

# 科学家揭示棉花基因组进化历程

作者：writer 来源：爱科学

本文原地址：<https://www.iikx.com/news/progress/9255.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！

科学家揭示棉花基因组进化历程。

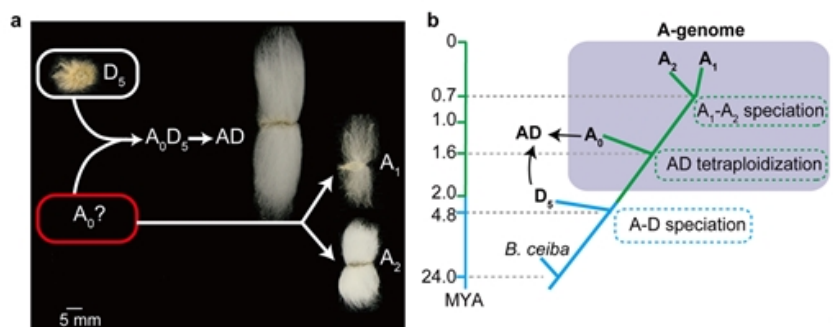


图1 棉花A基因组起源与进化模型 (a) 和重要进化事件 (b)。

图片来源：武汉高等研究院官网 我国是全球最大的棉花生产国，年均棉花产量占全球产量的24.4%，单产高于全球平均水平的一倍。与此同时，我国也是全球最大的棉花进口国之一，且近年来进口量愈发增大。随着人们生活水平和品质的提高，对高品质棉纺织品的需求量也越来越大，但与国外相比，我国棉花的品质还不够好。因此，提高棉花品质，是我国棉花研究中非常重要的课题。中国科学院院士、武汉大学高等研究院院长朱玉贤告诉《中国科学报》。近日，朱玉贤团队解析了世界上首个高精度的草棉参考基因组，解决了困扰已久的棉花A基因组进化起源问题，为棉花遗传改良提供了宝贵的基因组资源，将加快培育高产优质的棉花新品种的进程。相关研究成果发表于《自然—遗传学》。棉花起源之争 棉花是由一根根细小的纤维组成，其中，棉纤维长度是判断棉花品质的重要指标之一，纤维越长，品质越好。朱玉贤指出，我国生产的棉花纤维平均长度在3厘米以下，而北美产的棉花平均长度则在3厘米以上。短短数毫米之差，会带来很大影响。因为，在纺织品中必须有1/3以上的棉花纤维长度大于30毫米，否则纤维之间的接头太多，生产出来的棉布品质会很差。

破解棉纤维生长之谜，还要从棉花的起源与进化历史谈起。棉花共有四个栽培种，其中经济价值较高的是两个四倍体栽培种——陆地棉和海岛棉，它们是由A基因组和D基因组共同组成的异源四倍体AD基因组，AD基因组是由D基因组和A基因组通过自然杂交和染色体加倍而形成。已有研究表明，约95%以上的棉纤维来源于异源四倍体陆地棉（AD1基因组）。从分布来看，长期从事棉花生物技术及育种应用研究、华中农业大学教授张献龙告诉《中国科学报》，A基因组二倍体棉花主要分布在亚洲和非洲、D基因组的棉花主要分布在美洲，而野生异源四倍体棉花是在美洲发现。但A和D两个基因组的祖先种如何结合形成四倍体这是棉花研究中一直未阐明的问题。过去，其一直有着单系起源和多系起源之争，此外，A基因组只有1.3-1.5厘米的纤维长度，D基因组仅有像蒲公英一样0.3厘米的绒毛。这很有趣，两个‘祖先’中都没有棉纤维，组成四倍体以后怎么就产生了3厘米以上的棉纤维呢？控制棉纤维生长发育的基因是如何存在的？朱玉贤说。草棉（A1基因组）和亚洲棉（A2基因组）为现仅存的两个二倍体A基因组，历史上为重要

---

的栽培棉。朱玉贤希望从二倍体祖先中找到解锁陆地棉（AD1基因组）棉纤维基因的钥匙。过去大量研究已形成共识，即在陆地棉（AD1基因组）中，D亚基因组供体是雷蒙德氏棉（D5基因组），但A亚基因组的供体是谁，则是未解之谜，长期处在争论之中。有学者认为是来源于A2基因组亚洲棉，也有学者认为来自A1基因组草棉。

供体即可把自身优良性状提供给另一个生物体的个体。A1、A2在异源四倍体棉花中有哪些差别、它们与棉纤维的生长发育有怎样的关系、决定棉纤维的基因群究竟是哪个？这些是我们想研究清楚的。朱玉贤说。找到祖先之一早在2014年，朱玉贤团队便解析了全长1700兆碱基对的亚洲棉基因组；2019年，他们进一步通过系统地、全方位地转录分析阐释亚洲棉基因组，获得了迄今为止科学家对亚洲棉基因组的复杂转录全貌所进行的最准确的注释，为进一步的功能基因组研究提供了宝贵的资源。也是在2019年，朱玉贤开始组织团队解析亚洲草棉基因组。与6年前相比，高通量基因组测序技术、组装技术和计算机计算速度等得到了快速发展，我们才能获得更高质量的棉花基因组。朱玉贤说。经过一年的努力，研究人员利用PacBio长片段测序技术和染色质高级结构捕获技术（Hi-C）解析了世界上首个高精度的草棉参考基因组，并对陆地棉和亚洲棉基因组进行质量升级。升级后基因组在准确性和完成度等方面具有明显的提高，填补了大量的基因组漏洞，可作为参考基因组。朱玉贤表示，质量越高的基因组包含的棉花基因越完善，为棉花的基因组研究提供了新的数据资源，这为挖掘控制优良性状基因研究奠定了基础。接着，研究人员通过构建物种系统发育树发现，A亚基因组的供体可能来源于A1和A2的共同祖先A0。为了进一步证实这个猜想，我们对棉花基因组最大的组分转座子进行了系统分析。在朱院士的启发和激励下，我们从根本上革新了棉花转座子的遗传演化算法。区别于以往的算法，我们将棉花基因组完整的和大量碎片化的转座子全部扫描出来，发现棉花转座子不同批次的爆发事件。论文第一作者、武汉大学生命科学院副教授吴志国告诉《中国科学报》。在以往的研究中基因组最大的组分转座子被称为‘垃圾DNA’或者‘暗物质’，归咎于其巨大的重复数量以及随机的碱基替换事件。吴志国进一步解释，此研究中革新的转座子遗传演化算法完美的匹配了其巨大数量和随机性等特征，使植物基因组的完整系统分析成为了可能。

随后，研究进一步证实了陆地棉At、草棉A1、亚洲棉A2来源于共同祖先A0。从进化关系上看，A1基因组比A2基因组更接近于A0。研究揭示出，A0与D5基因组雷蒙德氏棉大约在1.6百万年前（MYA）形成异源四倍体，随后，A0基因组在大约0.7百万年前分化形成现存的A1和A2基因组。我们阐明了亚洲棉与草棉独立起源的问题，即亚洲棉不起源于草棉，而是与草棉同时起源于A0，并独立驯化。吴志国说。值得一提的是，该研究还发现位于基因区附近的变异位点改变了重要基因的表达，可能最终导致了四倍体陆地棉相较于二倍体棉在纤维品质等性状上有明显的改良与提升。优质棉花未来可期 该研究发现现存的非洲棉和亚洲棉可能都不是陆地棉A亚组的供体种，而真实的供体种可能已经不存在，这是棉花基因组进化方面的一大创新性结论。此外，非洲棉和亚洲棉是在陆地棉形成之后才独立分化出来，二者没有祖先与后代的关系。这厘清了长期关于陆地棉A亚组供体种的争论，也是本研究最显著的科学价值。张献龙说。事实上，自2011年起，朱玉贤受中国工程院院士喻树迅的邀请，帮助解决四倍体棉花基因组关键难题至今，他带领团队逐渐建立起了中国棉花基因组理论的框架体系，在雷蒙德氏棉基因组、亚洲棉基因组和群体、陆地棉基因组和亚洲棉转录组等领域均取得了重要突破。

但在朱玉贤等专家们看来，优质棉花之路依然任重道远。优质纤维比较缺乏、了解枯萎病和黄萎病等棉花严重病态的抗性机理、野生棉花种质资源挖掘，这是当前我国棉花的三大难题。朱玉贤说。张献龙也表示，两个异源四倍体栽培棉种的基因组序列虽然已有很大改进，但仍需要进一步矫正错误和填补缺口；多倍化后，在驯化过程中转座子起什么作用值得探讨；亚基因组之间存在不对称的转录调控，这一现象在重要农艺性状形成中发挥了什么作用，以及如何利用这一特征从头驯化棉种或改良现有品种等也值得去探讨。朱玉贤表示，未来，培育高产优质品种，一方面可通过单基因转入育种，目前相关技术手段已较为成熟，且正在试验的品种也显示出良好的效果；另一方面，则是借助分子育种手段，通过鉴定相关功能基因，解析优良性状的分子模块，设计培育高品质棉花的新品种。值得一提的是，目前备受关注的陆地棉与海岛棉基因的研究。

---

陆地棉产量高、适应性强，但纤维品质差，海岛棉纤维品质好，但栽培区域性强、产量低。若能研究清楚各个性状分别由哪些基因控制，则可以将海岛棉中控制优异纤维品质的基因导入到陆地棉，或者将陆地棉中控制产量的基因导入海岛棉中，实现产量品质的‘双得’朱玉贤说。（来源：中国科学报 韩扬眉） 相关论文信息：<https://doi.org/10.1038/s41588-020-0607-4> 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：朱玉贤等 来源：《自然—遗传学》

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