

Analysis reveals cotton genome stability across global lineages

April 20 2020



In the United States, 95 percent of the cotton grown is *Gossypium hirsutum*, known as Upland cotton. This image complements a news release from the DOE Joint Genome Institute regarding a *Nature Genetics* paper published April 20, 2020 reporting that a multi-institutional team has now sequenced and assembled the genomes of the five major cotton lineages. The genomes are available on JGI's plant data portal Phytozome. Credit: Cotton Inc.

Come harvest time, the cotton fields look like popcorn is literally growing on plants, with fluffy white bolls bursting out of the green pods in every direction. There are 100 million families around the world whose livelihoods depend on cotton production, and the crop's annual economic impact of \$500 billion worldwide underscores its value and importance in the fabric of our lives.

In the United States, [cotton production](#) centers around two varieties: 95 percent of what is grown is known as Upland [cotton](#) (*Gossypium hirsutum*), while the remaining 5 percent is called American Pima (*G. barbadense*.) These are two of the five major lineages of cotton; *G. tomentosum*, *G. mustelinum*, and *G. darwinii* are the others. All of these cotton lineages have genomes approximately 2.3 billion bases or Gigabases (Gb) in size, and are hybrids comprised of cotton A and cotton D genomes.

A multi-institutional team including researchers at the U.S. Department of Energy (DOE) Joint Genome Institute (JGI), a DOE Office of Science User Facility located at Lawrence Berkeley National Laboratory (Berkeley Lab) has now sequenced and assembled the genomes of these five cotton lineages. Senior authors of the paper published April 20, 2020 in *Nature Genetics* include Jane Grimwood and Jeremy Schmutz of JGI's Plant Program, both faculty investigators at the HudsonAlpha Institute for Biotechnology.

"The goal has been for all this new cotton work, and even the original cotton project was to try to bring in molecular methods of breeding into cotton," said Schmutz, who heads JGI's Plant Program. He and Grimwood were also part of the JGI team that contributed to the multinational consortium of researchers that sequenced and assembled the simplest cotton genome (*G. raimondii*) several years ago. Studying the cotton genomes provides breeders with insights on crop improvements at a genetic level, including why having multiple copies of

their genomes (polyploidy) is so important to crops. Additionally, cotton is almost entirely made up of cellulose and it is a fiber model to understand the molecular development of cellulose.

Cotton Genomes on Phytozome

The genomes of all five cotton lineages and of cotton D are available for [comparative analysis](#) on JGI's plant data portal Phytozome, which is a community repository and resource for plant genomes. They are annotated with the JGI Plant Annotation pipeline, which provides high quality comparisons of these genomes within themselves and to other plant genomes.

"Globally, cotton is the premier natural fiber crop of the world, a major oilseed crop, and important cattle feed crop," noted David Stelly, another study co-author at Texas A&M University. "This report establishes new opportunities in multiple basic and applied scientific disciplines that relate directly and indirectly to genetic diversity, evolution, wild germplasm utilization and increasing the efficacy with which we use natural resources for provisioning society."

The comparative analysis of the five cotton genomes identified unique genes related to fiber and seed traits in the domesticated *G. barbadense* and *G. hirsutum* species. Unique genes were also identified in the other three wild species. "We thought, 'In all of these wild tetraploids, there will be many disease resistance genes that we can make use of,'" Schmutz said. "But it turns out there isn't really that kind of diversity in the wild in cotton. And this is amazing to me for a species that was so widely distributed."

In the field, growers can easily distinguish the cotton species by traits such as flower color, plant height, or fiber yield. To the team's surprise, even though the major cotton lineages had dispersed and diversified over

a million years ago, their genomes were "remarkably" stable. "We thought we were sequencing the same genome multiple times," Schmutz recalled. "We were a little confused because they were so genetically similar."

Benefits of High Impact Science

"The results described in this *Nature Genetics* publication will facilitate deeper understanding of cotton biology and lead to higher yield and improved fiber while reducing input costs. Growers, the [textile industry](#), and consumers will derive benefit from this high impact science for years to come," said Don Jones, who handles variety improvement for Cotton Incorporated, the research and marketing company representing upland cotton funded by U.S. growers of upland cotton and importers of cotton and cotton textile products, often referred to as the dirt-to-shirt value chain.

Assembling cotton's large and complex [genome](#) means being selective in choosing which team to financially support, Jones added. "We must be careful who we ask to take on these projects due to their difficulty and complexity, but we have been extremely pleased with Jeremy, Jane and their team. Many groups assemble genomes, but very few do it so well that it stands the test of time and is considered the gold standard by the world cotton community. This is one such example."

Jones noted that he talks to growers about Cotton Inc.'s long-term investment in crop research. "What I have told our growers is, 'Think of these reference genomes as a surgeon's knowledge, and of gene editing as a new tool. In order to know exactly where to use your incredibly precise tool, you have to know where to use it, which exact base or series of bases you have to alter.' Why should we invest in something that may not be an immediate benefit to us for a decade? We believe this basic research has to occur in order to drive the research. Oftentimes, these

things take not five or eight years, but sometimes 10 or 15 years, because the technology develops over time."

More information: Genomic diversifications of five *Gossypium* allopolyploid species and their impact on cotton improvement, *Nature Genetics* (2020). [DOI: 10.1038/s41588-020-0614-5](https://doi.org/10.1038/s41588-020-0614-5) , www.nature.com/articles/s41588-020-0614-5

Provided by Lawrence Berkeley National Laboratory

Citation: Analysis reveals cotton genome stability across global lineages (2020, April 20)
retrieved 7 May 2024 from
<https://phys.org/news/2020-04-analysis-reveals-cotton-genome-stability.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.