
多组学研究助力玉米种质资源创新利用

作者：writer 来源：爱科学

本文原地址：<https://www.iikx.com/news/progress/17384.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！

多组学研究助力玉米种质资源创新利用。

近日，中国农业大学农学院、国家玉米改良中心王向峰教授在《科学通报》（英文版）（Science Bulletin）上发表了方法学研究论文。

农作物种质资源精准鉴定与基因挖掘是从源头上实现种业创新、保障粮食安全的根本路径，是生物育种重点攻关的关键技术之一。全基因组关联分析（GWAS）是通过推断基因型与表型的关联显著性，实现挖掘种质资源的常规手段。但是，基于基因型-表型的传统GWAS分析存在诸多不足。随着转录组、代谢组、蛋白组、表观遗传组以及表型组等各种组学技术的飞速发展与检测成本的大幅降低，聚焦一套公共核心种质资源开展全方位多组学研究将是种质资源研究领域的重点发展方向。

大规模、多维度的组学数据的急速积累催生了高维生物学（HDB）研究领域。由于多组学数据具有规模大、维度高、噪音大、异质性强等特点，传统关联分析中常用的混合线性模型很难用于HDB数据的高效、精准解析。王向峰教授团队开发的MODAS软件运用多种先进数据分析技术解决上述问题，实现多组学数据在群体水平上的关联分析与因果推断。MODAS包含以下六大功能模块或分析步骤：

第一步，MODAS利用Jaccard index、DBSCAN与PCA算法对基因型数据进行降维；将全基因组范围内数百万个SNP的基因型数据，抽象成由6万个基因组区段代表群体的遗传变异，并生成伪基因型索引文件；该文件用于分子性状（即：基因表达、代谢物等等）的过滤，以及关键基因与代谢物的初步筛选。该步骤是大幅度提升多组学关联分析的关键步骤。

第二步，MODAS利用伪基因型文件首先将分子性状与基因组区间的关联分析；在获得显著关联的分子性状与区间后，提取区间内的SNP在进行第二部的分子性状与SNP基因型之间的关联分析，确定显著性；通过以上两步，实现分子性状的初步筛选。

第三步，由于代谢物数据、表达数据可能存在较大的冗余，MODAS对共同关联到相同基因组区段的分子性状进行降维，降低关联分析的冗余。

第四步，MODAS将筛选出来的具有潜在生物学意义的分子性状进行基因表达-全基因组关联分析（eGWAS）或代谢物-全基因组关联分析（mGWAS），确定显著相关的分子性状与QTL。

第五步，MODAS将所有曼哈顿图以及QTL内基因信息进行注释与整合，生成可以网页式浏览的可视化数据库。

第六步，MODAS应用孟德尔随机化（MR）算法 [2, 3, 4]，推断遗传变异、转录因子、目标基因、基因表达、代谢物含量、表型性状两两之间的因果关系。MR因果推断的结果可以更好的辅助生物学家建立可验证的分子通路假设，对挖掘获得的候选基因开展下游的实验验证。

中国农业大学的刘松誉、徐峰博士生为该论文的共同第一作者，也是MODAS（Multi-Omics Data Association Analysis）软件的主要开发人。

本项目受到合成生物学国家重点研发计划子课题抗逆回路在底盘作物中的智能重建与育种应用的资助。（来源：中国科学报张晴丹）

相关论文信息：<https://doi.org/10.1016/j.scib.2022.01.021>

版权声明：凡本网注明来源：中国科学报、科学网、科学新闻杂志的所有作品，网站转载，请在正文上方注明来源和作者，且不得对内容作实质性改动；微信公众号、头条号等新媒体平台，转载请联系授权。邮箱：shouquan@stimes.cn。

作者：王向峰等 来源：《科学通报》

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