

## Sequencing of the Cultivated Tetraploid Cotton Genome - *Gossypium hirsutum*

YU Shu-xun<sup>1</sup>, WANG Kun-bo<sup>1</sup>, LI Fu-guang<sup>1</sup>, et al,  
KOHTEL Russell J<sup>2</sup>, PERCY Richard G<sup>2</sup>, YU John Z<sup>2</sup>, et al

(1. *Cotton Research Institute, Chinese Academy of Agricultural Sciences; Key Laboratory of Cotton Genetic Improvement, Ministry of Agriculture, Anyang, Henan 455000, China*; 2. *USDA-ARS, Southern Plains Agricultural Research Center, College Station, TX 77845, USA*)

Cotton is an important cash crop in the world, and it plays an irreplaceable role in China's national economy. Cultivated upland cotton (*Gossypium hirsutum* L.) represents 95% of the world's cotton production, but it has a complex allotetraploid genome that contains at least 30000 genes in 2500 Mb DNA. Sequencing of the cultivated upland cotton will make it possible that fiber yield and quality could be improved directly at the molecular or gene level. Many genes, related to fiber yield and quality, are located in this genome including a pair of homoeologous chromosomes 12 and 26. China National Cotton Research Institute (CAAS-CRI), in cooperation with Southern Plains Agricultural Research Center (USDA-ARS), launched a project of sequencing chromosomes of the upland cotton genetic standard TM-1. With the funding support from China's Ministry of Finance, an initial 10 million RMB was invested with Beijing Genomics Institute to sequence 400 minimum tilling path (MTP) BAC clones that were developed from 3258 BAC clones on chromosomes 12 and 26 by the BAC to BAC method. Significant progress was made in DNA sequence assembly and gene annotation of these clones. Detailed results from the upland cotton genome sequencing project and related genomics research areas will be reported.

**Key words:** upland cotton; genome sequence; fiber yield and quality