



TBtools——大数据时代下的国产生物软件

陈程杰^{1,2*} 夏瑞^{1,2*}

1 华南农业大学亚热带农业生物资源保护与利用国家重点实验室 广州 510640

2 华南农业大学园艺学院 广州 510640

* 通信作者 E-mail: ccj@scau.edu.cn; rxia@scau.edu.cn

DOI: 10.15978/j.cnki.1673-5668.202206006

中国的现代生物学研究起步晚于西方发达国家 20 年。近年，我们在生物学领域取得了系列突破性学术成果，科技论文数量更是跃居全球第一。然而我们快速进步的关键基础是以发达国家为模板，这样发展起来的生物学研究不可避免依赖于国外开发的研究体系和软硬件设备。其中，分子生物学实验室常用软件中几乎没有国产软件。绝大多数生物软件来源海外，其中不乏商业软件，如 MEGA, DNAMAN, DNASTAR, CLC Workbench 和 Geneious。近期，国际竞争日趋激烈，全球化合作前景堪忧；西方发达国家不断阻挠我国科研和产业发展，主要途径不仅包括限制先进科研硬件的出口，更包括限制高端专业软件的授权使用。其中，典型案例即是数学建模领域的 MATLAB 软件以及芯片设计领域的 EDA 软件。这些限制对我国现有科研体系和产业结构造成巨大冲击。此番遭遇，归根结底是我们起步晚，且创新力不够，相关工作难以在短时间内取得突破。

2000 年以来，生物学进入大数据时代，高通量技术普及，海量生物数据产生。系列优异的生物信息学软件被开发出来，辅助科研工作人员进行生物大数据挖掘，并涌现了一批高质量学术发现。在此背景下，绝大多数生物信息学软件聚焦于大数据处理与加工。这些软件可以有效地将大数据整理为人工可查看的结果，但几乎没

有一款软件可以辅助一线科研人员快速获取和解读这些结果，加速科研假设的提出与验证。

基于这一需求，我们课题组立足于自身生物数据解读过程中积累的经验和需要，独立研发了一款适用于各类生信数据分析，解读以及可视化的软件——TBtools。目前，这一软件已开发七年，用户覆盖了国内绝大多数分子生物学实验室以及部分海外实验室。基于软件更新检测信息统计，每天有超过 3000 名科研人员使用 TBtools 开展学术研究工作。可以说，TBtools 一时间成为了分子生物学实验室最常用软件之一。

鉴于课题组主要研究对象为植物，TBtools 绝大多数功能开发的出发点和着力点均面向植物相关研究，但不少功能也适用于其他生物甚至非生物领域的数据分析（图 1）。软件文稿于 2020 年 8 月发表于中国科学院分子植物科学卓越创新中心与中国植物生理与植物分子生物学学会共同主办的 *Molecular Plant* 期刊。发表后，受到广泛认可与关注，目前被引用次数达 2487 次（基于 Web of Science 数据库，检索时间为 2022 年 9 月 9 日），成为迄今该杂志引用次数最高的论文，并入选 2020 年度“中国百篇最具国际影响力学术论文”。

相比于其他生物信息软件，TBtools 重点关注湿实验科研工作者日常的数据分析需求。其主要特点是：

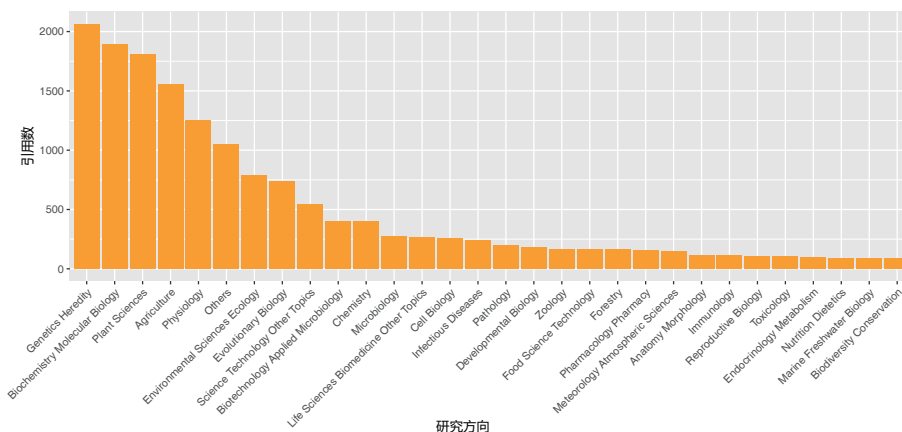


图1 TBtools 论文引文主要关注领域分布

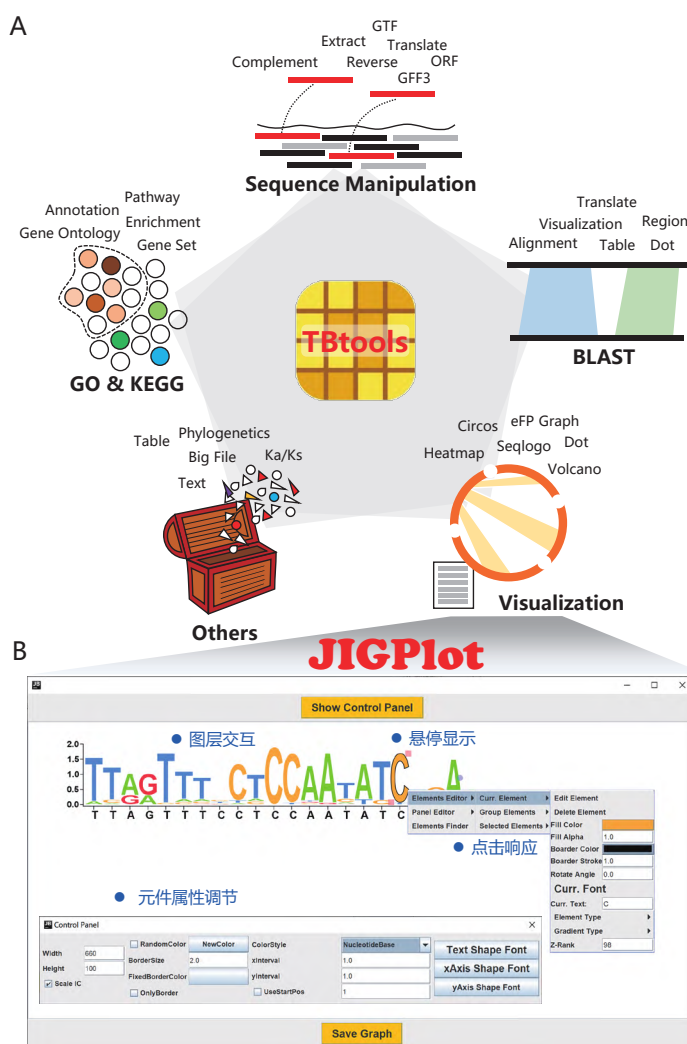


图2 TBtools 功能概况与 JIGplot 交互式绘图引擎

注：A. TBtools 主要功能划分为五大类：序列处理、BLAST 界面化、GO 和 KEGG 分析、数据可视化以及其他；B. JIGplot 绘图引擎输出交互式图稿，支持拖拽放置、点击响应、悬停显示以及元件属性调节等。



(1) 用户友好的操作界面和高兼容性(跨平台, 兼容不同操作系统); (2) 统一的功能使用逻辑, 即 IOS (Input-Output-Start, 输入-输出-启动); (3) 全面的多功能集成, 覆盖了绝大多数日常生信数据解读需要(图 2A); (4) 独立自主研发的绘图引擎(JIGplot: Java Interactive Graph Plotting Engine, 图 2B); (5) 强大的数据交互功能, 可以实现点击实时响应, 悬停菜单, 元件编辑等功能; (6) 庞大的用户和开发者互助社群, 目前已拥有 20 万用户, 超 1 万名用户在线实时交流。整体上, TBtools 软件并不是为用户开发某一个特定功能, 而是尝试为众多数据分析需求提供一个简单高效的“一站式”分析解决方案。

回首过往, TBtools 的开发和持续维护并非一帆风顺。首先 TBtools 软件的创新性一开始无法得到认可, 成果的价值也没有得到客观的评价, 但我们选择了坚持, 并坚信, 便捷性和全面性也是软件开发的关键方面。不少课题与工作的价值虽不能快速体现, 但大多数人认可的工作一定是有意义的。在海外软件已经覆盖了所有分子生物学实验室的背景下, TBtools 软件的意义体现更需要时间, 先追赶, 后超越, 实现弯道超车。其次是可用于 TBtools 开发的相关资源少。目前为止, 作为一款公益学术科研软件, TBtools 开发耗费了课题组大量时间和人力, 软件测试消耗了大量计算资源。这些投入均无直接研究经费支持。对于一个要取得持续发展的工作

来说, 稳定的经费支持和持续的人力投入极为重要。这或许是目前包括 TBtools 在内的国产生物软件研发的困境之一。

可以说, 相比绝大多数国产生物软件而言, TBtools 是幸运的。目前软件在国内已经被普遍接受与认可, 而且也得到越来越多的海外科研工作者的关注和使用。展望未来, 我们认为以下几个方向的工作最值得开展: (1) 实现“插件”模式, 对于非普适性但也有一定需求的功能, 通过插件的方式实现, 让用户按需安装; (2) 从“开发者即用户”走向“用户即开发者”, 实现实用简单的“插件”开发功能, 赋予用户开发和分享自我研发 TBtools 插件的能力, 群策群力, 实现软件平台化; (3) 紧跟国家“种业振兴”的战略需求, 继续发挥开发团队的植物学背景优势, 开发系列与种业科研相关的功能, 服务于产业需要; (4) 不忘初心, 持续维护和开发, 坚持做好 TBtools 开发这一件小事, 为其成为一款让所有生物学家都认可的生物大数据分析工具而努力。

综上, 路漫漫其修远兮, 吾将上下而求索。TBtools 开发上, 我们已经从零走到一, 接下来是从一到二, 到三, 再到一百。我们有充分的信心, 做成一款由我国科研人员自主开发和维护的, 拥有完全自主知识产权的, 具有良好国内外影响力的生物软件。我们也期待未来生物软件开发领域能出现更多更优秀的国产作品。

作者简介

夏瑞, 华南农业大学教授, 博士生导师, 2016 年入选国家级人才项目回国工作, 获得广东省特支计划科技创新领军人才项目支持。目前主要研究无患子科植物花性别分化机制以及南方主要水果花果发育调控机理。

陈程杰, 华南农业大学副教授, 目前以生物信息为研究切入点, 关注香蕉枯萎病抗性机理解析与应用。