Data in Brief 54 (2024) 110293

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Genomic dataset of a multiple-drug resistant *Pseudomonas* sp. strain RAC1 isolated from a flacherie infected Nistari race of *Bombyx mori* L.☆

Rittick Mondal[a], Paulami Dam[a], Joydeep Chakraborty[b], Shubhajit Shaw[a], Sayantan Pradhan[a], Sandip Das[a], Jannatun Nesa[c], Khemraj Meena[d], Amit Ghati[e], Sandip Dev Chaudhuri[a], Debjoy Bhattacharjee[a], Vivekananda Mandal[f], Biraj Sarkar[g,*], Amit Kumar Mandal[a,*]

[a] *Department of Sericulture, Raiganj University, North Dinajpur, West Bengal 733134, India*
[b] *Department of Microbiology, Raiganj University, North Dinajpur, West Bengal 733134, India*
[c] *Department of Zoology, Gangarampur College, Dakshin Dinajpur, West Bengal 733124, India*
[d] *Department of Biotechnology, School of Life Sciences, Central University of Rajasthan, Kishangarh, Rajasthan 305817, India*
[e] *Department of Microbiology, Barrackpore Rastraguru Surendranath College, Barrackpore, West Bengal 700120, India*
[f] *Plant and Microbial Physiology and Biochemistry Laboratory, Department of Botany, University of Gour Banga, Malda, West Bengal 732103, India*
[g] *Faculty of Allied Health Sciences (FAHS), The ICFAI University, Tripura; Kamalghat, Mohanpur, West Tripura 799210, India*

ARTICLE INFO

Article history:
Received 4 January 2024
Accepted 29 February 2024
Available online 5 March 2024

ABSTRACT

Species belonging to the genus *Pseudomonas* is a rod shaped Gram-negative bacteria emerged as an important silkworm pathogen with broad-level multi-drug resistance. The

☆ Data Availability: This whole-genome shotgun project was deposited in NCBI GenBank (accession number (NZ_JAUTXS000000000). The version described in this paper is NZ_JAUTXS000000000.2, and this version consists of sequences JAUTXS020000001 to JAUTXS020000038. The BioProject and BioSample accession numbers are PRJNA224116 and SAMN36766282, respectively. The raw data are available from the Sequence Read Archive (SRA) under accession number SRR25580540.

* Corresponding authors.
E-mail addresses: birajsarkar@iutripura.edu.in (B. Sarkar), amitmandal08@gmail.com (A.K. Mandal).

https://doi.org/10.1016/j.dib.2024.110293

Dataset link: Pseudomonas sp. RAC1, whole genome shotgun sequencing project (Original data)

Keywords:
Silkworm
Antimicrobial resistance (AMR)
Illumina NovaSeq 6000
Genetic diversity

extensive usage of antimicrobials in sericulture farming is gradually leading to the emergence of multi-drug resistance (MDR) strains, posing a significant threat to the well-being of both *Bombyx mori* L. and serifarmers. *Pseudomonas* spp. with MDR level may gets transmitted from the infected silkworm to human handlers either via direct contact or through contaminated feces. To understand the emerging concern of antimicrobial resistance (AMR) in *Pseudomonas* spp. provides insights into their genomic information. Here, we present the draft genome sequence data of *Pseudomonas* sp. strain RAC1 isolated from a flacherie infected Nistari race of *Bombyx mori* L. from the silkworm rearing house of Raiganj University, India and sequenced using the Illumina NovaSeq 6000 platform. The estimated genome size of the strain was 4494347 bp with a G + C content of 63.5%. The *de novo* assembly of the genome generated 38 contigs with an N50 of 200 kb. Our data might help to reveal the genetic diversity, underlying mechanisms of AMR and virulence potential of *Pseudomonas* spp. This draft-genome shotgun project has been deposited under the NCBI GenBank accession number NZ_JAUTXS000000000.

Specifications Table

Subject	Microbiology: Bacteriology
Specific subject area	Prokaryotic genomics
Data format	Raw, Analyzed
Type of data	Genomic sequences, tables and figures
Data collection	The isolate RAC1 was isolated from the hemolymph sample of flacherie infected 5th instar larvae of *Bombyx mori* L. and maintained in Cetrimide agar plate. Later the isolate was sequenced employing Illumina NovaSeq6000 platform followed by annotation of the genome assembly the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v6.6.
Data source location	Institution: Raiganj University • City/Town/Region:Raiganj • Country: India • Latitude and longitude for collected samples/data: 25.6071° N, 88.1306° E
Data accessibility	The reported data here is available at NCBI GenBank under the accession number NZ_JAUTXS000000000 (https://www.ncbi.nlm.nih.gov/nuccore/NZ_JAUTXS000000000.2) and/or https://0-www-ncbi-nlm-nih-gov.brum.beds.ac.uk/Traces/wgs/JAUTXS02

1. Value of the Data

- *Pseudomonas* sp. RAC1 genome sequence data will be useful to provide insight into the genotypic diversity of the isolate and to undermine the host-microbe interactions, molecular basis of pathogenesis, emerging antimicrobial resistance, etc.
- Researchers working on silkworm disease model will be benefited from this data in conducting studies related to advanced diagnostics kit development and in acquiring deep

Fig. 1. Infected silkworm specimen.

learning on disease-free silkworm harvesting system which can enhance the silk quality in real time, along with the reduction in silk crop losses.

- Researchers focusing on specific genes may conduct comparative genome analysis and to explore evolutionary and contrasting changes depending on the geographical region and host.

2. Background

Pseudomonas spp. is a Gram-negative opportunistic pathogen belongs to the *Gammaproteobacteria*, emerged as a suitable model bacterium to study bacterial virulence trait. It is ubiquitous in almost any human/animal-impacted environment [1]; such as, silkworms rearing farms. Almost 30–40 % of crop losses is associated with bacterial flacherie caused by *Pseudomonas* spp. [2,3]. Pathogenic microbial infections in *Bombyx mori* L. lead to a shift in metabolic profiles and disrupt enzymatic activities, resulting in cocoon spoilage and decrease in silk quality [4]. It is established that the genetic background of silkworms is highly homologous to specific genes related to human hereditary diseases [5]. Thus, there is a potential threat of silkworm pathogens to human health. So far, *Pseudomonas* spp. has emerged as an opportunistic human pathogen and the infections caused by such pathogen are difficult to treat due to its various antibiotic resistance mechanisms [1]. Therefore, it is crucial to understand the genomic insights of *Pseudomonas* spp. to interpret the plausible risk associated with MDR bacteria while changing its host. Nevertheless, this will help the researchers to undermine how genomic plasticity provides a suitable habitat to grow and to survive in antibiotics pressure.

3. Data Description

In this study, we present the draft genome sequence data of *Pseudomonas* sp. RAC1, isolated from the hemolymph sample of flacherie-infected 5th instar larvae of *Bombyx mori* L (Fig. 1). *De novo* assembly using Illumina NovaSeq 6000 platform was used for the draft genome sequencing of the isolated strain RAC1, and it was followed by the annotation of the genome assembly using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v6.6. The assembly produced a genome sequence of 4494347 bases in length, encompassing 38 contigs. The N50 length is 200 kb, whereas, the L50 count is found to be 7. The estimated G+C content is 63.5% and 434.71x coverage. A total of 4,021 coding sequences were annotated, including 7 rRNA genes (5 5S, 1 16S and 1 23S) and 58 tRNAs (Table 1). A whole genome based phylogenetic tree was constructed using Type (Strain) Genome Server (TYGS) to observe the strain to be distantly related with *Pseudomonas wayambapalatensis* RW3S1 (Fig. 2), as they were found to be placed in the same clade. Further study of the genome sequence of *Pseudomonas* sp. RAC1 might increase our understanding about the various potential properties of the organism like various antimicrobial resistance, host-pathogen interactions, etc. as apparent from the circular genome map obtained using BV-BRC web resources v. 3.33.16 (Fig. 3).

Table 1
Features (genome and annotation statistics) of the assembled sequence of *Pseudomonas* sp. RAC1.

Genome statistics	
Length of the genome	4,494,347
Genome coverage	434.71x
Assembly level	Contig
Total number of contigs	38
Contig N50	200 kb
Contig L50	7
GC percent	63.5
Annotation statistics	
Genes (total)	4,090
CDSs (total)	4,021
Genes (coding)	3,975
CDSs (with protein)	3,975
Genes (RNA)	69
rRNAs	5, 1, 1 (5S, 16S, 23S)
tRNAs	58
ncRNAs	4

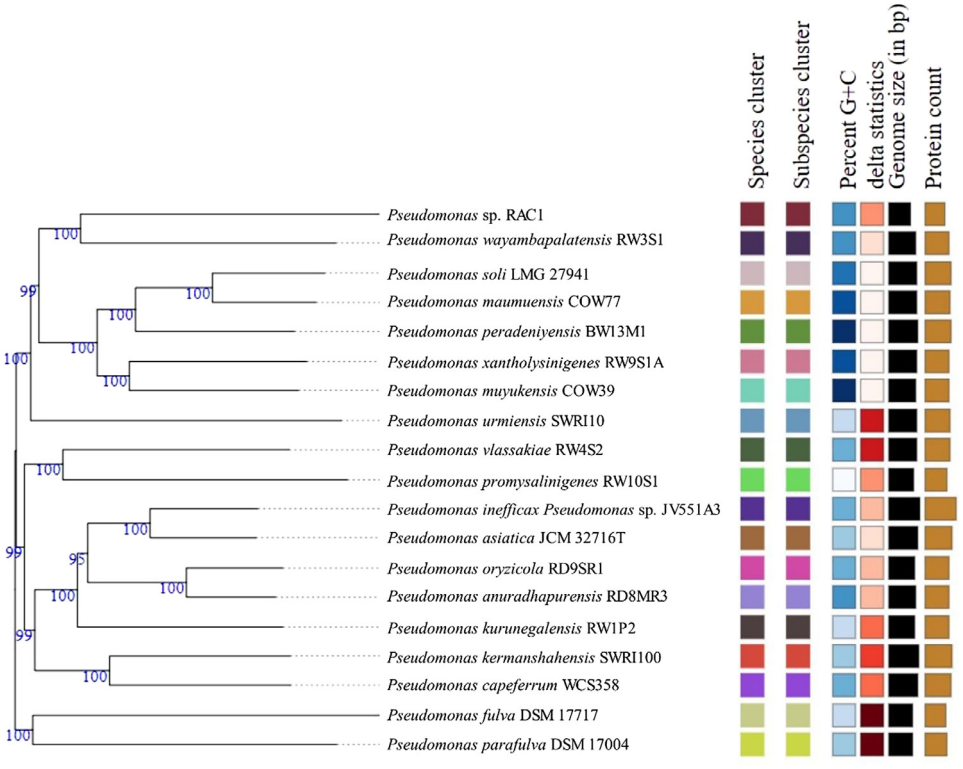

Fig. 2. GBDP tree (Whole genome sequence based).

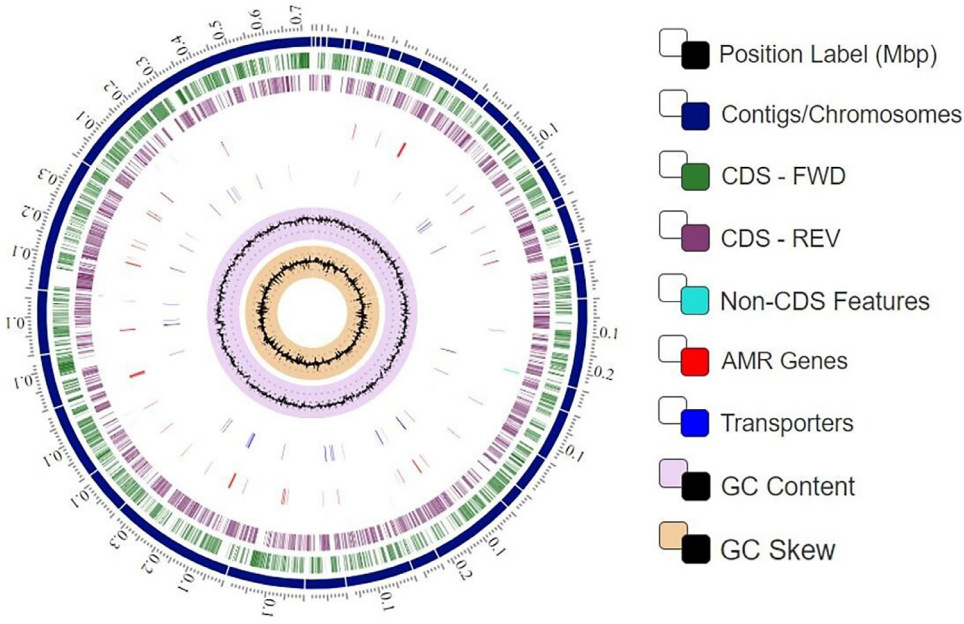

Fig. 3. Genome circular view. This image illustrates the Position label (Mbp), the contigs, forward and reverse CDS, Non-CDS features, presence of probable AMR genes, transporters, the GC content and GC skew of the reported draft genome sequence of *Pseudomonas* sp. RAC1.

Tree inferred with FastME 2.1.6.1 [6] from GBDP distances calculated from genome sequences. The branch lengths are scaled in terms of the GBDP distance formula d_5 [7]. The numbers above the branches are GBDP pseudo-bootstrap support values $>$ 60 % from 100 replications, with an average branch support of 99.5 %. The tree was rooted at the midpoint [8,9].

4. Experimental Design, Materials and Methods

4.1. Isolation of Sample and Culture, Maintenance and Preservation

Infected 5th instar larvae of *Bombyx mori* L. were collected in a sterile container from the silkworm rearing house of Raiganj University, India. The collected larvae were washed with 70% ethyl alcohol and surface sterilized with 5% (w/v) sodium hypochlorite for 5 min and finally rinsed three times with sterile deionized water. Hemolymph was collected in a pre-chilled microtube by rupturing the first abdominal leg of infected larvae (3rd day of 5th instar larvae) with a sterile super-fine 31G short needle [10]. After the collection of hemolymph sample from flacherie infected Nistari race of *Bombyx mori* L., the specimen was spread onto a Cetrimide agar media plates (Cetrimide Agar, Sisco Research Laboratories Pvt. Ltd., India). The inoculated agar plate was then incubated overnight at 30 °C. After incubation, individual single colonies were picked and transferred to fresh cetrimide plates to obtain pure colonies. The pure cultures were regularly maintained in fresh cetrimide plates and preserved at -80 °C in 20% glycerol.

4.2. Genomic DNA Isolation, Draft Genome Sequencing and Assembly

The standard phenol:chloroform procedure was used to extract chromosomal DNA [11]. Initially, the DNA extraction was carried out using DNeasy Ultraclean Microbial Kit. Later, the extracted DNA quantity is confirmed using Nanodrop 1000.

The library preparation was carried out using Ultra II DNA lib Preparation for Illumina Kit (NEBNext #E7645S/L). The final library was quantified using a Qubit 4.0 fluorometer (Thermofisher #Q33238) using DNA HS assay kit (Thermofisher #Q32851) following manufacturer's protocol. In order to determine the insert size, high-sensitivity D1000 screentapes (Agilent # 5067–5582) in accordance with the manufacturer's protocol were used to query the library on Tapestation 4150 (Agilent).

Raw data quality assessment was performed using FastQC v.0.11.9 [12]. The report generated by FastQC was summarized using MultiQC v.1.9 [13]. The data was preprocessed using Fastp v.0.20.1 (parameters: -f 3 -l 50 -c -g -q 30) [14]. Post filtering cleaned data were re-assessed using FastQC and were summarized using MultiQC. The processed paired-end reads were mapped to the pre KMA-indexed NCBI 2019 Genome Build (http://dx.doi.org/10.25910/5cc7cd40fca8e) database using KMA [15].The processed paired-end reads were assembled using MEGAHIT v.1.2.9 [16]. The prokaryotic assembled genomes were further binned using MetaBAT2 v.2.1 (default parameters) [17]. Estimation of the genome completeness and rate of contamination were assessed using CheckM [18].

4.3. Genome Annotation and Phylogenetic Analysis

The annotation of the genome assembly was performed using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v6.6 with the methods "best-placed reference protein set" and "GeneMarks-2+" [19].

The circular genome map was viewed using BV-BRC webresources v. 3.33.16 to observe the different contigs, CDSs (both forward and reverse), antimicrobial resistance (AMR) genes, transporters; GC content and GC skew.

The genome sequence data was uploaded to the Type (Strain) Genome Server (TYGS), a free bioinformatics platform available under https://tygs.dsmz.de, for a whole genome-based taxonomic analysis [20]. The phylogenomic tree was constructed using FastMe version 2.1.6.1 [6] from the GBDP (genome blast distance phylogeny) distances calculated from genome sequences. The branch lengths are scaled in terms of the GBDP distance formula d_5 [7]. The tree was rooted at the midpoint [8].

4.4. Antimicrobial Susceptibility Testing

Kirby-Bauer disc diffusion test was performed to check the antibiogram profile of *Pseudomonas* sp. RAC1 as per the earlier study [10]. Different classes of antibiotics hexa discs were purchased from HiMedia Laboratories PVT. LTD., India. The freshly grown overnight culture of isolated strain was spread and the discs were placed on top of Mueller-Hinton agar (MHA, Himedia Laboratories PVT. LTD., India) plates. The plates were then incubated overnight at 37 °C. Antibiotics to which isolated strains were sensitive formed halo inhibition zones, while those showed no halo inhibition zone indicated as resistant towards such antibiotics (Supplementary table S1).

Limitations

Not applicable.

Ethics Statement

Authors have read and follow the ethical requirements for publication in Data in Brief and confirming that the current work does not involve human subjects, animal experiments, or any data collected from social media platforms.

Data Availability

Pseudomonas sp. RAC1, whole genome shotgun sequencing project (Original data) (NCBI GenBank).

CRediT Author Statement

Rittick Mondal: Conceptualization, Data curation, Methodology, Software, Writing – original draft; **Paulami Dam:** Writing – review & editing; **Joydeep Chakraborty:** Methodology, Software; **Shubhajit Shaw:** Formal analysis; **Sayantan Pradhan:** Formal analysis; **Sandip Das:** Investigation, Visualization; **Jannatun Nesa:** Writing – review & editing; **Khemraj Meena:** Investigation, Visualization; **Amit Ghati:** Validation, Visualization; **Sandip Dev Chaudhuri:** Visualization; **Debjoy Bhattacharjee:** Visualization; **Vivekananda Mandal:** Validation, Supervision; **Biraj Sarkar:** Methodology, Software, Writing – original draft, Supervision; **Amit Kumar Mandal:** Conceptualization, Validation, Supervision.

Acknowledgments

Rittick Mondal would like to acknowledge the Department of Science & Technology, Government of India, for the DST-INSPIRE Ph.D. Fellowship (DST-INSPIRE-SRF; INSPIRE Code- IF190457). Paulami Dam was provided with an independent Ph.D. fellowship for financial assistance (UGC-SRF; NTA Ref.No. 201610181190). Amit Kumar Mandal acknowledges DST-SERB India for financial assistance (sanction order number: EEQ/2021/000058).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary Materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.dib.2024.110293.

References

[1] S.P. Diggle, M. Whiteley, Microbe profile: pseudomonas aeruginosa: opportunistic pathogen and lab rat, Microbiology 166 (2020) 30.
[2] D. Naskar, R.R. Barua, A.K. Ghosh, S.C. Kundu, Introduction to silk biomaterials, in: Silk Biomaterials for Tissue Engineering and Regenerative Medicine, Woodhead Publishing, 2014, pp. 3–40.
[3] P.C. Santha, S.K. Bhargava, S.S. Sindagi, C.K. Kamble, Bacterial flacherie of silkworm, Bombyx mori and it's control by the application of antibiotics, J. Exp. Zool. India 10 (2007) 1–7.
[4] A. Javaid, M. Hussain, K. Aftab, M.F. Malik, M. Umar, T. Iqbal, Isolation and characterization of bacteria associated with silkworm gut under antibiotic-treated larval feeding, Braz. J. Biol. 84 (2021) e249664.

[5] G. Li, X. Xia, S. Zhao, M. Shi, F. Liu, Y. Zhu, The physiological and toxicological effects of antibiotics on an interspecies insect model, Chemosphere 248 (2020) 126019.
[6] V. Lefort, R. Desper, O. Gascuel, FastME 2.0: a comprehensive, accurate, and fast distance-based phylogeny inference program, Mol. Biol. Evol. 32 (2015) 2798–2800.
[7] J.P. Meier-Kolthoff, A.F. Auch, H.P. Klenk, M. Göker, Genome sequence-based species delimitation with confidence intervals and improved distance functions, BMC Bioinformatics 14 (2013) 60, doi:10.1186/1471-2105-14-60.
[8] J.S. Farris, Estimating phylogenetic trees from distance matrices, Am. Nat. 106 (1972) 645–668.
[9] C.M. Gaffney, Identification and Antimicrobial Properties of Pseudomonas from Soil, Sam Houston State University, 2020 Doctoral dissertation.
[10] S. Some, B. Sarkar, K. Biswas, T.K. Jana, D. Bhattacharjya, P. Dam, R. Mondal, A. Kumar, A.K. Deb, A. Sadat, Bio–molecule functionalized rapid one-pot green synthesis of silver nanoparticles and their efficacy toward the multidrug resistant (MDR) gut bacteria of silkworms (Bombyx mori), RSC Adv. 10 (2020) 22742–22757.
[11] J. Marmur, A procedure for the isolation of deoxyribonucleic acid from micro-organisms, J. Mol. Biol. 3 (1961) 208–218.
[12] S. Andrews, FastQC: a quality control tool for high throughput sequence data, (2010).
[13] P. Ewels, M. Magnusson, S. Lundin, M. Käller, MultiQC: summarize analysis results for multiple tools and samples in a single report, Bioinformatics 32 (2016) 3047–3048.
[14] S. Chen, Y. Zhou, Y. Chen, J. Gu, fastp: an ultra-fast all-in-one FASTQ preprocessor, Bioinformatics 34 (2018) i884–i890.
[15] P.T.L.C. Clausen, F.M. Aarestrup, O. Lund, Rapid and precise alignment of raw reads against redundant databases with KMA, BMC Bioinform. 19 (2018) 307.
[16] D. Li, C.M. Liu, R. Luo, K. Sadakane, T.W. Lam, MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph, Bioinformatics 31 (2015) 1674–1676.
[17] D.D. Kang, F. Li, E. Kirton, A. Thomas, R. Egan, H. An, Z. Wang, MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies, PeerJ 7 (2019) e7359.
[18] D.H. Parks, M. Imelfort, C.T. Skennerton, P. Hugenholtz, G.W. Tyson, CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes, Genome Res. 25 (2015) 1043–1055.
[19] T. Tatusova, M. DiCuccio, A. Badretdin, V. Chetvernin, E.P. Nawrocki, L. Zaslavsky, A. Lomsadze, K.D. Pruitt, M. Borodovsky, J. Ostell, NCBI prokaryotic genome annotation pipeline, Nucleic Acids Res. 44 (2016) 6614–6624.
[20] J.P. Meier-Kolthoff, M. Göker, TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy, Nat. Commun. 10 (2019) 2182.