Journal of Applied and Natural Science
17(2), 663 - 670 (2025)
ISSN : 0974-9411 (Print), 2231-5209 (Online)
journals.ansfoundation.org

Research Article

Complete genome sequence of *Niallia* sp. SS-2023 isolated from oil-contaminated soil in Mosul city, Iraq

Sahira I. H. Al-Sanjary*
Department of Biology, College of Science, University of Mosul, Mosul, Iraq
Shafak Tarik Burhan
Department of Biology, College of Science, University of Mosul, Mosul, Iraq

*Corresponding author. E-mail: sahira.alsanjary@uomosul.edu.iq

Article Info
https://doi.org/10.31018/jans.v17i2.6518
Received: January 05, 2025
Revised: May 17, 2025
Accepted: May 22, 2025

How to Cite
Al-Sanjary, S. I. H. and Burhan, S.T. (2025). Complete genome sequence of *Niallia* sp. SS-2023 isolated from oil-contaminated soil in Mosul city, Iraq. *Journal of Applied and Natural Science*, 17(2), 663 - 670. https://doi.org/10.31018/jans.v17i2.6518

Abstract
Oil contamination in soil poses a significant environmental challenge, affecting microbial communities and ecosystem functions. Certain bacterial species have evolved mechanisms to survive and even degrade hydrocarbons in such polluted environments. In this context, the genus Niallia has attracted scientific interest due to its potential in bioremediation. The current research was focused on characterizing and conducting genomic analysis of a novel strain of *Niallia* sp., accordingly labeled as *Niallia* sp. SS-2023, obtained from oil-polluted soil alongside electric generators in the Mosul Governorate. Isolates were diagnosed using conventional methods and biochemical tests. The 16S rRNA gene sequencing identification was followed by whole-genome sequencing and bioinformatics analysis. The genome contains 86 contigs, representing 3.78 million base pairs with a GC content of 39.2% and 3,852 protein-coding sequences supported by *60 tRNA* genes. **R**apid **A**nnotation using **S**ubsystem **T**echnology (RAST) server analysis identified numerous genes involved in carbohydrate, amino acid, protein, and nucleoside metabolism. Genes linked to aromatic compound metabolism, such as *SalA, FAHF, QuiB*, and *BenK*, were found, indicating roles in the degradation of aromatic, anti-inflammatory for salicylate, hydroxylated and metabolic for gentisate, as well as carboxylated and preservative for benzoate. On the phylogenetic tree, *Niallia* sp. SS-2023 falls very close to *Niallia circulans* FDAARGOS_343 and *Niallia Taxi* M5HDSG1-1T, indicating a high degree of genetic similarity. From these molecular genomic data to phylogenetics, the information reassures that *Niallia* sp. SS-2023 is promising for application in bioremediation and bioproducts development.

Keywords: *16S rRNA,* aromatic genes*, Niallia* spp.*,* whole genome sequencing .

INTRODUCTION

Major recent advancements in microbiology have occurred, particularly bacterial taxonomy and genomics. With this 2020 reclassification of the genus Niallia, a new light is being shed on the Bacillaceae family. This reclassification, in which Niallia is individualized from the extended genus of Bacillus, highly signifies detailed phylogenetic relationships of these common bacteria. The name " Niallia " has been given in honor of Professor Niall A. Logan as a result of his extensive and seminal research in the field of Bacillus taxonomy and comprehension, which has been central to the redevelopment of the taxonomic structure of the Bacillaceae family (Gupta *et al.*, 2020).

Among the main characteristic features, the Bacillaceae family possesses the ability to be motile, biochemically able to grow both in aerobic and anaerobic conditions of growth and capable of forming endospores. Otherwise, these organisms have recently been found to possess their specific forms of niche ecological adaptation and biochemical properties. For example, some *Niallia* species have been found to grow in almost any environment that presents in great amounts with sewage and soil, having a temperature range of 30-37 degrees Celsius as the optimum for their growth (Thorat *et al*., 2022).Also, the molecular evidence based on two identified Conserved Signature Indels (CSIs) found in the proteins confirmed that Niallia has molecular features that differentiate it from other Bacillaceae and establish its classification in the previously described family genome species (Srivastava and Dafale, 2024).

The genomic investigation of members of the genus Niallia is important for deepening our existing

knowledge of both microbial evolution and ecological dynamics. Whole-genome sequencing of microorganisms can serve to uncover many novel metabolic pathways and disclose unique gene functions that give insights generally important for biotechnological applications in bioremediation, bioenergy production, and innovative bioproducts (Varjani, 2017; Bohra *et al*., 2019; Behrendorff, 2021). Apart from making an in-depth analysis with regard to microbial ecology and biodiversity, this also illustrates applications in different scientific and industrial fields. The present research focused on isolating the genus Niallia from heavily oil-contaminated soils and, consequently, their identification by means of *16S rRNA* gene sequencing. Besides that, the study was also focused on the deep whole-genome analysis of *Niallia* sp. with a view to characterization, organization, and functionality of genes involved in the metabolism of aromatic compounds.

MATERIALS AND METHODS

Sample collection and processing

Oil-contaminated soil samples were collected in the Mosul governorate close to electrical generators. Samples from three locations were collected twice. The samples were collected in sterile containers before their delivery to the laboratory. In the lab, one gram of each soil sample was added to 9 ml of sterile distilled water, thoroughly mixed, and then serial dilutions were prepared. Serial dilutions in concentrations 10^3, 10^4, and 10^5 were prepared, and 0.1 mL from each was aseptically transferred and inoculated onto separate plates containing nutrient agar. The plates were incubated at 37°C for 24-48 hours to promote microbial growth (Wojtowicz *et al*., 2023).

Extraction and purification of genomic DNA

The DNA was extracted from the bacterial isolates using Geneaid Genomic DNA Purification kit. The concentration of the DNA was measured using the Nanodrop spectrophotometer and the purity was measured by taking the absorption reading at 280/260 nm.

Genome sequence submission

The genome sequence of *Niallia* sp. was deposited as *Niallia* sp. SS-2023 into the NCBI GenBank repository under the accession number JASSVR000000000. This was done through the NCBI online submission portal, ensuring that the data were formatted in FASTA format with associated metadata for completeness.

Genome assembly and annotation

Depending on Bankevich *et al*. (2012), the raw sequence reads were assembled into contigs de novo using the SPAdes 3.5 bioinformatics tool, incorporating k-mer lengths of 21, 33, 55, and 77. Assembly statistics, such as N50 - L50, total contig count, and genome coverage were produce using QUAST software (Gurevich *et al*., 2013). The annotation of the assembled genome was conducted using RASTserver (Aziz *et al*., 2008), and the SEED tool was employed to predict functional genes in subsystem categories (Overbeek *et al*., 2014).

Phylogenetic analysis

To explain the evolutionary relationships of *Niallia* sp., the Type Strain Genome Server (TYGS) was used to construct a whole-genome-based phylogenetic tree in comparison with its closest related strains, SS-2023 (Meier-Kolthoff and Göker, 2019). The genome sequence in FASTA format was submitted to the TYGS platform using default settings. Further, the phylogenetic tree was generated using the FastME 2.0 program (Lefort *et al*., 2015). The resultant tree was examined for branching patterns and inter-strain distances that give an idea about the evolutionary linkages and common ancestry of *Niallia* sp.

In silico DNA-DNA Hybridization analysis

The DNA-DNA hybridization values were retrieved using the Genome-to-Genome Distance Calculator (GGDC) web tool based on WGSs for *Niallia* sp. SS-2023 and related strains, delineating their exact relatedness both quantitatively in genetic studies and taxonomically (Meier-Kolthoff *et al*., 2022).

Analysis of phylogenetic trees based on the *16S rRNA* gene

Homologous sequences of the *16S rRNA* gene from *Niallia* sp. SS-2023 were obtained by nucleotide sequence searching using the BLASTn tool against the NCBI GenBank database. The MEGA-11 software was used for constructing the *16S rRNA*-based phylogenetic tree, and the bootstrap analyses were performed 100 times to ensure robust phylogenetic inference. The resultant analysis has yielded a detailed comparison of evolutionary relatedness based on conserved ribosomal gene sequences (Tamura *et al*., 2021).

Genome comparisons

The **BLAST Ring Image Generator** (BRIG) software was utilized to compare the genome sequences of *Niallia* sp. SS-2023 with closely related Niallia species. This generated a visual representation highlighting both the similarities and differences between the *Niallia* sp. SS-2023 genome and other related genomes (Alikhan *et al*., 2011).

RESULTS

Several of the key features of the Niallia sp. SS-2023 strain genome is shown in Table 1, which gives the

Table 1. General genome features of *Niallia* sp. SS-2023 generated using QUAST software and RAST server

Feature	Value
Genome total length (pb)	3,776,428
Number of contigs	86
Largest contig (pb)	277,889
Smallest contig (bp)	217
GC content (%)	39.2
Total of protein-coding sequences (CDSs)	3,852
Number of tRNA genes	60
N50	143,700

general genome characteristics generated by QUAST software and the RAST server. QUAST software and the RAST server. The GC content is 39.2% in the genome. The order of magnitude for the length of the genome is approximately 3.78 million base pairs. Eighty-six contigs constituted this genome, so the genome has some degree of complexity. This strain's genetic potential is further put into perspective with 3,852 protein-coding sequences and 60 tRNA genes present in this organism, which might indicate strong potential for protein synthesis. Finally, the N50 value is 143,700, indicating that there were longer contigs and a high-quality assembly.

In Fig.1, the bar graph representation of data from the RAST server more significantly augments this analysis by separating subsystem proteins with green bars from non-subsystem ones in blue bars. This analysis gives the proportion of the genome devoted to certain metabolic functions and forms the whole picture of any organism's metabolic and physiological characteristics.

The comparative genomic analysis presented in Fig. 2 and Table 2 provides important insights into the position of *Niallia* sp. SS-2023 in relation to other type strains. The data show that *Niallia taxi* M5HDSGI-1T has a relatively high in silico DNA-DNA hybridization (DDH) value of 38.2% with *Niallia* sp. SS-2023, suggesting a closer genetic relationship within the *Niallia* genus. The similar GC content further supports the close relationship between these two strains. In contrast, *Neobacillus notoginsengisoli* JCM 30743 exhibits a much lower DDH value of 13.1% with *Niallia* sp. SS-2023, indicating significant genetic distance. The considerable difference in GC content further corroborates the distinction between these strains.

The phylogenetic tree presented in Fig. 3 illustrates the evolutionary relationships among the *Niallia* sp. SS-2023 (indicated by the black circle) and its closely related strains based on *16S rRNA* gene sequences. The data reveals that *Niallia* sp. SS-2023 is closely related to *Niallia taxi* M5HDSG1-1T and *Niallia circulans* FDAARGOS 343, as evidenced by the high bootstrap values of 71 and 82, respectively. These values, along with the branch lengths, suggest that *Niallia* sp. SS-2023 is more closely related to these strains than others, such as *Bacillus mesophilum* IITR-54, which is on a distinctly separate branch with a bootstrap value of 100, indicating a distant relationship.

Table 3 further elucidates the genetic relationships among *Niallia* sp. SS-2023 and related species. The data shows that *Niallia circulans* (FDAARGOS_343) exhibits a 99.80% similarity, indicating an extremely close genetic relationship, which suggests a recent common ancestor. Similarly, *Niallia taxi* (M5HDSG1-1T) has a similarity of 99.52%, also pointing to a very close relationship within the same genus. As the similarity percentages decrease further down the list, the strain's relationship becomes more distant, as exemplified by the 97.06% similarity with *Bacillus mesophilum* (IITR-54).

Table 2. Genome Pairwise comparisons of Niallia sp. SS-2023 genome vs. type strain genomes based on isDDH, GC content, δ- value, Genome size and Number of proteins.

	Digital *is*DDH value (%)	Percent G+C (%)	δ- value	Genome Size (pb)	Number of proteins
Niallia taxi M5HDSG1-1T	38.2	37.5	0.40	5,791,551	5411
Neobacillus notoginsengisoli JCM 30743	13.1	43.78	0.46	4,776,323	4594
Niallia alba MCC 3998	14.6	35.6	0.43	5,241,236	4695
Bacillus vallismortis DV1-F-3	13.1	43.7	0.37	3,871,829	4036
Bacillus spizizenii TU-B-10	13.3	43.8	0.42	4,207,222	4297
Bacillus infantis DSM 19098	13.3	45.6	0.45	4,978,725	5134
Niallia circulans NBRC 13626	14.3	35.5	0.49	5,097,246	4964
Robertmurraya siralis JCM 12213	13.2	37.9	0.44	5,003,679	4953
Robertmurraya kyonggiensis JCM 17569	13.3	38.1	0.39	4,772,772	4593
Bacillus yapensis XXST-01	13.3	38.1	0.40	4,601,158	4620
Bacillus benzoevorans DSM 5391	13.2	40.1	0.47	4,698,353	4639
Neobacillus niacini NBRC 15566	13.2	38.1	0.46	6,183,708	5928
Peribacillus huizhouensis DSM 105481	13.2	37.2	0.37	5,022,604	4749
Bacillus cihuensis FJAT-14515T	13.1	37.0	0.35	5,429,919	5037
Cytobacillus oceanisediminis CGMCC 1.10115	13.3	40.8	0.43	5,627,752	5438
Mesobacillus zeae DSM 103964	13.2	43.2	0.45	4,255,717	4439
Niallia nealsonii FO-92	14.6	34.6	0.46	4,691,494	4253

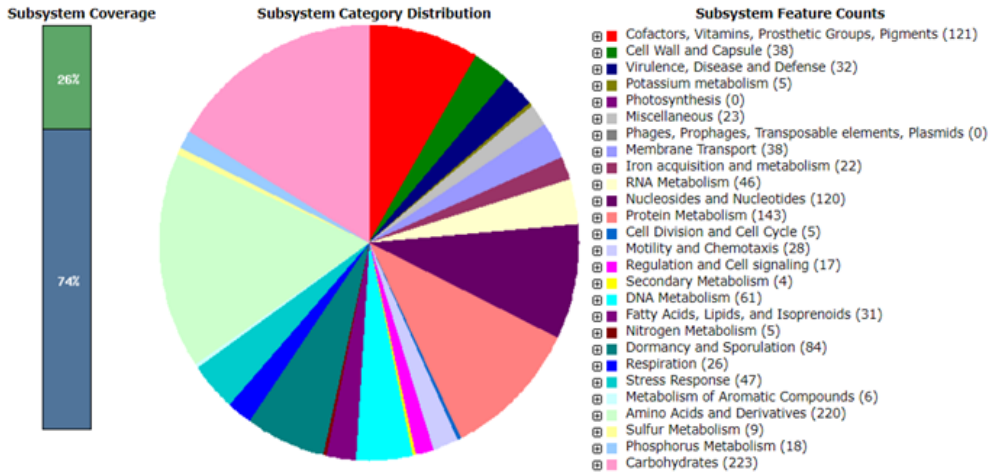

Fig. 1. *Subsystem category distribution statistics of Niallia sp. SS-2023. The genome was annotated using the Rapid Annotation System Technology (RAST) server. The pie chart showed the count of each subsystem feature, and the subsystem coverage was displayed using SEED viewer. The green bar of the subsystem coverage corresponds to the percentage of the proteins included in the subsystems, while the blue bar corresponds to the percentage of the proteins that are not included in the subsystems*

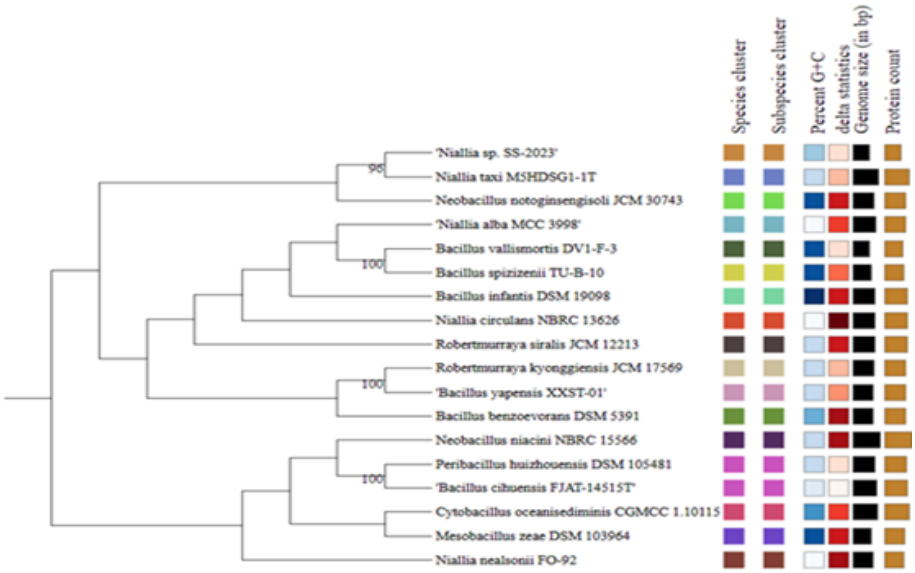

Fig. 2. *Phylogenetic taxonomy tree of Niallia sp. SS-2023 using TYGS server. The final tree was constructed with FastME 2.0 approach based on the balanced minimum evolution method (100X pseudo-bootstrap support values). Labels on leaves are indicated by association to species and subspecies clusters, genomic GC percent, δ-values, overall genome size and the total number of proteins*

The variety of genes involved in the metabolism of aromatic compounds is emphasized in Table 4. The *SalA* gene, acting as a salicylate hydroxylase, plays a pivotal role in the catabolism of salicylate and gentisate and in the degradation of salicylate esters. Likewise, the *BenK* gene, functioning as a benzoate MFS transporter, is essential for the breakdown of benzoate. The presence of genes like *FAHF* and *QuiB* in multiple pathways underscores the interconnected nature of aromatic compound metabolism, suggesting these genes may serve versatile roles in various biochemical processes. *SalE* is a significant secondary metabolite produced by NRPS/PKS gene clusters. The biosynthesis of catecholate-type siderophores, which chelate iron in iron-limited conditions, is often connected to these compounds, which support bacterial survival and growth.

A comprehensive genomic analysis of *Niallia* sp. SS-2023, set against the backdrop of its close relatives, is presented in Fig. 4. At the center of the figure is the genome of *Niallia* sp. SS-2023, represented in grey, indicates its scale up to 3.5 megabase pairs (Mbp) as a reference point. Surrounding this, a black ring illustrates the GC content, highlighting the proportion of guanine and cytosine and identifying regions of high GC concentration.

Furthermore, the purple and light green rings show the GC skew, providing perspectives on the equilibrium of guanine and cytosine throughout the DNA and suggest-

Table 3. Most related Niallia species with their accession numbers show homology with *Niallia* sp. SS-2023 retrieved from NCBI database

Species name	Strain name	Accession No.	Similarity (%)
Niallia circulans	FDAARGOS_343	RIBP01000004.1	99.80
Niallia taxi	M5HDSG1-1T	MK355518.1	99.52
Niallia nealsonii	DSM 15077	EU656111.1	99.46
Niallia circulans	ATCC 4513	AY724690.1	98.98
Niallia sp.	mt7	FZRH01000015.1	98.64
Bacillus dakarensis	Marseille-P3515	LT707409.1	97.42
Bacillus sp.	IDA3504	AJ544784.1	97.41
Niallia oryzisoli	1DS3-10	KT886063.1	97.21
Cytobacillus gottheilii	WCC 4585	FN995266.1	97.14
Bacillus benzoevorans	DSM 5391	D78311.1	97.07
Bacillus mesophilum	IITR-54	JN210567.1	97.06
Cytobacillus oceanisediminis	H2	GQ292772.1	97.05
Robertmurraya kyonggiensis	NB22	JF896450.1	97.03

ing replication trends by showcasing positive and negative skews. The outer layers of the figure feature colored rings that represent the genomes of four linked Niallia species. These colored segments denote genetic similarities and differences when compared to the reference genome of *Niallia* sp. SS-2023, allowing for a visual evaluation that highlights both genomic conservation and diversity between these species.

DISCUSSION

This approach is expected to extend present knowledge of the genomic and metabolic potential of this ecologically and industrially important bacterial genus. Molecular identification through *16S rRNA* gene sequencing and whole-genome analysis confirmed the isolate as *Niallia* sp. SS-2023, with its genomic features indicating strong metabolic potential.

The genome sequencing of *Niallia* sp. SS-2023 identified 86 contigs spanning 3.78 million base pairs, and the GC content of 39.2% includes 3,852 protein-coding sequences and *60 tRNA* genes. These characteristics are consistent with its ability to adapt to various environments by regulating complex metabolic processes, as Hou *et al.* (2021) noted in their study of related Niallia species. RAST analysis revealed a key metabolic speciality, showing 223 features dedicated to carbohydrate metabolism, 220 linked to amino acids, and notable activity in protein metabolism 143 features and nucleoside/nucleotide processing 120 features. These results suggest effective energy production, storage, and protein turnover, which is in agreement with the work of Fuentes *et al*. (2014), which demonstrated the diversity of microbial metabolic processes in environmental adaptation.

Metabolic analysis revealed that *Niallia* sp. SS-2023 degradation ability of aromatic compounds, which are recognized as carbon and energy sources for different bacteria. In particular, identifying genes such as *SalA, FAHF, QuiB*, and *BenK* confirm their involvement in pathways for aromatic degradation.

The *SalA* gene, responsible for salicylate hydroxylase, is important and required in the microbial degradation of esters of salicylates, an important step in degradation pathways of aromatics. Enzymatic activity can generate compounds that may be of real importance to bioremediation practices.This potential environmental use includes detoxifying heavily contaminated sites, as

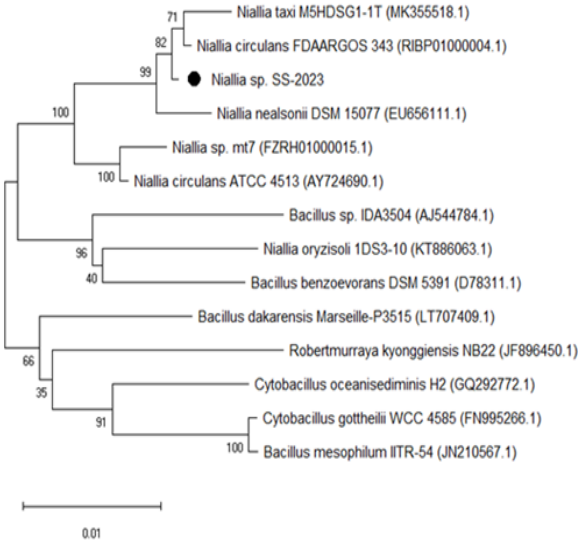

Fig. 3. *Neighbor-Joining phylogenetic trees showing the relationship between Niallia sp. SS-2023 (indicated in black circle) and the closely related strains based on 16S rRNA sequences using MEGA-11 software with a scale length of 0.01. The percentage of replicate trees in which the associated strains clustered together in the bootstrap test (100 replicates) are shown next to the branches*

Table 4. Genes involved in the metabolism of aromatic compounds

Gene	System	Function
SalA	Salicylate and gentisate catabolism Salicylate ester degradation	Salicylate hydroxylase (EC 1.14.13.1)
FAHF	Gentisate degradation Salicylate and gentisate catabolism	Fumarylacetoacetate hydrolase family protein
QuiB	Common Pathway For Synthesis of Aromatic Compounds (DAHP synthase to chorismate) Quinate	3-dehydroquinate dehydratase I (EC 4.2.1.10)
BenK	Benzoate degradation	benzoate MFS transporter BenK
SalE	Salicylate and gentisate catabolism Salicylate ester degradation	Salicylate and gentisate catabolism Salicylate ester degradation

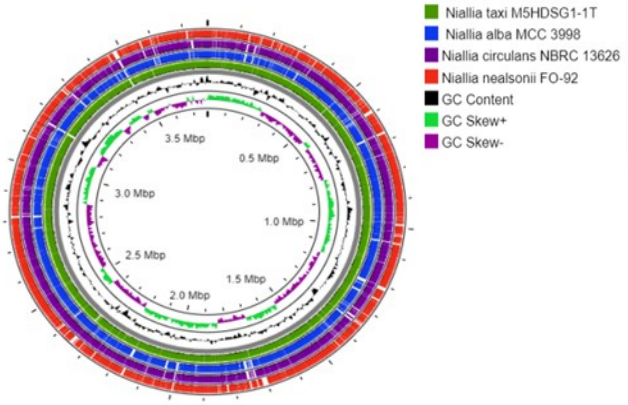

Fig. 4. *Showing Niallia sp. SS-2023 genome compared against four close related Niallia species. The innermost grey circle represents the genome of Niallia sp. SS-2023. The rings show GC content (black) and GC skew (purple/ light green). The next rings represent the genomes of other Niallia species, which are indicated in different colors. Regions without color in the ring indicate the absence of the region and the difference among the genome sequences*

pointed out by Costa *et al*. (2019). The enzyme's function underlines its relevance by allowing microorganisms to engage in the effective metabolism of generally toxic compounds of complex structure, hence justifying a wider application in environmental biotechnology.

Similarly, the *FAHF* gene encoding fumarate acetoacetate hydrolase acts at the terminal step of gentisate degradation by catalyzing the conversion of fumaryl acetoacetate to fumarate and acetoacetate. This reaction has been considered important for the complete mineralization of aromatic compounds to nontoxic byproducts, emphasizing its role in microbial pathways linked to environmental sustainability. In a similar case, fumarate acetoacetate hydrolase, as presented in several works, for example, the one done by Fuchs *et al*. (2011), has been proven to be an extremely important enzyme which secures the efficiency of degradation processes, hence central for the destruction of aromatic contaminants.

The *QuiB* gene encodes 3-dehydroquinate dehydratase, an enzyme implicated in both the biosynthesis and degradation of aromatic compounds. It catalyzes the transformation of DAHP into chorismate and is involved in the breakdown of quinate, further indicating metabolic versatility in *Niallia* sp. Such a dual function is important in cellular homeostasis and optimally utilizing available nutrients, as reported by Maes *et al*. (2004). This versatility underscores the adaptability of the organism to various ecological conditions, where metabolic efficiency is vital for its survival and use of resources.

The ability of *Niallia* sp. SS-2023 to degrade hydrocarbons is important for bioremediation in ecosystems contaminated with oil. Hydrocarbon-degrading bacteria, which enrich in such environments by having specific enzymatic pathways, vary in degradation mechanisms according to hydrocarbon complexity and different enzymes involved, as observed by Xu *et al*. (2018). Based on the comparative genomic analyses for strains of Niallia, such as those done by Gupta *et al*. (2020), *Niallia* sp. With such efficiency in hydrocarbon degradation, SS-2023 will probably be a similarly capable agent for oil spill cleanup.

Co-metabolic interaction with aromatic compounds is one of the metabolic pathways of *Niallia* sp. SS-2023, enhancing its potential for biotechnological applications. For example, the *BenK* gene, a benzoate MFS transporter, supports the degradation of benzoate and further underlines the utility of the organism in bioremediation. The works of Wojtowicz *et al*. (2023) and Patel *et al*. (2022) have demonstrated that Niallia species work effectively under both aerobic and anaerobic conditions, with adaptability advantageous in the oxygen-limited environments typical of oil-contaminated regions. The synergistic microbial consortia further increased the bioremediation potential of *Niallia* sp. it forms, especially to tackle complex oil pollutants. According to Bidja Abena *et al*. (2019), cooperative interactions in such microbial consortia have been noted to enhance the efficiency of bioremediation processes of marine and terrestrial environments that have been contaminated with crude oil. These collaborative microbial networks thus provide a formidable framework to

address challenges in hydrocarbon contamination and further advance the effectiveness of microbial remediation strategies. The enzymatic and genomic potential of *Niallia* sp. SS-2023 makes it one of the most adaptable and efficient candidates for environmental biotechnology. The degradation of hydrocarbons and aromatic pollutants justifies its use for environmental pollution, especially in ecosystems suffering from oil spills.

Conclusion

Detailed genetic characterization of " *Niallia* sp. SS-2023" decries revealed its genetic makeup and potentiality for bioremediation purposes, especially in the degradation of hydrocarbon pollutants in oil-polluted environments. When coupled with metabolic features, the organism's genetic features place it as an essential resource in the natural attenuation of pollutants, which is a promising avenue for future applications in biotechnology-related environmental cleanups. Further research is needed to fully understand and harness the potential of " *Niallia* sp. SS-2023" in works of eco-restoration and pollution control.

ACKNOWLEDGEMENTS

The researchers thank the laboratories of the College of Science/ Department of Biology/ University of Mosul and the journal's editorial board for publishing this research with appreciation.

Conflict of interest

The authors declare that they have no conflict of interest.

REFERENCES

1. Alikhan, N. F., Petty, N. K., Ben Zakour, N. L. & Beatson, S. A. (2011). BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC genomics*, 12, 1-10. DOI: 10.1186/1471-2164-12-402
2. Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., Formsma, K., Gerdes, S., Glass, E. M., Kubal, M. & Meyer, F. (2008). The RAST Server: rapid annotations using subsystems technology. *BMC genomics*, 9, 1-15. https://doi.org/10.1186/1471-2164-9-75
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. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, 19, 455-477. DOI: 10.1089/cmb.2012.0021
4. Behrendorff, J. B. (2021). Reductive cytochrome P450 reactions and their potential role in bioremediation. *Frontiers in Microbiology*, 12, 649273. https://doi.org/10.3389/FMICB.2021.649273 .
5. Bidja Abena, M. T., Sodbaatar, N., Li, T., Damdinsuren, N., Choidash, B., & Zhong, W. (2019). Crude oil biodegradation by newly isolated bacterial strains and their consortium under soil microcosm experiment. *Applied Biochemistry and Biotechnology*, 189, 1223-1244. DOI: 10.1007/s12010-019-03058-2
6. Bohra, V., Tikariha, H. & Dafale, N. A. (2019). Genomically defined Paenibacillus polymyxa ND24 for efficient cellulase production utilizing sugarcane bagasse as a substrate. *Applied Biochemistry and Biotechnology*, 187, 266-281. https://doi.org/10.1007/s12010-018-2820-5.
7. Costa, D. M., Gómez, S. V., de Araújo, S. S., Pereira, M. S., Alves, R. B., Favaro, D. C., Hengge, A. C., Nagem, R. A. & Brandão, T. A. (2019). Catalytic mechanism for the conversion of salicylate into catechol by the flavin-dependent monooxygenase salicylate hydroxylase. *International Journal of Biological Macromolecules*, 129, 588-600. DOI: 10.1016/j.ijbiomac.2019.01.135
8. Fuchs, G., Boll, M. & Heider, J. (2011). Microbial degradation of aromatic compounds—from one strategy to four. *Nature Reviews Microbiology*, 9, 803-816. https://doi.org/10.1038/nrmicro2652
9. Fuentes, S., Méndez, V., Aguila, P. & Seeger, M. (2014). Bioremediation of petroleum hydrocarbons: Catabolic genes, microbial communities, and applications. *Applied Microbiology and Biotechnology*, 98, 4781–4794. https://doi.org/10.1007/s00253-014-5684-9
10. Gupta, R. S., Patel, S., Saini, N. & Chen, S. (2020). Robust demarcation of 17 distinct Bacillus species clades, proposed as novel *Bacillaceae* genera, by phylogenomics 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, 5753-5798. DOI 10.1099/ijsem.0.004475
11. Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics*, 29, 1072-1075. DOI: 10.1093/bioinformatics/btt086 .
12. Hou, Y. J., Guo, Y., Li, D. F. & Zhou, N. Y. (2021). Structural and biochemical analysis reveals a distinct catalytic site of salicylate 5-monooxygenase NagGH from Rieske dioxygenases. *Applied and Environmental Microbiology*, 87, e01629-20. https://doi.org/10.1128/AEM.01629-20.
13. Lefort, V., Desper, R. & Gascuel, O. (2015). FastME 2.0: a comprehensive, accurate, and fast distance-based phylogeny inference program. *Molecular Biology and Evolution*, 32, 2798-2800. DOI: 10.1093/molbev/msv150
14. Maes, D., Gonzalez-Ramirez, L. A., Lopez-Jaramillo, J., Yu, B., De Bondt, H., Zegers, I., Afonina, E., Garcia-Ruiz, J. M. & Gulnik, S. (2004). Structural study of the type II 3-dehydroquinate dehydratase from Actinobacillus pleuropneumoniae. *Biological Crystallography*, 60, pp.463-471. https://doi.org/10.1107/S090744490302969X.
15. Meier-Kolthoff, J. P. & Göker, M. (2019). TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nature Communications*, 10, 2182. https://doi.org/10.1038/s41467-019-10210-3
16. Meier-Kolthoff, J. P., Carbasse, J. S., Peinado-Olarte, R. L. & Göker, M. (2022). TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. *Nucleic Acids Research*, 50, D801-D807. DOI: 10.1093/nar/gkab902

17. Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T., McNeil,L. K., Paarmann, D., Osterman, A. L., Meyer, F., Formsma, K., Kubal, M., Gerdes, S., Glass, E. M., Prjibelski, A. D., Aziz, R. K., DeJongh, M., Ben Zakour, N. L. & Stevens, R. (2014). The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Research*, *42*, D206-D214. DOI: 10.1093/nar/gkt1226
18. Patel, D., Patil, K. S. & Madamwar, D. (2022). Electrogenic degradation of reactive red 152 dye by Niallia circulans DC10 and its genome sequence analysis reveals genes mediating dye degradation and anodic electron transfer. *Journal of Water Process Engineering,* 47, 102690. https://doi.org/10.1016/j.jwpe.2022.102690.
19. Srivastava, S. & Dafale, N.A. (2024). Genomic dissection of *Niallia* sp. for potential application in lignocellulose hydrolysis and bioremediation. *Archives of Microbiology*, 206, 2. https://doi.org/10.1007/s00203-023-03728-0.
20. Tamura, K., Stecher, G. & Kumar, S. (2021). MEGA11: molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution*, 38, 3022-3027. https:// doi: 10.1093/molbev/msab120.
21. Thorat, V., Kirdat, K., Tiwarekar, B., Dhanavade, P., Karodi, P., Shouche, Y., Sathe, S., Lodha, T. & Yadav, A. (2022). *Paenibacillus albicereus* sp. nov. and *Niallia alba* sp. nov., isolated from digestive syrup. *Archives of Microbiology*, 204, p.127. https://doi: 10.1007/s00203-021-02749-x.
22. Varjani, S. J. (2017). Microbial degradation of petroleum hydrocarbons. *Bioresource Technology*, *223*, 277-286. https://doi.org/10.1016/j.biortech.2016.10.037.
23. Wojtowicz, K., Steliga, T., Kapusta, P. & Brzeszcz, J. (2023). Oil-Contaminated soil remediation with biodegradation by autochthonous microorganisms and phytoremediation by Maize *(Zea mays). Molecules* , *28*.6104. https://doi.org/10.3390/molecules28166104.
24. Xu, X., Liu, W., Tian, S., Wang, W., Qi, Q., Jiang, P., Gao, X., Li, F., Li, H. & Yu, H. (2018). Petroleum hydrocarbon-degrading bacteria for the remediation of oil pollution under aerobic conditions: a perspective analysis. *Frontiers in Microbiology*, 9, 2885. https://doi.org/10.3389/fmicb.2018.02885.