
蛋白质组学，是时候告别“基因推论”回归“真实”了——期刊主编深度反思 MDPI Proteomes

作者：writer 来源：科学网

本文原地址：<https://www.iikx.com/news/progress/38780.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！

蛋白质组学，是时候告别“基因推论”回归“真实”了——期刊主编深度反思 MDPI Proteomes。论文标题：Proteomics—The State of the Field: The Definition and Analysis of Proteomes Should Be Based in Reality, Not Convenience

论文链接：<https://www.mdpi.com/2227-7382/12/2/14>

期刊名：Proteomes

期刊主页：<https://www.mdpi.com/journal/proteomes>

一篇发表在Proteomes的观点文章，标题直指学科核心问题：蛋白质组学的研究现状：蛋白质组的定义和分析应基于现实，而非便利。

这篇文章由期刊主编Jens R. Coorsen教授和Matthew P. Padula副教授共同撰写，对当前蛋白质组学研究的方法论基础提出了深刻反思。文章发表后引起了广泛讨论，因为它触及了一个学科长期以来回避的问题：我们真的在研究真实的蛋白质吗？

一、文章核心观点：我们研究的是真实的蛋白质吗？

目前大多数蛋白质组学研究，依赖从基因组序列推导出的参考蛋白质组数据库进行质谱数据搜库。这种做法虽然便捷，但默认了一个前提：蛋白质就是基因序列预测的那个样子。

然而，事实远非如此简单。

一个基因可以通过可变剪接产生多种mRNA，每种mRNA翻译出的蛋白质还可以经历数十种翻译后修饰——磷酸化、糖基化、乙酰化、泛素化、甲基化、SUMO化等等。这些不同形式的蛋白质被称为蛋白质变体（proteoforms）。理论上，每个细胞类型可能存在超过100万种蛋白质变体，而我们实验检测到的，只是冰山一角。

文章指出，如果我们只研究由基因序列推论出来的蛋白质，而不是真实存在的蛋白质变体，可能会错过真正关键的生物学信息。作者用了一个生动的比喻：这就好比根据建筑设计图来研究一座建筑，却忽视了建筑实际使用过程中留下的痕迹、改造和磨损——而这些往往才是理解建筑功能

的关键。

真正的蛋白质组学，应该研究这些真实存在的蛋白质变体（proteoforms），而不是数据库里的标准答案。

二、为什么这个问题值得关注？

这个问题不仅仅是学术思辨，它直接关系到研究的可靠性和可重复性。

想象一下：如果一种疾病的标志物是某个蛋白质的特定磷酸化形式，而您检测的是该蛋白的总量，结果可能就是无显著差异——即使真正相关的那个变体已经发生了巨大变化。同样，如果一个药物靶点的功能依赖于某种特定修饰，而您在筛选药物时只关注蛋白总量，就可能错过真正有效的候选分子。

这正是很多标志物研究和药物靶点研究难以复现的可能原因之一：用检测总量的方法，去衡量本应关注变体的指标。

文章进一步指出，蛋白质变体的复杂性还体现在另一个层面：不同个体之间，同一蛋白质的修饰谱可能存在差异。这种差异可能是个体化表型的分子基础，也可能是个体对药物反应差异的原因。如果研究停留在通用蛋白层面，这些个体化信息就被完全抹去了。

作者呼吁，未来的研究需要在蛋白质变体水平上进行更精细的分析。这不仅需要技术上的突破（更灵敏的检测方法、更好的富集策略），更需要研究思路的转变——从找到了多少蛋白转向找到了什么形式的蛋白。

三、如何实现这种转变？

文章虽然没有给出具体的技术方案，但指出了几个可能的方向：

第一，数据库的改进。现有的参考蛋白质组数据库需要向更全面的蛋白质变体数据库升级，纳入已知的修饰信息和变体形式。

第二，搜库算法的优化。需要开发能够更好匹配修饰肽段的算法，减少因修饰导致的漏检。

第三，实验设计的调整。在研究特定生物学问题时，应有意识地设计针对特定修饰类型的富集策略，而不是只做全蛋白组扫描。

第四，数据报告的标准化。未来的研究论文应该更详细地报告所检测到的蛋白质变体信息，而不是只报告基因水平的鉴定结果。

四、学科启示：回归真实，是蛋白质组学的必由之路

We choose to go to the Moon in this decade and do the other things, not because they are easy, but because they are hard.

文章引用肯尼迪总统的这句名言，正是对蛋白质组学未来使命的最好诠释：

解析蛋白质组的真实复杂性，无疑是生命科学领域一项艰难的任务，但这份艰难的背后，是更精准的疾病诊断、更可靠的生物标志物、更有效的药物研发，以及对生命本质更深刻、更接近真相的认知。

只有放弃便捷主义，回归

真实世界，蛋白质组学才能真正走向成熟，为生命科学与医学带来真正颠覆性的价值。

五、延伸阅读：本期推荐的两篇相关文章

如果您对这个问题感兴趣，以下两篇同期发表的文章也从不同角度提供了有价值的视角：

1. 转化临床研究：高通量组学整合的力量

<https://www.mdpi.com/2227-7382/12/3/25>

这是一篇系统性的综述，发表于2024年9月。文章探讨了如何通过高通量多组学整合技术来应对生物系统的复杂性，推动临床研究的转型。

文章详细介绍了当前主流的高通量组学技术平台：用于基因组学的下一代测序（NGS）、用于转录组学的RNA测序（RNA-Seq）、用于蛋白质组的质谱技术、用于代谢组的核磁共振（NMR）波谱等。这些技术能够在不同分子层面产生海量数据，但如何将这些数据整合起来，形成对生物系统的整体理解，是当前面临的核心挑战。

文章重点介绍了多种数据整合策略：

- 基于相似性的方法：如相似性网络融合（SNF），通过构建各组学层的相似性网络，再整合为一个统一网络，识别共同通路
- 基于差异性的方法：如多组学因子分析（MOFA），识别不同组学层中与疾病相关的独特特征
- 机器学习方法：如LASSO和随机森林，用于从多组学数据中筛选最相关的特征

文章还介绍了多个实用的多组学分析工具，如OmicsNet（用于生物网络的可视化分析）和NetworkAnalyst（提供数据过滤、标准化、统计分析和网络可视化功能）。这些工具为研究者整合多组学数据、挖掘生物学意义提供了便利。

文章最后指出，将多组学数据与电子健康记录（EHR）整合，结合云计算和大数据分析技术，是未来精准医学发展的重要方向。尽管目前仍面临数据复杂性、技术局限性和伦理问题等挑战，但高通量组学整合技术在推动个性化医疗、改善临床预后方面的潜力值得期待。

2. 来自蛋白质组学工作台的观察

<https://www.mdpi.com/2227-7382/12/1/6>

这是一篇实践经验型的观察性文章，发表于2024年2月。作者以第一人称的视角，分享了一位资深蛋白质组学研究者日常实验中积累的观察与反思，文章语言朴实、紧贴实验细节。

文章涉及的核心内容包括：

- 样品制备环节：不同裂解液对特定类型蛋白的提取效率差异，如何根据研究目的选择合适的方法
- 质谱参数设置：一些容易被忽略但影响数据质量的细节，如离子传输管的清洁频率、校准液的更新周期
- 数据处理陷阱：常见的数据解读误区，如如何区分真正的低丰度蛋白和噪声信号
- 实验设计建议：针对不同研究目的（标志物发现、机制研究、验证实验）的质谱策略选择

作者特别强调了一个观点：蛋白质组学不仅仅是用质谱跑样本，对样本本身的理解往往决定研究的成败。比如，组织样本的取材部位、存储时间、冻融次数，都可能对蛋白质谱产生显著影响，这些细节需要在论文中充分报告，以保证研究的可重复性。

这篇文章对正在开展实验的研究生尤其有参考价值，可以帮助他们少走一些弯路。作者在文末写道：希望这些来自实验台的观察，能让新手们避开我们曾经踩过的坑。

Proteomes 期刊介绍

主编：Jens R. Coorssen, Brock University, Canada; Matthew P. Padula, The University of Technology Sydney, Australia

期刊专注于蛋白质组分析的各个方面，特别关注蛋白质组在蛋白质形式（proteoforms）、功能性生物学单元以及标准氨基酸序列水平上的定量和表征。目前已被Scopus、ESCI (Web of Science)、PubMed等数据库收录。

2024 Impact Factor: 3.6

2024 CiteScore: 7.2

Time to First Decision: 28.6 Days

Acceptance to Publication: 5.6 Days

特别声明：本文转载仅仅是出于传播信息的需要，并不意味着代表本网站观点或证实其内容的真实性；如其他媒体、网站或个人从本网站转载使用，须保留本网站注明的“来源”，并自负版权等法律责任；作者如果不希望被转载或者联系转载稿费等事宜，请与我们接洽。

来源：Proteomes

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](http://iikx.com)转发