
植物所开发出基因组组装质量评估和改进新方法

作者：writer 来源：中国科学院

本文原地址：<https://www.iikx.com/news/progress/24658.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！

近年来，随着测序技术和算法的开发，大量动植物基因组被陆续测序和组装，但基因组组装质量参差不齐，存在不同程度的组装错误，影响后续的相关研究。高质量的参考基因组对于基因的精准注释和功能研究以及比较基因组学和调控元件的挖掘等至关重要。目前已有一些基因组组装质量评估的方法，而多数方法仅提供概述性的评估值，未能针对特定区域设置特定碱基给出精准度的评估。

中国科学院植物研究所焦远年研究组研究开发了新的不依赖参考基因组的组装质量评估工具CRAQ（Clipping information for Revealing Assembly Quality），可以在单碱基水平检测和评估基因组序列的精准度，并提供了相关纠错方案。

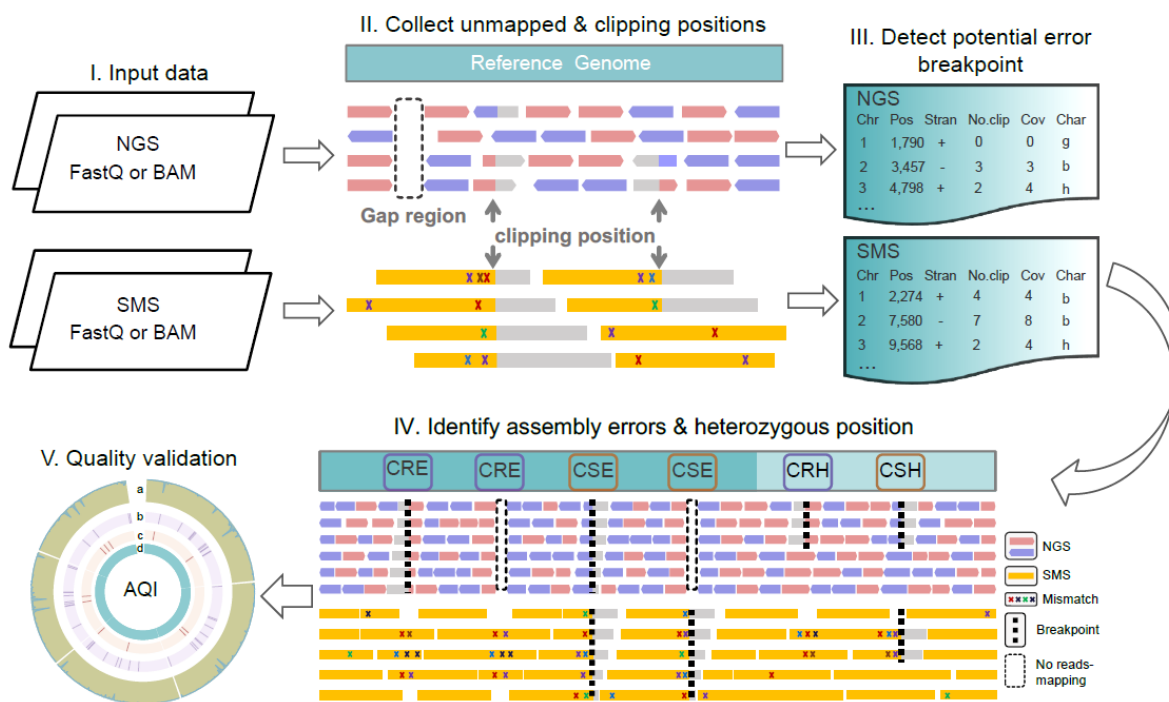
CRAQ通过将原始测序序列比对到组装的基因组上，基于序列比对产生的有效剪切对齐（clipping alignment）信息，精准地检测基因组中存在的组装错误。结合SMS长读长测序片段和NGS短读长测序片段与基因组比对的特征，CRAQ可以识别基因组内小规模的区域组装错误（CREs）和大范围的结构组装错误（CSEs）。不同类别的错误数量经过统计和标准化处理后被转化为两个组装质量评估指标——R-AQI和S-AQI，以反映不同层面的基因组组装质量。

此外，CRAQ能够将组装错误与基因组内的高杂合区域或单倍型差异区分开来，并在单碱基分辨率下指示低质量组装区域和潜在错误断点的位置。因此，CRAQ能够帮助研究人员识别基因组中存在的嵌合片段，并将这些片段准确地拆分，以利于结合光学图谱或构象捕获（Hi-C）技术进一步构建结构更加准确的参考基因组。

10月17日，相关研究成果发表在《自然-通讯》（Nature Communications

）上。研究工作得到国家重点研发计划、国家自然科学基金和中国科学院青年交叉团队项目等的支持。

[论文链接](#)



参考基因组组装质量评估的新方法

研究团队单位：植物研究所

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](https://www.iikx.com)转发