方便相关领域研究应用该工具，课题组提供了docker供学术界用户免费使用Met4DX（https://hub.docker.com/r/zhulab/met）。该工作开发的四维代谢组学精准分析技术M4dx-ret4DX已申请国家发明专利和国家软件著作权。研究工作得到国家自然科学基金、科技部、中科院和上海市科学技术委员会等的支持。

论文链接

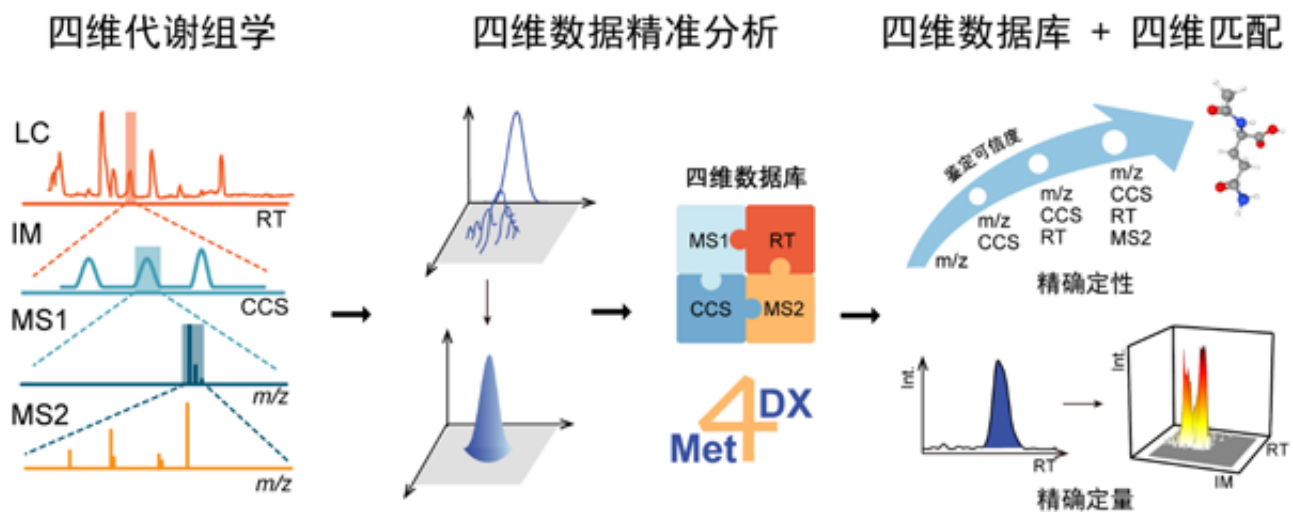


图1.四维代谢组学精准数据分析技术Met4DX

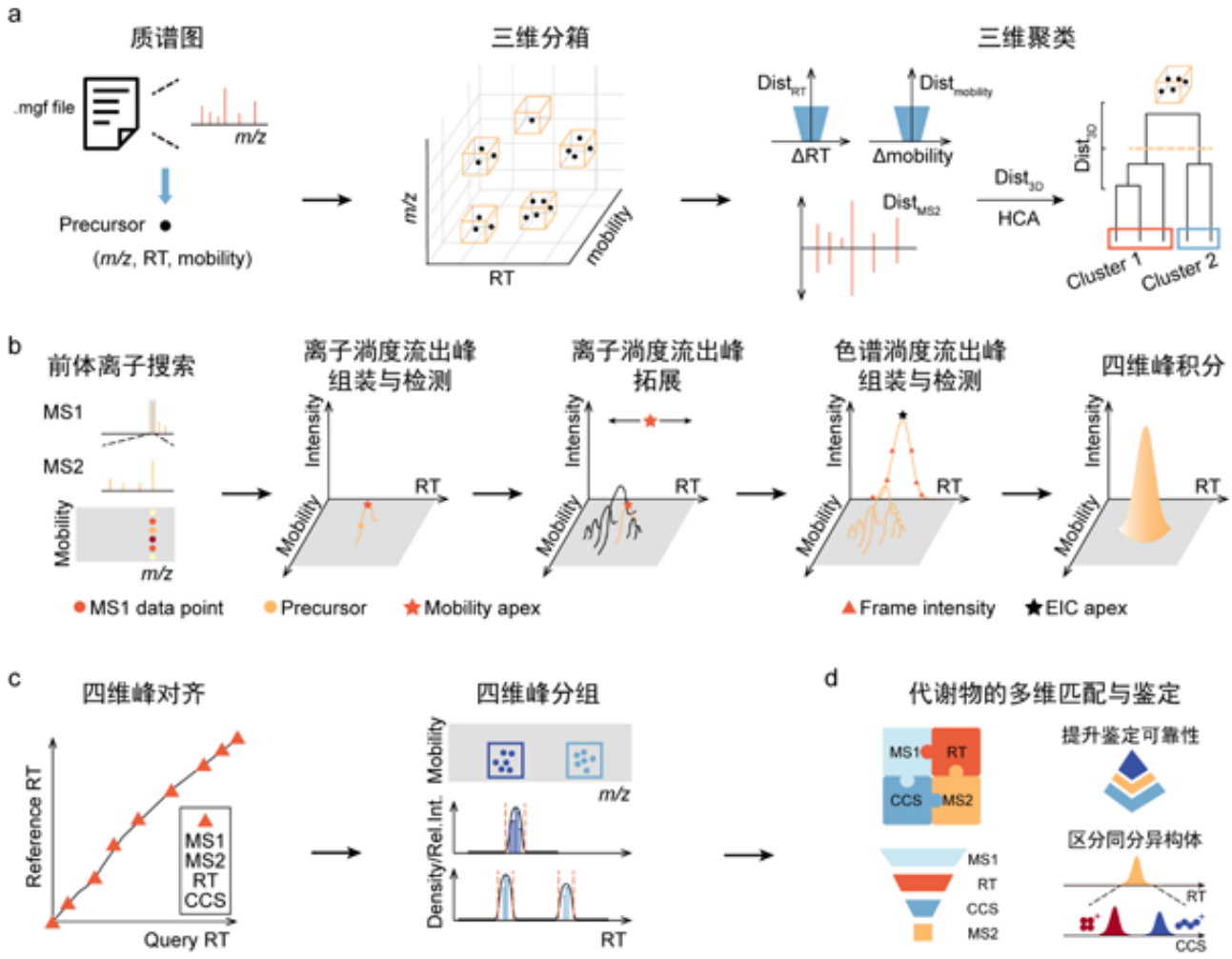


图2.自下而上的峰组装算法和Met4DX数据分析流程（图片修改自《自然-通讯》）

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