

应用纪要

利用Progenesis QI和天然产物数据库 Natural Product Atlas从头发现天然产物

Suraj Dhungana, David Heywood, Jeff Goshawk, Giorgis Isaac

Waters Corporation



仅供研究使用，不适用于诊断。

这是一份应用简报，不包含详细的实验部分。

摘要

科学界和制药界对发现具备生物和药理活性的新型化合物有着浓厚的兴趣。多年来，研究人员一直在寻找各种来源的天然产物（即植物、微生物、海洋生物等来源），目的是发现新的疗法或对健康有益的化合物。鉴定新化合物或现有化合物的新活性需要全面的工作流程。工作流程早期需要优化提取方法并进行活性检测，而完成工作流程需要使用复杂的分离、检测和表征工具。液相色谱与质谱联用(LC-MS)是一种有效的方法，可以分离和表征天然产物提取物中存在的化合物。具有生物活性的提取物通常都很复杂，在高分辨率Q-Tof仪器上进行串联质谱分析之前，需要先进行在线色谱分离。完成色谱分离之后，在气相条件下利用离子淌度技术分离分子离子，证明分离度提高和碰撞截面(CCS)值增加对化合物鉴定有用，更加有利于复杂样品的分析。手动解析复杂的天然产物数据集不具备可行性，需要借助信息学解决方案以及完善的化合物数据库来简化数据处理和化合物鉴定流程。本文重点介绍的发现工作流程讨论了使用Natural Product Atlas数据库（收录了>25,000种化合物）结合Progenesis QI信息学解决方案发现和自动鉴定天然产物化合物的方法。

优势

- 整合UPLC Q-Tof质谱仪和离子淌度工作流程与NP Atlas数据库，用于天然产物研究
- 使用Progenesis QI中的NP Atlas数据库开发完整的发现解决方案
- 使用批量搜索自动鉴定化合物
- 建立自定义数据库，实现可靠的化合物鉴定
- 进行聚类分析、节点分析，或查询GNPS数据库

简介

推进天然产物的研究对于寻找具有治疗价值的新型化合物和了解天然产物用作传统药物的健康益处非常关键。由于样品的复杂性，因此需要通过复杂的流程来表征天然产物中的目标分子。分离特定提取物是降低分析复杂性的标准第一步。尽管已执行多次分离，但天然产物提取物的馏分中仍可能包含数百种分子，需要使用先进的分析工具进一步分离和检测，还需要简化的工作流程和相关谱库以实现高效表征。

在天然产物发现中使用液相色谱与高分辨率质谱联用(LC-MS)工作流程（图1），可提供获得化学和结构复杂性信息所需的分析分离和高分辨率串联质谱数据。我们通过此方法将复杂的天然产物提取物分离成单独的组分，获取有关母离子和碎片离子的精确质量数信息用于结构解析。使用基于离子淌度技术的仪器在漂移时间维度上进一步分离分子离子，此方法可与色谱分离正交，提供碰撞截面(CCS)信息用于化合物鉴定。捕获数据后

，将数据导入Progenesis QI软件进行分析，通过适当的谱库匹配鉴定已知化合物，或通过数据库搜索和网络/聚类分析以缩小从头发现候选物的范围。

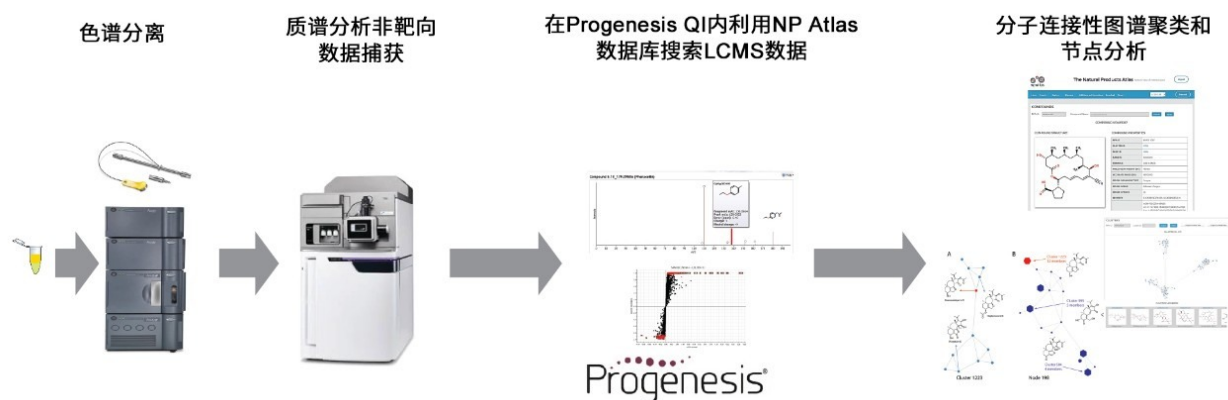


图1.使用Progenesis QI信息学解决方案中的天然产物数据库进行天然产物研究的从头发现工作流程。

可访问且兼容LC-MS数据的谱库和数据库对于及时、可靠地完成天然产物发现项目至关重要。Natural Products Atlas ([npatlas.org](http://www.npatlas.org) <<http://www.npatlas.org/>>)来自世界各地的数据管理者联合创建，并由西蒙弗雷泽大学Roger Linington教授研究小组的研究人员负责维护，是天然产物研究领域的重要资源，为天然产物研究的推进提供了一个巨大的知识共享平台¹。Natural Products Atlas (NP Atlas)中涵盖经同行评审的主要科学文献中发表的所有微生物来源的天然产物，截至2021年6月，其最新版本中含有超过>29,000种化合物（约11,000种细菌来源的化合物和18,000种真菌来源的化合物），该数据库中包含分类信息，可用于查询不同物种来源化合物的分布。NP Atlas还兼具搜索、探索和发现功能，可以帮助研究人员进一步了解相关知识。用户可以根据化合物结构、化合物名称、分子量、化学式、InChiKey、SMILES等信息，或使用结构绘图工具来检索化合物。NP Atlas上的探索功能可以对结构相似化合物进行聚类 and 节点分析，而发现功能则可以从作者、期刊等其他角度查询化合物。

开放式NP Atlas数据库虽然是用于天然产物发现的巨大资源库和工具，但用户无法使用天然产物提取物生成的LC-MS数据进行直接查询。沃特世与NP Atlas合作开发了与Progenesis QI和UNIFI软件兼容的桌面或本地版NP Atlas数据库。Progenesis QI和UNIFI版本的NP Atlas库均可在marketplace.waters.com <<http://marketplace.waters.com>> 网站下载。这一开发项目实现了NP Atlas搜索和化合物鉴定过程的自动化，赋予了推进天然产物研究迫切需要的灵活性。除NP Atlas之外，还可以在Waters Marketplace免费下载其他天然产物数据库，并用于Progenesis QI和UNIFI。

结果与讨论

在从头发现期间，工作流程中的关键步骤之一是化合物鉴定。化合物鉴定是天然产物工作流程中的瓶颈，使用一个综合数据库对于有效鉴定化合物至关重要。Progenesis QI版本的NP Atlas专门针对此工作流程开发，旨在实现鉴定过程自动化。将Progenesis QI版本的NP Atlas设置为化合物数据库，并指定母离子质量数偏差和理论碎片偏差的搜索参数，二者可以在一个步骤中实现（图2）。Progenesis中的搜索功能还可以根据化学式计算理论同位素分布，并在化合物鉴定和评分过程中加以考虑。图3A显示了使用精确质量数、理论碎片和同位素相似性评分鉴定红霉素G的示例。完成鉴定后，导出红霉素G的实验数据，生成内部自定义数据库，该数据库现在包含实验得到的MS和MS/MS谱图、LC保留时间(RT)，以及在基于离子淌度技术的HRMS实验期间生成的CCS信息（图3B）。

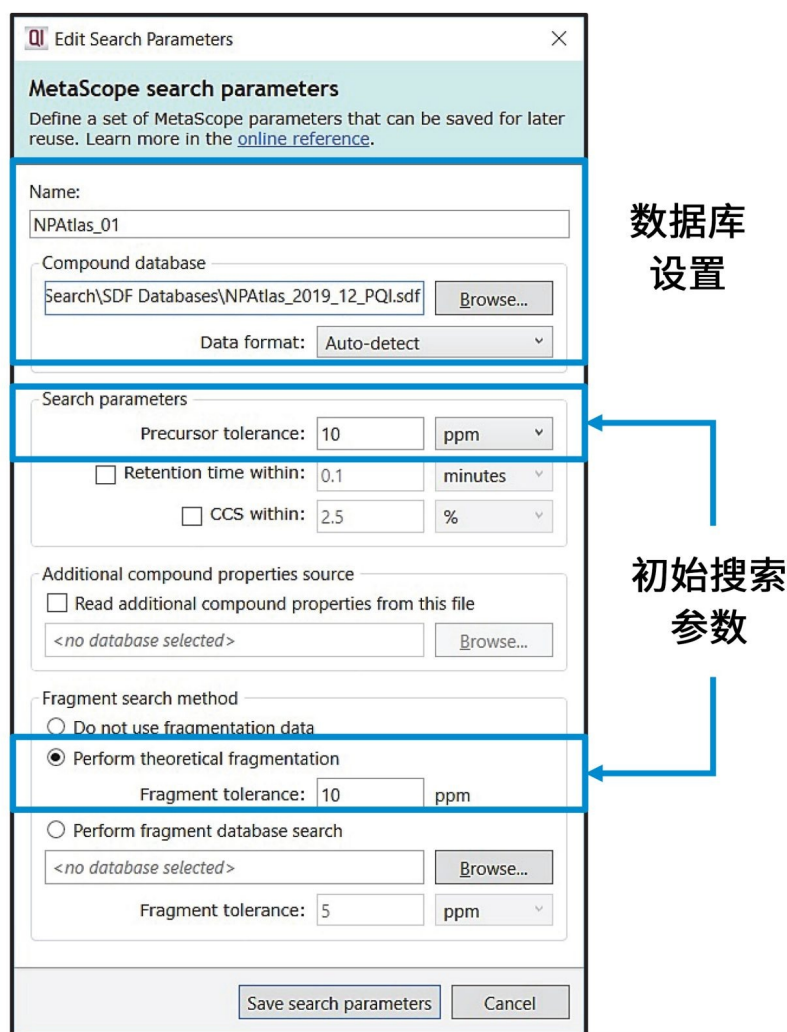


图2.在Progenesis QI中添加NP Atlas作为数据库并设置初始搜索参数的简单方法。

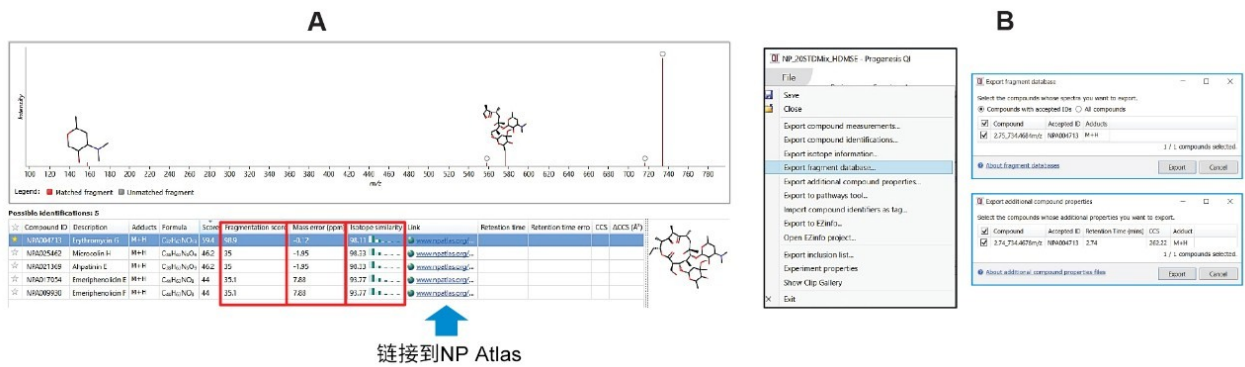


图3. A)使用NP Atlas作为Progenesis Q1中的数据库进行未知化合物鉴定，B)导出实验得到的MS/MS谱图、RT和CCS参数以创建内部自定义数据库。

创建自定义数据库后，可以根据自定义数据库搜索下一组发现样品，进行可靠的化合物鉴定。在基于离子淌度技术的工作流程中，自定义数据库使用了5个参数来鉴定化合物：母离子精确质量数、同位素相似性评分、RT、精确质量数碎片离子信息和CCS。如果在未使用离子淌度技术的情况下生成自定义数据库，则使用4个参数进行鉴定。添加RT和CCS参数（在离子淌度实验中）可提供正交信息，增加鉴定结果的可信度。图4展示了一个使用自定义创建的数据库执行搜索的示例，本示例再次观察了红霉素G的鉴定过程，但使用的是自定义数据库来执行搜索。该搜索的数据库匹配参数包括母离子精确质量数、同位素相似性评分、RT、精确质量数碎片离子信息和CCS值。确认鉴定结果后，可以通过Progenesis中的超链接访问NP Atlas网站上的特定化合物页面，利用NP Atlas中的所有可用功能探索和发现化合物，进一步解析化合物。NP Atlas还提供其他外部链接，例如GNPS和MIBiG。图5显示了NP Atlas与GNPS连接的示例，用于进一步查询GNPS内的化合物，进行分子网络分析（图5）。

Edit Search Parameters

MetaScope search parameters
Define a set of MetaScope parameters that can be saved for later reuse. Learn more in the [online reference](#).

Name: NPAAtlas_02

Compound database
Search\SDF Databases\NPAAtlas_2019_12_PQI.sdf
Data format: Auto-detect

Search parameters

Precursor tolerance: 5 ppm

Retention time within: 0.2 minutes

CCS within: 2.0 %

Additional compound properties source
 Read additional compound properties from this file
arch\SDF Databases\NPAAtlas_std_RTandCCS.csv

Fragment search method
 Do not use fragmentation data
 Perform theoretical fragmentation
 Perform fragment database search

Fragment tolerance: 10 ppm

arch\SDF Databases\NPAAtlas_std_Fragments.msp

Fragment tolerance: 5 ppm

搜索参数

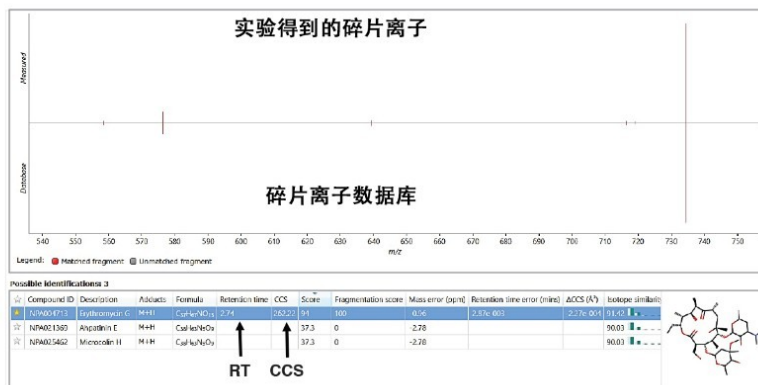


图4.在基于离子淌度技术的仪器上设置搜索参数并使用自定义创建的数据库执行搜索。

物鉴定总共使用4个参数，而不是前面讨论的5个参数。

结论

Natural Product Atlas为天然产物研究领域提供了优质资源。为了进一步发挥NP Atlas的优势，开发了桌面版NP Atlas数据库。桌面版的主要优点是可以兼容LC-MS数据，并且能够直接使用NP Atlas搜索LC-MS数据进行化合物鉴定，而OpenAccess版本目前不提供该功能。该桌面版数据库的兼容性还体现在可以根据母离子精确质量数、理论碎片离子和理论同位素分布、保留时间和通过离子淌度实验获得的CCS值进行化合物批量搜索和自动鉴定。桌面版NP Atlas数据库免费提供，用户可以访问marketplace.waters.com <<http://marketplace.waters.com>> 网站进行下载。

参考资料

1. Santen, Jeffrey A. van, Grégoire Jacob, Amrit Leen Singh, Victor Aniebok, Marcy J. Balunas, Derek Bunsko, Fausto Carnevale Neto, et al. “The Natural Products Atlas: An Open Access Knowledge Base for Microbial Natural Products Discovery.” ACS Central Science 5, no.11 (November 27, 2019): 1824–33.<https://doi.org/10.1021/acscentsci.9b00806> <<https://doi.org/10.1021/acscentsci.9b00806>> .

致谢

感谢西蒙弗雷泽大学的Roger Linington教授、Jeffrey van Santen和Linington研究小组在本项目上与我们的合作。

特色产品

ACQUITY UPLC I-Class PLUS系统 <<https://www.waters.com/134613317>>

SYNAPT XS高分辨率质谱仪 <<https://www.waters.com/waters/nav.htm?cid=135020928>>

Xevo G2-XS QToF四极杆飞行时间质谱仪 <<https://www.waters.com/134798222>>

MassLynx MS软件 <<https://www.waters.com/513662>>

Progenesis QI <<https://www.waters.com/134790652>>

720007297ZH, 2021年6月

© 2021 Waters Corporation. All Rights Reserved.