GENOME EDITING OF *TAS4, MIR828* AND TARGETS *MYBA6/A7*: A CRITICAL TEST OF *XYLELLA FASTIDIOSA* INFECTION AND SPREADING MECHANISMS IN PIERCE'S DISEASE

Principal Investigator: Chris Rock Texas Tech University Dept of Biological Sciences Lubbock, TX 79409-3131 chris.rock@ttu.edu

Research Associate:

Sunitha Sukumaran Texas Tech University Dept of Biological Sciences Lubbock, TX 79409-3131 sunitha.sukumaran@ttu.edu

Co-Principal Investigator:

Leo De La Fuente Auburn University Dept Entomology Plant Pathology, Auburn, AL 36849 Izd0005@auburn.edu

Research Associate:

Sy Mamadou Traore Auburn University Dept Entomology Plant Pathology, Auburn, AL 36849 sytraore@gmail.com

Collaborator: David Tricoli University of California, Davis Plant Transformation Facility Davis, CA 95616 dmtricoli@ucdavis.edu

Graduate Research Asst:

Fakhrul Azad Texas Tech University Dept of Biological Sciences Lubbock, TX 79409-3131 fakhrul.azad@ttu.edu

Graduate Research Asst:

Sayani Mallick Texas Tech University Dept of Biological Sciences Lubbock, TX 79409-3131 sayani.mallick@ttu.edu

Reporting Period: The results reported here are for work conducted from March, 2016- July 26, 2016

Introduction

A renewal application for 2016 funding was provisionally approved on March 30, 2016, and executed on July 26, 2016.

We hypothesize that novel target MYB transcription factors (VvMYBA6/A7) in grape are effectors of anthocyanin accumulation and potentially glass winged sharpshooter (GWSS) feeding preference determinants important for PD etiology. The model postulates microRNA828 and evolutionarily-related *Trans-Acting Small-interfering locus4* (*TAS4*) activities silence target *VvMYBA6/A7* and other homologous *MYB*s expression in response to XF infection, mediated through inorganic phosphate (P_i) and plant stress hormone abscisic acid (ABA) signaling crosstalk. Anthocyanin induction in vegetative tissues may serve as antagonists to feeding by GWSS and to colonization by XF. We are currently testing the XF infection/spread hypothesis directly by "knocking out" the key genes using a new genome editing technology- Clustered Regularly Interspaced Short Palindromic Repeats (CRIPSR/Cas9)^{2, 3} that the CDFA-PD Board nominated as a feasible, high-priority approach to engineering PD resistance.

We have added a new, value-added independent approach component: *Nicotiana benthamiana* transient assays⁴ for CRISPR/Cas9 activities, going forward that complements and leverages the previously described surrogate tobacco XF infection system developed by the Co-I (De La Fuente)⁵ to quickly assess susceptibility to XF infection of a transgenic tobacco line⁶ (Myb237) that over-expresses the Arabidopsis homolog of VvMYBA6/A7: *PRODUCTION OF ANTHOCYANIN PIGMENT2/MYB90*. We elaborate here our preliminary results from this facile method that demonstrates Cas9 effector protein is expressed well in plants from our backbone vector. Applying this transient assay opens the possibility to conduct complimentary experiments using CRISPR/Cas9 editing technology to target endogenous *MIR828*, *TAS4*, and target *MYB* loci in tobacco, including

stable transformants for rapid and independent tests of the working model. The PI conceived of this approach in June 2016, after the previous progress report submission. Deployment of this transient assay experimental system accelerates one facet of the project; namely a proof in principle that Objective I will be successful.

OBJECTIVES (as funded)

- I. Test the miR828, *TAS4*, and target *MYBA6*/7 functions in PD etiology and XF infection and spreading by genome editing using CRISPR/Ca9 transgenic technology.
- II. Characterize tissue-specific expression patterns of *TAS4* and *MIR828* primary transcripts, sRNAs, and *MYB* targets in response to XF infections in the field.

Description of activities conducted

I. Test the miR828, *TAS4*, and target MYBA6/7 functions in PD etiology and XF infection and spreading by genome editing using CRISPR/Cas9 transgenic technology

The PI's lab and greenhouse was certified for Biosafety Level II work with XF by the USDA-APHIS (permit# P526-160120-034 issued 04/06/16) and by his Institutional Biosafety Committee on June 1, 2016. This clears the way for the PI to assume sole responsibility for the XF work going forward, as proposed in the 2016 renewal application, and for the Co-I De La Fuente to change his status in the project to Cooperator upon completion of the ongoing experiments in October 2016 and transfer of associated data to the PI. The Co-I shipped XF Temecula-1 and WM-1 strains to the PI on May 4, 2016 and will send them again as a backup contingency as soon as possible.

Engineered binary T-DNA *Agrobacterium* vectors designed to genome edit the grapevine *VvMIR828*, *VvTAS4ab*, and target *VvMYBA6* /*VvMYBA7*, and *Phytoene Desaturase* (*PDS*) loci (the latter as an independent test of editing efficiency) were sent to the Collaborator David Tricoli's lab under APHIS BRS permit # 15-231-102m in Nov., 2015 and were described in the first and second Progress Reports. Dr. Tricoli reported to the PI in May 2016 that the regenerants from Agrobacterium co-cultivation of these constructs were developing unusually slowly, despite setting up two independent transformation runs for each construct in November and December 2015. **Figure 1** shows a representative progress of regeneration from somatic embryos; evidence of photobleached sectors for *PDS* constructs, or any purple anthocyanin-pigmented sectors for candidate editing events of *TAS4* and *MIR828*, which would be de facto evidence for high-efficiency production of bi-allelic target

deletions, is still lacking due to slow regeneration. Transformants of 101-14 rootstock appear more advanced than Thompson Seedless, which was transformed with the *PDS* constructs because of its ease and rapidity of regneration. The reason for the slow regeneration times is currently unknown- dozens of grape transformations in the Cooperator's pipeline are responding as expected so we do not suspect a media problem etc. Dr. Tricoli set up a backup transformation for all constructs in late May, 2016 and is currently transforming tobacco with the subject vectors to troubleshoot the grape regeneration problem. It is too early to ascertain the status of the newer materials. Carrying out this repeat transformation of grapevine could not be done sooner because of a lack



Fig. 1. Regeneration status of a representative project construct VvMYBA7.1s (right) compared to a unrelated construct (left), July 2016.

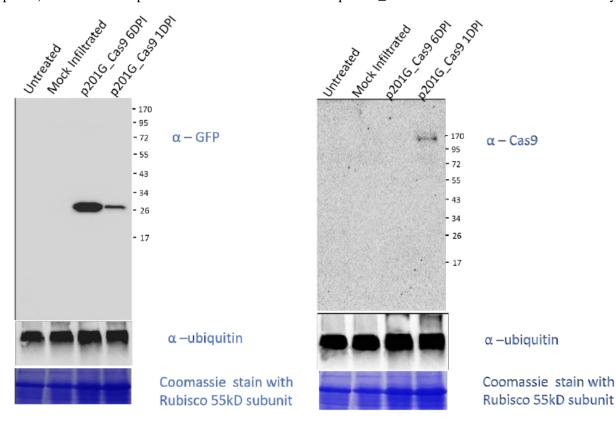
of sufficient numbers of stock grapevine embryogenic culture starting materials, derived from anthers of immature flowers harvested in the spring. It can take eight or more months to generate enough embryogenic culture materials for transformations.

Validation of editing events going forward will be by PCR cloning and sequencing of target genes, and PAGE-based genotyping⁷. We are in the process of setting up mock editing assays, in order to be ready for genotyping the bona fide grapevine samples, by using a 15 nt deletion of the phytochrome *PHYD-1* gene of Arabidopsis ecotype Wassilewskija (Ws-1) to 'dope' with different tracer amounts of genomic DNA from Ws-1 the bulk gDNA (with wild type *PHYD* allele) from control Ws-0 extracts. This allows us to create a 'needle in a haystack' mock experiment for optimizing the genotyping assays and determining the limits of detection for editing events (and thus editing activities/efficiencies) using the *PHYD-1* deletion allele in pilot experiments.

As an independent, partial test of the hypothesis, we have initiated work on *Nicotiana benthamiana*, which is a facile system for high level protein expression in plants⁸ by infiltration of Agrobacterium harboring T-

Table I. Synthetic guide sequences currently being assayed for transient CRISPR-Cas9 editing of Nb-MIR828 and Nb-						
TAS4a-b genes. Off targets candidates computed at http://cbi.hzau.edu.cn/cgi-bin/CRISPR						
Gene.test	Engineered guide sequence	Relative genome position	Off targets, seed (12)NGG?			
Nb-MIR828.1a	GGAATACTCATTTGAGCAAGAGG	Mature miRNA, antisense	4			
Nb-TAS4a.1a*	GAA <u>GGTCC</u> GAGGTTGAGGTTGG	D4 phase, antisense	17			
* restriction site AvaII for mutant screening by Cleaved, Amplified Polymorphism (CAP)-PCR.						

DNA vectors engineered from the same starting backbone vector p201N_Cas9. **Table I** lists synthetic guide sequences being engineered as T-DNA vectors using starting materials² described in Progress Report 2. **Figure 2** shows an immunoblot result demonstrating humanized (codon optimized for translation and shown previously to work in plants) Cas9 is well-expressed from the vector backbone p201G_Cas9 in *N. benthamiana* transient assay



Α

Coomassie stained gel images are also shown as loading control.

Fig. 2. Immunoblot of protein extracts harvested at 1 or 6 days post infection (DPI) from transiently transformed *N. benthamiana* leaves co-cultivated with Agrobacterium strain GV2260 harboring CRISPR/Cas9 vector p201G_Cas9 expressing green fluorescent protein (GFP) and humanized SpCas9 under control of the 2x35S plant promoter. A) Blot probed with anti-GFP monoclonal antibody (Invitrogen), showing GFP accumulation over time, as previously observed. B) Blot probed with anti-Cas9 (Novusbio), showing the Cas9 protein is subject to turnover after 1 DPI but is nonetheless expressed. Anti-Ubiquitin (PD41, Abcam) antibody was used to re-probe blots to validate equal loading of proteins, and the

В

after on day one post infection (DPI). In the future we will express vectors targeting endogenous *N. bethamiana* $MIR828^9$ and $TAS4ab^1$ genes (Table I) to easily assay editing efficiencies of our adopted vector system *in planta*. The same constructs could be used to stably transform tobacco for future XF challenge experiments that would complement the ongoing tobacco transgenic experiments with Myb237 Hmo and Hmi lines.

II. Characterize tissue-specific expression patterns of *TAS4* and *MIR828* primary transcripts, sRNAs, and *MYB* targets in response to XF infections in the field.

In the previous progress report we characterized and correlated molecular phenotypes of XF titres, TAS4-3'D4(-) small RNA abundances by RNA blot estimation, and anthocyanin quantities extracted from the transgenic tobacco line Myb237 overexpressing AtMYB90 challenged with XF in the greenhouse, and from PD-infected and

symptomless Merlot leaves and petioles collected from the 'Calle Contento' vineyard in Temecula CA, and the Black Stock vineyard in Dahlonega, Lumpkin Co., GA. Those compelling results were consistent with our working model of XF interaction with anthocyanin biosynthesis regulation by the host during PD progression and showed significant differences in accumulation of anthocyanins in XFinfected vs. control leaves from the field and greenhouse samples. Furthermore, the results with the homozygous AtMYB90 overexpressing transgenic tobacco line (Hmo) showed significantly greater (57% of leaves) disease symptom development five weeks after XF challenge than either non-transgenic (SR1) or hemizygous transgenic (Hmi) genotypes (22-27% of leaves), which was inversely correlated with anthocyanin accumulation in leaves of the transgenic Hmo and Hmi genotypes (more disease~ less anthocyanins, with similar titres of XF found across the experiment). Below we present characterization of Illumina small RNA libraries generated from the same tobacco- and California PD-infected and control samples, sequenced by the UC-Riverside Institute for Integrative Genome Biology (fee for service). The results of Figures 3 and 4 are strong evidence that XF infection triggers up regulation of TAS4 siRNAs, supporting our working model. We now have data (not shown) supporting the claim that XF induces TAS4 siRNAs in non-transgenic SR1 control tobacco, which is compelling evidence supporting our model. In the previous progress report, we showed small RNA blot evidence that XF

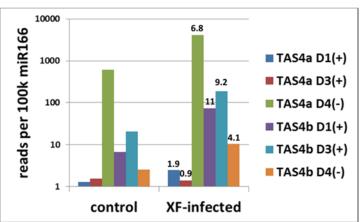


Fig. 3. *VvTAS4a* and *VvTAS4b* siRNAs are elevated by XF infection in field samples from Temecula, CA. Numbers above bar: –fold effect by XF on specific species of *TAS4* tasiRNA accumulation, normalized to the most abundant miRNA in plants, miR166. *TAS4a D4(-)* (green bar) is the predominant active trigger of *VvMYBA6/A7* phased siRNA production¹.

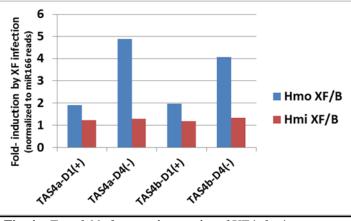


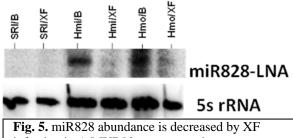
Fig. 4. >Two-fold **phenotypic severity** of XF infection symptoms previously shown for homozgous AtMYB90-OX tobacco line 237Hmo (vs 237Hmi or SR1 control) symptoms **correlates with AtMYB90 negative effector** *TAS4ab* **3'D4(-) abundance.**

infection caused an increase in TAS4-3'D4(-) abundance in Hmo genotype, which correlated with disease symptom severity compared with Hmi or SR1 infected genotypes, a result independently verified conclusively by deep sequencing results shown in Fig. 4. **Figure 5** shows a preliminary result suggesting that miR828 is upregulated in MYB90 OX genotypes and down regulated by XF infection, which is evidence that an autoregulatory

loop operating in tobacco as described for Arabidopsis TAS43'-D4(-) and PAP1/MYB75¹⁰, which we have also identified in other species (see presentation citation below). Thus, our working hypothesis is that an analogous autoregulatory loop also operates on the *Nt-MIR828* locus when AtMYB90 is overexpressed in tobacco. Velten et al (2012)⁶ also observed that miR828 was strongly elevated in sectors of tissue from an independent transgenic tobacco line (Myb27) homozygous for MYB90 when it undergoes spontaneous transgene silencing. These data taken together are evidence consistent with our interpretation that MYB90 overexpression acts on endogenous autoregulatory loops for *TAS4* (**Fig. 4**) and *MIR828* (**Fig. 5**) that we hypothesize antagonize each other to maintain pathway homeostasis and which XF targets for mis-regulation by unknown mechanisms.

Table II shows the parameters of the CA grapevine field and tobacco greenhouse small RNA libraries.

The tobacco Myb237 XF challenge experiment is currently being repeated by the De La Fuente lab and is scheduled for completion in a few weeks, with data analyses completed in Oct. 2016. We have found additional compelling evidence in the literature supporting our phosphate-regulation XF etiology model: in Arabidopsis infected with XF genomewide transcriptome analysis showed *TAS4* siRNA target *MYB PRODUCTION OF ANTHOCYANIN PIGMENT1/MYB75* and another phosphate-regulated locus, At5g20150/SPX DOMAIN which is a positive regulator of cellular responses to phosphate starvation, are both strongly down regulated by XF infection¹¹.



infection in AtMYB90 overexpressing genotypes, consistent with a negative autoregulatory loop between AtMYB90 and endogenous Nt-MIR828 that is affected by XF infection.

Furthermore, SPX1 messenger RNA is mobile in the vasculature¹², which is relevant to XF growth habitat. These serendipitous findings constitute 'smoking guns' supporting our working model and warrant further study.

Table II. Parameters of sequenced libraries constructed from tobacco greenhouse studies of XF-infected genotypes, and Merlot Vitis samples from Temecula CA 'Calle Contento' vineyard characterized by ShortStack^{13, 14} miRNA discovery software

Sample library^	Number of trimmed reads (millions) >17nt mapping to reference genome*	miR166 lib total reads%	MIRNA loci empirically by ShortStack	Ratio <i>MIRNAs</i> found/10 ⁶ reads
Tobacco SR1-B	8.94	3.3%	31	3.5
Tobacco SR1-XF	8.47	2.8%	46	5.4
Hmo MYB90 OX-B	15.10	1.0%	18	1.2
Hmo MYB90 OX-XF	7.30	0.2%	7	0.9
Hmi MYB90 OX-B	9.18	1.1%	12	1.3
Hmi MYB90 OX-XF	7.69	1.0%	17	2.2
Vitis-C from field	1.66	11.0%	13	7.8
Vitis-XF from field	3.77	4.1%	22	5.8

 $^{\text{B}}$ B = Buffer mock infection; XF= Xylella infected, validated > 10⁸ cfu/gfw by qPCR; C= control field sample, validated XF titre < 10³ cfu/gfw petiole.

* for Vitis, NCBI RefSeq GCF_000003745.3_12X, updated 6/14/16. For tobacco, Burley TN90 draft assembly of 256k contigs (http://solgenomics.org)

We have purchased an Illumina Trueseq stranded mRNASeq kit for mRNA-Seq (Obj. II, Method 2) and library construction from grapevine samples is currently underway which will permit digital measurement of primary transcripts including *MIR828*, *TAS4* ncRNAs, and *MYB* targets as well as all other differentially expressed genes deranged by XF in grapevine. This will allow a systems approach to discover other etiological effectors/reporters of PD and network analyses of gene interactions affecting primary and secondary metabolism in the process.

Publications

Mallick S. (July, 2016) Characterization of Arabidopsis Pyrabactin-Like ABA Receptor (PYL4 and PYL7) and transcription factor (RAV and ABI5) activities in transiently transformed *Nicotiana benthamiana* and stable transgenic lines of cotton (*Gossypium hirsutum*). M.Sc. thesis: Texas Tech University. https://ttu-etd.tdl.org/advisor/75IO0sPPh8s/review.

Publications in preparation

Ibrahim RK, Rock CD. "A22246: Phenylpropanoid metabolism (version 2.0)" *Encyclopedia of Life Science:* eLS. In prep. Deadline for manuscript submission: September 2016.

Sukumaran S, Tricoli D, Rock CD. "Efficacy of CRISPR/Cas9 in grapevine based on select guide sequences targeting *Phytoene Desaturase*." *Computational and Structural Biotechnology Journal*, in prep. Research article solicited by Editor Gianni Panagiotou. Deadline for manuscript submission: September, 2016

Sukumaran S, Traore SM, Azad F, De La Fuente L, Rock CD. "small RNA profiles in grapevine variety Merlot infected with *Xylella fastidiosa*." In prep. Solicited research article for special edition of *Frontiers in Plant Science*: 'Omics and systems approaches in grapevine fruit composition to understand responses to environmental factors and agronomical practices.` Eds: José Tomás Matus, Simone Diego Castellarin, and Giovanni Battista Tornielli. Deadline for manuscript submission: October, 2016.

Presentations

Invited seminar by C Rock: Dept of Genetics, Botucatu Institute of Biosciences, Sao Paulo State University-Botucatu, Brazil. Nov. 4, 2015. "Plant Polyphenolics, small RNAs, and Darwin's 'Abominable Mystery.""

Sukumaran S, Traore S, Azad Md.F, De La Fuente L, Rock C. "Conservation of an autoregulatory feedback loop regulating anthocyanin biosynthesis in dicots." Plant Biology 2016: Annual Meeting of the American Society of Plant Biologists. July 9-13, 2016. Austin, TX. poster #1000-063.

Research relevance statement

The general research objective (within the scope of Years 1-2 seed funding) is to test the hypothesis that specific *trans*-acting small interfering RNAs (*ta*-siRNAs) produced by grape are regulators of the Pierce's Disease process. The long-term goal is to establish a new technology in grapes that will allow genetic manipulations that will not carry the negative connotation of "GMO." This is because the vector transgenes can be removed by conventional backcrosses to the transgenic lines, or by editing out (transiently) the effector transgenes, resulting in only mutated endogenous effector genes in progeny and vegetative regenerants. These proof-in-principle experiments could result in a new paradigm for host-vector-pathogen interactions in PD for the advancement of the grapevine biotechnology and breeding sectors.

Layperson summary of project accomplishments

We are on track to achieve our Objectives. In future applications to PD/GWSS, contingent upon satisfactory progress towards Objectives 1 and 2, we will characterize the changes in control versus edited genotypes for xylem inorganic phosphate (P_i) and other macronutrients, and polyphenolic levels of XF-infected stems. We will conduct XF challenge experiments with genome-edited transgenic plants. We have within scope to conduct insect diet preference and XF growth assays with candidate polyphenolics that arise from our results. It is noted that no host genes are yet known that normally function to *enhance* host susceptibility; altering host gene (e.g. PD resistance) activities may result in increased susceptibility to infections. Thus engineering PD resistance is likely to be by incremental advances from characterizing molecular mechanisms.

Status of funds

Salaries are encumbered through Aug. 2016 and remaining funds including fringe will be completely spent by Sept 1, 2016. Two graduate students (Sayani Mallick, Md. Fakhrul Azad) were supported for the summer semester from June 1- July 15th, 2016. All travel funds are spent after trips to CA to collecte PD infected materials in June 2016 and presentation of results from the project at an international conference (ASPB, Austin TX) July 9-13, 2016 (see "presentations" above). \$1,953 in publication charges remain and are budgeted for two manuscripts, in prep. \$2,030 in sequencing fee-for-service is encumbered to be paid to UC Riverside Institute for Integrative Genome Biology facility upon invoicing for sequencing currently in the pipeline. This is enough for 1.5 lanes of HighSeq2500 single end 50 bp runs which will cover 24 smRNA libraries (half are prepared and awaiting completion of the other 12 libraries for pooling) and 12 stranded mRNA-Seq libraries, which are in preparation. There is ~\$800 remaining in Maintenance and Operations. The budget allocated for the Co-I De La Fuente (\$20,851) has been completely used to partially pay the salary of the postdoc S. Traore and for supplies for ongoing greenhouse experiments with transgenic tobacco. All that remains is for invoicing from the Co-I.

Summary and status of intellectual property associated with the project

The PI previously documented in the last progress report a disclosed "Subject Invention". The pending patent application awaits First Office Action by the USPTO. The PI will disclose in due course a second subject invention that documents the reduction to practice using specific sequences from engineered vectors, to be described.

Literature Cited

- 1. Rock CD. (2013) *Trans-acting small interfering RNA4*: key to nutraceutical synthesis in grape development? *Trends Plant Sci* 18: 601-610
- 2. Jacobs T, LaFayette P, Schmitz R, Parrott W. (2015) **Targeted genome modifications in soybean with** CRISPR/Cas9. *BMC Biotech* **15**: 16
- 3. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. (2013) **RNA-guided** human genome engineering via Cas9. *Science* 339: 823-826
- 4. Llave C, Kasschau KD, Carrington JC. (2000) Virus-encoded suppressor of posttranscriptional gene silencing targets a maintenance step in the silencing pathway. *Proc Natl Acad Sci USA* **97**: 13401-13406
- De La Fuente L, Parker JK, Oliver JE, Granger S, Brannen PM, van Santen E, Cobine PA. (2013) The bacterial pathogen *Xylella fastidiosa* affects the leaf ionome of plant hosts during infection. *PLoS ONE* 8: e62945
- 6. Velten J, Cakir C, Youn E, Chen J, Cazzonelli CI. (2012) Transgene silencing and transgene-derived siRNA production in tobacco plants homozygous for an introduced *AtMYB90* construct. *PLoS ONE* 7: e30141
- 7. Zhu X, Xu Y, Yu S, Lu L, Ding M, Cheng J, Song G, Gao X, Yao L, Fan D, Meng S, Zhang X, Hu S, Tian Y. (2014) An efficient genotyping method for genome-modified animals and human cells generated with CRISPR/Cas9 system. Sci Rep 4: 6420
- 8. Shamloul M, Trusa J, Mett V, Yusibov V. (2014) **Optimization and utilization of** *Agrobacterium*mediated transient protein production in *Nicotiana*. *J Vis Exper : JoVE*: 51204
- 9. Frazier T, Xie F, Freistaedter A, Burklew C, Zhang B. (2010) Identification and characterization of microRNAs and their target genes in tobacco (*Nicotiana tabacum*). *Planta* 232: 1289-1308
- 10. Luo Q-J, Mittal A, Jia F, Rock CD. (2012) An autoregulatory feedback loop involving *PAP1* and *TAS4* in response to sugars in Arabidopsis. *Plant Mol Biol* 80: 117-129
- 11. Rogers EE. (2012) **Evaluation of** *Arabidopsis thaliana* as a model host for *Xylella fastidiosa*. *Mol Plant-Microbe Interact* **25**: 747-754
- 12. Thieme CJ, Rojas-Triana M, Stecyk E, Schudoma C, Zhang W, Yang L, Miñambres M, Walther D, Schulze WX, Paz-Ares J, Scheible W-R, Kragler F. (2015) Endogenous Arabidopsis messenger RNAs transported to distant tissues. *Nature Plants* 1: 15025
- 13. Axtell MJ. (2013) ShortStack: Comprehensive annotation and quantification of small RNA genes. *RNA* **19**: 740-751

14. Johnson NR, Yeoh JM, Coruh C, Axtell MJ. (2016) **Improved placement of multi-mapping small RNAs**. *G3: Genes/Genomes/Genetics* **6**: 2103-2111

FUNDING AGENCIES

Funding for this project was provided by the CDFA Pierce's Disease and Glassy-winged Sharpshooter Board.