Phase 2 of high-throughput sequencing as a CCPP routine diagnostic tool for variety introduction

G. Vidalakis, University of California Riverside, vidalg@ucr.edu
K. Cardwell, Oklahoma State University, kitty.cardwell@okstate.edu
A. Espindola, Oklahoma State University, andres.espindola@ okstate.edu
T. Dang, University of California Riverside, tdang004@ucr.edu
S. Bodaghi, University of California Riverside, sohrab@ucr.edu
I. Lavagi-Craddock, University of California Riverside, irenela@ucr.edu
F. Osman, University of California Riverside, fmosman@ucdavis.edu

Year 2 of 4 (60% Complete)

Problem and Significance

High-throughput sequencing (HTS) is a powerful technology that combines molecular biology and computer sciences that has been applied to diagnostics. The current tools to detect graft transmissible citrus pathogen are polymerase chain reaction (PCR), enzyme-linked immunosorbent assay (ELISA), and biological indexing. The consensus from the scientific and regulatory community is that HTS will soon become the new gold standard for diagnostics. HTS allows for the simultaneous detection and identification of multiple pathogens from a plant sample without the need to perform multiple tests. The sequenced data can be stored long term with minimal footprint and reanalyzed as new pathogens are discovered and characterized.

HTS is a promising tool, however there are challenges with the data analysis. The analysis requires dedicated personnel with in-depth background in bioinformatics analysis and expensive computer infrastructure. In addition, the lack of standardization for data analysis between diagnostic laboratories can result in inconsistent results. A promising tool that can manage the HTS data analysis is the E-probe Diagnostic Nucleic Acid Analysis (EDNA) technology developed by Oklahoma State University (OSU) Institute of Biosecurity and Microbial Forensics. EDNA provides a user-friendly online interface to analyze raw HTS data for citrus pathogens without the need for a dedicated bioinformatician. The e-probes are *in silico* probes that are designed to detect presence or absence pathogens. Multiple e-probes can be used in parallel for the simultaneous detection of multiple pathogens in a single sample, which will reduce time and cost.

When e-probes are designed for a specific pathogen, they will need to undergo extensive validation to prove that EDNA will work equally well or better than current regulatory approved diagnostics assay and determine the limit of detection (LOD) for each e-probe assay.

Benefit to Industry

The EDNA technology can be transferable to various diagnostic laboratories because of the online platform and ease of use. The Citrus Clonal Protection Program (CCPP) and other citrus variety introduction programs can utilize this technology to benefit the citrus industry by providing access to pathogen tested budwood for citrus grove establishments. In addition, this technology will reduce time for the indexing pipeline and greenhouse footprint for the CCPP. This will allow for a more efficient pipeline for the introduction and distribution of new and established citrus varieties in California.

Progress Summary

In year two, we developed and completed the validation for citrus tristeza virus (CTV), citrus exocortis viroid, and '*Candidatus* Liberibacter asiaticus' e-probes. The detailed validation of these 3 e-probe sets will set the path on how subsequent e-probes will be validated.

We have focused most of our time in refining the statistical analysis for EDNA scores. The statistics are important to determine whether a sample is positive or negative for the targeted pathogens.

Scores are generated based on the e-probe and target match and the diagnostic calls are determined by t-test between the target and decoy score within a sample. If the p-value < 0.05 the sample is positive for the targeted pathogen. An issue arose where the healthy sample scores had no variance (Figure 1) and provided undefined values, as result the LODs could not be calculated. We overcame this issue by designing specialized citrus housekeeping gene internal control e-probes that consistently generated sufficient variance used to accurately calculate the LOD of an e-probe and without obtaining false positive results (Figure 1).

We have continued the development of e-probes for citrus pathogens with and without complete genomes from publicly available databases. We have completed mining the pathogen genome sequences from GenBank and assembling genomes, and we have sequenced various graft transmissible citrus pathogens from the CCPP disease bank. So far, we have developed e-probes for citrus tatter leaf virus (CTLV), citrus vein enation virus (CVEV), *Xanthomonas citri* (canker, *X. citri*), and *Xylella fastidiosa* (variegated chlorosis, *X. fastidiosa*), *Spiroplasma citri* (stubborn, *S. citri*). Validation of these e-probes has begun with simulated Illumina data and will continue with real life samples as established with the original set of e-probes for CTV (tristeza), CEVd (exocortis), and CLas (huanglongbing).

For this technology to be adopted for routine use in citrus pathogen testing and variety introduction, it must go through regulatory approval.

During the Western Extension Research Activity - WERA 20 “Virus and Virus-Like Diseases of Berries, Fruit and Nut Trees, and Grapevines” conference on May 2021, organized by Prof. G. Vidalakis, the use of HTS based technologies as a diagnostic tool for National Clean Plant Network (NCPN) centers was a critical topic of discussion. We learned that many NCPN and United States Department of Agriculture (USDA) laboratories have begun to use on a limited basis HTS for diagnostics and release of quarantine propagative materials in some crops. Currently, the main obstacle for the deployment of these methods is to develop standards and guidelines for the uniform computer data analysis. This is a perfect opportunity for EDNA to address some of the standardization problems with HTS data analysis which can help with regulatory approval since EDNA results are based on an automated statistical analysis and not on personalized interpretation of data by a bioinformatician.

From our discussions with USDA, we have learned that the Animal and Plant Health Inspection Service Plant Protection and Quarantine (APHIS PPQ) Science and Technology (S&T) as well as USDA Germplasm Field Op have begun working on creating labs dedicated on using HTS for routine plant diagnostics and have been highly receptive in using this technology. In addition, we have maintained direct discussions with regulators and scientists at USDA regarding the use of EDNA technology. We have arranged an online workshop (October 28th, 2021) with USDA Animal and Plant Health Inspection Service (APHIS) to present our findings and provide hands on training with the online EDNA platform. In addition, we have scheduled hands on training through the American Phytopathological Society Plant Health Workshop series (November 16th, 2021 and to be determined date in February, 2022). This will allow for more exposure of the technology and potential adoptions for other laboratories.

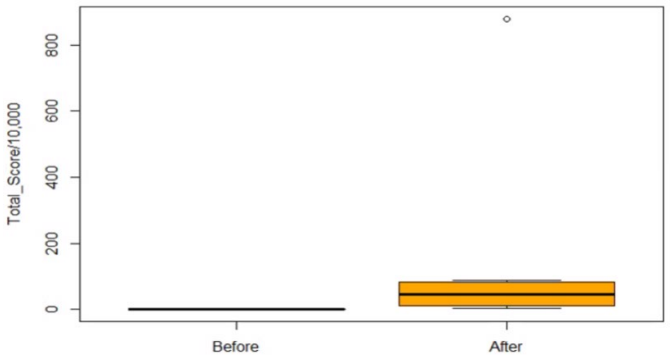

Figure 1. E-probes targeting citrus exocortis viroid e-probe tested against healthy samples (n=10) before and after the addition of special internal control.

CRB Project # 5300-205

Publications and Presentations

T. Dang, H. Wang, A. Espindola, J Habiger, G. Vidalakis, K. Cardwell. (2021) 'Determining Limit of Detection of High Throughput Sequencing Diagnostics by Including Internal Controls in E-probes.' American Phytopathological Society. Online Conference.

Vidalakis G. (2021) 'Manage of Disease Caused by Systemic Pathogens in Temperate and Sub-tropical Fruit Crops and Woody Ornamentals', Western Extension Research Activity (WERA 20) Conference.