大连化物所发布代谢组学数据库1.0版

作者：writer 来源：科学网

本文原地址：https://www.iikx.com/news/progress/43137.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！

大连化物所发布代谢组学数据库1.0版。近日，中国科学院大连化学物理研究所（以下简称大连化物所）研究员刘心昱与研究员许国旺团队，在前期三个月试运行的基础上，正式发布代谢组学数据库（Met-SC）1.0版。该数据库由大连化物所自主建立，集数据标准化存储、AI-就绪（AI-ready）质谱资源构建、在线数据分析与疾病导向知识图谱挖掘于一体，实现了代谢组学数据的存—算—析一体化。该数据库为代谢组学数据的存储管理、共享应用与规模化智能分析提供了基础设施支撑。相关成果发表在《核酸研究》。

随着高分辨质谱技术的飞速发展，代谢组学数据正以指数级速度迅速增长。HMDB、MassBank等数据库已极大推动了代谢组学数据的存储、共享与分析。然而，数据存储、分析及功能解析仍分散在不同资源中，研究者往往需在多个数据库系统间反复搬运数据与结果。跨平台操作不仅抬高了可重复分析的门槛，也制约了跨研究整合。随着代谢组学数据规模持续膨胀、人工智能应用的快速渗透，存—算—析割裂的格局已成为制约生物学洞见高效产出的瓶颈，亟需一个将数据提交、在线分析与生物学解释贯通的一体化数据库。

针对上述问题，研究团队在国家基因组科学数据中心的批准下，承担了代谢组学分中心的建设，自主研发了Met-SC数据库，并构建了基于FAIR原则的标准化数据汇交框架。团队整理整合了199万种化合物的775万张参考质谱图，以及10万余份真实生物样本中的7300万张未知特征质谱，构成面向AI-ready的代谢组学质谱数据底座。该数据库内置零代码在线分析模块，降低了组学数据分析门槛，还整合了结直肠癌代谢知识图谱，将分散文献证据与多组学数据贯通串联，为疾病机制发现与靶点挖掘提供系统化知识支撑。

Met-SC 1.0版的发布，实现了数据存储、谱图资源管理、在线计算与知识图谱探索的一体化集成，在一定程度上缓解了跨平台数据壁垒问题。（来源：中国科学报 孙丹宁）

相关论文信息：https://doi.org/10.1093/nar/gkag1127

作者：刘心昱等 来源：《核酸研究》

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