
泛基因组 (Pan-genome) 是一个物种内所有基因组信息的总和，它比单一参考基因组涵盖了更多的遗传多样性。

近年来，科学家已经获得了多个作物的泛基因组。如何利用这些更为全面的基因组信息培育下一代良种，成为科学家们关心的课题。

北京时间2022年6月8日，《自然》在线发表了中国农业科学院深圳农业基因组研究所（以下称基因组所）黄三文团队有关泛基因组的两项研究成果。

一篇论文中，他们首次获得了番茄的图泛基因组，并借此找回了番茄育种中丢失的遗传力，为解析生物复杂性状的遗传机制提供了新思路。论文评审专家认为，这项工作是对图泛基因组概念最全面的分析。图泛基因组将可能成为基因组分析和作物基因组育种的标准，在这个意义上，这篇论文是奠基性的 (foundational)。

另一篇论文中，他们首次解析了二倍体马铃薯的泛基因组，研究了茄科茄属 (Genus Solanum) 的物种进化，破解了马铃薯如何结薯的分子机制，并为杂交马铃薯育种改良提供了丰富的遗传变异信息。《自然》同期发表的观点性文章指出，组装数量如此之多的异交和高杂合马铃薯高质量基因组，是一项非凡的成就 (remarkable feat)。泛基因组全面鉴定了参与重要生物学过程的基因，能够赋能马铃薯育种。

意味着更准确更详尽

在一个物种内，有些基因是某些个体所特有的。比如任意两个番茄所含有的基因是不完全相同的，基因不同可能导致口感不同。论文通讯作者黄三文在接受《中国科学报》采访时解释道，番茄有很多品种，如大果番茄、樱桃番茄和醋栗番茄等，如果只是用大果番茄的参考基因组进行研究，那其中不会包含樱桃番茄和醋栗番茄特有的基因组信息。

黄三文说，把一个物种所有的基因都找到，并按照顺序排列，就形成了泛基因组。而图泛基因组 (graph pangenome) 是借助数学和计算机中被称为图 (graph) 的数据结构来展示一个物种所有的基因排列和结构。

论文第一作者、基因组所副研究员周姚告诉《中国科学报》，由于泛基因组比单一参考基因组更加准确和详尽地代表了整个物种的遗传多样性，而详尽和准确的变异对下游的遗传学分析提供了完美的基础，所以借助泛基因组做研究不容易漏掉重要的基因和信息。

此前有很多作物都获得了泛基因组图谱，如水稻、玉米、小麦、大麦、棉花、番茄、油菜等。这次我们首次获得了番茄的图泛基因组，并第一次完成了二倍体马铃薯的泛基因组。黄三文说。

在关于番茄图泛基因组的研究中应用了图这个概念。图是数学和计算机科学中一种常见的数据结构。已有研究发现，利用该数据结构整合泛基因组，可以减少单一基因组带来的参考基因组偏差问题。周姚说，为了构建一个准确的图泛基因组，他们首先利用高准确率三代测序技术重新组装了骨架基因组，其组装质量在完整性、连续性和准确性等指标上均优于之前的版本。

随后，他们选取了31份具有代表性的番茄材料进行了组装，并鉴定出相关的遗传变异；在整合已公布的结构变异和短片段测序信息后，最终构建了来自838个番茄基因组的图泛基因组。

法国农业科学研究院研究员Mathilde Causse认为，图泛基因组资源对番茄遗传和基因组研究非常重要，促进了后续的基因定位。这篇文章将成为番茄研究领域里的基石。

二倍体马铃薯泛基因组论文的第一作者、基因组所博士生唐蝶告诉《中国科学报》，为了获得高质量的二倍体马铃薯泛基因组，他们挑选了地方栽培种、野生种、近缘野生种等44份具有代表性的二倍体马铃薯种质进行了重测序和基因注释，最终完成了第一个二倍体马铃薯泛基因组。

审稿人评价说，马铃薯泛基因组的构建和分析结果令人印象深刻。这项研究展示了基于广泛选材的马铃薯组和类马铃薯组泛基因组学的力量，能为其他作物泛基因组研究提供参考方法。

好吃的番茄：找回丢失的遗传力

黄三文告诉记者，作物的性状由遗传因素和环境因素的共同作用而决定。科学家用遗传力这个概念表示一种作物的性状受遗传调控的比例。遗传力越高，说明性状的决定过程中遗传因素占比越大，环境因素占比越小。

在未来的作物育种中，基因组选择技术将被广泛应用。当番茄还是幼苗的时候，就对它的基因组进行测序，以此来预测它是否抗病、产量如何、好吃不好吃等。

然而，现在的预测还不是很准。这是因为调控上述复杂性状的基因有很多，有些基因的作用明显，其遗传力容易被检测到；而另一些基因的作用比较微弱，这些基因难于检测到的遗传力被称为‘丢失的遗传力’。黄三文说，遗传力丢失是一个经典的数量遗传学问题，即通过遗传标记估计的遗传力以及通过全基因组关联分析（GWAS）发现的所有相关基因所贡献的遗传力总和均低于实际的遗传力。

找回这些丢失的遗传力，将有助于理解复杂性状的遗传机制。周姚说，遗传力是研究基因型与表型相关性的基础，如何更准确和系统地找到更多的决定表型的遗传变异是作物育种中的基础问题。无论是分子标记辅助育种还是分子设计精准育种，都需要大量的已知功能的遗传变异来指导品种选择或改良。



受访者供图

目前认为遗传力丢失的主要原因有以下几个：遗传标记与真实影响表型的基因之间存在着不完全连锁，导致模型估计产生偏差；稀有等位基因的贡献不易被察觉；某个基因的不同突变表现出相同的表型使得遗传分析遗漏；基因与基因的互作和基因与环境的互作增加了分析的复杂性。传统短片段测序可对长度较小的变异进行鉴定，但对复杂结构变异的检测能力较差，无法全面评估复杂结构变异对遗传力的具体影响。论文共同第一作者、基因组所张智洋说。

此前，在人类中遗传力丢失问题研究最为丰富。但受限于技术，这些研究主要关注单核苷酸变异与表型之间的联系，而忽视了更多的隐藏的大结构变异对表型的影响。

由于番茄的遗传资源丰富，其驯化历史、表达调控以及风味代谢方面都已有了相关的研究基础，为进一步研究遗传力丢失奠定了基础。

论文共同第一作者、基因组所鲍志贵说，通过构建番茄图泛基因组，他们的研究准确鉴定了番茄基因组中的结构变异，并发现大的结构变异是遗传力丢失的关键原因之一，为遗传力丢失问题提供了新的解决思路。

进一步研究发现，与利用单一参考基因组相比，基于图泛基因组的遗传变异可将估计的遗传力提高24%，展现了图泛基因组在找回丢失的遗传力上的重要作用。

以影响番茄产量和糖度的重要代谢物可溶性固形物为例，采用上述分析方法，该团队共鉴定出2个潜在的与可溶性固形物含量高度相关的结构变异，可以用于未来的分子标记辅助选择。通过对影响番茄风味的33种代谢物进行分析发现，利用所有的结构变异作为分子标记进行基因组选择的效果最佳。

因此，他们进一步精心构建了一个包含近2.1万个结构变异的数据集。如果利用该数据集设计育种芯片，评估基因组选择的准确率可能超过利用全部的单核苷酸多态性。

西湖大学教授杨剑长期从事数量遗传学研究。他在接受《中国科学报》采访时说，遗传力丢失的问题不只局限在番茄里，它是各个物种，包括人类中广泛存在的数量遗传学问题。这篇论文很好地利用番茄这个作物回答了这个广义的普遍性问题。

该团队通过图泛基因组鉴定了大量结构变异后，提出了一个重要问题：结构变异是否能解决遗传力丢失的问题？结果是肯定的。结构变异的神秘面纱被揭开了，它们才是遗传力这辆‘汽车’的‘司机’，发挥着主导作用，而简单变异可以说是‘乘客’。杨剑说，这篇论文将让大家更加重视对结构变异的研究。

培育优薯：破解结薯密码

作为世界第三大主粮作物，传统马铃薯栽培以四倍体为主，依靠薯块无性繁殖。然而，四倍体遗传分析复杂，育种不可积累；薯块运输成本高，易感染病虫害。

为彻底打破产业发展中的障碍，2017年，在农业农村部、深圳市和中国农科院的支持下，黄三文联合国内外优势单位发起了优薯计划，旨在用基因组学和合成生物学指导马铃薯产业的绿色革命，即用二倍体替代四倍体，用杂交种子替代薯块，对马铃薯育种和繁殖方式进行颠覆性创新。

马铃薯种质资源丰富，自然界中70%的马铃薯是二倍体，其中大部分是野生材料，充分利用这些资源中的优异性状，有利于加快马铃薯的遗传改良。此外，马铃薯无性繁殖方式对马铃薯基因组的影响以及薯块形成的遗传演化机制还没有被充分解析。黄三文说，目前已发表马铃薯的基因组序列只捕获马铃薯有限的生物多样性，不足以全面了解马铃薯基因组以用于育种指导。



约70%的马铃薯是二倍体。受访者供图

唐蝶介绍，在构建二倍体马铃薯泛基因组的同时，他们还挑选了马铃薯姊妹类群——类马铃薯组（Section *Etuberosum*）的两个种进行基因组的组装和注释。同番茄一样，*Etuberosum*也是马铃薯的近缘物种，其植株外型和马铃薯非常相似，也会形成地下分枝，不同的是*Etuberosum*不会产生薯块。以往的分类学研究关于马铃薯、番茄和*Etuberosum*的系统发生关系一直存在争议。

高质量的基因组为解析马铃薯及近缘物种的系统发生关系提供了契机。唐蝶说，他们发现马铃薯与近缘物种番茄、*Etuberosum*之间，以及马铃薯类群内部，都存在广泛的种间杂交和不完全谱系分选现象，说明马铃薯类群经历了复杂的演化历史。

利用高质量的泛基因组，他们发现相比于番茄和*Etuberosum*，马铃薯的抗病基因拷贝数明显扩张。我们推测，这是由于马铃薯依赖于生长在土壤中的薯块进行无性繁殖，其相比于种子更容易受到病原菌的侵染。马铃薯的无性繁殖可能促使了抗病基因数量的扩张以应对病原菌对薯块的侵染。论文共同第一作者、基因组所博士生李宏博说。

马铃薯类群、番茄类群和*Etuberosum*类群是进化距离很近的近缘物种，但只有马铃薯演化出了薯块这一重要的生物学性状。论文共同第一作者、基因组所博士后贾玉鑫说，*Etuberosum*和马铃薯都会产生地下分枝，但*Etuberosum*的地下分枝向上生长发育成新的植株；而马铃薯的匍匐茎向下生长，并在顶端膨大形成薯块；番茄不含有地下分枝，也不形成薯块。因此我们推测*Etuberosum*是薯块形成的过渡态。

通过对上述三者的多组学比较分析，该团队鉴定到一个可能在薯块发育过程中发挥关键作用的TCP转录因子。

进一步地，他们在二倍体马铃薯中创制了上述转录因子的基因纯合缺失突变体。表型观察发现，相比于野生型，突变体匍匐茎顶端无法正常膨大形成薯块，转而发育成了侧枝。这证明该基因在薯块发育的起始时期发挥关键作用。论文共同第一作者、中国农科院蔬菜花卉所助理研究员张金喆说。该基因被命名为薯块身份基因。



马铃薯和不结薯种都存在地下分枝，前者发育成匍匐茎并在顶端膨大形成薯块，后者向上生长发育成新的植株。受访者供图

进一步发现栽培马铃薯内部共线性缺失现象，这说明栽培马铃薯材料中广泛的遗传多样性。在对马铃薯进行杂交育种时，必须谨慎考虑这些共线性缺失片段，以及可能导致共线性缺失的大结构变异带来的连锁累赘等影响。他们发现，马铃薯基因组中存在很多大的结构变异，而马铃薯的无性繁殖方式很难将这些结构变异清除出去的。

其中，马铃薯3号染色体的倒位事件与薯块中控制类胡萝卜素积累的基因紧密连锁，在自交后代中该倒位区域重组率显著下降。这意味着在育种中选择黄肉薯块这个重要的营养性状，就选择了该倒位区间的所有基因，这可能带来严重的连锁累赘现象。

水稻、玉米等作物是二倍体，用种子来繁殖，育种改良相对方便。我们平时吃到的马铃薯都是四倍体，块茎繁殖，遗传改良一直很慢。上海师范大学生命科学学院教授、水稻基因组专家黄学辉告诉《中国科学报》，2021年《细胞》发表的黄三文团队培育的第一代高纯合的二倍体马铃薯自交系和杂交马铃薯品系优薯1号改变了这种情况。

作为优薯计划的后续，充分挖掘二倍体马铃薯中的等位变异将是进一步育种改良的关键。该团队在高质量马铃薯泛基因组基础上，获得了一些重要的基因功能线索，尤其是成功鉴定了马铃薯结薯的核心调控基因，发现其在薯块发育过程中发挥关键作用。

优良的变异要怎样聚合到新品种中？如何优中选优？这份成果获得的大量遗传多样性信息将对未来二倍体马铃薯改良产生重要作用。黄学辉说。

瑞典斯德哥尔摩大学生态、环境与植物科学系Juanita Guti é rrez-Valencia和Tanja Slotte在《自然》同期发表的观点性文章中说，该研究大大扩展了马铃薯的基因组资源，利用泛基因组鉴定的遗传

变异无疑会推进其基础和应用研究。

了解与薯块形成相关的基因可以帮助育种家选育高产马铃薯品种，而对抗病基因的全面分析使得从野生种中定向引入抗性基因成为可能，从而提高马铃薯的抗病性。文中写道，该研究提供的组学信息资源将助力基因组学辅助育种。

该研究将为马铃薯研究供丰富的基因组大数据支持，加深对马铃薯重要生物学性状的理解，有力地推动杂交马铃薯育种，并加速马铃薯作为重要主粮作物的育种进程。(来源：中国科学报李晨)

相关论文信息：

<https://doi.org/10.1038/s41586-022-04822-x>

<https://doi.org/10.1038/s41586-022-04808-9>

<https://doi.org/10.1038/d41586-022-01419-2>

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：黄三文等 来源：《自然》

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