

# 青年人才成长故事

——将“大”数据应用于“小”麦：梦想照进现实

据世界粮农组织统计，小麦（wheat, *Triticum* L.）供养着 30% 以上的世界人口；同时，小麦也是我国两大主粮作物之一。然而，正如李振声院士所述，“普通小麦是异源六倍体，基因组庞大且复杂（是水稻基因组的 40 倍、人类基因组的 5.5 倍），其功能基因组学研究远远落后于水稻和玉米，复杂的遗传背景一直以来也是制约重要农艺性状的基因克隆和小麦改良的瓶颈”。从 2014 年 7 月《*Science*》杂志推出小麦专辑并报道了小麦全基因组草图，近几年来，不同倍性小麦的多个版本全基因组测序结果陆续释放，到 2018 年 8 月，《*Science*》杂志发表六倍体小麦模式种“中国春”的参考基因组，预示着小麦研究进入了后基因组时代。随着六倍体小麦参考基因组的公布，高分辨率的转录组、表观修饰组、SNP 育种芯片等高通量小麦组学大数据呈指数级增长趋势；但是这些大数据，信息量大，分析需要专业的技术，对于以小麦为研究对象的普通科研工作者和育种工作者来说，很难直接使用。更重要的是，后基因组时代的小麦功能基因研究策略和方法将会发生革命性的变化，这些研究新思路 and “研究路线图”亟需系统的提炼。鉴于此，被托举人王萌博士做了大量的工作和尝试：

## 兵马未动，粮草先行

大数据终归只是一堆数据，只有普通科研工作者可以方便的使用它，这些数据才是活的、有意义的。针对上述问题，王萌博士在中国科协“青年人才托举工程”等项目的支持下，作为创始人之一参与搭建了我国首个麦族多组学大数据可视化网站 *Triticeae Multi-omics*

Center (<http://202.194.139.32>)。该网站页面友好，操作简单，可以快速、直接地查看和提取小麦基因组中任一基因的序列信息、基于转录组的表达信息、表观修饰信息、突变体信息、SNP 育种标记等详细地数据。目前该网站已有超过十万次点击量，用户遍布全球，得到了国内外专家学者的广泛认可和支持。“国际小麦全基因组测序联盟”发起人之一、英国 John Innes Centre 的 Cristobal Uauy 教授在社交媒体对该网站进行了专门的推荐；国家“万人计划”领军人才、河南农业大学陈锋教授在《*Plant Biotechnology Journal*》发表的综述中专门点评该网站，认为该网站为小麦基础研究和育种工作提供了很多有用的信息和工具。为了更好地发挥该网站的作用，王萌博士主持举办了“Wheat Omics 第一届青年学术研讨会”（江苏南京，2019），并先后赴中国农业大学、西北农林科技大学、山东农业大学、山东大学等单位，走进校园，向青年教师和研究生讲解该网站的使用攻略。



图 1. 王萌博士做培训性报告

## 授之以鱼，不如授之以渔

尽管不同倍性、不同品系的小麦全基因组信息在近三年来得到陆续公布，如前所述麦族多组学大数据可视化网站也已经建立，但是“授之以鱼不如授之以渔”-如何将这些大数据应用于小麦的基础研究和育种工作，相关的策略亟需进行系统地梳理、比较和总结。鉴于此，王萌博士与合作者从基因克隆、功能验证、进化分析、育种芯片开发等全方位的视角，分析了小麦全基因组信息公布之后如何对研究“提速”和“提质”；同时，针对正向遗传学和反向遗传学这两种经典的方法，比较了全基因组公布之前和之后的策略差异，并基于小麦遗传复杂性的特点，着重提出正、反向遗传学相结合的思路与方案，详细阐述了后基因组时代如何快速、精准地挖掘小麦重要农艺性状功能基因；根据六倍体小麦不同于水稻、玉米等二倍体作物，拥有大量二倍体、四倍体祖先种和近缘物种等特色遗传材料的特点，梳理了小麦特色遗传材料用于小麦功能基因挖掘和验证的思路，强调了传统特色遗传材料与新技术创制材料相结合来实现小麦重要农艺功能基因的快速鉴定。上述成果先后发表《*Trends in Plant Science*》、《*Trends in Biotechnology*》等国际顶级综述期刊（王萌博士为第一作者）。成果得到了“国际小麦全基因组测序联盟”的官方报道和推荐；包括西北农林科技大学康振生院士团队等在研究论文中引用了我们的观点。

## 只说不练是个假把式

前述“研究路线图”和大数据资源得到国内外其他小麦研究团队的认可并在研究中成功应用，与此同时，申请人也围绕自己的研究方向-小麦在盐渍化土壤中适应性和氮高效性的分子机制，鉴定到一个在小麦基因组中存在特异串联重复扩增的盐响应基因 *TaCYP81D5*,

过表达该基因，不仅可以提高苗期小麦的耐盐性，更可以提高小麦在盐渍土中的产量，说明该基因对于小麦抗盐育种具有潜在应用价值；同时，借助小麦参考基因组信息，全套克隆了小麦中高亲和力硝酸根转运蛋白 NRT2 及其伴侣蛋白 NAR 两个基因家族的成员，并发现由 ABA-GE 快速转化产生的 ABA 可以作为硝态氮的信号“传导器”甚至“放大器”，通过激活根部特异 *TaNRT2* 的表达，提高小麦的氮素吸收效率，为提高小麦氮营养吸收能力提供了改良靶点和提效策略。上述成果分别发表在《*Plant Biotechnology Journal*》、《*Plant Physiology*》等植物学国际权威期刊，得到了国内外专家的高度好评。例如，植物营养领域国际知名专家、美国佛蒙特大学 Jeanne M. Harris 教授评价到“看到小麦中建立 ABA 与硝态氮的直接关系让人兴奋”。英国剑桥大学 Stephanie Swarbreck 博士也对上述研究进行了专门推荐，并邀请王萌博士共同为植物学权威综述期刊《*Trends in Plant Science*》撰写了题为“A Roadmap for Lowering Crop Nitrogen Requirement”的文章。

纸上得来终觉浅，小麦作为粮食作物，研究不能只停留在实验室层面，要回到大田中去。在康振生院士的牵头下，王萌博士参与到小麦田间表型 GWAS（全基因组关联分析）研究平台的搭建，该平台在全国布置了 8 个试验点，通过多年多点的大田实验，结合 SNP 芯片基因型测定数据，对小麦重要农艺性状进行 GWAS 分析。第一部分关于小麦耐盐性相关的成果已经在国产“卓越行动”重点资助期刊《作物学报-英文版》发表（并列第一作者）。

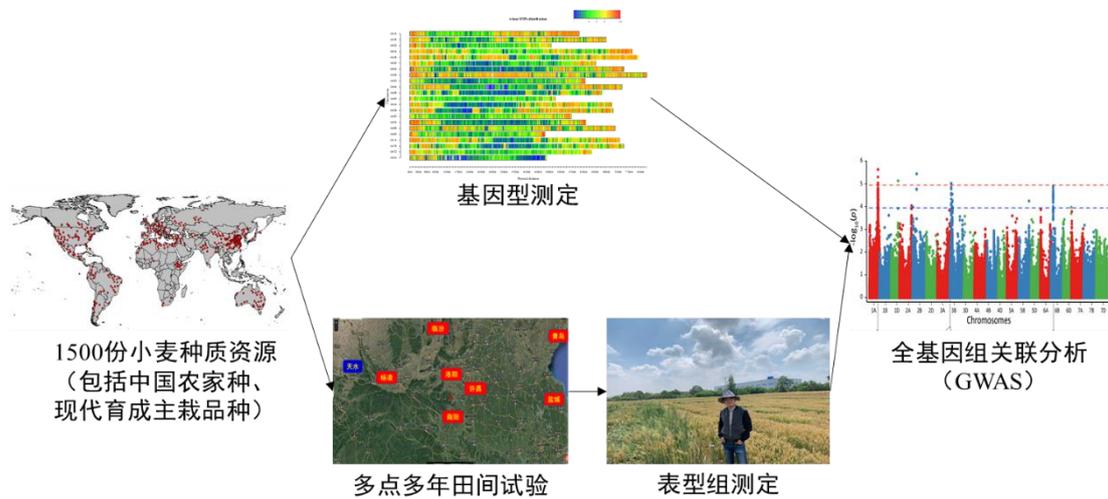


图 2. 王萌博士参与的小麦田间表型 GWAS 平台

小麦作为一个基因组庞大的作物，研究人员和育种家一直期待它的基因组可以被破解；随着参考基因组的公布，这个梦想已经实现，但这动辄几十 G 的大数据，对普通的科研工作者（尤其是年龄较大、对生物信息学不熟悉的科研工作者）十分不友好。从研究生开始，王萌博士已经从事小麦研究十个年头，深知组学大数据对于小麦研究和育种的重要性。作为青年科研工作者，他一直关注小麦组学大数据的发展趋势，并敢于学习掌握大数据这一新鲜事物。在中国科协“青年人才托举工程”等项目的支持下，王萌博士参与搭建小麦组学大数据可视化网站、撰写后基因组大数据时代的小麦“研究路线图”，并主动尝试把组学大数据应用于他的日常研究中；更重要的是，他主动到多所高校、研究所，与青年科研工作者和学生分享他在小麦大数据方面的使用体会和经验，用具体研究案例鼓舞科研工作者-小麦大数据并不是遥不可及，梦想也可以照进现实。