



ISBN 978-85-66836-16-5

SEQUENCING, ASSEMBLING AND ANNOTATING THE GENOME OF THE BIOCONTROL AGENT *Serratia marcescens* STRAIN N4-5¹ / Sequência completa do genoma de *Serratia marcescens* strain N4-5. L. C. Ferreira², M. V. C. Viana³, T. J. Sousa³, D. Roberts⁴, J. Maul⁴, V. A. C. Azevedo³, J. T. de Souza². ²Department of Plant Pathology, Federal University of Lavras, Lavras, Brazil / ³Laboratory of Cellular and Molecular Biology, Federal University of Minas Gerais, Belo Horizonte, Brazil / ⁴Agricultural Research Service, United States Department of Agriculture. E-mail: larissacarvalhoferreira@gmail.com

Serratia marcescens is a multifaceted bacteria, able to colonize endophytically plant tissues and confer benefits such as growth promotion and control of plant pathogens. It is also able to cause disease in insects, plants and humans, where it is an opportunistic pathogen associated with the urinary and respiratory tracts. *Serratia marcescens* is a major producer of prodigiosin, a secondary metabolite that has antimicrobial, nematocidal and anticancer properties. The whole genome of the biocontrol agent *Serratia marcescens* strain N4-5 was sequenced, assembled and annotated. Sequencing was performed on four lanes of an Illumina NextSeq-500 using the run kit "Illumina NextSeq® 500/550 High Output Kit v2". The sequencing resulted in 22789104 reads and the quality was checked by Fastqc v0.11.5. A *de novo* assembly was performed using SPAdes and the contigs were united into a scaffold with CONTIGuator using *S. marcescens* strain B3R3 (accession number CP013046.2) as a reference. The gaps were closed in CLC Genomics Workbench 7 using the B3R3 genome as the reference. Genome annotation was done using RASTtk. The assembly resulted in a single circular chromosome and one extrachromosomal plasmid. The chromosome sequence was 5078487bp containing 59.69% GC, 4836 CDS, 23 rRNAs and 84 tRNAs genes. Analysis of the genome in PATRIC (Pathosystems Resource Integration Center, www.patricbrc.org) detected two CRISPR-CAS arrays, with 25 spacers and 27 repeats. Moreover, N4-5 has 4185 proteins with functional assignments, 651 hypothetical proteins, 208 genes associated with virulence factors and 36 antibiotic resistance genes. The whole genome sequence of *Serratia marcescens* N4-5 will contribute to increase our understanding of the diversity of ecological functions in this species. It will also shed light on the role of the metabolites produced by this strain in the biocontrol of plant pathogens.

Key words: Prodigiosin; NGS; Structural genomics

¹ Grant information: FAPEMIG, CAPES, CNPq.