Data Article

Draft genome sequence data of a T7like phage 3A_8767 isolated from wastewater of a butcher house near Palar river

Avtar Sain, N.S. Jayaprakash*

Centre for Bio-separation Technology, Vellore Institute of Technology, Vellore 632014, Tamilnadu, India

A B S T R A C T

Phage 3A_8767 is a newly isolates phage from river water sample against *Salmonella* typhi 8767 (MTCC). The genome of the phage is linear, double stranded and 38,821 bp long in size. A total 49 functional ORF (open reading frame) were annotated and no tRNA was predicted. Phage 3A_8767 has icosahedral shaped head with stubby tail which comes under family *Podoviridae*, and genera T7 like virus.

© 2020 Published by Elsevier Inc.
This is an open access article under the CC BY-NC-ND license. (http://creativecommons.org/licenses/by-nc-nd/4.0/)

Specifications table

Subject	Biology
Type of data	Whole genome sequence data
How data were acquired	Next generation sequencing (NGS) of bacteriophage genome was done on Illumina platform
Data format	Raw, Analyzed
Parameters for data collection	Water samples were collected for isolation of bacteriophage against multidrug resistant *Salmonella* and DNA was isolated for the sequencing
Description of data collection	Genome sequence of phage 3A_8767
Data source location	Vellore, Tamilnadu, India (12.939196 N, 79.132771E)
Data accessibility	https://www.ncbi.nlm.nih.gov/nuccore/MH382198.1
Related research article	none

* Corresponding author.
  E-mail address: nsjayaprakash@vit.ac.in (N.S. Jayaprakash).

Value of the Data

- Data provides information of genes and proteins encoded by those genes.
- Data ensures of no toxic gene carried by phage which, makes it a potent agent for applications.
- Data provides information about lytic enzymes which could be used as to design new antimicrobials to ease the burden of chemical antimicrobial pollution on environment.
- Data can be considered by researchers for the genomic, proteomic and pattern of evolutionary studies
- This study may assist researchers in sequence comparison.

1. Data description

Bacteriophage 3A_8767 was isolated from butcher house waste water sample near Palar river, Vellore district, Tamilnadu, India at geographical location (12.939196 N, 79.132771E). The Palar river collects waters from a vast geographical landscape of the Tamilnadu which makes it a thriving source of large and diverse number of organisms including phage battery. Transmission electron micrograph in figure-1[A] showed the morphology of phage 3A_8767 having icosahedral head (∼52 nm in diameter) and stubby tail which is a typical characteristic of *Podoviridae* [1]. The DNA polymerase gene was used to study the evolutionary pattern and found that the phage 3A_8767 comes in the clade of T7 like virus from Fig.1[B]. The blast result showed 92.84% identity to bacteriophage T7 (Accession no. V01146.1) with 92% query coverage and 0.0 E – value. The Table 1 represents the genome sequence and annotation summary comprising of size, $G+C$ content, open reading frames and predicted functional genes. The average fragment size of the library was 449 bp. Data was generated on illumine platform which provided 7302,941 number of reads.

2. Experimental design, materials and methods

The phage isolation and enumeration was carried out against *Salmonella* typhi host strain 8767 MTCC [2]. Luria Bertani broth and agar were used for host strain propagation. Sample swabs collected from collection site were enriched in host strain culture, prior to isolation. DNA isolation was carried out as per the protocol mentioned in the DNA and RNA isolation kit (Gene-JET viral DNA and RNA purification kit, Thermo Scientifics), after plaque purification and phage enumeration. The negative staining of the phage particles was performed using 2% w/v phosphotungstic acid [3]. QIAseq FX DNA library kit was used to generate the paired end DNA sequence library with initial concentration of 100 ng gDNA [4]. Smaller fragments were obtained by enzymatic shearing followed by continuous end repair to add adenosine nucleotide at 3′ end, which facilitates the adapter binding. A high fidelity amplification step was performed using HiFi PCR master to ensure the maximum yield.The amplified library was analyzed in Bioanalyzer 2100

Table 1
Genome sequence characteristic of phage.

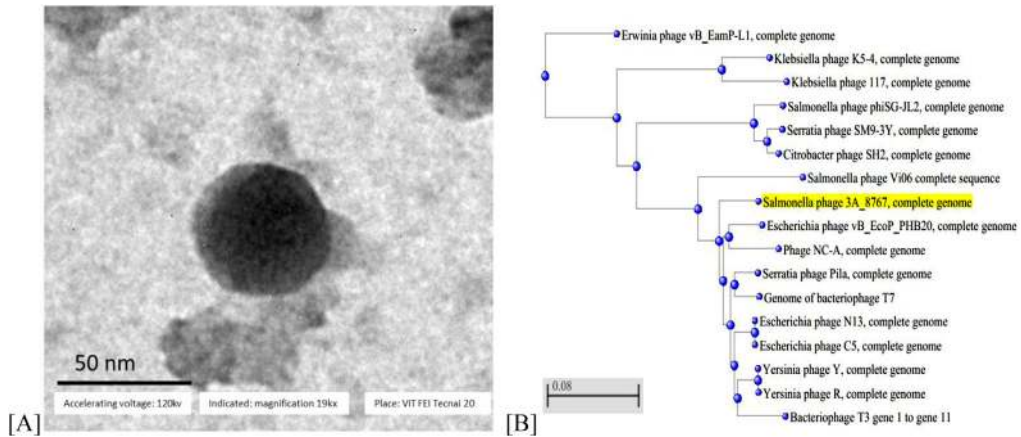| Index	Value
Genome size	38821bp
GC content	49.25%
ORF	175
Predicted no. of genes	49
Accession no.	MH382198

Fig. 1. [A] Phage morphology by transmission electron micrograph of bacteriophage 3A_8767. The icosahedral head with diameter 52 nm and size bar 50 nm is visible in image. [B] Phylogenetic tree was constructed based on phage DNA polymerase of selected bacteriophages with NCBI tools using neighbor joining method.

(Agilent Technologies) using High Sensitivity (HS) DNA chip. The *de novo* assembly from the high quality reads was obtained using CLC genomics workbench v6.0 [5].

The assembled genome of bacteriophage 3A_8767 was annotated with the help of Rapid Annotation using Subsystem Technology (RAST) [4,6] and verified by NCBI blast for each gene individually. Open reading frames were determined by ORF finding tool from NCBI. Each CDS was confirmed by BLASTn tool and each protein was determined for its function by BLASTp [7]. The tRNAscan-SE tool [8] was used to search tRNA in the genome sequence. The *Salmoenlla* phage 3A_8767 genome was analysed for the presence of tandem repeats as well as insert sequences by using the online tool Tandem Repeat Finder v-4.09 and Repeat Masker respectively [9].

Nucleotide sequence accession number – MH382198

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships which have, or could be perceived to have, influenced the work reported in this article.

Acknowledgments

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.dib.2020.105446.

References

[1] M. Krupovic, B.E. Dutilh, E.M. Adriaenssens, J. Wittmann, F.K. Vogensen, M.B. Sullivan, J. Rumnieks, D. Prangishvili, R. Lavigne, A.M. Kropinski, J. Klumpp, A. Gillis, F. Enault, R.A. Edwards, S. Duffy, M.R.C. Clokie, J. Barylski, H.W. Ackermann, J.H. Kuhn, Taxonomy of prokaryotic viruses: update from the ICRV bacterial and archaeal viruses subcommittee, Arch. Virol. 161 (2016) 1095–1099 https://doi.org/10.1007/s00705-015-2728-0.
[2] K.S. Sritha, S.G. Bhat, Genomics of salmonella phage ΦStp1: candidate bacteriophage for biocontrol, Virus Genes 54 (2018) 311–318 https://doi.org/10.1007/s11262-018-1538-3.
[3] K.S. Sritha, J. Augustine, S.G. Bhat, Draft genome sequence data of T-5 like salmonella bacteriophage ΦSP3 with demonstrated therapeutic potential, Data Brirf 27 (2019) https://doi.org/10.1016/j.dib.2019.104606.
[4] V.V. Kadnikov, A.V. Mardanov, Y.A. Frank, A.V. Beletsky, O.V. Karnachuk, N.V. Ravin, Genomes of three bacteriophages from the deep subsurface aquifer, Data Brief 22 (2019) 488–491 https://doi.org/10.1016/j.dib.2018.12.045.
[5] QIAGEN, CLC Genomics workbench user manual, 2016.
[6] K. McNair, R.K. Aziz, G.D. Pusch, R. Overbeek, B.E. Dutilh, R. Edwards, Phage genome annotation using the RAST pipeline, Methods Mol. Biol. 1681 (2018) 231–238 https://doi.org/10.1007/978-1-4939-7343-9_17.
[7] X. Yu, Y. Xu, Y. Gu, Y. Zhu, X. Liu, Characterization and genomic study of "phiKMV-Like" phage PAXYB1 infecting pseudomonas aeruginosa, Sci. Rep. 7 (2017) https://doi.org/10.1038/s41598-017-13363-7.
[8] T.M. Lowe, S.R. Eddy, TRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence, Nucl. Acids Res. 25 (1996) 955–964 https://doi.org/10.1093/nar/25.5.0955.
[9] Q. Peng, Y. Yuan, Characterization of a newly isolated phage infecting pathogenic escherichia coli and analysis of its mosaic structural genes, Sci. Rep. 8 (2018) https://doi.org/10.1038/s41598-018-26004-4.