

## Complete Genome Sequencing of Phenol Degrading *Niallia Taxi* Strain S2 , Isolated from Contaminated Soil in Northern Algeria

Adjoudj Fatma <sup>1</sup> , Maghnia Djamila <sup>2</sup> , Bouchikh Yamina <sup>3</sup> , Guessas Bettache <sup>1</sup> , Hafidh Zemour <sup>4</sup> , Berrani Abdelkader <sup>4</sup> , Ammam abdelkader <sup>5</sup> , Belmamoun Ahmed Reda <sup>6</sup> , Fathi Abdellatif Belhouadjeb <sup>7</sup>

1. Laboratory for Experimental Biotoxicology of Biodepollution and Phytoremediation, University of Oran Ahmed Ben Bella 1 . Algeria
2. Laboratory of applied microbiology university, University of Oran Ahmed Ben Bella 1. Algeria
3. National institute of agronomic research INRAA Sidi bel abbes ,Algeria
4. Faculty of Natural and Life Sciences, Ibn Khaldoun University, BP 14000, Tiaret, ALGERIA.
5. Laboratory of Pharmacognosy, Bio toxicology, and Biological Valorization of Plants, Moulay Tahar University, BP 20000, Saida, Algeria.
6. Department of Agricultural Sciences, Faculty of Nature and Life Sciences, Djillali Liabes University, Sidi-Bel-Abbes 22000, Algeria
7. Center of Research in Agropastoralisme (CRAPAST), Djelfa, ALGERIA

**Author corresponding :** fatma.adjoudj@univ-mosta.dz

Received: 05/02/2023; Accepted: 11/08/2023; Published: 04/09/2023

### Abstract

This study reports the complete genome sequencing of *Niallia taxii* strain S2, which was isolated from contaminated soil in Algeria using phenol as a sole carbon source. The genome assembly consists of 22 scaffolds/contigs with a total length of 5,606,246 bp and an average G + C content of 38.43%. The genome annotation was conducted using the NCBI Prokaryotic Genome Annotation Pipeline, RAST server with SEED database, Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), and Clusters of Orthologous Groups (KOG) databases. The genome of strain S2 contains more than 20 hydroxylase genes related to catalyzing the oxidation of various organic compounds, such as phenolic compounds. As a result of this study, we gain insight into the

## Complete Genome Sequencing of Phenol Degrading *Niallia Taxi* Strain S2 , Isolated from Contaminated Soil in Northern Algeria

biodegradation of phenol from the environment, which will be key to the bioremediation of xenobiotics from the environment.

**Key words:** complete genome sequencing , *Niallia* , degrading phenol , phenol hydroxylase ,phenolic compounds

**Tob Regul Sci. <sup>TM</sup> 2023;9(1): 4894 - 4898**

**DOI: [doi.org/10.18001/TRS.9.1.340](https://doi.org/10.18001/TRS.9.1.340)**

### Body :

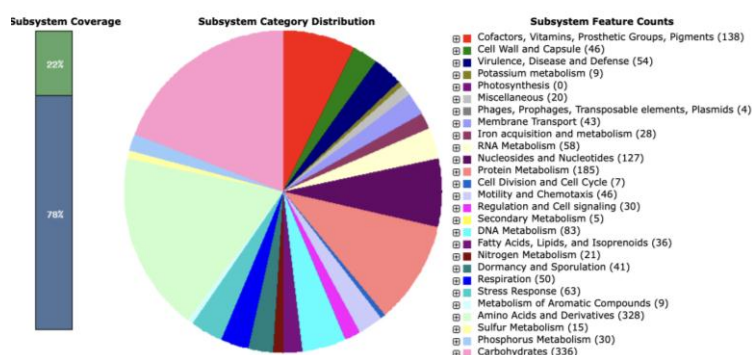
A wide range of industries use phenolic compounds, that are synthesized, such as dyes, drugs, pesticides, herbicides, insecticides, explosives, and other chemical processes. The degradation of phenolic substances is actively carried out by numerous bacterial genera, as shown in a study by (Deng et al.,2018). *Bacilli*, among the bacterial genera, are recognized as potential bioremediator agents capable of breaking down several toxic substances (Arora et al., 2014, Singh *et al.* , 2008 , Liu *et al.*, 2016) . *Niallia* is a Gram-positive, rod-shaped bacterial genus belonging to the family *Bacillaceae* within the order *Bacillales*. Notably, members of *Niallia* were formerly classified under the genus *Bacillus*. As of May 2021, *Niallia* comprises a total of five species with validly published names, as documented in the study by (Gupta *et al.*, 2021). In taxonomic investigations, this genus, *Niallia*, was identified as a distinct and monophyletic clade, with no significant phylogenetic relation to other *Bacillus* species. In our research, we present the comprehensive genome sequencing of *Niallia taxi* strain S2, which was isolated from contaminated soil in northern Algeria. The isolation was achieved through enrichment cultivation using phenol as the sole carbon source. Furthermore, a time-course experiment demonstrated the strain's ability to thrive in a liquid medium with various concentrations of phenol as the sole carbon and energy source and to use other phenolic compounds as substrate (including o-crésol , m- crésol , p-crésol , 3,4- DMP, Ethylbenzène ,benzene, toluene, and Xylène) (Djamila *et al.* , 2022 ). A commercial DNA isolation kit was employed to extract genomic DNA (Pure link kit, thermo Fisher Scientific), and the sequencing was performed using the Illumina HiSeq 2000 platform by Novogene Bioinformatics Technology. The assembly of reads and the initial results were obtained using the SMRT Link (Ardui *et al.*, 2018, Reiner *et al.*, 2018) utilizing a whole-genome shotgun (WGS) approach. The assembly of sequences was performed using SOAP denovo (version 2.0) (Li *et al.*,2008, 2010), SPAdes ( Bankevich *et al.*, 2012), ABySS software (Simpson *et al.*, 2009), and CISA software (Lin *et al.*, 2013) . The total sequence length has been counted to be 5,606,246 bp consisting of 22 scaffolds/contigs, a contig N50 of 601,385 and a contig N90 of 238,245. The genome size of *Niallia taxi* strain S2 was found to be 5,606,246 bp with an average G+C content of 38.43%. Table 1 displays the obtained results.

Table 1 : whole genome statistics

Feature type	Value
Genome Size(bp)	5,606,246
Gene Number	5,8
Gene Length	4,743,615
G+C Content (%)	38.43
% of Genome(Genes)	84.61
Gene Average Length	818
Nbre of cotings	22
Genome accession number	JAPEVP000000000

The protein sequences were aligned with the corresponding genome sequences using BLAST, and then GeneWise was used to predict gene structure based on reliable alignments (evalue < 1e-5). Then, the coding genes were predicted by Augustus (version 2.7) (Stank *et al.*,2008) with homologous evidence. The analysis revealed that there were 5800 predicted coding genes, with a total length of all the coding genes was 4,743,615bp and an average length of the coding genes was 818 bp. Approximately 84.61% of the total genome length consisted of coding regions. The genome annotation was conducted using NCBI Prokaryotic Genome Annotation Pipeline , and additional functions of the predicted genes were conducted by RAST server with SEED database , Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) and Clusters of Orthologous Groups (KOG) databases.

Based on RAST annotation of the draft complete genome, coding sequences were grouped into 311 subsystems. A total of 336 coding sequences were found in the subsystem of Carbohydrates, followed by 328 sequences in the subsystem of Amino Acids and Derivatives, and 9 genes involved in aromatic compound metabolism, including (quinate and benzoate degradation, salicylate and gentisate catabolism.



aromatic amino acid catabolism and gentisate degradation). (figure 1)

Figure 1: Subsystem category distribution for *Niallia taxii* Strain S2 was annotated using the RAST server

## Complete Genome Sequencing of Phenol Degrading *Niallia Taxi* Strain S2 , Isolated from Contaminated Soil in Northern Algeria

The bacteria that degrade phenol under aerobic conditions have enzymatic systems that allow them to respond to contamination. Among these enzymes, there are phenol hydroxylases (PH), which are involved in the initial attack of phenols. They catalyze phenol into catechol. Bacterial PHs belong to two distinct enzyme families: multicomponent hydroxylases (mPH) and single component hydroxylases (sPH) (Shingler, 2003). The genome of strain S2 contains more than 20 hydroxylase genes related to catalyzing the oxidation of various organic compounds, such as phenolic compounds such as phenol, toluene, p-cresol, 4 hydroxybenzoate, phenylacetate, naphthalene, and benzoate. *Niallia taxii* strain S2 genome information will provide insight into the biodegradation of phenol, which will be important in the bioremediation of xenobiotic substances from the environment.

*Niallia taxii* strain S2's entire genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession JAPEVP000000000. The version described in this paper's version JAPEVP010000000

under BioProject number PRJNA846062 and BioSample number SAMN28867505.

### References:

- [1] Ardui, S., Ameer, A., Vermeesch, J. R., & Hestand, M. S. (2018, February 1). Single molecule real-time (SMRT) sequencing comes of age: applications and utilities for medical diagnostics. *Nucleic Acids Research*, 46(5), 2159–2168. <https://doi.org/10.1093/nar/gky066>
- [2] Arora, P. K., & Bae, H. (2014, August 12). Biotransformation and chemotaxis of 4-chloro-2-nitrophenol by *Pseudomonas* sp. JHN. *Microbial Cell Factories*, 13(1). <https://doi.org/10.1186/s12934-014-0110-7>
- [3] Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., & Pevzner, P. A. (2012, May). SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology*, 19(5), 455–477. <https://doi.org/10.1089/cmb.2012.0021>
- [4] Deng, Y., Li, B., & Zhang, T. (2018, March 2). Bacteria That Make a Meal of Sulfonamide Antibiotics: Blind Spots and Emerging Opportunities. *Environmental Science & Technology*, 52(7), 3854–3868. <https://doi.org/10.1021/acs.est.7>
- [5] Djamila, M., Fatma, A., abdesslam, A. A., Linda, M., Abdelkader, A., & Reda, B. (2022, December 24). A novel *Niallia nealsonii* Bacteria Degrading Phenol Isolated From Oum Ghellaz Lake Shore Soil in Northern Algeria. *Egyptian Academic Journal of Biological Sciences. C, Physiology and Molecular Biology*, 14(2), 501–509. <https://doi.org/10.21608/eajbsc.2022.289666>
- [6] Gupta, R. S., Patel, S., Saini, N., & Chen, S. (2020, November 1). Robust demarcation of 17 distinct *Bacillus* species clades, proposed as novel Bacillaceae genera, by phylogenomics

Complete Genome Sequencing of Phenol Degrading *Niallia Taxi* Strain S2 , Isolated from Contaminated Soil in Northern Algeria

- and comparative genomic analyses: description of *Robertmurraya kyonggiensis* sp. nov. and proposal for an emended genus *Bacillus* limiting it only to the members of the *Subtilis* and *Cereus* clades of species. *International Journal of Systematic and Evolutionary Microbiology*, 70(11), 5753–5798. <https://doi.org/10.1099/ijsem.0.004475>
- [7] Li, R., Zhu, H., Ruan, J., Qian, W., Fang, X., Shi, Z., Li, Y., Li, S., Shan, G., Kristiansen, K., Li, S., Yang, H., Wang, J., & Wang, J. (2009, December 17). De novo assembly of human genomes with massively parallel short read sequencing. *Genome Research*, 20(2), 265–272. <https://doi.org/10.1101/gr.097261.109>
  - [8] Lin, S. H., & Liao, Y. C. (2013, March 28). CISA: Contig Integrator for Sequence Assembly of Bacterial Genomes. *PLoS ONE*, 8(3), e60843. <https://doi.org/10.1371/journal.pone.0060843>
  - [9] Liu, S. H., Zeng, G. M., Niu, Q. Y., Liu, Y., Zhou, L., Jiang, L. H., Tan, X. F., Xu, P., Zhang, C., & Cheng, M. (2017, January). Bioremediation mechanisms of combined pollution of PAHs and heavy metals by bacteria and fungi: A mini review. *Bioresource Technology*, 224, 25–33. <https://doi.org/10.1016/j.biortech.2016.11.095>
  - [10] Reiner, J., Pisani, L., Qiao, W., Singh, R., Yang, Y., Shi, L., Khan, W. A., Sebra, R., Cohen, N., Babu, A., Edelmann, L., Jabs, E. W., & Scott, S. A. (2018, January 22). Cytogenomic identification and long-read single molecule real-time (SMRT) sequencing of a Bardet–Biedl Syndrome 9 (BBS9) deletion. *Npj Genomic Medicine*, 3(1). <https://doi.org/10.1038/s41525-017-0042-3>
  - [11] Shingler, V., Powlowski, J., & Marklund, U. (1992, February). Nucleotide sequence and functional analysis of the complete phenol/3,4-dimethylphenol catabolic pathway of *Pseudomonas* sp. strain CF600. *Journal of Bacteriology*, 174(3), 711–724. <https://doi.org/10.1128/jb.174.3.711-724.1992>
  - [12] Simpson, J. T., Wong, K., Jackman, S. D., Schein, J. E., Jones, S. J., & Birol, N. (2009, February 27). ABySS: A parallel assembler for short read sequence data. *Genome Research*, 19(6), 1117–1123. <https://doi.org/10.1101/gr.089532.108>
  - [13] Singh, R. K., Kumar, S., Kumar, S., & Kumar, A. (2008, June). Biodegradation kinetic studies for the removal of p-cresol from wastewater using *Gliomastix indicus* MTCC 3869. *Biochemical Engineering Journal*, 40(2), 293–303. <https://doi.org/10.1016/j.bej.2007.12.015>
  - [14] Stanke, M., Diekhans, M., Baertsch, R., & Haussler, D. (2008, January 24). Using native and syntenically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics*, 24(5), 637–644. <https://doi.org/10.1093/bioinform>