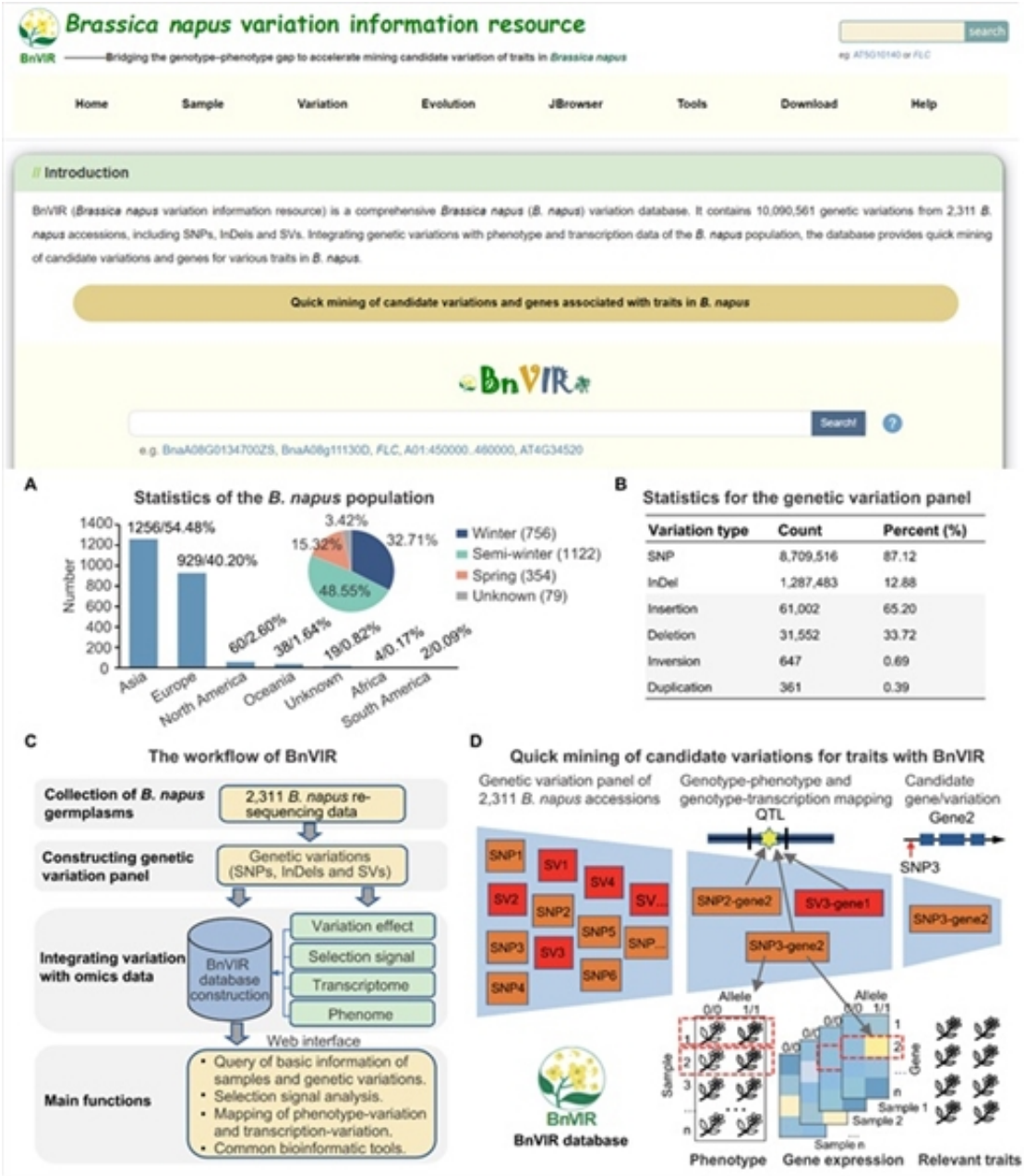
多组学数据库“齐聚荟”助力油菜“全能”竞技

作者：writer 来源：爱科学

本文原地址：https://www.iikx.com/news/progress/17364.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！

多组学数据库“齐聚荟”助力油菜“全能”竞技。



BnVIR数据库主页、构建流程和概貌。 课题组供图

油菜产业集观赏、蜜源、生态、经济作物等多种功能于一身，是新兴的全能选手。围绕油菜，我国科学家近年来实现了从理论、技术、产品到转化的链式创新，推动了我国油菜基础研究与应用步入世界第一方阵。

然而，如何将基础研究中长期积累的资源和多组学数据优势转化成育种优势，仍是当前乃至未来很长一段时期油菜遗传育种的核心研究内容之一。

华中农业大学信息学院教授杨庆勇团队发布油菜群体变异数据库BnVIR，建立了甘蓝型油菜基因型与基因表达、表型的关联，为快速高效地挖掘油菜候选变异和基因提供了平台。这是继该团队构建首个油菜泛基因组及其数据库BnPIR、油菜转录组数据库BnTIR、植物遗传变异参考面板数据库Plant-ImputeDB后又一重要工作，对油菜种质资源精准鉴定、评价、利用以及分子育种具有重要推动作用。相关论文2月8日在线发表于《分子植物》。

寻找遗传变异从图开始

油菜是世界四大油料作物之一，我国年均油菜生产和消费居世界首位，常年种植面积1亿亩左右。近年来，油菜的油、菜、花、饲、肥、蜜等多种功能得以全面开拓，是三产融合发展的多面手。油菜华丽变身的背后，少不了科技力量的支持。

然而，当前我国油菜大多以常规育种与分子标记辅助选择育种技术为主，与国外大规模应用的以基因组选择为核心的分子育种技术相比，在品种育成速度、成功率及效率等方面仍存在较大差距。论文通讯作者杨庆勇指出。

对于遗传变异非常丰富的甘蓝型油菜，研究人员如何解析其中的基因密码，为挖掘重要的功能基因提供便利？

遗传变异是物种表型多样性形成的物质基础，包括单核苷酸多态性（SNP）、小片段的插入/缺失（InDel）以及大规模结构变异（SV）。目前，遗传变异已经被广泛应用于人类疾病、动植物重要农艺或经济性状相关遗传位点的鉴定、重要功能基因的克隆、分子标记辅助选择育种等工作。

杨庆勇向《中国科学报》进一步解释说，其中，对于作物遗传育种，准确了解遗传变异对基因的调控效应以及与表型分化的关联，对种质资源的精准鉴定和优异等位基因的挖掘具有重要作用。

传统的遗传变异检测手段主要通过基因组重测序，然后将重测序的读段比对到参考基因组，可以称之为按图索异。杨庆勇介绍，这里的图是指参考基因组。其团队很早就提出了构建一份系统全面的图，以快捷方便地检索、使用相关基因组资源的研究思路和策略。

团队首先提出整合多个代表性种质资源的基因组信息，构建了首个油菜泛基因组和比较基因组生物信息平台BnPIR。相关论文2020年10月发表于《植物生物技术杂志》。用杨庆勇的话说，泛基因组数据库仅仅是个开端。

挖掘与表型相关的候选变异

在这项新研究中，杨庆勇团队首先构建了一套由2311个甘蓝型油菜核心种质鉴定的10090561个遗传变异的数据集，其中包括SNP、InDel和不同种类的SV。这是目前已发表的最为系统、最为完整的甘蓝型油菜遗传变异集合。

之后，该研究以其中1703个材料的18个性状的表型数据和309个材料的RNA-seq数据为基础，构建出甘蓝型油菜变异信息资源数据库BnVIR。

论文共同第一作者、华中农业大学已毕业博士生杨植全介绍，该数据库提供了群体水平的遗传变异基因型与表型和基因表达量的关联，可以辅助快速挖掘与油菜性状相关的候选变异和基因。

也就是说，BnVIR架起了油菜基因型与表型的桥梁。例如，通过挖掘与控制开花3~5个位点/单体型（即具有特定功能的基因型），在育种中加以充分利用，可获得各种不同开花时间的油菜品种。再将不同开花时间的油菜按需播种，可使油菜观花旅游期从现有的不足一个月延长至2~3个月。

杨庆勇表示，传统油菜育种侧重油用，对其他多功能利用（菜用、观花、绿肥等）所需的性状关注相对较少。借助BnVIR平台，研究人员可以加快挖掘油菜多功能利用中所需的位点、标记、基因、种质资源等。这也是加快油菜全产业链建设、推动油菜花农旅融合发展，乃至油菜生产区域实现乡村振兴的一个路径。

论文共同第一作者、华中农业大学博士生梁聪园说，此外，该数据库为甘蓝型油菜种质资源查询、群体遗传学和进化分析以及关联分析等提供了多种工具。

最后，该研究通过挖掘一些已报道的和新鉴定的候选基因和功能变异的案例分析，以及独立表型数据集的检验分析等系列工作，系统全面地展示了该数据平台在挖掘与表型相关的候选变异和基因的潜力、可靠性与稳健性。

与现有的甘蓝型油菜数据库不同，BnVIR提供了更加全面、更加丰富的变异及其与多个组学数据的关联，以及更加友好和丰富的用户界面，可以显著提高研究者挖掘候选变异和基因的效率，并为甘蓝型油菜的分子育种开发提供有价值的标记。杨庆勇说。

搭建油菜参考基因组生态

随着后基因组时代的到来，作物科学领域已积累了海量数据。以十字花科作物育种为例，我国科学家在该领域发表外文论文总量排名世界第一，并且产生了大量优良的遗传材料以及基因组、转录组和代谢组等多组学数据。这些数据如何才能真正利用和共享？

利用生物信息技术将大规模的多组学数据进行系统地整合和分析，并构建更友好的图形化、可视化数据库，将极大促进重要农艺性状的功能基因的发掘及机制解析，推动优异种质资源的开发和利用，加快品种选育。这对发展高效、精准的新一代生物育种，打好种业翻身仗具有重要意义。杨庆勇表示。

其团队近年来聚焦油菜基因组学与分子育种研究方向，先后与国内多个研究团队合作，通过系统整合和分析甘蓝型油菜的多组学数据，构建了首个油菜泛基因组及其数据库BnPIR、油菜转录组

数据库BnTIR、油菜遗传变异数据库BnVIR和植物遗传变异参考面板数据库Plant-ImputeDB，搭建了我国自主可控的油菜参考基因组生态。

具体来说，泛基因组数据库构建相对系统全面的基因集合，转录组数据库了解这些基因的表达模式及初步的调控关系，遗传变异数据库则是更加系统全面的变异集合，以及这些变异与表型的潜在关联。

这些研究不仅为甘蓝型油菜今后的遗传育种提供了重要平台，同时也为其他作物进行整合和利用多组学数据推动育种产业发展提供重要参考。

杨庆勇指出，一个值得注意的问题是，当前我国生物育种工作所用的数据标准、参考基因组等大多来自国外资源。搭建我国自主可控的油菜参考基因组生态，也是为生物育种国家战略作长期打算。在做好参考基因组的同时，对配套的基因注释（即泛基因组数据库）、转录组、遗传变异组、表观组和代谢组等数据库进行完善，是我们团队努力的一个方向。（来源：中国科学报王方）

相关论文信息：

https://doi.org/10.1016/j.molp.2022.02.002

https://doi.org/10.1038/s41477-019-0577-7

https://doi.org/10.1111/pbi.13665

https://doi.org/10.1111/pbi.13491

https://doi.org/10.1093/nar/gkaa953

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：杨庆勇等 来源：《分子植物》

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转发