
研究开发基于深度学习的RNA多类型修饰解析算法

作者：writer 来源：科学网

本文原地址：<https://www.iikx.com/news/progress/38395.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！

研究开发基于深度学习的RNA多类型修饰解析算法。中国科学院动物研究所研究员赵方庆团队开发基于纳米孔RNA直接测序技术与深度学习策略的RNA修饰图谱解析算法ORCA。相关研究发表于《自然-通讯》。

基于深度学习的RNA修饰系统识别与注释模型 论文作者供图

RNA修饰对RNA的剪接加工、出核转运、以及RNA的稳定性和翻译效率有着重要的调控作用。但现有研究难以实现多种RNA修饰的同时检测。如何在同一转录本上系统解析不同修饰的分布模式，探究修饰间的协同或竞争关系，并阐明它们如何共同调控RNA剪接加工等关键生物学过程，仍是该领域面临的重要挑战。

近年来，新兴的纳米孔RNA直接测序技术为系统检测RNA修饰等核酸化学结构变化提供了重要

技术基础。然而，现有基于纳米孔RNA直接测序的修饰识别算法多依赖于体外转录合成的修饰训练集，仅能识别有限修饰类型；或基于不同样本间电信号水平变化进行修饰差异比较，难以实现对RNA修饰图谱及其互作模式的全景检测，极大限制了RNA修饰研究的深入开展。

ORCA通过创新的信号多态性特征提取与域对抗学习策略，克服了RNA修饰检测对特定修饰类型训练集的依赖问题。ORCA仅需使用6种体外转录合成的RNA修饰数据集进行训练，即可在mRNA及核糖体RNA中对超过15种修饰类型的系统识别及化学计量比预测，大幅拓展了现有DRS数据中可解析的RNA修饰类型范围。

在此基础上，研究团队进一步建立了基于迁移学习的RNA修饰类型注释方法，实现了多种已知RNA修饰类型的可靠注释。同时，通过对背景修饰位点的随机采样作为负样本训练，有效避免了对未知新型修饰类型的错误分类。在保证结果可靠性的同时，显著扩充了目前已知RNA修饰位点的数量。

研究团队进一步利用SGNex项目中多个细胞系的RNA直接测序数据，构建了跨细胞类型的RNA修饰全景图谱，发现RNA修饰位点在转录本上呈现系统的成簇分布。并进一步建立了修饰簇互作识别模型，在单分子水平上对邻近修饰位点间的协同或互斥发生进行系统鉴定，发现转录本异构体特异的修饰位点附近存在剪接调控因子及修饰相关RNA结合蛋白结合位点的显著富集，提示RNA修饰与剪接加工过程之间存在系统性的功能关联，为深入研究转录组的表观修饰多样性及其复杂的调控机制提供了重要的研究思路。

ORCA方法通过纳米孔RNA直接测序技术与人工智能算法的深度融合，实现了RNA修饰检测种类与数量的显著提升。该研究进一步揭示了RNA修饰位点的系统性邻近分布及其协同调控模式，为解码RNA修饰的动态互作及其在不同生物学背景下的调控机制提供了具有高度通用的计算生物学工具。（来源：中国科学报 王兆昱）

相关论文信息：<https://doi.org/10.1038/s41467-026-68419-y>

作者：赵方庆等 来源：《自然—通讯》

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](https://www.iikx.com)转发