Plan Overview

A Data Management Plan created using DMPTool

Title: Investigating the Genetic Diversity of Pantoea ananatis strains endemic to Georgia

Creator: Spencer Stumpf

Affiliation: University of Georgia (uga.edu)

Principal Investigator: Spencer Stumpf

Data Manager: Spencer Stumpf

Funder: United States Department of Agriculture (usda.gov)

Funding opportunity number: 29159

Template: USDA - NIFA: National Institute of Food and Agriculture

Last modified: 10-24-2017

Copyright information:

Investigating the Genetic Diversity of Pantoea ananatis strains endemic to Georgia

Quantitative and qualitative data will be ascertained for the purposes of this study. Experiments will be conducted in lab settings. Strains isolated from 1997 to 2015 from symptomatic onion, onion seed, various weed species, *Frankliniella fusca* throughout southern Georgia will be analyzed. DNA sequences from genes fusA, gyrB, leuS, pyrG, rplB, and rpoB will be gathered and analyzed by extracting genomic DNA from strains, amplifying specific genes, and submitting amplicons to Eurofins for Sanger Sequencing. Data will be visualized in Geneious for quality and phylogenetic analysis will be conducted using MEGA software. The Tamura-Nei model for assessing phylogeny will be used. Assays to assess ice nucleation and copper tolerance will be performed and repeated in the lab. For the ice nucleation assay, each strain will be replicated five times with a total of two experiments. In the event of mixed results (some water droplets solidify and some do not), a threshold of >6 droplets freezing will be required for strains to be considered ice nucleation positive. Each strain will be spread plated onto 3 plates of copper amended media one plate of non amendia media. Colonies must be present on all plates to be considered copper tolerant to account for a chance mutation on one plate not previously existent in the genome. Existing data such as isolation year, host, location, and pathogenicity will be considered for the purposes of assessing diversity. All of this data (excluding phylogenetic tree) will be arranged in an excel file. Ice nucleation and copper tolerance will be demarcated as simply a '+' or a '-'.

Sequences in FASTA format will be submitted to GenBank for public access upon publication. Phenotypic data will be gathered using MS Excel (.xls), Portable Document Format (.pdf), Joint Photographic Experts Group (.jpg), sequence (fasta), and Tagged Image File Format (.TIFF). Experiment conditions will be saved in a .pdf format outlining the specifc procedural methods for replication.

Culture suspensions are stored in 15% glycerol solution at -80 degrees Celsius and filed in the Molecular Diagnostic Lab and available for routine culture if necessary. Genomic DNA gathered will be submitted to Genbank for public access. The data backup and preservation must be maintained by the overseer of the Molecular Diagnostic Lab for smooth transition from one graduate student to the next.

There will be no permission restrictions placed on data gathered form this project. Data will be made available upon publication. No issues should arise from our intention to share data publicly and data does not intend to be witheld.

In order to successfully share information with collaborators, information will be deposited in a secure cloused-based access such as DropBox. Once we have completed our work, our data will be accessible to the public. We will use Genbank as the standard repository for datasets. Collaborators working on Whole Genome Sequencing will submit those results to Genbank, while I will submit the phenotypic data and housekeeping gene sequences. An umbrella project number will be assigned to all data to quickly access information regarding the project.

We acknowledge that data gathered from this project will be monitored as specified by NIFA. Principal investigators will be review and revise this data management plan.
