Genome Sequencing of *Xanthomonas citri* pv. *malvacearum* and the PCR-based Detection of the Cotton Bacterial Pathogen from Seed and Plant Materials

Shien Lu, Department of Biochemistry, Molecular Biology, Entomology and Plant Pathology, Mississippi State University

Bacterial blight of cotton caused by *Xanthomonas citri* pv. *malvacearum* (Xcm), is considered one of the most serious diseases of cotton and results in substantial yield losses. Use of pathogen-free seeds is the key practice to prevent introduction and development of the cotton disease. The objectives of this study were to draft the genome sequence a local strain (MSCT1) and to develop a PCR-based pathogen detection method. Illumina sequencing was conducted to generate the genome draft of strain MSCT1. The resulting assembly had a sum length of 5,142,224 base pairs distributed between one large circular scaffold of 49 contigs and 13 unplaced scaffolds. Sequencing depth was 247.7 and 39.5 genome equivalents for the Illumina PCR-free DNA library and Mate Pair library, respectively. The MSCT1 genome average nucleotide identity (ANI) analysis showed strain MSCT1 was most closely related to the Xcm race 18 isolates. A total of 4,247 protein coding genes were predicted by Prodigal. Gene ontology analysis yielded a total of 2,252 molecular function, 2,181 biological processes, and 882 cellular component annotations on 2,614 proteins. The rRNA segments were comprised of two copies of each of the 23S, 5S, and 16S rRNA subunits. At least one tRNA for each of the 20 basic amino acids was identified in the 54 predicted tRNA loci. The program RepeatModeler predicted one DNA insertion element and two uncategorized repeats. The DNA insertion element was found in 20 locations in the genome collectively occupying 8,480 bp. The two uncategorized repeats were identified in 84 locations and all together these repeats account for 40,419 bp of the genome. Based on genome-wide comparisons, more than 50 pairs of PCR primers were identified and synthesized. After comprehensive analyses of specificity, a novel and highly sensitive TaqMan real-time PCR detection protocol was designed from a genetic locus coding for a hypothetical ATP-binding protein. The developed primers are specific to five different cotton pathogenic Xcm races tested in this work but not other *Xanthomonas* species that are nonpathogenic to cotton. The TaqMan based real-time PCR assay was optimized and employed for detection and quantification of Xcm in plant samples. The efficiency of the assay was evaluated with the cotton leaf and seed samples carrying Xcm. The efficiency was 100 cells from 1 gram fresh leaf samples and the detection limit of the reaction was 37cells from 1 gram seeds. The newly developed TaqMan real-time PCR protocol provides a rapid accurate method for diagnosis of the disease and detection of the pathogen.