Draft genome sequence of lysogenic Pseudomonas sp. strain VB3 isolated from Vashisht hot spring, India

Draft genome sequence of lysogenic *Pseudomonas* sp. strain VB3 isolated from Vashisht hot spring, India

Running title: Mitomycin-C inducible prophage of *Pseudomonas* sp. from the hot spring.

NICOLA FARIA and *SANJEEV C. GHADI

School of Biological Sciences & Biotechnology, Goa University, Taleigao Plateau, Goa-403206, India.

*Correspondence: saga@unigoa.ac.in

ORCID identifiers: Nicola Faria, 0000-0001-9315-7362; Sanjeev C.Ghadi 0000-0002-0007-390X

Abstract:

Bacterial strain VB3, an aerobic bacterium, was isolated from the water of Vashisht hot spring, Himachal Pradesh, India. The lysogenic bacterial strain VB3's draft genome sequences comprise approximately 6.06 MB DNA and 5607 predicted genes with a G+C content of 60.56%. The 16S rRNA sequencing assigned the bacterial strain VB3 to the genus *Pseudomonas*. *Pseudomonas* sp. VB3 was a Gram-negative, rod-shaped bacteria demonstrating catalase, oxidase, citrate utilization, cellulase, and lipase activities. Genome mining revealed fourteen gene clusters coding for the biosynthesis of bioactive secondary metabolites. PHASTER analysis identified two intact and three incomplete putative prophage-related sequence regions in the draft genome. The reduction in the growth rate of the strain VB3 on mitomycin-C treatment was concurrent with the detection of the clearance zone in the spot assay.

Keywords: Draft genome, Hot spring, Lysogenic, *Pseudomonas* sp. VB3, Secondary metabolites, Prophage.

Abbreviations:

ANI: Average nucleotide identity

APE Vf: Arylpolyene cluster.

BGCs: Biosynthetic gene clusters

dDDH: digital DNA-DNA hybridization

GSI: Geological Survey of India.

NAGGN: N-acetylglutaminylgutamine amide

NRPS: Non-ribosomal peptide synthase

ONPG: α ortho-nitrophenyl β-galactoside

PAH: Polyaromatic hydrocarbons

PHASTER: PHAge Search Tool Enhanced Release

RiPP-Like: Ribosomally synthesised and post transcriptionally modified peptide product

TYGS: Type strain Genome server

UV: Ultraviolet

Statement and Declarations

Conflict of interest: The authors declare no conflict of interest in the publication.

Ethical approval: This article does not contain any studies with human participants or animals performed by any authors.

Nucleotide sequence accession number:

The 16S rRNA sequence (1368 bp) of the strain VB3 with accession number MZ520013 is accessible from GenBank database. This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the

accession JAPQMF000000000 (BioProject PRJNA899365). The version described in this paper is version JAPQMF010000000.

Acknowledgments: Nicola Faria is thankful to the University Grants Commission, New Delhi, India for providing financial support under the Maulana Azad National Fellowship (MANF-2015-17-GOA-62859).

Authors' contributions: NF performed the experimental work, analysis, and manuscript writing. SCG designed the work, supervised and reviewed the manuscript.

Introduction

Geothermal springs from India are grouped into seven geological provinces: Himalaya, Sohana, Cambay, West Coast, Sona-Narmada Tapi (SONATA), Godavari, and Mahanandi (Chandrasekharam, 2000). The Geological Survey of India (GSI) has reported around 400 thermal springs based on tectonic elements and geothermal gradients across India. (Chandrasekhar & Chandrasekharam, 2010). Several hot springs are located in Himachal Pradesh's Kullu, Mandi, and Kinnaur districts. The Vashisht hot spring from the Kullu district is three kilometers from Manali, on the other side of the Beas River. The curative properties of the hot spring make it one of Manali's most popular tourist destinations. Although a wide distribution of hot springs across India exists, limited attention has been accorded to microbiological analysis.

Migula first discovered the genus *Pseudomonas* in 1894, and around 586 species have been described in the list of prokaryotic names with standard nomenclature (Parte, 2020). *Pseudomonas* species are ubiquitous and isolated from diverse environments such as soil, plants, freshwater, clinical specimens, and marine environments (Peix et al. 2009). Additionally, *Pseudomonas* species and associated bacteriophages have been extensively studied for their medical and ecological importance (Ceyssens & Lavigne, 2010). Two novel *Pseudomonas* species have been reported and studied from Taiwan's hot springs, (Liu et al. 2013, Wong and Lee, 2014). The metagenomic studies have also revealed the presence of *Pseudomonas* species from Indian hot springs (Saxena et al. 2017, Sharma et al. 2018, Poddar and Das, 2018). Although lysogeny is a favourable lifestyle for bacteriophages in hot springs (Breitbart et al. 2004), particularly in Indian hot springs, they are the least explored. Based on this, the study aimed to screen for temperate bacteriophages from hot springs by inducing the host organism with several inducing agents such as ultraviolet (UV)-light radiation, polyaromatic hydrocarbons (PAH) and Mitomycin-C.

The present study reports the draft genome sequences of a lysogenic *Pseudomonas* sp. strain VB3 isolated from the Vashisht hot spring, Himachal Pradesh, India harboring gene clusters for bioactive secondary metabolites and putative mitomycin-C inducible prophage-related sequences.

Material and Method:

Water samples were collected directly from the sprout of the Vashisht hot spring in Himachal Pradesh, India (GPS coordinates: (32.26684 N,77.18753 E) where no anthropogenic activity was observed. Temperature, pH, electrical

conductivity, and TDS of water samples were measured on-site using a multi-parameter (CyberScan PC 650, Eutech Instruments, Thermo Fisher Scientific, India). Culturable bacteria were isolated within 48 h of collecting the water sample.

Bacterial strain isolation and identification

Around 100 µl of the water sample from the hot spring was aseptically plated on the basal medium (BM) agar plates, (g/L) containing KH_2PO_4: 1.36, $(NH_4)_2SO_4$:1.0, $MgSO_4.7H_2O$: 0.2, $FeSO_4$: 0.001, NaCl: 2.0, yeast extract, 1.0 and supplemented with 0.2% sodium succinate and were incubated at 50°C. Morphologically distinct colonies were isolated and purified. The bacterial isolates from the hot spring served as potential hosts for screening temperate bacteriophages. As a preliminary screening for prophage induction, all bacterial isolates were either treated with UV-light radiation, polycyclic aromatic hydrocarbons (PAH), or mitomycin-C (Jiang and Paul, 1996). A bacterial isolate known as strain VB3 was selected for further study since mitomycin-C treatment resulted in prophage induction. The colony morphology and the Gram staining of bacterial strain VB3 was determined. The biochemical and growth characteristics (temperature and pH tolerance) of bacterial strain were also studied. The growth of bacterial strain VB3 was determined between 4 and 50°C, whereas growth at various pH was evaluated between 5 and 11. The tolerance of the strains VB3 to different concentrations of NaCl was tested by supplementing BM medium with NaCl in the range of 0-3% (w/v)._Biochemical tests for catalase, oxidase, nitrate reduction, amylase, cellulase, protease, and lipase activity of strain VB3 were determined as per standard protocols. The carbohydrate utilization pattern was determined by HiCarbo Kit KB009 (Himedia, India)._The bacterial strain VB3 was preserved in BM containing 15% (v/v) glycerol at -80°C.

Genomic DNA isolation and 16S rRNA sequencing

Bacterial strain VB3 was inoculated into BM (Basal Medium) and was incubated at 30°C for 24 h. The bacterial genomic DNA of strain VB3 was extracted with the Genei bacterial DNA isolation Kit (Genei, Bangalore, India). The 16S rRNA gene sequence was determined after PCR amplification using 27F and 1429R universal primers (Frank et al. 2008). A phylogenetic tree was constructed by the Neighbor-joining tree method using MEGA-X (Molecular Evolutionary Genetics Analysis across computing platforms) software v.10.1.8 (Kumar et al. 2018).

De Novo whole genome sequencing and annotation

For next-generation sequencing, the genomic library was prepared from bacterial strain VB3 genomic DNA using the Nextera XT as per the manufacturer's instructions (Cat #FC-131-1024). The Nextera XT adapters were added to the DNA fragments and the sequences were read on the Illumina HiSeq platform (150 x 2 chemistry). A total of 8,269,162 million reads with a mean fragment length of ~358 bp were obtained. The quality check, trimming, assembly, and annotation for WGS were done on Galaxy Australia version 20.05 web server (https://usegalaxy.org.au/) platform (Jalili et al. 2020). The Illumina pair-ended reads were quality-checked using Fast QC v 0.11.8 (Andrews, 2010). The quality reads were trimmed using Trim Galore v.0.6.3. The high-quality reads were then assembled into contigs using SPAdes assembler (v 3.12.0) (Bankevich et al. 2012) with the default settings. The kmer value was set to 77K-mer. The assembled contigs were checked using Bandage Info v. 0.8.1 (Wick et al. 2015) and the assembly statistics were determined by QUAST software (Gurevich et al. 2013). The functional annotation of *Pseudomonas* sp. VB3 was also done by Rapid annotation using subsystem technology (RAST) web service (Aziz et al. 2008). The Prokka v1.14.5 (Seemann 2014) pipeline determined the gene prediction and functional annotation. For genomic comparison, the average nucleotide identity (ANI) was calculated using the OrthoANIu calculator on the EZbiocloud server (https://www.ezbiocloud.net/tools/ani) (Yoon et al. 2017). The assembled draft bacterial genome was uploaded to the Type strain Genome server (TYGS) server (https://tygs.dsmz.de/) (Meier-Kolthoff et al. 2022) and the digital DNA-DNA hybridization (dDDH) was performed. The antiSMASH 6.1.1 server was used to identify the biosynthetic gene clusters (BGCs) for various secondary metabolites (Blin et al. 2021). The genomic circular map was constructed by using the Proksee server (https://proksee.ca/) (Stothard et al. 2019).

Prophage prediction was done using PHASTER (PHAge Search Tool Enhanced Release) application program interface (API) (http://phaster.ca/) (Arndt et al. 2016). It was used to predict putative prophages in the genome of *Pseudomonas* sp. VB3.

Prophage induction assay with Mitomycin-C and spot assay.

Bacterial strain VB3 was grown in BM medium until the optical density at 600 nm (OD_{600}) was 0.3, after which mitomycin-C was added at 1 µg/mL (Chen et al. 2006). The OD 600 nm was recorded every two hours. The culture was incubated at 30°C for 18 h and clarified by centrifugation at 5000 ×g for 10 min, followed by filtration through

a 0.22 µm filter. 10µl of filtered lysate was spotted on the lawn of *Pseudomonas* sp VB3 host (Sambrook et al. 1989). Similarly, 10µl of lysate was spotted on the VB3 strain grown in the absence of mitomycin-C as control.

Results and Discussion

The temperature and pH of the water sample collected from the Vashisht hot spring were recorded as 50ºC and 7.4 respectively. The conductivity and TDS measured *in situ* were 162 µS cm^{-1} and 780 mg/L, respectively. While investigating the culturable bacterial diversity in the water sample of the Vashisht hot springs on BM plates, five morphologically distinct colonies were isolated on the BM plates. All 5 bacterial isolates were tested for preliminary prophage induction using UV-light radiation, PAH, and mitomycin-C treatments. However, PAH or UV-light radiation were unable to induce any prophages. Out of the five bacterial isolates, one of the bacterial isolate designated as strain VB3 demonstrated a decrease in cell density after treatment with Mitomycin-C. This bacterial strain VB3 was indicative of a lysogenic strain.

Bacterial strain VB3 grew at temperatures ranging from 25 to 50°C, with 30°C being the optimal temperature for growth (Fig. S1). Growth of strain VB3 was observed between pH 6-10, with pH of 7 being optimal for growth (Fig. S2). Strain VB3 grew on BM medium containing 0.1-0.2% NaCl, no growth was observed above 3% NaCl. The rod shaped bacterial strain VB3 was Gram-negative and motile. Strain VB3 was positive for the catalase and oxidase activities The bacteria demonstrated positive activities for citrate utilization, cellulase, and lipase. Nitrate reduction and ONPG (α ortho-nitrophenyl β-galactoside) tests were negative. The morphological and other biochemical characteristics are described in Table S1.

The 16S rRNA gene sequence of bacterial strain VB3 (MZ520013) demonstrated pairwise sequence similarity to *Pseudomonas gessardii* strain DSM 17152^T (99.63%), *Pseudomonas synxantha* strain DSM 18928^T (99.55%), *Pseudomonas libanensis* strain CIP 105460^T (99.48%), *Pseudomonas mucidolens* strain LMG 2223^T (99.40%), *Pseudomonas azotoforman* strain DSM 18862^T (99.33%), *Pseudomonas carnis* strain B4-1^T (99.33%), *Pseudomonas paralactis* strain DSM 29164^T (99.33%), *Pseudomonas lactis* strain DSM 29167 ^{T} (99.25%), and *Pseudomonas bernneri* strain CFML 97-391 ^{T} (99.25%) and lower sequence similarity to other species of the genus *Pseudomonas*. The phylogenetic tree constructed using the sequence of closely related bacterial type strains by the Neighbour-Joining tree method indicated that the strain VB3 formed a coherent cluster with the clade that comprises the genus-type strains *Pseudomonas* (Fig. 1). This topology relationship was supported by the Maximum-Likelihood tree also. Thus, strain VB3 is closely related to *Pseudomonas gessardii* and was identified as *Pseudomonas* sp. VB3.

The genomic DNA of *Pseudomonas* sp. strain VB3 was sequenced using Illumina HiSeq 2500 platform and approximately 8,269,162 million reads with a mean fragment length of ~358 bp were obtained (59.69 % G+C content). The quality of the raw data was evaluated using FastQC. The contigs were assembled by *De novo* assembly using SPAdes assembler (v 3.12.0) (Bankevich et al. 2012) at the Galaxy platform. The genome assembly size of *Pseudomonas* sp. strain VB3 is 6,065, 353 bp distributed over 256 long contigs and has a G+C content of 60.5% which is comparable to other *Pseudomonas* species. The largest contig length was 1, 97,576 bp, and the N50 of the genome was 56,396 bp. The draft genome of *Pseudomonas* sp. strain VB3 is 6,065, 353 bp (Table S2). The general features of the *Pseudomonas* sp. strain VB3 are summarised in Table S3, ensuing the Minimum information about any (X) sequence as per (MIxS) standard checklist.

The genome sequence of *Pseudomonas* sp. VB3 includes 5607 protein-coding genes, four rRNAs, and 58 tRNAs, in 256 contigs. Applying the RAST, a total of 5607 protein-coding genes were assigned to 388 subsystems (Fig. 2). The number of protein-coding sequences within the genome is about 5607, with 62 non-coding RNAs (including tRNA (58), and rRNA (4) (Fig. 3). The maximum values of ANI and digital DNA-DNA hybridization (dDDH) of the genome were 90.45 % and 49.5%, respectively (Table S4). The ANI threshold for species delineation is recommended at 95–96% (Richter & Rosselló-Móra, 2009) and ANIu between *Pseudomonas* sp. strain VB3 and phylogenetically closest relatives is ≤ 90%. The ANI values were significantly lower than the threshold for describing prokaryote species (95-96%). The dDDH values of *Pseudomonas* sp. strain VB3 were predicted by using the formula d4 to *P.gessardii* DSM 17152 ^{T} and *P. synxantha DSM 18928* ^{T} were 49.5.6% and 29.1% respectively.

A whole genome-based taxonomic analysis performed in Type strain Genome server (TYGS) constructed a GBDP (Genome BLAST Distance Phylogeny) tree on genome data of *Pseudomonas* sp. strain VB3 illustrated that it forms a distinct clade with *P.gessardii* DSM 17152 ^{T} (Fig. S3). However, the dDDH analysis obtained from TYGS stated that the strain VB3 separate from *P. gessardii* DSM17152 as the dDDH values between them were 49.5%, which again is less than the recommended threshold (<70%). The dDDH values of other *Pseudomonas* sp. were ≤ 29.1% is much lower than the species threshold of 70% recommended for species differentiation (Meier-Kolthoff et al. 2013) (Table S4).

Genome mining revealed that *Pseudomonas* sp. strain VB3 harbours 14 gene clusters coding for the biosynthesis of bioactive secondary metabolites (Table 1). Based by antiSMASH 6.1.1 analysis, the genome of *Pseudomonas* sp.

strain VB3 harbours NRPS (Non-ribosomal peptide synthase), arylpolyene, RiPP-like, redox-cofactor, betalctone, butyrolactone, hserlactone, NAGGN (N-acetylglutaminylgutamine amide), NRPS-like producing secondary metabolite biosynthetic gene clusters (BGC). The antiSMASH 6.1.1 prediction suggested that *Pseudomonas* sp. strain VB3 may produce small molecules like Pyoverdine, Lankacidin C, and fengycin. The antiSMASH 6.1.1 identified four BGCs for pyoverdine (Clusters 2,7,11 and 13) produced by fluorescent *Pseudomonas* strains and represent a major class of siderophores that possess high-affinity iron chelating abilities (Greenwald et al. 2007). The BGC for lankacidin C is known to have antitumor and immunosuppressive activities on cancer cell lines (Oostu et al. 1975). Additionally, fengycin has been reported as a biocontrol against fungal disease in crop plants (Zeng et al. 2021). Further characterization of these pathways and secondary metabolites from *Pseudomonas* sp. VB3 would unravel novel compounds with unique biological properties.

Prediction of Prophage and Preliminary induction assay using Mitomycin-C

The putative prophage DNA sequences were detected using PHAge Search Tool Enhanced Release (PHASTER) method. Five putative prophages-like elements were identified from the genome of *Pseudomonas* sp. strain VB3. The genome sizes of *Pseudomonas* VB3 prophages ranged from approximately 7.4- 47.6 kb and GC content varied between 52.49 and 60.53% (Table 2). Prophage search by PHASTER showed the presence of coding sequences (CDS) predictive for two intact and three incomplete prophage sequences in the genome of *Pseudomonas* sp. VB3.

After mitomycin-C treatment of *Pseudomonas* sp. strain VB3, cell lysis was detected by a decrease in the optical density of the bacterial culture. Mitomycin-C treated induced decreased growth curves was indicative of lysogeny in *Pseudomonas* sp. VB3 (Fig. 4). An initial increase followed by a sharp decrease in optical density in the mitomycin-C treated culture compared to the respective control was an obvious indication of the presence of prophage. In spot assay, lysis was observed on the host lawn as clearance on the lawn (Fig. 5). Prophages have been identified as widespread and varied genetic components by studying bacterial genomes. Under certain adverse conditions, active prophages can be induced during host cell division. Prophages, as external genetic material, contribute significantly to the evolution of the host strain by transferring their useful genes (Casjens, 2003). These could make the bacterial host population more able to adapt to a certain environment (Zeng et al. 2016). Further characterization of the induced virus particle through TEM imaging and NGS sequencing will give a better understanding of which putative temperate prophage from strain VB3 is inducible by Mitomycin-C

Conclusion:

In this study, we determined the draft genome sequence of *Pseudomonas* sp. VB3 from Vashisht geothermal spring in the Himalayan geothermal province. This lysogenic bacterial strain VB3 harbours 5 putative prophage regions. Genome mining also revealed that *Pseudomonas* sp. strain VB3 harbours 14 gene clusters coding for the biosynthesis of bioactive secondary metabolites. Further characterization of these pathways and metabolites will support the goal to identify novel compounds with biological properties. Characterization of the Mitomycin-C inducible phage will give a better understanding of prophages and phages from the hot springs.

Nucleotide sequence accession number:

The 16S rRNA sequence (1368 bp) of the strain VB3 was submitted to the GenBank database with accession number MZ520013. This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession JAPQMF000000000 (BioProject PRJNA899365). The version described in this paper is version JAPQMF010000000.

Acknowledgements: Nicola Faria is thankful to the University Grants Commission, New Delhi, India for providing financial support under the Maulana Azad National Fellowship (MANF-2015-17-GOA-62859).

Authors' contributions: NF performed the experimental work, analysis, and manuscript writing. SCG, designed the work, and supervised and reviewed the manuscript.

Statement and Declarations

Conflict of interest: The authors declare no conflict of interest in the publication.

Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors.

References

Andrews, S. (2017) FastQC: a quality control tool for high throughput sequence data. 2010. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS (2016) PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res 44: W16-21. https://doi.org/10.1093/nar/gkw387

Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F (2008) The RAST Server: rapid annotations using subsystems technology. BMC Genomics 9:75. https://doi.org/10.1186/1471-2164-9-75

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455-77. https://doi.org/10.1089/cmb.2012.0021

Blin K, Shaw S, Kloosterman AM, Charlop-Powers Z, Van Wezel GP, Medema MH, Weber T (2021) antiSMASH 6.0: improving cluster detection and comparison capabilities. Nucleic Acids Res 49: W29-35. https://doi.org/10.1093/nar/gkab335

Breitbart M, Wegley L, Leeds S, Schoenfeld T, Rohwer F (2004) Phage community dynamics in hot springs. Appl Environ Microbiol *70*:1633–1640. https://doi.org/10.1128/AEM.70.3.1633-1640.2004

Casjens S (2003) Prophages and bacterial genomics: what have we learned so far? Mol Microbiol 49 :277-300. https://doi.org/10.1046/j.1365-2958.2003.03580.x

Ceyssens PJ, Lavigne R (2010) Bacteriophages of Pseudomonas. Future Microbiol 5 :1041-55. https://doi.org/10.2217/fmb.10.66

Chandrasekhar V, Chandrasekharam D (2010) Energy independence through CDM using geothermal resources: Indian scenario. In Proceedings World Geothermal Congress.

Chandrasekharam D, Chandrasekhar V (2000) Geothermal energy resources of India: country update. In Proceedings World Geothermal Congress 2000 (pp. 133-145).

Chen F, Wang K, Stewart J, Belas R (2006) Induction of multiple prophages from a marine bacterium: a genomic approach. Appl Environ Microbiol 72:4995-5001. https://doi.org/10.1128/AEM.00056-06

Frank JA, Reich CI, Sharma S, Weisbaum JS, Wilson BA, Olsen GJ (2008) Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes. Appl Environ Microbiol 74(8):2461-70. https://doi.org/10.1128/AEM.02272-07

Greenwald J, Hoegy F, Nader M, Journet L, Mislin GL, Graumann PL, Schalk IJ (2007) Real time fluorescent resonance energy transfer visualization of ferric pyoverdine uptake in *Pseudomonas aeruginosa*: a role for ferrous iron. J Biol Chem 282 :2987-95. https://doi.org/10.1074/jbc.M609238200

Gurevich A, Saveliev V. Vyahhi and, N., Tesler G (2013). QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072-5. https://doi.org/10.1093/bioinformatics/btt086

Jalili V, Afgan E, Gu Q, Clements D, Blankenberg D, Goecks J, Taylor J, Nekrutenko A (2020) The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2020 update. Nucleic Acids Res 48: W395-402. https://doi.org/10.1093/nar/gkaa554

Jiang S C, and Paul J H (1996) Occurrence of lysogenic bacteria in marine microbial communities as determined by prophage induction. *Mar. Ecol. Prog. Ser.*, *142*, 27-38. https://doi.org/10.3354/meps142027

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 35:1547-1549. https://doi.org/10.1093/molbev/msy096

Liu YC, Young LS, Lin SY, Hameed A, Hsu YH, Lai WA, Shen FT, Young CC (2013) *Pseudomonas guguanensis* sp. nov., a gammaproteobacterium isolated from a hot spring. Int J Syst Evol Microbiol 63: 4591–4598. https://doi.org/10.1099/ijs.0.047712-0

Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M (2013) Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics 14:60. https://doi.org/10.1186/1471-2105-14-60

Meier-Kolthoff JP, Carbasse JS, Peinado-Olarte RL, Göker M (2022) TYGS and LPSN: a database tandem for

fast and reliable genome-based classification and nomenclature of prokaryotes. Nucleic Acids Res 50: D801-7. https://doi.org/10.1093/nar/gkab902

Ootsu K, Matsumoto T, Harada S, Kishi T (1975) Antitumor and immunosuppressive activities of lankacidin-group antibiotics: structure-activity relationships. Cancer Chemother Rep 59:919-28.

Parte AC, Carbasse JS, Meier-Kolthoff JP, Reimer LC, Göker M (2020) List of Prokaryotic names with Standing in Nomenclature (LPSN) moves to the DSMZ. Int J Syst Evol Microbiol 70 :5607. https://doi.org/10.1099/ijsem.0.004332

Peix A, Ramírez-Bahena MH, Velázquez E (2009) Historical evolution and current status of the taxonomy of genus *Pseudomonas*. Infect Genet Evol. 9 :1132-47. https://doi.org/10.1016/j.meegid.2009.08.001

Poddar A, and Das S K (2018) Microbiological studies of hot springs in India: a review. Arch Microbiol, *200*: 1–18. https://doi.org/10.1007/s00203-017-1429-3

Richter M, Rosselló-Móra R (2009) Shifting the genomic gold standard for the prokaryotic species definition. Proc Nat Acad Sci U S A 106 :19126-31. https://doi.org/10.1073/pnas.0906412106

Sambrook J, Fritsch EF, Maniatis T. (1989) *Molecular cloning: a laboratory manual*. New York: Cold spring harbor laboratory press.

Saxena R, Dhakan DB, Mittal P, Waiker P, Chowdhury A, Ghatak A, Sharma VK (2017) Metagenomic Analysis of Hot Springs in Central India Reveals Hydrocarbon Degrading Thermophiles and Pathways Essential for Survival in Extreme Environments. Front Microbiol, *7*: 2123. https://doi.org/10.3389/fmicb.2016.02123

Seemann T (2014) Prokka: rapid prokaryotic genome annotation. Bioinformatics 30 :2068-9. https://doi.org/10.1093/bioinformatics/btu153

Sharma A, Schmidt M, Kiesel B, Mahato NK, Cralle L, Singh Y, Richnow HH, Gilbert JA, Arnold W, Lal R (2018) Bacterial and Archaeal Viruses of Himalayan Hot Springs at Manikaran Modulate Host Genomes. Front Microbiol, *9*: 3095. https://doi.org/10.3389/fmicb.2018.03095

Stothard P, Grant JR, Van Domselaar G (2019) Visualizing and comparing circular genomes using the CGView family of tools. Brief Bioinform 20 :1576-82. https://doi.org/10.1093/bib/bbx081

Wick RR, Schultz MB, Zobel J, Holt KE (2015) Bandage: interactive visualization of de novo genome assemblies. Bioinformatics. 31:3350-2. https://doi.org/10.1093/bioinformatics/btv383

Wong B T, and Lee D J (2014) *Pseudomonas yangmingensis sp. nov*., an alkaliphilic denitrifying species isolated from a hot spring. J Biosci Bioeng. *117*: 71–74. https://doi.org/10.1016/j.jbiosc.2013.06.006

Yoon SH, Ha SM, Lim J, Kwon S, Chun J (2017) A large-scale evaluation of algorithms to calculate average nucleotide identity. Antonie van Leeuwenhoek, *110*: 1281–1286. https://doi.org/10.1007/s10482-017-0844-4

Zeng Q, Xie J, Li Y, Chen X, Gu X, Yang P, Hu G, Wang Q (2021) Organization, evolution and function of fengycin biosynthesis gene clusters in the *Bacillus amyloliquefaciens* group. Phytopathology Res 3:26. https://doi.org/10.1186/s42483-021-00103-z

Zeng Z, Liu X, Yao J, Guo Y, Li B, Li Y, Jiao N, Wang X (2016) Cold adaptation regulated by cryptic prophage excision in *Shewanella oneidensis*. ISME J 10:2787-800. https://doi.org/10.1038/ismej.2016.85

Table 1: Secondary metabolite biosynthetic gene clusters in *Pseudomonas* sp. strain VB3 as predicted by antiSMASH 6.1.1.

Clusters	Type	Most similar know cluster with MIBiG	MIBIG-BGC-ID (Gene location predicted by antiSMASH)
Cluster 1 (Region 5.1)	Arylpolyene	APE Vf biosynthetic gene cluster (35% of gene show similarity)	BGC0000837 (10,693 - 54,267 nt)
Cluster 2 (Region 10.1)	NPRS	Pyoverdine biosynthetic gene cluster (9% of gene show similarity)	BGC0000413 1 - 45,167 nt.
Cluster 3 (Region 16.1)	RiPP-Like	N.D	-[a] 8,363 - 19,244 nt
Cluster 4 (Region 17.1)	Redox-cofactor	Lankacidin C biosynthetic gene cluster (13% of gene show similarity)	BGC0001100 39,232 - 61,379 nt
Cluster 5 (Region 21.1)	RiPP-Like	N.D	- 42,061 - 52,906 nt.
Cluster 6 (Region 30.1)	betalactone	Fengycin biosynthetic gene cluster (13% of gene show similarity)	BGC0001095 1 - 17,637 nt
Cluster 7 (Region 35.1)	NRPS	Pyoverdine biosynthetic gene cluster (12% of gene show similarity)	BGC0000413 1 - 29,749 nt.
Cluster 8 (Region 43.1)	Butyrolactone	N.D	13,441 - 26,923 nt.
Cluster 9 (Region 47.1)	Hserlactone	N.D	17,929 - 38,507 nt.
Cluster 10 (Region 70.1)	NAGGN	N.D	17,929 - 38,507 nt.
Cluster 11 (Region 73.1)	NRPS	Pyoverdine biosynthetic gene cluster (9% of gene show similarity)	BGC0000413 1 - 28,761 nt.
Cluster 12 (Region 105.1)	NRPS-like	N.D	1 - 17,231 nt.
Cluster 13 (Region 111.1)	NRPS	Pyoverdine biosynthetic gene cluster (3% of gene show similarity)	BGC0000413 1 - 15,624 nt.
Cluster 14 (Region 151.1)	NRPS	N.D	1 - 7,765 nt.

N.D., no defined known clusters are available; NAGGN, N-acetylglutaminylgutamine amide; NRPS, non-ribosomal peptide synthase; RiPP-Like, other unspecified ribosomally synthesised and post transcriptionally modified peptide product; APE Vf arylpolyene cluster.

[a]'-' indicates no match found.
[b] The similarities indicate the percentage of genes in BGC similar to other known clusters predicted by antiSMASH.

Table 2: Putative prophage DNA sequences were detected using PHAge Search Tool Enhanced Release (PHASTER).

Prophage	Score	Region Length	No of proteins	Completeness	Closest relative	%GC
Prophage region 1	130	20.5kb	29	Intact	PHAGE_Vibrio_VP882_NC_009016 (14)	60.53%
Prophage region 2	150	47.6kb	45	Intact	PHAGE_lodoba_phPLPE_NC_0011142(14)	56.88%
Prophage region 3	20	10.3kb	19	Incomplete	PHAGE_Pseudo_PMG1_NC_016765(3)	57.38%
Prophage region 4	30	24.2kb	12	Incomplete	PHAGE_Escher_500465_1_NC_049342 (6)	58.75 %
Prophage region 5	30	7.4kb	10	Incomplete	PHAGE_Shigel_sf11_NC_021857 (2)	52.49%

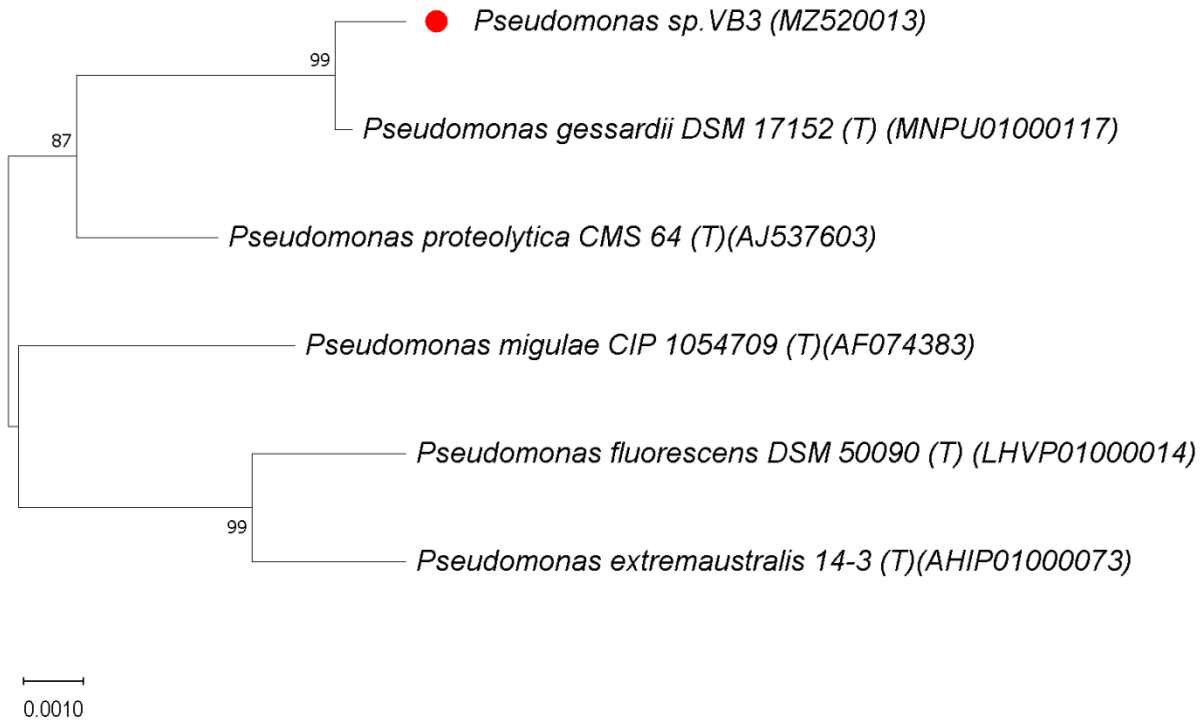

Fig. 1: Neighbor-joining phylogenetic tree based on 16SrRNA gene sequencing of *Pseudomonas* sp. strain VB3 (with a Bootstrap value of 1000 replication).

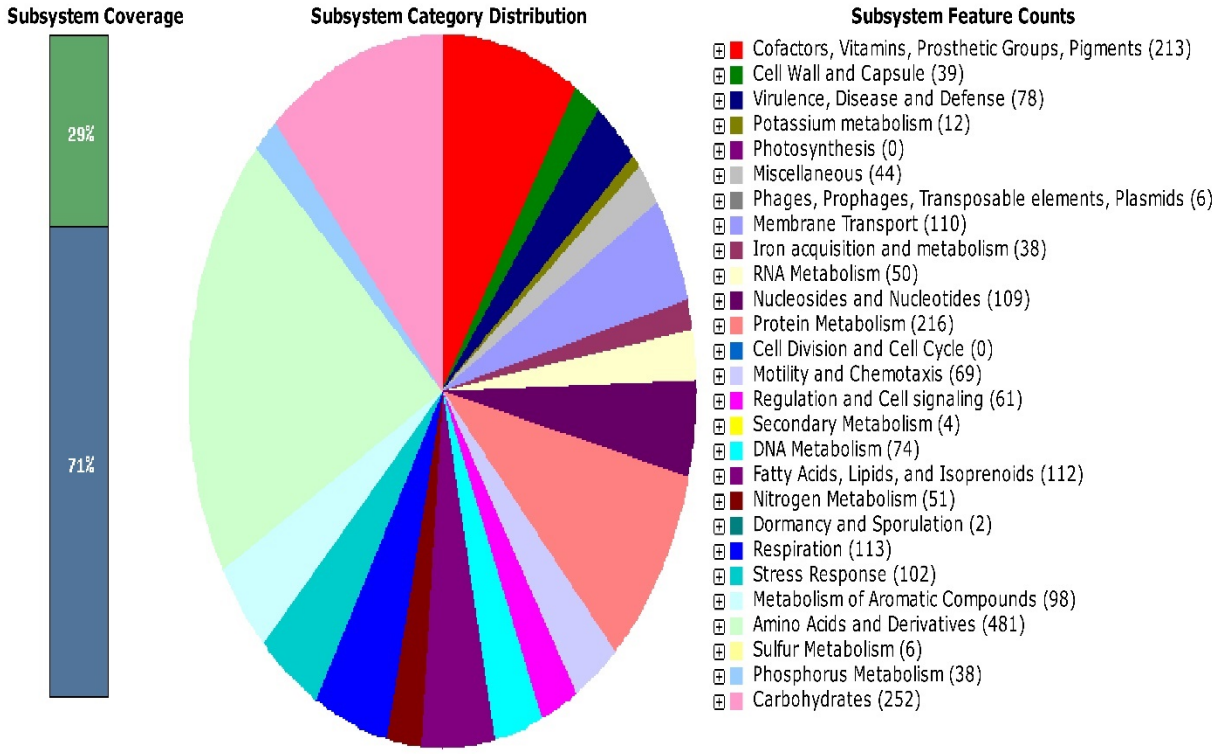

Fig. 2: RAST annotation of *Pseudomonas* sp. VB3: The subsystem category distribution shows the percentage distribution of the genes in different pathways, that are labelled with different colours and the number of genes in each are mentioned in the bracket.

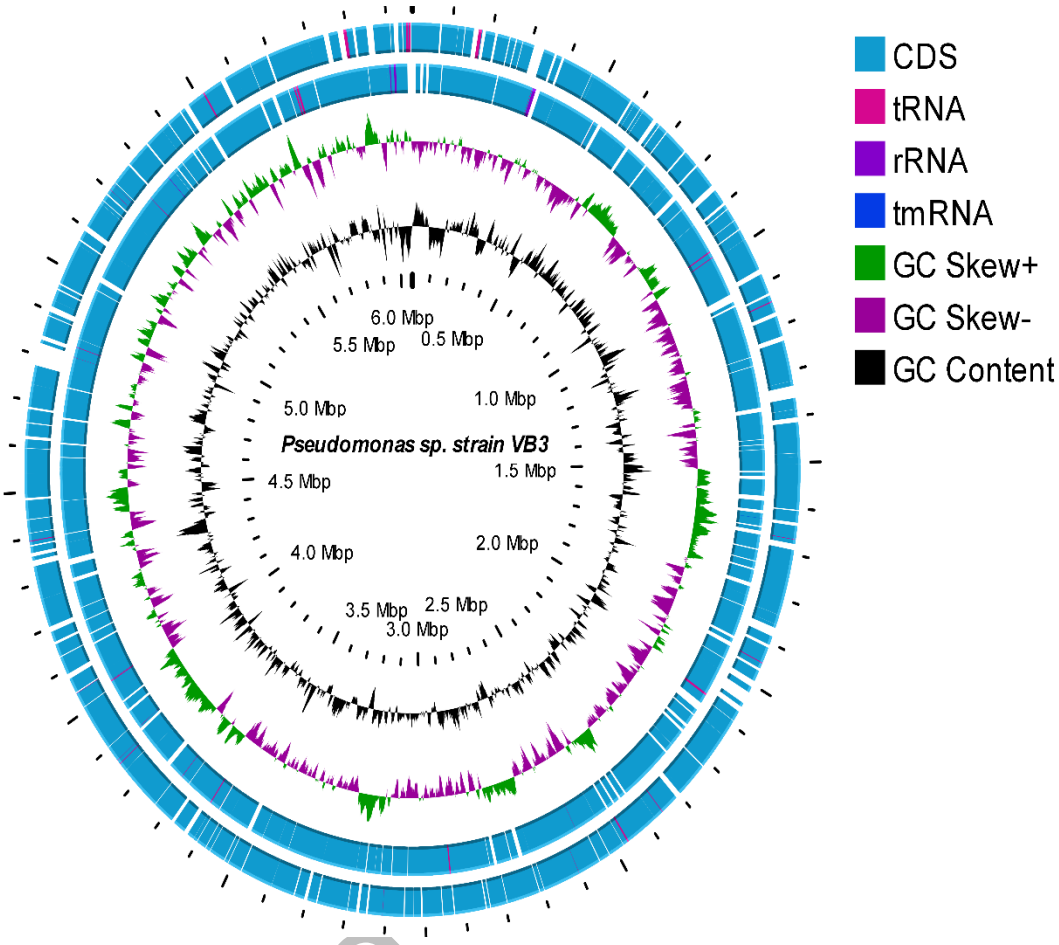

Fig. 3: Circular representation of *Pseudomonas* sp. strain VB3 contigs using Proksee (https://proksee.ca/). The scale is shown in megabases on the inside central circle. Moving outward, two outer blue circles show forward and reverse strand CDSs, respectively. The tRNAs (magenta), rRNAs (purple), tmRNA (dark blue) are shown in CDSs circles. The next circle shows GC content (black) followed by GC Skew (dark green and purple).

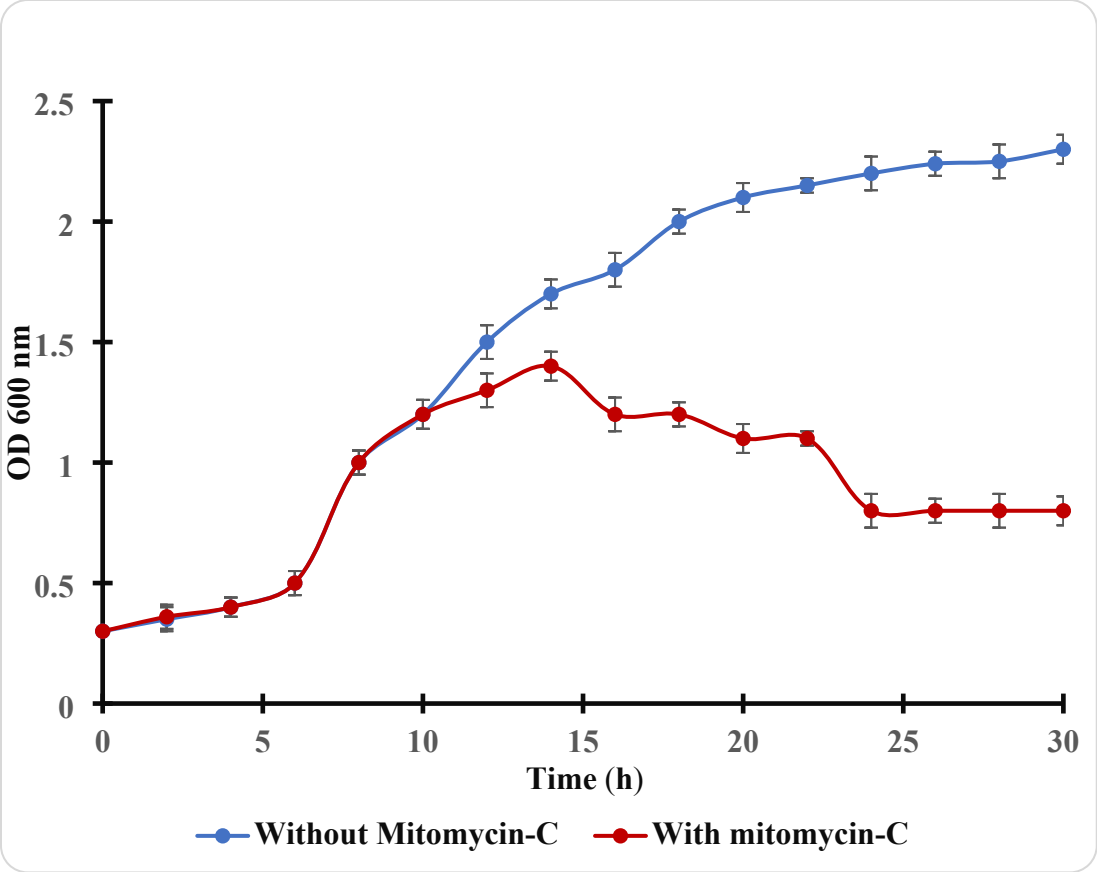

Fig. 4: Prophage induction in bacterial strain VB3 with and without Mitomycin-C.

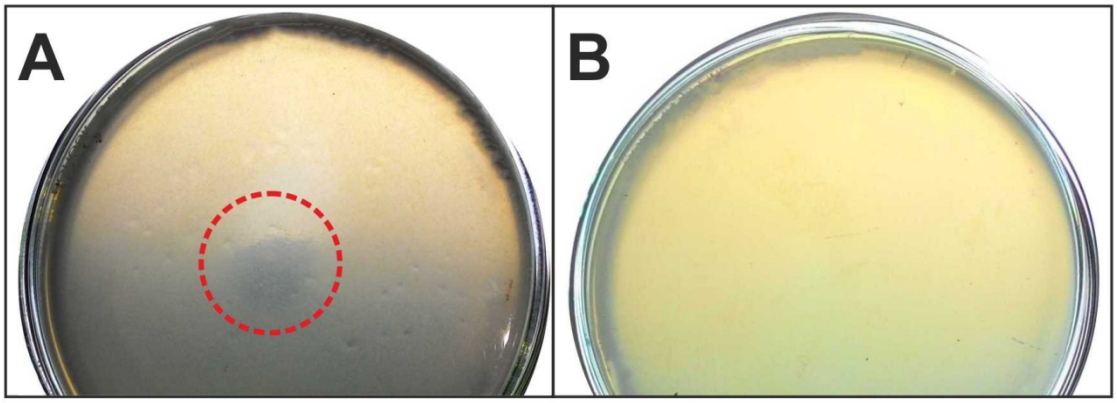

Fig 5: Spot test after induction with Mitomycin-C on the lawn of *Pseudomonas* sp. VB3.

(**A**) Mitomycin-C treated, Arrow depicts clearance zone, (**B**) Control, without Mitomycin-C.