Arizona Iceberg Lettuce Research Council
2017-2018 Final Report

Project Title: **Genomic analysis of field soil infested with *Fusarium oxysporum* f.sp. *lactucae***

Principle Investigators:
Paul E. Brierley
Executive Director, Yuma Center of Excellent for Desert Agriculture, University of Arizona

Barry M. Pryor
Professor of Plant Pathology and Microbiology, School of Plant Sciences, University of Arizona

Introduction.
Since its initial detection in Arizona in 2001, Fusarium wilt of lettuce, caused by the pathogenic soilborne fungus *Fusarium oxysporum* f.sp. *lactucae* (FOL), has continued to spread in the soil of fields in the Yuma and Wellton areas, as well as production areas in California. Much has been learned about the epidemiology of Fusarium wilt, including the genetics of FOL and the occurrence of seedborne inoculum. Methods of control by modification of cultural practices have also been investigated, which provide a certain level of disease management through avoidance of infested fields during optimal periods for disease development. However, FOL is an opportunistic soil inhabitant and once introduced in an area, it can persist in the soil almost indefinitely. As such, regional management of Fusarium wilt will require a broader approach to include long-term strategies for management of soil populations of the pathogen. Moreover, sustainable disease management strategies for Fusarium wilt of lettuce in Arizona will require multi-pronged approaches including strategies based on improved host genetics, novel cultural practices, new chemical/biological control, and reducing soil population levels. Until these solutions are developed and proven, predicting and managing fields with high soil populations will be an important management tool. This project aims to lay the groundwork for utilization of genetic soil testing methods to predict disease pressure prior to making planting decisions.

Trace Genomics is a company that focuses on microbial analysis of soil and tackles the challenging problems of profiling the soil microbiome with emphasis on consistency, reliability, and accuracy. They are a resident of the Western Growers Center for Innovation and Technology. Their goal is to make it cost effective to add microbiome analysis to any field trials or R&D and QC pipelines. Trace Genomics provides unbiased DNA extraction from the environmental complex, effective and affordable sequencing library construction, scalable bioinformatics analytics of DNA, validated curation of microbial genome databases, functional analysis of microbial candidates, and scientific interpretation of microbial communities.

Objectives.
This study will provide fundamental data on the soil population levels and distribution pattern of the FOL pathogen in Arizona soils. The long-range goal of this study is to develop an effective strategy for quantifying and managing field populations of the Fusarium wilt pathogen in Arizona. Our immediate objectives are three-fold:

1. Conduct hierarchical sampling of 3 lettuce production fields with a history of Fusarium wilt and determine soil population of each sample using conventional soil plating techniques.
2. Submit a portion of each of the samples to Trace Genomics for analysis of the microbial community composition using next-generation genomic sequencing technologies.
3. Combine results from conventional plating and genomic sequencing and describe the soil population distribution of the Fusarium wilt pathogen in production fields in Arizona.

Materials and Methods.

Objective 1. In September and December, 2017, three fields were selected for sampling following discussion with collaborators at J-V Farms, The selection was based upon prior history of Fusarium wilt and a desire to samples fields with varying levels of potential pathogens; low, medium, and high. Each sample site was divided into ½ acre sections and in each section 3 soil cores were taken and composited. All samples obtained were dried, homogenized, and stored at 4 C until testing.

Original plans were to plate each sample onto an all-purpose nutrient agar and enumerate total *Fusarium* colonies (all different *Fusarium* species). In discussion with other *Fusarium* experts, it was concluded that the plating should be done on a modified semi-selective Komada's agar, which was developed by UC Davis plant pathologists for FOL research and differentially selects for only *Fusarium* colonies. Moreover, isolates of FOL appeared unique on this medium and could be identified and counted in the presence of other common isolates of non-pathogenic *Fusarium* species based upon distinct morphological characteristics of the fungus in culture. This enabled a quantitative assessment of fungal populations and both total *Fusarium* counts and counts of FOL per gram of soil were obtained.

Objective 2. Concurrent with our plating work, a portion of each sample was sent to Trace Genomics for fungal population determination via DNA sequencing. Upon receipt of samples, Trace Genomics extracted all DNA from the soil and constructed a DNA fragment library that was used as a template for sequence amplification and determination. The DNA sequences generated were analyzed through the Microbial Discovery and Analytics platform at Trace Genomics, and the results presented as percentages of pathogens of interest compared to total fungi recovered. In addition to FOL, the Trace Genomics platform also quantified several other soil borne pathogens. Data for all pathogens have been compiled and summarized in an estimation of overall soil health per soil core. A copy of these reports for each field sampled is included in this final report. Moreover, a comparison of these data are placed in comparison with soil dilution plating results.

Objective 3. Knowledge of the population distribution of soil pathogens provides the groundwork for recommended sampling patterns and for obtaining the resolution necessary to predict FOL disease pressure in a given field with reasonable certainty. Based upon results from both methods of soil population estimations (soil dilution plating and genomic sequencing) and using statistical methods for describing population distributions, distribution pattern of FOL populations in each field will be determined and described. Based upon distribution descriptions, recommendations for soil sampling will be developed to obtain the most statistically accurate estimate of pathogen levels in each lettuce production field. With this

information, field-specific decisions regarding Fusarium wilt management can best be developed.

Results.

Objective 1. In our original preparations of the UC Davis formulation of Komada's medium, no fungal colonies grew, and it was determined that the medium was too selective. Following several iterations of recipe adjustments, a further modification of the growth medium formulation was developed with less fungicide in the recipe that would allow for selective growth of all *Fusarium* species while suppressing the growth of most other fungi found in native soils.

Results from the soil dilution plating can be seen in Table 1, and both total *Fusarium* counts and putative FOL count were obtained. The counts obtained for total soil *Fusarium* were without complication and were reproducible as determined by replicate plating. As can be seen, populations of total *Fusarium* were high in all samples, which reveals that the modified medium is likely allowing the recovery of all *Fusarium* fungi present in soil and which is consistent with what we know about *Fusarium* fungi in general; they are some of the most common fungi present in most soils found around the world. What can also be seen is that the total Fusarium counts showed no correlation with either the FOL population rankings obtained by Trace Genomics no with historical assessment of past incidence of Fusarium wilt. This of course is no surprise as most *Fusarium* species do not cause Fusarium wilt disease. In fact, the field with the highest total *Fusarium* counts had the lowest percentile ranking by the Trace Genomic results and the lowest ranking based upon prior history of Fusarium wilt incidence.

In contrast, the results obtained for FOL populations were not robust and were not reproducible based on replicate plating. Although previous researchers found FOL to be distinct on this medium, those finding were based upon work in artificially inoculated research micro-plots or in greenhouse studies working with pure cultures. No previous work had been attempted using this medium to describe FOL population in native soils. The problem with its use in native soils became obvious in this study because of the diversity of *Fusarium* species and unique strains within species often found in a single field. Moreover, not all FOL strains looked identical on this medium, which was not anticipated. As a result, the counts obtained from the putative FOL counts did not correlate with either the Trace Genomic results or the historical incidence data. In fact, similar to the total *Fusarium* counts, the field with the highest putative FOL counts had the lowest percentile ranking by the Trace Genomic results and the lowest ranking based upon prior history of Fusarium wilt incidence.

Objective 2. Trace Genomics generated population profiles and provided a relative population ranking of FOL for each sample, the relative population percentile (%) (see Table 1 and subsequent Trace Genomic reports). The value "relative population percentile" is not a quantitative value, but instead is a value that is based upon comparisons between the number of FOL DNA sequencing reads (a proxy for abundance) from each sample from the experimental plots vs the FOL sequencing reads from all other soil samples taken in the Yuma area over the last several years. These other samples (over 200 samples from both FOL positive and FOL negative fields) were analyzed for FOL as part of on-going data set building efforts for different agricultural regions across the western states. As such, the relative population percentile value

can be used as a qualitative measure for assessing how individual soil samples compare to overall samples taken from the same general area.

While not a quantitative measurement, the qualitative value of relative population percentile can be used to assess the relative disease potential of a sample site (field). In the case of FOL, the percentile can be used to determine whether a soil sample has a relatively high level of FOL, a medium level, or a low level. For practical purposes, a high level corresponds to levels higher than 75% of all other fields; a medium level corresponds to levels higher that 25% of all other fields, and a low level corresponds to levels below 25% of all other fields (see Table 1). Results reveal that the Research Field had the highest recover of FOL, with 50% of all samples ranked as high FOL levels and the other 50% ranked as medium levels. Next ranked was the Sun Valley fields, with 25% of samples ranked as high, 25% ranked as low, and the remainder as medium. And the field with the lowest ranked disease potential was the Snyder Ranch field, with 50% of all samples ranked as low and the other 50% ranked as medium. It is important to note that all samples that ranked as low actually had no FOL recovering at all.

To validate in a more robust manner the relative ranking of fields as having a high, medium, or low disease potential, each sample ranking was correlated with a ranking of disease history based upon the incidence of disease in the most recent lettuce crop. For the Research Field and the Sun Valley Field, the disease history was ranked high, and the relative population percentile was ranked on average high and medium, respectively. The Snyder Ranch field (east side) disease history was ranked low, as was the relative population percentile ranking.

Included in the Trace Genomics report, the same samples were also evaluated for the presence of seven other pathogens. While these evaluations were not included in the objectives of this study, they provide a value-added addition to potential disease assessments for each field. Moreover, the determination of addition pathogen percentiles in conjunction with those of FOL may provide insight into subsequent disease development when disease development does not follow predicted patterns. It is well know that multiple pathogens can interact in synergistic ways to exacerbate disease potential beyond what each pathogen can contribute alone. The knowledge of the entire pathogen population complex in each field may ultimately give a more accurate assessment of disease potential and the strongest tool for crop disease management.

Objective 3. Quantitative data from soil dilution plating failed to correlate with field disease histories and therefore had limited predictive utility. In contrast, qualitative data from Trace Genomic sequencing showed relative population percentile ranking that generally correlated with field disease history rankings. Therefore, values from qualitative scores were used to calculate population distribution patterns for each field. Three indices of dispersion; the variance to mean ratio (V/X), Lloyd's index of patchiness, and Morisita's index of dispersion were calculated for each field data set to test for deviation from unity (Table 2). Values that are significantly greater than one indicates aggregated spatial pattern. Lloyd's index of patchiness was calculated by using the following formula LIP = X*/Xbar, where X* is the mean crowding and X bar is the mean. Mean crowding (X*) is calculated by $X^* = Xbar((S^2/Xbar)-1)$ where Xbar = mean and S^2 = variance. Morisita's index of dispersion was calculated by using the following formula: $I_d = n((\Sigma(x^2) - \Sigma x) / ((\Sigma x)^2 - \Sigma x))$, where n is number of sampling units and x is the propagule population.

The variance to mean ratio for fields Research, Sun Valley, and Snyder were 27.2, 67.6, and 3 3.2, respectively. The Lloyd's index of patchiness for fields Research, Sun Valley, and Snyder were 148, 1031, and 478, respectively. For both the variance/means ratio and Lloyd's index, the greater the value above 1, the greater the degree of patchiness or aggregation. Both indices describe the populations in all three fields as aggregated, with the Sun Valley field as the most aggregated. The Morisita's index for Research, Sun Valley, and Snyder were 1.03, 1.41, and 2.21, respectively. For the Morisita's index, values that deviate from 1.0 indicate increasing aggregation. For this index, aggregation decreased from Snyder Ranch to Research Field, which correlates to increasing population density. This is consistent with standard theory of distribution of soil borne pathogens, populations tend to increase from a very patchy low-density distribution to more uniform high-density populations as the pathogen is spread around the field with general farming practices.

By all measures of aggregation, all fields exhibited a patchy distribution of FOL. This has important impacts on how each field is sampled for future determination of soil population of FOL and for assessments of disease during the growing seasons. Decisions regarding collecting samples along a single diagonal transect, an X pattern, or a W pattern are strongly impacted by how aggregate the soil populations are. However, the densities of samples taken in the 2017 sampling (number of samples per field) were not sufficient in number to determine the size of the individual aggregates, and this information is important for calculating how many individual soil cores need to be taken along each arm of the sampling transects. For example, the higher the aggregation pattern, the higher the number of sample that need to be taken to insure a statistically accurate estimation of field soil population. To develop a robust soil sampling protocol for the determination of field populations of FOL, an additional year of soil sampling research needs to be conducted.

Future Directions.

Results of this study can be summarized as follows.

1. Quantitative determination of FOL soil populations using soil dilution plating techniques was not successful in providing a robust assessment of FOL in soil, and did not provide numbers that could correlate with either pre-sampling field disease history or Trace Genomic sequencing data. Although this technique has been used successfully for quantifying FOL in small research micro-plots and greenhouse studies in which the pathogen was artificially introduced into the soil, it has not been used successfully in quantifying FOL in native soils containing other native *Fusarium* species.

2. Trace Genomics amplified and sequenced DNA from native soils and provided qualitative data on the relative abundance of FOL in sampled soil, as well as the relative abundance of seven other lettuce pathogen. These values of relative abundance, recorded as low, medium, or high, correlated roughly with the pre-sampling history of the incidence of Fusarium wilt in each field. A more robust correlation with Trace Genomic data may be obtained with a more robust assessment of disease incidence per field, obtained either from the crop immediately preceding the sampling or from the crop immediately following the sampling.

3. Because of the poor quality of quantitative data obtained from direct plating of soil, correlation with Trace Genomics qualitative data was not possible. In addition, the resolution of Trace Genomic data was not high enough to permit a more robust estimation of the size of each FOL population aggregate in each field. Without a more accurate estimation of aggregate size, the development of a statistically accurate sampling strategy for estimating soil population, and subsequently disease potential, is not possible. To continue with the development of a robust soil sampling strategy that is accurate for fields with vary pathogen population densities, additional sampling at higher soil core densities will need to be conducted, with subsequent analysis of pathogen numbers.

4. The success of Trace Genomic sequencing efforts indicates that DNA-based methods for estimation of pathogen number in each soil core may be a more effective and economical means for soil analysis than traditional culture-based methods. Importantly, there are a number of cutting-edge technologies for DNA-based detection and analyses recently developed, or are in development, that are being integrated into diagnostics programs. It is anticipated that in the next few years, these methods will be commonly used for nearly all diagnostic activities, whether these activities involve plant material, water, or soil. As such, it is anticipated that these new technologies will be fully integrated into lettuce disease management programs.