Title: Fortifying plant disease diagnostics with whole genome sequencing for emerging *Xylella fastidiosa*

Introduction and Rationale. Plant diseases caused by pathogens limit crop production across the world. Identification of unknown diseases or rapidly diagnosing invasive pathogens is critical for effective crop management (1). Traditional diagnostic methods use culture-based, microscopic, serological or molecular approaches to define the causal agents of an epidemic (1). Although valuable, these approaches are often time-consuming and may not specifically identify an invasive or emerging pathogen (1). Next Generation (NexGen) sequencing technologies allow for whole genome sequencing (WGS) analysis and provide precise identification of pests and pathogens (2). However, plant diagnosticians are only beginning to unlock their potential to revolutionize the identification, characterization and tracking of disease-causing microbes (3).

The plant pathogenic bacterium *Xylella fastidiosa* limits crop and ornamental production for farmers in the US and across the globe (4). The recent emergence of *X. fastidiosa* in Italy on olive has affected the European economy and poses a potential threat to other industry such as grape (4). This emerging disease also produces significant challenges for management as it has not traditionally been experienced by the European community (4). Therefore, management tools like diagnostics are needed to better understand and best characterize future outbreaks. Here we propose to develop and expand WGS diagnostics for *X. fastidiosa*.

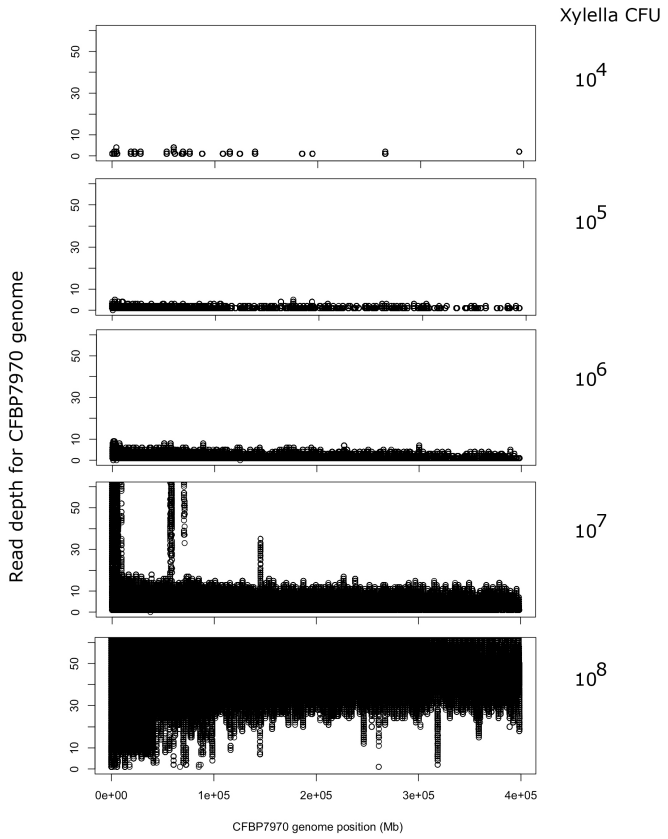

Figure. *X. fastidiosa* genome coverage of serial bacterial dilutions and reads to the reference pathogen genome. Plant samples were spiked with known concentrations of bacteria. The threshold for qPCR diagnostics is 10^4 CFU, and we were able to map reads specifically to the reference genome (x-axis). Individual dots represent mapped reads (Dupas, Roman-Reyna unpublished data).

Host laboratory and desired applicant. The Emerging Infectious Disease Ecology Laboratory (TEIDEL) led by Prof. Jonathan M. Jacobs (Ohio State) focuses on the basis of bacterial colonization of plants and translational approaches to improve management of phytobacterial diseases. The team comprises a lab manager, technician, two postdocs, five graduate students and three undergrads. This team is a dynamic, supportive group focused on *Xanthomonas, Xylella* and *Ralstonia* research. The potential COST STSM visiting researcher or student will hopefully have either basic bacteriology, molecular biology, microbiome or bioinformatics experience. This project is structured so that a candidate with bioinformatics expertise will learn important bench skills or a bench scientist will expand to learn bioinformatics pipelines for diagnostics.

Preliminary data. TEIDEL recently invested in an Illumina iSeq-100 machine to provide diagnosticians a fast, cost-effective tool to

identify pathogens with whole genomes in ways not possible until now. This proposal idea results from preliminary research funded by PI Jacobs to develop WGS diagnostics for plant pathogens, including *X. fastidiosa*. We partnered with Dr. M.A. Jacques (INRA, Angers, France), Dr. G. Marchi (University of Florence, Italy) and Prof. Francesca Peduto Hand (OSU, Columbus, USA) to use Illumina iSeq-100 next-generation WGS as a molecular tool for *Xylella fastidiosa* detection and diversity analysis.

A current standard for *X. fastidiosa* identification is by qPCR, which has a threshold limit of 10^4 CFU (5). We were able to specifically map reads of bacterial genes from mixed plant samples with serial dilution from 10^8 to 10^4 CFU, which is the threshold for qPCR (Figure). qPCR allows for specific identification of known genotypes of a pathogen. The unique advantage of WGS is that the metagenome is analyzed, and this allows for identification of known or undefined, emerging *X. fastidiosa* pathotypes.

We have preliminary data the demonstrates that we can detect and differentiate single or mixed-infected plant samples at concentrations as low as 0.05 pg/uL (data not shown). We also demonstrated that our method is a strong alternative as we detected *Xylella* reads from samples that tested negative with real-time PCR (data not shown), a technology, until now and as mentioned above, considered the goal standard of diagnostics. Overall, we developed the preliminary data of a successful diagnostics pipeline (discussed below) that could be implemented for emerging plant threat surveillance in crops. This tool development and training through a COST STSM will translate the technology to other researchers in Europe in hopes that WGS will be a possible option for identification of *X. fastidiosa* and other Xanthomonadaceae pathogens.

Objectives. This proposed project describes a novel, low-cost and user-friendly diagnostic tool that allows for quick (1-2 days) and specific identification of long-standing and emerging *X. fastidiosa* diseases. The objectives of this proposal are to: 1) prepare genomic libraries from DNA collected from naturally infected samples from Italy and across the US for WGS and 2) analyze collected samples with WGS to validate our diagnostics pipeline.

Objective 1: Prepare genomic libraries from DNA collected from naturally infected samples from Italy and across the US for WGS. We have collected genomic DNA from plants infected with *Xylella* from researcher partners in Europe (Jacques [France] and Marchi [Italy]) and across the US (MA Hansen, Virginia Tech). We plan to use these samples to validate our bioinformatics pipeline with actual field samples from known crop hosts (e.g. olive, grape, etc) and ornamental hosts (e.g. oak, etc). We can also imagine that the partner laboratory applying for this STSM can bring genomic DNA from hosts suspected to harbor *X. fastidiosa* from their region. We will also collect leaf and stem tissue in Ohio and across the Midwest before the student/researcher arrives to OSU.

We will extract DNA with a standard CTAB protocol, or DNA previously extracted will be shipped to the OSU TEIDEL team. We will follow manufacturer protocols for Illumina Nextera Library preparation. We will sequence the extracted DNA from infected plants and use the iSeq-100 output reads for downstream analysis. The visiting research will gain experience in genomic DNA extraction, library preparation and QC analysis and running the low-cost sequencer. These results will provide invaluable support to validate our experimental pipeline that will be adopted by the Ohio State Plant Disease Diagnostic Clinic.

Objective 2: Analyze collected samples with WGS to validate our diagnostics pipeline. The visiting student/researcher will directly analyze genomes developed from their iSeq100 experiments. We will use the bioinformatic toolkit Kraken and metaSPAdes for the analysis (6). We will align and assign taxonomic labels to the iSeq-100 reads using a custom-made database of known *X. fastidiosa* pathogens genomes. Then, we will plot percentage of reads that correspond to *X. fastidiosa*. This research will be supported by postdoctoral researcher Dr. Roman-Reyna, who established the pipeline in collaboration with PhD student Enora Dupas (Jacques' lab, INRA, France). For all sequencing experiments, we will have a data

collection pipeline that will be open eventually to the public. This diagnostic pipeline will also be publicly available so that students, growers, and extension educators can implement it in their programs.

Contribution to EuroXanth COST and significance. The exciting potential of diagnostics with metagenomics by WGS is that we have a documented, digital record of pathogens analyzed. This project has both basic and applied aspects to Xanthomonadaceae family bacteria, which includes *X. fastidiosa*. This research directly supports working group 1 of the objectives: "To develop, implement, compare and standardize methods of pathogen detection by coordinating research on molecular diagnostics of plant-pathogenic Xanthomonadaceae". This proposed diagnostic method could help support the EPPO diagnostic pipeline for subspecies identification (7). We can use these sequences to hopefully detect variation in the pathogen population as the epidemic emerges or trace back how the pathogen has potentially changed. A strength of The Ohio State University's Department of Plant Pathology and diagnostic clinic is that we work directly with farmers and extension educators. We will be one of the first diagnostic clinics in the US to provide low-cost sequencing through Illumina iSeq100 for farmers and industry. The goal is the Ohio State can use these data to make recommendations for plant disease control. It will be exciting to partner with European researchers to support this initiative.

Planned period of time: 3 months (September to November 2020)

Cost estimate for STSM ($3400 = 3100 EUR):

1) Housing ($500/mo): $1500
2) Travel (flight and taxi): $1000
3) Subsistence ($300/mo): $900
   *The Jacobs laboratory will cover experiment fees and additional costs for visit.

References

1. S. A. Miller, F. D. Beed, C. L. Harmon, Plant Disease Diagnostic Capabilities and Networks. *Annu. Rev. Phytopathol.* (2009) https:/doi.org/10.1146/annurev-phyto-080508-081743.
2. C. C. Chrystoja, E. P. Diamandis, Whole genome sequencing as a diagnostic test: Challenges and opportunities. *Clin. Chem.* (2014) https:/doi.org/10.1373/clinchem.2013.209213.
3. M. E. Mechan Llontop, *et al.*, Strain-Level Identification of Bacterial Tomato Pathogens Directly from Metagenomic Sequences. *Phytopathology™*, PHYTO-09-19-0351-R (2019).
4. R. P. P. Almeida, *et al.*, Addressing the New Global Threat of Xylella fastidiosa. *Phytopathology* (2019) https:/doi.org/10.1094/PHYTO-12-18-0488-FI.
5. E. Dupas, M. Briand, M. A. Jacques, S. Cesbron, Novel Tetraplex Quantitative PCR Assays for Simultaneous Detection and Identification of Xylella fastidiosa Subspecies in Plant Tissues. *Front. Plant Sci.* (2019) https:/doi.org/10.3389/fpls.2019.01732.
6. D. E. Wood, S. L. Salzberg, Kraken: Ultrafast metagenomic sequence classification using exact alignments. *Genome Biol.* (2014) https:/doi.org/10.1186/gb-2014-15-3-r46.
7. PM 7/24 (4) Xylella fastidiosa. *EPPO Bull.* **49**, 175–227 (2019).