CITRUS ADVANCED TECHNOLOGY PROGRAM

QUARTERLY & FINAL REPORTS: Control of Citrus Greening, Canker & Emerging Diseases of Citrus

SELECT PERIOD June 2013

(●) *Quarterly Report* ( ) *Final*

Proposal Title

Screening and Cloning of Resistance Related Genes by RNA-Seq in Huanglongbing (HLB) Resistant and Susceptible Citrus Breeding Lines

Today's Date	Sponsoring Organization (drop-down)	Category (drop down)
7/12/13	Citrus Research and Development Foundation	Citrus Genomics

REPORT UPDATE (500 words-It is not necessary in this public report to disclose your institution's proprietary information or intellectual property.)

The objectives of this project are: 1) to generate transcriptome profiles of both susceptible and resistant citrus responding to HLB infection using RNA-Seq technology; 2) to identify key resistant genes from differentially expressed genes and gene clusters between the HLB-susceptible and HLB-resistant plants via intensive bioinformatics and other experimental verifications such as RT-PCR; and 3) to create transgenic citrus cultivars with new constructs containing the resistant gene (s).

The first group of 5 samples for RNA-Seq, including resistant/tolerant vs. susceptible plants are nearly complete. The second group of 10 samples is in the sequencing process.

We mapped the RNA-Seq data to a reference genome, C. clementina using the bioinformatics program STAR. About 85% of the raw reads could be uniquely mapped. The transfrags of each library were assembled with cufflinks and merged with cuffmerg. 24,275 genes of the originally predicted genes had been found to be expressed and a total of 10,539 novel transfrags were identified with cufflinks, which were missing from the original reference genome annotation.

Some of the NBS genes were found to be expressed. For C. clementine and C. sinensis, there were 118,381 and 214,858 mRNAs or ESTs deposited in GenBank and 93 out of 607 and 221 out of 484 NBS related genes match one or more ESTs respectively. The number of ESTs varied from 1 to 25. The expression abundance of each gene was measured by FPKM. The distribution curves of density of FPKM of 5 samples are very similar, indicating that the gene expression is similar and the quality of sequencing is high.

We also performed the principal component (PC) analysis study on the expressions of five samples. The plot of the first PC against the second PC showed that R2017 and R20T18 clustered together (the resistance group) and R19T23, R19T24 and R20T24 clustered together (the susceptible group). This result showed that the gene expressions were significantly different in resistant vs. susceptible citrus. Using cuffdiff, a total of 821 genes were identified as difference expressed genes (DE genes) between the two groups using both p-value and an FDR threshold of 0.01. Among them, 306 genes are up-regulated and 515 are down-regulated in resistant citrus. Using the program iAssembler, a total of 53,981 uni-transfrags were obtained. Most of the assembled uni-transfrags should be novel genes, compared with the citrus reference genome.

To reveal the differences in resistance, we also identified the exon variations (SNP/INDEL). A total of 612,618 SNP/INDELs were identified using the mpileup method employed in samtools. We focused on two types of mutations that could contribute to the resistance difference. The first type of mutation is a mutation in the genes of susceptible citrus leading to pseudogenes (Type1) and the other type of mutation is a mutation in resistant citrus genes that may gain a new function (Type2). Type1 mutations should have a homozygote mutation genotype in the susceptible citrus. We identified 146 candidate genes having Type1 mutations, which produced high impact variations such as frame shifts, splice site acceptors, splice site donors, a start lost, a stop gain or stop lost, and 3,578 genes with Type2 mutations. We expect that as the number of libraries being sequencing increases, the number of candidates will be reduced to a reasonable number allowing for further validation.

We identified a few LRR-PKs genes for further comparative study based on the RNA-Seq data. The results indicated sequence variations of these genes in different varieties are indeed due to SNP/indels, and some of them were annotated as putative pseudogenes because of a truncation. Further verification is underway.

PI First Name Yong-Ping

PI Last Name Duan

Email Yongping.duan@ars.usda.gov

Phone 772-462-5640

Organization USDA-ARS-USHRL

Project Number 523

Project Duration (years) 3 *Year of Project* 2

Total Funds (current year) $155,345.00