

## Comparative genomics of subspecies of *Pantoea stewartii*

Zaky Adam

Agriculture and Agri-Food Canada, Canada

Plant-associated *Pantoea* are either epiphytes or pathogens and constitute the majority of validly described species of this genus. *Pantoea stewartii* includes two subspecies (*stewartii* and *indologenes*) delineated based on indole production. However, a multi-locus sequence analysis using six housekeeping genes identified two strains (DOAB 230 and DOAB 231), classified as *Pantoea stewartii* ssp. *indologenes* based on indole production, which clustered distinctively from the type strain LMG 2632. The objective of this work was to use comparative genomics to understand better the delineation of subspecies of *Pantoea stewartii*. The genomes of DOAB 21 (subspecies *stewartii*), DOAB 213 (subspecies *indologenes*) and DOAB 230 (subspecies nov.) were sequenced using Illumina HiSeq technology. De novo assembly was performed because there is no publicly available genome of *Pantoea stewartii*. DOAB 213 sequencing utilized 150 bp paired-end reads with 360 bp inserts. DOAB 230 and DOAB 21 sequencing utilized 101 bp paired-end reads with 300 bp inserts. The genomes were optimally assembled based on various k-mer values and using multiple assemblers including Velvet and Abyss. The resulting scaffolds were produced using SSPAGE, and gaps between contigs within scaffolds were filled using GapFiller. All the draft genomes assembled in this study showed highest similarity with *P. ananatis* LMG 5342. Percentage similarities between *Pantoea ananatis* LMG 5342 and DOAB 213 (73%) or DOAB 21 (76%) were comparable, while strain DOAB 230 (87%) showed the highest similarity. Using Glimmer, the number of genes predicted in *P. stewartii* ssp. *indologenes* DOAB 213, *P. stewartii* ssp. *stewartii* DOAB 21, and *P. stewartii* ssp. nov. DOAB 230 were 5396, 6129, and 4727 respectively. These genome sequences will provide a better delineation of subspecies of *P. stewartii*.

Zaky.Adam@agr.gc.ca