**Pathogen Surveillance: A nextflow computational pipeline for automated whole genome sequence-based diagnostics of pathogens**

ZSL Foster1, F Iruegas-Bocardo2, M Sudermann2, C Parada-Rojas2, R Alcalá Briseño2, AJ Weisberg2, JH Chang2, and NJ Grünwald1

1 Horticultural Crops Disease and Pest Management Research Unit, USDA ARS, Corvallis, Oregon, USA

2 Dept. Botany and Plant Pathology, Oregon State University, Corvallis, Oregon, USA

Computational analysis of genome sequence data for epidemiology is currently complex. Our objective is to develop an automated computational workflow that allows rapid and fine scale analysis of whole genome sequences to diagnose pathogen species and characterize the diversity within species (races, pathovars, lineages). We use a reproducible bioinformatic workflow in Nextflow to develop genome-driven analysis in the ‘PathogenSurveillance’ pipeline. Developing a proof-of-concept for *Xanthomonas* and *Phytophthora* pathogens. PathogenSurveillance allows for rapid prototyping, reproducibility, portability, and parallelism. Furthermore, this workflow can be executed in any local, cloud, or cluster computer. The pipeline automatically processes short read data and determines an initial taxonomic placement using *k-mer* searches, finds and downloads the appropriate reference genomes, assembles and annotates genomes, and performs a core genome phylogenetic analysis. In addition, the PathogenSurveillance pipeline maps reads to a reference, calls variants, and provides fine-scale population genomics insights. This allows for placing phylogenetically unknown species into a biological context for diagnosticians and researchers. The current version reproduces published data on *Xanthomonas* pathogen identity and genetic placement. We are currently expanding the pipeline to process multiple samples of different bacterial, fungal, or oomycete taxa, as well as assembling genomes. The Nextflow workflow promises to fully leverage and accelerate the analysis of whole genome sequence data and allow analysis by non-bioinformatic expert users.