

---

# 研究开发出高效环形RNA检测与定量工具

作者：writer 来源：中国科学院

本文原地址：<https://www.iikx.com/news/progress/35939.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！**

## 研究开发出高效环形RNA检测与定量工具

。近日，中国科学院北京基因组研究所（国家生物信息中心）高远团队与动物研究所赵方庆团队，开发出面向TB级转录组数据的高效环形RNA检测与定量工具CIRI3，通过反向剪接序列比对与跨样本整合算法设计，实现了TB级数据的超高速处理，并可高灵敏识别低丰度及非传统剪接信号的新型环形RNA，突破了环形RNA大规模检测方法的灵敏度与计算效率瓶颈。

针对现有环形RNA计算工具效率低、内存占用大等问题，CIRI3采用动态多线程任务划分与分块查找策略，实现了TB级测序数据的超高效处理（处理21TB SAM文件仅需105小时，峰值内存占用45.85 GB）。CIRI3可实现对大规模样本的联合分析。相较于单一样本分析，联合分析可有效防止“漏检”，使得环形RNA的检出率提升19.3%，并提升了低丰度环形RNA的检测灵敏度与定量精度。同时，CIRI3对非经典剪接信号等关键特征具有高度的可扩展性，实现了对内含子自连接型环状RNA等新亚型的系统性识别。综合性能评估表明，CIRI3在环形RNA的识别与定量方面均优于现有工具，可为大规模环形RNA研究及疾病生物标志物的挖掘提供技术支撑。

基于CIRI3的大规模数据挖掘能力，研究团队对涵盖30种癌症类型的2535个人类癌症及正常样本转录组测序数据进行整合挖掘，构建出迄今规模最大的AI-ready肿瘤相关环形RNA数据资源CIRI onco。评估表明，CIRI3计算的环形RNA剪接比率在不同批次和不同来源的样本间表现出高度的稳健性，其批次效应低于传统的反向剪接读段数，使其成为跨批次、跨平台整合分析的更优指标。进一步，基于差异剪接的环形RNA，团队构建了基于人工智能的组织类型溯源与疾病表型预测模型，实现了样本来源器官系统、组织类型及疾病水平的精准预测，展示了环形RNA作为疾病分子分型与早期诊断生物标志物的潜力。该模型可在用户小规模数据集上迁移学习，实现优于从头训练模型的预测性能，为解析环形RNA的生成调控与功能机制提供了计算工具。

CIRI3是可在TB级人群队列转录组数据中高效检测和精确定量环形RNA的工具；CIRI onco数据库为癌症相关环形RNA的研究与功能探索提供了AI-ready数据集与计算模型，并为环形RNA在癌症分型和精准诊断中的应用提供了方法学工具和数据基础。

相关研究成果发表在《自然-生物技术》（Nature Biotechnology）上。

研究工作得到国家重点研发计划、国家自然科学基金、中国科学院战略性先导科技专项等的支持。

[论文链接](#)

---

[CIRI3源代码、在线分析网页](#)

[CIRIlonco数据库](#)

CIRI3方法原理

CIRIlonco数据库及组织类型溯源与疾病表型预测模型

研究团队单位：北京基因组研究所

---

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](http://www.iikx.com)转发