Unravel the key genes potentially related to high strength of cotton fiber by comparative phenotypic and genomic analyses

The demand of high strength of cotton fibers has been increased dramatically with the advent of modern high speed spinning technology for producing yarn. Cotton fiber is a gigantic single cell which consists of almost pure cellulose. Thus, the strength of cotton fibers can be determined by the factors regulating the content, amount, and order of the cellulose. Individual fiber strength is the most critical quality attribute that greatly influences the strength of the yarn spun from cotton fibers. Fiber strength is usually measured from bundles of fibers due to the difficulty of reliably measuring strength from individual cotton fibers.

Fig. 1. Development of near isogenic cotton lines, MD 52ne and MD90ne that are genetically almost similar, but differ in fiber strength.

However, bundle fiber strength is not always correlated with the strength measured from individual fibers and yarns since it is affected by multiple fiber properties involved in fiber-to-fiber interactions within a bundle in addition to the individual fiber strength. Cotton researchers have tried to improve this trait in Upland cotton (*Gossypium hirsutum*) genetic backgrounds. Molecular mechanisms responsible for regulating fiber strength remain unknown. *G. hirsutum* germplasm near isogenic lines (NILs), MD52ne and MD90ne were developed through backcross (BC) breeding. MD90ne is the recurrent parent and MD52ne is a BC_6 having high BFS gene (Fig. 1). As results, the MD52ne obtained a gene promoting fiber strength became genetically almost identical to the MD90ne.

This study was conducted to unveil the molecular mechanisms behind the formation of superior individual fiber strength, which is correlated with yarn strength, in MD52ne using RNA-seq technology. Comparative fiber property analyses revealed that MD52ne have higher (~20%) fiber strength both single and bundle than its NIL MD90ne (Fig. 2A). Comparative genome mapping analyses using F_2 population derived from a cross between two NILs discovered that loci associated with high fiber strength is located on Upland cotton chromosome A03 at 84 mega base pairs (Fig. 2B). Comparative transcriptomic analyses were conducted following microarays, RNA sequencing (RNA-seq) and reverse transcribed quantitative PCR (RT-qPCR) (Fig. 2C).

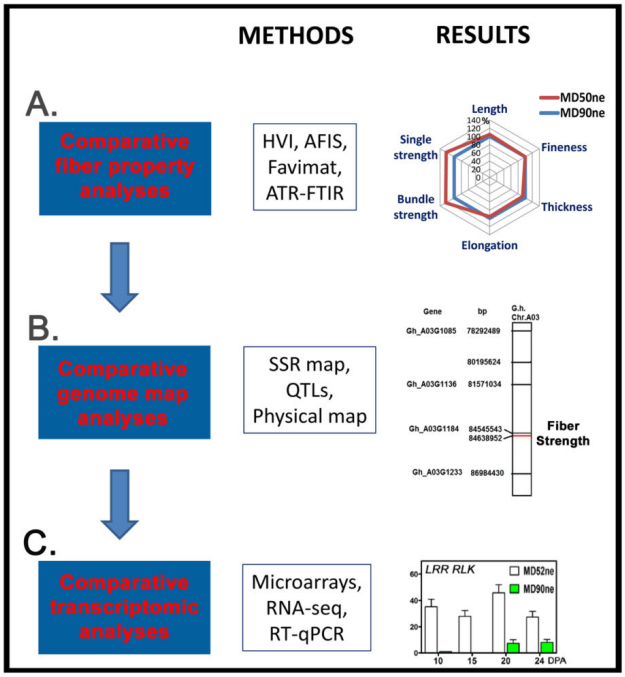

Fig. 2. Comparative analyses between MD52ne and MD90ne cotton lines. Abbreviated names included in this figure are: AFIS, Advanced fiber information system; ATR-FTIR, Attenuated total reflectance-Fourier transform infrared spectroscopic; HVI, Heavy volume instrument; LRR RLK, Leucine rich repeat receptor like kinase; QTL, Quantitative trait loci; SSR, Simple sequence repeat;, RT-qPCR, reverse transcribed quantitative polymerase chain reaction.

RNA-seq results revealed that differential expressions of the genes involved in crystalline cellulose assembly, ethylene phytohormone and *receptor-like kinase* (RLK) signaling pathways between the MD52ne and MD90ne developing fibers. Transcripts related to ethylene and its networking signaling pathways for promoting fiber elongation were highly abundant in MD52ne. Ethylene gas is known as a major phytohormone stimulating fiber development. Ethylene and its phytohormonal network might promote the elongation of MD52ne fibers and indirectly contribute to the bundle strength by potentially improving fiber-to-fiber interactions. Interestingly, multiple RLKs localized in

plasma membranes were differentially expressed in the developing fibers of the stronger cotton line, MD52ne as compared with the weaker line, MD90ne and localized in genomic regions encompassing the strength quantitative trait loci. Among the various classes of RLKs, the leucine rich repeat (LRR) RLKs containing three domains (LRR ligand binding motif, a transmembrane region and a kinase domain) were most frequently abundant. The gene related to LRR RLKs also located in the genomic region of fiber strength loci on chromosome A03. The LRR RLKs have been recently suggested as a novel signaling pathway regulating plant cell wall integrity maintenance and cellulose deposition in other plants. An *LRR RLK* (*Gh_D08G0203*) was one of the DEGs that were most highly up-regulated in developing MD52ne fibers at both 15 DPA (377 fold) and 20 DPA (702 fold). Several candidate genes involved in crystalline cellulose micro unit assembly were also up-regulated in MD52ne fibers while the cellulose was actively produced in cotton fibers. Comprehensive analyses of the NILs revealed that the superior bundle strength of MD52ne fibers resulted from high individual fiber strength with minor contributions from greater fiber length and the RLKs is candidate genes for regulating cotton fiber cell wall assembly and strength.

Md Sariful Islam, David D. Fang, Hee Jin Kim
USDA-ARS, Southern Regional Research Center, Cotton Fiber Bioscience Research Unit, New Orleans, LA 70124, USA

Publication

Comparative fiber property and transcriptome analyses reveal key genes potentially related to high fiber strength in cotton (Gossypium hirsutum L.) line MD52ne.
Islam MS, Fang DD, Thyssen GN, Delhom CD, Liu Y, Kim HJ
BMC Plant Biol. 2016 Feb 1