

COTTON BREEDING: HIGH TECH AND NATURAL

Kater Hake

Anyone who has seen wild cotton growing in Mexico knows how efficient cotton growers have been in the domestication of this plant. From a sparse fruiting tree with less than 1% of its dry weight in harvestable products man created our modern varieties that have 40% of their dry weight as lint and seed.

At one point all the continents were connected in a super continent called Pangea. Cotton's ancestors being hardy plants spread throughout the warm regions of Pangea which started drifting apart about 175 million years ago into the current continents. Scientist believe that only 1 of these cotton plants, out of trillions, subsequently floated across the Pacific ocean and fertilized (crossed with) a distant cotton relative to create the progenitor of Upland (*Gossypium hirsutum*) and Pima (*Gossypium barbadense*) cottons. These male and female plants were different species that had evolved on separate continents – one in Africa and the other in America. This rare ocean voyage and crossing would have failed immediately but not for another extremely rare event, that of chromosome doubling from an infertile diploid to produce the fertile tetraploid that evolved into all Upland and Pima cottons.

This one wild tetraploid plant species over the next 20 million years evolved into the diversity of tetraploid plants which man stumbled upon and started breeding approximately 6,000 years ago. From this beginning, cotton breeding included 2 essential steps (1) the selection of superior plants for replanting, and (2) the expansion of diversity within which to select. Simple, but powerful, and getting more so, as science creates even better selection methods and greater diversity within which to select.

We could consider that “natural” stopped long ago, perhaps when primitive ocean bacteria released oxygen creating the great oxidation over two billion years ago, or perhaps when early man started wildfires to herd game, perhaps when agriculture denuded European forests 8,000 years ago, perhaps when the Conquistadors started mining silver in South America, perhaps when coal burning in England powered their textile mills, or perhaps when the first layer of radioactive Cesium 131 from atmospheric testing coated the earth. All of these are evident in the geological record and all can be used to delineate geological epochs¹. Geologist debate which of many permanent geological marks should be used for the transition from the Holocene to the Anthropocene². All of man impacts are being considered, since geologist 100,000 years from now will find these marks and more when digging in the earth.

¹ Science 2016 *The Anthropocene is functionally and stratigraphically distinct from the Holocene* 351: 137

² Nature 2015 *Defining the Anthropocene* 519: 171-180

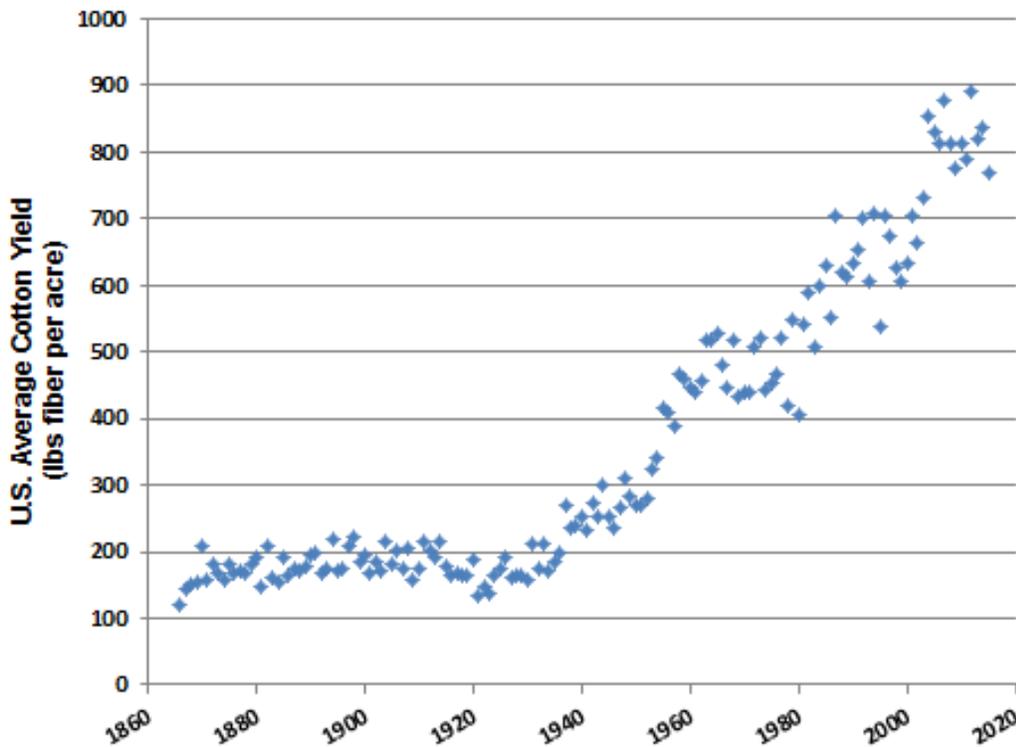
So how does this relate to “Cotton Breeding: High Tech and Natural”? These two essential steps in breeding (1) selection and (2) new diversity, have been on a trajectory that has allowed man’s global population to explode during the last 10,000 years. And by borrowing innovations from medical sciences for plant breeding improvements, there is no question about the ability of the natural fiber from the cotton plant to supply the world’s textile needs for innovative garments for another 10,000 years. In recent years synthetic fibers have made substantial gains in innovation because their “breeding” process can be accomplished in one lab in one week. Various proprietary chemicals (referred to as comonomers) are blended with a base polymer such as Polyethylene terephthalate to add diverse functionality to polyester fibers³. On the other hand, plant breeding currently requires a decade to make slight changes in the fiber and then thousands of acres to evaluate whether these changes are improvements over previous plants. The direct synthetic path to innovation in fiber (pour in a different batch of molecules and extrude novel fiber) versus the indirect path to innovation in plant made fibers (create diverse plants and select the improvements based on field testing) has put cotton at a disadvantage when consumer preferences shift rapidly, as they have with athletic wear. Only by bringing new high tech advancements to cotton breeding can we rebuild cotton’s innovation pipeline to a level that it competes with synthetic fibers.

The geological record of cotton’s domestication has been largely destroyed by the warm wet conditions which promote rapid decomposition of cellulose in regions where cotton has been grown. We are fortunate that dry conditions in the Peruvian Andes have preserved some samples of yarn that date to 5490 before the present.⁴ Since these samples were of sufficient quality for fishing nets, we must assume that cotton was domesticated before that time. From that early domestication 6,000 years ago until the 1920s cotton was grown without the distinction of varieties but rather based on the origin of production. Cotton farmers selected locally adapted land races from within their fields, their neighbors, and from immigrants bringing new seeds. In the 1920s the Mendelian laws of heredity which had recently been rediscovered around the world were applied to plant breeding. The results were spectacular as seen in U.S. cotton yields in the graph below. The deliberate introduction of diversity by crossing different varieties, versus just selecting from within existing land races, created novel varieties that deserved varietal distinction and subsequent propagation on millions of acres. Varieties such as Deltapine 16 were introduced into Asia and Australia launching a resurgence of cotton production that continues today. The selection process was also enhanced by statistical methods of R.A. Fisher in 1920⁵. Instead of identifying superior plants solely by eye, field experiments were conducted and mathematical principles applied to assess varietal performance.

³ Polymer 2002 Analysis of comonomer content and cyclic oligomers of poly (ethylene terephthalate). 43 (6) 1797-1804

⁴ Science 2007 Preceramic Adoption of Peanut, Squash, and Cotton in Northern Peru. 316: 1890-1893

⁵ R.A. Fisher 1925 *Statistical Methods for Research Workers* (ISBN 0-05-002170-2)



High tech statistical tools for selection came along with instruments to measure fiber quality - stelometer to measure bundle strength and fibronaire to measure fiber micronaire. High technology was also applied to the diversity side of plant breeding with radiation induced mutations and chemical induced mutations in the 1930s to 1980s. In the 1990's the use of wild cotton relatives to increase genetic diversity was demonstrated by the movement of Reniform resistance from the diploid *Gossypium longicalyx* into the tetraploid *G. hirsutum* Upland cotton.⁶ This was accomplished using repeated crossing with diverse cotton species to find a suitable bridging species since *G. longicalyx* and *G. hirsutum* are not compatible using traditional crossing.

High tech selection was boosted with the discovery of polymerase chain reaction (PCR) in 1983 but did not find applications in cotton breeding until the 2000s with the first publication of cotton Simple Sequence Repeat (SSR) markers.⁷ Although the cotton genome had not been sequenced at that time it was possible using large recombinant inbred populations to associate an SSR marker with a major phenotypic trait. Cotton breeding has benefited from the investments in medical research to lower the cost of PCR by combining multiple tests in one run. This method is used by breeders of biotech crops to identify the presence or absence of GMO traits. PCR methods are just now being applied to selection of cotton germplasm with superior disease tolerance and will rapidly be applied to fiber quality as well.

⁶ J. Nematology 2005 *Introgression of resistance to Rotylenchulus reniformis into Gossypium hirsutum from G. longicalyx* 37: 391.

⁷ J Heredity 2000 *Chromosomal assignment of microsatellite loci in cotton* 91:326–332.

The latest iteration of high tech diversity has been biotechnology. Biotechnology has allowed diversity to expand beyond that which is accessible using crosses - beyond the genus *Gossypium*. Biotechnology was first demonstrated in the mid 1980's and planted in commercial cotton fields in 1995. Using a natural process of plant infection that creates crown galls in the wild, scientists have been able to move genes from bacteria into individual cotton cells which are then grown into fertile plants. Once the novel gene has been stably inserted into a fertile plant it can be moved using the low tech crossing methods employed in the 1920's.

Two high tech methods that have not yet been applied to cotton, but used in other crops, include the selection of superior plants using Genomic Selection and the creation of novel diversity using Genome Editing. As both of these names imply, they rely on the plant's genome – the sequence of 4 different DNA bases (A,C,T and G) that control most of an organism reproduction and development. The Upland cotton genome is 3 billion DNA bases long and broken into 52 segments or chromosomes. It was just made available in 2015 for one variety, TM1. ^{8 9}

Genomic selection is a method to use the genome of new varieties to predict which are superior even before field testing¹⁰. It relies on computer models build on previous training populations of hundreds of varieties which have been scrutinized for their phenotype in the field and their genome. Once a Genomic Selection model has been created it can be improved with each passing year and will be a powerful breeding tool to improve cotton when adopted.

Genome editing is an even more recent tool to expand diversity¹¹. A natural defense mechanism that bacteria employ to destroy viruses (phage) called CRISPR has been adapted to a wide variety of plants. CRISPR genome editing avoids the random mutation and insertions of radiation and biotechnology by creating a DNA break at precise points in a plant's genome. At this time, CRISPR use in cotton has not been announced, but applications in crops are expanding rapidly. Although previous methods of genome editing have been used in some crops (zinc finger nucleases, TALENS, meganucleases) only CRISPR offers the speed of application and ease of use that allows companies, Institutes and Universities around the world to employ it in breeding crops for which a genome has been published. Early applications of CRISPR will be similar to radiation and chemical mutagenesis in altering single DNA bases with the huge difference that these random changes in a genome of 3 billion DNA bases require massive populations scrutinized over decades to find the one desired DNA base change. CRISPR tools allow precise editing of a single DNA base in one plant in a just a few weeks. Considering the recent discovery of CRISPR and its rapid improvement, the full capability to create novel diversity and accelerate plant breeding is not known at this time.

⁸ Nature Biotechnology 2015 *Genome sequence of cultivated Upland cotton (Gossypium hirsutum TM-1) provides insights into genome evolution* 33:524-530

⁹ Nature Biotechnology 2015 *Sequencing of allotetraploid cotton (Gossypium hirsutum L. acc. TM-1) provides a resource for fiber improvement* 33:531-537

¹⁰ Crop Science 2009 *Genomic Selection for Crop Improvement* 49:1-12

¹¹ Science 2012 *A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity* 337:816-821

On the horizon are even more powerful techniques to expand the diversity within which plant breeders can select high yielding, better adapted and improved fiber cotton varieties – organelle breeding, epigenetic breeding and metabolic engineering. Organelle breeding is the improvement of Upland cotton's metabolic organelles (mitochondria and plastids) by accessing organelles from wild cotton species.¹² Wild cotton's have retained higher levels of heat and drought tolerance which may be transferable into Upland and Pima cottons either directly with traditional crossing or through genome editing. Epigenetic breeding relies on the natural tagging of DNA bases which determines how a plant creates different plant tissues with just one genome. Methods to create epigenetic diversity have only recently been published in model plants¹³.

A third source of diversity is substantially further away from application in any crop – photosynthetic engineering. The highest yielding crops (maize, sorghum, sugarcane and grain amaranth) use an efficient photosynthetic machinery referred to as C4, for the number of carbon atoms in the molecule first created by capturing carbon from the air. Less efficient crops such as cotton, wheat, soybean and rice utilize a C3 carbon capture method. Significant investment is ongoing to transfer C4 carbon capture into rice. Until this is achieved other crop applications will remain in the distant horizon.

The basics of cotton breeding will be the same in 2020 as it was over 6,000 years ago when man selected superior plants for replanting from within the diversity he saw in his fields. Selection and diversity are the two elements that are unchanged. What has changed is our ability to select superior plants and our ability to create greater diversity within which to select from. Whether these advancements are natural or not is subject to discussion, but their ability to add value to the entire textile chain is undeniable. Advancements in medical research and basic science are being applied to agriculture at a rapid pace. We are just now seeing the first of these applications in cotton breeding with dramatically improved fiber length, strength and yield. When the full suite of tools currently employed in other crops becomes available for cotton breeding, these gains will accelerate. When we access the exciting tools on the horizon, breakthrough advances that we cannot now envision will likely become commonplace. Cotton fiber textile consumers have much to look forward to in greatly improved fiber, thus developing relationships with plant breeders at Universities, Institutes and companies to encourage varieties that satisfy specific textile needs is timely and prudent.

¹² Cell Systems 2016 *Evolutionary Inference across Eukaryotes Identifies Specific Pressures Favoring Mitochondrial Gene Retention* 2:101-111

¹³ Nature Communications 2015 *Arabidopsis MSH1 mutation alters the epigenome and produces heritable changes in plant growth* 6:1-9