COMPLETE GENOME SEQUENCING OF TWO *ERWINIA* STRAINS

Zlatohurska M.[1], Gorb T.[1], Romaniuk L.[1], Wagemans J.[2], Lavigne R.[2], Kropinski A.[3], Tovkach F.[1]
[1] D.K. Zabolotny Institute of Microbiology and Virology of the NAS of Ukraine;
[2] KU Leuven; [3] University of Guelph
e-mail: zlatohurska@gmail.com

Advances in genome sequencing have produced hundreds of thousands of bacterial genome sequences, many of which have integrated prophages derived from temperate bacteriophages. These prophages play critical roles by influencing bacterial metabolism, pathogenicity, antibiotic resistance, and defense against viral attacks (Gauthier et al., 2022). Here we aimed to provide the complete genome sequences and *in silico* prophage content analysis of two *Erwinia horticola* strains, the causative agent of beech black bacteriosis in Ukraine.

The genomic DNA of *E. horticola* 60-2n and 43II was isolated using the DNeasy UltraClean microbial kit (Qiagen). The short reads library was prepared using the Nextera Flex DNA Library Kit (Illumina). The library quality was verified using the Bioanalyzer 2100 (Agilent) (High-sensitivity DNA kit). Both library preps were sequenced using the MiniSeq Mid Output flowcell (300 cycles; 2*150 bp reads). Basecalling and demultiplextion were done using the Illumina Miniseq system. In parallel, long sequencing reads were obtained by means of the MinION sequencer from Oxford Nanopore Technologies (Flowcell R9.4.1) using the Rapid Sequencing Barcoding Kit. Guppy (v3.1.5) was used for the basecalling. The genomes were assembled by Unicycler and annotated by the RAST (Wick et al., 2017, Brettin et al., 2015). Preliminary genome analysis was performed using the Comprehensive Genome Analysis Service (https://www.bv-brc.org/). The assembled genomes were considered for prophage prediction using PHASTER (Arndt et al., 2016).

The genome of *E. horticola* 60-2n consists of a circular chromosome of 5,001,269 bp with an average GC content of 55.07 %. This genome has 4,729 protein-coding sequences (CDS), 79 transfer RNA (tRNA) genes, and 22 ribosomal RNA (rRNA) genes. The annotation included 769 hypothetical proteins and 3,960 proteins with functional assignments. A total of four prophages were identified. Out of those, three prophage elements (11,732 bp, 31,389 bp and 35,799 bp) were incomplete and one prophage sequence (43,298 bp) was intact. The genome of *E. horticola* 43II has a chromosome of 5,061,335 bp (55.03% GC content) and one plasmid (183,568 bp). Annotation predicted 4,707 CDS, 78 tRNA genes, and 22 rRNA genes. Annotated genes encode 893 hypothetical proteins and 4,026 proteins with predicted functions. One intact and one incomplete prophages were identified (53,313 bp and 15,227 bp, respectively). Phylogenetic core genome analysis revealed that both *Erwinia* strains were close to *E. billingiae.* The Average Nucleotide Identity between the studied strains and *E. billingiae Eb661* was 97.6 %, suggesting that these genomes belong to the same species.

As a result, whole genome sequencing of two *Erwinia* strains was obtained. The general characteristics of genomes and mobile genetic elements were identified. Preliminary phylogenetic analysis indicated that the studied strains should be assigned to *E. billingiae* species.