

COORDINATED RESOURCES LINKING THE GENOME TO COTTON IMPROVEMENT

Don C. Jones

Cotton Incorporated, Cary, North Carolina, USA

The USA was once the world's leading cotton producer, but it is now third behind India and China. To mitigate the decline, the US industry must continue to improve yield using improved grower decisions, precision agricultural tools, and knowledge gleaned from big data analyses. Equally important for cotton improvement is increasing the rate of genetic gain, especially in regards to fiber quality. One approach to increase genetic gain is understanding the connections between genomics, genetics, and breeding, and then coordinating those resources to drive improvements sought by downstream customers.

Elite cultivated cottons come from only a tiny subset of the cotton gene pool, and a host of valuable traits are well known to exist in wild relatives. However, genetic improvements must be precisely targeted, altering only essential genes to avoid unanticipated and undesirable effects. In order to achieve the required level of precision, genomic knowledge and tools must be readily available to cotton scientists. These are the building blocks of success. One such building block is a reference grade Upland cotton genome sequence, often times called a gold standard. Planning for this commenced in earnest at the International Cotton Genome Initiative Conference in 2002. After a decade of sustained effort, the cotton community's first gold standard genome assembly, the diploid progenitor *Gossypium raimondii*, was published in December, 2012. Using that as a foundation, two draft Upland cotton (*Gossypium hirsutum*) genomes were developed and eventually published in May, 2015. A gold standard Upland genome sequence assembly is nearing completion as this abstract is being prepared, and it should facilitate cotton improvement for years to come.

Just as sequencing technology has advanced and helped drive development of genome assembly, so too has understanding of breeding designs advanced. A promising opportunity to enhance the genetic diversity of domesticated Upland cotton are use of wild relatives with valuable attributes. This is needed because cotton has long been recognized as being among the most genetically depauperate of the world's major crops. US cotton germplasm needs an injection of new diversity to re-invigorate genetic gains, and this in turn will help improve competitiveness for producers. A national partnership involving applied and basic scientists in nine states, USDA-ARS, National Plant Germplasm System, and Cotton Incorporated began to invigorate US cotton improvement through a new initiative starting in 2011. Over the past 8 years, the US cotton community has invested significant resources to develop genetic populations that will allow for identification of valuable genes that were 'left behind' in the process of cotton domestication and improvement. This effort is called the Nested Association

Mapping project, or NAM. Over the next few years, these populations will undergo DNA fingerprinting and evaluated in detail for fiber quality performance and breeding value across environments, with careful analysis of the results to guide effective deployment of these valuable genes in mainstream cotton improvement. An expanded winter nursery was required to accelerate its deployment, plus improved archival and seed distribution infrastructure. Enriched statistical and computational resources will be required for efficient data utilization. Outcomes of this effort are expected to include substantially improved fiber quality, and knowledge of branches of the 'cotton family tree' from which new genetic innovations are identified.

Paterson, Andrew H, et al. 2012. Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibres Nature 492, 423–427

Li, Fuguang, et al. 2015. Genome sequence of cultivated Upland cotton (*Gossypium hirsutum* TM-1) provides insights into genome evolution Nature Biotechnology. 33: 524–530

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

McMullen MD, et al. 2009. Genetic properties of the maize nested association mapping population. Science 325: 737-740