CRB Project No. 5100-135

REVISED

CITRUS RESEARCH BOARD

PROJECT PLAN - RESEARCH GRANT PROPOSAL FOR FY2010-2011

Fiscal Year: 2010-11 **Anticipated Duration of Project:** 3 Years

This Project is: _____ **New or** XX **Ongoing (Year** 3 **of** 3 **)**

Project Leader: Abhaya M. Dandekar, Cristina E. Davis, Oliver Fiehn and Raissa D'Souza
Name
Plant Sciences Department, University of California, Davis CA, 95616
Mechanical and Aerospace Engineering, University of California, Davis CA, 95616
Genome Center, University of California, Davis CA, 95616
Location: Mechanical and Aerospace Engineering, University of California, Davis CA, 95616
Address *(University department, if applicable)*
Mailing Address (if different): _____

530-752-7784	530-752-8502	amdandekar@ucdavis.edu
530-754-9004	530-752-4158	cedavis@ucdavis.edu
530-754-8258	530-754-9658	ofiehn@ucdavis.edu

Phone: 530-754-8405 **FAX:** 530-754-9089 **E-Mail:** raissa@cse.ucdavis.edu

Cooperating Personnel: _____

Project Title: A generalized reagentless sensor to detect citrus plant and fruit responses

Keywords: Reagentless detection, citrus, fruit quality, asymptomatic diagnostics, host responses, volatile organic compounds

Abstract *(limit 200 words):(clearly and succinctly state what your project is about why you are doing it and expected out come and how the industry will use these outcomes)*

In this project we will develop a novel detection system based on the Differential Mobility Spectrometer (DMS), a portable and field-mobile device that analyzes the induced volatiles emitted by plants when they respond to stress induced by pathogens and insect pests. We have already tested HLB infected samples with the DMS under laboratory conditions using field samples employing mathematical algorithms to identify particular patterns of volatile compounds and we aim to test this instrument in the field to optimize and develop a robust and easy-to-use tool for citrus growers. The first outcome of this project will be a version of this device in the orchard to accurately distinguish different disease states under field conditions in real-time. The second outcome will be the understanding of the mechanisms of plant responses to pathogen attack. Over the remaining year, our team will identify the specific responses in citrus trees to the three high-priority diseases in citrus (HLB, CTV and CVC) and we will develop a repository of information on plant-pathogens interactions using mathematical and statistical approaches that integrate genes and metabolites in an interconnected network of interactions. This biological regulatory network will be useful to the industry to understand the mechanisms that regulate the plant response to disease and their relationship to emitted volatiles.

Problem and its Significance*: *(include literature review)*

*Use as much space as necessary; attach additional pages as needed; budget info and signatures will appear on last page.

CRB Project Plan – Research Grant Proposal for FY 2010-2011

Project Title: A generalized reagentless sensor to detect citrus plant and fruit responses

Disease detection and control are the #1 priority of the citrus stakeholder community. Development of new or improved diagnostic techniques is a standing request of the citrus industry. Exotic diseases, complex biology, and epidemiology of existing pathogens (e.g., low titer and uneven distribution in the tree) are constant threats to the sustainability of citriculture in the U.S.

The goal of this project is to improve the efficiency and accuracy of disease scouting through the deployment of a Differential Mobility Spectrometer (DMS) a portable technology for pathogen detection in the orchard. The DMS will provide a real time diagnostic tool to rapidly determine the type of pathogen or to determine the presence of disease and disease causing organisms at a pre-symptomatic stage where growers and stakeholders have conventional options to control disease spread. Therefore, the proposed approaches will help detect diseases before the occurrence of secondary spread in the orchard, and potentially with the first appearance of infected insects based on induction of volatile organic compounds (Fig.1).

Practical application of these technologies requires an understanding of the mechanisms that regulate bacteria and plant responses. This understanding helps validate the diagnostic systems and provides new early disease biomarkers detectable by these technologies. Another important application is the development of new therapeutic strategies that limit disease spread by creating plants that resist a certain disease or engineering them to, repel or attract the vectors, in order to improve disease management and reducing hazardous pesticide use. A systems-based approach is essential to understand the biological network that regulates important pathways associated with bacteria-host responses to disease.

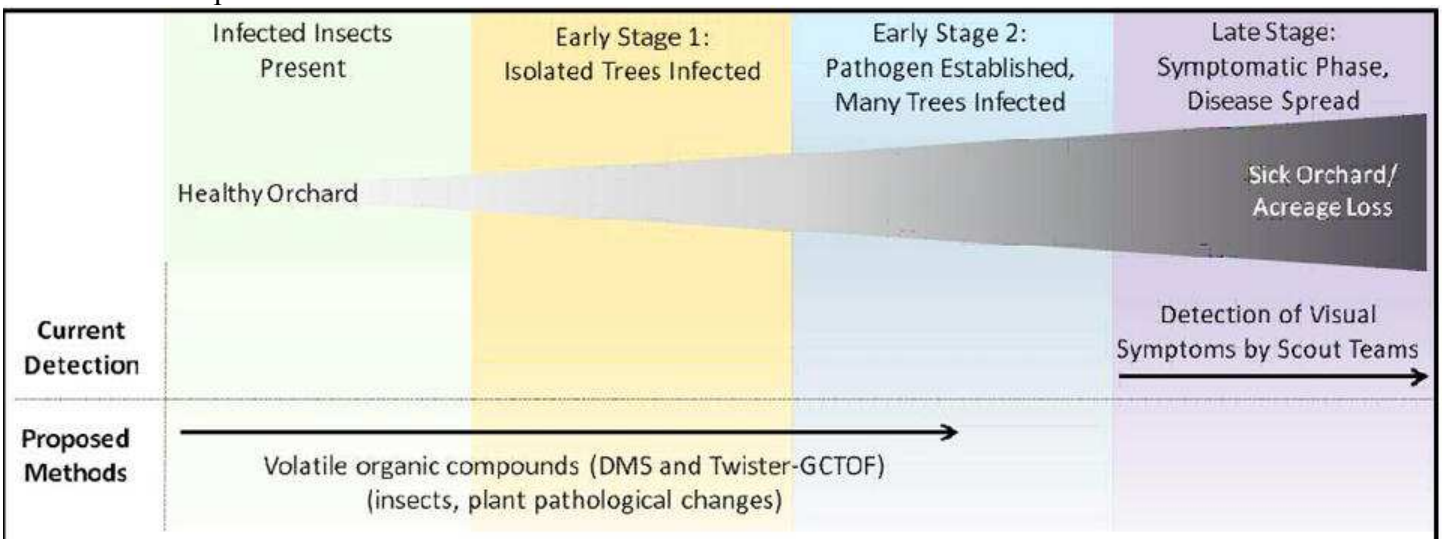

Figure 1: Proposed methods for real time field detection and timing of their use during disease progression.

Technologies such as deep transcriptome sequencing allow more thorough examination of the transcriptome than microarrays or cDNA subtractive libraries (‘t Hoen et al., 2008) These data can be correlated with the IVOC profile analysis and database developed using the Twister-GC-TOF to investigate interactions between genes, RNAs, proteins, and emitted metabolic and volatile organic compounds (VOCs). We will integrate the complex data sets obtained from our experiments with known libraries that catalogue biological interactions to define the underlying biological regulatory network governing interactions between plants and pathogens. We will develop a public web-database, which will be available to any researcher studying pathogen-host responses in citrus. This information will also be useful to study similar pathogens and disease mechanisms in other specialty crops.

*Use as much space as necessary; attach additional pages as needed – budget info. and signatures will appear on the final page.

CRB Project Plan – Research Grant Proposal for FY 2010-2011

Project Title: A generalized reagentless sensor to detect citrus plant and fruit responses

Objectives*: (*succinctly state each objective and milestone, ie the time expected to successfully complete an achieve each objective)*

We propose to build an early pre-symptomatic disease detection system that has two components: (1) the development, optimization and field testing of a novel IVOC detection device, the Differential Mobility Spectrometer (DMS) to enable real time disease detection in orchards; (2) the deployment of a Twister-GC-TOF technology concomitant with an IVOC mass spectrometry database to enable the analysis and interpretation of IVOC profiles across species and stress and disease conditions and (3) the development of a biological regulatory network (BRN) based on transcriptome analysis to interpret and predict citrus pathogen/defense mechanisms, IVOC pathways and to enhance the discover new disease specific biomarkers.

Goal: Develop a framework for the comprehensive disease surveillance and management of high-priority diseases affecting citrus crops, which will include design and field deployment of detection/diagnostic devices and the understanding of the biological regulatory networks of the host responses to pathogen infection.

In support of this goal, we will focus on achieving the following milestones:

(1) Develop and deploy an inexpensive DMS electronic chemical sensor and an IVOC database to track disease at early pre-symptomatic stages via disease-specific volatiles to improve disease diagnosis, detection and scouting.
(2) Develop an IVOC database to more effectively profile IVOCs associated with various disease states using Twister-GC-TOF technology to enhance the identification of disease specific IVOCs.
(3) Develop a biological regulatory network to evaluate the transcriptome profile associated with various disease states to (a) identify new pre-symptomatic disease-specific biomarkers, (b) provide a greater understanding of regulatory pathways that guide plant-pathogen/vector-pathogen interactions.

To accomplish these milestones we propose the following specific aims:

Specific Aim 1: Development and field testing of chemical sensor technologies for early disease detection.
Activity 1. Development of DMS sensor systems for orchard disease sampling for the detection of disease and pest specific volatile biomarkers.
Activity 2. Field and greenhouse testing of DMS sensor systems for asymptomatic disease detection.

Specific Aim 2: Discovery and validation of disease specific biomarkers.
Activity 1. Development of a database to analyze and interpret volatile profiles obtained from field and greenhouse samples using Twister-GC-TOF methodology.
Activity 2. Develop a biological regulatory network to correlate and visualize gene expression data obtained by deep transcriptome sequencing.

Project's Benefit to the Industry*: *(How will the industry utilize your research results or product)*

A hand-held Differential Mobility Spectrometer (DMS) will be a novel diagnostic tool trained to detect specific plant and fruit responses and interpret them using a knowledgebase that integrates plant gene expression and metabolite profiles to derive specific biomarkers for specific problems. An enhanced knowledgebase derived from relationships between genes and metabolites that are regulated in citrus fruit tissues will provide a real-time view of the underlying biology associated with fruit development, injury and disorders. This biology-based understanding provides a foundation for a new tool (DMS) to assist development of chemical, biological, and genetic strategies for

*Use as much space as necessary; attach additional pages as needed – budget info. and signatures will appear on the final page.

control of disease, pests and other fruit-related disorders. Understanding how specific genes respond to citrus fruit development or disorders will allow more precise measurements of internal fruit quality based on the expression pattern of specific genes and metabolites; this will enhance management of citrus fruit in the field and after harvest. Greater precision in fruit quality management will strengthen good agricultural practices (GAPs) and improve citrus fruit production providing greater consistency in different fruit quality attributes. Management of internal fruit quality will also enhance postharvest management and allow better control of fruit quality during production, harvesting, storage, and marketing of citrus fruit. Over the long term, identification of critical genes in the knowledgebase can be linked to natural variation among fruit-specific genes in citrus germplasm leading to development of improved markers to enhance breeding programs. We previously demonstrated the design and application of such DMS platforms for detecting human pathogens. The outcome of the proposed research is to identify biomarkers suitable for reagentless diagnostic systems to monitor citrus trees and fruit in orchards during production and post-harvest storage. The technology has broad applicability to a variety of citrus production problems as explained above and can be used by growers, managers, researchers and regulators in the field to monitor disease, pest and physiological disorder occurrence, prevention, and treatment. Some potential outcomes of this project are:

1 A quantitative approach toward understanding plant and fruit biology and plant-environment and host-pathogen interactions.
2 Identification of unique biomarkers for disease monitoring and prediction.
3 Development of a unique knowledgebase to identify and exploit therapeutic targets for disease prevention and orchard management.
4 Development of inexpensive diagnostics for use in the field by growers, researchers, and regulators to monitor plant health status, detect pathogen infections, and validate the success of therapy and eradication programs.
5 Development of inexpensive diagnostic tools for monitoring the health status of orchards, including nutrient status, environmental stress, and interaction with beneficial organisms in addition to detection of potential pathogens, pests, and threat agents.
6 Development of computational tools to monitor fruit growth and development and to monitor the efficacy of field interventions or treatments.

Research Collaboration* (*be specific*):

Abhaya Dandekar: Plant Sciences Department; UC Davis, CA 95616 – profiling of gene expression and development of knowledgebase to integrate VOCs and gene-based biomarkers.

Cristina Davis: Aeronautical and Mechanical Engineering Department; UC Davis, CA 95616 – Development, training and field testing of DMS device.

Oliver Fiehn: Genome Center, UC Davis, CA 95616 – Metabolite profiling of VOCs using Twister-GC-TOF, biomarker identification and linking diseases to biochemical pathways.

Raissa D'Souza: Aeronautical and Mechanical Engineering Department; UC Davis, CA 95616 – Development network tools to link gene expression data to metabolite profiles and to define unique biomarkers for disease.

Plans and Procedures* *(use this section to describe your experimental design site location(s)and elaborate on objectives and milestones)*

For each of the two specific aims and activities, key persons involved, research design and methods, analysis and interpretation, advantages of the proposed methodologies, limitations and alternative strategies, milestones and safety concerns are described below.

*Use as much space as necessary; attach additional pages as needed – budget info. and signatures will appear on the final page.

Specific Aim 1: Development and field testing of chemical sensor technologies for early disease detection.

Activity 1. Development of DMS sensor systems for orchard disease sampling for the detection of disease/pest specific volatile biomarkers.

Key persons involved: Activity leader Cristina Davis will lead the DMS sensor and IVOC algorithm development. The volatile profile of infected plant samples will be collected for CTV in collaboration with Ray Yokomi as well as working at the Lincove Field station in Exeter (CA) with CTV infected trees. For the detection of CVC detection experiments will be done in collaboration with John Hartung at the citrus exotic disease facility in Beltsville, MD.

Research design and methods: First, we will develop hardware and instrumentation for a field-mobile chemical detection system for volatile plant biomarkers of disease. Second, we will train our sensor system to detect specific volatile compounds, some of which will likely be identified and verified by the Fiehn laboratory using the Twister-GCTOF technology and the BinBase volatile database (Specific Aim 2, Activity 1). We will also develop an IVOC database that will help us detect DMS biomarkers for each disease using the DMS platform as a biomarker discovery platform. It is likely that different subsets of chemical species may become detectable using the DMS instrumentation approach over any given range of sensitivity and that other differences may be detectable with the Twister sampling. By utilizing both methods and approaches for biomarker discovery, we are likely to establish the most robust biomarker library for field testing. Building a reliable IVOC database using GC/DMS will consist of two major steps. The first step is to locate a chemical compound of interest with the DMS spectra outputs, and then generate a spectrum of the compound at different concentrations. This allows us to generate a concentration curve to establish a limit of detection. For a chemical compound that can be detected or resolved from GC profiles alone, the corresponding GC/DMS spectral data can be used for chemical identification and database construction. This is the same process that we follow when performing chemical identification in traditional GC/MS. For a chemical that needs to be resolved or located by both retention time from the GC and compensation voltage from the DMS, which in other words means GC profile alone can not provide a high resolution to resolve this chemical, we will employ a third dimension the RF voltage from the DMS to construct the database. The second step is to establish a speedy database search process, especially required for a large database. The spectrum search can be either based on spectral similarity defined by the dot product of two spectra or upon a pattern recognition process. The traditional method using one-against-all comparisons for a query spectrum will be modified to speed the search process. Through this, we will identify previously undetected HLB, CTV and CVC disease biomarkers.

Analysis and interpretation: We will explore our GC/DMS data sets with several typical machine learning algorithms, including (1) a genetic algorithm (GA) that mimics natural evolution (Cavill et al., 2009) to select the most discernable biomarkers, (2) principal component analysis (PCA) and partial least squares regression (PLSR) to deal with the possible co-linearity among multiple features, reduce data dimension, and provide powerful linear classification tools, and (3) neural network-based approaches to examine the separability of nonlinear classification systems. We will test for detection limits, linearity of the DMS sensor response over several log concentration ranges, and confounding effects from other chemical and biological substances. These models will be statistically validated through independent test sets that were not used to generate the model.

Milestones: First, we will adapt the DMS sensor into a field-mobile system for citrus diagnostics. The overall footprint will be reduced to approximately 2' x 2' x 1' and we will incorporate commercial off-the-shelf (COTS) air sample introduction modules. As we discover chemical IVOC signals that define specific infections, we will train our sensor system to detect those specific compounds for each of the three citrus diseases studied and we will determine the detection limits of each compound and produce specifications for sensor placement in citrus orchard

tree canopies. These results represent the milestones that will be achieved at the end of the 12 months (for HLB), 18 months (CTV), 24 months (CVC).

Advantages: Currently, diseases like HLB are diagnosed by scout teams looking for disease-specific visual symptoms, which often appear too late to prevent secondary spread. Real-time PCR methods are unable to identify the presence of the bacteria in asymptomatic trees due to the phloem-specific nature of the bacteria and limits of detection. The use of DMS will represent a revolutionary tool for growers for early and pre-symptomatic detection avoiding secondary spread of infections and to quickly perform eradication procedures to better contain the spread of the disease.

Limitations and alternatives approaches: The sensitivity of our DMS platform is frequently better than traditional bench-top analytical chemistry equipment. If the number of disease-specific IVOCs identified by Prof. Oliver Fiehn's laboratory is limited or if the GC/MS variability of these compounds is too great to be diagnostically useful, we will seek alternative methods for biomarker discovery by exploring the information rich GC/DMS spectra. We could use our sensor system to determine alternative biomarker chemicals which may have been missed or not identified using the TOF platform. If these data cannot be resolved by simple signal peak analysis, alternative computational strategies are in place to examine those results. For example, we can normalize within our spectra to spiked control chemicals, if needed. Also, we can use techniques such as wavelet Fourier transform analysis to identify features of interest in our spectra.

Safety concerns: Over the course of the program, the GC/DMS instrument will be used to validate volatile organic compounds obtained and shipped via chemical stabilization methods. It is not necessary to ship or move infected material between labs to test our equipment. This greatly minimizes the risk of unintentional transfer of pathogens.

Activity 2. Field and greenhouse testing of DMS sensor systems for asymptomatic disease detection.

Key persons involved: Activity leader Cristina Davis will work closely with some citrus growers in California and Florida. Timothy Spann is an extension specialist at the Citrus Research and Education Center at in the University of Florida.

Research design and methods: It is not enough to test new technology in a laboratory-created "virtual orchard" or even on individual trees held under laboratory conditions. For widespread acceptance in the agriculture community, the platform must be validated in commercial orchards under realistic environmental conditions and disease pressures. Once the detection system is optimized (see milestones of activity 1) we will test the DMS in the field or greenhouse for each disease. Plant VOCs will be collected from HLB symptomatic and asymptomatic infected trees, apparently healthy PCR-negative trees in the same orchard, and healthy trees from another disease-free location in Florida. In the first half of the second year, we will conduct the field testing in the orchard for HLB in Florida. The GC/DMS sensors will be placed in an orchard to test whether they can detect the canopy profile of individual healthy and diseased trees and whether the sensor systems can scan individual tree canopies while on a platform that is stationary and/or in motion and differentiate between healthy and diseased canopy. Also, we will compare volatile profiles of healthy and unhealthy trees, gained from Twister technology, with results from the DMS technology. In the second year we will test the DMS for early detection of CTV and CVC (both in the second 6 months) at the USDA ARS facility in Beltsville. In California, field validation of the effects of pests on the DMS sensor output will be validated on potted plants infested with individual pest species (e.g. Asian Citrus Psyllid, California red scale, citrus red mite, citrus thrips, various aphid species, citricola scale, and/or citrus leafminer) reared at the University of California Kearney Agricultural Center. We will evaluate whether the sensor can distinguish between pest

infested trees, and will describe the background noise that pest populations may create. If either sensor system can detect specific insect pests, we will field-validate the sensor for pests in Florida and California to determine if it can accurately sense an infestation of Asian citrus psyllid.

Analysis and interpretation: We will design and analyze field experiments using well accepted statistical methods. The samples will be run under a splitless injection mode on the GC/MS. The GC/MS data of six replicates for the four categories (symptomatic, asymptomatic, apparently healthy and healthy from disease-free location) will be converted into GC total ion count (TIC) profiles. Six samples from each group will be used to examine the separability of the four groups. Instead of working on the original GC profiles, an auto-regressive (AR) model will be applied to extract the feature of each GC profile. The principal component analysis of the AR coefficients for each sample will examine the separation between the groups and value the efficacy of the sensor for early HLB diagnosis The AR model based feature extraction method significantly reduce the data dimension and solves the possible signal alignment problem in the GC profiles. The visual clustering result of PCA will demonstrate the classification effect of this novel feature extraction method on disease diagnosis. We expect a more accurate and robust diagnosis system by employing the machine learning methods with self-learning ability.

Milestones: The milestones of these activities will be the validation of the ability of the DMS to detect in the field citrus trees infested with the diseases Huanglongbing (*Candidatus Liberobacter asiaticus*), Citrus tristeza virus (CTV), and Citrus variegated chlorosis (CVC) particularly when those trees are asymptomatic. We expect that our statistical analysis will separate the four categories of trees analyzed and in particular will be able to distinguish asymptomatic trees from healthy ones in the same orchard and in the greenhouse.

Advantages: The field testing of the DMS will verify the effectiveness of the sensor to detect the disease early under field conditions and will address future work on the optimization protocol for a more rapid and consistent detection. Therefore, the GC-DMS will represent an invaluable tool for growers to detect diseases before the occurrence of secondary spread in the orchard, and potentially with the first appearance of infected insects based on induction of volatile organic compounds.

Limitations and alternative strategies: Pitfalls of this approach will be represented by the difficulties to fully characterize trees and clearly distinguish between asymptomatic trees and healthy ones. PCR testing will be conducting periodically and healthy trees of a proven disease-free location will be included in the study. The variability in volatile measurements will be reduced by analyzing the different sample categories at the same time during the day and in optimal environmental conditions.

Safety concerns: It is not necessary to ship or move infected material between labs to test our equipment. This greatly minimizes the risk of unintentional transfer of pathogens. SPME fibers used for the field analysis will be carefully handled, stored at 4ºC and the shipping across the U.S. will not pose any identifiable safety concern.

Specific Aim 2: Discovery and validation of disease specific biomarkers.

Activity 1. Development of a database to analyze and interpret volatile profiles obtained from field and greenhouse samples using Twister-GC-TOF technology.

Key persons involved: Co-PI Oliver Fiehn is an expert in chemical analysis of plant metabolism and bioinformatics. Our collaborators Kim Bowman in Florida for HLB, along with John Hartung in Beltsville for CVC and Ray Yokomi in Fresno CA for CTV will expose Twisters to infected plant materials and then mail these back by

overnight mail to the Fiehn lab.
Research design and methods: We will investigate how early such volatile emissions can be detected, which compounds and pathways contribute to pathogen responses, which signals differentiate infections in different species, and how reliably early infections can be identified under orchard conditions. Volatiles are trapped by 1-cm magnetic 'twister' bars that are easily placed at suitable locations within the canopy. Twisters are mailed in simple containers with ice packs and volatiles are released for detection in the lab by thermodesorption. One-hour passive absorption, or potentially only minutes with GCxGC technology, is sufficient for citrus volatile detection. Identification of detected volatile compounds will be achieved by matching mass spectra and retention indices to entries from the Adams volatile library and custom mass spectral databases (Fiehnlab). A public database of identified plant volatiles and novel compounds (Fiehn lab) will allow us to distinguish specific from generic plant responses.

Analysis and interpretation: GC-TOF mass spectra will be housed in the BinBase database, established at the Fiehn laboratory. Volatiles will be identified by commercial and Fiehn Lab libraries. Quantitative data will be analyzed by chemometrics experts (Davis and D'Souza) and statistical software (Statistica DataMiner and TIGR's 'multi-experiment viewer MeV'). Metabolic alterations will be mapped to biochemical pathways and Tanimoto substructure networks.

Milestones: This project will deliver data on volatile profiles from healthy and infected plants at different stages of pathogenesis and from insects carrying pathogen vectors. We will diagnose early, pre-symptomatic infections of citrus diseases, starting with HLB and then CTV and CVC. We will link the volatile emission profiles detected by the Twister-GC-TOF to those obtained with the DMS and to molecular pathways to allow early detection and create a better understanding of pathogenesis in the field. This activity has three milestones, one for each disease: the identification of emitted volatiles specifically induced in early and pre-symptomatic stages of disease. Each milestone will be obtained during a year. During the first year we will focus on HLB In the mean time in the second half of the first year we will start the analysis on CTV that will finish in one year and the entire second year will be also dedicated to CVC.

Advantages: The advantage to use the Twister-GC-TOF technology will be represented by the validation of the IVOC profile detected by the DMS, the discovery of new biomarkers for early detection in the field and the identification of the DMS biomarkers with the development of an IVOC database. We will be able to diagnose early, pre-symptomatic infections of the three citrus diseases. Data obtained by the volatile profiling will be correlated with transcriptomic analysis for the construction of a biological regulatory network that will identify important biomarkers for disease detection and potentially future therapeutic strategies.

Limitations and alternative strategies: The biggest barriers to validating early biomarkers for infection are obtaining access to controlled infection studies and identifying the severity and stage of pathogenesis before symptoms arise. The consortium has experts from several states, ensuring cutting-edge knowledge about prevalent bacteria and viruses. Integration of gene expression and metabolomic data will ensure that pathogenesis stages can be accurately described at the molecular level. Although the technologies have been established in different laboratories of this consortium, integration of data analysis between technologies has not yet been tested. Each technology alone (GCTOF and DMS,) adds to the validation of early pest detection, cross-validating results even if signals cannot be unambiguously annotated as known metabolites.

Safety concerns: The experts collaborating in this project are well aware of the dangers of spreading plant diseases. Volatile trapping and shipping is a safer method of transport than sending actual plant materials. Hence, safety

concern for shipping and analysis of samples will be limited.

Activity 2. Develop a biological regulatory network to correlate and visualize gene expression data obtained by deep transcriptome sequencing.

Key persons involved: Activity leader Abhaya Dandekar's laboratory will build libraries to sequence cDNA and focus on mRNA populations. Field samples for HLB will be provided by Timothy Spann and Kim Bowman. Dawei Lin will manage and disseminate raw sequences from the genome and transcriptome to assemble a gene set for sweet orange. He will perform bioinformatics analysis and create mining tools to investigate the RNA profile. We will collaborate with Norman Schaad for the longitudinal study of HLB and John Hartung for the study of host response to CVC and with Ray Yokomi for CTV. Raissa D'Souza will analyze the structure of networks and interpret their significance.

Research design and methods: Sweet orange plants (cv. Valencia) for this study will be obtained from 4 randomly chosen orchards in different locations in Florida. To obtain better identification of host response biomarkers, both immature and mature/fully expanded leaves have been sampled. Five different biological replicates composed of pooling of leaves from individual trees will be analyzed for each treatment. For all the three citrus diseases, our experimental design includes the following four categories of samples: symptomatic and asymptomatic leaves in infected trees, leaves of apparently healthy trees in same orchard and healthy leaves from healthy trees in disease-free orchard. We will also conduct a longitudinal study of host response to CaLas infection to gain insights into the early molecular events associated with host response at different time points of HLB infection. 10 different trees of sweet orange (cv. Valencia) have been collected and acclimatized in pots at the optimal conditions for the inoculation. Five trees will be infected with CaLas and the other 5 with a bacteria-free solution. Branches of the 5 different trees will be chosen for the three treatments (infected leaves, adjacent not infected leaves and leaves from uninfected trees). Three different time points after infection will be analyzed (3, 15, 21 days after infection) pooling 3-4 leaves. A total of 15 RNA isolations (5 biological replicates x 3 time points) will be performed for each of the three treatments. All RNA samples will be analyzed with Agilent Bioanalyzer and validated with Real time analysis using recently identified genes associated with different stages of disease. The Dandekar lab will create RNA libraries for sequencing on an Solexa Genome Analyzer II following the protocol developed by Illumina. Each run provides 16 GB DNA sequence information. Since transcribed regions are a subset of all genes, actual coverage will be greater. Image files will be transferred to the bioinformatics facility. They will be processed by the pipeline provided by ABI to obtain raw sequence calling and quality scores. These data will be deposited in a Web-based database at the UC Davis Genome Center Bioinformatics Core. For each library, reads will go through quality control to make sure base calling was accurate, then be extracted, clustered, and enumerated. Each sequence is a tag for a particular transcript. Initially, we will use the citrus EST database at NCBI with 449,394 entries as a reference to cluster reads. Raw sequence reads will be further filtered and aligned to reference sequences by a customized pipeline developed by Dawei Lin based on Maq (http://maq.sourceforge.net) and some other in-house developed software tools. Contigs from the assembly will be compared to the NCBI non-redundant database for correspondence to genes that are not annotated in citrus. Policies set by the NCBI GEO (http://www.ncbi.nlm.nih.gov/projects/geo/info/seq.html) will guide our deposition of SOLEXA sequencing data. Once a contig is matched we use the total number of reads associated with that contig as the metric to measure expression of that gene.Differentially expressed genes will be categorized in pathways for the functional analysis using software such as Blast2GO (Conesa et al 2005) and MapMan (Thimm et al., 2004). Results will be validated with Real Time analysis. We will combine our data from transcriptome sequencing and metabolomics with publicly available libraries of signal transduction, metabolomics and regulatory pathways, and protein-protein association databases to identify and integrate all interactions into a coherent network. To visualize and quantify the networks

and identify "hub" proteins, we will use existing software packages, novel approaches from D'Souza's knowledge of network theory, and recent advances for error quantification (Kossinets, 2006; Clauset et. al., 2008). A Web-based cloud server developed by Lin's group will be used to manage and share data among collaborating labs. A platform-independent software system, "Gaggle" (Shannon et al 2006), will be used to seamlessly integrate various bioinformatics software and databases.

Analysis and interpretation: Pathogen-induced mRNA will be validated by Taqman Real Time PCR. The specificity of the mRNA will be analyzed by comparison to other samples challenged with different pathogens. Dawei Lin will provide expertise in integration of large, complex datasets. Existing software packages, together with methods developed in D'Souza's lab (Achlioptas et al 2009, D'Souza and Roy 2008), will be used to analyze structural differences networks and their biological significance will be interpreted via close collaboration with biologists (e.g., Dandekar and lab group).

Milestones: The first milestone will be the construction of a biological regulatory network that will identify important points of regulation and deliver validated biomarkers for early disease recognition, detected by handheld devices like the DMS. Also biomarker for screening tests in facility machines (i.e real time PCR) could allow early diagnosis of citrus infections, potentially decreasing the need for eradication programs. The primary application of the product would be rapid, onsite screening for the presence of a panel of pathogens or host biomarkers of concern. Messenger RNAs that are induced in a rapid and specific manner are potential early diagnosis biomarkers and useful for pre-symptomatic disease diagnosis for eradication strategies. Another milestone will be the development of a web-knowledgebase that will be publicly available for the three pathogen species at various disease stages. Ultimately, comparing interaction networks of healthy and diseased plants of the same species will identify key proteins to target for future therapeutic strategies.

Advantages: We will build a knowledgebase of the biological regulatory network for each disease under investigation at various stages of disease development in both host and pathogens. These will aid early disease detection and help develop therapeutic strategies that can rectify a deficient network. The study of their interactions will clarify the nature of three diseases and facilitate their management.

Limitations and alternative strategies: A limitation of the annotation assignments of the differentially regulated sequences will be represented by the incomplete availability of the sequence of the entire citrus genome. However a high number of ESTs haved been deposited in public databases and 15808 unigenes are at the moment present for *Citrus sinensis* in NCBI database. Another big challenge will be the determination of the functions of the mRNA differentially expressed between healthy and unhealthy plant material for each specific disease. A major emphasis of the work will be focused on sequence comparisons and literature search through existing databases of model plants (Arabidopsis in particular).

Safety concerns: None only those associated with the use of molecular biology techniques and recombinant DNA technologies.

Cited Literature:

Achlioptas D, D'Souza RM, Spencer J. 2009. Explosive Percolation in Random Networks. Science. 323 (5920): 1453 – 1455.

Cavill R, Hector CK, Holmes H, Lindon JC, Nicholson JK, and Ebbels TMD. 2008. Genetic algorithms for

simultaneous variable and sample selection in metabonomics. Bioinformatics. 25 (1): 112-118.

Clauset A, Moor C, Newman MEJ. 2008. Hierarchical structure and the prediction of missing links in networks. Nature. 453: 98-101.

Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, et al. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21: 3674-3676.

D'Souza RM, Roy S. 2006. Network Growth with Feedback. Physical Review. E, 78.

Kossinets G. 2006. Effects of missing data in social networks. Social Networks. 28(3): 247-268.

Shannon PT, Reiss DJ, Bonneau R, Baliga NS. 2006. "The Gaggle: An open-source software system for integrating bioinformatics software and data sources", BMC Bioinformatics. 7: 176.

Thimm O, Blaesing O, Gibon Y, Nagel A, Meyer S, Krüger P, Selbig J, Müller LA, Rhee SY, Stitt M. 2004. MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. Plant Journal. 37 (6): 914-39.

't Hoen PA, Ariyurek Y, Thygesen H, Vreugdenhil E, Vossen RHAM, de Menezes RX, Boer JM, van Ommen G-JB, den Dunnen JT. 2008. Deep sequencing-based expression analysis shows major advances in robustness, resolution and inter-lab portability over five microarray platforms. Nucleic Acid Research. 36: e141.

Other Funding Sources for this Project *(current, pending, potential; can this project be used as matching funds for other funding sources)*

This project has obtained a match for one year with UC Discovery. A project #itl07-10167, entitled, "Development and testing of a generalized reagentless chemical sensor for the real-time detection citrus plant and fruit responses" was approved for the project period 05/01/08 to 04/30/09 for $120,000 providing a 1:1 match to CRB funding. We were successful in obtaining a second round of funding by UC Discovery grant number BIO09R-156555 for 2 years (11/01/2009-10/31/2011). UC Discovery has matched 1:0.7 or up to a max of 70% of the CRB funds that we received 2009-2010 ($180,000). This proposal was submitted to UC Discovery July 14, 2009 and got a start date of Nov 1, 2009.

Technology Transfer* *(include any potential intellectual property issues; steps necessary for grower utilization extension/communication component):*

Prof. Cristina Davis works very closely with a small venture capital-based start-up company, Sionex Corporation (Bedford, MA), that is currently commercializing the DMS technology. Sionex closed their C round of venture financing on 12/07/2005, and they have a very successful small-business strategy for producing chemical detection systems. On 04/24/2006, they established a strategic partnership and financial relationship with In-Q-Tel, the venture funding arm of the US Central Intelligence Agency (CIA). As a strategic investor, In-Q-Tel mixes product development funding with equity investments to accelerate development of technologies of value to the US intelligence community. To date, the company has moved three products to market production: (1) The "Egis™ Defender", a briefcase-sized port-of-entry screening unit used for explosives and chemical agent detection at commercial airports (licensed and produced by ThermoElectron, Inc.); (2) The "CP4900 DMD MicroGC" a portable

sensing unit to detect sulfur compounds for oil industry applications (licensed and produced by Varian, Inc.); and (3) OEM sub-systems for trace chemical detection (produced by Sionex).

List of patents generated by Cristina Davis and her colleagues related to sensor development:

1. Miller RA, Nazarov EG, Zapata AM, **Davis CE**, Eiceman GA, Bashall AD. Systems for differential ion mobility analysis. Issued 06/06/2006. US Patent #7,057,168.
2. Miller RA, Nazarov EG, Zapata AM, **Davis CE**, Eiceman GA, Bashall AD. Systems for differential ion mobility analysis. (Filed with USPTO on 03/10/2005) US Patent Application # 20050051719.
3. **Davis CE***, Borenstein JT, Zapata AM, Gelfand JA, Callahan MV, Stair TO, Miller RA. Non-invasive breath analysis using field asymmetric ion mobility spectrometry. (Filed with USPTO on 04/21/2005) US Patent Application # 20050085740.
4. **Davis CE***, Tingley RD, Krebs MD. Alignment and autoregressive modeling of analytical sensor data from complex chemical mixtures. (Filed with USPTO on 01/26/2006) US Patent Application #20060020401.
5. Miller RA, Nazarov EG, Zapata AM, **Davis CE**, Eiceman GA, Bashall AD. Systems for differential ion mobility analysis. (Filed with USPTO on 07/13/2006) US Patent Application #20060151687.
6. Zapata AM, Kim ES, Agrawal P, Krebs MD, **Davis CE***. Apparatus and systems for processing samples for analysis via ion mobility spectrometry. (Filed with USPTO on 07/27/2006) US Patent Application #20060163471.
7. Merrick WF, Zeskind JE, Krebs MD, **Davis CE***. Monitoring drinking water quality using differential mobility spectrometry. (Filed with USPTO on 05/04/2006) US Patent Application #11/417,897.
8. **Davis CE***, Borenstein JT, Zapata AM, Gelfand JA, Callahan MV, Stair TO, Miller RA. Non-invasive breath analysis using field asymmetric ion mobility spectrometry. (Filed with EU on 10/21/2004) PCT Patent Application #EP1627223 (World Patent Reference #WO 2004/090534 A1).
9. Miller RA, Nazarov EG, Zapata AM, **Davis CE**, Eiceman GA, Bashall AD. Systems for differential ion mobility analysis. (Filed with EU on 09/23/2004) PCT Patent Application #EP1601948 (World Patent Reference #WO 2004/081527 A2).

Our university-based team is in the uniquely suited to perform the cutting-edge research and development necessary to adapt this powerful chemical sensor for use in agricultural applications. Together with industry collaborators, we will then be able to transition this technology into production resulting in an actual commercial product that will be available for citrus growers to purchase.

Budget Justification:

Personnel.
Abhaya M. Dandekar, Ph.D. (PI)
Salary: $0; Benefit Rate 26%: $0 (calendar year salary, 11 month)
8% effort (1.0 month, calendar year)
Prof Dandekar will coordinate the overall effort; interact more closely with CRB and orchard systems to obtain biological sampling for all of the interacting labs. His lab will lead the effort of sampling the transcriptome and work with D'Souza to build a biological regulatory network working with the Davis and Fiehn labs.

Cristina Davis, Ph.D. (Co-PI)
Base Salary: $9,544; Benefit Rate 16.7%: $1,594 (calendar year salary, 9-month academic year)
8% effort (1.0 month, calendar year)

Prof. Davis will lead the efforts relating to the DMS detection, integration of technology with the IVOC database developed by Prof. Fiehn's group and the transcriptional library developed by Prof. Dandekar's group.

Oliver Fiehn, Ph.D. (Co-PI)
Salary: $14,167; Benefit Rate 26%: $3,683 (calendar year salary, 9-month academic year)
8% effort (1.0 month summer salary, calendar year)
Prof Fiehn will evaluate data, supervise the postdoctoral scientist, design studies, interact with PI Dandekar and co-PIs, Davis and D'Souza, visit collaborators in California and Florida, write manuscripts and disseminate results to growers and the scientific community.

Raissa D'Souza, Ph.D. (Co-PI)
Base Salary: $9,556; Benefit Rate 24%: $1,214 (calendar year salary, 9-month academic year)
8% effort (1.0 month, calendar year)
Prof. D'Souza will lead the efforts to the creation of a biological regulatory network based on the analysis of the transcriptome and work closely with the Dandekar, Fiehn and Davis labs on the integration of diverse datasets.

Weixiang Zhao, Ph.D. (Staff Research Scientist)
Base Salary $55,464, Benefit rate 34%; $18,858
75 % effort (9.0 months, calendar year)
Dr. Zhao will analyze all of the DMS data sets and provide novel quantitative methods for evaluating IVOCs with the sensor system. He will also work with Prof. Fiehn's group to coordinate our data base development with their discovery methods, and to choose the "best" and most robust biomarkers for subsequent system use.

Abhinav Bhushan, Ph.D. (Postdoctoral Fellow)
Base Salary $41,500, Benefit rate 34%; $14,110
75 % effort (9.0 months, calendar year)
Dr. Bhushan will work to develop the IVOC database with the DMS system and to prepare the instrumentation for portable unit testing in field conditions.

Kirsten Skogerson. (Graduate Student)
Base Salary: $35,000; Benefit Rate 26%: $9,100
100 % effort (12.0 months, calendar year)
A full time graduate student in the Fiehn lab performs the bulk of the work, from method development in Twister GCxGC-TOF mass spectrometry to compound identification by commercial libraries, reference compounds and the Adams library to interact with database programmers and the UCD Genome Center bioinformatics core staff. The student will further interact with scientists in the PI's and collaborator laboratories as well as growers and researchers in other states.

Federico Martinelli, Ph.D. (Postdoctoral Fellow)
Base Salary: $22,573; Benefit Rate 21%: $4,704
50 % effort (6.0 months, calendar year)
Dr Martinelli will work in the Dandekar lab and be responsible for harvesting the biological samples for the transcriptome analysis. He will also help coordinate the biological samples for the Twister and DMS. Dr Martinelli will extract RNA construct libraries for the DNA sequencing of the transcriptome profile. He will work closely D'Souza to build the biological regulatory network and to be involved in the interpretation and literature analysis.

Russell Reagan, MS. (Assistant Specialist)
Base Salary: $10,748; Benefit Rate 51%: $5,310
25 % effort (3.0 months, calendar year)
Russell Reagan will work in the Dandekar lab and be responsible for cDNA analysis pipeline. He will work closely with the bioinformatics group in the Genome Center and assist Dr Martinelli developing a robust transcriptome profile. He will also participate in the construction of the BRN network and be responsible for data visualization and annotation.

Supplies, Materials and Services.
A total of $67,676 is budgeted for supplies in the first year and $60,326 in year 2. The Fiehn lab will use about $5,050 in year 1 and $2,691 in year 2 from the supply budget to conduct their studies, with major items being fractional costs for computational hardware ($1,500 p.a.), fractional costs for mass spectrometer maintenance ($1,500 p.a.) and bioinformatics support ($1,500 p.a.). These fractional costs are allocated to each project in the Fiehn laboratory based on per-sample information stored in the SetupX study design database and are estimated based on the experiences in the first year of UC Discovery support for citrus volatile investigations. Further consumable costs cover GC-TOF and Twister materials, mainly syringes, liners, Twister bars, O-rings, vials, caps, columns, liquid nitrogen, Helium gas, solvents, vial inserts and chemical standards. The Davis lab will use a total of $37,357 (year 1) and $33,978 (year 2) of the supply budget to perform the research outlined in this proposal. The proposed expenses include a 1-year instrumentation maintenance service contract for GC/MS ($9,640 per year), *based on quote #29729708133 from Varian, and their current 2009-2010 agreement and PO. Consumables for headspace VOC testing ($800/month, $9,600 per year). SPME fibers and holders, borosilicate vials, Teflon septa, graphite ferrules, gas tight syringes, etc. Compressed He gas tanks and liquid N2 dewars for analysis ($500/month, $6,000 per year). General laboratory consumables ($100/month, $1,200 per year) and small parts and mechanical connection supplies for machine build ($2,882). The remainder of the supplies budget for $25,269 in year 1 and $23,657 in year 2 will be used for the transcriptome analysis conducted by the Dandekar lab. This includes supplies for the isolation of RNA, purification of mRNA and construction of cDNA libraries ($3,500 per year). The transcriptome profile will be determined by deep sequencing using a SOLEXA at the Genome Center. We expect to run 4 to 8 lanes of a 120 bp paired end run which cost $2,800 per lane for a total cost around $11,200 to $22,400. For assembly we will use the bioinformatic services at the Genome Center and we are budgeting $3,850 for this service to process the almost 10-20GB of DNA sequence data. In addition there are lab consumables and supplies to run real time PCR validation for specific genes and pathways.

Travel.
A total of $10,000 in year 1 and $10,500 in year 2 is budgeted, of this $3,000 per yr is for the Davis lab, $1,000 year 1 and $1,500 year 2 for the Fiehn lab and $6,000 per yr for the Dandekar lab. About 60% will be used to sample collection and processing in Florida and the remainder for the researchers to attend conferences and events sponsored by the California Citrus Research Board (CRB).

CRB Project Plan – Research Grant Proposal for FY2010-2011
Project Title: A generalized reagentless sensor to detect citrus plant and fruit responses

Project Budget

Department Account Number: *(if applicable)*

	Year: **2010-2011**	Year: **2011-2012**	Year: **2012-2013**
Salaries and Benefits:			
Postdocs/Research Assistants	$112,778.50		
SRA's	$		
Lab/Field Assistance			
Benefits	$ 34,912.50		
Supplies and Expenses:	$ 27,060.00		
Equipment:	$		
Operating Expenses and Travel:	$ 5,250.00		
Lindcove Recharges:			
Lindcove Packline:			
Other:	$		
ANNUAL TOTAL:	$180,000.00		

*****No funds are requested only a 1 year "no cost extension" of existing CRB funds to maintain the match with our current UC Discovery grant

Specifics regarding contract (i.e., "split" funding to more than one PI):

Signatures

Project Leader: **Date:** July 31, 2010

Date:

Date:

Dept. Chair: **Date:**

*Use as much space as necessary; attach additional pages as needed – budget info. and signatures will appear on the final page.